Predictive Analytics of Image Descriptors for Students Biometric Authentication System

Babatunde Taiwo OLOMOLA
LCU/PG/002445

Being a MSc Thesis Submitted to the Department of Computer Science, Faculty of Natural and Applied Sciences, Lead City University, Ibadan, Oyo State, Nigeria

In Partial Fulfillment of the Requirements for the Award of Master of Science Degree (MSc) in Computer Science

2023

Certification

This is to certify that Babatunde Taiwo OLOMOLA with matriculation number LCU/PG/002445 carried out this research work titled "An Ensemble Predictive Analytics of Image Descriptors for Student Authentication Using Bimodal Biometrics" in the Department of Computer Science, Faculty of Natural and Applied Sciences, Lead City University, Ibadan, Oyo State, for the award of Master of Science Degree (MSc) in Computer Science and that this has not been previously submitted.

_____ _____

Dr. Azeez Waheed Date
(Supervisor)

_____ _____

Dr. Wilson Sakpere Date
(Head of Department)

Dedication

This research work is dedicated to the owner of all knowledge and wisdom, the Almighty.

Acknowledgement

This is to acknowledge the Management and entire academic community of LEAD CITY UNIVERSITY, Ibadan, for the enabling environment and privileges to carry out this research work.

Profound gratitude goes to my project supervisor, Dr. Azeez Waheed, who painstakingly guided me throughout the period of this work. Also, to all lecturers in Computer Science Department, who have imparted a great measure of knowledge in this programme: Dr. Wilson Sakpere (Head of Department), Dr. Akintayo Ayoade, Dr. Wumi Ajayi, Dr. Rahmon Badru, Dr. Taye Adeniran, Mrs. Oluseyi Afe and others.

I acknowledge with gratitude the key role played by Mr. Olusegun Popoola (CITM, Yaba College of Technology, Lagos), Mr. Ayo Ademoroti and Mr. Oluwaseye J. Ige (OTM Department, Yaba College of Technology, Lagos) and Mr. Taiwo Olaleye (Federal University of Agriculture, Abeokuta)

Lastly, Pastor and Mrs. Daniel Akinyemi, my darling wife Olajumoke Olatunmbi Olomola and my lovely children Heritage, Hephzibah and DavidKing for their great contributions to the overall success of this work.

"Though the above-mentioned institutions and persons have assisted in the process of this research work, I alone stand responsible for the errors, if any, found in the work".

Abstract

Nowadays, our education as well as other sectors is in a verge where accuracy of its authentication process is vital so as to close door tightly against impostors and impersonators. This thesis therefore improved the effectiveness and accuracy of biometrics based authentication models already in use for students' attendance. The authentication framework is a five-phase biometric-based student attendance verification system that combined iris and fingerprint recognition attributes for the purpose of training deep learning models. The first phase entails the image acquisition. The acquired image inputs were then subjected to the feature extraction phase where attributes were extracted from the image inputs in form of numeric image descriptors. Data resampling were done in order to ensure a balanced training set for the machine learning-based study, which were consequently deployed for the deep learning phase after which performance evaluation was carried out in phase five. Three learner algorithms of Decision Tree (DT), Support Vector Machine (SVM), and Sequential Minimal Optimization (SMO) were trained with numeric vectors extracted from both fingerprint and iris biometrics of students for a student attendance authentication system. The numeric vectors were extracted using the SqueezeNet, InceptionV3, VG16, VG19, and Painters image embedders to return five distinct databases. The performances of the three base learners' algorithms were evaluated alongside the performance of a Vote ensemble model after the five databases are subjected to a synthetic minority oversampling. Experimental results returned the Vote ensemble as the best model for student authentication which is followed by the SMO. The F1 score of Vote ensemble outperforms other models across the five datasets, with accuracy score as high as 0.999. The synthetic minority oversampling of the training sets further improved the performance of the models through data resampling. Consequently, Vote ensemble machine learning is better deployed for student authentication systems with any of the five image embedders.

Keywords: Ensemble Machine, Information Security, Biometric Authentication, Biometric Recognition, Database Management System

Word Count: 273

Table of Contents

Chapter Five: Conclusion

List of Tables

List of Figures

List of Figures Cont'd

Chapter One

Introduction

1.1 Background to the Study

A security concept known as an authentication system works as a preventative mechanism to foresee and stop illegal access to a service or data[1]. As an authentication system, biometrics captures the distinctive biological characteristics of users in what is frequently referred to as "*what the user is*," in addition to the "*what a user knows*" identity authentication criteria of username and/or password[1]. Due to the differentiating characteristic of being more representative of who a person is than any other secondary means of verification, biometrics have also demonstrated to be more effective than the usage of tokens or hardware things (what a user owns)[2].

A person can be positively identified by their biological, physiological, or behavioural characteristics, bearing in mind their individuality. The fingerprint, face, or retinal scans, as well as occasionally the voice for voice recognition authentication systems, are common biometric traits used for authentication systems[3]. For verification purposes, the same biological characteristics are recorded, preventing impersonation, identity theft, identity mismatch, etc. In this scenario, users are usually given access to data or services when the retrieved biological attributes and those in the database are successfully matched. Therefore, the science of biometrics involves the identification and verification of people using their quantifiable characteristics[2]. Before granting or denying access, a system with biometric functionality might uniquely identify users based on their measurably distinct characteristics. Diverse studies have shown that the iris is one of the characteristics that is most crucial for ensuring more accuracy, identification, and efficiency in identity recognition[3]. It is also a quick and direct opinion from that iris

feature, which uses numeric feature vectors to express human identification, is thought to have more than 250 distinct pieces[4].

Studies have also shown that feature vectors are distinct even amongst identical twins[5]. What more, as detailed by that Iris is reputable to be hard to spoof hence the wide acceptance for user authentication. Another biological factor for user authentication is the finger print. Finger print recognition are deployed across industries to authenticate humans for verification purposes. It is of a wide opinion in the documents of that fingerprint properties used in conjunction with feature extraction techniques enable predictive analytics to identify and authenticate persons[6]. A lofty performance measure for verification and attendance management is anticipated if a multimodal approach is employed for authentication. The deployment of fingerprint and iris as user authentication elements should then hold great potentials for recognition authentication purposes. The fusion guarantees a decrease in verification failure for class attendance management in addition to enhancing the performance of a verification system[7]. The concatenated numeric properties feature vectors in this case could be used in a classification technique for modelling before deployment for tracking class attendance. The generated feature vector data can be used for data mining, and the ensemble guarantees a system for user authentication for a bimodal approach.

The combination of iris and fingerprint recognition for authentication could thus well serve the goal of data mining for user identifying pattern recognition given the growing requirement for a trustworthy authentication system. As a result, data mining might be used to decipher statistics, machine learning, and mathematical algorithms to identify patterns from the iris and fingerprint data in a detailed quantity. The information might then be transformed into something understandable by applying pattern recognition to it.

The feature extraction, a type of picture preprocessing, would take the two-factor inputs and extract feature vectors from them to create the biometric database for an authentication system. This means that the processes of image selection, preprocessing, transformation, data mining, and evaluation are all part of the knowledge discovery process from the fingerprint-iris input.

Machine learning makes it possible to evaluate, as it did in this study. The retrieved feature vectors of the biological features for modelling would be applied in the course of training the algorithm for machine learning, which would then be used for user authentication. Because machine learning is effective at processing large amounts of data for pattern detection, its functionality becomes crucial in this study. Therefore, during the testing stage of the machine learning methodology, a pupil might be simply identified using supervised machine learning technology. A trustworthy data mining strategy for user authentication is promised. Therefore, iris and fingerprint biological features will be combined in this study's pattern recognition and feature extraction processes. The retrieved feature vectors would be used to train deep learning algorithms, which would then be used to train user authentication and evaluation systems. The capabilities that biometrics and data mining techniques provide for user authentication is generic, but specifically, the following factors drive this study's motivation - the potential for improving user authentication through a multimodal method using biological characteristics of the iris and fingerprint; the need to enhance current research using machine learning for multimodal users' authentication, including the use of synthetic minority oversampling to solve data over-fitting; the study is additionally driven by a desire to advance the work by deploying deep image embedding through transfer learning methodology for feature extraction purposes.

1.2 Statement of the Problem

Low performance metrics are a result of the flaws discovered in primary studies, which have been shown in the literature to have disastrous effects on a trustworthy user authentication system. This is partially as a result of methods used for feature extraction in image-based authentication. When compared to deep learning-based algorithms, filters produced far lower performance evaluation metrics[8]. Also, data mining techniques that do not take into account the under-fitting/over-fitting trade-off result in a user authentication system is not at its best[9]. There is a need therefore to determine whether and which of deep image embedding network for feature extraction work best with deep learning algorithm as an attendance authentication management application

1.3 Aim and Objectives of the Study

The aim of this study is to implements an authentication system through the instrumentality of bimodal biometrics and data mining functionalities which will improve the trustworthiness and accuracy of the biometrics authentication system already in use. The specific objectives are to:

i. acquire iris and fingerprint biometrics images
ii. deploy five (5) image embedding models for feature vector extraction from iris and fingerprints biometrics images.
iii. implement ensemble predictive analytics while deploying synthetic minority oversampling technique for addressing the under-fitting and over-fitting trade-off.
iv. evaluate the effectiveness of the five (5) image embedding models and deep learning prototypes for an efficient attendance management system

1.4 Significance of the Study

This study determined the efficiency of deep image embedders for feature extraction tasks on finger print and iris-based input images. It evaluated the performance of the embedders coupled with their ensemble with deep learners for attendance management purposes. The significant role of the synthetic minority oversampling technique on a bimodal feature attributes will be established in the study. The study will ultimately ascertain various data science factors suitable for a biometric-based attendance management solution using predictive analytics.

1.5 Scope of the study

This work implements an authentication system through the instrumentality of bimodal biometrics and data mining functionalities. A multimodal iris and fingerprint factors are captured and then subjected to multi-filtering feature extraction to return feature vectors comprising of the biological traits inherent in the two-factor. The extracted feature vectors are deployed to train deep learning algorithms in a supervised machine learning framework for subsequent student authentication.

1.6 Limitation of the study

This work concentrates on the use of a bi-modal biometric attributes for attendance management system through predictive analytics. also, the study evaluates the performances of deep image embedder networks for feature extraction, and the performance of deep learning algorithms for pattern recognition.

1.7 Operational Definition of Terms

i. Biometrics: Biometric is sometimes defined as an area focusing on measuring and analyzing a person's unique characteristics. Biometric uses technologies that verify or recognize the identity of a living person based on their physiological and behavioral characteristics.

ii. Multimodal biometrics: Multimodal biometrics is basically a blending of two or more biometric methods or system in an automatic Verification system.

iii. Spoofing: - This is illegal situation where data can be unauthentic. This happens to be a major challenge in Unimodal biometrics.

iv. Information Security: This can be referred to as a method employed in securing information or data from an illicit user to ensure the confidentiality, integrity and availability of the data.

v. Authentication: It is the process of authenticating the originality of an approved or a legal user.

vi. Application- A computer software developed for handling a specialized task

vii. Repository- a central safe where data is kept where data is managed. The data of data will be combined in a storehouse.

Endnotes

1. B. Ammour, L. T. Bouden, & M. Ramdani. "*Face-Iris Multimodal Biometric Identification System*." **Electronics**, 2020: 1-18.

2. M. Faundez-Zanuy, J. Fierrez, M. A. Ferrer, M. Diaz, R. Tolosana, & R. Plamondon. *"Handwriting biometrics: Applications and future trends in e-security and e-health."* **Cognitive Computation** 12, 2020. 940-953.

3. M. C. Chiu, G. J. Hwang, L. H. Hsia, & F. M. Shyu. "*Artificial Intelligence-Supported Art Education: A Deep Learning-Based System for Promoting University Students' Artwork Appreciation and Painting Outcomes.*" **Interactive Learning Environments,** 2022: 1-19.

4. N. Dong, L. Zhao, C. H. WuJ, & F. Chang. "*Inception V3 Based Cervical Cell Classification Combined With Artificially Extracted Features*." **Applied Soft Computing 1**, no. 93 2020: 106311.

5. S. Bhattacharya, G. S. Nainala, P, Das, & A, Routray. "*Smart Attendance Monitoring System (SAMS): A Face Recognition Based Attendance System for Classroom Environment.*" ***IEEE 18th International Conference on Advanced Learning Technologies (ICALT)***. IEEE, 2018.

6. S. Krishnmaoorthy, L. Rueda, S. Saad, & H. Elmiligi. "*Identification of User Behavioral Biometrics for Authentication Using Keystroke Dynamics and Machine Learning.*" **2nd International Conference on Biometric Engineering And Applications**, 2018, 50-57.

7. H. S. Maghdid, A. T. Asaad, K. Z. Ghafoor, A. S. Sadiq, S. Mirjalili, & M. K. Khan. "*Diagnosing COVID-19 Pneumonia From X-ray And CT Images Using Deep Learning and Transfer Learning Algorithms.*" **Multimodal Image Exploitation and Learning**, 2021: 99-110.

8. N. A. Wirdiani, T. Lattifia, I. K. Supadma, B. K. Mahar, D. N. Taradhita, & A. Fahmi. "Real-time Face Recognition with Eigenface Method." ***International Journal Image, Graphic Signal Process*** 11, no. 11, 2019. 1-9.

9. M., Sajjad, S., Khan, T., Hussain, K., Muhammad, A.K., Sangaiah, A., Castiglione, C. Esposito, & S.W. Baik. *"CNN-based Anti-Spoofing Two-Tier Multi-Factor Authentication System".* **Pattern Recognition Letters**, 126, 2019. 123-131.

10. S. Satpathy. "*Overcoming Class Imbalance Using SMOTE Techniques."* **Analytics Vidhya.** 2021 **Home Page**.

11. V. Seelam, A. K. Penugonda, B. P. Kalyan, M. B, Priya, & M. D, Prakash. "*Smart Attendance Using Deep Learning and Computer Vision*." **Materialstoday***: **Proceedings.*** **Elsevier**, 2021, 4091-4094.

12. P. P. Shinde & S. Shah. "*A Review of Machine Learning and Deep Learning Applications.*" ***Fourth International Conference on Computing Communication Control and Automation*** **(ICCUBEA. IEEE)**, 2018, 1-6.

Chapter Two

Literature Review

2.1 Conceptual Review

This section described concepts relevant to this study and outline relationship between them; biometrics technology and attendance system, authentication methods, database management and security, machine learning and empirical reviews were all discussed in details.

2.1.1 Historical Background of Biometric Technology

It has been reported in the years past that Biometrics are a set of biological measurements that serve as a means or medium through which individuals can be identified. For instance, issues of fingerprinting and mapping, the use of facial recognition mechanisms, as well as the use of eye retina scanning are reported to be viable forms biometric technology solutions commonly or most recognized approaches among diverse methods of individuals' recognition techniques[10]. Biometrics methods of individuals' recognition system have been in existence thousands of years ago. Biometrics have been around for over one millennium thus experiencing diverse form of synthesis in its classification of being an instrument of identity authentication using a wide range of modalities to predicting human behaviour through already registered biometric features.

In addition to this, the chronicles of biometrics dates back to as far as 500BC during the Babylonian empire, the actual and the initial record of biometric identification system was first witnessed around 1800s, in Paris, France. The first record was heralded by the work of Alphonse Bertillon, who created a special and a specific sets of instruments that were used for the cataloguing and measurement of a level of resemblance of criminals.

He developed a near perfect system that encouraged the use of biological characteristics in identity authentication.

It was opined by that the use of fingerprint in identification of diverse individual features popularly known as "Fingerprinting" came up 1880s[11]. This time it did not only serve as an avenue of identifying criminals rather a form of seal on contracts signing that confirms the authenticity of any document(s) that carries such signature. It was acknowledged at this time that a fingerprint was a clear representation of a person's characteristics and that at any point in time, one could be held responsible for it. Though there are different forms of debates on who actually commenced fingerprinting for recognition of individual's unique features. Edward Henry is recognized as a professional who instigated the use fingerprinting as a universal instrument often referred to as the Henry Classification System. This is seen as the system for identification based on the genuine designs of fingerprints. This system was immediately adopted by law enforcement agents thus quickly replacing Bertillon's system and becoming the standard for criminal recognition and as a standard for initiating the arrest of culprits. This set up a new century with high substance of research in the area of unique physio-biological characteristics can be applied for unique identification of individuals

Around the early 1960s, quasi-automated face identification and recognition systems were originated forcing administrators to examine and scan facial characteristics around the concept of the existing or real time image and extract meaningful and measurable feature points[12]. These developments require more of manual application and testing of images compared to the ones that are often used in opening our phones these days. As at 1969, the facial recognition and fingerprint were also extensively used by various bodies

of law enforcement agents in the United States of America[13]. Specifically, the FBI released funds in this direction for the purpose of development of automated system. This was, indeed, a massive encouragement for the improvement on the more complex and high-tech sensors for capturing biometric data, extraction and classification.

In the earlier part of 1980s, the Institute for National Standards and Technology, Maryland, United States of America came up with a dedicated Speech group who were saddled with the responsibility of studying and recommending the progressions for speech recognition technology. These works formed the basics and foundation for automatic recognition systems as well as voice command that are still being used until this day. More importantly, in addition to this work, in the year 1985, the concept of the use of facial appearance and features, fingerprints, irises, among others, had unique characteristics of every individual that was proposed. In year 1994, the very first and basic algorithm for reading and recognition of iris data was born and it commenced its inroad into the world of biotechnology. While this was in progress, then the discovery of pattern of blood vessels in the eyes were found to distinctive in every human being and could be used for authentication as well. By mid-1991, the facial detection tools were developed. This made it possible for real time image recognition to be read, captured, classified and identified uniquely for every human in question[14]. Though these developments were not without their challenges, it was an eye opener for renewed attention in the field facial recognition biometrics especially for the purpose of authentication and security surveillance.

Just around the end of 2000s, a very sizeable number of biometric authentication recognition methods and algorithms were fully and effectively developed and sold out in

commercial quantities in the United States of America. Government settings no longer develop biometrics system. The sale was done in high large quantities and were also sold at large scale within and outside the United States of America.

In the Last Decade, research in the direction of biometric technology has continued to move at jet speed with biometrics technology moving form just fresh technology to becoming part of our daily life. Then in the year 2013, Apple involved fingerprint as a means of unlocking iPhone. Thus championing a high level of acknowledgement of biometrics as a platform for authentication and identification. These days, most mobile phones in use now possess biometric features competencies and so many applications now go into the application of biometrics as an authenticator for everyday activities.

From our own point of view, looking at the speed of biometric technology and despite all the successes recorded in this field, potentials of biometric authentication and identification is not likely to be exhausted any time soon. As the wave of biometric research continues, continues, a merger between artificial intelligence and biometric technologies is seriously envisaged. This will grant a giant leap to the technology for the purpose of biometric devices and systems constructs that can acquire information, adapt to the supplied information and its users thus creating a seamless and contact-free authentication experience. The consistent growth of this technology brings an end to identification by proxy of individuals. One may now use himself as proof of his own identity, one may not necessarily need to carry around keys and cards anymore. The next big thing could now be a future that has a rightfully identified society with contactless transactions and automatic exchanges as well as secured access control.

2.1.2 Techniques in Biometric Authentication Methods

Biometrics authentication is evolving by the day and indeed becoming a field of controversy among scholars and researchers around the world. It is of the opinion that Civil liberties groups have seriously expressed diverse areas of concerns specifically in areas of privacy and identity issues[15]. Today, biometric laws and regulations are in evolving every day. Biometric industry standards are being tested for face recognition, fingerprinting, but with constant technological progress and with the threat of terrorism, researchers and biometric developers will stimulate this security technology for the twenty-first century.

It has been understood that various biometrics characteristics have been used in the recent times which are, but not limited to fingerprints, the eye (retina and iris), the vein pattern in the palm or finger, the face, voice and speech[16]. These have series of advantages unlike pass cards, keys and passwords, biometric data cannot be transferred from one person to another. They are much less inclined to fraud. There is every tendency that a person's identity can always be established with absolute assurance. They are also not easy to remove since they are physical characteristics. An attempt to do so may lead to destruction of physical characteristics in an individual. It is a convenient technology that produces a high level of security. The combination of different types of biometrics technology paves ways for a high security be it in applications or an environment.

Biometric features can be categorized into "two main classes - physiological and behavioural biometrics. Physiological biometrics are related to the shape of the body and thus it varies from person to person. Fingerprints, face recognition, hand geometry and iris recognition are some examples of this type of Biometric. The behavioural biometrics are related to the behavior of a person. Some examples in this case are signature,

keystroke dynamics and voice. Sometimes voice is also considered to be a physiological biometric as it varies from person to person.

Biometrics has a strong security characteristic. Recently, a new drift has been developed that merges human perception to computer database in a brain-machine interface[17]. This approach has been referred to as cognitive biometrics. Cognitive biometrics is based on specific responses of the brain to stimuli which could be used to trigger a computer database search.

In connection with the above, fingerprint recognition and iris scanning identification methods are perhaps, part of the most popular biometric security methods applied and often used by most institutions and industries in the last decade. This is due to nothing but its effectiveness and reliability in the recognition of specific characteristics in individuals. In some other places around the world, finger print, palm print as well as pattern recognition are also becoming very popular and relevant in present day research endeavours.

A. Fingerprint Recognition

Figure 2.1: Fingerprint[49]

The recognition system identified above generally looks for direct and unique identification characteristics in an individual that strictly goes in direct connection with fingerprint features as shown in figure 2.1, It in turn, looks for special, non-confusable and recognizable features on the finger of an individual. Such features, when found are treated as recognition item. The divergences, ridge endings and other machine noticeable features are combined together to form the lines of patterns that are recognizable by the system and makes it unique to every individual[18]. These are then kept in the system storage in order to form unique image that are saved as biometric images or pictures. This makes enrolment and identification phase easier than can be imagined

In the recent research outcomes, there have been diverse clearly stated challenges of capturing images of this nature. No matter how an image is stored, it still remains an image. Therefore, it can be compared. In principle, its code can be replicated. It is now a common practice to model fingerprints with the use of relatively easy and quite accessible technology. Alternatively, an important area to consider is that the actual finger utilized for the purpose of recognition does not importantly need to be a component of a body as at the time of the recognition. This is one of the giveaways of using this recognition methods[19]. Fingerprints may also get worn out as age sets in or in actual sense, as one ages, if one does a lot of DIY (Do it yourself) or one carries out most of the basic assignment by himself without little or no assistance at all in the area of biometric recognition and security based applications, it may become more difficult to recognize or read some individual's fingerprint (false rejection) or their fingers are so difficult to enroll or recorded. It has also been recorded that there are hereditary disorders that result in people being born without fingerprints.

On hereditary disorder called "Adermatoglyphia", - "*the Genetic Disorder of People Born Without Fingerprints*", this exceptionally unusual disease does not cause any problem on their own aside the unusual difficulties with the authorities/government. Joseph Stromberg, had this recent report for *Smithsonian Magazine*:

> *"In 2007, dermatologist Peter Itin was contacted by a Swiss woman with an unusual quandary: She was having trouble entering the U.S. because she had no fingerprints. Regulations require all non-residents to be fingerprinted when they enter the country, and the authorities were baffled when the woman said that she simply hadn't been born with any. When Itin looked into the case, he found that eight other members of the woman's extended family also had been born printless. Ultimately, working with Israeli dermatologist Eli Sprecher and other colleagues, Itin tracked down three other unrelated families that included people with* ***Adermatoglyphia****, which they dubbed "immigration delay disease," and successfully located the single gene mutation responsible in 2011"*

In the other way round, fingerprint identification is already now very popular and has welcome a high level of acceptance from different users. The technology is also very remarkably cheap. It should be decently noted, however, that the class of recognition system may vary from one system to the other, with an acceptable level of considerable disagreement between systems most especially in terms of false acceptance and false rejection rates.

Recently, the issue of biometric spoofing came to being and it is an attempt garnered towards the production of a falsified biometric features with the principal aim of being identified as a person that one is not[20]. This may sometimes involve the use of copied or falsified fingerprint or the use of replicated contact lens with a falsified iris pattern. This biometric risk is only application on superficial external characteristics.

b. Recognition of Faces

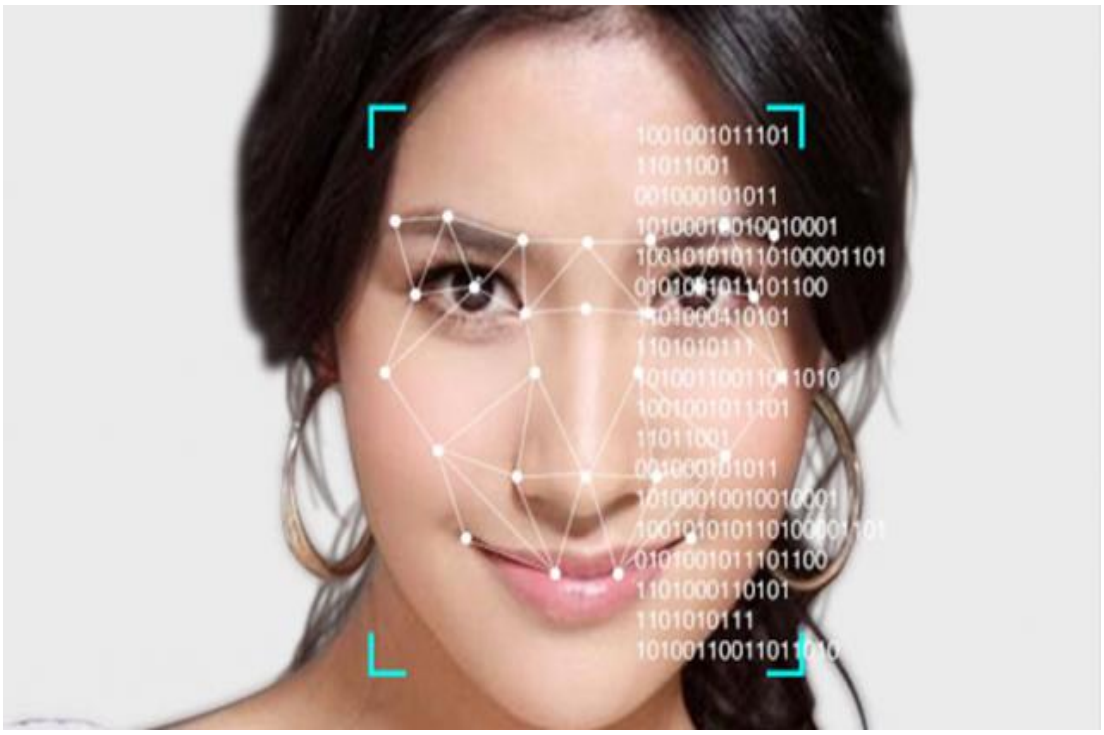

Figure 2.2: Face Recognition[50]

As revealed in figure 2.2, different parts and patterns as well as shapes on the face are properly recognized to determine a balance. What is also taken to account are what is referred to as surface features. These are skin, are also sometimes taken into account.

Face detection technology has been recognized as a result of using facial recognition technology for security purposes. This is, most of the times, used to identify faces in wide range of images. Technology has led to a different important results in the recent time and it is therefore a reliable method to use when a remote recognition and identification is needed[21].

Another inspiration in this direction is that this trending technology provides 'negative identification', or face exclusion. This makes it a more wonderful and individuals are

easily recognized among a large crowd. This advantage can be helpful when a suspicious individual is being trailed.

It is of the opinion that this technology is not without its own shortcomings[22]. A major example is its focus mainly on the face rather than other unique features around the face, that is, the hairline down. This means that face cannot just be captured on the move, the individual may have to look straight to the camera. This is just a shared rigidity which may need to be looked into. Secondly, the technology is still evolving. The degree of security it offers has not been able to match those of iris scanning or vein pattern recognition.

c. Iris Identification and Scanning

Figure 2.3: Iris Identification and Scanning[51]

This scanning method identifies special and recognizable part on an individual's iris which are shown in figure 2.3. These are in turn transformed to what is usually to an encoded or bar code. This is indeed, a much more dependable and reliable security

method. It becomes more secure when it is performed with the use of infrared light[23]. In the light of the above, the major setback with this method is the tough stubborn and most at times, an unexpected resistance from users. Only a very small number of people number of people often find it rather pleasant to have their eyes scanned. A particular stance or position may have to be maintained in order to make the scanner is able to read one's iris, because of certain uneasiness that may arise in the course of this. Hygiene is another frequently cited shortcoming, because many users are required to position their chin on the scanner's chin rest which uncountable number of individuals have used before their turn.

Lastly, as important as this method is, the concept of speed may be the opportunity cost of the high level of security in this system. Incidentally, systems that can read iris in a relatively short times have been developed in the recent times.

d. Finger Vein Pattern Recognition

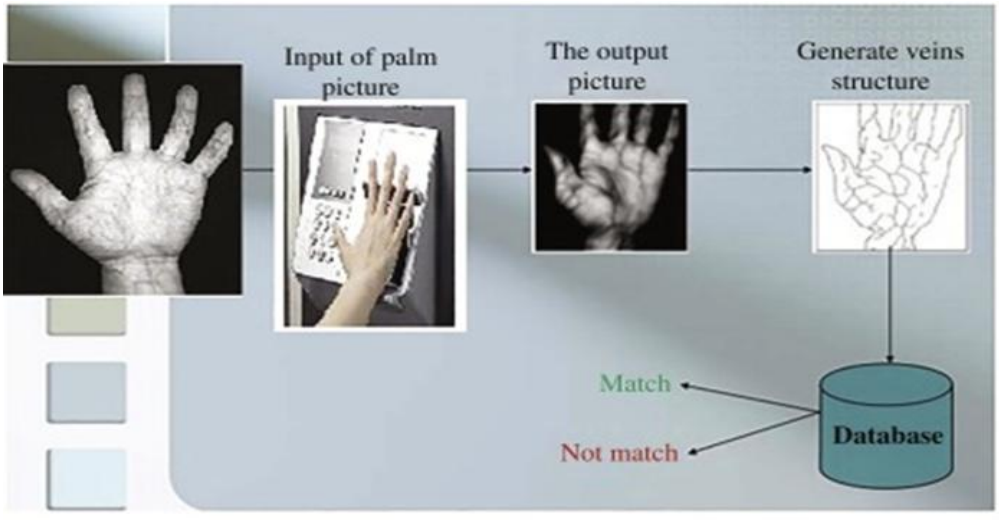

Figure 2.4: Finger Vein Recognition[52]

When a biometric method that captures the vein pattern recognition beneath the body is discussed, then all targets should in the directions of "Finger Vein Pattern Recognition". Figure 2.4 shows how this method works, ending points and divergences of the veins in the finger are essentially captured in the form of an image, digitized and converted into an encrypted code. When these methods are brought together, veins are found deep beneath the finger print rather than the surface of the skin. This is a proof that indicates the security of this technology and empowers it as a more secured biometric technology when compared with the fingerprint-based identification[24]. It is also very fast and more and more convenient for the users. Its major disadvantage at the moment is that it is expensive. It is very important to note that cold fingers and 'dead' fingers (especially those suffering from Raynaud's syndrome-a situation where it just becomes difficult to read through the veins of an individual's) extremely hard to read with the use of finger vein pattern recognition. For now, this seems to be one of the most recognized shortcomings of this method. However, it is still baffling that this technology has not been relatively explored in full thus making it relatively unpopular.

e. Palm Vein Pattern Recognition

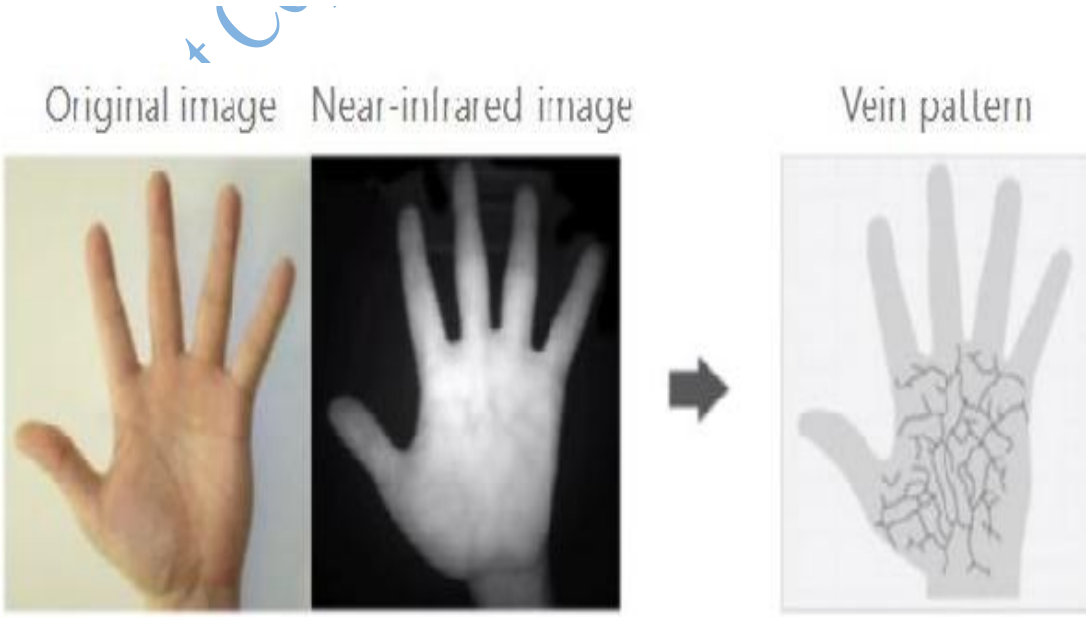

Figure 2.5: Palm Vein Recognition[53]

As displayed in figure 2.5, unique patterns in the veins are clearly recognized. Apart from this, more recognition points are taken into consideration when scanning is being done. This makes it a simpler and a rather more reliable and secured identification method.

