Available online at www.qu.edu.iq/journalcm
JOURNAL OF AL-QADISIYAH FOR COMPUTER SCIENCE AND MATHEMATICS
ISSN:2521-3504(online) ISSN:2074-0204(print)

# A secure banking transaction system for special needs people

***Rooa K. Uraiby [a], Abbas Akram Khorsheed [b] ,Thekra Abaas[c*]***

*Department of Computer Science, Mustansiriyah University, Baghdad, Iraq. Email: aroozy121@uomustansiriyah.edu.iq[a]*

*Email: abbasarab2000@uomustansiriyah.edu.iq[b]*

ARTICLE INFO

*Article history:*
Received: 14 /08/2025
Rrevised form: 25 /11/2025
Accepted : 01 /12/2025
Available online: 30 /12/2025

Keywords:

Voice Recognition Methods

Secure digital Banking System

Fraud detection

Role of AI in voice recognition

ABSTRACT

This paper synthesizes previous research efforts focusing on the development of innovative banking transaction systems designed to empower individuals with special needs through secure and accessible financial management tools. These systems leverage advanced technological solutions and sophisticated authentication strategies to guarantee the confidentiality and integrity of sensitive financial data. The creating a customizable interfaces and settings tailored to a variety of cognitive and physical abilities are also integrated by User-centered design principles. [10] Testing and User feedback elucidate the significant positive impact of these systems on the banking experiences of individuals with special needs, offering a reliable, secure, and accessible means for managing their financial well-being.[1],[4],[6]

MSC..

https://doi.org/10.29304/jqcsm.2025.17.42566.

## 1. Introduction

While the digitalization of banking has improved overall accessibility, certain challenges persist for individuals with special needs.[1],[2] To bridge this gap, voice recognition technology offers a compelling solution by enabling secure and user-friendly banking experiences for people with mobility or visual impairments. Banking systems incorporating voice recognition allow users to authenticate, transact, and retrieve account information simply by using their voices. The enhanced security and accessibility afforded by this technology are particularly beneficial for clients with special requirements. Machine learning and deep learning algorithms, especially neural networks, are fundamental to the sound recognition capabilities of these banking systems. These algorithms are trained to identify each individual's unique voice characteristics, enabling accurate speaker verification with minimal error. This paper explores these technical and practical facets, analyzing the algorithms and applications of voice-enabled banking for people with special needs. Transactions of Secure banking systems are garnering increasing attention,

∗Corresponding author Rooa K. Uraiby

Email addresses: aroozy121@uomustansiriyah.edu.iq

Communicated by ‘sub etitor’

propelled by the imperative to simplify financial inclusion and safeguard vulnerable populations from financial risks [4],[6]. This need presents a particularly salient challenge when addressing individuals with special needs, who may encounter amplified barriers in accessing and securely navigating conventional banking services. Consequently, this paper endeavors to explore and Discuss secure banking transaction systems specifically tailored to meet the unique requirements of this demographic, with a pronounced emphasis on striking a delicate equilibrium between user-friendliness and robust protection against potential security threats.[1],[2],[3],[4],[5]

***2. RELATED WORK***

Related literature examines the creation and implementation of secure banking transaction systems specifically tailored to meet the needs of customers with disabilities.[5],[1] These studies aim to simulate the unique challenges and barriers this demographic faces when conducting financial transactions, including restricted mobility, cognitive impairments, and visual or auditory deficits.[3],[5] For developing secure and user-friendly banking platforms that protect the privacy and security of users with special needs, research investigates a range of technologies and strategies, such as voice recognition, biometric authentication, and accessible user interfaces.[9] The findings address ethical considerations regarding consent and privacy in the design of assistive technologies, while also advancing financial inclusion and promoting greater independence for this underserved population.[9],[10]

### 2.1-Recognition of voice and speech using NLP Techniques

This paper explains a novel approach to automated biometric voice and speech recognition, that are employing a hybrid methodology incorporating Natural Language Processing (NLP), Convolutional Neural Networks (CNN), and Gaussian Mixture Models (GMM). For developing security systems, extracted features were taken usefulness from the (GMCC). The system's architecture increases features extracted from the Mel-Frequency Cepstral Coefficients (MFCC) model to enhance security and limit breaches by pre-recording user voice data in a secure database. While the CNN was implemented to attain high-accuracy voice discrimination, yielding a success rate of approximately 95%, which significantly enhances the system's performance. In Addition, a GMM was incorporated to generate probability ratios for voice recognition tasks, supporting the system's capacity to accurately distinguish between speakers. The results are expounding that the system has an improvement of security by exploiting the unique nature of customer's voice biometrics, making it less susceptible to breaches. Also, the combination of CNN and GMM proved effective, achieving satisfactory accuracy in distinguishing different voice data. Future work may explore the integration of additional biometric methods to develop a more comprehensive and robust voice recognition system.[1]

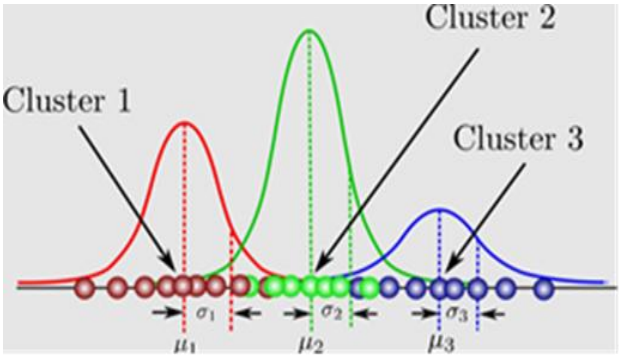


Fig 2.1: - Gaussian Mixture Models.[11]

