e-ISSN: 2582-5208
International Research Journal of Modernization in Engineering Technology and Science
(Peer-Reviewed, Open Access, Fully Refereed International Journal)
Volume:06/Issue:04/April-2024 Impact Factor- 7.868 www.irjmets.com

RECOMMENDATION SYSTEM USING SPEECH RECOGNITION BY CREATING USER PROFILES

Mrs. S. Satya Sudha[*1], K. Shiva Sai Venkat[*2], K. Jayanth[*3], K. Thirumal Reddy[*4]

[*1]Assistant Professor, Department Of CSE Artificial Intelligence & Machine Learning ACE Engineering College, Hyderabad, Telangana, India.

[*2,3,4]IV B.Tech Students, Department Of CSE (Artificial Intelligence & Machine Learning ACE Engineering College, Hyderabad, Telangana, India.

DOI : https://www.doi.org/10.56726/IRJMETS51922

ABSTRACT

This project presents an innovative approach to recommendation systems by seamlessly integrating speech recognition technology with user profiling. The system, comprising a sophisticated user interface, speech recognition module, recommendation engine, user profiles, keyword extraction, and a robust database system, is designed to redefine user interactions in the digital realm. The user interface facilitates intuitive navigation, personalized profile management, and visually appealing recommendation displays. Leveraging advanced speech recognition, the system converts spoken words into text, enabling a natural and effortless mode of communication. The recommendation engine employs collaborative and content-based filtering, dynamically adjusting suggestions based on user profiles and extracted keywords from spoken content. User profiles store evolving preferences, demographic details, and interaction histories, adapting over time through continuous user feedback. Keyword extraction processes transcribed speech to extract meaningful terms, enhancing recommendations with contextual relevance. The database system ensures efficient storage and retrieval of user data, content details, and keyword associations, contributing to a responsive and dynamic user experience.

I. INTRODUCTION

Recommendation system based on speech recognition using user profiles is focused on creating a recommendation system that entices the interests of users in an intricate manner using speech as the mode. Traditional speech recognition systems have always been too complex even for single task like transcribing text from speech. They lacked diversity in the content recommended that is lack of new content discovery. They use collaborative filtering that is items are recommended based on similar tastes. Speech recognition is the most crucial part and traditional systems often faced challenges with limited integration or if at all used they used simple keyword based commands rather than extracting context-rich information from spoken content. Depth of user profiling is also limited to user inputs. The System proposed in this project is Recommendation System using Speech Recognition by creating user profiles where the speech content is analyzed and filtered based on Speech Profile. The recommendations are the critical keywords associated with each profile these are gathered using NLP libraries using Spacy, NLTK. These help in understanding the preferences of user in detail.

Our proposed system focuses on alleviating these issues and making recommendation systems more streamlined. Speech is innate form of communication and has been used since before the existence of writing. The main premise of using speech for the proposed recommendation system is speech conveys much more information than text.

II. LITERATURE REVIEW

The research centered on a thorough investigation into the effectiveness of various methodologies within the domain. By scrutinizing pertinent research papers, the aim was to assess a multitude of approaches and techniques employed in these areas. This process sought to reveal the nuanced intricacies and advancements within the field.

ZhenhuanYang et al. [1]: The research paper titled "Fairness-aware Differentially Private Collaborative Filtering" explores the integration of differential privacy mechanisms with collaborative filtering algorithms to protect user privacy while ensuring fair recommendations. The authors present the DP-Fair framework, which addresses the trade-offs between privacy protection and algorithmic fairness in recommendation systems. For my project, this research paper is valuable as it addresses critical challenges in privacy-preserving machine

learning, specifically in the context of collaborative filtering-based recommendation systems. By incorporating differential privacy and fairness constraints, the DP-Fair framework offers a comprehensive solution to mitigate privacy risks while ensuring fair treatment across user groups. The experimental results presented in the paper, based on datasets from Amazon and Etsy, demonstrate the superior performance of the DP-Fair framework compared to vanilla differentially private stochastic gradient descent (DP-SGD) methods. This performance improvement is observed in regards to both overall accuracy and user group fairness, highlighting the effectiveness of the proposed approach. Overall, this research contributes to advancing the field of privacy-preserving recommendation systems by offering a practical framework that balances privacy protection with algorithmic fairness. By leveraging the insights and methodologies presented in this paper, my project can benefit from enhanced privacy guarantees and improved fairness in recommendation outcomes.