This technology is unique in its own right. It is regarded as the most reliable in the field of biometric security. Palm scanning operates with high speed and extremely accurate and the methods also offers a much more reliable level of user convenience.

Palm vein pattern recognition methods are highly expensive. Hence the technology is often used within sectors that have a high demand for security such as banking sector, justice system as well as in government. There seems to be a mix up between this method and hand geometry. Though, it is an old type of biometrics which is tailored in the direction of the shape of the hand and it only encompasses very few exceptional features when compared to other fingerprint recognition methods.

Other Methods Include

Palmprint: This is essentially a very different and unique application of biometric technology. Its own scanning method uses optical readers that are relatively close to those used for fingerprint scanning. It is far better in size. This size serves as a limitation in its application in workstations and mobile devices but it can be one of the leading features if research is directed towards this method in developing a way to scan a smaller and view unique characteristics on it.

Hand Vein: This technology sees vein pattern as distinctive. The veins under skin vein is read with the use infrared light. The pattern taken through this method are darker and

deeper. This area of technology is still being explored. This technology will be a good security instrument if the breakthrough in the field becomes a reality.

DNA: DNA sampling is rather invasive at the moment because it requires a form of bodily sample that people find difficult to leave behind or donate. If there has to a remarkable breakthrough in this direction, the method for obtaining sample must be meaningfully improved. Though DNA analysis is now being made an automatic activity at least to a reasonable extent but must still be improved upon. DNA analysis of a typical human is now possible in less than 10 minutes. As research is being tailored towards this direction. It now becomes far easier to do DNA marching in real time. As of now, Biometric Systems DNA is majorly applied in crime detection as it has recorded more successes in this area. It will continue to be relevant in the law enforcement area for now because of its uniqueness and effectiveness in determining the unmatchable deep rooted features in every individual[25].

Thermal Imaging: This is closely related to hand vein geometry. It also makes use of an infrared source of camera and light principally for the production of the vein pattern in the face or in the wrist[26].

Ear Shape: This basically refers to as an identification of persons with the shape of ear. The law enforcement agent oversea has had to use this human features to detect criminals in most crime scenes. It is still seriously considered as one of the many human physical characteristics to be universally taken into consideration for use in identification of persons. Ear markings that are detected in most crime spots are deeply dissected and critically examined for the detection of criminals. This is an application for the future as more research endeavor is directed towards its success. The ear shape detector called "*Optophone*" was manufactured and patented by one of the French companies called *ART*

Techniques. It is a phone-like handset with cameras and lighting unit. It is designed to capture dual images of the ear per time.

Body Odour Biometrics: The technology used in this direction is directed towards the fact that each human being has his own unique smell. This biometric characteristic is normally retrieved with the use of special sensors that are capable of obtaining odour from open parts of the human body. These parts of the body are normally non-invasive. Parts of the body such as back of the hand, wrist area among others. Research efforts are still ongoing in this direction. One of the unverified cases in this direction is that "every human smell is made up of chemicals known as volatiles and they are extracted by the system and converted into template". It has also been postulated that there is possibility of diagnosing some diseases or undertakings in the last hours by analyzing one's body odour.

Keystroke Dynamics: This method verifies the uniqueness of an individual by categorizing them as trained typist and an amateur two-finger typists. It will be possible for the system to recognize them from the log-in stage. Their biometric systems can also be monitored a speculated period in order to identify them deeply than just a surface identification. These systems are cheap to install as only software packages are needed.

Signature Verification Technique: Its uniqueness is based on the chances that every signature is has its own genuineness. Rather than a strict assessment of the signature itself, certain unique features in the signature are taken into consideration and are seen as key areas around such signature. The uniqueness in itself can now be assessed from the point of *"pressure, direction, acceleration and the length of the strokes, dynamics number of strokes and their interval."* The strength of this method is that fraudsters will find it difficult to decipher the very techniques involved in getting a signature ready by just examining the ones written earlier. Different types of instruments are applied in

capturing the signature dynamics. A good number of those devices are more traditional tablets and at some other times, special purpose devices.

Summarily, table 2.1 below showed the characteristics strength and weaknesses of four (4) basic biometrics as portrayed by Dargan and Kumar.

Table 2.1 Overview of Characteristic Features of Biometrics as Detailed by Dargan and Kumar[54]

Characteristic	**Facial Recognition**	**Iris Scanning**	**Fingerprint Identification**	**Voice Verification**
Accuracy	Medium-low	High	High	Medium
Devices Required	Camera	Camera	Scanner	Microphone
Interference	Lighting, Glasses, Hair, Moving, age	Lighting, Inaccurate eye positioning in relocation to the device	Dryness, dirt, deep injury, age	Noise, colds
Attack's precaution	Average	Very high	High	Average
Verification time	About 3 seconds	Generally less than 5 seconds	Less than 3 seconds	About five seconds

2.2 Database Management System and Biometric System Environment

The concept of database is perhaps, one of the appealing and popular technologies in the information system community. It involves a diverse application of technical and managerial issues and features that are at the centre of today's information system. The collection of information that are organized in way which are easily accessible, managed and updated are known as database. While Database Management System (DBMS) itself is a software package designed majorly for the definition, building, sustainability and manipulation of data in the database.

Database management system involves in the manipulation of the data and the general format of data like the field name, record structure as well as the file structure in the database. Database management system takes diverse instructions from the other end of the database administrator and thus processes it accordingly. The instruction may be to update, retrieve or modify an existing data in the database. It can also be to store a new set on data which can then be manipulated by database administrator when necessary. For example, a user can update the GSM number of a client in a client database. Database applications are most widely used today in a business to model and manage business objects within corporate databases[27]. The revolution that occurs in 21st century exposed many of the corporate bodies and businesses to the application of database management system in their business operations though database security in authentication and authorization is still major area of concern till today.

It is of the opinion that development of database management system and the positive changes in the world of state-of-the-art computers came up as a result of society's recognition of the key importance of improved systems with diverse qualities that make their works faster and easier that what has been obtainable in the past. Features such as storing, managing and retrieving the rapid expanding volumes of business data are taken into consideration. In accessing database, some security features have to be put in place so as to control who is who in accessing the organizational data.

User Identification/Authentication: User that wants to access a data in the database must definitely establish a way to connect to the database management system through the use of database management application(s) and then sought to establish a user session with the database server. Such user must first establish a unique identity to the database server. This can be achieved by assigning some user identifier and other relevant and 'not easily

identifiable' entry codes in form of passwords, pins and other unique modes of identity for a genuine authentication which is done by the operating system to avoid unauthorized user from accessing data in the database. The traditional ways of securing such voluminous data such as password and PIN code can be easily forged and forget as it contains names, address and unique identification number. Despite their popularity, passwords are considered by a wide range of users to have shown a very low usability and low level security because they are dependent majorly on the users to remember them and this often make them difficult to remember. A means whereby an individual user does not need to remember any password or an identification PIN but is able to authenticate successfully is the use of biometric technology. It identifies and verifies an individual user based on the features used to enroll at the enrolment stage, having the features accepted at the matching process authenticate such user as the authentic user to be granted access to use the system.

System Privileges: This allows the execution of database operations such as create, alter and drop objects. This operation is sole assignment of a user often referred to as Database Administrator who can later allow other users the privileges to operate as he does[28]. Prior studies from see database as a way of assembling data into a highly organized collections and its high-level software that is in charge of databank itself and manages the database environment which is commonly referred to as 'database management system'[29]. Database Management System is software that makes it easy for organizations to centralize data efficiently and provide access to data application programs. It also acts as a bridge between application programs and physical data files.

Computer was invented because of the need to help in processing. Also it has served massively in the use for handling of high amount of data that have been accumulating for

years. Highlights the very opinion that people have been interested in data for thousands of years past while today the concepts of data are often associated with computer which brought about database itself[30]. The size and degree of complexity of database can be of any degree and its mode of maintenance may be maintained in a manual form or by the application computer software that makes it automatic and faster than the application of manual method. Traditional electronic database includes those of airline reservation system, many library catalogues, magazine subscription for most of the large scale publishing houses, patient tracking applications and systems in largest hospitals and inventories for supermarket and big box stores. These databases normally has its contents as text and numeric based while the newer trends of databases normally include different modes of data such as multimedia data and formulas for data and analysis.

When the discussion of effectiveness of a database system comes to play says that it is often measured by the degree to which it affords flexibility of access to data, integration and consistency of data with reduced redundancy of data in order to help in the management of some other enterprise within the organisation[31]. Any database management system used must give high level of flexibility of data access while the integrity of data is not compromised and its consistency is placed in high esteemed with zero data redundancy. This is because management will benefit a lot once this is in place.

2.2.1 Database Management System

Before the database concept was developed, recounted that all data in information system was stored in a single linear file. Some application and their programs required data from only single file while other required data from diverse files[32]. Several of the more complex application used data extracted from another file. Generally, files were created for a single application and were used only for that application. Barter and the use of

money to trade comes in which the keeping of records were encouraged. As time went on, data keeps increasing which are kept for future purpose such as census data, marriage records, inventories are tracked and so on. It was in 17th century that the issue of data started drawing attention of people to the fact that devices could automatically process data. Blaise Pascal in France during 1940s, produced one of the best-known devices called calculator while Jacquard's loom invented punched card as a means of storage device. In 1950s, when commercial use of computer started, a large organization in all the files, then each file would have different files for different purpose. If a file needs an update of information available in all the files, then each file would have to be handled separately and based on its merit. Data redundancy became rampant since data are not often shared among different program that needs them.

2.2.2 Basic Components of Database Management System

Universally, there are five basic constituents of database management system while others are used to extend the core functionality of the database and its environments. Each component components have a specific function in the database management system community.

The five (5) major ones include:

The Concept of Hardware: This refers to everything electrical and physical devices such as computers and their peripheral. It includes the input/output device, storage device and other visible component of the computer system. Databases are created, accessed and managed with the use of various machines from micro-computer to main frames. When any DBMS such as MYSQL runs on the computer system, computer parts like the mouse, keyboard, the memory, hard-disk and so on are used as DBMS hardware.

Software: It refers to well guided multiple lines of programs used for the management and control of the entire database and its structure. This includes the database software, operating system, network software used to share the data with other users, and the applications software which helps user to access the data.

Data: Data are basic raw facts that need proper organisation and processing to make it more meaningful. Database dictionaries are used to consolidate document. It is used for control, and coordination of the use of data within an organization. A database is a repository of information about a database (also called metadata).

Procedures: Procedures are the instructions used in a database management system. It involves everything from instructions to setup and install, login and logout. It involves the management of the day-to-day operations. It takes backups of data, and generate diverse reports.

Database Access Language: This language is regularly used to write commands for the purpose of accessing, updating, and deleting of data stored within a database. Users may write commands with the use of diverse existing Database Access Language before they are submitted to the database for execution. Albeit utilizing the language, users can as well create new databases, tables, insert data, and can also delete data.

2.2.3 Types of Database Management System

Database Management System is a collection of software application used to define, manipulate, store, retrieve and manage data in database. It involves in the manipulation of the data and the general format of data like the field name, record structure as well as the file structure in the database. Database management systems receive instruction from the database administrator and process it accordingly. The instruction may be to update,

retrieve or modify an existing data in the database. It can also be to store a new set on data which can then be manipulated by database administrator when necessary. There are four main types of database management system which include: relational database, hierarchical database, object -oriented database and flat database.

Hierarchical Database: This is one of the oldest and the simplest of all the database models. In hierarchical database, accessing and updating of data is very fast because the relationship has been predefined. It fixed itself well to the tape storage system used by mainframes in 1960s -1970s and it is still in some types of passenger reservation system and inventory and account system by many banks, companies, hospitals and government department.

Distributed Database: this is made up of two or more files placed in diverse sites. This may be within the same network or on an entirely different networks. What is most important here is that parts of the database of are stored up in different but multiple physical location where processing is distributed among workstations that are normally referred as the database nodes. Some of its storage devices may not be attached to common central processing unit rather, its data may be stored in a multiple computer located in the same physical location or may be spread over a network of computers. In this database, the database server is created and located at different locations rather than at a single location. It is very possible to manage and define the core of an organization's information centrally while the local branches only hold on data that are relevant to their basic operations at the local level. Organisations tend to function effectively when it learns to manage its information in a meaningful and an organised manner. For example, branches can access the central core information in a controlled manner while the

headquarters can easily access data of all other branches in order to have information about the overall performance of the organization.

Network Database: This type of database uses network structure to create association amongst certain and different entities. Its application is more relevant on large scale digital computers. In this database Management system, a network mode can have a relationship with multiple entities and it was invented by Charles Bachman. Examples are Integrated Data store (IOS), Integrated Database Management system and so on.

Relational Database: This is a database that uses table to represent both the data and the relationship which exists between the data. It has Data Manipulation Language (DML) and Data Definition Language (DDL) embedded in it. The relational database model is primarily used today for commercial data processing application packages which are used for storing their personnel data, financial records, and many more. It is preferred to other data models like the network model or the hierarchical model due to its simplicity and it is easy to handle by programmer.

Object-Oriented Database: This Database uses object-oriented programming techniques such as polymorphism and inheritance which makes the interaction between the objects, a inconsequential task. It helps to deals with complex data structure. Data is stored as a collection of queries and objects than can easily be performed by following the pointer from parent object to its children.

Database Management System has its own functions such as:

Maintaining Data Integrity: DBMS reduces data redundancy or repetition and increases data consistency hence information appears just once thereby freeing up more storage capacity.

Data Storage Management: The newer database management system has the capacity of warehousing various data types ranging from text to video and images. DBMS has a procedure and processes that ensures that data is stored accordingly as the data is keyed into the system.

Data Security: This is the process involved in securing data from an unauthorized user(s) to ensure the availability, confidentiality and integrity of the data.

2.2.4 Database Security

This is often referred to as the collection of procedures used to safeguard and defend the security of a database management software from illegitimacies. This varies from malicious practices and attacks, unauthorized access, and any activity that poses a threat to the workability and integrity of the database. This security comprises of diverse and high number of procedures, tools and advanced methods and methodologies that certifies the security and integrity of the database and its environment. It houses and imposes total and diverse methodical security on all areas and components of database that is stored in the database, database server, database management system and other database workflow application.

2.2.5 Analysis and Implementation of Database Security

The method of restricting unauthorized access and use is often a very complex process that involves adequate care with principles and processes that are properly following through the implementation of a very reliable and securely controlled access and data management.

2.3 Identity Authentication Systems

Systems for data or system safety that need inputs for access or functioning are known as authentication systems. Therefore, methods and processes used to verify whether a person or object is who or what they say they are fall under the category of authentication. By tying user inputs to data in the database for correlation and verification purposes, authentication systems could also offer access control, ensuring a secure procedure and system. When a user enters credentials, such as a password, for user identity reasons, a user identification number is distributed. This number must correspond to the user's information stored in the database. A two-factor authentication system would need an additional input factor, such as a special code, whereas a single-factor authentication system would only utilize the user identification number and password. Multifactor authentication refers to the use of biometrics in addition to a user ID, password, and maybe a personal question. The terms authentication and authorization are frequently used interchangeably in literature, and even though they are frequently combined, they serve two different purposes and have different characteristics.

However, authentication is the process of verifying a user's or processes purported identity before allowing access to networks or systems with significant security hazards. Additionally, authorization is a more precise process that verifies the veracity of a claimed user identification before approving access. Access control refers to how tightly a resource's access is restricted while yet being regulated. As a result, the authorization step

comes after the authentication process. The usage of single factor, double factor, or multifactor authentication systems ensures that only authorized users are granted access to resources. In deployments that simply require a user ID and password for sign-on, single factor authentication is more common. Organizations use authentication to regulate access to networks and resources from an information technology standpoint. Businesses also use verification to give remote workers safe access to their networks and applications. As was already said, several factors, such as user identification and passwords, are employed for authentication; nonetheless, these may not be considered to be a robust authentication system. Therefore, a system that employs at least two authentication factors and requires the use of strong passwords with at least eight characters, including special symbols and digits, is known as having strong authentication. An authentication factor is a data attribute that is used to verify and genuinely confirm the rightness of the identity of a person who requests access to the application or the system on ground.

2.3.1 Authentication Factors

Knowledge-Based Factors

The knowledge-based factor, often known as the "something you know" factor, consists of authentication credentials that require knowledge that a user knows, such personal identification number, user-generated password, username, mother's maiden/secret name, etc.

Possession-Based Factor

The term "possession-based factor" refers to something that the user possesses, such as item-based credentials that they personalize and take along with them, such as hardware devices, security tokens, or even mobile devices that were used to send texts or create PINs or one-time passwords.

Inherence-Based Factor

The inherent factor is anything about the user that is typical of a biometric identity system that includes fingerprints, facial recognition, thumbprints, retina scans, and other similar features.

Location-Based Factor

With the aid of global positioning systems that have varying degrees of accuracy, the location factor is frequently used as a plus to other factors. The location factor is rarely sufficient for authentication on its own, but it supports other factors by offering a way to reject a few other requests. For example, it might stop an unauthorized user from impersonating a user if they are situated in a certain remote location.

Time-Based Factor

The time-based element, which refers to when you are authenticating, is comparable to the previously described location-based component. It is frequently used, nevertheless, as a supplemental strategy to close up an attacker who may have intention to access a resource within a particular period of time during which the resource is not actually accessible to the authorized user. This is a frequent practice that is done in conjunction with the location-based factor. Even if the location- and time-based elements are distinctive and effective, they are not sufficient for a generally reliable authentication system; therefore, one of the first three factors must be combined with them in order to authenticate a user.

2.3.2 Approaches to Attendance Marking

The most challenging aspect of putting in place an attendance monitoring system is making sure that all the methods are current. A few of the most popular technologies used in attendance management systems in this post, along with how they can benefit systems is discussed in this section[33].

2.3.3. Biometric Logins

In recent times, to login in through biometric means have become regularly used[34]. The features of fingerprint are extracted and scanned, and if it there is positive match with the fingerprint data on the system, access is granted and the attendance is recorded[26]. In our opinion, these technologies should be perfect in the area of protecting any valuable process, system and or technology in order to replace the old login and logout out system. Nevertheless, they represent a powerful tool of avoiding forms of cheating that would naturally take place among diverse human nature. Attendance by proxy is also carefully avoided[35]. This is because personal identification is a crucial instrument in determining an individual's genuine identity. The identity of a person can be verified using biometric or conventional techniques. Traditional approaches can be divided into two categories: token-based and knowledge-based identifications. Token based identification identifies an individual through the possession of physical identification instruments such passport, driver's license, identity (ID), Credit card, among others. Although possession of these identity instrument can be very convenient, but may be vulnerable to copying, theft, and even loss.

Alternatively, the knowledge based method depend solely on the knowledge of our authentication details such the use of password and/or personal identification number (PIN) created previously by each user for proper unique identification. However, it may be very easy to forget. This is mostly experienced under the similitude of multiple logins

into the different applications within the same environment. Adopting biometrics is another alternate strategy that takes into account a person's unique physical or behavioural traits. Physical traits relate to attributes that are naturally present in a human body part. These consist of the iris, the face, and the fingerprints. Contrarily, behavioural characteristics focus on traits seen in human behavior and related gestures such Gait (a person's way or manner of walking), voice, and signature are a few examples of human gestures and related action. The above-mentioned issues with traditional approaches can be resolved by applying biometric techniques. Currently, biometrics are used across a wide range of industries. In 2018 research, German and Barber from the Center of Identity at the University of Texas at Austin identified financial services, technology, and government as the top three industries that use biometric technologies[26].

This is accompanied by the employment, leisure, and healthcare sectors, with education seeing the least usage. The technology of biometrics is often applied in the diverse offices and in educational institutions for behavioural analysis, access control, and attendance tracking. The salary of an employee and the eligibility of a student to take an exam are both based on attendance records. Only individuals with the proper authorization can enter a building thanks to access permission. Additionally, behavioural biometrics is used to screen each individual's level of concentration in a classroom or an office setting. As shown at *Universal Studio Hollywood*, which requested tourists to leave their fingerprints in order to enter the theme park, the recreation sector has also embraced biometric technology. Another illustration can be found in Wuzhen, a historical theme park in China. To enter numerous attractions, visitors must register using facial biometrics.

RFID (Radio Frequency Identification) Tags

Employees are issued personal I.D. cards with RFID tags embedded in them, in this approach[26]. There exist built-in antennae. With this built in antennae, the card its reader is able to interconnect perfectly. The card reader is turned on by the card's chip, it also confirms the card as being active and then marks the user's attendance if the card is brought inside the EMF (Electro-Magnetic Field) of the card reader. The card user now has access to everything they were using it for. RFID tags are significantly safer than biometric login since they operate on a specific frequency. However, it is simple to copy RFID tags, and the necessary equipment may be purchased from any online vendor. Additionally, RFID cards are pricey and difficult to replace if they are misplaced. Frequency changes made on a regular basis can help to prevent these issues. One of the most important responsibilities in a school, college, or university is managing attendance[36]. Teachers can monitor their pupils' activities thanks to daily student attendance. Additionally, it helps parents determine whether their kids are paying attention in class and are regular attendees. However, it can be difficult for teachers to keep track of thousands of pupils' attendance each day. Traditional methods of recording daily student attendance can lead to mistakes and a lot of manual labour. It automates the students' attendance process and makes it simple for teachers and parents to track and monitor students' activities, an RFID-based attendance system can be a fantastic way to tackle such difficulties.

Bluetooth Attendance System

Although it is a relatively new idea, a Bluetooth attendance system is nevertheless useful. Employee must have an application installed on their mobile device for it to function. The staff must also keep their Bluetooth turned on after the attendance process has started[37]. A

different Media Access Control (MAC) address is assigned to each device. This enables the manager to avoid any possibility of duplication, mistakes, or fake attendance.

Bar Code Scanning

A barcode scanner operates by, well, scanning barcodes, much like how RFID technology functions. These barcodes, which are used to track attendance, are often in the form of stickers applied to employee identification cards[38]. These barcodes are typically quite safe and function similarly to the ones used to identify foods. These are also inexpensive to repair and replace as necessary. This technology will allow teachers to preserve all of their class time and eliminate the need to take attendance during each class period. To check attendance, both students and teachers will have individual logins. The teacher will have access to change or edit the student information[39]. Parents of absent students will be notified by email and text message if their child is not in class. The cloud will be used to store all of the student's data. The teachers would also have quick access to the data stored in the cloud. Teachers can access the system online from any location; they are not required to do so from their actual school or college.

GPS Technology

Has clearly identified global positioning system (GPS) is a feature of company-issued mobile phones and tablets[40]. This enables the business to monitor the device's position and location to make sure an employee isn't lying about where they are[26]. However, it's incredibly simple to trick this technology. Location spoofing apps are widely available, and some of the better ones can even somewhat imitate how people naturally move their maps. This is a straightforward deterrent; therefore, it will stop someone who is determined to get around it. These days, tracking attendance and calculating working hours is crucial for practically every organization. The software will be a GPS service-

based application that will enable us to determine the staff members exact geographical position based on their most recent location.

Facial Recognition

Another specialized recognition technique in our age is now in the world of identity management is Face recognition. In his observation [41] addressed the issue of face recognition as though it is not often used for attendance management. Its usage for the purpose of security management and prevention of illegal access is however becoming more popular in recent times. Most senior members of various organizations often see it safe means of avoiding unauthorized access. Biometric facial recognition is widely employed in our environment. “An individual can be recognized by their face traits thanks to this sophisticated, automatic, and practical identification method. A digital camera is used to take the face image, a computer is used for processing and analysis, and an output device is used to show the identification outcome. Any person can easily be recognized from various facial photos[26]. The face recognition system is a quick and trustworthy piece of technology. Because it can accurately identify persons, this biometric is incredibly safe. While looking at the development of face recognition, the face recognition is regarded as the first step in developing biometric access control-based application scenarios, in which the persons' biometric traits are extracted [42]. Numerous institutions as well as government agencies and other organizations make use of this technology.

2.4 Types of Authentication Systems

Password-Based Authentication Systems

The time-based component, which pertains to when you are authenticating, is comparable to the location-based factor discussed before. It is frequently used as a supplemental technique, though, to sift out attackers who plan to access a resource at a specified time when the authorized user is not allowed to do so. Along with the location-based component, it is frequently used. Despite the distinctiveness and effectiveness of the location- and time-based factors, they are not adequate by themselves for a largely trustworthy authentication system; therefore, one of the first three factors must be combined with it in order to authenticate a user. Furthermore, in today's modern environment, apps that access resources across various systems would need repeated or duplicated authentication processes. Smarter usernames and passwords anchored on regulations that provide a set minimum password length and complexity, which encourages the use of capital letters and other unique symbols, are just a few ways to reduce these restrictions.

Two-Factor Authentication Systems

By requesting a second verification factor from the user in addition to the password element, this type of user authentication adds an additional layer of security to the entire procedure. The user of such systems must provide a one-time password verification code, which is frequently received via text message through a pre-registered mobile phone and device or through a code generated by a verification application.

A Multifactor Authentication System

Under this authentication system, users are expected to present more than one authentication factor to complete the verification process. These factors include a

biometric factor that ranges fingerprint to facial recognition element and a possession factor, such as a security key that is personal to the individual involved. A token generated by an authenticator application will also be one of the major deals that makes this authentication system much more reliable.

One-Time Password (OTP) System

The term "one-time password" refers to a password that is automatically created and consists of a string of numeric or alphabetic characters for user authentication. The password is normally offered to new sets of users or those users who are finding it difficult to remember their passwords and are instead issued an OTP for login. The password is typically good for just one login session, or transaction, as the case may be.

Three-Factor Authentication System

Three authentication factors are used in this particular sort of multi-factor authentication system. A knowledge factor (like a password) is regularly paired with what is often referred to as the possession factor. This might be a security token or a hidden number known to you and an inherited factor (like a biometric feature).

Biometrics

This kind of authentication system is frequently tabled as an alternative or a second or third authentication factor, as well as a standalone authentication system. The fingerprint, face, or retinal scans, as well as occasionally the voice for voice recognition authentication systems, are common biometric traits used for authentication systems[3].

2.4.1 Biometric Authentication System

As noted, many authentication factors can be used as a two-factor or a multi-factor authorization access control system to secure an entity against unwanted attacks. The biometric authentication system, however, provides a more sophisticated and dependable way for user authentication. Biometric system authentication uses distinctive physiological or behavioural characteristics of a person as an attribute for verification and authentication rather than depending on a common secret or key. Therefore, it becomes a very easy task to uniquely identify genuine users with unique biometric features, such as fingerprints, ear shapes, irises, etc**.** The biometric system might include a biometric sensor, such a fingerprint reader, that reads in the biometric data. A feature extraction system extracts the most pertinent feature set that accurately and comprehensively captures the user's data points. A biometric authentication system based on fingerprints takes into account the minute characteristics that make a fingerprint unique. For the purpose of authentication, a matcher correlates the derived properties with versions in the database. The level of user authentication is based on how closely the recently extracted sample matches the pre-stored template in the database. Because biometric features cannot be lost, forgotten, or stolen, biometric systems are thought to be more dependable than conventional password-based systems.

Any of the following biological characteristics could be used to install various types of biometric authentication systems:

i. Eyes – extraction of iris attributes through scanners
ii. Eyes – extraction of retina attributes through scanners
iii. Fingerprint attributes
iv. Hand geometry attributes

v. Facial recognition attributes

vi. Vein recognition attributes

The following eight categories of biometric system attacks can be used to classify the vulnerability issues with biometric authentication systems, including those related to accuracy, such as false positive and negative matches, etc.

i. Fake biometric including a falsified fingerprint attributes through reproduction systems

ii. Replay attack that involves replays of an old recorded signal that may include presentation of an old fingerprint image copy for authentication

iii. Use of Trojan horse for an override feature extraction to produce new feature sets of the genuine user.

iv. Override matcher is used by attackers to tamper with the system matcher in order to produce artificially high or low match rates

v. Replacement of extracted feature set with a synthesized one with same feature representation

vi. More often, attacker tamper with the database holding the template features through modification of inherent features for a fraudulent authorization of an individual

vii. Communication channels are sometimes attacked to tamper with the template enroute storage database to the matcher like a man-in-the middle fraudulent attack

viii. Final authentication decisions could likewise be changed as a decision override attack when the result of the tested recognition system is falsified at the very last step

2.4.2 Iris Biometric Recognition System

Iris recognition certainly resides in the method of recording a high quality image of a person's iris though the use of visible and near-infrared light for identification purposes. Iris recognition falls into the same category as fingerprint or facial recognition technologies as a form of biometric system use case for user identification[9]. Through the use of scanning technology, professionals compare the iris of suspects with a pre-registered set in the database to determine or validate the user's identification. For the ensuing verification technology, iris data might be obtained covertly even without a person's knowledge. Iris scanning analyses the distinctive patterns in the colourful circles that make up a person's iris.

By detecting and excluding other unintended eye features like eyelashes, eyelids, and specular reflections that naturally block sections of the iris, the scanners first illuminate the iris with infrared light that is invisible to the naked eye in order to uniquely identify and keep record of certain patterns that are not readily visible to human eye. Various outcomes of this is a collection of pixels that are unique to the iris. For the purpose of extracting bit patterns that encode iris information, the eye line patterns and colours are further examined.

For the purpose of authentication, the extracted bit pattern is converted to digital form and put through what is akin to a one-to-one matching with already-existing pre-stored instances in the database. Iris recognition scanners are known to collect 240 or more biometric features. This combination is quite unique to each eye. The extracted data is later converted into a digital format by the scanners. This digital database contains the numeric representation of the data that was taken from the iris image. A more effectively managed system that would minimize compromise is needed because the system of

having pre-stored digital features of the iris for user verification through matching, as previously described, is believed to be susceptible to manipulations, especially when the third party database is compromised[9].

In light of the aforementioned, predictive analytics has been largely considered to be a superior technique for managing biometrics. Since the resulting intelligent model is a mathematical function that encapsulates all components without a clear definition or any storage system for dataset contained therein, the concept—which comprises statistical techniques, machine learning, data mining, etc. promises an authentication system that cannot be manipulated.

Right and left eyes each have different iris patterns, which proves that a person's iris traits vary not only between eyes but also inside an individual. The technique has several benefits, including the fact that it is a very accurate method of identifying people[8]. Others include:

Accuracy

Iris Recognition is one of the most accurate biometric authentication methods available, Iris Recognition technology is the most accurate, achieving the highest matching accuracy evaluation in the Iris Exchange (IREX) IX assessment conducted by the U.S. National Institute of Standards and Technology (NIST).