### 2.2 An AI Based Chat Bot for Banking Applications using Intelligent Chat Bots

The key points explained in this paper were :(AI-Powered Assistant: the development of a chat and voice assistant that uses artificial intelligence (AI) technology. That assistant is specifically designed for integration into banking applications, supporting user interaction and service delivery , while the Human-like Interaction represented the ability of the assistant to simulate human conversations was One of the significant advancements highlighted and that is achieved through sophisticated data translation techniques that agree both text and audio interactions to feel more natural and engaging for users, and Technological Advancements: that a wide improvements in the fields of machine learning and natural language processing (NLP).in which These advancements enable the assistant to learn from past interactions, which provides to more human-like conversations and responses , in additional the Versatile Functionality: wherever that many tasks such as calculating loan interests, verifying transaction details, and checking account balances can be processed by the user with help of The AI assistant is capable of performing a variety of banking-related tasks . That versatility makes it a valuable tool for enhancing customer service in banking applications. at last Evolution of Assistants: The paper certains the evolution of voice and chat assistants since their inception. The continuous technological progress has transformed these tools from basic functionalities to more complex systems capable of understanding and responding to user requirements effectively).[2]

### 2.3 Real Time Voice Recognition System using Tiny ML on Arduino Nano 33 BL

An audio recognition structure that utilizes TinyML that the scientific article's goal for creating competencies that allow machine studying items to art work on source-constrained gadgets, for instance, the Arduino Nano 33 BLE. A digital assistant, speech-to-text translation, car navigation, audio biometrics, and domestic automation are combined as the machine's applications of the system. The performance of this structure entails schooling a TinyML version making use of the (Edge Impulse) framework, which has an impressive accuracy of 97%, which lets in the machine to efficiently be used. This task no longer most effective shows technical feasibility however moreover emphasizes the developing significance of audio recognition in regular technology.[3]

### 2.4- PAYV - Payment Voice: A Platform using Voice Recognition to Enable Payment Transactions

The development of the PAYV system is detailed in this paper, which aims to improve the accessibility of digital payment transactions, especially for individuals with disabilities. the design and functionality that supports this goal by central focus was an important part of the general framework. The system uses a group of advanced technologies, consisting of voice recognition, natural language processing (NLP), face recognition, and fingerprint recognition, to ensure a secure and efficient user experience during payment transactions. A key component of PAYV is its robust user authentication process, which employs multiple biometric modalities (face and fingerprint recognition) alongside voice commands. This approach is designed for securing transactions and permit users to authenticate themselves easily across various devices. Additionally, the integration of artificial intelligence (AI) employs to streamline payment processes by decreasing the need for extensive user intervention, thereby accelerating transactions and enhancing efficiency, a vital consideration in the digital economy. Moreover, the design of PAYV is adaptable to ongoing advancements in speech recognition technology, ensuring its continued relevance and effectiveness. The improvement of security, efficiency, and inclusivity of digital transactions were the significance of this initiative resides in its potential, representing a substantial advancement in the evolution of payment systems by enhancing accessibility for a wider audience.[4]


**Fig 2.3: - Natural Language Processing (NLP). [10]**

### 2.5 Implementation of the Voice Based Control and Detection of the Currency in ATM Transaction

This paper introduces presenting of a voice-controlled system designed for facilitating various user commands related to ATM transactions. By modern banking, automated teller machines (ATMs) are essential since they provide clients with easy access to their accounts and banking services. While the Traditional ATMs, however, predominantly depend on touch screens and physical buttons, which can pose limitations for certain segments of the population. A minimum than ideal banking experience could result from ATMs' traditional user interfaces' inability to adequately accommodate all users, especially those with impairments or those who have trouble using touch screens. For mitigating these problems, this paper introduces an innovative approach that integrates voice recognition technology, supporting user interaction with ATMs by enabling users to control transactions and detect currency denominations using voice commands. This is completed through a software interface that's employing an advanced speech recognition algorithms and machine learning models, enabling the system to accurately interpret user commands and identify different currency denominations during transactions. A key feature of the proposed solution with existing ATM hardware is the compatibility, in which it is ensuring seamless integration of the voice-based control and detection system into current banking infrastructures without necessitating substantial modifications. The applying of this voice-based system is projected to transform the ATM experience, enhancing both user-friendliness and security. The proposed solution has the potential to greatly enhance the general banking experience for a variety of users by strengthening accessibility and security.[5]

### 2.6 MFCC and VQ voice recognition-based ATM security for the visually disabled

A system designed This paper presents the providing a secure and accessible ATM access for visually impaired customers by voice recognition technology. That system authenticates users through analyzing distinctive vocal traits, thereby circumventing the reliance on PIN codes, which can be challenging to use and vulnerable to theft. Core technologies employed contain Vector Quantization (VQ) for voice matching and user modeling, and Mel-Frequency Cepstral Coefficients (MFCC) for feature extraction. The authors suggest system that leverages voice processing and pattern recognition (MFCC+VQ) for developing a secure and user-friendly ATM authentication solution for blind and visually impaired customers. This approach points some of the usual limitations of conventional verification methods while enabling visually impaired customers to securely and freely access ATMs through recording and identifying distinctive speech patterns.[6]

## 3. METHODLOGIES

**An adopted strategies and methodologies tailored to banking transactions in the previous studies, emphasizing the using of state-of-the-art voice recognition techniques and other advanced methods for**

**individual's authentication and identity verification, thereby ensuring the security of person's information. The maigor role of artificial intelligence in developing intelligent models for optimizing performance has been consistently highlighted. In the following divisions, we will discuss the most prominent methods employed in the aforementioned paper: -**