Yong Zheng [2] The research paper titled "Context-Aware Collaborative Filtering Using Context Similarity: An Empirical Comparison" explores the application of context-aware recommender systems (CARS) in alleviating information overload by tailoring recommendations to user preferences in different contextual situations. The paper investigates the use of context similarity to address the sparsity problem inherent in CARS, where users may not rate items in various context situations, leading to challenges in building accurate recommendation models.

This research paper offers valuable insights into the design and implementation of context-aware collaborative filtering methods, particularly focusing on context similarity. By leveraging context information to adapt recommendations to diverse user preferences across different situations, the paper provides a framework for enhancing the relevance and effectiveness of recommendation systems. The paper presents an empirical comparison of context-aware collaborative filtering methods using multiple datasets, highlighting the efficacy of context similarity in mitigating sparsity issues and improving recommendation accuracy. By summarizing the findings of this empirical comparison, the paper offers practical guidance for selecting and implementing context-aware recommendation algorithms in real-world scenarios. In summary, this research paper contributes to the advancement of recommender systems by introducing context-aware collaborative filtering techniques and evaluating their performance through empirical analysis. By incorporating the methodologies and insights outlined in this paper, my project aims to leverage context similarity to enhance the personalization and adaptability of recommendation systems, ultimately improving the user experience and mitigating information overload.

Douglas O'Shaughnessy [3] The paper discusses the design of automatic speech recognition (ASR) systems to effectively utilize the discriminative information encoded in human speech. It contrasts with recent machine learning approaches that apply general recognition architectures to signals without considering the nature of the input. The paper examines the unique characteristics of speech as a complex signal driven by quasi-periodic and/or noisy sources, aiming to communicate diverse information using human vocal systems. It explores the extraction of pertinent features from speech and the reliable identification of different units of oral language. Additionally, the paper reviews the history of ASR attempts, successful methods, and advancements leveraging data availability and computational power, particularly focusing on deep neural networks. The paper concludes by suggesting ways to enhance ASR accuracy and efficiency.

Amarildo Rista et al. [4] I This research paper presents a systematic literature review aimed at providing a comprehensive understanding of various techniques within the domain of Speech Recognition. The abstract highlights the importance of speech recognition in facilitating human-machine communication and its widespread applications across different fields. It acknowledges the challenges inherent in speech recognition due to the complexity and variability of natural languages. The paper aims to introduce significant and relevant techniques to guide future research in this area.

Ali Bou Nassif et al. [5] The introduction provides a foundational understanding of deep learning, elucidating its core principles and architectural components. It highlights the evolution of deep learning algorithms, which have enabled more efficient feature extraction and representation learning from raw data. Specifically, the authors underscore the superiority of DNNs over traditional met hods, such as Gaussian Mixture Models (GMMs) and Hidden Markov Models (HMMs), in handling complex, non-linear relationships inherent in speech signals. The paper then transitions to discussing the historical progression and recent

developments in deep learning-based speech recognition systems. It outlines notable studies and benchmarks that demonstrate the superior performance of DNN-based models in various speech-related tasks, including automatic speech recognition, emotional speech recognition, speaker identification, and speech enhancement. Moreover, it discusses the challenges and limitations associated with traditional approaches and how DNNs have addressed these issues, particularly in handling temporal dependencies and modeling complex acoustic features. The methodology section elucidates the systematic literature review process undertaken by the authors, following established guidelines for comprehensive research synthesis. Through meticulous data collection and analysis, the authors identify and categorize relevant research papers published over the specified time frame. They extract key information from these papers, including types of speech data, training databases, language variations, environmental conditions, feature extraction techniques, and publication sources. Subsequently, the paper presents the findings of the systematic review, offering insights into the trends and advancements observed in DNN-based speech recognition research over the years. It synthesizes the extracted data into statistical representations, showcasing the distribution of research topics, methodologies, and emerging research directions. Additionally, it identifies research gaps and areas for future exploration, aiming to guide and inspire future researchers in the field.