Contactless

Although closeness to the iris scanner is required, iris recognition offers a contactless solution that makes it less invasive and hygienic to operate[8].

Flexible and Scalable

Iris recognition is incredibly flexible and scalable. Scanners can be utilized in the dark or at night thanks to the employment of an infrared camera. Iris Recognition has been considered and applied by various organisations and governments all around the world because of its adaptability and great scalability.

Liveness Detection

It is one of the areas in which NEC is a pioneer. The iris scanner's technology uses movement of the iris to identify a person's liveliness and lower the chance of falsified access by producing high-quality static images[8].

Fast Matching

Iris Recognition has also been recognized as one of the quickest methods of biometric identification if a person is already registered with the system.

2.4.3 Fingerprint Biometric Recognition System

Another application of biometric authentication is fingerprint authentication, which uses a user's fingerprints to uniquely identify them based on their biological characteristics. For forensic criminal investigations, financial services, digital identity systems, and other use cases, the notion of finger print authentication has been applied. It is difficult to fake or alter a fingerprint for harmful intentions because of its reputation as an inherent factor, or "what you are." The combination of fingerprint input with additional elements like a password or PIN guarantees the best security system and user experience[4]. Benefits of this combination include:

i. Measured as one of the safest validation methods
ii. Much tougher to false than identity cards
iii. Convenient and easy for the user

iv. Can’t be predicted, forgotten, or misplaced

v. Most accepted and mature biometric authentication method

A new, reliable biometric authentication mechanism has been created for handheld electronic devices as a result of the quick improvement in sensor technology. Utilizing next-generation sensor consumer electronics (CE) devices and classifiers like Random Forest (RF) and Hidden Markov Model, it has been implemented by intelligently combining the movements of the finger in three dimensions with mental activity (HMM). TRSG (True Random and Timestamp Generator) has been used in a revolutionary biometric identification system that uses FPGA to perform quick and accurate fingerprint authentication. Systems for reading fingerprints inspect a finger put against a flat surface. The ridges and valleys on the finger are scanned, and a collection of unique locations where the ridges and valleys meet or part are referred to as minutiae. These small details are what the fingerprint recognition system compares. The dips (minutiae) and ridges on the surface tips of a human finger are utilized to identify a person. Numerous biometric applications have used face and iris identification. The combination of both biometrics offers various benefits in addition to enhancing verification performance, including expanding user population coverage and lowering enrollment failure. The second approach is to use a classifier like Fisher's discriminant analysis or a neural network with radial basis function (RBFNN) to determine whether the matching distances of the face and iris classifiers are genuine or an impostor. A well-known biometric technology, fingerprint recognition is a common security feature in the most recent smartphones.

There have been comparisons between these two biometric characteristics since Microsoft added iris recognition to its smartphones. Both of these biometric technologies, as well as their capabilities and security aspects, will be covered. When compared to other biometric

qualities, iris and fingerprint identification both have superior accuracy, reliability, and simplicity. Due to these characteristics, iris and fingerprint recognition functions more effectively and offers a viable security solution in modern culture[25]. The iris and fingerprint images are first captured, and any noise effects are then removed during pre-processing. The differentiating characteristics are then mined and compared to determine whether the two feature sets are identical. The judgment module uses the matching scores produced by the various recognizers to determine if a person is real or an impostor.

Everybody has a different iris pattern, which is continuous over the course of a person's lifespan. In the middle of the eyeball is a circular, black disc always regarded as the pupil, which in turn, enlarges in the presence of light and shrinks in the presence of darkness. As a result, the pupil's size varies depending on how much light it is exposed to. The iris, an annular ring with a great deal of intricate intricacies, is situated between the sclera and pupillary boundary. The physical structure of the iris is also incredibly data-rich and carries the floral design that is particular to each person. With age, this trend doesn't change. The rest of the acquired ocular image is used to extract this particular floral design, which is then converted into a strip and subjected to a pattern matching algorithm[25]. The acquired iris image must be rich in texture because this will affect every phase of the iris recognition process. Multi-scale quadrature wavelets are utilized to extract the iris' structural information in a suggested automated iris recognition system. A 2048-bit iris code is created, and the difference between two iris representations is assessed using the XOR operator to evaluate their Hamming distances. Zero-crossing representation of a 1-D wavelet transform that describes the texture of the iris has been determined at different resolution levels of a virtual circle on an iris image[25].

2.4.4 Features of Biometric Attendance System

Fingerprint Storage Capacity

The biometrics system requires storage and validation of a large number of fingerprints. The gadget should be able to save all of the students' fingerprint information as the organization expands and admit more students.

Durability and Doggedness

A sturdy and long-lasting gadgets are required for the system. This means that they ought to be resistant to severe temperatures, mud, and dust.

360 Degree Verification Angle

No matter if they are entering or leaving the premises, employees will always be rushing. In order to swiftly identify employee fingerprints irrespective of how their fingerprints are placed on the device's optical sensor, the desired biometric time trackers must offer a 360-degree verification angle so that the rate of effectiveness can be of high degree.

Access Control

An attendance system with biometric capabilities can be used in area of security considerations in addition to time tracking. It stops unauthorized individuals from accessing secure locations. As a result, one must take into account gadgets with access control features.

Fingerprint Identification Time

When choosing a biometric system, one of the most crucial factors to take into account is the accuracy of fingerprint identification. The biometric sensor should have the ability to instantly detect and transmit fingerprints. This is not a complete list of biometric

fingerprint attendance system features. Other features include interaction with the payroll system, Internet access, battery backup, and after-sales support from the system manufacturer.

When utilizing a biometric fingerprint attendance system, organizations must consider a number of crucial factors, including:

Data Safety

The most crucial consideration when utilizing a biometric attendance system is data security. Since the system leverages the employees' private information, data security needs to be guaranteed to avoid identity theft.

How Usable Should the System Be?

Any deployed system is expected to be clearly simple and reliable. The system must be simple to use, efficient in accomplishing the task at hand, and offer the optimum user experience. Any business must do this in order to accomplish its objectives and obtain the necessary results.

2.4.5 Importance of a Biometric Authentication System

Due to the numerous advantages they provide to enterprises, the usage of biometric time and systems developed in attendance related direction is expanding. A biometric system uses a person's fingerprint, face, or iris to recognize him or her and determine if they are authorized to enter the grounds of a company, building, etc. In addition to access control, it aids in monitoring staff attendance and working hours. The following points highlight the significance of putting in a biometric attendance management system[43]:

Prevents Proxy Attendance

It often assists to avoid proxy attendance and makes sure that persons with permission are allowed to enter any place[44]. Since each person must share their fingerprint, iris pattern, or facial characteristic in order to be detected by the biometric system, one employee cannot promote the attendance of their coworkers.

Ensures Accuracy of Data

The HR software of a company can be directly connected with biometric time and attendance systems. Furthermore, the system will provide correct time and attendance information. Calculations done manually are prone to mistakes[45].

Payroll Deduction

A biometric system not only makes the process of processing employee salaries simpler and more efficient, but it also makes it possible to maintain accurate leave records because it stays connected to both the access control and HRM software[16].

Low Cost

A very economical technology is a biometric fingerprint attendance system (NEC 2022). Since most employees lose their chip cards and have to buy new ones, chip cards frequently wind up being pricey. It is clear from the foregoing that installing a biometric attendance system has many advantages.

2.4.6 Use Case Industries of Biometric Authentication Systems

The world of biometrics is rapidly evolving, as is the rate of adoption. Systems for tracking identity entry and exit times that use biometric fingerprints are recommended and regarded as the most effective and cost-effective option. The following are some sectors that heavily rely on biometric fingerprint identification methods[46].

Banking

Developers of banking software make considerable use of this to precisely identify customers of banking services. This technology is also used to access ATMs and in regions where KYC compliance is required, in the areas where PINS and cards are preferred[25].

Government Agencies

Government agencies' in charge of public safety often use biometric systems for visitor management and citizen identification in safe zones like border control, prisons, sensitive department entry points, locations where public functions are held, voter registration, among others.

2.5 The Concept of Machine Learning

In order to deploy machine learning techniques as set out in this study, it is crucial to first have a comprehensive awareness of the workings of machine learning methodologies, proficiencies, and limitations[47]. Machine learning (ML) is the process of creating intelligent computers that automatically increase their knowledge based on experiences found in the dataset that was used to train the computer. It is at the foundation of data science and artificial intelligence, which covers clustering, classification, and data analytics. It is one of the technical domains with the fastest growth rates, following computer science and statistics.

The creation of sophisticated algorithms and theory, as well as the accessibility to massive data for computer intelligence, have spurred recent advancements in ML. Improvements in evidence-based decision-making across use cases, including the health industry, manufacturing sector, academics, financial technology, security, and marketing,

can be seen as a result of the implementation of data-intensive machine learning frameworks in science, business, and technology[10]. What fundamental statistical, computational, data-evidence, and theoretical rules govern all learning methods, including computers, humans, and organisations, are two primary consistent concerns that the machine learning discipline addresses.

The field of machine learning (ML) has advanced rapidly from academic curiosity to a practical technology with broad industrial use. The preferred way for creating practical software for computer vision, speech recognition, natural language processing, robot controller, and other top-tier applications has arisen within the Artificial Intelligence (AI) subfield of computer science. AI system developers are aware that training a computer system with instances of anticipated input-output behaviour, as opposed to doing so manually, is simpler[48]. This enables the computer system to predict the desired response on all likely inputs. A variety of businesses managing data-intensive paradigms, including consumer and public relation services, fault detection in complex systems and the medical field, recommender systems, logistics and retailing, etc., have benefited greatly from the favourable effects of ML. Based on the size and purpose of the data mining study, ML is essentially separated into two distinct approaches: supervised and unsupervised learning.

The terms "Classical Machine Learning" frequently refer to both supervised and unsupervised learning. A machine learning algorithm uses data supplied by an analyst to calibrate its parameters, learning (fitting) the model and refining it as additional data is gathered. The inscription "supervised learning" arose from the fact that analyst gives the computer's algorithm a training set with clearly labelled input variables and clearly

labelled output or expected variables, guiding (and hence supervising) it as it calibrates parameters.

Consider anticipating the relationship between market returns and other macro variables, such as the price of oil, the value of the dollar, volatility, consumer mood, rates, etc. To calculate the beta of market returns to these variables, an analyst would use linear regressions in conventional statistics. An analyst could compute exposures using sophisticated regression models that take into account outliers, deal with a large number of variables in a robust manner, distinguish between correlated input variables, take into account potential non-linear effects, etc. through the use of machine learning. This is made possible by new algorithms created by computer scientists for linear regressions[20]. One extension, lasso regression, for instance, selects the smallest, most essential group of input variables.

Another technique, known as logistic regression, is made specifically for handling data and produces outputs with binary values, such as "buy" or "sell." Understanding the variables influencing asset price is another frequent task in financial analysis. For instance, an equities analyst would try to link stock returns to those of the whole market, a particular sector, a particular fashion trend, etc. An analyst would use a method called Principal Component Analysis in traditional statistics (PCA). Machine learning is compatible with PCA (or related techniques like Independent Component Analysis - ICA) without any modifications. In the context of machine learning, algorithms like PCA fall under the heading of "Unsupervised Learning.". The phrase "unsupervised learning" derives from the fact that the computer is simply provided with the whole set of asset return data; it has no concept of independent variables or outputs or dependent variables.

Another approach to unsupervised learning is clustering. It entails grouping together smaller sets of variables based on some premise of resemblance.

2.5.1 Supervised Learning

In supervised learning, algorithm is given various historical data (input and output variables) and is tasked with identifying the association that contains the highest level of predicted accuracy for data that is not included in the sample. Methods of regression and methods of classification are additional categories for supervised learning techniques. The goal of classification methods is to categorise or group output. For instance, depending on a number of variables, one might want a model's output to be a group of 1s and 0s often referred to as binary such as "purchase" or "sell." The forecast of a regression is a continuous number (e.g., the market will increase by 1%, 1.15 percent, or two percent, etc.), but the forecast of a classification is a discrete number (e.g., buy=1, sell=0, or the volatility regime will be: high=+1, medium=0, low=-1)[29]. The model changes its weights as input data is fed into it until the model has been properly fitted, which takes place as part of the cross validation process. Such as classifying spam in a different folder from your email, supervised learning assists enterprises in finding scalable solutions to a number of real-world issues. Various compute methods and algorithms are applied during supervised machine learning operations.

The most popular learning techniques are briefly described below, usually calculated using software like R or Python. It is common to discuss both supervised and unsupervised machine learning together. Unsupervised learning makes use of unlabeled data as opposed to supervised learning[17]. It extracts patterns from such data to help with clustering or association issues. When subject matter experts are unaware of common characteristics within a data set, this is especially helpful. Gaussian mixture models, k-

means, and hierarchical clustering techniques are frequently used. In semi-supervised learning, only a portion of the incoming data is labelled. Since having total confidence in domain expertise to categorise data accurately for supervised learning can be time-consuming and expensive, unsupervised and semi-supervised learning may be more tempting alternatives.

2.5.2 Deep Machine Learning

Artificial intelligence (AI) and machine learning techniques called deep learning, model how systems acquire specific types of information. Data science, which also encompasses statistics and predictive modelling, contains deep learning as a key component[22]. Deep learning makes the process quicker and simpler, which is very advantageous to data analytics which involves gathering, analysing, and interpreting massive amounts of data. Deep learning can be viewed as a means to automate predictive analytics at its most basic level[49], which are piled in a hierarchy of increasing complexity and abstraction, as opposed to conventional machine learning algorithms, which are linear. The learning process in typical machine learning is supervised, hence requiring explicit coding expertise during the pattern recognition-based methodology.

The computer's success rate in this painstaking procedure, known as feature extraction, entirely hinges on the accurate measure of the class labelling in the supervised learning methodology. In addition to being quicker, deep learning has the advantage that the software develops the feature set independently and without supervision. The deep learning algorithm may initially be supplied training data, such as a collection of images for which each image has been tagged with metatags to indicate the classification category of input images[50]. The computer develops a feature set for the image input and a

predictive model using the information it learns from the pattern recognition task when training the data. A deep learning-based machine learning algorithm sorts through millions of images after being shown a training set, accurately identifying class categories in a matter of minutes.

Deep learning systems need access to massive size of training data and processing power in order to attain an acceptable degree of accuracy. Until the age of big data and cloud computing, neither of these resources was readily available to programmers[1]. Deep learning programming is able to produce precise predictive models from enormous amounts of unlabeled, unstructured data because it is capable of producing complicated statistical models directly from its own iterative output. This is crucial as the internet of things (IoT) spreads further because the majority of data generated by people and devices is unstructured and unlabeled.

2.5.3. Modes of Deep Learning Implementation:

Transfer Learning

This method involves refining a model that was initially trained, it then calls for an access into the network's internal functionalities.[24]. The new users are expected to first add new data including previously unclear classifications to the already-existing network and once the network has been modified, future prediction can be carried out with more accurate categorization expertise. The benefit of this approach is that it uses a lot less data than others, which cuts down computation time to minutes or hours.

Training from Scratch

This is the second alternative to deep learning approach. In this approach, an extensively labelled data set must be put together, while a network architecture that can learn the

features and model must not only be developed but properly ascertained.[23]. Applications with massive output categories and new applications can both benefit greatly from this strategy but since it requires a lot of data and takes days or weeks to train, it is often a less popular strategy in literature. This approach goes through the training, testing, and evaluation phases of the traditional machine learning phases for any pattern recognition task in data mining.

2.5.4 Deep Learning vs. Classical Machine Learning

Sketchy evidence from viewing winning entries at data science competitions suggest that techniques like XGBoost and Random Forests are the most effective at analysing structured data[25]. Deep Learning is only used to analyse text or images in winning entries. Tools for deep learning still need a lot of data to train. Research on how to train on small sample sizes using so-called generative adversarial models is still in its early stages. Because huge sample sizes are required, it is possible that deep learning will be used in intraday or high-frequency trading before it is used at lower frequencies. In a roundabout way, deep learning is immediately useful to portfolio managers. Convolutional neural networks and other Deep Learning architectures are used to count automobiles in parking lot photos.

Deep Learning architectures (such lengthy short-term memory) are used to evaluate text on social media in order to detect sentiment. As was demonstrated in earlier sections of this paper, such traffic and mood data can be immediately incorporated into quantitative tactics. The calculation of such signals will be contracted out to specialised companies, who will create a custom neural network architecture for the job[12]. We must examine the concept of a neuron in order to comprehend deep learning.

A neuron can be thought of as a straightforward calculator that computes the weighted average of the inputs before producing the result if the value is higher than a predetermined threshold. These neurons might be layered and connected by computer scientists to create the neural network depicted below. The first layer's neurons process the user's inputs to create outputs that are passed on to the second layer[25]. The third layer receives its output from the second layer, which in turn feeds it to the third layer, and so forth. The final layer provides the user with an output[51]. A data scientist can use a training set comprising input-output pairs, just like in traditional machine learning, to calibrate the weights of each neuron in the neural network. Although the concept of a neural network itself is not new, recent advances in computer science have made it possible to effectively calibrate the link weights.

In some specific applications, such as picture and text analysis, computer scientists have discovered that neural networks perform well not only at approximating functions on the training set, but also at generalising to previously unseen examples. It's unclear why there is such generalisation and why there isn't overfitting even when there are many factors. The input layer is the leftmost layer of the neural network depicted in the image above that handles input. The output layer is the layer that is furthest to the right. In essence, the user's training set is used to supply values for the input and output layers[7]. The middle levels are referred to as hidden layers. The number of hidden layers and the number of neurons in each hidden layer are both choices available to the neural network builder. Deep learning is the application of neural networks with a large number of hidden layers. The term "many" is a relative term; in reality, 3 or more layers are frequently referred to be "Deep Learning."

Systems with more than 10 hidden layers are commonly referred to as "extremely Deep Learning" systems. It is typical to apply a non-linear activation function to a weighted average of inputs within a single neuron, such as max(x,0)52. One must define the number of hidden layers and the number of neurons in each layer after choosing an activation function. The dataset we are examining fixes the number of neurons in the input and output layers. We would have 10 neurons in the first layer and 1 in the last if, for instance, we were predicting returns for an asset using 10 signals. The designer has the option of selecting the number of neurons in each hidden layer; this number is often set to fall between that of the input layer and that of the output layer. It was usual practise to employ a single hidden layer before the invention of deep learning. It can be demonstrated that a neural network with a single hidden layer may accurately estimate any continuous function[12]. The addition of hidden layers enables the network to learn abstract concepts and features from incoming data, according to new results on very large data sets. There are other characteristics that define a neural network in addition to the number of layers and neurons. The table below is a list of some of the well-known ones.

2.5.5 Ensemble Machine Learning

This is a hybrid of various predictions models; ensemble learning is a very broad somehow complicated approach to machine learning that intends to improve predictive performance. Three major techniques rule the world of ensemble learning, even though there are diverse and infinite amount of ensembles that can be created for predictive modelling[52]. Its study has become so popular today that it has given to various technical and more specialized approaches rather than just creating algorithms per se. The three primary classes of ensemble learning techniques are bagging, stacking, and boosting and

it is very important to have a deep understanding of each one of them. One should also consider mastery these techniques to be able to function as necessary in any predictive modelling project. Ensemble methods in statistics and machine learning combine several learning algorithms to achieve higher predicted performance than any one of the individual learning algorithms could.

A machine learning ensemble, which normally allows for considerably more flexible structure to exist among those alternatives, only consists of a specific finite collection of different models, unlike a statistical ensemble in statistical mechanics, which is typically infinite. Empirically, ensembles often produce superior outcomes when the models are significantly diverse, and as a result, many ensemble approaches work to encourage variation among the models they combine. More random algorithms (such random decision trees) can be utilized to create a stronger ensemble than extremely purposeful algorithms, which may seem counterintuitive (like entropy-reducing decision trees).

However, it has been demonstrated that powerful learning algorithms work better than methods that aim to dumb down the models in order to encourage diversity. When training a model, it is feasible to boost variety by employing information metrics like cross entropy for classification tasks or correlation for regression tasks[6]. What is the best classifier for a specific classification problem? This is arguably the main reason ensemble based systems are utilized in practice. There are two possible answers to this query: 1) Which type of classifier, out of a wide range of competing models, such as the multilayer perceptron (MLP), support vector machines (SVM), decision trees, and naive Bayes classifier, should be used? 2) Given a specific classification algorithm, which realization of the algorithm should be used? For instance, different initializations of MLPs can result in different decision boundaries, even if all other parameters are held constant.

Unfortunately, the most widely used method—choosing the classifiers with the smallest error on training data—has flaws. Even when calculated using a cross-validation method, performance on a training dataset can be deceptive in terms of the classification performance on the previously unexplored data. Which classifier, then, should be picked out of all the (potentially infinite) ones that might all have the same training data or even the same (pseudo) generalization performance as calculated on the validation data (part of the training data left unused for classifier performance evaluation)? One might be tempted to pick at random, all else being equal, but that selection carries the danger of picking a particularly subpar model. Instead of selecting just one, using an ensemble of these models and aggregating their outputs by, for example, simply averaging them, can lessen the chance of picking an unlucky classifier that performs exceptionally poorly. It is crucial to stress that there is no assurance that using several classifiers will always result in higher performance than using the best individual classifier in the ensemble. There is no guarantee that the ensemble will perform better than it typically does, with certain exceptions.

Therefore, combining classifiers may not always outperform the performance of the top classifier in the ensemble, but it definitely lowers the probability of making a generally bad choice. The experts must differ from one another in some way for this procedure to be successful, as will be covered in more detail later in this article. Individual classifiers can produce various decision limits within the context of classification because of the diversity in the classifiers, which is commonly achieved by employing distinct training settings for each classifier. If adequate diversity is established, each classifier will make a separate error, which may then be strategically combined to lower the overall error. Even when calculated using a cross-validation method, performance on a training dataset can

be deceptive in terms of the classification performance on the previously unexplored data. Which classifier, then, should be picked out of all the (potentially infinite) ones that might all have the same training data or even the same (pseudo) generalization performance as calculated on the validation data (part of the training data left unused for classifier performance evaluation)? One might be tempted to pick at random, all else being equal, but that selection carries the danger of picking a particularly subpar model.

Ensemble-based systems can be helpful when working with vast amounts of data or insufficient data. The data can be purposefully divided into smaller subgroups when the volume of training data is too great to make training a single classifier challenging. Then, using a suitable combination rule, each partition can be utilized to train a different classifier (see below for different combination rules). Conversely, if there is insufficient data, bootstrapping can be used to train various classifiers with various bootstrap samples of the data. Each bootstrap sample is a random sample of the data drawn with replacement and is treated as though it were independently drawn from the underlying distribution[53].

2.5.6 Importance of Ensemble Learning

Divide and Conquer

There are some issues that a particular classifier simply cannot address. In actuality, the decision targets that distinguishes data from other classes may be overly complex or lay outside the range of functions that the selected classifier model can execute[54]. A linear classifier, one that is capable of learning linear boundaries, cannot learn this complex non-linear boundary in the context of the two-dimensional, two-class problem. However, any non-linear boundary can be learned by properly combining a group of such linear

classifiers. Consider the case where we have a classifier model that can produce circular borders. This complex non-circular boundary can be learned with ease using a decision based on the majority voting of enough of these classifiers (provided that the classifier outputs are independent and that at least half of the classifiers correctly classify an instance; for proof and a detailed analysis of this method, see the discussion on voting based combination rules). By breaking up the data space into smaller, simpler segments that are easier to learn, the classification system is essentially using a divide-and-conquer strategy. Each classifier only learns one of the simpler partitions. An adequate mix of various classifiers can then be used to approximate the underlying complex decision boundary.

Data Fusion

It is common to acquire data from several sources that may provide supplementary information in many applications that ask for automated decision making. Data or information fusion, the appropriate combining of such data, can increase classification decision accuracy when compared to a choice based solely on one of the distinct data sources. For instance, a neurologist may use the electroencephalogram (one-dimensional time series data), magnetic resonance imaging (MRI), functional MRI, or positron emission tomography (PET) scan images (two-dimensional spatial data), as well as the demographics of the subject, such as age, gender, education level, etc., to diagnose a neurological disorder (scalar and or categorical values).

Confidence Estimation

An ensemble-based system's inherent structure makes it easy to give a level of confidence to the choice it makes. Think about training a group of classifiers to solve a classification problem. A result can be regarded as the ensemble having high confidence in its

conclusion if the vast majority of the classifiers agree with it. However, if half of the classifiers reach one conclusion and the other half reach a different conclusion, this could indicate that the ensemble is not confident in its choice. It should be noted that just because an ensemble has high confidence in its decision doesn't necessarily mean that decision is right. Likewise, a decision with low confidence doesn't necessarily mean that decision is wrong. It has been demonstrated, though, that an ensemble choice that has been properly educated is typically correct if its confidence is high and typically incorrect if its confidence is low. Therefore, using such a method, the ensemble decisions can be utilized to calculate the classification decisions' posterior probabilities.

Diversity

The diversity of the classifiers that make up an ensemble system is crucial to its effectiveness, or its capacity to remedy the mistakes of some of its members. After all, if all classifiers produced the same results, it would be impossible to fix a potential error. As a result, each classifier in an ensemble system must commit a different error on every occasion. Therefore, it follows that if each classifier has a unique error, combining these classifiers strategically should lower the overall error—a concept comparable to low pass filtering of the noise. An ensemble system needs classifiers, namely, whose judgement boundaries are sufficiently distinct from those of others. It is referred to as a diversified group of classifiers. Utilizing various training parameters for multiple classifiers is another strategy to achieve diversity. Multilayer perceptron (MLP) neural networks, for instance, can be trained using a variety of weight initializations, layer and node counts, error targets, etc. By changing these values, one can influence how unstable each classifier is and so increase the diversity of the classifiers. Different types of classifiers, including MLPs, decision trees, closest neighbour classifiers, and support vector machines, can also be coupled for more variety. Finally, utilizing various features or

various subsets of already existing traits is another way to achieve diversity. The random subspace approach[8], which is what is used to create various classifiers using random feature subsets, is actually detailed later in this article.

2.6 Statistical Data Resampling through Synthetic Minority Oversampling Technique (SMOTE)

The basis of applied machine learning is data. It is crucial that it be efficiently gathered and used as a result[55]. When picking observations from a domain with the goal of estimating a population parameter, statistical procedures are referred to as data sampling. Data resampling, on the other hand, refers to techniques for economically utilizing a gathered dataset to enhance the estimate of the population parameter and aid in quantifying the estimate's level of uncertainty[5]. A challenge involving predictive modelling necessitates the use of both data sampling and data resampling techniques. One frequently do not have access to all potential observations while working with data. This could be for a variety of reasons, like:

1. Making further observations could be challenging or expensive.
2. It could be difficult to compile all observations into one place.
3. Future observations are anticipated to increase.

Identifies this method as one whose domain's observations are samples of a larger, idealized population of all potential observations that could be made there[56]. One can distinguish between and relate the data and the idealized population thanks to this helpful conception[5]. Statistical sampling is the practise of choosing subsets of instances from a population. This is with the sole aim of approximating the population's characteristics. It is an active type of sampling. The intended outcome of this practice is to estimate

population characteristics, and to ascertain sampling methodology is under control. This control, unlike doing an experiment, is unable to affect the process that results in each observation. As a result, the field of sampling fits well between unadulterated observation and supervised experimentation.

The classifier's job in binary classification problems is to assign each sample to one of two classes[5]. The dataset is deemed unbalanced when there are significantly more samples in the majority (larger class) than in the minority (smaller class). This skewness presents a challenge because both the classifier training procedure and the widely used metrics to gauge classification quality are frequently skewed in favour of the dominant class. Many metrics, such as the Brier score and the area under the receiver operating characteristic curve, have been proposed to address the problems with imbalanced binary classification (AUC). Each one of them is non-symmetric and assigns a greater loss for incorrectly identifying a minority sample than a majority sample. Modern machine learning classifiers often seek to forecast the underlying probabilities by optimizing the symmetric loss functions, despite being interested in non-symmetric loss functions. To resolve this mismatch, balancing strategies that improve the data's balance before training the classifier were developed. The simplest balancing techniques include either undersampling the majority class or oversampling the minority class by duplicating minority samples. Re-sampling is a set of techniques used to reconstitute training sets and validation sets of your sample data sets. It might in some way offer more varied sample sets that are "helpful" for the learning process.

Re-sampling is a technique that involves taking additional samples from the initial data samples. Re-sampling is a non-parametric technique for drawing conclusions from statistical data. To put it another way, the resampling method does not rely on the use of

general distribution tables (such normal distribution tables) to estimate approximations of p probabilities[5]. In resampling, cases are chosen at random with replacement from the original data sample so that each sample number contains a certain amount of cases that are representative of the original data sample. The amount of samples drawn for the resampling procedure is made up of recurrent cases as a result of replacement. Re-sampling produces a distinctive sample distribution based on the real data. The unique sampling distribution is produced by the resampling approach using experimental methods rather than analytical ones[5]. Because the resampling method is based on unbiased samples of all potential outcomes from the data under study, it produces estimates that are free of bias. Re-sampling is sometimes referred to as Monte Carlo Estimation or Bootstrapping. The phrases Bootstrapping and Monte Carlo estimate must be understood in order to comprehend the idea of resampling:

The test statistic is computed using repeated samples from the original data sample in the bootstrapping method, which is identical to resampling. The resampling findings are obtained using Monte Carlo estimation, which is also equivalent to the bootstrapping method[5]. In resampling, there are assumptions including:

i. The parametric assumptions used by this resampling technique typically neglect the underlying data distribution's characteristics. As a result, non-parametric assumptions form the basis of the procedure.

ii. There is no set sample size requirement for resampling. Therefore, the confidence intervals produced by the resampling procedure are more trustworthy the larger the sample.

iii. Overfitting noise in the data is more likely to happen. Combining the procedure of cross-validation with the approach of re-sampling makes it simple to address this kind of challenge.

In SMOTE, where synthetic minority samples are produced by the interpolation of pairs of the original marginal points, the concept of adding synthetic minority samples to tabular data was first put up[15]. Later, dozens more SMOTE variations and expansions were put forth. Real-world datasets frequently contain a large percentage of typical cases and a very small number of unusual or intriguing examples. The cost of mistaking an aberrant example for a normal example is likewise true. It has been suggested that undersampling the majority (normal) class is a useful way to make a classifier more sensitive to the minority class[15]. The classifiers can perform better than just under-sampling the majority class if they over-sample the minority (abnormal) class and under-sample the majority (normal) class. Instead of oversampling with replacement, SMOTE utilizes an oversampling strategy that oversamples the minority class using artificial examples. SMOTE is an efficient oversampling technique to resolve the over-fitting issue caused by a too-narrow decision interval[15]. It can manage the quantity of examples and distribution to meet the goal of balancing the dataset using synthetic additional examples. In SMOTE, each sample in the minority class and its closest neighbours are used to create new samples, with each synthesized sample Xsyn being given by:

$$Xsyn = Xi + rand(0, 1) \times Xi - Xneighbour$$

where Xi is a predetermined sample from the minority class and Xneighbour is a randomly chosen sample from the sample's K closest neighbours A random number between 0 and 1 is represented by Xi and rand as 0; 1. Avoiding overfitting is the key benefit of utilising SMOTE for oversampling rather than other approaches like Random Over Sampling (ROS). This can be accomplished by creating fresh samples from the minor class rather than duplicating the existing ones as is done in ROS[15]. Despite its popularity and strong performance in a variety of application areas, SMOTE still has a lot

of issues because nothing is ever completely faultless. First, despite oversampling the minor class, most classifiers are unable to improve their performance due to the high dimensionality problem of the datasets employed. Another issue is the neglect of major class neighbours, which results in the development of minor samples, which causes greater class overlap, especially when the minor class is scarce[15].

The creation of noisy samples as a result of SMOTE's inherent unpredictability, as is evident in the equation above, is its most significant shortcoming, independent of the size of the dataset. This is in addition to any noise that might have been present in the original dataset and that might be utilized to create new samples, causing more noise to spread and grow. When datasets are oversampled using SMOTE, these noisy samples serve as outliers that prevent many classifiers from improving their performance. Therefore, when SMOTE is used, it is a recommended practice to utilize one of the noise removal procedures[15]. As a result, the performance of the classifiers used to categorize the datasets is improved. This helps to lessen the effect of noise produced by SMOTE or noise that may exist in the original datasets[15].