**3.1** The paper**" Recognition of voice and speech using NLP Techniques"** was employing many core methodologies in its development of a robust voice and speech recognition system utilizing Natural Language Processing (NLP) techniques. The structure of system focused on a combined biometric voice recognition system, integrating both voice recognition and speech recognition models to enhance total performance and security. Specifically, the Speech Recognition Model used Grid Mel-Frequency Cepstral Coefficients (GMCC) as features, which the researchers identified as an effective method for discerning speech patterns by extracting salient characteristics from audio sources. For processing these GMCC features, a Convolutional Neural Network (CNN) architecture with two layers was executed, capitalizing on the aptitude of CNNs for treating grid-like data such as audio spectrograms. The synergistic effect of combination of GMCC features with CNN architecture proved considerable in which promises to improve speech recognition efficacy. An integrated circuit represented through, the Voice Recognition Model gathered a Gaussian Mixture Model (GMM) to process Mel-Frequency Cepstral Coefficients (MFCC), thereby facilitating speaker recognition. GMMs, was employed as statistical models, are well-suited to elucidating the distribution of features within a dataset. While MFCCs provides a crucial role in capturing the distinctive attributes of an individual's voice. In additional face the system featured a User Interface that authorized users to register their voice and other biometric data. This interface simplifies the construction of a customized database, securely storing customer information to allow for rapid access and recognition during subsequent interactions. A key tenet of this paper was an assurance on security. The employing methodologies sought for supporting the biometric system's security across all dimensions, mitigating the risk of security breaches. This security highlight represents a critical component, ensuring the system's resilience against potential attackers.[1]

The main layers and equations that are involved in CNNs: -

**1. Convolutional Layer:**

$$\text{Output}[i, j] = \Sigma\,\Sigma\ \text{Input}[i + m, j + n] * \text{Kernel}[m, n]$$

**2. Activation Function:**

$$\text{Output}[i, j, k] = \max(0, \text{Input}[i, j, k])$$

**3. Pooling Layer (Max Pooling):**

$$\text{Output}[i, j, k] = \max(\text{Input}[x, y, k]) \text{ for } x \text{ in region, } y \text{ in region } [10]$$

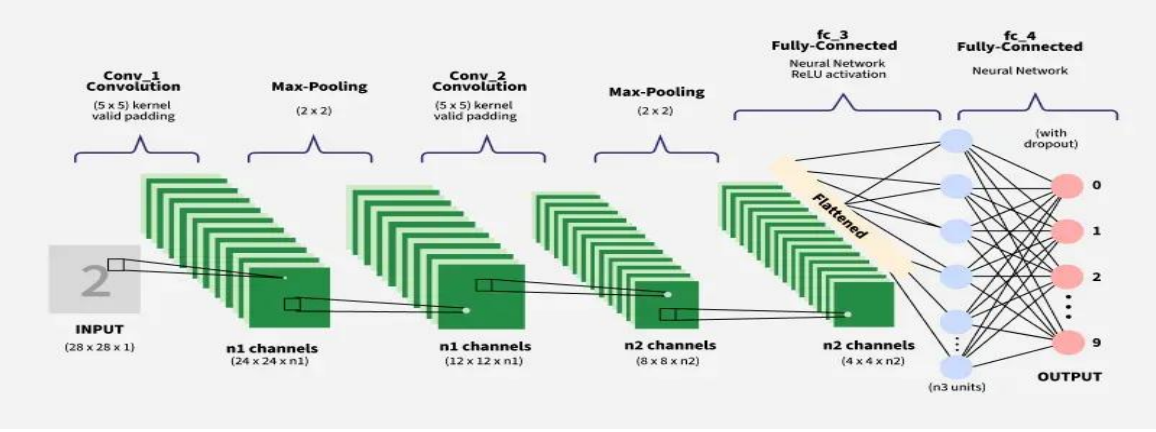


**Fig 3.1 Convolutional Neural Network (CNN).[10]**

**3.2** In the paper**, "An AI Based Chat Bot For Banking Applications using Intelligent Chat Bots"** authors applied the following key methodologies in its development: Artificial Intelligence Integration: A core methodology contained the integration of AI technologies into the banking application. This presumably included the utilization of

machine learning methodologies, enabling the chatbot to assimilate knowledge from user interactions and enhance its efficacy over time, resulting in higher effectiveness. and natural conversations representing by a Natural Language Processing (NLP): the NLP techniques were employed for enabling the chatbot to accurately interpret user queries and respond in a conversational and intuitive manner. Voice Recognition Technology: The system was incorporating voice recognition technology for permitting customers to interact via spoken commands, thereby enhancing accessibility and improving the overall user experience. Data Transformation Techniques: here Data transformation techniques were implemented for converting the raw data into a format readily understood and processed by the AI, facilitating smoother and more efficient interactions. eventually User Experience (UX) Design: UX design principles were employed to generate a user-friendly interface. This entailed meticulously designing the user interface and interaction flow for simplifying seamless user engagement with the chatbot for diverse financial processes.[2]


**Fig 3.1 Chat Bot For Banking Applications [12]**

**3.3"Real Time Voice Recognition System using TinyML on Arduino Nano 33 BLE**," implemented the following key methodologies: TinyML Framework: A TinyML framework was holding down a job, in which a representing of specialized approach to machine learning designed for execution on low-energy and resource-constrained devices. This was major for the Arduino Nano 33 BLE, given its inherent limitations in processing power and memory. Model Training: The "Edge Impulse framework" was used for training the voice recognition model. This platform is well-suited to the constraints of the Arduino Nano 33 BLE, as it is specifically designed to facilitate the construction and deployment of machine learning models on edge devices. Keyword Recognition: The system was highlighting specifically on recognizing the keywords "ON" and "STOP." These keywords are necessary for system operation, enabling it to perform actions in response to user commands. Deployment on Hardware: while in this aspect the hardware setup handled the system for capturing voice commands in real-time and respond accordingly. Action Triggering: onto the recognition of the designated keywords, specific actions were turned on, such as activating a lightbulb connected to the Arduino via a relay. That was an illustration of the practical application of the voice recognition technology in a home automation context. Performance Evaluation: finally, the evaluated of the voice recognition system's performance, exhibiting a noteworthy accuracy rate of 97%. That's high level of accuracy explains the effectiveness of the model training and implementation techniques employed.[3]