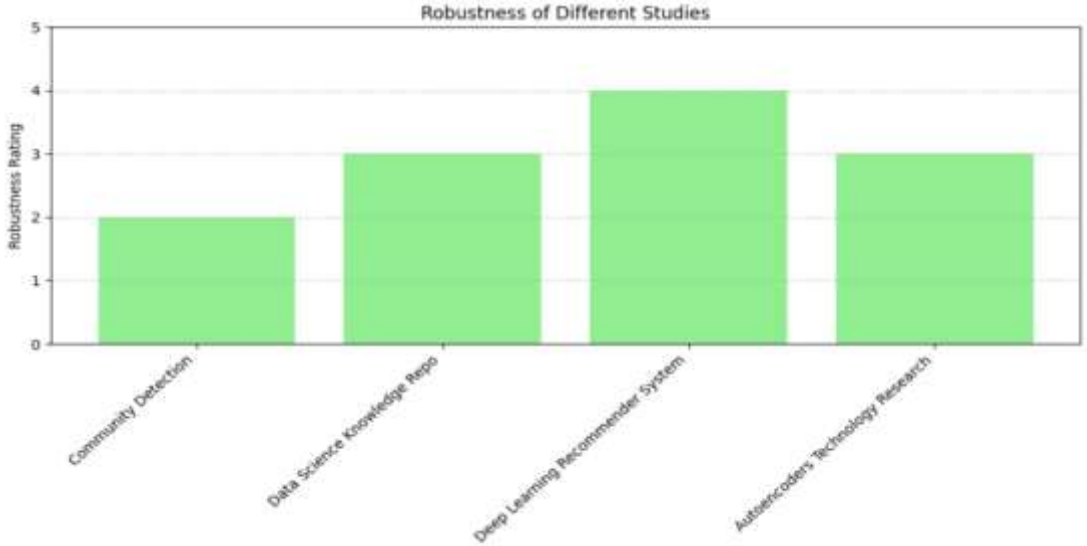
III. METHODOLOGY

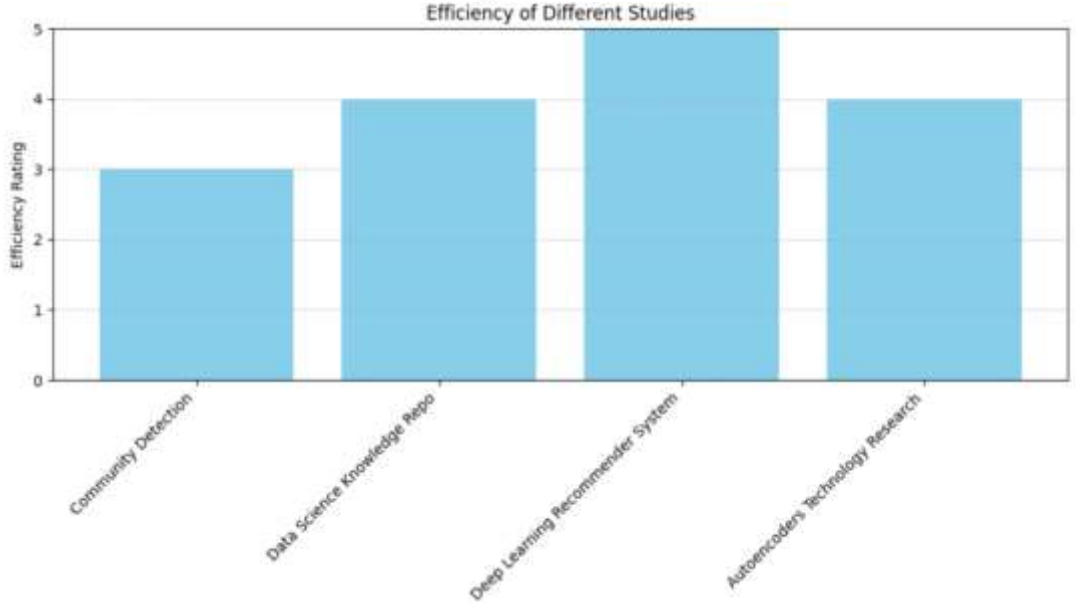
Techniques/Methodology	Pros	Cons
1. Fairness-aware Differentially Private Collaborative Filtering	• **Privacy Protection:** The paper introduces a novel approach, DP-Fair, which combines differential privacy mechanisms with fairness constraints to safeguard user privacy while ensuring fair recommendations • **Superior Performance:** Empirical results demonstrate that the proposed DP-Fair method demonstrates superior performance in both overall accuracy and user satisfaction compared to vanilla DP-SGD. • **Real-world Application:** The experimental evaluation is based on datasets from Amazon and Etsy, two prominent e-commerce platforms. This indicates the applicability of the proposed method in real-world scenarios	• **Real-world Application:** The experimental evaluation is based on datasets from Amazon and Etsy, two prominent e-commerce platforms. This indicates the applicability of the proposed method in real-world scenarios • **Algorithmic Fairness Trade-offs:** While DP-Fair aims to address fairness concerns, there may still be trade-offs between privacy protection, fairness, and recommendation accuracy.
2. Context-Aware Collaborative Filtering Using Context Similarity: An Empirical Comparison	• **Context-Aware Recommendations:** The paper introduces context-aware recommender systems (CARS), which adapt recommendations to users' preferences in different contextual situations which involves incorporating contextual information to enhance recommendation accuracy and relevance. • **Addressing Information Overload:** CARS can alleviate information overload by delivering tailored recommendations based on contextual factors such as	• **Sparsity Issue**: Context-aware collaborative filtering suffers from the sparsity problem, where users may not rate items in various context situations. This limitation can impact the effectiveness of recommendation models, especially when relying on context similarity to alleviate sparsity. • **Complexity in Modelling:** Building recommendation models based on context similarity requires measuring the similarity of contexts and selecting

	companion, occasion, or location. This capability is valuable in scenarios where users face overwhelming choices and need personalized guidance. • **Empirical Comparison:** The paper provides an empirical comparison of context-aware collaborative filtering methods using context similarity. This comparison offers insights into the effectiveness of different approaches and can guide the selection of methods for implementation.	appropriate rating profiles for model construction. This process may introduce complexity in model development and interpretation • **Generalizability of Findings:** The empirical comparison is based on multiple context-aware datasets, but the scope may not cover all possible contextual scenarios. As a result, the findings may have limited generalizability to diverse contexts or domains, requiring further validation in other settings.