Other Resampling Approaches

Cross Validation Approach

A statistical technique for validating a predictive model is cross-validation. A model is fitted to the remaining data (a training set) and used to make predictions for the validation set. Subsets of the data are held back for use as validating sets. An overall measure of prediction accuracy is obtained by averaging the quality of the predictions across the validation sets. It is common practise to use cross-validation while creating decision trees[2]. One type of cross-validation, known as the jackknife, excludes one observation at a time. Another method, known as K-fold cross-validation, divides the data into K subsets, each

of which serves as the validation set in turn. This is frequently used to choose the number of predictor variables to include in a regression. Adding predictors always lowers the residual sum of squares without cross-validation (or possibly leaves it unchanged). In contrast, the cross-validated mean-square error will typically go down when adding useful predictors, but go up when adding useless ones.

Sub-Sampling

An alternate technique for roughly approximating an estimator's sampling distribution is subsampling. The resample size is smaller than the sample size, and the resampling is carried out without replacement, which are the two main variations from the bootstrap. In comparison to the bootstrap, subsampling has the advantage of being valid under much weaker conditions. The limiting distribution must also be continuous, and the resample (or subsample) size must approach to infinity together with the sample size but at a slower rate such that their ratio converges to zero. These are only a few examples of the sufficient criteria. When resampling time series data, one resamples blocks of following data rather than individual data points, although subsampling was initially recommended primarily for the situation of independent and identically distributed (iid) data[2]. For instance, in situations when the rate of estimator convergence is not the square root of the sample size or when the limiting distribution is non-normal, subsampling yields correct inference while bootstrapping does not. The bootstrap is usually more accurate when subsampling and consistency are both present. A well-liked subsampling algorithm is RANSAC.

Permutation Test

Resampling the initial data while assuming the null hypothesis is how permutation tests work. How probable the original data is to occur under the null hypothesis can be deduced from the resampled data.

2.7 Review of Empirical Studies

In a work title 'Multimodal Biometrics: Applications, Challenges and Research Area', The author addressed the universality of Biometrics methods through a generalized model. It was later discovered that biometrics systems can be all encompassing with one model that can take care of many applications but not without its weakness of not being able to address the uniqueness of individual and psychological changes in humans as a result of unforeseen factors and that it did not factor in evolving biometric methods

A related development on Overview of Multimodal Biometrics - Signal & Image Processing- by Sanjekar et al, this work was placed side by side of unimodal biometrics and rated using ANN metrics. Though there was an intention to develop a universal system with its target of security as the primary key point around a better improvement on unimodal system. The work did not achieve its optimal goal of maintaining an all-round security in biometric-based environment.

In the study of 'Multimodal Biometric System Iris and Fingerprint Recognition Based on Fusion Technique', two major traits were applied basically the fingerprint and iris for feature extractions. It decomposed the features using KNN and Fuzzy Logic. The result had a high level of accuracy, security and speed of processing but imbalanced data reduced the performance metrics.

Elhoseny et al in 'Cascade Multimodal Biometric System Using Fingerprint and Iris Pattern' looked at overview of fusion with focus on data quality, soft biometrics combination and improved biometric recognition accuracy with the strength of faster processing of images but non-guarantee data security and difficulty in Real world application.

Also, in 'Machine Learning for Biometrics. Concepts, Algorithms and Applications'. A volume in Cognitive Data Science in Sustainable Computing. This work used algorithm ranking for fusion based methods. Thus, making a way for more than two methods to be used per time with a very strong stream of security with emphasis more on cognitive data science but its result cannot be predicted

'Multimodal Biometric Recognition using Iris Feature Extraction and Palmprint Features', utilizing Support Vector algorithms with strengths in provision for higher level exploration of other related Artificial Intelligence Methods but The quality of production and analysis is weak i.e. a weak and unreliable result.

Gangwar et al worked on DeepIrisNet: Deep Iris representation with applications in iris recognition and cross-sensor iris recognition- Used logistic regression approach by applying Gabor Wavelength. Reliable result because of diverse methods that are used but feature selection was not done prior to model,

In Multimodal Biometric Authentication with Secured Templates — A Review**,** The author used an adaptive multimodal matrix with a Particle Swarm Optimisation Method and Wavelenth Based Kernel that improved the quality of the output were used. They came up with an improved identification module and a secured environment of 89% accuracy which was better than what was initially available. The system was not totally reliable at the fusion level as there were lots of inconsistencies at this stage as a result of involvement of too many biometric factors.

A Comprehensive Survey on the Biometric Recognition Systems based on Physiological and Behavioral Modalities, used multimodal matrices using face and palm print though

Daubechies wavelength. It also proposed a universal method that can fit into most applications but fusion here became a challenge in the application. For every other application that does not fit into this design, there may be need to develop a fresh application.

'Multi-Biometric Fusion Authentication in wireless Multimedia Environment using Dynamic Bayesian Method', it was proposed that a multi-biometric fusion authentication solution using dynamic Bayesian decision method[31]. The fusion of fingerprint and voiceprints is done by using matching layer score. Fingerprint features are captured and extracted such features are the attribute point, position and direction field value of the feature point. The extracted feature point information is then serialized to generate the biometric template. For the voice recognition, the features extracted are the acoustic feature of the speech and voiceprint model by Gaussian Mixture Model (GMM) which are then used as template. Matching score is used to calculate the voiceprint. Thereafter, the Dynamic Bayesian method is then used to fuse the features recognition in the score layer and in the decision layer based on the fingerprint and voiceprint single mode feature extraction. But did not look at any physical factors that may also identify the individual.

Secured Cryptographic Key Generation from Multimodal Biometric: Feature level Fingerprint Image features i.e. minutia point is obtained using segmentation, orientation field estimation and morphological operators while Iris image features are acquired by segmentation, estimation of Iris boundries and Normalization. Both minutia point and iris texture extraction are fused at feature level to build the multimodal biometric template. Concatenation, shuffling and merging processes are means by which the fusion is achieved. The multi-biometric template generated is used to generate 256 bits cryptographic key which is used to enable user authentication and security.

In another work which focuses on DL model with multi-modality using Residual Network (ResNet-50) for extracting features from face, finger and ECG. The network architecture is divided into three parts: feature extraction, fusion layer and task layer. Each modality feature is extracted using a different method as each modality has different characteristics. The data type of fingerprint and facial is a static image while ECG is a temporal biological signal. ResNet. Is widely used feature extraction trained on the ImageNet database which is the largest database. For Fingerprint extraction: 7,200 images are obtained from 150 individuals which consist of four different sensors; Electric field sensor, Optical Sensors, Thermal sweeping sensor and SfinGe v3.0 having 1,800 images per sensor. Each database is divided into two subsets, A and B having DB1-A to DB4-A. Each subset contains 140 subjects with 12 images, resulting into 1,680 images. For other subsets DB1-B to DB4-B, the image resolution of the four sensors was 96x96 pixels, 400x560pixels, 400x500 pixels and 288x384pixels.

In another review on Multi-Modal Biometrics: Applications, Strategies and Operations. Though, it is very difficult to properly fuse three different modalities at the input level. Comparison was made based on the performance between score level fusion and the feature level fusion. The three methods of score level fusion, which are sum, product and max rules were tested. The feature level fusion is a method before matching phase that consolidates features of modalities to one feature vector, score-level fusion is a method used after matching that calculates the degree of similarity using rules i.e. sum, product and max., from each output of modality to the final output vector.

Although the number of data used was not big enough, it is possible to substitute other multiple biometric data since the feature extracting method is independent on the type of

data than other previous studies. The fusion layer enables the model to learn and infer even when there are N modalities.

It was concluded that applying multiple biometric data in a single model has the advantage of using parameters efficiently as well as to also enhance the security of the system since it can be processed on a single device. The use of ECG, facial image and fingerprint datasets are used as the multimodal having multitask learning which is robust to noise. The model developed is able face the spoof attack problems faced with most of the models based on single modality.

Feature Extraction and Classification by Machine Learning Methods for Biometric Recognition of Face and Iris by M. Oravec, 2018. Face recognition system and Iris recognition system were fused together with the help of some algorithm that was able to detect an individual face, eye and other parts of the face in order to authenticate an individual.

Bimodal Biometric Verification Mechanism using Fingerprint and face images was conducted in 2016 where the research work shows that fingerprint and face biometric recognition system were used as a verification mechanism. Normalization techniques such as Z-score, two quadrice and fusion techniques such as Max score, match weighing were all used to achieve the model.

In the work of, an automatic system for attendance solution was implemented with a single-factor face recognition using deep learning[57]. The students' faces is captured and is used to train the Multi-Task Cascaded Convolutional Neural Network (MTCNN) and the public data set of FaceNet. The testing set consist of 108 pictures with a 100% accuracy for face detection and recognition. The ensemble of private and public datasets for the

training ensured a robust intelligence level for the MTCNN deep learning training hence the optimal accuracy result. However, the model was not tested for variance error.

The main thrust of the work of another author is to ensure a reliable authentication system for individual recognition[1]. The model is trained on four public data repositories using electrocardiogram data. The deep learning model employed synthetic minority oversampling to address the class imbalance problem while no feature selection technique is employed to ensure optimal performance of the model.

An optimized individual authentication system with high privacy was conducted in another study. By using minutiae point quality for the optimal 3D shell form template, the study varied from the normal deployment of minutiae point location. For the study, nine fingerprint repositories are used. The optimization of the minutiae point system and the homogeneous nature of training set proved to be optimal steps for the model while adoption of a single-factor authentication was a drawback.

A smart attendance monitoring system using face recognition is the main thrust of the work of Bhattacharya et al. The study employed Convolutional Neural Network on the private dataset and a normalized sharpness technique is incorporated to enhance the performance of the model. The image size and resolution were also optimized for the deep-learning framework outweighs state-of-the-arts. Similarly, a deep one-shot learning deep learning system is carried out for an efficient attendance management system. The HOG image filtering technique is employed in the study to extract features which is fitted by the CNN. In three iterations, the model achieved accuracy and a F1 score of 97% and 98.4% respectively on a camera phone-enabled framework, while 91.9% and 94.8% is

achieved respectively with a Moto G camera. A webcam-based image sourcing achieved a 51.2% and 61.1% respectively.

The work of Seelam et al. implemented a smart attendance solution using deep learning with computer vision. The framework is implemented on a raspberry pi which is a portable solution that enhanced the performance of the framework. The Haar cascade machine learning algorithm is trained with the data for the Adaboost-based ensemble framework. A 80:20 training and testing split is implemented with a state-of-the-art experimental result.

A biometric system of fingerprint factor is implemented in the framework of which seeks to authenticate end users of a healthcare electronic system. The approach ensured reduction in the computational cost while also ensuring speed performance using the fast stereo algorithm. The single factor authentication system did not yield an optimal performance just as the framework could be better with a data mining approach of predictive analytics.

Iris factor is incorporated with a steganography-based user authentication system for Internet-of-Things in another study. By integrating and applying the information-hiding technique known as steganography and protecting the public key from key exposure-related compromising assaults, the suggested model strengthens the security aspects of the system. The system was greatly improved by the proposed cancelable biometric technique used in the study. the distribution of record multiplicity-based attacks and the utilization of cancelable biometric methodology returns an optimal performance for the study whereas, the vulnerability of the biometric system and the single factor approach were not mitigated in the study using a double-factor input parameters.

In a CNN-based two-tier and multi-factor authentication system for an anti-spoofing framework is proposed[58]. The system includes both soft and hard biometric schemes, with the Tier I biometrics including palm vein print extraction, facial recognition, and fingerprints. The Tier II biometric scheme uses a CNN model to identify spoofing on the test dataset. The fingerprint is hashed, and following successful verification, a similar procedure is applied to the additional input factors. The multimodal authentication system with the multi-factor input greatly returns an optimized result while the homogeneous nature of the training set and limited feature vectors used for the implementation were major drawbacks.

A hybrid approach is implemented by a study using a big data analytics for the user identification study. In order to create real-time, highly recognizable information for user verification, this approach analyses the dynamic user activities. The user authentication method using the recommender system was effective. Use of user's personality traits as attributes aided the identity verification system while the big data used for the predictive analytics likewise enhanced the framework. The vulnerability of the proposed system due to non-availability of independent authentication and the inability to adopt the model for large scale verification are the major drawbacks.

Moreover, a counterpart approach using machine learning techniques for an attendance and feedback system was presented[18]. In another study, two years of attendance of students in the computer science and engineering departments at Aurora's Technological and Research Institute, Parvathapur, Uppal, Hyderabad, are collected to produce a sizable dataset. The characteristics are discovered through surveys that ask students why they missed a particular class, as well as by patterns in past attendance records that have been

noted. This analysis of historical attendance data and student surveys identifies the characteristics of the dataset on which machine learning will be performed for the label and class attendance. The study employed linear regression and Random forest for the modelling of the predictive analytics. Inconsistent outcomes from the linear regression model defied both intuition and verification using actual class strength as input data. It was evidently discovered that the linear model was not suitable for use for attendance prediction after utilizing both the random forest regression model and real data for prediction. On the other hand, the predictions made by the Random Forest Regression Model were more accurate. The random forest regression model and was successful in predicting attendance 8 out of 10 times with an accuracy of at least 90% when both models were evaluated with real data and attendance.

An automatic student attendance marking system is likewise proposed where different stages are involved in the conceptual framework including the image acquisition where the suggested solution makes use of a camera that is strategically placed to take pictures of students' faces in a safe setting. Image processing where Histogram equalization is used to improve the quality of input images; Face detection when the Viola-Jones and HAAR Cascade algorithms are used to identify faces in the image; Feature extraction involving the LDA technique to extract the features and minimize the dimensionality of the feature vectors; Face recognition using the LDA, SVM, and KNN algorithms. Marking of attendance completes the process by matching in the database, the student's attendance will be automatically marked on the Excel sheet according to the lecture's name and time. Otherwise, it will show an error and the attendance won't be recorded.

2.8 Summary of Literatures Reviewed

This section reviewed literatures on the varying subjects contained within the study for a concise synopsis of the techniques employed in the study. The subject of user authentication was reviewed and tailored down to user identification and as well as the different types and approaches to user authentication. The biometric authentication system is also examined considering the double factors of iris and finger print inputs, with their benefits. The data science use case of predictive analytics is discussed in the later part of the section, which leads to the subject of ensemble and deep machine learning concepts. The section is concluded by reviewing related studies including the methodologies, strengths and weaknesses of the studies. It is observed that the bias and variance tradeoff consideration is seldom implemented by the studies just as all the studies employed image filters for their feature extraction phases. The two gaps were covered in this study while evaluating the performances of the deep embedder alternative for feature extraction.

Endnotes

1. *"A Brief History of Biometrics"*. **Bioconnect.** 2021: https://bioconnect.com/2021/12/08/a- brief-history-of-biometrics

2. R. Agrawal & R. Goyal. "*Developing Bug Severity Prediction Models Using Word2vec.*" **International Journal of Cognitive Computing in Engineering**, 2021: 104-115.

3. I. Al-Amoudi, R. Samad, N. R. Hasma-Abdullah, M. Mustafa, & W. I. Pebrianti. *Automatic Attendance System Using Face Recognition with Deep Learning Algorithm*. **Springer**, 2022: 573-588

4. A. A. AbdElaziz. "*A Survey of Smartphone-Based Face Recognition Systems for Security Purposes*." **Kafrelsheikh Journal of Information Sciences** 2, no.1, 2021.1-5

5. P. Ghimire, S. Piya, & A. M. Gurung. "*Comparative Study of Face Mask Recognition Using Deep Learning and Machine Learning Classifiers*." **2021 International Conference On Innovative Computing, Intelligent Communication And Smart Electrical Systems (ICSES). IEEE**, 2021: 1-9.

6. I. Gaytán-Campos, W. Morales-Castro, B. Priego-Sánchez, E. Fitz-Rodríguez, & R. Guzmán-Cabrera. "*Automatic Classification of Images With Skin Cancer Using Artificial Intelligence*." **Computación y Sistemas 26**, no. 1 2022.

7. S. Pal & Z. Jadidi. "*Protocol-Based and Hybrid Access Control for the IoT: Approaches and Research Opportunities*." **Sensors** 21, no. 20, 2021. 6832.

8. S. K. Choudhary & A. K. Naik. "*Multimodal Biometric Authentication with Secured Templates — A Review,*" **3rd International Conference on Trends in Electronics and Informatics (ICOEI),** 2019. 1062-1069, doi: 10.1109/ICOEI. 2019.8862563.

9. T. Kumar, S. Bhushan, & S. Jangra. "*Ann Trained and WOA Optimized Feature-level Fusion of Iris and Fingerprint.*" **Materials Today: Proceedings** 51, 2022. 1-11.

10. A. S. Mustafa, A. J. Abdulelah, & A. K. Ahmed. *"Multimodal Biometric System Iris and Fingerprint Recognition Based on Fusion Technique."* **International Journal of Advanced Science and Technology**, 29, no. 3, 2020. 7423-7432.

11. H. S. Maghdid, A. T. Asaad, K. Z. Ghafoor, A. S. Sadiq, S. Mirjalili, & M. K. Khan. "*Diagnosing COVID-19 Pneumonia from X-ray and CT Images using Deep Learning and Transfer Learning Algorithms*." **Multimodal Image Exploitation And Learning**, 2021, 99-110.

12. A. Kumar, A. Kaur, & M. Kumar. "*Face Detection Techniques: A Review*". **Artificial Intelligence Review** 52, 2019. 927–948

13. NEC. "*What Is Iris Recognition and How Does It Work*?" 2022: **https://www.nec.co.nz/market-leadership/publications-media/what-is-iris-recognition-and-how-does-it-work/**

14. A. R. Shahzad & A. Jalal. "*A Smart Surveillance System for Pedestrian Tracking and Counting Using Template Matching*". **International Conference on Robotics And Automation in Industry (ICRAI), IEEE**, 2021. 1-6.

15. M. Sajjad, S. Khan, T. Hussain, K. Muhammad, A.K. Sangaiah, A. Castiglione, C. Esposito, & S.W. Baik. *"CNN-Based Anti-Spoofing Two-Tier Multi-Factor Authentication System"*. **Pattern Recognition Letters**, 126, 2019. 123-131.

16. A. S. Mustafa, A. J. Abdulelah, & A. K. Ahmed. "*Multimodal Biometric System for Iris and Fingerprint Recognition Based on Fusion Technique*." **International Journal of Advanced Science And Technology** 29, no. 3, 2020. 7423-7432.

17. A. Saha, J. Saha, & B. Sen. *"An Expert Multi-Modal Person Authentication System Based on Feature Level Fusion of Iris and Retina Recognition."* **International Conference on Electrical, Computer and Communication Engineering (ECCE),** 2019. 1-5.

18. G. Carvalho, S. Mykolyshyn, B. Cabral, J. Bernardino, & V. Pereira. *"Comparative Analysis of Data Modeling Design Tools."* ***IEEE Access*** 10, 2021. 3351-3365.

19. D. M. Kroenke, D. Auer, S. L. Vandenberg, & R. C. Yoder. "*Database Concepts 9th Edition.*" **Pearson**. 2019.

20. K. W. Brian & C. S. Stacey. "*Using Information Technology*" ***A Practical Introduction to Computers and Communications 12th Edition***. 2019.

21. B. Joseph. *"The Roles of Information and Communication Technologies (ICTs) and E-Commerce as Agents of Nigeria's Economic Development: Review of Challenges and Prospects."* **Wireless Engineering And Technology** 10, no. 03, 2019.

22. W. Yang, S. Wang, J. Hu, G. Zheng, & C. Valli. *"Security and Accuracy of Fingerprint-Based Biometrics: A review"*. **Symmetry**, 11, no. 2, 2019. 141.

23. C. Huifan. "The *Difference Between Iris Recognition and Fingerprint Recognition."* **Chongquing Huifan Technology**. 2022.

24. Z. Truce. "*Technologies Used in the Attendance Managemen*t." ***Technology Insights.*** https://www.zimyo.com/insights/technologies-used-in-the-attendance-management, 2020

25. M. Szymkowski, P. Jasiński, & K. Saeed. "*Iris-Based Human Identity Recognition with Machine Learning Methods and Discrete Fast Fourier Transform*." **Innovations In Systems And Software Engineering**, 2021: 309-317.

26. S. M. Zaidawi, K. Al, M. H. Prinzler, J. Lührs, & S. Maneth. "*An Extensive Study of User Identification Via Eye Movements Across Multiple Datasets.*" ***Signal Processing: Image Communication*****,** 2022, 116804.

27. W. Yang. "*A Cancelable Iris- and Steganography-Based User Authentication System for the Internet of Things.*" **MDPI**, 2019.

28. P. P. Shinde & S. Shah. "*A Review of Machine Learning and Deep Learning Applications.*" **4th *International Conference on Computing Communication Control And Automation* (ICCUBEA), IEEE**, 2018. 1-6.

29. U. Sharma, P. Tomar, S. S. Ali, N. Saxena, & R. S. Bhadoria. "*Optimized Authentication System with High Security and Privacy.*" **Electronics**, 2021: 1-23.

30. H. U. Suru & P. Murano. "*Security and User Interface Usability of Graphical Authentication Systems – A Review*." **International Journal of Engineering Trends and Technology (IJERT)** 67, 2019. 17-36.

31. S. Krishnmaoorthy, L. Rueda, S. Saad, & H. Elmiligi. "*Identification of User Behavioral Biometrics for Authentication using Keystroke Dynamics and Machine Learning* ." **2018 2nd International Conference on Biometric Engineering and Applications** . 2018. 50-57.

32. T. Olaleye, O. Arogundade, C. Adenusi, S. Misra, & A. Bello. "*Evaluation of Image Filtering Parameters for Plant Biometrics Improvement Using Machine Learning.*" **International Conference on Soft Computing and Its Engineering Applications. Springer,** 2021. 301-315.

33. M. S. M. Suhaimin, M. H. A. Hijazi, C. S. Kheau, & C. K. On. *"Real-Time Mask Detection and Face Recognition Using Eigenfaces and Local Binary Pattern Histogram for Attendance System."* **Bulletin of Electrical Engineering and Informatics** 10, no. 2, 2021. 1105-1113.

34. V. Seelam, A. K. Penugonda, B. P. Kalyan, M. B. Priya, & M. D. Prakash. "*Smart Attendance using Deep Learning and Computer Vision*." ***Materialstoday: Proceedings,*** **Elsevier**, 2021. 4091-4094.

35. Z Truce. "*Technologies Used in the Attendance Management*." **Technology Insights**. https://www.zimyo.com/insights/technologies-used-in-the-attendance-management, 2020.

36. M. A. Lavadkar, P. K. Thorat, A. R. Kasliwal, J. S. Gadekar, & D. P. Deshmukh. *"Fingerprint Biometric Based Online Cashless Payment System."* **IOSR Journal of Computer Engineering (IOSR-JCE),** *e-ISSN*: 2019. 2278-0661.

37. N. S. Reddy, M.V. Sumanth, & S. S. Babu. "*A Counterpart Approach to Attendance and Feedback System using Machine Learning Techniques."* **Jetir-*International Journal of Emerging Technologies and Innovative Research*,** 2018. 190-193.

38. T. O. Olaleye & O. R. Vincent. "*A Predictive Model for Students' Performance and Risk Level Indicators using Machine Learning.*" ***International Conference in Mathematics, Computer Engineering and Computer Science*. Lagos: IEEE**, 2020. 1-7.

39. S. Krishnmaoorthy, L. Rueda, S. Saad, & H. Elmiligi. "*Identification of User Behavioral Biometrics for Authentication Using Keystroke Dynamics and Machine Learning* ." **2nd International Conference on Biometric Engineering and Applications** . 2018. 50-57.

40. M.I. Jordan & T. M. Mitchell. "*Machine Learning: Trends, Perspectives, And Prospects."* **American Association for the Advancement of Science**, 2015. 255-260.

41. S. Aishwarya "*A Comprehensive Guide to Ensemble Learning (with Python Code)"* **Analytic Vidhya**. 2018. https://www.analyticsvidhya.com/blog/2018/06/comprehensive-guide-for-ensemble-models

42. N. Rahimi, F. Eassa, & L. Elrefaei. *"An Ensemble Machine Learning Technique for Functional Requirement Classification."* **Symmetry** 12, no. 10, 2020, 1601.

43. S. K. Choudhary & A. K. Naik. "*Multimodal Biometric Authentication with Secured Templates — A Review*," **3rd International Conference on Trends in Electronics and Informatics (ICOEI),** 2019, 1062-1069, Doi: 10.1109/Icoei.2019.8862563.

44. D., Devikanniga, A., Ramu, & A. Haldorai. "*Efficient Diagnosis of Liver Disease using Support Vector Machine Optimized with Crows Search Algorithm*". **EAI Endorsed Transactions on Energy Web**, 7, no. 29, 2020. e10-e10.

45. S. Satpathy. "*Overcoming Class Imbalance using SMOTE Technique*". **Analytics Vidhya**. 2020. http://www.analyticsvidhya.com/blog/ 2020/10/ overcoming-class-imbalance-using-smote-techniques

46. Y. Elor & H. Averbuch-Elor. "*To Smote, or Not To Smote?*" **Arxiv:**2201.08528v2, 2022. 1-22.

47. S. Bhattacharya, G. S. Nainala, P. Das, & A. Routray. "*Smart Attendance Monitoring System (SAMS): A Face Recognition Based Attendance System for Classroom Environmen*t." **IEEE 18th International Conference on Advanced Learning Technologies (ICALT). IEEE,** 2018.

48. M. Shaha, & M. Pawar. "*Transfer Learning for Image Classification*." ***2nd International Conference on Electronics, Communication and Aerospace Technology*** **(ICECA).** IEEE, 2018. 656-660.

49. J. E. Hoover, " *FingerPrint Anatomy*". **Encyclopaedia Britannica,** 2023.

50. B. Mesnik, *"How Does Facial Recognition Work?*". **Kintronics**, 2017.

51. T. Cho, "*Iris Recognition Reaches the Mainstream for Identification, Authentication"*. **Biometric Update**. 2022.

52. Suresh , K. Logeswaran , P. Keerthika , R. Manjula, Devi, K. Sentamilselvan, G.K. Kamalam, & H. Muthukrishnan. **"***Contemporary Survey on Effectiveness of Machine and Deep Learning Techniques for Cyber Security in Machine Learning for Biometrics, Concepts, Algorithms and Applications*" **ScienceDirect,** 2022, 177-200.

53. R. Parihar. " *Palm Vein Recognition System for Human Authentication: A Review"*. **International Journal for Research in Applied Science and Engineering Technology**, 2019.

54. S. Dargan & M. Kumar, "*A Comprehensive Survey on the Biometric Recognition Systems Based on Physiological and Behavioral Modalities*," **Expert Systems with Applications,** p. 113114, 2019.

Methodology

3.1 Research Approach

This section describes techniques, approaches and stages involved in the implementation framework, aimed at improving on the existing models that were reviewed in the literature.

The bimodal authentication framework is a five-phase biometric-based student attendance verification system that combines iris and fingerprint recognition attributes for the purpose of training deep learning models. The first phase entails the image acquisition which acquires the iris and fingerprints of both authentic and unauthentic students for the purpose of the study. The acquired image inputs are then subjected to the feature extraction phase where attributes are extracted from the image inputs in form of numeric image descriptors. The phase three of the framework is a data resampling phase in order to ensure a balanced training set for the machine learning-based study. Balanced training set is consequently deployed for the deep learning phase of four after which the performance evaluation is carried out in phase five.

The framework is as captured in:

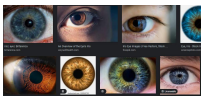

Figure 3.1: Bimodal Authentication Framework (Researcher, Olomola, B. 2023)

3.2 Data Acquisition

The first phase of the framework is the image acquisition phase when dual image inputs of iris and fingerprints of legitimate students and others is captured as instances on the dataset. The images belong to authorized users' category and the non-authorized users, each having a pair of fingerprint and iris representations. Instances for each data feature has more than one representation on the dataset to avail the predictive model an armful opportunity of a robust pattern recognition task. The sizes of the images notwithstanding,

two sets are prepared for this study including the training set and the test set. The training set would be deployed for the machine learning phase while the test set would be deployed for the user verification or testing phase. The test set is used to test the model to ascertain its accuracy for user authentication. Consequently, for a particular student, both iris and fingerprints are captured, and there will be more than one representation of the particular student (for both iris and finger print) on the training set.

The training and test set for this study comprised of both primary and secondary dataset. The secondary dataset were 800 fingerprints and 800 iris biometrics images obtained from the internet. Meanwhile, the primary data were 80 fingerprints and iris biometrics images obtained from individuals randomly chosen and captured with the aid of an Android Infinix Note 8i mobile device. Professional ethical issues were taken into consideration while collecting these dataset. The fingerprints are captured on white background papers, while the iris are captured directly all in broad day light. The entire set is split between training and test set. In order to address a germane objective of this study which is to test the performance of the model with a balanced dataset. The images are all in Joint Photographic Expert Group (JPEG) format and their sizes ranges from 27kb to 20kb, in greyscale format. The acquired fingerprint and iris dataset in JPEG are imported into the Orange data mining toolkit for feature extraction. The five deep learning networks for image embedding, described earlier in Table 3.1, are employed for the feature extraction. The images are passed through each of the five embedders to generate numeric vector attributes to return five different and unique numeric datasets, which will be used in turn to train the machine learning algorithms.

3.3 Feature Extraction Through Image Embedding

In this stage of the conceptual framework, image embedding networks are used to obtain numerical image descriptors. By using transfer learning technology, image embedding derives feature vectors from image inputs. Deep neural networks, which return a vector of numbers on each individual image and are referred to as numerical image descriptors, are used to do this. A feature vector is created by the deep learning neural network model, which then outputs an improved data table with columns. This method is used in this study because it has been discovered that typical image filters for feature vectors work insufficiently well. As a result, five (5) distinct deep neural network image embedders — SqueezeNet, Painters, InceptionV3, VGG16, and VGG19—are used in this study. For feature extraction, the various numbers of photos representing a specific student's iris and finger prints are fed into each of the five embedders. The name of the specific student, as found in the training set, is labelled on the numeric image descriptors (for both iris and finger print) received by the embedding. The distribution of extractable features across the embedders is presented in Table 3.1.

3.3.1 The SqueezeNet Network

In this case, the iris and finger print signals on each image instance are returned as 1000-no numeric vector attributes by the deep neural network Squeezenet. In comparison to other deep embedders, the SqueezeNet is reputed to be a quick and compact embedder for image identification with significantly lower computational overheads[3]. The model is pretrained on the well-known imageNet and achieves accuracy comparable to AlexNet with 50 times less parameters[5]. Three stages of a smaller network are built on top of its fire module, which is accomplished by exchanging 3x3 filters for 1x1 filters[5]. The remaining 3x3 filters and, finally, the network's late down-sampling receive an input reduction.

3.3.2 The InceptionV3

The InceptionV3 design outperforms V1 and V2. Batch normalization layers are used to speed up training, and factoring convolutions with larger spatial filters is also added, considerably increasing computing efficiency[2]. The inception module, which comes in 1x1 and 3x3 kernel sizes, is the primary construction piece. From each input image, the network extracts 2048 feature vectors.

3.3.3 VGG 16 and 19

Convolutional layers and the activation function rectified linear unit are used by the embedder family of 16 and 19[8]. With a filter dimension of 3x3, the VGG-16 and VGG-19 embedders, respectively, have 16 and 19 layers. From each signal input, both embedders extract 4096 feature vectors.

3.3.4 Painters

A deep convolutional network embedder called Painters has been trained to identify painters from artwork images. Its inclusion is justified by its capacity to recognize patterns in creative image signals[1], which might be used to describe finger print images in this situation. For embedding purposes, the network's penultimate layer is activated. The deep embedder extracts 2048 feature properties from each input of iris and fingerprint data.