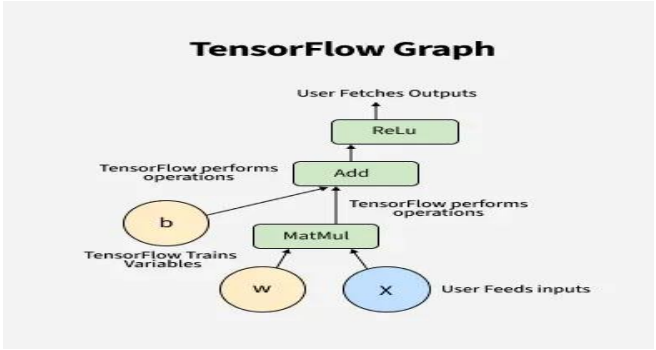


**Fig 3.3 Tensor Flow Graph**

3.4 The paper, "**PAYV - Payment Voice: A Platform using Voice Recognition to Enable Payment Transactions,**" leverages several key points for facilitating an efficient and secure payment transactions by the development of a virtual voice assistant. The most important methodologies were employed in this system consist

of: Voice Recognition Technology: a state-of-the-art voice recognition technology was incorporated in the system to accurately interpret and process user commands. This approach is enhancing accessibility and user-friendliness by enabling customers to start financial transactions using their voice. Natural Language Processing (NLP): NLP techniques are employed for facilitating seamless and intuitive interactions between the user and the system. This allows the platform to understand the intent behind user requests, even with variations in phrasing and syntax. Biometric Authentication: both fingerprint and face recognition technologies were employed in PAYV for robust user authentication. This multi-factor authentication strategy is supporting transaction security by requiring customers to verify their identity using iometric data, which is inherently difficult to forge. Cross-Device Compatibility: The design of system was generated to operate seamlessly across a variety of devices, ensuring broad accessibility. This is particularly crucial for persons with disabilities, as it enables authentication and transaction completion regardless of the user's preferred platform. Artificial Intelligence (AI) Integration: Furthermore, AI was playing the significant role in streamlining payment processes. The system operates transactions with a rate of effectiveness by reducing the need for human intervention, which is particularly advantageous in a rapidly evolving digital environment. Technological Adaptability: the PAYV's architecture was designing to adapt to ongoing advancements in voice recognition and related technologies. This inherent flexibility ensures the system's continued effectiveness and modernity, while providing consumers with a reliable and future-proof payment option.[4]

The main steps and its equations that are included in NLP: -

1. **Bag of Words (BoW):**

   **TF (t, d) = count(t) in d …… (for Term Frequency)**

2. **TF-IDF (Term Frequency-Inverse Document Frequency):**

   **TF (t, d) = count(t) in d …. ( for Term Frequency)**

   **IDF (t, D) = log (|D| / | {d ∈ D: t ∈ d} |) ……(for Inverse Document Frequency (IDF))**

   **TF-IDF (t, d, D) = TF (t, d) * IDF (t, D) ……. (TF-IDF)**

3. **Cosine Similarity:**

   **Similarity *(A, B) = (A · B) / (||A|| * ||B||)***

4. **Neural Networks (Word Embeddings, Recurrent Networks, Transformers):**

   - **Recurrent Neural Networks (RNNs):**

     **h_t = f(W * h_{t-1} + U * x_t + b)**

   - **Transformers (Attention Mechanism):**

     **Attention(Q, K, V) = softmax( (Q K^T) / sqrt(d_k) ) V [10]**

**3.5 The paper, "Implementation of the Voice Based Control and Detection of the Currency in ATM Transaction,"** utilized the following methodologies: Voice Recognition Technology: The initial approach contained integrating voice recognition technology for enhancing ATM accessibility and user experience. This provides handling for customers interaction with ATMs using spoken commands, streamlining the transaction process. To implement the Speech Recognition Algorithms: for developing of the software interface in demanded the implementation of sophisticated speech recognition algorithms capable of accurately interpreting user voice commands in varied acoustic environments. Machine Learning Models: in the other side the Machine learning models were consolidated to improve the accuracy and efficiency of currency denomination detection. The ATM was qualified for identifying the quantity requested by the customer through voice commands as a result of the training of these models to distinguish and recognize a variety of currency notes. Software Interface Development: here the key component of the methodology was the generation of a user-friendly software interface in which seamlessly combines the voice recognition and machine learning units. The design of that inter face has been intuitive, for simplifying ATM system navigation for users interacting via voice. Compatibility Testing: The development of

system was processed by a strong emphasis on compatibility with existing ATM hardware. Extensive testing of the software interface was conducted to ensure seamless integration into current banking systems, minimizing the need for significant hardware modifications. User Testing and Feedback: At last the indispensable element of the development was represented with User testing.so its Gathering feedback from a diverse user base is crucial for refining voice recognition capabilities, optimizing overall user experience, and ensuring that the system effectively meets the needs of all customers.[5]