3. Speech Recognition Using Deep Neural Networks: Ali Bou Nassif, Ismail Shahin	• **High Accuracy:** DNNs can achieve significantly higher accuracy compared to traditional speech recognition methods like Hidden Markov Models (HMMs), especially in complex and noisy environments. • **Adaptability:** DNNs can adapt to diverse accents, dialects, and speaking styles thanks to their ability to learn from large amounts of training data. • **Robustness to Noise:** DNNs can extract relevant features from noisy audio signals, improving performance in real-world situations with background noise.	**Computational Cost:** Training and running DNN models can be computationally expensive, requiring significant hardware resources and potentially limiting their use in smaller systems. **Data Dependency:** DNN performance heavily relies on the quality and quantity of training data. Insufficient or biased data can lead to inaccurate or biased models. • **Privacy Concerns:** Collecting and storing large amounts of voice data raises privacy concerns that need to be addressed with transparent data practices and strong security measures.
4. Automatic Speech Recognition: A COMPREHENSIVE SURVEY	• **Comprehensive Coverage:** The paper covers a wide range of techniques and approaches in speech recognition, providing readers with a thorough understanding of the field. • **Clear Objective:** The objective of the paper, to guide future research in speech recognition, is clearly stated, helping readers understand the purpose and significance of the study. • **Interdisciplinary Approach:** By addressing speech recognition as an interdisciplinary subfield of natural language processing, the paper acknowledges its broader implications and applications, appealing to a diverse audience. • **Systematic Review**: The paper follows a systematic literature review procedure, enhancing the credibility	• **Limited Scope:** While the paper aims to introduce significant techniques, it acknowledges that it cannot cover all relevant topics due to the extensive nature of the subject. This may leave some readers wanting more detailed information on specific techniques or approaches. •**Lack of Novelty:** The abstract does not explicitly mention any novel findings or contributions, which may limit its appeal to readers looking for groundbreaking research in the field. **Assumption of Prior Knowledge:** The paper assumes some level of prior knowledge about speech recognition and its components, which may make it less accessible to readers who are new to the topic.