Table 3.1 below shows the number of extracted image descriptor by the five (5) image embedders from each signal of iris and fingerprint.

Table 3.1: Distribution of Attributes Across the Embedding Networks
(Researcher, Olomola B. 2023)

S/N	Pre-Trained Deep Embedder	Size of Extracted Image Descriptor From Each Image Signal	Eventual Training Set For Bimodal Features (Iris & Finger Print)
1.	SqueezeNet	1000	2000
2.	Inception v3	2048	4096
3.	VGG16	4096	8192
4.	VGG19	4096	8192
5.	Painters	2048	4096

3.4 Synthethic Minority Oversampling

In machine learning, resampling the training dataset is a crucial factor to take into account in order to overcome the issue of underrepresentation of a specific data feature, also known as over-fitting. This study will use the Synthetic Minority Oversampling Technique (SMOTE), an oversampling strategy that will scale-up the class distribution of the less-represented data characteristics. This eliminates the over-fitting issue that might occur in the training set. By concentrating on the feature space, SMOTE does a random oversampling to produce new examples by interpolating between nearby instances of the less-represented classes. The number of oversampling classes, *N*, is determined, and is susceptible to tweaking as necessary. A random example from a less-represented class is chosen to start the loop, and its *KNN* is then calculated. The freshly created synthetic instances are interpolated using the N of the K instances. The distance between a feature vector and its neighbours is calculated using the *aby* distance metric, multiplied by a random number between (0,1), and then added to the preceding feature vector.

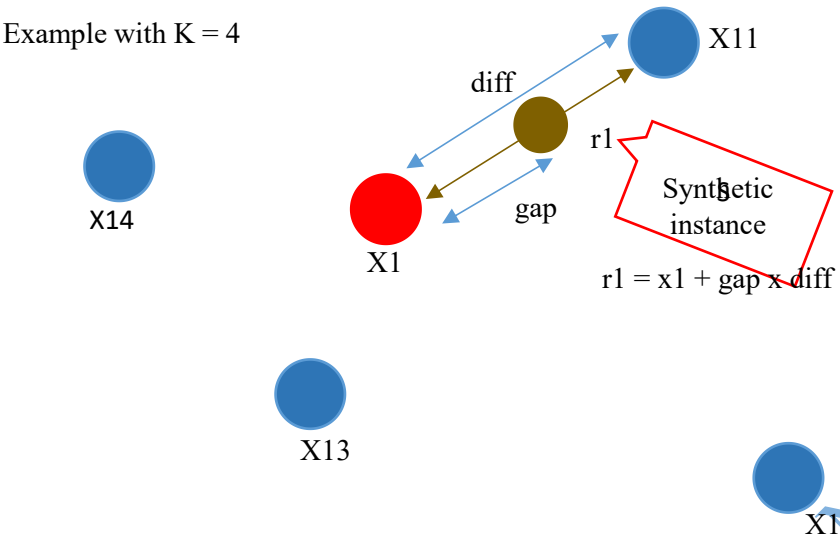

Figure 3.2: SMOTE Framework[9]

Algorithm 1. The Synthetic Minority Oversampling Technique

Input: number of minority class **T**; intended synthetic amount **N%**; Number of nearest neighbors **k**

Output: : **(N/100)** x **T** (synthetic versions of minority class)

4. **if** N < 100
5. **then** randomize the minority class (T)
6. T = (N/100) x T
7. N = 100
8. **Endif**
9. N = (int) x (N/100)
10. k = number of nearest neighbors
11. numattrs = attributes number
12. sample[ ][ ]: initial minority class in array

13. newindex: number of synthetic samples generated (initialized to 0)

14. synthetic[ ][ ]: synthetic sample array

15. for i ← 1 to **T**

16. Compute **k** nearest neighbors for i, and save the indices in the *nnarray*

17. Populate(N, i, nnarray)

18. Endfor

19. while N ≠ 0

20. Choose random number between 1 and **k**, call it **nn**. (This step chooses one of the k nearest neighbors of i.)

21. **for attr** ← 1 to **numattrs**

22. Compute: dif = sample[nnarray[nn]][attr] − sample[i][attr]

23. Compute: gap = random number between 0 and 1

24. synthetic[newindex][attr] = sample[i][attr] + gap ∗ dif

25. **Endfor**

26. **newindex++**

27. N = N – 1

28. Endwhile

29. return (∗ End of Populate. ∗)

3.5 Predictive Analytics For Attendance Authentication

In this stage of the study, machine learning is used to compare the deep image embedders' performance to that of their deep learning algorithm counterparts. The Multilayer Perception (MLP) deep learning algorithm, the Decision Tree method, and the Sequential

Minimal Optimization (SMO) learning algorithm are used to deploy the deep learning models of machine learning for the training and testing phases of the machine learning predictive analytics. The balanced numeric image descriptors received earlier from the SMOTE phase to train, which are all obtained from the five image embedding networks, will be used by the three learner algorithms. In order to train the three algorithms separately, MLP, Decision Tree, and SMO, five datasets obtained from each of the networks that comprise balanced numeric descriptors of each student's iris and finger print are employed.

Sequential Minimal Optimization (SMO) Algorithm

Support Vector Machines (SVM) have recently seen an increase in deployment in studies because they have been empirically shown to provide good generalization performance on a variety of problems, including text categorization, pedestrian detection, face detection, pedestrian recognition, and character recognition[6]. However, an enhanced version of the SVM called the SMO has been developed[4]. It is faster, easier to use, theoretically simpler, and has better scaling qualities for complex situations than the standard SVM training process[7]. The SMO is an improved version of the conventional secure vector machine (SVM), which was primarily developed to address the secure vector machine's optimization issue. It differs from SVM methods in that it does not require a quadratic programming solution, making it an efficient technique for training SVM on classification problems suitable for sparse data sets.

According to research, SVMs can be made more efficient by breaking down a large quadratic programming problem into several smaller mini-problems. Once no more progress can be made on any of the sub-problems, optimizing each mini-issue minimizes

the original quadratic programming problem. Thus, the first quadratic programming issue is resolved. Since each sub-problem can have a consistent size, optimization through decomposition is feasible. As a result, decomposition in SMO is far more effective than SVM's quadratic programming strategy. Since SMO uses a dual-sized subproblem, each subproblem has a methodical approach. SMO is the only optimizer that explicitly leverages the quadratic form of the objective function and simultaneously applies the analytical solution, which is employed in this methodology for a superior outcome. Other approaches to handling the quadratic programming problem of SVM hold considerable potential.

Algorithm 2. The Sequential Minimal Optimization Algorithm

1: **Input**: Ground truth labels (x_t, y_t), t = 1, ..., n, and a small constant f.

2: **Output**: optimal output α.

3: $i \leftarrow -1; j \leftarrow -1$

4: $\nabla\tilde{} f(\alpha) \leftarrow 0$

5: **while** α is not optimal **do**

6: Unsystematically permute samples.

7: for $t \leftarrow 1, .., n$ do

8: $\nabla\tilde{} f(\alpha)_t = w^T x_t - y_t$

9: if $\nabla\tilde{} f(\alpha)_t < \nabla\tilde{} f(\alpha) - f$ and $t \in I_{low}$ **then** $\measuredangle I_{low}$ is defined in (4).

10: $\nabla\tilde{} f(\alpha)i \leftarrow \nabla\tilde{} f(\alpha)_t$

11: $i \leftarrow t$

12: **else if** $\nabla\tilde{}\ f(\alpha)_t > \nabla\tilde{}\ f(\alpha) + f$ and $t \in I_{up}$ **then** ⊿ I_{up} is defined in (4).

13: $\nabla\tilde{}\ f(\alpha)_j \leftarrow \nabla\tilde{}\ f(\alpha)_t$

14: $j \leftarrow t$

15: **end if**

16: **if** $i \neq -1$ and $j \neq -1$ **then**

17: update α_i and α_j according to (13)

18: update w according to (14)

19: update $\nabla\tilde{}\ f(\alpha)$ according to (18)

20: $i \leftarrow -1; j \leftarrow -1;$

21: **end if**

22: **end for**

23: **end while**

Multilayer Perceptron (MLP)

A feed forward family of artificial neural networks includes multilayer perceptron machine learning algorithms (ANN). MLP is a vague term that is sometimes used to refer to any feed-forward ANN and is frequently used to describe networks that contain many layers of perceptron that are activated at a threshold. The input layer, hidden layer, and output layer are the three node layers that must be present in an MLP. Other nodes are a neuron that uses a nonlinear activation function in addition to the input nodes. Back-propagation for training, a subset of machine learning that focuses on supervised learning, includes MLP. MLP differs from a linear perceptron, which is renowned for

differentiating and classifying non-linearly separable datasets, because it has many layers and the non-linear activation feature.

The perceptron learns by changing the connection weights as soon as each data point is processed and is based on the amount of error in the output that is equivalent to the desired outcome. The linear perceptron is a streamlined version of the least mean squares algorithm. The formula ej(n) = dj(n) - yj(n), where d is the goal value and y is the perceptron-generated value, can be used to express the degree of error in an output node j in the nth data point.

The node's weights are then changed in accordance with corrections by minimizing the output's overall error, which is accomplished by

$$\Sigma(n) = \frac{i}{2} \sum_{j=0} e \quad ^2(n) \tag{2}$$

change in each weight is computed using gradient descent thus

$$\Delta wji(n) = -\eta \left(\frac{\delta \Sigma(n)}{\delta vj(n)} \right) yi(n) \tag{3}$$

where y_i represents output of previous neuron, with η as learning rate. The learning rate is designated to guarantee the weights eventually converge, without oscillations, to a response. The calculated derivative largely depends on induced local field v_j.

Decision Tree Algorithm

Another supervised learning technique is the decision tree algorithm, and unlike other classification learners, it may be used to address both classification and regression issues. By learning straightforward decision rules derived from previous data, a Decision Tree is used to build a training model that may be used to predict the class or value of the target

variable (training data). In decision trees, the prediction of the ground truth begins at the tree's root, and there is comparison between the root attribute's values and the record's attribute's values. Based on the comparison, the branch corresponding to that value is followed, and then the node after that is selected. The following are crucial decision tree parameters:

1. **Root Node:** The entire population is represented by the root node, and it is further divided into homogenous sets.
2. **Splitting:** A node is split by splitting it into sub-nodes.
3. **Decision Node:** A sub-node becomes a decision node when it divides into more sub-nodes.
4. **Terminal Node:** The terminal node is the node that could not be divided any further.
5. **Pruning:** Pruning, which is the opposite of splitting, is the removal of sub-nodes from a decision node**.**
6. **Sub-Tree:** A branch or sub-tree is a division of the overall tree.
7. **Parent and Child Node:** A parent node of sub-nodes is a node that has sub-nodes, whereas sub-nodes are the children of a parent node.

With a trainset $X = x_1, ..., x_n$ and ground truths $Y = y_1, ..., y_n$, a random sample with replacement of the trainset is selected and fits trees to the data sample thus:

Given $b = 1, ..., B$: which is an optimal number of trees,

1. Sampling with replacement is done with n training examples from X, Y referred to X_b, Y_b.
2. A classification methodology is then used to train the tree f_b on X_b, Y_b.

Consequently, predictions for samples x' is achieved by averaging predictions from trees on x' thus:

$$f = \frac{1}{B} \sum_{b=0}^{B} fb\,(x') \tag{4}$$

or by taking the majority vote in the case of classification trees.

Vote Ensemble Machine Learning

An ensemble machine learning model called a "vote ensemble" combines the functional predictions of various base learner algorithms in order to outperform the standard single model in terms of performance metrics. The various weighted sets of the training instance datasets are used to fit the individual base classifier predictors. In order to create a reliable binary classification model, the training, validation, and test datasets are used in a collaborative manner. Like image recognition, ensemble learning is a popular and extensive research topic in machine learning, which ultimately helps to improve performance capability. The predictors, which are frequently heterogeneous in type, fit a fresh data instance as shown in Figure 4, and then they each make their predictions.

The ensemble's vote would be based on the majority decision, and the result would be the ensemble model's anticipated value. Hard voting or soft voting could both use vote ensembles. In soft voting, the classification decision is returned as the majority vote after the classifiers' probabilities with regard to it are averaged. In hard voting, the vote ensemble's outcome is determined by the majority of votes, as seen in Figure 3. By integrating the MLP, Decision Tree, and SMO base learners into the Vote ensemble approach, this study applies ensemble machine learning to address the under-fitting issue that has been previously discovered in the literature.

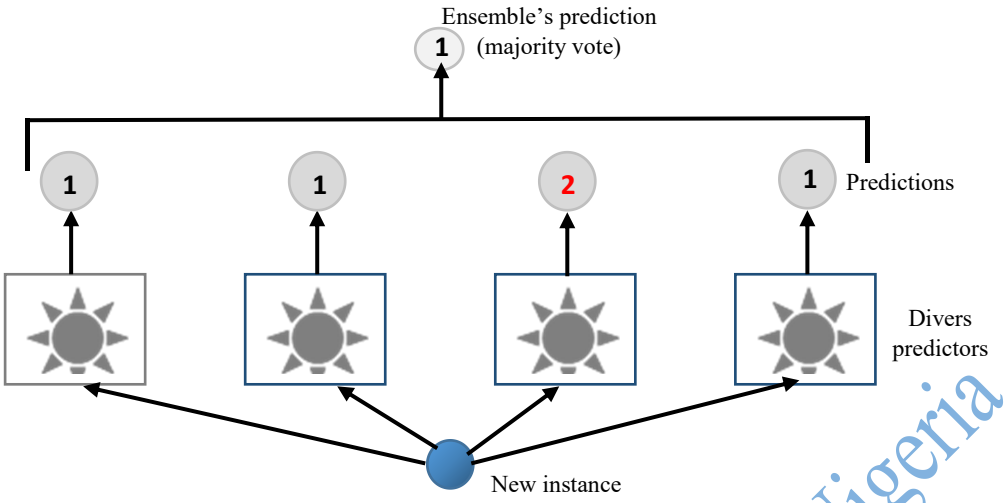

Figure 3.3: Framework of the Vote Ensemble (Researcher, Olomola. B. 2023)

3.5 **Performance Evaluation**

The final phase of the framework, as presented in Figure 1, is the evaluation phase. The study will be assessed using a dual method that includes model testing using data inputs and fundamental machine learning performance metrics like precision, accuracy, recall, F-measure, AUC and ROC, among others, as presented in Table3.2. The performances of Decision Tree, SMO, MLP, and the Vote ensemble of the three algorithms will be evaluated on each of the five datasets obtained from the five image embedding networks of SqueezeNet, InceptionV3, VGG16, VGG19, and the Painters.

Table 3.2: Model Evaluation Metrics

Metric	**Description**
Precision	Percentage of positive predictions correctly predicted
F-Measure	Harmonic mean of precision and recall. Examines trade-offs between the metrics

Received Operating Characteristic (ROC)	Curve plotted to facilitate investigating the cost-benefit of possibly optimal models
Area Under the ROC Curve (AUC)	Likelihood of sureness in a model to precisely predict positive outcomes for actual positive instanes
False Positive Rate (FPR)	Percentage of actual negatives predicted as positives. Signifies the significance level of a model
True Negative Rate (TNR)	Percentage of actual negatives that are correctly predicted. Represents the specificity of a model
False Negative Rate (FNR)	Percentage of actual positives predicted as negatives. Inversely relational to the statistical power of a model
Accuracy	Percentage of accurate predictions in the entire population. Not reliable for skewed class-wise data hence class balancing of this model

Researcher, Olomola B. (2023)

Endnotes

1. M. C. Chiu, G. J. Hwang, L. H. Hsia, & F. M. Shyu. "*Artificial Intelligence-Supported Art Education: A Deep Learning-Based System for Promoting University Students' Artwork Appreciation and Painting Outcomes*." ***Interactive Learning Environments*****,** 2022.1-19.

2. P. Ghimire, P. Sweekar & A. M. Gurung. "*Comparative Study of Face Mask Recognition using Deep Learning and Machine Learning Classifiers*." **International Conference on Innovative Computing, Intelligent Communication and Smart Electrical Systems (ICSES). IEEE,** 2021. 1-9.

3. B. Koonce. "*SqueezeNet." Convolutional Neural Networks with Swift for Tensorflow.*" **Apress,** 2021. 73-85.

4. A. K. Jumani, M. H. Mahar, F. H. Khoso, & M. Memon. *"Online Text Categorization System using Support Vector Machine."* **Sindh University Research Journal-SURJ (Science Series)** 50, no. 1, 2018. 85-90.

5. H. J. Lee, I. Ullah, W. Wan, Y. Gao, & Z. Fang. "*Real-Time Vehicle Make and Model Recognition with the Residual Squeeznet Architecture* ." **Sensors** , 2019.

6. M. B. Abdulrazzaq & J. N. Saeed. *"A Comparison of Three Classification Algorithms for Handwritten Digit Recognition."* **International Conference on Advanced Science and Engineering (ICOASE)**, **IEEE,** 2019. 58-63.

7. D., Devikanniga, A., Ramu, & A. Haldorai. "*Efficient Diagnosis of Liver Disease using Support Vector Machine Optimized with Crows Search Algorithm".* **EAI Endorsed Transactions on Energy Web**, 7, no. 29, 2020. e10-e10.

8. M. Shaha & M. Pawar. "*Transfer Learning for Image Classification."* **2nd International Conference on Electronics, Communication and Aerospace Technology** *(ICECA).* IEEE, 2018. 656-660.

9. S. Satpathy. *Home page.* **http://www.analyticsvidhya.com/blog/2020/10/overcoming-class-imbalance-using-smote-techniques**, 2020.

Chapter Four

Results and Discussion of Findings

4.1 Implementation Environment

This chapter discusses the implementation procedure and techniques used for the execution of the conceptual framework presented in chapter three. The results of the predictive analytics are likewise discussed in the chapter, while the evaluation of the model is discussed using the performance metrics of the machine learning model.

The machine learning framework is implemented using libraries of the object-oriented Python programming language, executed on the Jupyter notebook environment of the Anaconda Navigator. The TKinter library is used to design the user interphase of the framework while the Sklearn library is employed for the machine learning and binary classification problem of this study. The model evaluation is also executed by the library. Numpy library is deployed for handling the multi-dimensional array and matrices occasioned by the image embedding with a high level of mathematical functions to operate on the arrays. The TensorFlow library is also employed in the study as its supports traditional machine learning models. The Pandas library is employed for data analysis including transformation, cleaning and visualizing data points while the Scikit-learn (Sklearn) is the most useful and robust statistical machine learning modelling tool for clustering and classification.

4.2 Data Acquisition

The training and test set for this study which were divided into 880 and 580 respectively were 800 fingerprints and 800 iris biometrics images obtained from

the internet. Meanwhile, the primary data were 80 fingerprints and iris biometrics images obtained from individuals randomly chosen and captured with the the aid of an Android Infinix Note 8i mobile device. The fingerprints are captured on white background papers, while the iris are captured directly all in broad day light. The entire set is split between training and test set in order to address a germane objective of this study which is to test the performance of the model with a balanced dataset. The images are all in Joint Photographic Expert Group (JPEG) format and their sizes ranges from 27kb to 20kb, in greyscale format. The acquired fingerprint and iris dataset in JPEG are imported into the Orange data mining toolkit for feature extraction. The five deep learning networks for image embedding, described earlier in Table 3.1, are employed for the feature extraction. The images are passed through each of the five embedders to generate numeric vector attributes to return five different and unique numeric datasets, which will be used in turn to train the machine learning algorithms.

4.2.1 Hardware Requirements

The Microsoft Windows 10 operating system is used for the implementation in the integrated development environment. The following are the minimum requirements to run the software;

i. Core i3 processor
ii. 8GB RAM
iii. 100GB Hard disk
iv. CD-ROM Drive (80X)
v. SVGA Monitor
vi. Standard or Enhanced Keyboard

4.2.2 Software Requirements

i. Microsoft Windows 7 Operating System

ii. Anaconda Navigator IDE

iii. Jupyter notebooks

iv. Python libraries

4.3 Result

This section introduces stages of the Python implementation procedure with various graphical representations of the diagnosis system. Screenshots are used to display the results of different segments of the framework. The fingerprint and iris images of each collected user samples are sent for image embedding for the extraction of numeric feature vectors by the five deep learning networks earlier presented in Table 3.1, before the machine learning phase. The feature extraction is implemented on the Orange Data Mining Toolkit whose execution widgets are shown in the screenshot of Figure 4.1. The various Python libraries imported for the machine learning phase as earlier described in Chapter 3 are also shown in this section alongside computations of the Confusion_matrix, for model evaluation. The importation of the various libraries for implementation can be observed on the screenshot in Figure 4.2. The code snippet for the implementation of the TKinter user interphase is depicted on the screenshot of Figure 4.3, while the Pandas library that imports the numeric vectors of the training set for each of the five training sets is presented in the code snippet of Figure 4.4. Figure 4.5 presents the code snippet for the Synthetic Minority Oversampling Technique (SMOTE) and the machine learning and training phases are implemented through the code snippet presented in Figure 4.6. The

Vote Ensemble Machine Learning Model is implemented with the code snippet presented in the screenshot of Figure 4.7.

Figure 4.1: Screenshot of the Orange Data Mining Toolkit for Image Embedding (Researcher, Olomola, B. 2023)

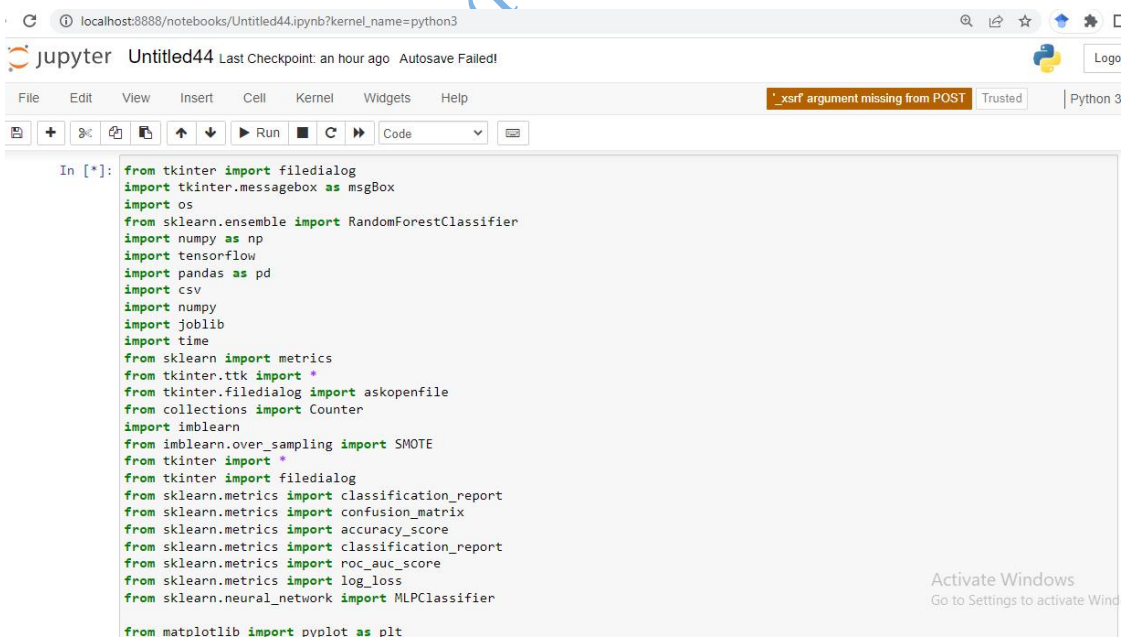

Figure 4.2: Code Snippet of Python Libraries Imported for the Implementation (Researcher, Olomola, B. 2023)

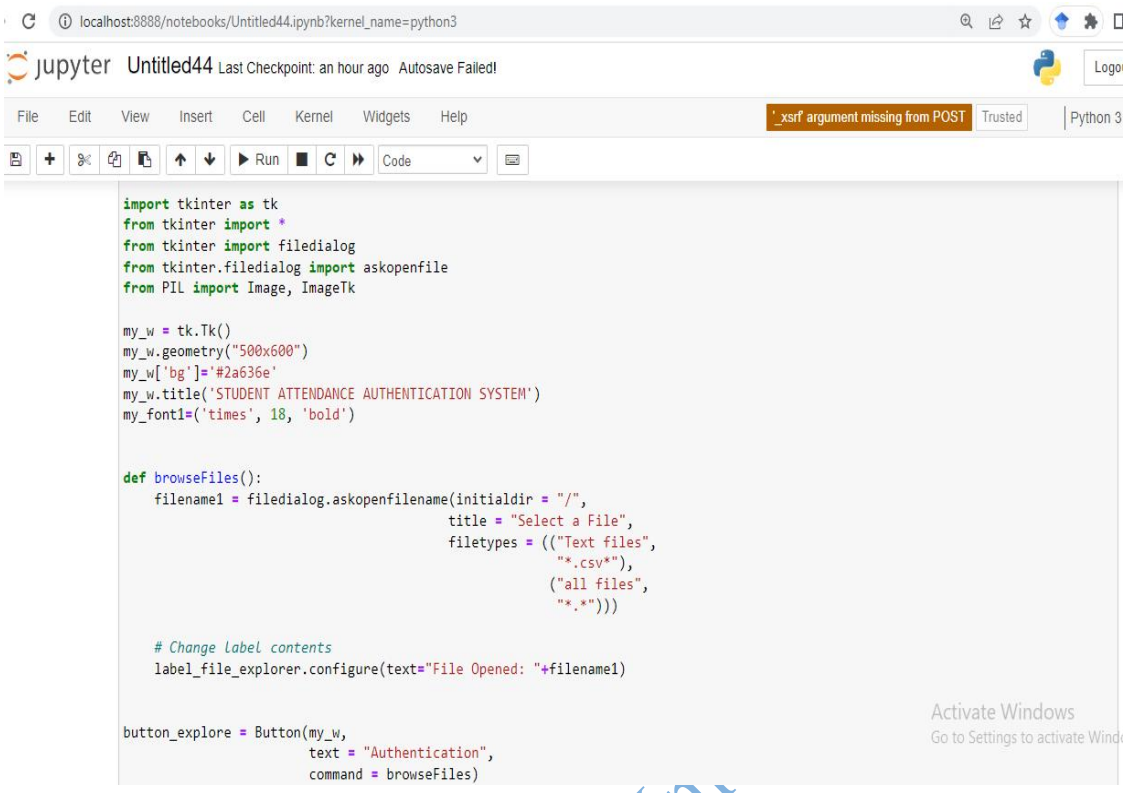

Figure 4.3: Code Snippet for the Graphical User Interphase Implementation (Researcher, Olomola, B. 2023)

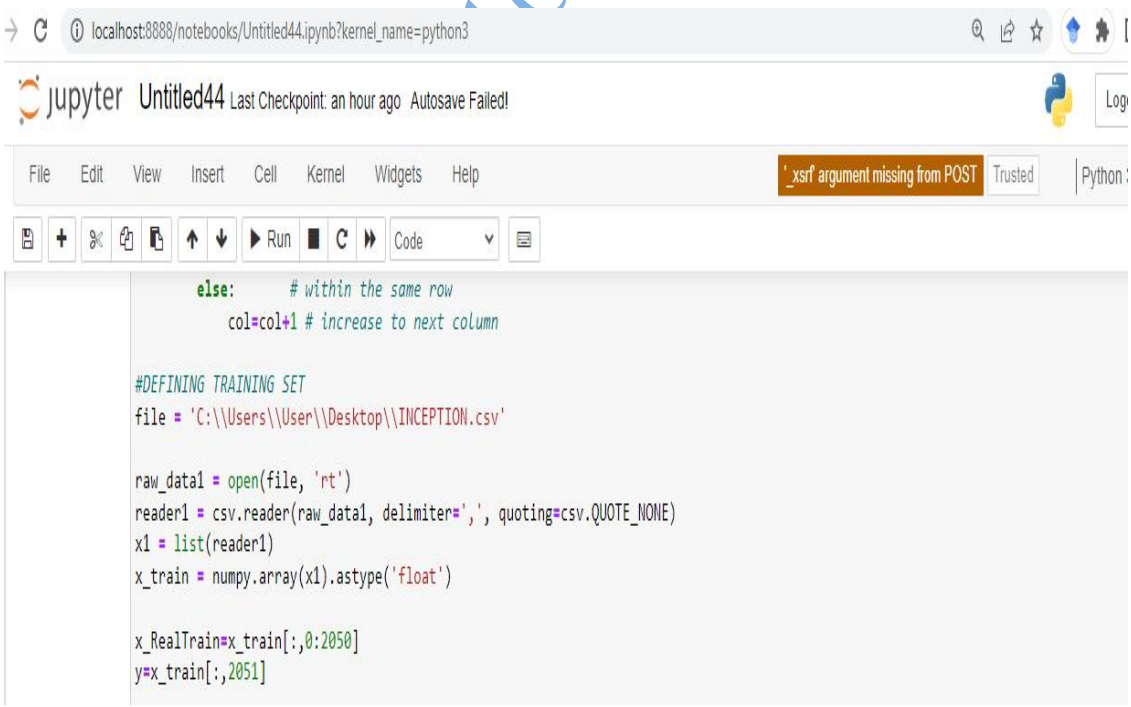

Figure 4.4: Code Snippet of the Pandas Library Used to Import the Training Set (Researcher, Olomola, B. 2023)

```
counter=Counter(y)
print('Before',counter)

smt=SMOTE()
x_sm,y_sm=smt.fit_resample(x_RealTrain,y)

counter=Counter(y_sm)
print('After',counter)

clf=RandomForestClassifier(n_estimators=100)

#FIT DATA ONTO THE MODEL
clf.fit(x_sm,y_sm)
```


Figure 4.5: Code Snippet of the SMOTE Function for Data Resampling
(Researcher, Olomola, B. 2023)

```
#FIT DATA ONTO THE MODEL
clf.fit(x_sm,y_sm)

#.....
#DEFINING TEST_SET

file_Test = 'C:\\Users\\User\\Desktop\\TEST.csv'
raw_data2 = open(file_Test, 'rt')
reader2 = csv.reader(raw_data2, delimiter=',', quoting=csv.QUOTE_NONE)
x2 = list(reader2)
x_test = numpy.array(x2).astype('float')
x_RealTest=x_test[:,0:2050]

y_pred = clf.predict(x_RealTest)

prediction=pd.DataFrame(y_pred, columns=['Attendance Identity']).to_csv('C:\\Users\\User\\Desktop\\INCEPTPREDICT.csv')
```

Figure 4.6: Code Snippet for the Training and Testing of the Imported Training Set
(Researcher, Olomola, B. 2023)

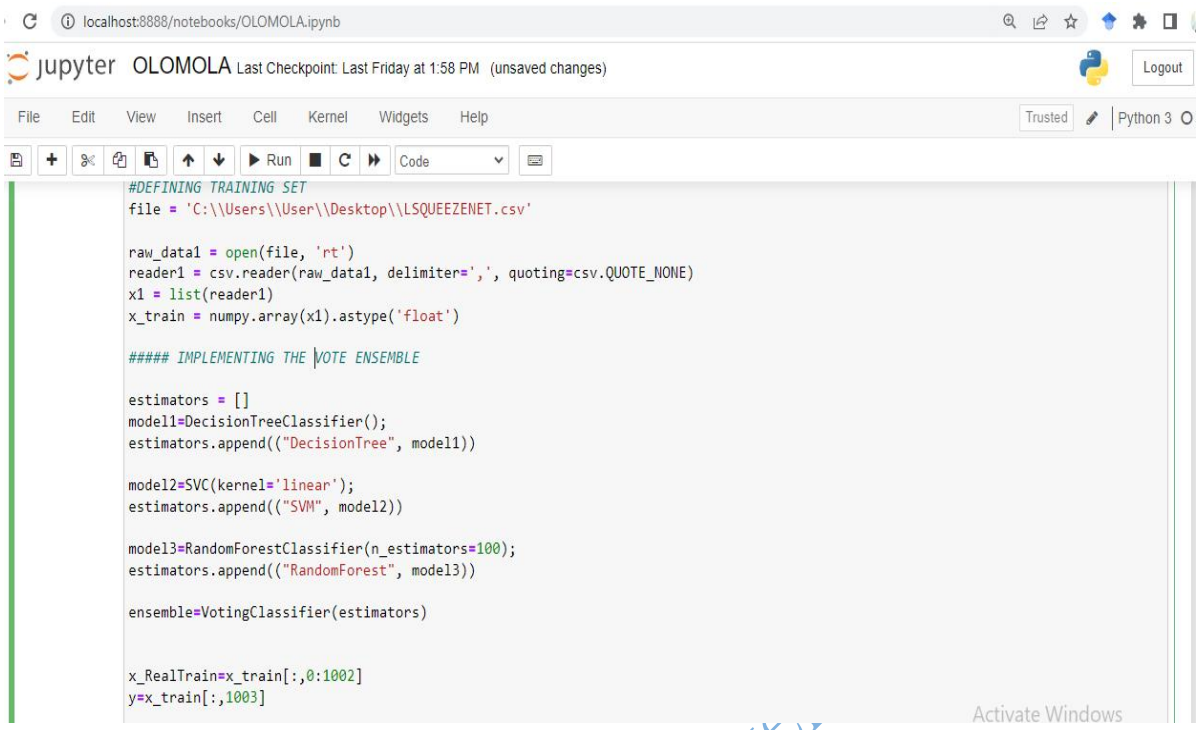

Figure 4.7: Screenshot of Code Snippet for the Vote Ensemble Modelling (Researcher, Olomola, B. 2023)

4.4 Discussion of Findings

Subsequent to the collection of fingerprint and iris biometric features of users, with sample presented in the screenshot of Figure 4.8, the data is forwarded to the severs of the five image embedders to extract numeric feature vectors through the Orange data mining toolkit whose widget is earlier presented in Figure 4.1. The distribution of the acquired fingerprint and iris are unbalanced across the biometrics of students captured as observed in Figure 4.9, before the image embedding is carried out. The extracted numeric vectors from the pair of fingerprint and iris of each unique student, as extracted by each of the five image embedders, is presented in Figure 4.10.