**Fig 3.4: - PAYV - Payment Voice.[11]**

3.6 The paper**, "MFCC and VQ voice recognition-based ATM security for the visually disabled,"** employed the following fundamental methodologies: Voice Input Collection for Authentication: the customer prompted by this system for articulating a specific passphrase or password into the ATM's microphone to collect a distinct audio sample. This audio sample is subsequently implemented with using Mel-Frequency Cepstral Coefficients (MFCC) for extracting features that represent the unique characteristics of the customer's voice. Vector Quantization-Based Codebook Creation: The produced MFCC features are then employed in conjunction with Vector Quantization (VQ) for generating a codebook in which it is unique to learn customer during registration. This codebook play role as a personalized representation of the user's voice. Verification Phase via Codebook Comparison: In the phase of verification, a new voice samples are acquired and compared against the customer's database (stored codebook). The point of similarity between the new sample and the codebook determines access to the ATM system. System Integration for Visually Impaired Users: The systemin this phase was developed in a way to be seamlessly integrated into existing ATM interfaces to provide secure authentication for visually impaired users through a combination of audio prompts and feedback, thereby enhancing accessibility and security.[6]

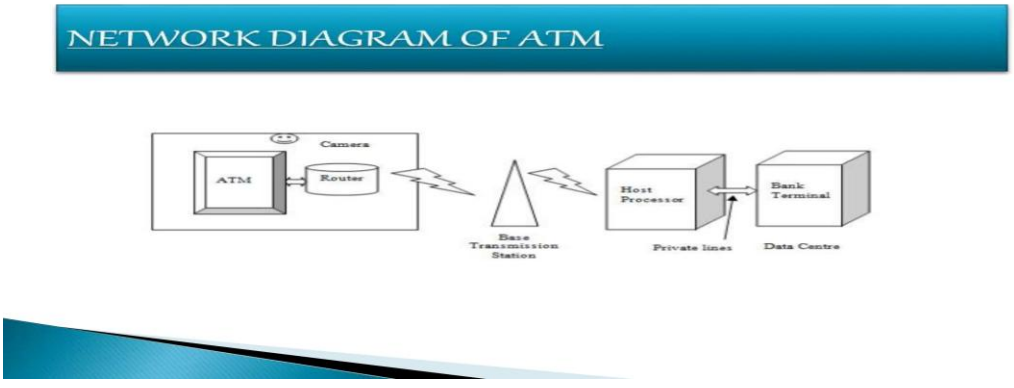


**Fig 3.5: - VOICE AIDED ATM MACHINE is used for visually challenged people | PPTX. [11]**

## 4. RESULTS AND DISCUSSION

### 4.1 RESULTS:

A synthesis of prior studies revealing a multifaceted comparative analysis spanning various key dimensions of voice recognition systems. fundamentally, the heterogeneous methodologies executed across many papers, encompassing a spectrum of algorithms and techniques, were rigorously evaluated in terms of their efficacy in achieving optimal user voice recognition accuracy and general system performance.[1],[9],[8], particularly the strategic choice of feature extraction methodologies (e.g., Mel-Frequency Cepstral Coefficients (MFCCs), Linear Predictive Coding (LPC), deep learning-derived feature embeddings), classification algorithms (e.g., Hidden Markov Models (HMMs), Gaussian Mixture Models (GMMs), Support Vector Machines (SVMs), Deep Neural Networks (DNNs)[1],[2][7],[6] and noise reduction paradigms (e.g., spectral subtraction techniques, Wiener filtering approaches, adaptive filtering methodologies) was systematically assessed to determine its impact on established performance metrics, such as Word Error Rate (WER) and Equal Error Rate (EER),[2],[7],[8], [9].In other hand , a quantitative benchmarking was performed, centered on the percentage accuracy of system outputs following matching with input data. Concurrently, the Critical evaluation was performed in each system's security postures, with a focus on how resilient they were to various spoofing assaults. (e.g., replay attacks, impersonation attempts, voice synthesis-based adversarial examples). [1],[2], The measurements effectiveness, such as anti-spoofing feature engineering, liveness detection procedures utilizing physiological or behavioral biomarkers, and cryptographically secure voiceprint storage solutions, underwent thorough examination.[4],[7],[9] Additionally, the level of user flexibility, accessibility, and inclusivity provided by each system was compared across a range of use-case scenarios, considering aspects such as intuitiveness of interface design, latency of response, and inherent adaptability to variations in accent, dialect, and speaking style.[1],[2],[7],[6],[9] Furthermore Challenges commonly materialized by these systems, encompassing sensitivity to non-stationary environmental noise profiles, computational resource constraints, and the imperative for large-scale, representative training corpora, were meticulously identified and critically discussed.[2],[3],[4], Future development and optimization possibilities were tested, involving the strategic establishment of transfer learning paradigms, federated learning architectures for privacy-preserving collaborative model training, and edge computing frameworks for low-latency, real-time processing at the network edge.[1],[2], [5]. Ultimately, the progressive role of artificial intelligence (AI), machine learning (ML), and deep learning (DL) in supporting and enhancing these systems was thoroughly examined. The mutual benefits of leveraging AI/ML/DL for critical missions such as automatic feature representation learning, adaptive noise suppression with enhanced robustness, and personalized voice modeling tailored to individual user characteristics were meticulously highlighted.[1],[2],[3],[7],[9]. Potential for algorithmic biases were included the concomitant ethical implications of deploying AI-driven voice recognition systems, the, data privacy violations, and the exacerbation of existing societal inequalities, were also carefully addressed from a responsible innovation perspective.[5],[10]