e-ISSN: 2582-5208
International Research Journal of Modernization in Engineering Technology and Science
(Peer-Reviewed, Open Access, Fully Refereed International Journal)
Volume:06/Issue:04/April-2024 Impact Factor- 7.868 www.irjmets.com

	and reliability of its findings.	
	• **Comprehensive Overview:** The paper provides a comprehensive overview of the challenges and advancements in automatic speech recognition, making it suitable for both newcomers and experts in the field. • **Insightful Analysis**: It offers insightful analysis into the unique properties of speech signals and how ASR systems can effectively leverage them to improve recognition accuracy and efficiency. • **Relevant References:** The paper references relevant literature and research, allowing readers to explore further details and delve deeper into specific topics related to ASR.	• **Lack of Mathematical Detail:** While the paper emphasizes explanations over mathematical details, some readers may find it lacking in-depth technical insights into the algorithms and methodologies discussed. **Limited Scope:** The paper primarily focuses on acoustic aspects of speech-to-text (STT), with brief discussions on related issues such as language modeling and rescoring. Readers seeking a broader understanding of ASR applications beyond STT may find the scope limited.





IV. CONCLUSION

In conclusion, our project underscores several key aspects and challenges addressed within the domain of speech recognition utilizing deep neural networks (DNNs). Through our systematic review, we have identified significant advancements and trends in this rapidly evolving field, shedding light on both the progress made and the remaining challenges. One of the primary challenges tackled by our project is the need for more robust and efficient speech recognition systems. Traditional approaches like Gaussian Mixture Models (GMMs) and Hidden Markov Models (HMMs), have limitations in modeling complex, non-linear relationships inherent in speech signals. DNN-based models offer a promising solution by leveraging deep learning techniques to extract meaningful features and capture temporal dependencies more effectively.

ACKNOWLEDGEMENT

We would like to thank our guide Mrs. S. Satya Sudha, Assistant Professor, Department of CSE (Artificial Intelligence & Machine Learning), ACE Engineering College .Also, we are extremely grateful to her for her continuous support and invaluable time.

V. REFERENCES

[1] Amarildo Rista, Dr. Arbana Kadriu, Automatic Speech Recognition: A Comprehensive Survey, Faculty of Contemporary Sciences and Technologies, 2020.

[2] Ali Bou Nassif, Ismail Shahin, Imtinan Attili, Mohammad Azzeh And Khaled Shaalan, Speech Recognition Using Deep Neural Networks: A Systematic Review, 2019.

[3] Douglas O'Shaughnessy, Trends and developments in automatic speech recognition research, 2021

[4] Yong Zheng, Context-Aware Collaborative Filtering Using Context Similarity: An Empirical Comparison, 2022

[5] Zhenhuan Yang, YingqiangGe, CongzheSu, DingxianWang, XiaotingZhao, YimingYing, Fairness-aware Differentially Private Collaborative Filtering, 2023