However, the SMOTE resampling technique is employed to ensure a balanced feature representation which is then presented in the graph of Figure 4.11. The three algorithms

of SMO, SVM, and Decision Tree are then trained with each of the five numeric vector sets. The Vote ensemble model is likewise trained on each of the five numeric vector databases. The resulting four intelligent models are activated for authentication phase through the Student Attendance Authentication System (SAAS) implemented by the TKinter earlier discussed, and presented in Figure 4.12. The *Upload Input* button of the SAAS takes in the numeric vector representations of students through the widget presented in Figure 4.13. The uploaded fingerprint and iris of students is displayed on the SAAS frame of Figure 4.14. The inputs subjected to image embedding also to extract their numeric vectors which is then uploaded into the SAAS by the *Authentication* button. The outcome of the student identity authentication is automatically saved into a *CSV* file which comes up automatically upon the press of the *Check Result* button as shown in Figure 4.15 for the SVM model. The result of the authentication system for SMO, Decision Tree, and Vote ensemble are likewise presented in the screenshots of Figure 19, Figure 4.16, and Figure 4.17 respectively. The evaluation of the models are consequently conducted to ascertain their performances and the outcome of the Confusion matrix is as presented in Figure 4.19 for SVM, Figure 4.20 for SMO, and Figure 4.21 for Decision Tree, and Figure 4.22 for Vote ensemble. The matrixes are a 5*5 matrix of the multi-classification model for SqueezeNet dataset only. It shows the Predicted values of the model and the True values (actual values) of the given observations of the training set. The columns represent the True Positive (TP), False Negative (FN), False Positive (FP), and the True Negative (TN) compartments of the matrix.

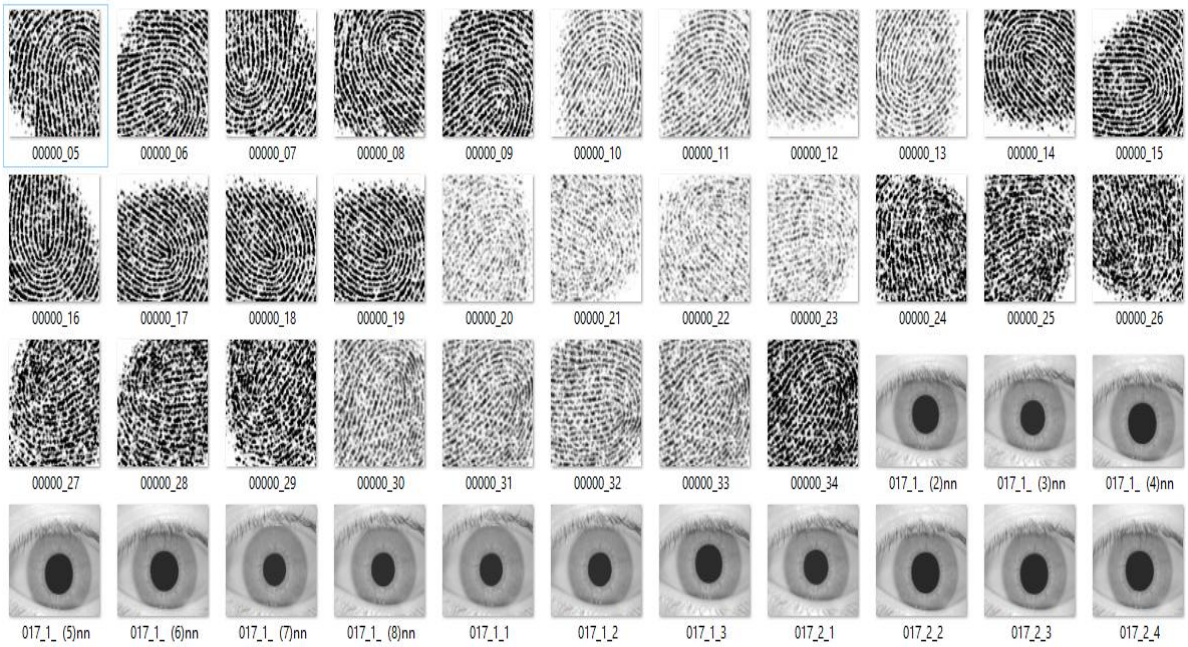

Figure 4.8: Screenshot of Acquired Fingerprint and Iris Dataset (Researcher, Olomola, B. 2023)

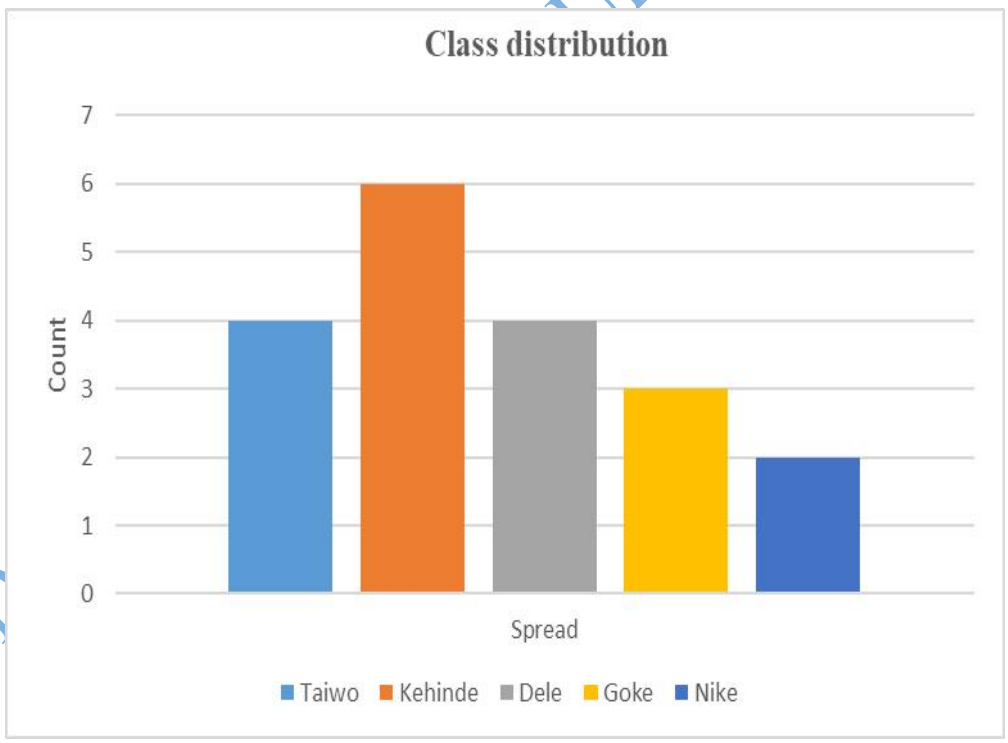

Figure 4.9: Graph Showing Unbalanced Class Distribution (Researcher, Olomola B. 2023)

3.121365072

A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T
1.621695	1.080107	-0.04582	-0.93544	4.993327	5.417585	3.87822	3.000086	2.865945	2.102606	0.219051	-1.11991	3.129022	1.732367	0.325192	2.661255	0.017301	4.037043	2.117922	-0.5
0.586875	-0.05379	-0.14816	-0.37768	5.077109	5.396108	5.094633	3.494626	2.886464	3.409795	1.99224	-0.13112	4.065132	2.459115	0.204643	2.488004	-0.13453	3.969847	4.117436	0.10
1.753717	3.044223	0.39014	-1.85587	4.421483	3.088923	3.285255	2.487846	2.970096	2.155586	-1.26349	-1.77441	3.029357	1.330288	-0.40461	2.501239	0.51086	2.543503	1.028806	-0.2
1.350713	3.62497	0.858815	-0.78051	6.609146	4.589913	4.156326	3.694974	3.830407	2.976902	-1.12262	0.375481	4.450412	0.411569	0.896197	2.043157	0.801629	4.414005	3.776101	0.2
1.047986	2.180639	-0.86254	-0.49301	4.733052	6.206831	3.47191	-0.1863	1.344296	0.394626	-1.9148	-1.92653	2.17422	1.514178	0.073686	0.2867	-0.40323	2.296195	4.234026	0.8
1.301711	2.292259	3.161784	1.674234	3.76217	5.947043	3.904703	-0.41791	-0.62705	0.351722	-0.04815	-0.92427	2.304631	3.250471	-0.20237	-0.41937	0.402484	2.089395	4.152262	4.43
1.243203	2.162687	2.020679	1.786509	4.453234	3.947092	1.413818	-1.89626	-2.42883	0.619186	0.96069	0.601286	0.842996	2.196859	-1.25018	-2.19089	-0.63271	0.461439	5.873984	3.89
2.374996	2.740097	2.463334	2.764909	3.905271	4.608775	1.789367	-2.48459	-1.79728	-1.65827	0.093531	0.651821	0.8107	2.610948	-2.26036	-2.59168	-1.22189	-0.35545	4.58707	1.78
2.571822	1.146581	2.267234	4.12666	6.192764	5.125176	2.492995	-1.90296	-1.96763	-0.92966	0.523287	1.209141	1.290028	2.254958	-2.8141	-3.4776	-1.84632	0.870271	5.312079	2.7
3.53105	0.74128	2.540934	3.987367	6.056505	5.240571	3.756372	-1.27626	-0.65949	0.425674	2.109539	1.968042	1.865282	3.318442	-2.33259	-1.37164	0.012341	2.050857	4.564716	3.20
1.107335	1.746091	1.394297	2.547097	6.582766	11.22733	10.95803	4.174493	6.251522	-0.38015	1.784641	1.304514	4.324563	3.619431	2.378	5.519973	0.907265	6.99444	5.343109	-0.9
1.63401	3.565997	1.957072	1.626199	9.337096	13.44643	10.42706	3.820699	6.123128	0.617153	-2.6908	0.24977	2.801645	2.208392	-1.87464	1.83208	-0.09909	3.59289	5.947627	-0.9
1.379685	3.22664	3.718019	3.271621	10.54876	13.68062	11.82469	3.753007	5.868204	1.912522	-2.50616	0.013485	1.513987	1.845888	-1.05607	-0.90469	0.343747	3.701868	4.81284	-0.4
1.686907	3.789677	5.438348	4.637577	10.76899	14.25843	12.48899	6.31203	8.105814	0.548572	-0.3409	2.554553	2.932244	3.696689	0.355556	1.075263	2.459245	6.734091	7.428319	2.07
1.865339	2.231119	5.061323	4.704207	8.640826	10.27599	10.35729	1.661172	4.513297	2.768399	-1.01082	0.438993	4.574837	2.882907	0.782924	2.016985	1.141555	3.867666	4.222229	1.50
1.523772	1.264634	3.202396	3.733238	8.062444	9.119662	9.098503	1.163248	3.855943	4.578848	-1.00396	1.416264	4.388765	1.861274	0.39225	1.388954	1.111071	3.122963	4.445051	1.81
2.435993	-0.12176	3.173773	4.850889	6.554578	8.898167	9.76737	1.296381	2.942876	5.443063	-0.37451	0.346414	3.566977	0.044941	-0.2546	1.596604	0.558614	2.648629	2.357885	-0.9
1.017074	1.30288	1.412287	2.440622	6.475621	8.919294	8.459331	4.167303	6.048209	-1.28574	1.226253	1.697269	3.401123	3.559049	1.478607	3.350242	0.818509	6.53416	4.980091	0.13
2.199433	1.962357	2.406196	3.474508	8.503108	11.74725	11.55846	4.171801	6.064548	0.835705	0.178439	0.95074	1.561018	3.136724	-0.21578	2.467348	0.575832	5.005235	6.589384	-1.2
2.107163	1.61867	1.390846	2.674432	7.697319	10.51632	9.503696	3.039096	4.328369	0.800406	-0.77903	-0.17863	0.392701	2.170711	-0.59935	1.996364	-0.23573	3.742217	5.30304	-1.0
2.084172	3.183272	1.680622	1.001858	8.694349	9.131172	7.236485	3.979826	5.988275	-0.75253	-1.18381	0.732374	3.26431	2.281527	0.51013	1.473094	0.700108	5.087208	6.836217	0.25
2.268039	3.683559	1.763063	1.259579	8.278327	8.964585	6.957859	3.065949	4.683384	-0.99867	-1.95323	-0.64372	2.239162	1.754493	-0.18925	1.140248	0.419241	4.171541	5.573948	0.15
2.06132	4.880807	1.632613	1.534298	8.749708	12.60818	10.12444	3.414337	4.884673	0.227815	-2.19978	-0.3611	1.951759	0.991183	-1.00263	1.98248	-1.4267	3.841252	5.591115	-1.9
1.54774	3.818959	0.923659	1.065075	8.982079	11.15941	9.139111	4.073384	5.182752	-0.72482	-1.77876	0.030323	2.438529	0.936112	-0.61248	0.829951	-0.65496	5.262624	6.937415	-1.5
1.327946	3.001758	1.207104	1.065084	9.018515	12.18772	8.810103	3.507488	4.892316	0.43261	-3.35809	-0.11805	2.081322	1.952101	-1.48231	0.696671	-0.83293	3.066837	4.684425	-1.6

Figure 4.10: Screenshot of Numeric Vectors after Feature Extraction
(Researcher, Olomola B. 2023)

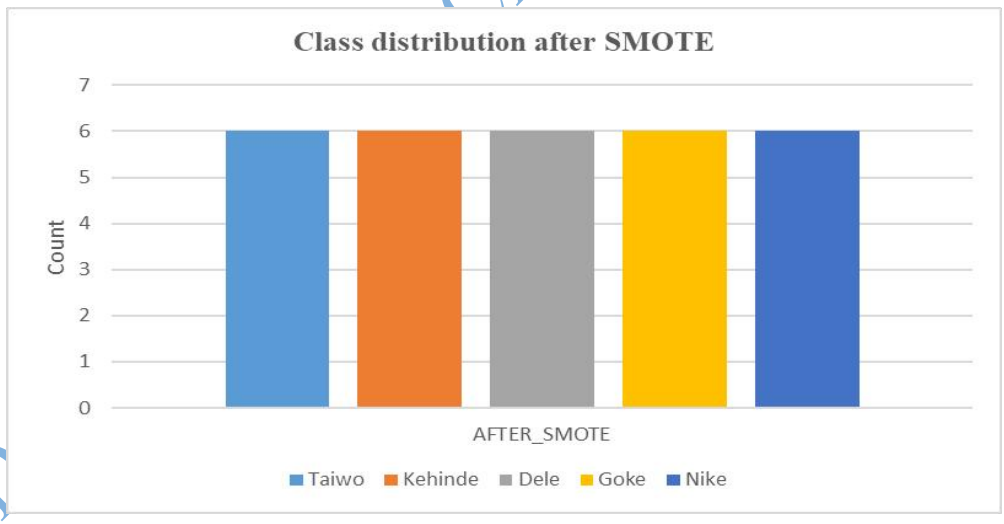

Figure 4.11: Graph Showing a Balanced Class Distribution after SMOTE Application
(Researcher, Olomola, B. 2023)

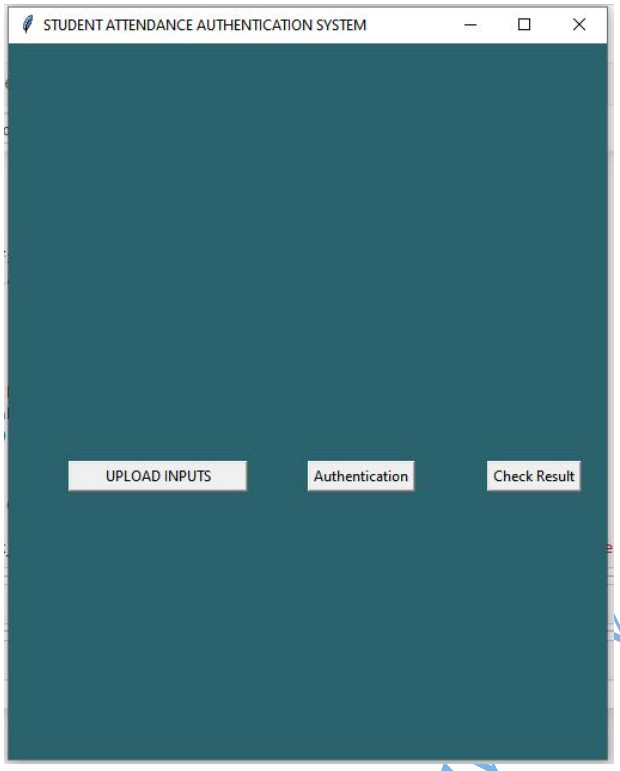

Figure 4.12: GUI of the Student Attendance Authentication System
(Researcher, Olomola B. 2023)

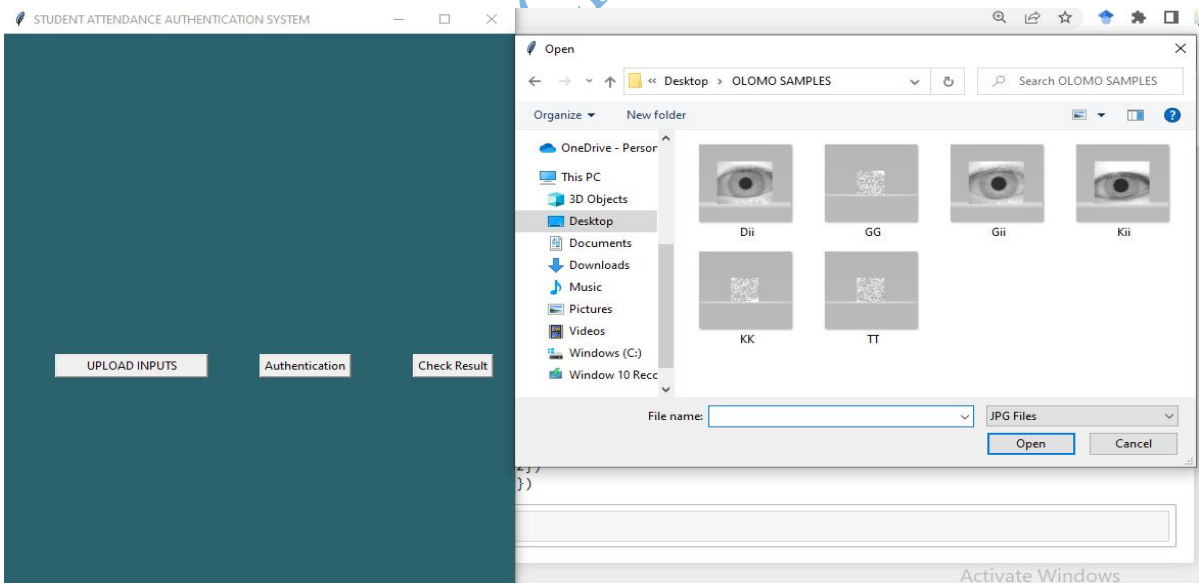

Figure 4.13: The Dataset Uploading Widget of the GUI (Researcher, Olomola B. 2023)

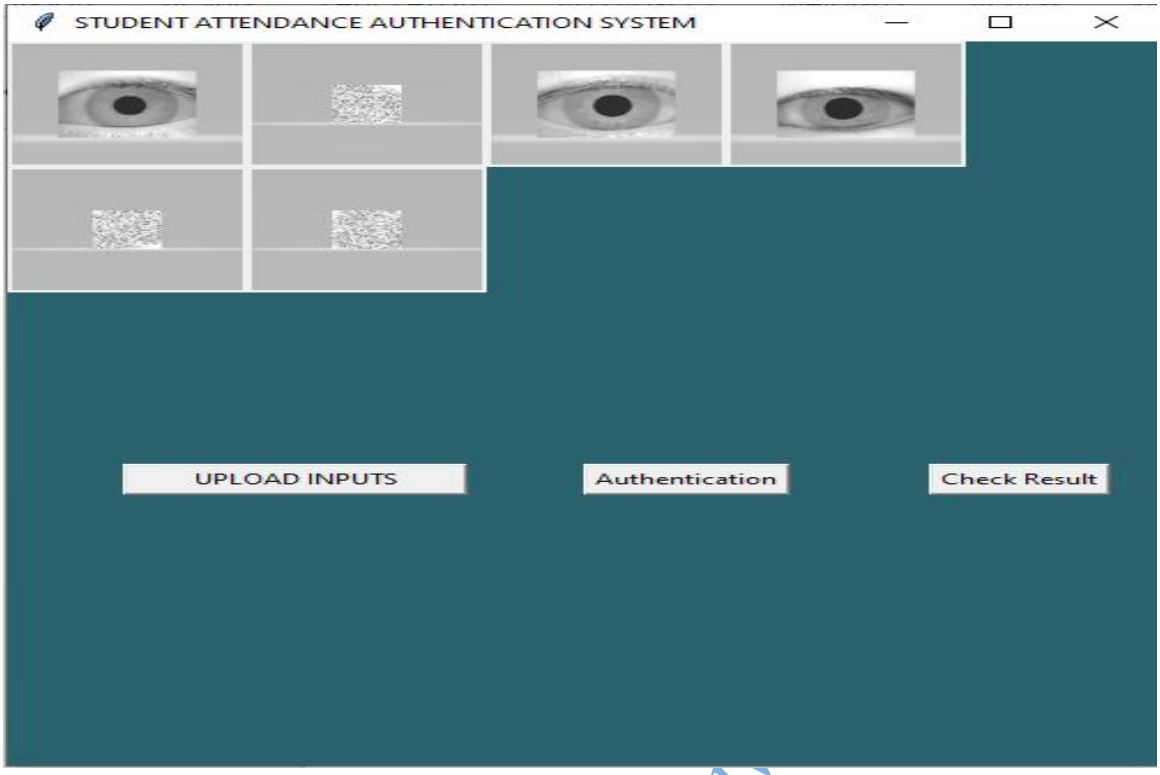

Figure 4.14: The GUI Frame Showing the Fingerprint and Iris Input for Authentication
(Researcher, Olomola B. 2023)

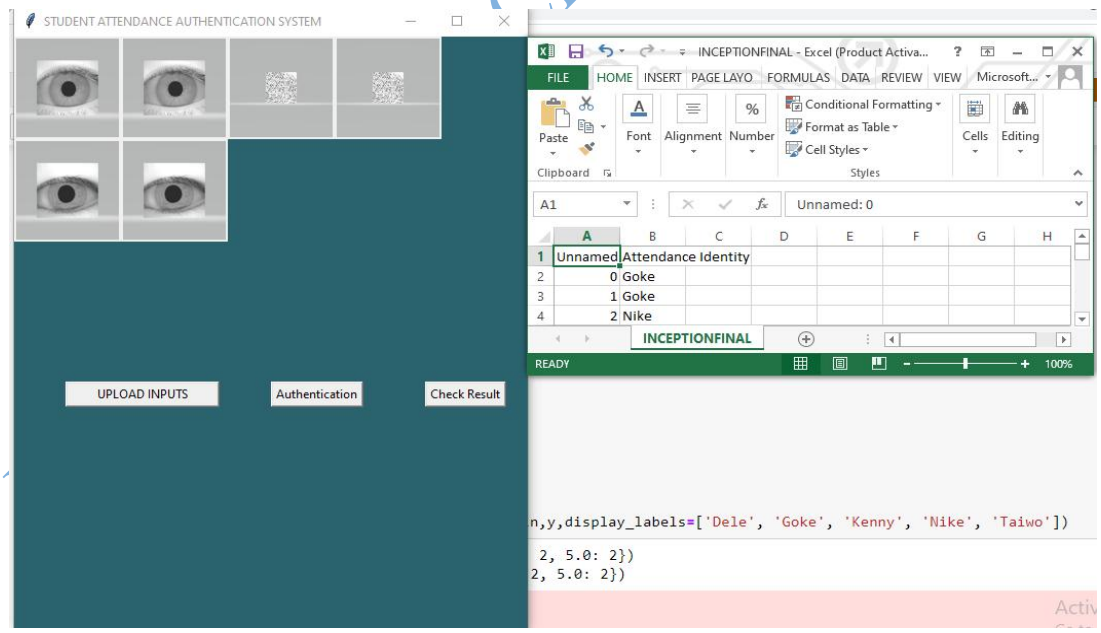

Figure 4.15: The Result of the SVM for Student Identity Authentication
(Researcher, Olomola B. 2023)

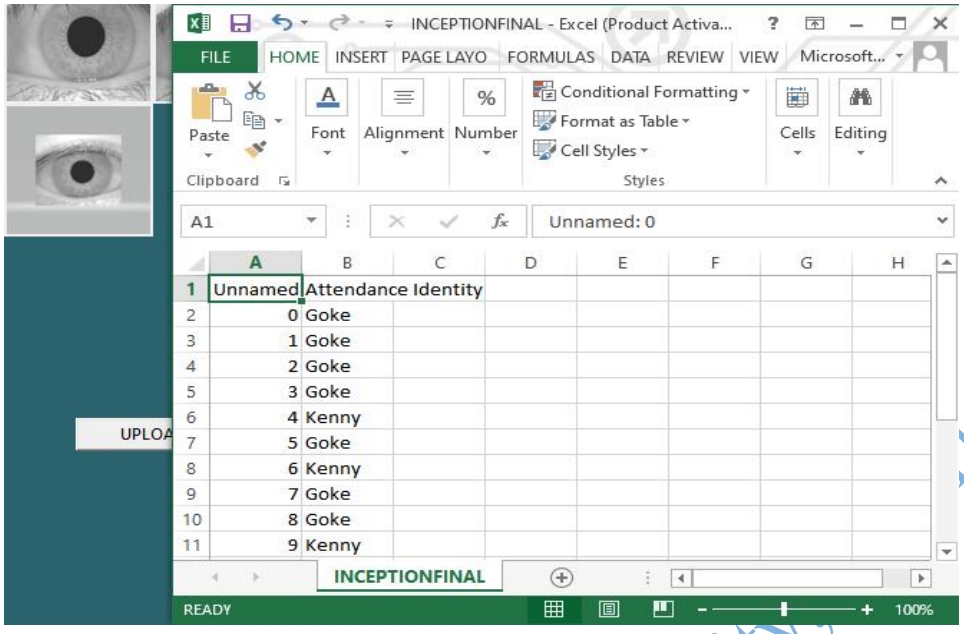

Figure 4.16: The Result of the SMO for Student Identity Authentication (Researcher, Olomola, B. 2023)

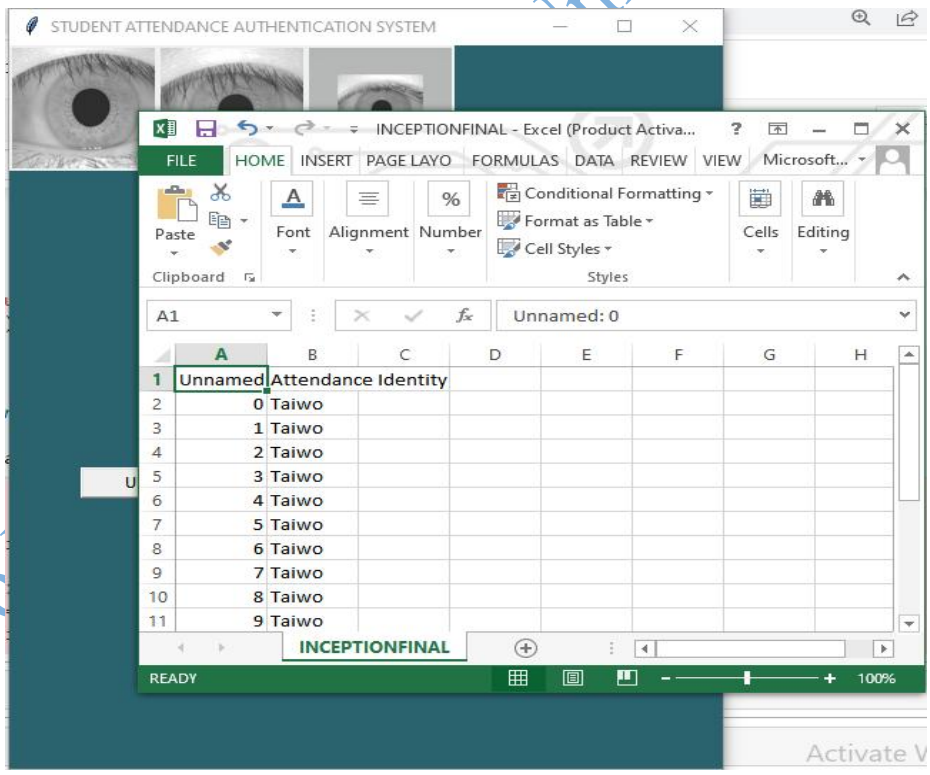

Figure 4.17: The Result Of The Decision Tree For Student Identity Authentication (Researcher, Olomola, B. 2023)

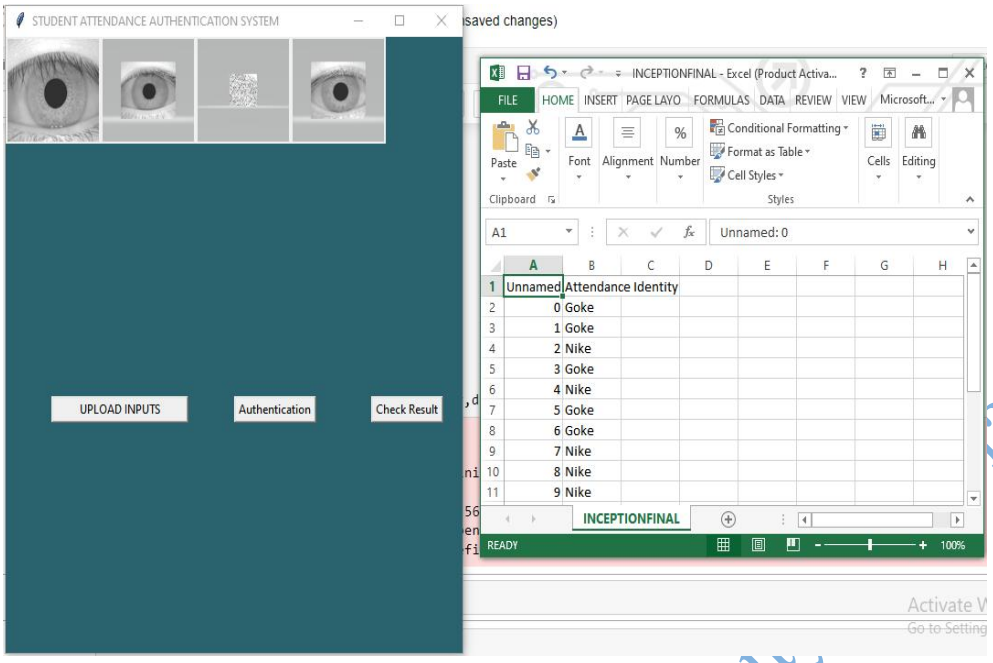

Figure 4.18: The Result Of The Vote Ensemble For Student Identity Authentication (Researcher, Olomola, B. 2023)