### 4.2 Discussion

The proliferation of smart banking systems in the contemporary digital landscape offers customers enhanced convenience and efficiency in financial management. However, the escalating prevalence of cyber threats necessitates a paramount focus on the robust security of these systems. Key considerations for achieving this include: Provision of Advanced Data Protection Mechanisms: Smart banking systems must implement state-of-the-art encryption technologies to safeguard sensitive customer data and prevent unauthorized access. Robust encryption protocols are essential to maintaining data confidentiality and integrity .[1],[6],[7],[15],[16]Adherence to Global Privacy and Security Standards: Compliance with globally recognized privacy and security standards, such as GDPR, CCPA, and PCI DSS, is crucial to ensure the comprehensive protection of customer data and maintain user trust .Development and Integration of Intelligent Technologies: Banking institutions should prioritize the development and integration of advanced artificial intelligence and data analytics technologies. These technologies should be designed to optimize the user experience while simultaneously enhancing security and ease of use. Examples include AI-powered fraud detection systems and biometric authentication methods.[6],[8],[9],[ 10],[13] Comprehensive Training and Education Programs: Banking institutions must provide comprehensive and ongoing training programs for their personnel to enhance their expertise in managing security technologies and effectively responding to emerging cyber threats. By addressing these critical aspects and proactively mitigating current challenges, smart banking systems can achieve enhanced accuracy and security, effectively meeting customer needs while ensuring the preservation of their privacy and sensitive data.[1],[3],[6],[10],[16]

### *4.3 Table 1: Table of Comparing previous studies of several performance evaluation and system usability*

| Papers name & year | The using methods | Accuracy Rate % | Security level | Flexibility | challenges |
|---|---|---|---|---|---|
| 1." Recognition of voice and speech using NLP Techniques" 2024 | CNNs, GMM, NLP, HMMs, DNNs, RNNs | -Baseline (Acoustic Model only): 10% - 25% (depending on the dataset, noise, etc.)<br>- WERs in the range of 5% - 15% | Few security vulnerabilities:<br>-Vulnerability to Adversarial Attacks<br>-Privacy Leakage & Spoofing Attacks | -Adaptation to Different Domains<br>-Robustness to Accents and Speaking Styles<br>-Contextual Understanding<br>-Error Correction | -Complexity of Natural Language<br>- Data Sparsity<br>-Computational Cost<br>-Real-Time processing<br>-Lack of Robustness<br>-Bias |
| 2. "An AI Based Chat Bot For Banking Applications using Intelligent Chat Bots" 2023 | AI-powered chat and voice assistant integration,<br>Machine learning and natural language processing advancements | Voice Command Recognition: 90% - 99% accuracy is a realistic goal, especially with constrained vocabularies and clear prompts. | a multi-layered issue, and the evaluation isn't as simple as a single percentage.<br>-Voice-Based Identification Bypass<br>- Adversarial Knowledge<br>-Imperfect Accuracy | It should adapt to different users, languages, environments, and transaction types, while maintaining a high level of security and reliability. The more adaptable the system, the more likely it is to be adopted by a wider user base. | Achieving High Security<br>•Maintaining Imperceptibility<br>•Balancing Security and Practicality<br>•Data Availability<br>•Ethical Considerations |
| 3. "Real Time Voice Recognition System using Tiny ML on Arduino Nano 33 BLE " 2024 | TinyML model trained using EdgeImpulse framework,<br>Real-time voice recognition system implemented on Arduino Nano 33 BLE | Very Simple Case (2-3 words, quiet environment, speaker-dependent, optimized code): 75%-90% accuracy. This is the best-case scenario. | Level of Security (Extremely Limited):<br>• Lack of Robust Authentication.<br>•Vulnerability to Replay Attacks<br>•Vulnerability to Spoofing<br>•No Secure Storage<br>•Simple Security Measures | Level of Flexibility (Very Limited) | • Memory Constraints<br>• Processing Power<br>• Energy Consumption<br>• Feature Extraction<br>•Model Size Optimization<br>•Real-Time Performance<br>• Limited Dataset<br>• Hardwar Limitations |
| 4. "PAYV - Payment Voice: A Platform using Voice Recognition to Enable Payment Transactions" 2024 | Voice Recognition Technology, Natural Language Processing (NLP), Biometric Authentication, AI Integration | Rate of 95% to 99% or even higher for voice authentication, in the range of 85% to 95% of Realistic Estimate for Real-World Accuracy. | Weak | Low Flexibility: - Limited support for different accents or languages.<br>- Requires very specific voice commands. | Balancing Security and Usability<br>Preventing Replay Attacks, Detecting Voice Mimicry, Robustness to Noise |

| Papers name & year | The using methods | Accuracy Rate % | Security level | Flexibility | challenges |
|---|---|---|---|---|---|
| 5. "Implementation of the Voice Based Control and Detection of the Currency in ATM Transaction" 2024 | Voice Recognition Technology<br><br>-Speech Recognition Algorithms<br><br>- Machine Learning Models<br><br>-Software Interface Development | Ideal Conditions (Quiet Environment, Clear Speech): 95% - 99% accuracy can be achievable with modern speech recognition engines.<br><br>-Real-World Conditions (Noisy. | Voice Authentication Weaknesses<br><br>-If the system relies solely on voice, security is very low.<br><br>With robust MFA, end-to-end encryption, and fraud detection, security could be moderate. | The system's adaptability to various user needs, environmental conditions, and future extensions | • Security vs. Usability<br><br>• Replay Attack Mitigation<br><br>• Voice Mimicry Detection<br><br>•Noise Robustness<br><br>•Accent and Language Support |
| 6. "MFCC and VQ voice recognition-based ATM security for the visually disabled" 2017 | MFCC, VQ for feature vector compression and codebook generation<br><br>and, either a HMMs or Dynamic Time Warping based classifier to make the final decision. | -Quiet environment<br>~95%<br>~2%<br>~3%<br><br>-Noisy environment<br>85–90%<br>~5%<br>~8%<br><br>-Voice variation<br>80–90% | Very low | Flexibility refers to the system's ability to work under varying conditions:<br><br>•Speaker Variability<br><br>•Noise Robustness<br><br>•Phrase Variation | Balancing Security and Accessibility<br><br>-Replay Attack Mitigation<br><br>-Technical Difficulty<br><br>-Noise Robustness |