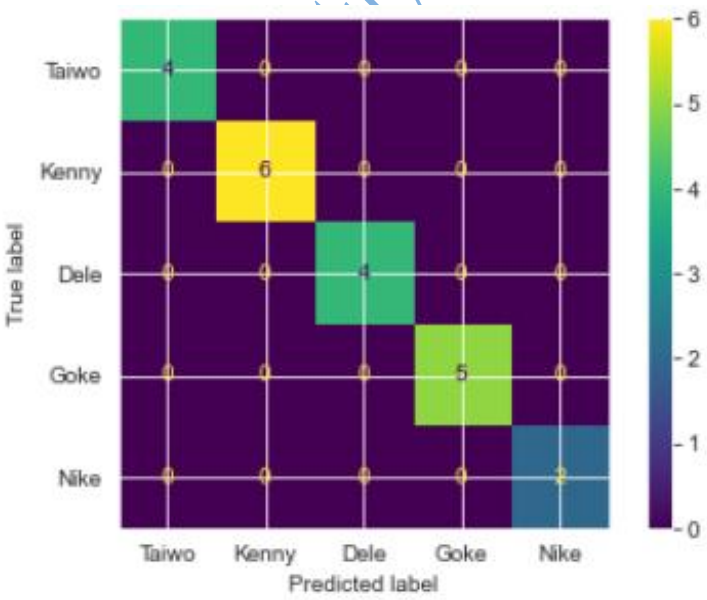

Figure 4.19: Confusion Matrix of SVM (Researcher, Olomola, B. 2023)

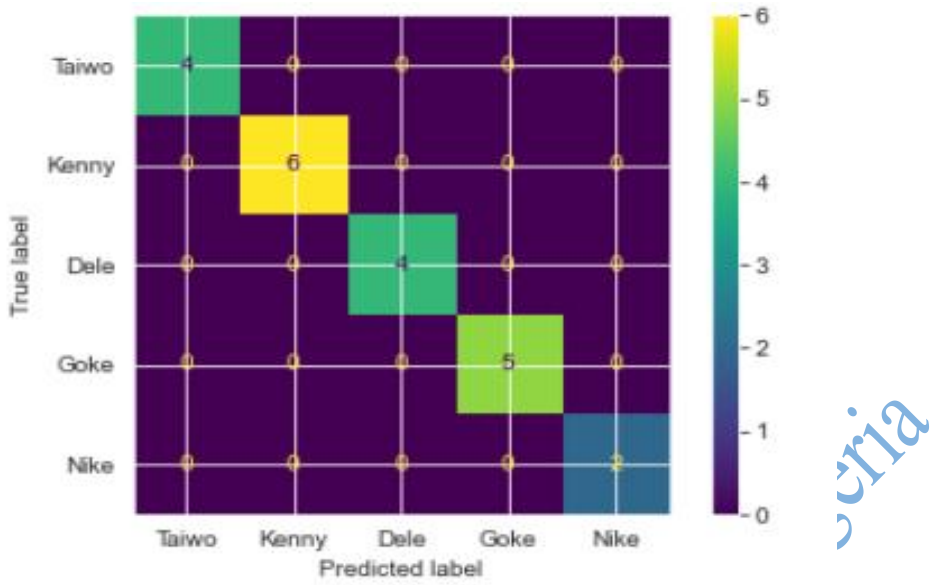

Figure 4.20: Confusion Matrix of SMO (Researcher, Olomola, B. 2023)

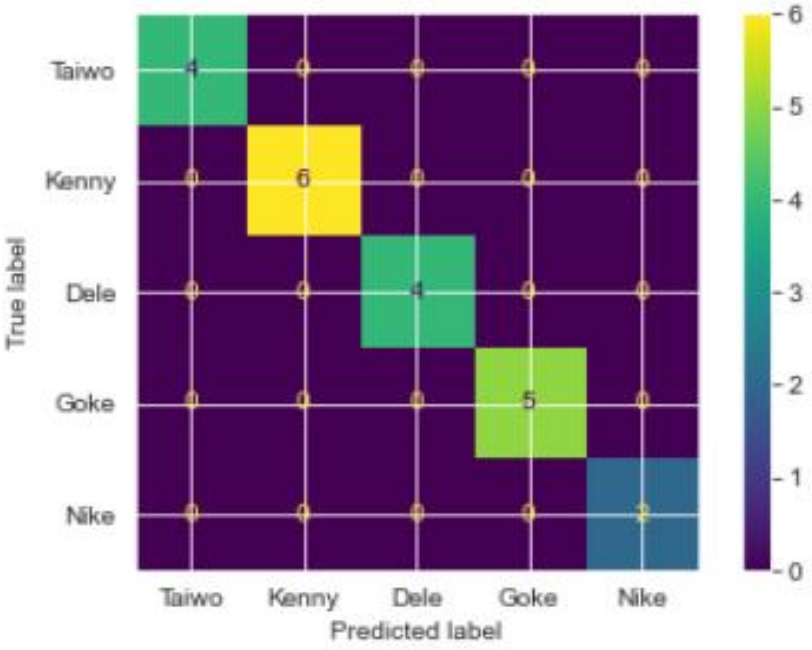

Figure 4.21: Confusion Matrix of Decision Tree (Researcher, Olomola, B.T. 2023)

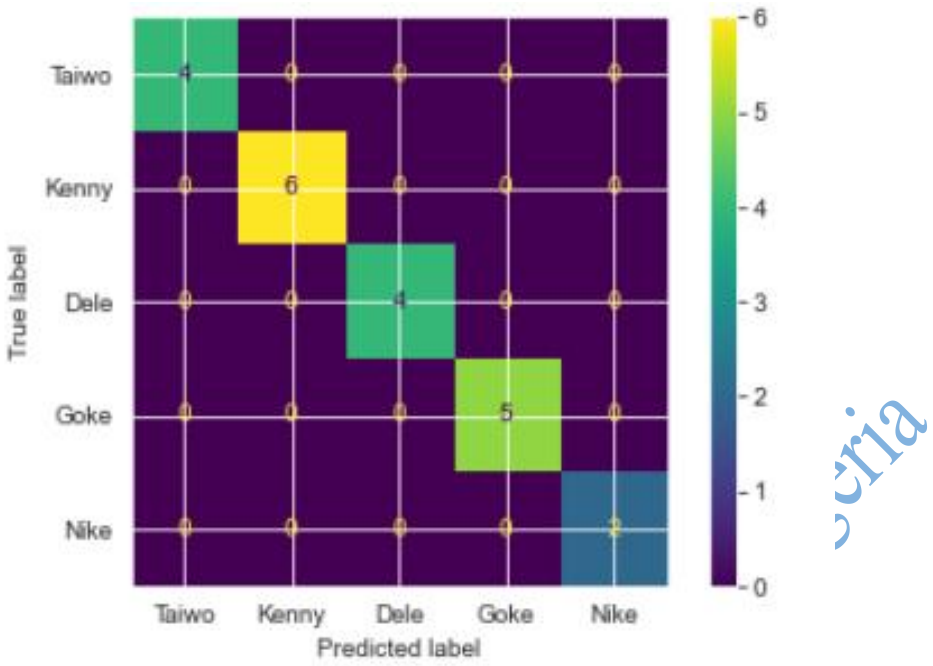

Figure 4.22: Confusion Matrix of Vote Ensemble (Researcher, Olomola, B. 2023)

4.5 Performance Evaluation

The performance evaluation of the three learner algorithms of SMO, SVM, and Decision tree are determined, including the Vote ensemble to discover how well the models could rightly authenticate student identity as implemented above. The calculation of F1 measure as recommended in Khaleel et. al. (2016) and Kong and Yan (2013), with other performance metrics will help determine the dependability and trustworthiness of the models. The True Positive (TP), True Negative (TN), False Positive (FP) and False Negative (FN) elements of the Confusion Matrix presented in Figure 4.19 to Figure 4.22 are used. The performance metrics are computed with their outcomes discussed below:

i. The **Accuracy** parameter returns the model's overall precision, which reveals the percentage of positive cases overall that the learner accurately predicted.

ii. **Precision**, however, establishes the percentage of identity authentication that establish the true identity of a student

iii. The **Recall** is the proportion of positive cases that are really expected to be positive

iv. The **Specificity** is the percentage of expected negative authentication

v. A better evaluation statistic than Accuracy is the **F1 score**, which combines Precision and Recall into a single metric.

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN}$$

$$Precision = \frac{TP}{TP+FP}$$

$$Recall = \frac{TP}{TP+FN}$$

$$F_1 = 2 * \frac{Recall*Precision}{Recall+Precision}$$

$$Misclassification\ rate\ (Classification\ Error) = 1 - Accuracy$$

$$Specificity = \frac{TN}{TN+FP}$$

where TN is the indicator of the model's outcome that successfully verified a student, TN is the indicator of the model's outcome that was correctly predicted, FP is the indicator of the model's outcome that was incorrectly predicted, and FN is the indicator of the model's incorrectly predicted outcome that was negative. From Table 4.1, weighted averages of the model's performance metrics are presented across the five datasets. The performance of the models in authenticating the identity of students clearly shows the better expertise of Vote ensemble over the other three base learners. However, Vote ensemble returned the best weighted average with the SqueezeNet database with the Precision, Recall, Specificity, and F1 Score netting 0.994, whereas the ensemble nets 0.999 with both

SqueezeNet and InceptionV3 respectively. The harmonic mean of Precision and Recall, which is a more trustworthy performance metric than Accuracy, is found via the F1 measure. From the experimental result, Vote ensemble returned the best F1 score performance across the datasets, which is closely followed by the SMO.

Table 4.1: Distribution of Weighted Averages across the Image Descriptors and Classifiers

Image Descriptors Dataset	Classifers	Accuracy	Precision	Recall	Specificity	F1 Score
SqueezeNet	*SMO*	0.990	0.991	0.991	0.991	0.991
	SVM	0.989	0.989	0.989	0.989	0.989
	DT	0.960	0.907	0.907	0.907	0.907
	Vote	0.999	0.994	0.994	0.994	0.994
InceptionV3	*SMO*	0.997	0.995	0.995	0.995	0.995
	SVM	0.989	0.989	0.989	0.989	0.989
	DT	0.959	0.960	0.960	0.960	0.960
	Vote	0.999	0.992	0.992	0.992	0.992
VG16	*SMO*	0.980	0.981	0.981	0.981	0.981
	SVM	0.979	0.979	0.979	0.979	0.979
	DT	0.960	0.907	0.907	0.907	0.907
	Vote	0.989	0.982	0.982	0.982	0.982
VG19	*SMO*	0.980	0.981	0.981	0.981	0.981
	SVM	0.979	0.979	0.979	0.979	0.979
	DT	0.960	0.907	0.907	0.907	0.907
	Vote	0.989	0.982	0.982	0.982	0.982
Painters	*SMO*	0.889	0.881	0.881	0.881	0.881

SVM	0.955	0.945	0.945	0.945	0.945
DT	0.959	0.945	0.945	0.945	0.945
Vote	0.959	0.960	0.960	0.960	0.960

(Researcher, Olomola, B. 2023)

Chapter Five

Conclusion

5.1 Summary of Findings

From Table 4.1, weighted averages of the model's performance metrics are presented across the five datasets. The performance of the models in authenticating the identity of students clearly shows the better expertise of Vote ensemble over the other three base learners. However, Vote ensemble returned the best weighted average with the SqueezeNet database with the Precision, Recall, Specificity, and F1 Score netting 0.994, whereas the ensemble nets 0.999 with both SqueezeNet and InceptionV3 respectively. The harmonic mean of Precision and Recall, which is a more trustworthy performance metric than Accuracy, is found via the F1 measure. From the experimental result, Vote ensemble returned the best F1 score performance across the datasets, which is closely followed by the SMO.

5.2 Conclusion

In this study, three learner algorithms of Decision Tree, Secure Vector Machine, and Sequential Minimal Optimization are trained with numeric vectors extracted from both fingerprint and iris biometrics of students for a student attendance authentication system. The numeric vectors are extracted using the SqueezeNet, InceptionV3, VG16, VG19, and Painters image embedders to return five distinct databases. The performances of the three base learners are evaluated alongside the performance of a Vote ensemble model after the five databases are subjected to a synthetic minority oversampling. The intelligent models trained by the five databases were tested on a training set of students' concatenated fingerprint and iris and the result is evaluated using performance metrics of the machine

learning models. Experimental results return the Vote ensemble as the best model for student authentication which is followed by the SMO. The F1 score of Vote ensemble outperforms other models across the five datasets, with accuracy score as high as 0.999. The synthetic minority oversampling of the training sets further improved the performance of the models through data resampling. Consequently, Vote ensemble machine learning is better deployed for student authentication systems with any of the five image embedders.

5.3 Recommendations

The followings are the recommendations from the study:

i. Multimodal Biometrics approach (at least Bimodal) is recommendable for any attendance authentication system.

ii. Experimental results from this study return the Vote Ensemble as the best model for student authentication, Vote Ensemble Model is therefore recommended for deployment to get better results in any Biometrics Authentication System.

iii. The pair of Squeeze Net image embedder with Vote Ensemble Model is discovered to be best suited for implementation of student's authentication system when using Iris and Fingerprint Biometrics and are therefore recommended.

5.4 Contributions to Knowledge

The study contributes to knowledge in the following ways:

i. Appraisal of existing studies on student authentication systems

ii. The realization that Vote Ensemble learning model with deep image embedder networks reliably returns a machine learning-based student authentication system

iii. The pair of SqueezeNet and Vote ensemble or SMO is better suited for fingerprint and iris-based authentication system

5.4. Suggested Areas For Further Studies

i. Future work could further evaluate the performances of Bagging and other Machine Learning Ensemble Models on the five deep image embedder networks i.e. The SqueezeNet, The InceptionV3, VGG 16, VGG 19 and Painters.

ii. This research took a bimodal approach and utilized Iris and fingerprint biometrics, further studies that will consider other biometrics approach e.g. Facial, Voice etc. is suggestible.

Bibliography

Books

Bolle, R. M., Connell, J.H., Pankanti, S., Ratha, N. K., & Senior, A. W. "*Guide to Biometrics*." Springer Science & Business Media, 2013. XXX, 364

Brian, K. W. & Stacey, C. S. "*Using Information Technology*" A Practical Introduction to Computers and Communications 12th Edition, 2019.

Cho, T. *"Iris Recognition Reaches the Mainstream for Identification, Authentication".* Biometric Update. 2022.

Dargan, S. & Kumar, M. "*A Comprehensive Survey on the Biometric Recognition Systems Based on Physiological and Behavioral Modalities*," Expert Systems with Applications, p. 113114, 2019.

David, M. K. & David, J. A. "*A Review of Database Concepts"* Prentice Hall New Jersey. 2016.

Dong, N., Zhao, I., WuJ, C.H., & Chang, F. "*Inception V3 Based Cervical Cell Classification Combined with Artificially Extracted Features*." Applied Soft Computing 1, no. 93 2020: 106311.

Faundez-Zanuy, M., Fierrez, J., Ferrer, M.A., Diaz, M., Tolosana, R. & Plamondon, R. *"Handwriting Biometrics: Applications and Future Trends in E-Security and E-Health."* Cognitive Computation 12, 2020. 940-953.

Graeme, C. S. & Graham, C. W. "*Data Modeling Essentials"*. San Francisco: Morgan Kaufmann". 2004: 560

Hoover, J. E. " *FingerPrint Anatomy*". Encyclopaedia Britannica, 2023.

Kroenke, D. M., Auer, D., Vandenberg, S.L. & Yoder, R.C. "*Database Concepts 9th Edition."* Pearson. 2019.

Mesnik, B. *"How Does Facial Recognition Work?*". Kintronics, 2017.

Conference Paper

Abdulrazzaq, M. B. & Saeed, J. N. *"A Comparison of Three Classification Algorithms for Handwritten Digit Recognition."* International Conference on Advanced Science and Engineering (ICOASE), IEEE**,** 2019. 58-63.

Al-Amoudi, I., Samad, R., Hasma-Abdullah, N. R., Mustafa, M., & Pebrianti, W. I. "*Automatic Attendance System using Face Recognition with Deep Learning Algorithm*." Springer Singapore, 2022. 573-588.

Ammour, B., Bouden, L. T. & Ramdani, M. "*Face-Iris Multimodal Biometric Identification System***."** Electronics**,** 2020: 1-18.

Bhattacharya, S., Nainala, G. S., Das, P., & Routray, A. "*Smart Attendance Monitoring System (SAMS): A Face Recognition Based Attendance System For Classroom Environment."* IEEE 18th International Conference on Advanced Learning Technologies (ICALT). IEEE, 2018.

Carvalho, G., Mykolyshyn, S., Cabral, B., Bernardino, J. & Pereira, V. *"Comparative Analysis of Data Modeling Design Tools." IEEE Access* 10, 2021. 3351-3365.

Chiu, M. C., Hwang, G. J., Hsia, L. H., & Shyu, F. M. "*Artificial Intelligence-Supported Art Education: A Deep Learning-Based System for Promoting University Students' Artwork Appreciation and Painting Outcomes."* Interactive Learning Environments, 2022: 1-19.

Choudhary, S. K., & Naik, A. K. "*Multimodal Biometric Authentication with Secured Templates — A Review.*" 3rd International Conference on Trends in Electronics and Informatics (ICOEI)**,** 2019, 1062-1069, doi: 10.1109/ICOEI.2019.8862563.

Devikanniga, D., Ramu, A., & Haldorai, A. "*Efficient Diagnosis of Liver Disease using Support Vector Machine Optimized with Crows Search Algorithm".* EAI Endorsed Transactions on Energy Web, 7, no. 29, 2020. e10-e10.

Elor, Y. & Averbuch-Elor, H. *"To SMOTE, Or Not To SMOTE*?" arXiv:2201.08528v2, 2022: 1-22.

Ghimire, P., Piya, S., & Gurung, A. M. "*Comparative Study of Face Mask Recognition using Deep Learning and Machine Learning Classifiers*." International Conference on Innovative Computing, Intelligent Communication and Smart Electrical Systems (ICSES). IEEE, 2021. 1-9.

Ibrahim A., & Ouda, A. "*Hybrid-Based Filtering Approach for User Authentication."* IEEE 30th Canadian Conference on Electrical and Computer Engineering (CCECE) (IEEE), 2017.

Joseph, B. *"The Roles of Information and Communication Technologies (ICTs) and E-Commerce As Agents of Nigeria's Economic Development: Review Of Challenges and Prospects."* Wireless Engineering and Technology 10, no. 03, 2019. 41.

Jordan, M. I. & Mitchell, T.M. "*Machine Learning: Trends, Perspectives and Prospects."* American Association for the Advancement of Science, 2015. 255-260.

Kanade, S., Petrovska-Delacretaz D., & Dorizzi, B. "*Obtaining Cryptographic Keys Using Feature Level Fusion of Iris and Face Biometrics for Secure User Authentication***."** Computer Society Conference on Computer Vision and Pattern Recognition Workshops. IEEE**,** 2011. 138-145.

Koonce, B. "*Squeezenet." Convolutional Neural Networks with Swift for Tensorflow*." Apress, 2021: 73-85.

Krishnmaoorthy, S., Rueda, L., Saad, S., & Elmiligi, H. "*Identification of User Behavioral Biometrics for Authentication using Keystroke Dynamics and Machine Learning*." 2nd International Conference on Biometric Engineering and Applications, 2018, 50-57.

Kumar, T., Bhushan, S. & Jangra. S. "*Ann Trained and WOA Optimized Feature-level Fusion of Iris and Fingerprint.*" Materials Today: Proceedings 51, 2022. 1-11.

Kumar, A., Kaur, A., & Kumar, M. "*Face Detection Techniques: A Review*". Artificial Intelligence Review 52, 2019. 927–948

Latha, L., & Thangasamy, S. "*A Robust Person Authentication System Based on Score Level Fusion of Left and Right Irises and Retinal Features,*" Procedia Computer Science**,** 2, 2010. 111-120,

Lee, H. J., Ullah, I., Wan, W., Gao, Y., & Fang, Z. "*Real-Time Vehicle Make and Model Recognition with the Residual Squeeznet Architecture* ." Sensors **,** no. 5, 2019. 982.

Maghdid, H. S., Asaad, A. T., Ghafoor, K. Z., Sadiq, A. S., Mirjalili, S., & Khan, M. K. "*Diagnosing COVID-19 Pneumonia from X-Ray and CT Images using Deep*

Learning and Transfer Learning Algorithms." Multimodal Image Exploitation and Learning, 2021, 99-110.

Olaleye, T. O. & Vincent, O. R. "*A Predictive Model for Students' Performance and Risk Level Indicators Using Machine Learning*." International Conference in Mathematics, Computer Engineering and Computer Science. Lagos: IEEE, 2020. 1-7.

Olaleye, T., Arogundade, O., Adenusi, C., Misra, S., & Bello, A. "*Evaluation of Image Filtering Parameters for Plant Biometrics Improvement using Machine Learning*." International Conference on Soft Computing and Its Engineering Applications. Springer, 2020. 301-315.

Pal, S., & Jadidi, Z. "*Protocol-Based and Hybrid Access Control for the IoT: Approaches and Research Opportunities."* Sensors 21, no. 20, 2021. 6832.

Rahimi, N., Eassa, F., & Elrefaei, L. *"An Ensemble Machine Learning Technique for Functional Requirement Classification."* Symmetry 12, no. 10, 2020, 1601.

Rahman, M. N., Rahman, A. A., Seyal, A. H., & Kamarudin, N. "*Facial Recognition using Eigenfaces Approach.*" International Conference on Global Economy, Commerce and Service Science . 2014.

Saha, A., Saha, J. & Sen. B. *"An Expert Multi-Modal Person Authentication System Based on Feature Level Fusion of Iris and Retina Recognition."* International Conference on Electrical, Computer and Communication Engineering (ECCE), 2019. 1-5.

Sajjad, M., Khan, S., Hussain, T.K., Muhammad, A.K., Sangaiah, A., Castiglione, C., Esposito, & Baik, S.W. "*CNN-Based Anti-Spoofing Two-Tier Multi-Factor Authentication System*". Pattern Recognition Letters, 126, 2019. 123-131.

Sasidhar, K., Kakulapati, V. L., Ramakrishna, K., & Kailasarao, K. "*Multimodal Biometric Systems Study to Improve Accuracy and Performance*," Arxiv Preprint Arxiv:1011.6220, 2010.

Seelam, V., Penugonda, A. K., Kalyan, B. P., Priya, M. B., & Prakash, M. D. "*Smart Attendance using Deep Learning and Computer Vision.*" Materialstoday: Proceedings. Elsevier, 2021. 4091-4094.

Shaha, M., & Pawar, M. "*Transfer Learning for Image Classification*." 2nd International Conference On Electronics, Communication And Aerospace Technology (ICECA). IEEE, 2018. 656-660.

Shahzad, A. R., & Jalal, A. "*A Smart Surveillance System for Pedestrian Tracking and Counting Using Template Matching"*. International Conference on Robotics and Automation in Industry (ICRAI), IEEE, 2021. 1-6.

Sharma, U., Tomar, P., Ali, S. S., Saxena, N., & Bhadoria, R. S. "*Optimized Authentication System with High Security and Privacy*." Electronics, 2021: 1-23.

Shinde, P. P., & Shah, S. "*A Review of Machine Learning and Deep Learning Applications*." 4th International Conference on Computing Communication Control and Automation (ICCUBEA), IEEE, 2018. 1-6.

Suhaimin, M.S.M., Hijazi, M.H.A., Kheau, C.S. & On. C.K. *"Real-Time Mask Detection and Face Recognition Using Eigenfaces and Local Binary Pattern Histogram for Attendance System."* Bulletin of Electrical Engineering and Informatics 10, no. 2, 2021. 1105-1113.

Suresh , Logeswaran, K., Keerthika, P., Manjula, R., Devi, Sentamilselvan, K., Kamalam, G.K. & Muthukrishnan, H. "Contemporary Survey on Effectiveness of Machine and Deep Learning Techniques for Cyber Security in Machine Learning for Biometrics, Concepts, Algorithms and Applications" ScienceDirect, 2022, 177-200.

Szymkowski, M., Jasiński, P., & Saeed, K. "*Iris- Based Human Identity Recognition with Machine Learning Methods and Discrete Fast Fourier Transform*." Innovations in Systems and Software Engineering, 2021: 309-317.

Wilson, P. "*Importance of Biometrics to Business ."* Dublin Business School, *2001.*

Yang, W., Wang, S., Hu, J., Ibrahim, A., Zheng, G., Macedo, M.J., Johnstone, M.N. & Valli, C. *"A Cancelable Iris-And Steganography-Based User Authentication System for the Internet of Things."* Sensors 19, no. 13, 2019. 2985.

Yang, W., Wang, S., Hu, J., Zheng, G. & Valli, C. *"Security and Accuracy of Fingerprint-Based Biometrics: A review"*. Symmetry 11, no. 2, 2019. 141.

Zaidawi, S. M., Al, K., Prinzler, M. H., Lührs, J., & Maneth, S. "*An Extensive Study of User Identification Via Eye Movements Across Multiple Datasets*." Signal Processing: Image Communication, 2022, 116804.

Journal

AbdElaziz, A. A. *"A Survey of Smartphone-Based Face Recognition System for Security Purposes."* **Kafrelsheikh Journal of Information Sciences** 2, no. 1, 2021. 1-5.

Adeola, O. S. *"A Fuzzy-Based Multi-Modal Architectural Model for Electioneering in Nigeria."* **International Journal of Technology Diffusion (IJTD)** 11, no. 4, 2020. 1-26.

Agrawal, R., & Goyal, R. "*Developing Bug Severity Prediction Models Using Word2vec.*" **International Journal of Cognitive Computing in Engineering**, 2021: 104-115.

Ahmad, I., S. Khan, M., Naeem, M., Hayat, U. R., Azmi, S., Ahmed, & Irfan, M. "*Molecular Identification of Ten Palm Species Using DNA Fingerprinting.*" **International Journal of Pure & Applied Bioscience (IJPAB)** 7, no. 1, 2019: 46-51.

Akinrinmade, A. A., Adetiba, E., Badejo, J. A., & Atayero, A. A. *"Creation of a Nigerian Voice Corpus for Indigenous Speaker Recognition."* **Journal of Physics: Conference Series**, vol. 1378, no. 3, 2019. 032011.

Arjun, B. C., & Prakash, H. N. "*Multimodal Biometric Recognition: Fusion of Modified Adaptive Bilinear Interpolation Data Samples of Face and Signature using Local Binary Pattern Features*." **International Journal of Engineering and Advanced Technology** 9, no. 3, 2020. 3111-3120.

Babatunde, A. N., Oke, A. A., Babatunde, R. S., Ibitoye, O., & Jimoh, E. R. *"Mobile Based Student Attendance System using Geo-fencing with Timing and Face Recognition."* **Journal Advances in Mathematical & Computational Sciences** 10, no. 1, 2022.

Bansal, M., Kumar, M., Sachdeva, M., & Mittal, A. *"Transfer Learning for Image Classification using VGG19: Caltech-101 Image Data Set."* **Journal of Ambient Intelligence and Humanized Computing**, 2021. 1-12.

Chawla, N. V., Kevin W. B., Lawrence, O. H., & Kegelmeyer, W.P., "*SMOTE: Synthetic Minority Over-Sampling Technique.*" **Journal of Artificial Intelligence Research,** 16, 2002. 321-357.

Chen, J., Hongwei, H., Anthony, G. C., Dongming, Z., & Mingliang, Z. "*Machine Learning-Based Classification of Rock Discontinuity Trace: SMOTE Oversampling Integrated with GBT Ensemble Learning*." **International Journal of Mining Science and Technology** 32, No. 2. 2022. 309-322.

Desmond, A. *"Multimodal Biometric Authentication for a Computer-Based Test (CBT) Application*." **IRJCS: International Research Journal of Computer Science**, 7, no. 7, 2020. 179-196."

Dhabre, S. R. & Pol, R. S. *"Automated Face Recognition Based Attendance System Using lbp Face Recognizer."* **International Journal** 5, no. 4, 2020.

El-mehdi-Cherrat, R. A., & Bouzahir, H. "*A Multimodal Biometric Identification System Based on Cascade Advanced of Fingerprint, Finger Vein and Face Images*." **Indonesian Journal of Electrical Engineering and Computer Science** 18, no. 1, 2020. 1562-1570.

Esomonu, N. P. M., Esomonu, M. N., & Eleje, L. I. "Assessment Big Data in Nigeria: Identification, Generation and Processing in the Opinion of the Experts." **International Journal of Evaluation and Research in Education** 9, no. 2, 2020. 345-351.

Farayola, M., & Aman, D. "*A Proposed Framework: Face Recognition with Deep Learning."* **International Journal of Science and Technology Research** 9, no. 7 2020.

Folorunso, C. O., O. S. Asaolu, & O. P. Popoola. *"A Review of Voice-Base Person Identification: State-Of-The-Art."* **Covenant Journal of Engineering Technology** 3, no. 1, 2019.

Hamd, M. H. & Mohammed, M. Y. "*Multimodal Biometric System Based Face-Iris Feature Level Fusion."* **International Journal of Modern Education Computer Science** 11, no. 5, 2019. 1-9.

Hassen, O., N. Abu, & Z. Abidin. *"Human Identification System: A Review."* **The International Journal of Computing and Business Research (IJCBR)** 9, 2019. 1-26.

Hoo, S. C. & Ibrahim, H. *"Biometric-Based Attendance Tracking System for Education Sectors: A Literature Survey on Hardware Requirements."* **Journal of Sensors** 2019. 1-25.

Ibrahim, M., Shuaibu, I., Ibrahim, M. Y., & Abdulkadir, A. *"Design of An Automated Attendance System Using Face Recognition Algorithm."* **International Journal of Engineering and Technology** 6, no. 10, 2019. 1114-1119.

Isinkaye, F. O., Soyemi, J. & Arowosegbe, O. I. *"An Android-Based Face Recognition System for Class Attendance and Malpractice Control."* **International Journal of Computer Science and Information Security (IJCSIS)** 180, no. 1, 2020. 78-83.

Iwasokun, G. B., Opatoye, K. I., & Orunmuyi, B. O. *"Multi-Modal Biometrics Fusion Based on Component Analysis and Stationery Wavelet Transform."* **International Journal of Information Security Science** 9, no. 2, 2020. 114- 125.

Jothi, R.A., & Palanisamy, V. "A Robust and Efficient Fingerprint Minutiae Extraction in Post-Processing Algorithm". **International Journal of Biometrics**, 2022. 83-97.

Jumani, A. K., Mahar, M. H., Khoso, F. H., & Memon, M. *"Online Text Categorization System Using Support Vector Machine."* **Sindh University Research Journal-SURJ (Science Series)** 50, no. 1, 2018. 85-90.

Krishna, S. T., & Kalluri, H. K. *"Deep Learning And Transfer Learning Approaches for Image Classification."* **International Journal of Recent Technology and Engineering (IJRTE)** 7, no. 5, 2019. 427-432.

Kumar, T., Bhushan, S., & Jangra, S. *"An Improved Biometric Fusion System of Fingerprint and Face Using Whale Optimization."* **International Journal of Advanced Computer Science and Applications** 12, no. 1, 2021.

Lavadkar, M. A., Thorat, P. K., Kasliwal, A. R., Gadekar, J. S., & Deshmukh, D. P. *"Fingerprint Biometric Based Online Cashless Payment System."* **IOSR Journal of Computer Engineering (IOSR-JCE),** *e-ISSN*: 2019. 2278-0661.

Medjahed, C., Rahmoun, A., Charrier, C., & Mezzoudj, F. "*A Deep Learning-Based Multimodal Biometric System Using Score Fusion*." **IAES International Journal of Artificial Intelligence** 11, no. 1, 2022. 65.

Misganaw, A., & Seyoum, F. *"Assessment of Manual Perioperative Anesthesia Record-Keeping Practice: A Multicenter; Descriptive Cross-Sectional Study, Ethiopia."* **International Journal of Surgery Open**, 46, 2022.100526.

Mohammed, B. O. & Hashim, Z. M. *"Individuality Representation Using Multimodal Biometrics with Aspect United Moment Invariant for Identical Twins."* **Journal of Theoretical and Applied Information Technology** 98, no. 12, 2020. 2148-2157.

Mustafa, A. S., Abdulelah, A. J., & Ahmed, A. K. "*Multimodal Biometric System Iris and Fingerprint Recognition Based on Fusion Technique*." **International Journal of Advanced Science and Technology** 29, no. 3, 2020. 7423-7432.

Mustapha, M.F., Mohamad, N.M. & AB, S.H. *"A Survey on Video Face Recognition Using Deep Learning."* **Journal of Quality Measurement and Analysis JQMA** 18, no. 1, 2022. 49-62.

Nasr, M., & El-Fangary, L. *"Applications of IoT In Smart Grids Using Demand Respond for Minimizing On-Peak Load. "* **International Journal of Computer Science and Information Security (IJCSIS)** 19, no. 8, 2021.

Ogbodo, I. A. *"Exploring Access to EHR by Emergency Patients Using Multimodal Biometrics."* **International Journal of Latest Technology in Engineering Management & Applied Science** 9. 2020.

Ojo, O., Kareem, M. K., Odunuyi, S., & Ugwunna, C. *"An Internet-of-Things Based Real-Time Monitoring System for Smart Classroom."* **Journal of the Nigerian Society of Physical Sciences** 2022. 297-309.

Ojo, O., Yekini, N. A., Aigbokhan, E. E., & Onadokun, I. *"Deterring Malpractice in a Networked Computer Based Examination Using Biometric Control Attendance Register."* **International Journal of Advanced Networking and Applications** 10, no. 6, 2019. 4061-4064.

Okereafor, K., Osuagwu, O., Ayegba, S. F., & Adelaiye, O. *"Randomized Multi-Biometric Liveness Detection: Prospects and Applications for Secure Authentication."* **International Journal in IT & Engineering (IJITE)** 8, no. 10, 2020.

Oladele, T. O., Adeniyi, K. & Aro, T. O. *"Framework for User Authentication at a Distance for Mobile Phones Using Contactless Hand-Based Multimodal Biometric System."* **Journal of Computer Science and Control Systems** 12, no. 1, 2019. 24-27.

Olatunji, K. A., Oguntimilehin, A., Aweh, O. M., & Adeyemo, O. *"Feature Level Fusion Algorithm for Iris and Face."* **International Journal** 9, no. 12, 2021.

Omotosho, L., Ogundoyin, I., Adebayo, O., & Oyeniyi, J. *"An Enhanced Multimodal Biometric System Based on Convolutional Neural Network."* **Journal of Engineering Studies and Research** 27, no. 2, 2021. 73-81.

Onyema, E. M., Shukla, P. K., Dalal, S., Mathur, M. N., Zakariah, M., & Tiwari, B. *"Enhancement of Patient Facial Recognition Through Deep Learning Algorithm: ConvNet."* **Journal of Healthcare Engineering,** 2021.

Oyeniyi, J. & Oyeniran, O. *"Bimodal Biometric Recognition System Based on Score Level Fusion."* **International Journal of Engineering and Artificial Intelligence,** 1, no. 4, 2020. 71-76.

Parihar. R. " *Palm Vein Recognition System for Human Authentication: A Review"* **International Journal for Research in Applied Science and Engineering Technology**, 2019.

Patra, G.R., Mohapatra, S.K., & Mohanty, M.N., "*Applications of Deep Learning Algorithms in Biomedical Signal Processing - Pros and Cons*". **International Journal of Biometrics,** 2022. 98-124.

Piñeyro, L., Pardo, A., & Viera, M. *"Structure Verification of Deep Neural Networks at Compilation Time*." **Journal of Computer Languages**, 67, 2021. 101074.

Popoola, O. P., Folorunso, C. O., Asaolu, O. S., Joshua, J. J., & Oyeyemi, M. D. *"Person Identification System from Speech and Laughter using Machine Learning Algorithms."* **Journal of Engineering Research,** 25, no. 2, 2020. 173-190.

Prabakaran, E., & Pillay, K., "*Nanomaterials for Latent Fingerprint Detection: A Review*." **Journal of Materials Research and Technology,** 12, 2021: 1856-1885.

Reddy, N. S., Sumanth, M.V., & Babu, S. S. "*A Counterpart Approach To Attendance And Feedback System Using Machine Learning Techniques."* **Jetir- International Journal of Emerging Technologies and Innovative Research,** 2018. 190-193.

Sanjekar, P. S. & Patil, J. B. "*Multimodal Biometrics with Serial, Parallel and Hierarchical Mode at Decision Level Fusion*." **Indonesian Journal of Electrical Engineering and Computer Science** 16, no. 3, 2019. 1303-1310.

Sarangi, P. P., Nayak, D. R., Panda, M., & Majhi, B. "*A Feature-Level Fusion Based Improved Multimodal Biometric Recognition System Using Ear and Profile Face*." **Journal of Ambient Intelligence and Humanized Computing** 2022.1-32.

Sujana, S. & Reddy, V. S. K. "*Comparison of Levels and Fusion Approaches for Multimodal Biometrics*." **Indonesian Journal of Electrical Engineering and Computer Science** 23, no. 2, 2021. 791-801.

Sun, L., Zhong, Z., Qu, Z., & Xiong, N. "*Perae: An Effective Personalized Autoencoder for ECG-Based Biometric in Augmented Reality System*." **IEEE Journal of Biomedical and Health Informatics 26**, no. 6, 2022. 2435-2446.

Suru, H. U., & Murano, P. *"Security and User Interface Usability of Graphical Authentication Systems – A Review."* **International Journal of Engineering Trends and Technology (IJERT)** 67, 2019. 17-36.

Tyagi, S., Chawla, B., Jain, R., & Srivastava, S. *"Multimodal Biometric System Using Deep Learning Based on Face and Finger Vein Fusion."* **Journal of Intelligent & Fuzzy Systems** 42, no. 2, 2022. 943-955.

Ukpai, G. *"Integrating Learning Management System for Teaching and Learning in Nigeria Tertiary Institutions: A Need for 21st Century Education."* **Journal of Resourcefulness and Distinction** 18, no. 1, 2021.

Vijay, M. & Indumathi, G. *"Deep Belief Network-Based Hybrid Model for Multimodal Biometric System for Futuristic Security Applications."* **Journal of Information Security and Applications**, *58*, 2021. 102707.

Vijayakumar, T. "*Synthesis of Palm Print in Feature Fusion Techniques for Multimodal Biometric Recognition System Online Signature*." **Journal of Innovative Image Processing (JIIP)** 3, no. 02, 2021. 131-143.

Wirdiani, N.A, Lattifia, T., Supadma,I.K ., Mahar, B.K., Taradhita, D.V., & Fahmi. A. "Real-time Face Recognition with Eigenface Method." ***International Journal Image, Graphic Signal Process*** 11, no. 11, 2019. 1-9.

Yadav, A. K. "*Deep Learning Approach for Multimodal Biometric Recognition System Based on Fusion of Iris, Fingerprint and Hand Written Signature Traits*." **Turkish Journal of Computer and Mathematics Education** 12, no. 11, 2021. 1627-1640.

Yu, J., Hu, A., Li, G., & Peng, L. "*A Robust RF Fingerprinting Approach Using Multisampling Convolutional Neural Network*." **IEEE Internet of Things Journal 6**, no. 4, 2019. 6786-6799.

Yusuf, N., Marafa, K. A., Shehu, K. L., Mamman, H., & Maidawa, M. *"A Survey of Biometric Approaches of Authentication."* **International Journal of Advanced Computer Research** 10, no. 47, 2020. 96-104.

Electronic Source (Internet)

Aishwarya S *"A Comprehensive Guide to Ensemble Learning (with Python Code)"* Analytic Vidhya.. 2018. https://www.analyticsvidhya.com/blog/2018/06/comprehensive-guide-for-ensemble-models

BCAdmin, *"A Brief History of Biometrics"*. BioConnect.. 2021. https://bioconnect.com/2021/12/08/a-brief-history-of-biometrics

Gaytán-Campos, I., Morales-Castro, W., Priego-Sánchez, B., Fitz-Rodríguez, E., & Guzmán-Cabrera, R. "*Automatic Classification of Images with Skin Cancer Using Artificial Intelligence*." Computación y Sistemas 26, no. 1 2022. https://doi.org/10.13053/cys-26-1-4176

Huifan, C. "The *Difference Between Iris Recognition and Fingerprint Recognition."* Chongquing Huifan Technology. 2022. https://www.hfteco.com/news/Iris-recognition-and-fingerprint-recognition.html

NEC. "*What is Iris Recognition and How Does it Work*?". 2022. https://www.nec.co.nz/market-leadership/publications-media/what-is-iris-recognition-and-how-does-it-work/

Satpathy, S. "*Overcoming Class Imbalance Using SMOTE Technique"*. Analytics Vidhya. 2020. http://www.analyticsvidhya.com/blog/ 2020/10/ overcoming-class-imbalance-using-smote-techniques**.**

Truce, Z. "*Technologies Used in the Attendance Management*." Technology Insights. 2020. https://www.zimyo.com/insights/technologies-used-in-the-attendance-management/

Appendix I: Python Code

```
from tkinter import filedialog
import tkinter.messagebox as msgBox
import os
from sklearn.ensemble import RandomForestClassifier
import numpy as np
import tensorflow
import pandas as pd
import csv
import numpy
import joblib
import time
from sklearn import metrics
from tkinter.ttk import *
from tkinter.filedialog import askopenfile
from collections import Counter
import imblearn
from imblearn.over_sampling import SMOTE
from tkinter import *
from tkinter import filedialog
from sklearn.metrics import classification_report
from sklearn.metrics import confusion_matrix
from sklearn.metrics import accuracy_score
from sklearn.metrics import classification_report
```

```
from sklearn.metrics import roc_auc_score
from sklearn.metrics import log_loss
from sklearn.svm import SVC
from sklearn.tree import DecisionTreeClassifier
from sklearn import metrics

from matplotlib import pyplot as plt
import seaborn as sns

from sklearn import model_selection
from sklearn.ensemble import VotingClassifier

sns.set_style('darkgrid')

import tkinter as tk
from tkinter import *
from tkinter import filedialog
from tkinter.filedialog import askopenfile
from PIL import Image, ImageTk

my_w = tk.Tk()
my_w.geometry("500x600")
my_w['bg']='#2a636e'
my_w.title('STUDENT ATTENDANCE AUTHENTICATION SYSTEM')
my_font1=('times', 18, 'bold')
```

```
def browseFiles():
    filename1 = filedialog.askopenfilename(initialdir = "/",
                                           title = "Select a File",
                                           filetypes = (("Text files",
                                                         "*.csv*"),
                                                        ("all files",
                                                         "*.*")))

    # Change label contents
    label_file_explorer.configure(text="File Opened: "+filename1)


button_explore = Button(my_w,
                        text = "Authentication",
                        command = browseFiles)
button_explore.place(x=250,y=350)


b1 = tk.Button(my_w, text='UPLOAD INPUTS',
   width=20,command = lambda:upload_file())
b1.place(x=50,y=350)



def upload_file():
    f_types = [('JPG Files','*.jpg')]   # type of files to select
```

```
    filename1 = tk.filedialog.askopenfilename(multiple=True,filetypes=f_types)
    col=1 # start from column 1
    row=5 # start from row 3
    for f in filename1:
        img=Image.open(f) # read the image file
        img=img.resize((100,100)) # new width & height
        img=ImageTk.PhotoImage(img)
        e1 =tk.Label(my_w)
        e1.grid(row=row,column=col)
        e1.image = img
        e1['image']=img # garbage collection
        if(col==4): # start new line after third column
            row=row+1# start wtih next row
            col=1    # start with first column
        else:       # within the same row
            col=col+1 # increase to next column


#DEFINING TRAINING SET
file = 'C:\\Users\\User\\Desktop\\LSQUEEZENET.csv'

raw_data1 = open(file, 'rt')
reader1 = csv.reader(raw_data1, delimiter=',', quoting=csv.QUOTE_NONE)
x1 = list(reader1)
x_train = numpy.array(x1).astype('float')
```

```
##### IMPLEMENTING THE VOTE ENSEMBLE
estimators = []
model1=DecisionTreeClassifier();
estimators.append(("DecisionTree", model1))

model2=SVC(kernel='linear');
estimators.append(("SVM", model2))

model3=RandomForestClassifier(n_estimators=100);
estimators.append(("RandomForest", model3))

ensemble=VotingClassifier(estimators)

x_RealTrain=x_train[:,0:1002]
y=x_train[:,1003]

#SMOTE

#counter=Counter(y)
#print('Before',counter)

#smt=SMOTE()
#x_sm,y_sm=smt.fit_resample(x_RealTrain,y)

#counter=Counter(y_sm)
```

```
#print('After',counter)


#clf=DecisionTreeClassifier()

#clf=RandomForestClassifier(n_estimators=100)

#svclassifier = SVC(kernel='linear')


#FIT DATA ONTO THE MODEL


ensemble.fit(x_RealTrain,y)

#clf.fit(x_RealTrain,y)

#clf.fit(x_RealTrain,y)

#svclassifier.fit(x_RealTrain, y)


#.....

#DEFINING TEST_SET


file_Test = 'C:\\Users\\User\\Desktop\\LSQUEEZENETTEST.csv'

raw_data2 = open(file_Test, 'rt')

reader2 = csv.reader(raw_data2, delimiter=',', quoting=csv.QUOTE_NONE)

x2 = list(reader2)

x_test = numpy.array(x2).astype('float')

x_RealTest=x_test[:,0:1002]


y_pred = ensemble.predict(x_RealTest)

#y_pred = svclassifier.predict(x_RealTest)
```

```
prediction=pd.DataFrame(y_pred,                    columns=['Attendance
Identity']).to_csv('C:\\Users\\User\\Desktop\\INCEPTPREDICT.csv')


df = pd.read_csv("C:\\Users\\User\\Desktop\\INCEPTPREDICT.csv")
df['Attendance Identity'] = df['Attendance Identity'].replace({1 : 'Taiwo'})
df['Attendance Identity'] = df['Attendance Identity'].replace({2 : 'Kenny'})
df['Attendance Identity'] = df['Attendance Identity'].replace({3 : 'Dele'})
df['Attendance Identity'] = df['Attendance Identity'].replace({4 : 'Goke'})
df['Attendance Identity'] = df['Attendance Identity'].replace({5 : 'Nike'})


df.to_csv("C:\\Users\\User\\Desktop\\INCEPTIONFINAL.csv", index=False)


#.....


def openPre():
    predict=filedialog.askopenfilename()
    #my_label.config(text=predict)
      #Open the file
    os.system('"%s"' %predict)


button_prediction = Button(my_w,
                text = "Check Result",
                command = openPre)
button_prediction.place(x=400,y=350)
```

```
my_w.mainloop()  # Keep the window open

metrics.plot_confusion_matrix(clf,x_RealTrain,y,display_labels=['Taiwo', 'Kenny', 'Dele',
'Goke', 'Nike'])

print("Accuracy:", metrics.accuracy_score(x_RealTest,y_pred))
```

Bio-Data

Personal Data

Surname:	OLOMOLA
Other Name:	Babatunde Taiwo
Date of Birth:	July 12, 1973
Gender:	Male
Local Government Area	Ilesha West Local Government
State of Origin:	Osun State
Nationality:	Nigerian
Marital Status	Married
Religion	Christianity
Next of Kin:	Mrs. Olajumoke Olomola (08067080451)

Contact:

Residential Address:	13 Kumuyi Street, Mushin, Lagos
Portal Address:	Same as above
E-mail:	demilade2003@yahoo.com
Mobile No:	08035737944

Educational Background:

Institution attended with Dates and Qualifications:

- M.Sc. Computer Science, Lead City University, Ibadan. (in view)
- B.Sc. (Edu.) Computer Education, University of Nigeria, Nsukka. (2008)
- HND. Computer Science, The Polytechnic Ibadan. (1997)
- OND. Computer Science, The Polytechnic Ibadan. (1994)
- WAEC. Methodist Secondary Grammar School, Elekuro, Ibadan (1989)

Work Experience with Dates:

- Yaba College of Technology, Yaba, Lagos State March 2000 – Till date

  Designation: Chief Technologist (Academics)

Courses Taught:

- Webpage Design
- Desktop Publishing
- Management Information System (MIS)
- DataBase Management System (DBMS)
- Information and Communication Technology (ICT)

Other Activities

- Vice Chairman, Academic Staff Union of Polytechnics (Yaba College of Technology Chapter) 2018
- Chairman, Academic Staff Union of Polytechnics (Yaba College of Technology Chapter) 2019
- Coordinator of Part-Time Programme, OTM Department, Yaba College of Technology. 2023

Membership of Academic Professional Bodies:

- Member, Computer Society of Nigeria (NCS)
- Fellow, National Institute of Office Administrators and Information Managers (NIOAIM)

Published Journal Articles:

- Artificial Intelligence (AI) in an Organization: Its Opportunities and Threats (Authors: **Olomola, Babatunde T**; Akoyokun, Temitope S.; & Ige, Oluwaseye, J., 2019)
- Cloud Computing: Assessing Its Relevance to Office Managers' Job Performance and Organizational Overall Productivity (Authors: **Olomola, Babatunde T**; & Ige, Oluwaseye, J., 2019)

- Social Media: Its Impact on Lecture Delivery and Students' Academic Performance in Nigerian Tertiary Institutions (Authors: **Olomola, Babatunde T**; Akhademe, Abiose E.; & Ige, Oluwaseye, J., 2018)

Books Published:

- Understanding the Internet and Web Design using HTML/CSS (Authors: **Olomola, B. T.**; Ademoroti, A.; Irokanulo, C.; and Ige, O. J.)

- Understanding the Concept of Desktop Publishing (Authors: **Olomola, B. T.**; Ademoroti, A.; and Ige, O. J.)

Referees:

- Dr. Mrs. P. N. Ogadi,
  Dean School Of Management and Business Studies,
  Yaba College Of Technology, Lagos.

- Dr. Mrs. A.E. Akhademe
  Head of Department, Office Technology and Management,
  Yaba College Of Technology, Lagos.

Signature _____ **Date** _____

The University Compliance Certification

This is to certify that this thesis by **Babatunde Taiwo OLOMOLA** with Matric Number LCU/PG/002445 in the Department of Computer Science, Faculty of Natural and Applied Sciences, Lead City University, Ibadan, Oyo State is in FULL compliance with the approved University Format and Style.

Signature **Date**

This portion is not part of the work pls

S. Dargan and M. Kumar, "A Comprehensive Survey on the Biometric Recognition Systems based on Physiological and Behavioral Modalities," Expert Systems with Applications, p. 113114, 2019.

S. Dargan and M. Kumar, "A Comprehensive Survey on the Biometric Recognition Systems based on Physiological and Behavioral Modalities," Expert Systems with Applications, p. 113114, 2019.

---

[1] Bolle, R. M., J. H. Connell, S. Ratha, N. K., and A. W. "Guide to biometrics." *Springer Science & Business Media*, 2013.

[2] Bhattacharya, Shubhobrata, Gowtham Sandeep Nainala, Prosenjit Das, and Aurobinda Routray. "Smart Attendance Monitoring System (SAMS): A Face Recognition Based Attendance System for Classroom Environment." *IEEE 18th International Conference on Advanced Learning Technologies (ICALT)*. IEEE, 2018.

[3] Krishnmaoorthy, Sowndarya, L. Rueda, S. Saad, and H. Elmiligi. "Identification of user behavioral biometrics for authentication using keystroke dynamics and machine learning ." *2018 2nd International Conference on Biometric Engineering and Applications* . 2018. 50-57.

[4] Maghdid, H. S., A. T. Asaad, K. Z. Ghafoor, A. S. Sadiq, S. Mirjalili, and M. K. Khan. "Diagnosing COVID-19 pneumonia from X-ray and CT images using deep learning and transfer learning algorithms." *Multimodal image exploitation and learning*, 2021: 99-110.

[5] Rahman, Mohd Noah, Armanadurni Abd Rahman, Afzaal H. Seyal, and Nursuziana Kamarudin. "Facial recognition using eigenfaces approach." *International Conference on Global Economy, Commerce and Service Science* . 2014

[6] Sajjada, Muhammad, et al. "CNN-based anti-spoofing two-tier multi-factor authentication system."

[7] Satpathy, Swastik. *Home page.* October 6, 2020. http://www.analyticsvidhya.com/blog/2020/10/overcoming-class-imbalance-using-smote-techniques.

[8] Seelam, Vivek, Akhil kumar Penugonda, B. Pavan Kalyan, M. Bindu Priya, and M.Durga Prakash. "Smart attendance using deep learning and computer vision." *Materialstoday: Proceedings.* Elsevier, 2021. 4091-4094.

[9] Shinde, Pramila P., and Seema Shah. "A review of machine learning and deep learning applications." *Fourth international conference on computing communication control and automation (ICCUBEA.* IEEE, 2018. 1-6.

[10] https://bioconnect.com/2021/12/08/a-brief-history-of-biometrics

[11] Agrawal, Rashmi, and Rinkaj Goyal. "Developing bug severity prediction models using word2vec." *International Journal of Cognitive Computing in Engineering*, 2021: 104-115.

privilege

[12] Al-Amoudi, Ibrahim, Rosdiyana Samad, Nor Rul Hasma Abdullah, Mahfuzah Mustafa, and Dwi Pebrianti. "Automatic Attendance System Using Face Recognition with Deep Learning Algorithm." *Proceedings of the 12th National Technical Seminar on Unmanned System Technology* . Springer, 2021. 573-588.

[13] Al-Amoudi, Ibrahim, Rosdiyana Samad, Nor Rul Hasma Abdullah, Mahfuzah Mustafa, and wi Pebrianti. *Automatic Attendance System Using Face Recognition with Deep Learning Algorithm.* Springer, 2021.

[14] Chen, Shaxun, Amit Pande, and Prasant Mohapatra. "Sensor-assited facial recognition: an enhanced biometric authetication system for smartphones." *12th annual international conference on Mobile systems, applications, and services.* 2014. 109-122.

[15] Gaytán-Campos, ISrael, Wendy Morales-Castro, Belem Priego-Sánchez, Efren Fitz-Rodríguez, and Rafael Guzmán-Cabrera. "Automatic Classification of Images with Skin Cancer Using Artificial Intelligence." *Computación y Sistemas* 26, no. 1 (2022).

[16] Ghimire, Prashant, Sweekar Piya, and Anish Man Gurung. "Comparative study of Face Mask Recognition using Deep Learning and Machine learning classifiers." *2021 International Conference on Innovative Computing, Intelligent Communication and Smart Electrical Systems (ICSES).* IEEE, 2021. 1-9.

[17] S. K. Choudhary and A. K. Naik, "Multimodal Biometric Authentication with Secured Templates — A Review," *2019 3rd International Conference on Trends in Electronics and Informatics (ICOEI)*, 2019, pp. 1062-1069, doi: 10.1109/ICOEI.2019.8862563.

[18] Kanade, Sanjay, Dijana Petrovska-Delacretaz, and Bernadette Dorizzi. "Obtaining cryptographic keys using feature level fusion of iris and face biometrics for secure user authentication ." *2010 IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops.* IEEE, 2010. 138-145.

[19] R. Verma and A. Goel, "Wavelet application in fingerprint recognition," International Journal of Soft Computing and Engineering (IJSCE) ISSN, vol. 1, pp. 2231-2307, 2011.

[20] Maghdid, H. S., A. T. Asaad, K. Z. Ghafoor, A. S. Sadiq, S. Mirjalili, and M. K. Khan. "Diagnosing COVID-19 pneumonia from X-ray and CT images using deep learning and transfer learning algorithms." *Multimodal image exploitation and learning*, 2021: 99-110.

[21] Kumar, A., Kaur, A. & Kumar, M. Face detection techniques: a review. *Artif Intell Rev* **52,** 927–948 (2019). https://doi.org/10.1007/s10462-018-9650-2

[22] NEC. "What is Iris Recognition and how does it work?" June 15, 2022. https://www.nec.co.nz/market-leadership/publications-media/what-is-iris-recognition-and-how-does-it-work/ (accessed September 23, 2022).

[23] Oren, M., C. Papageorgious, P. Sinha, E. Osuna, and T. Poggio. "Pedestrian Detection Using Wavelet Templates." *Proc. Computer Vision and Pattern Recognition*, 193-199: 2017.

[24] Sajjada, Muhammad, et al. "CNN-based anti-spoofing two-tier multi-factor authentication system." *Pattern Recognition Letters*, 2019: 123-131.

[25] K. Sasidhar, V. L. Kakulapati, K. Ramakrishna, and K. KailasaRao, "Multimodal biometric systemsstudy to improve accuracy and performance," arXiv preprint arXiv:1011.6220, 2010.

[26] L. Latha and S. Thangasamy, "A robust person authentication system based on score level fusion of left and right irises and retinal features," Procedia Computer Science, vol. 2, pp. 111-120, 2010.

[27] Graeme, C. S & Graham C. W. (2014). Data modeling Essentials. San Francisco: Morgan Kaufmann.

[28] David M. K and David J. A (2016): A Review of Database concepts Prentice Hall New Jersey.

[29] Brian K.W and Stacey C. S (2019) - Using Information Technology – A practical introduction to Computers and Communications 12th Edition.

[30] Brian K.W and Stacey C. S (2019) - Using Information Technology – A practical introduction to Computers and Communications 12th Edition.

[31] Maes S.H., Navrátil J., Chaudhari U.V. (2017): Conversational Speech Biometrics, Chapter in "E-Commerce Agents Marketplace Solutions, Security Issues, and Supply and Demand," J. Liu and Y. Ye (Eds.): Springer Verlag, page 166 – 179.

[32] Keith A. Rhodes, (2016): Information Security: Challenges in using Biometrics.

[33] Thakkar, Danny. "Iris Recognition Scanners vs. Fingerprint Scanners: Compare and Contrast." *Bayometric.* 2019. https://www.bayometric.com/iris-recognition-scanners-vs-fingerprint-scanners/ (accessed September 15, 2022).

[34] Truce, Zimyo. "Technologies used in the Attendance management." *Technology Insights.* August 7, 2020. https://www.zimyo.com/insights/technologies-used-in-the-attendance-management/ (accessed October 4, 2022).

[35] Szymkowski, Maciej, Piotr Jasiński, and Khalid Saeed. "Iris-based human identity recognition with machine learning methods and discrete fast Fourier transform." *Innovations in Systems and Software Engineering*, 2021: 309-317.

[36] Zaidawi, S. M. K. Al, M. H. Prinzler, J. Lührs, and S. Maneth. "An extensive study of user identification via eye movements across multiple datasets." *Signal Processing: Image Communication*, 2022: 116804.

[37] Yang, Wencheng, et al. "A Cancelable Iris- and Steganography-Based User Authentication System for the Internet of Things." *MDPI*, 2019.

[38] Shinde, Pramila P., and Seema Shah. "A review of machine learning and deep learning applications." *Fourth international conference on computing communication control and automation (ICCUBEA.* IEEE, 2018. 1-6.

[39] Sharma, Uttam, Pradeep Tomar, Syed Sadaf Ali, Neetesh Saxena, and Robin Singh Bhadoria. "Optimized Authentication System with High Security and Privacy." *Electronics*, 2021: 1-23.

[40] Mihajlov, Martin, Borka Jerman-Blazic, and Marko LLievski. "ImagePass-Designing graphical authentication for security." *2011 7th International Conference on Next Generation Web Services Practices* . IEEE, 2011. 262-267.

[41] Krishnmaoorthy, Sowndarya, L. Rueda, S. Saad, and H. Elmiligi. "Identification of user behavioral biometrics for authentication using keystroke dynamics and machine

learning ." *2018 2nd International Conference on Biometric Engineering and Applications* . 2018. 50-57.

[42] Olaleye, Taiwo, Oluwasefunmi Arogundade, Cecelia Adenusi, Sanjay Misra, and Abosede Bello. "Evaluation of image filtering parameters for plant biometrics improvement using machine learning." *International Conference on Soft Computing and its Engineering Applications.* Springer, 2020. 301-315.

[43] Rahman, Mohd Noah, Armanadurni Abd Rahman, Afzaal H. Seyal, and Nursuziana Kamarudin. "Facial recognition using eigenfaces approach." *International Conference on Global Economy, Commerce and Service Science* . 2014.

[44] Seelam, Vivek, Akhil kumar Penugonda, B. Pavan Kalyan, M. Bindu Priya, and M.Durga Prakash. "Smart attendance using deep learning and computer vision." *Materialstoday: Proceedings.* Elsevier, 2021. 4091-4094.

[45] Truce, Zimyo. "Technologies used in the Attendance management." *Technology Insights.* August 7, 2020. https://www.zimyo.com/insights/technologies-used-in-the-attendance-management/ (accessed October 4, 2022).

[46] Wilson, Patrick. "Importance of Biometrics to Business ." 2001.

[47] Reddy, N.Sudhakar, M.V. Sumanth, and S.Suresh Babu. "A Counterpart Approach to Attendance and Feedback System using Machine Learning Techniques." *JETIR-International Journal of Emerging Technologies and Innovative Research*, 2018: 190-193.

[48] Olaleye, T. O., and O. R. Vincent. "A Predictive Model for Students' Performance and Risk Level Indicators Using Machine Learning." *2020 International Conference in Mathematics, Computer Engineering and Computer Science.* Lagos: IEEE, 2020. 1-7.

[49] Krishnmaoorthy, Sowndarya, L. Rueda, S. Saad, and H. Elmiligi. "Identification of user behavioral biometrics for authentication using keystroke dynamics and machine learning ." *2018 2nd International Conference on Biometric Engineering and Applications* . 2018. 50-57.

[50] Jordan, M. I., and T. M. Mitchell. "Machine learning: Trends, perspectives, and prospects." *American Association for the Advancement of Science*, 2019: 255-261.

[51] Gandhi, Vipul. *A comprehensive guide to Ensemble learning* . 2019. https://www.kaggle.com/vipulgandhi/a-comprehensive-guide-to-ensemble-learning.

[52] Zhang, Cha, and Yunqian Ma. "Ensemble machine learning: methods and applications." *Springer Science & Business Media*, 2012.

[53] S. K. Choudhary and A. K. Naik, "Multimodal Biometric Authentication with Secured Templates — A Review," *2019 3rd International Conference on Trends in Electronics and Informatics (ICOEI)*, 2019, pp. 1062-1069, doi: 10.1109/ICOEI.2019.8862563.

[54] Platt, John C. "Sequential Minimal Optimization: A Fast Algorithm for Training Support Vector Machines." *Microsoft Research*, 2018: 1-22.

[55] Satpathy, Swastik. *Home page.* October 6, 2020. http://www.analyticsvidhya.com/blog/2020/10/overcoming-class-imbalance-using-smote-techniques.

[56] Elor, Yotam, and Hadar Averbuch-Elor. "To SMOTE, or not to SMOTE?" *arXiv:2201.08528v2*, 2022: 1-22.

[57] Bhattacharya, Shubhobrata, Gowtham Sandeep Nainala, Prosenjit Das, and Aurobinda Routray. "Smart Attendance Monitoring System (SAMS): A Face Recognition Based Attendance System for Classroom Environment." *IEEE 18th International Conference on Advanced Learning Technologies (ICALT).* IEEE, 2018.

[58] Shaha, Manali, and Meenakshi Pawar. "Transfer learning for image classification." *2018 second international conference on electronics, communication and aerospace technology (ICECA).* IEEE, 2018. 656-660.

`