## 5. conclusion

In conclusion, this discussion has synthesized previous research that has explored diverse methodologies and techniques within the field of artificial intelligence and their application in constructing digital banking systems capable of providing immediate services to users [4],[5], [13],[14],[16]. However, these studies have acknowledged inherent vulnerabilities, limitations, and persistent challenges that warrant continued investigation to enhance system performance in future iterations. Approaches to mitigation may involve the implementation of alternative algorithms, comparative analyses of current and prior results, and iterative refinement of existing algorithms [7],[8],[15]. These strategies are particularly relevant given the rapid advancements in artificial intelligence technologies, machine learning, and deep learning.[1],[2],[3],[16] The review of prior studies and literature assuring the role of artificial intelligence, and its more advanced and impactful subfields such as ( ML and DL) [1],[6] , in building a digital banking systems discloses that the judicious selection of an operational strategy, encompassing the choice of the most effective and accurate technologies and algorithms, employs a significant influence on the attainment of favorable or reasonably satisfactory outcomes. However, a preponderance of existing systems involved challenges that warrant future resolution for facilitating development and achieve supported results, particularly in bolstering security measures. Given the paramount sensitivity of customer banking data, its preservation constitutes a fundamental objective in enhancing digital banking systems. This is especially critical in the last years characterized by the exponential advancement of digital technologies, particularly in the domains of cyberattacks and increasingly sophisticated fraud techniques, which create a constant threat of infiltrating financial systems through any existing vulnerabilities.[5],[10]

**Acknowledgements**

The author gratefully acknowledges the unwavering support and expert guidance of their research supervisors, Dr. Abbas Akram khorsheed and Dr. Thekra Abbas, whose insights were instrumental to the completion of this work. Deep appreciation is also extended to Al-Mustansiriya University.

## References

[1] Jadhav, D., Shirsath, C., Sahasrabuddhe, S., & Singh, P. (2024). Recognition of Voice and Speech Using NLP Techniques." 75–80. https://doi.org/10.1201/9781003433309-7

[2] Ajmeera, K., Dbk, K., Paulraj, D., Vidhya, V., & V, S. (2023). *An AI Based Chat Bot For Banking Applications using Intelligent Chat Bots*. 1–4. https://doi.org/10.1109/rmkmate59243.2023.10369062

[3] Gupta, N., & Gajjar, S. (2023). *Real Time Voice Recognition System using Tiny ML on Arduino Nano 33 BLE **. 385–388. https://doi.org/10.1109/ises58672.2023.00085

[4] Kambampati, P., Rane, S., Shoeb, A., & Dhannawat, R. (2024). *PAYV - Payment Voice: A Platform using Voice Recognition to Enable Payment Transactions*. 1–6. https://doi.org/10.1109/apcit62007.2024.10673442

[5] T. N, "Implementation of the Voice Based Control and Detection of the Currency in ATM Transaction," *2024 3rd International Conference for Innovation in Technology (INOCON)*, Bangalore, India, 2024, pp. 1-4, doi: 10.1109/INOCON60754.2024.10511632.

[6] E. D. Dimaunahan, A. H. Ballado, F. R. G. Cruz and J. C. Dela Cruz, "MFCC and VQ voice recognition based ATM security for the visually disabled," *2017IEEE 9th International Conference on Humanoid, Nanotechnology, Information Technology, Communication and Control, Environment and Management (HNICEM)*, Manila, Philippines, 2017, pp. 1-5, doi: 10.1109/HNICEM.2017.8269516.

[7] Bhavani, R., & Kathirvelu, K. (2024). *Advancements in Speech Recognition: Exploring Feature Extraction Techniques for Enhanced Accuracy and Human-Machine Communication*. 1–8. https://doi.org/10.1109/icaect60202.2024.10468885 In-text Citation (Bhavani & Kathirvelu, 2024)

[8] Chetalam, L. J. (2018). *Enhancing Security of Mpesa Transactions by Use of Voice Biometrics*. http://erepo.usiu.ac.ke/handle/11732/4113

[9] Oyewale, O. O., Hossain, M. D., Taenaka, Y., & Kadobayashi, Y. (2024). *Optimizing Voice Biometric Verification in Banking with Machine Learning for Speaker Identification*. 377–384. https://doi.org/10.1109/apcc62576.2024.10768085

[10] https://www.geeksforgeeks.org/

[11] https://www.google.com/

[12] https://automationedge.com/blogs/ai-chatbot-in-banking/

[13] O. Abdel-Hamid, A. -r. Mohamed, H. Jiang, L. Deng, G. Penn and D. Yu, "Convolutional Neural Networks for Speech Recognition," in *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, vol. 22, no. 10, pp. 1533-1545, Oct. 2014, doi: 10.1109/TASLP.2014.2339736.

[14] Pattern Recognition and Classification (Geoff Dougherty)

[15] R. Maganti, J. Gutla, Y. Mallimpalli and S. Yellisetti, "Innovative Currency Identifier for the Blind through Audio Output using Deep Learning," *2024 International Conference on Automation and Computation (AUTOCOM)*, Dehradun, India, 2024, pp. 358-363, doi: 10.1109/AUTOCOM60220.2024.10486196.

[16] A. S. Nimgaonkar and R. Kumbhar, "Usability - A Key to Security and Utility: Case of Core Banking System," *2024 International Conference on Emerging Smart Computing and Informatics (ESCI)*, Pune, India, 2024, pp. 1-6, doi: 10.1109/ESCI59607.2024.10497323.

[17] Yang, S.-B., & Liang, T.-W. (2023). Deep Multi-Layer Neural Network with Variable-Depth Output. *International Journal of Pattern Recognition and Artificial Intelligence*. https://doi.org/10.1142/s021800142359022x

[18] Zhang, S., Vassiliadis, V. S., Dorneanu, B., & Arellano-Garcia, H. (2023). *Hierarchical multi-scale parametric optimization of deep neural networks*. *53*, 24963–24990. https://doi.org/10.1007/s10489-023-04745-8

[19] Nazaret, A., & Blei, D. M. (2022). Variational Inference for Infinitely Deep Neural Networks. *International Conference on Machine Learning*, 16447–16461. https://doi.org/10.48550/arXiv.2209.10091

[20] Alturki, A., Bchir, O., & Ben Ismail, M. M. (2022). Depth-Adaptive Deep Neural Network Based on Learning Layer Relevance Weights. *Applied Sciences*, *13*(1), 398. https://doi.org/10.3390/app13010398

[21] Varma, V., & Saravanan, A. K. (2023). Advancements in Speaker Recognition: Exploring Mel Frequency Cepstral Coefficients (MFCC) for Enhanced Performance in Speaker Recognition. *International Journal For Science Technology And Engineering*, *11*(8), 88–98. https://doi.org/10.22214/ijraset.2023.55124

[22] Githuku, W. A. & Kinyuru, R. N. (2018). Digital banking and customer relationship in banking industry in Kenya. International Academic Journal of Human Resource and Business Administration, 3(2), 14-32

[23] Valsamidis, S., Tsourgiannis, L., Pappas, D., Mosxou, E. (2020). Digital Banking in the New Era: Exploring Customers' Attitudes. In: Horobet, A., Polychronidou, P., Karasavvoglou, A. (eds) Business Performance and Financial Institutions in Europe. Contributions to Economics. Springer, Cham. https://doi.org/10.1007/978-3-030-57517-5_6

[24] Ahmed HASSAN, Rasha ABD EL-AZIZ and Meer HAMZA (2020), "The Exclusion of People with Visual Disabilities from Digital Banking Services in the Digitalization Era", Journal of Electronic Banking Systems, Vol. 2020 (2020), Article ID 519078, DOI: 10.5171/2020.519078

[25] A. S. Naik, "Assistive Technology in Banking," 2024 15th International Conference on Computing Communication and Networking Technologies (ICCCNT), Kamand, India, 2024, pp. 1-6, doi: 10.1109/ICCCNT61001.2024.10724888.

[26] S. Anwar, G. Alburaiki, A. Ibrahimi and L. Qadri, "Digital Banking Issues Experienced by Older People," 2024 10th International Conference on Optimization and Applications (ICOA), Almeria, Spain, 2024, pp. 1-5, doi: 10.1109/ICOA62581.2024.10754510.

[27] A. L, R. P. Lakshmi and S. Pavithra, "Cybersecurity in Banking and Cloud Computing: Threats, Defenses, and Innovations," 2025 International Conference on Data Science, Agents & Artificial Intelligence (ICDSAAI), Chennai, India, 2025, pp. 1-6, doi: 10.1109/ICDSAAI65575.2025.11011646.

[28] T. Yao, "Mathematical Statistics and Analysis on the Path Mechanism of Protecting Personal Information Relying on Information Digitization and Big Data," 2021 IEEE International Conference on Emergency Science and Information Technology (ICESIT), Chongqing, China, 2021, pp. 729-733, doi: 10.1109/ICESIT53460.2021.9696460.

[29] S. Lilia, V. Yulia, G. E. Arslanovna and Y. Evgeniy Yurievich, "Research of the need for training financial literacy under economic uncertainty in the framework of the educational process at schools," 2022 8th International Conference on Energy Efficiency and Agricultural Engineering (EE&AE), Ruse, Bulgaria, 2022, pp. 1-7, doi: 10.1109/EEAE53789.2022.9831342.

[30] I. Kovačević, S. Groš and A. Đerek, "Automatically Generating Models of IT Systems," in IEEE Access, vol. 10, pp. 13536-13554, 2022, doi: 10.1109/ACCESS.2022.3147312

[31] A. Agape and M. Postolache, "Internet-enabled Access Control System using a Mobile Application," 2018 22nd International Conference on System Theory, Control and Computing (ICSTCC), Sinaia, Romania, 2018, pp. 244-249, doi: 10.1109/ICSTCC.2018.8540687.

[32] I. Samuel, F. A. Ogunkeye, A. Olajube and A. Awelewa, "Development of a Voice Chatbot for Payment Using Amazon Lex Service with Eyowo as the Payment Platform," 2020 International Conference on Decision Aid Sciences and Application (DASA), Sakheer, Bahrain, 2020, pp. 104-108, doi: 10.1109/DASA51403.2020.9317214.

[33] M. Pahwa, S. Agarwal and T. Rawal, "Revolutionizing Customer Experience: the Impact of AI-Driven Banking Automation," 2025 First International Conference on Advances in Computer Science, Electrical, Electronics, and Communication Technologies (CE2CT), Bhimtal, Nainital, India, 2025, pp. 371-376, doi: 10.1109/CE2CT64011.2025.10939341.

[34] Biswajit Basu; Kevinkumar Garala; Bhupendra G. Prajapati, "Role of HMI in the Drug Manufacturing Process," in Human-Machine Interface: Making Healthcare Digital, Wiley, 2024, pp.329-356, doi: 10.1002/9781394200344.ch13.

[35] K. Chandrasekhar, S. Das, N. Saibabu, N. Chaitanya, T. Bommali and F. Adamova, "A Machine Learning Driven Approach to Real-Time Fraud Detection in Modern Banking System," 2025 6th International Conference for Emerging Technology (INCET), BELGAUM, India, 2025, pp. 1-5, doi: 10.1109/INCET64471.2025.11140148.