# Chat-Bot Assisted Recommendation Platform for Computers and Repair Services

Shanila Thirimanne
*Department of Data Science*
*Sri Lanka Institute of Information Technology*
Malabe, Sri Lanka
shanilathiri@gmail.com

Wishwa Rathnaweera
*Department of Data Science*
*Sri Lanka Institute of Information Technology*
Malabe, Sri Lanka
wishwa5020@gmail.com

Amanullah Unais
*Department of Data Science*
*Sri Lanka Institute of Information Technology*
Malabe, Sri Lanka
unaisaman@gmail.com

Minindu Senadeera
*Department of Data Science*
*Sri Lanka Institute of Information Technology*
Malabe, Sri Lanka
minindusenadheera@gmail.com

Lakmini Abeywardhana
*Department of Data Science*
*Sri Lanka Institute of Information Technology*
Malabe, Sri Lanka
lakmini.d@sliit.lk

***Abstract*—In the age of computers, individuals used to look for computers through every avenue or ask for such assistance to find the ideal equipment based on the necessity. It is therefore important to have all the information under one roof. Hence, this novel approach have proposed a chat-bot recommendation system for computers, accessories and repair centers which provides those information at ease. This particular method uses Large Language Models to understand the user's preferences and requirements through a user friendly Chat - Bot interface. Furthermore, this platform utilized a state of the art ASR model to gather audio inputs. This study tends to give the best solutions towards users by collecting reviews of devices and repair centers by web-scraping certain websites to gather relevant data. In addition, image processing has been introduced in order to identify computer accessories without any difficulties. All these features will be available through the proposed chat bot system such that any individual with varying knowledge of electronic devices can get a solution without any hesitation.**

***Index Terms*—Chat - Bot, Image Processing, Recommendations, Automatic Speech Recognition**

## I. INTRODUCTION

In this fast-moving digitized world, having electronic devices has become a must in life. In such times, it is best to have the most suitable devices to cater your needs. Hence, in traditional days we tend to ask for such help or search for devices in all possible ways. Thus, it is best if we could have all in one place for us to identify the best device we need upon the requirement. Thus, this particular approach represent a recommendation system embedded into a Chat-Bot which enhances the human interactions and ease of use.

## II. LITERATURE SURVEY

In order to identify our primary motivation, is is important to understand what are the prevailing methods used to find a device, a computer accessory or a repair center. Ordinarily, main procedure is to search on web browsers for the items in need. Searching for brands, specific requirements and browse until desired output is gained. But in most cases, it is a lengthy process. Looking into different brands, device specs and customer reviews and so on. Else, another way is to ask someone else for help, a technician per say. Not always it's under the same roof. Thus, it would be beneficial for all of us if all were under the same roof.

Thus, our novel study introduces a novel approach of recommending devices according to their specific requirements and considering the emotions of the customer. Furthermore, identifying computer accessories using image processing and using online reviews, video reviews for recommendation purposes as well. In the realm of computer hardware identification, previous works have mainly focused on object detection and classification techniques. However, these approaches often lack the ability to accurately identify fine-grained components and associate them with specific specifications. The proposed research incorporates advanced image processing techniques to precisely identify and categorize various computer hardware components, enabling a more accurate and detailed understanding of a user's setup. Last but not least, all of this is introduced to the users via a Chat-Bot interface which is more user - friendly. This helps to retain the human nature whilst serving the need and fulfilling the motivation in the fast moving digitized world.

## III. Research/Market Gap

Throughout the history, there have been situations where Chat-Bot based E-commerce systems were implemented for various purposes. These were implemented to cater various needs and to make communication easier as well. Let us identify how this study stands out from the rest with the aid of the following table.

TABLE I
Overall Research Gap Comparison

| Research | Systems Used for Recommendation | | | | |
|---|---|---|---|---|---|
| | User reviews | Speech recognition | Video reviews | Image recognition | Chat Bot |
| [1] | ✗ | ✗ | ✗ | ✗ | ✓ |
| [2] | ✗ | ✗ | ✗ | ✗ | ✓ |
| [3] | ✓ | ✗ | ✗ | ✗ | ✗ |
| [4] | ✗ | ✓ | ✗ | ✗ | ✗ |
| Proposed | ✓ | ✓ | ✓ | ✓ | ✓ |

Such an example is “Design and Implementation of a chat bot for e-commerce” [1]. This system was mainly used for marketing purposes and make conversation faster. Thus, the proposed system also caters this need but in contrast, we will be specifying towards computers and the Sri Lankan market. Thus, an non - technical person will be able to find a device which cater their specific need. In addition, we will be looking towards, online reviews and providing customers with the most suitable devices. Another such example is “Development of an E-Commerce Chat bot for a University Shopping Mall” [2]. This system seeks to provide an easy, smart, and comfortable shopping experience for the Covenant University Community. Whilst our proposed system will also cater this need. This proposed approach aims to find users nearest and best service repair centers to cater their need. In contrast, this system will read reviews of other customers through online reviews and give the best recommendation towards the users. “A hierarchical recommendation system for online user reviews in e-commerce” is the subject of research [3]. Approach primarily focuses on e-commerce platforms where a system uses online reviews to propose products to users based on their greatest user insights was created in a hierarchical manner, with many levels filtering applied to the final recommendation, and it made use of ML algorithms to examine the users’ prior browsing and purchasing history. In this particular research, “Machine Learning Based Card-less ATM Using Voice Recognition Techniques” [4] uses speech to text in a Banking ATM system. This is a very useful feature since in this proposed system, uses speech to text in order to gather useful keywords and data from online reviews. Thus, making the recommendation system even more capable.

## IV. Methodology

### A. Chat - Bot Component and User Interface

The chat bot is the initial interface where users interacts with the proposed system as previewed in figure 1. Thus, this should seamlessly work with a user friendly interface. What should be noted is that in this modernized world, there are vast amounts of chat bots yet most of them fail to deliver the human nature of a conversation. Main reason for this is due to rule based chat bots. For example in the particular Research ( [5] Sinhala Chat bot with Recommendation System for Sri Lankan Traditional) it is a rule based chat bot. By doing so, the human nature of the conversation is drastically eliminated. Thus, in order to resolve this problem, novel approaches such as Large Language Models are incorporated with prompt engineering to build the chat bot.

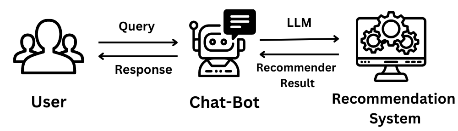


Fig. 1. Chat Bot Architecture

Large language models are advanced artificial intelligence models that have been trained on massive amounts of text data. These models are designed to generate human-like text based on the input they receive. They use a type of deep learning architecture called a transformer, which allows them to understand and generate coherent and contextually relevant text. [6]

Yet, the most crucial part of building our custom chat bot is carefully prompting to get the desired output. This involves providing specific instructions, context, or cues to guide the model’s output. Effective prompt engineering can significantly influence the quality and relevance of the generated responses.

The evaluation of chat bots is a complex process that can be biased towards qualitative analysis rather than quantitative analysis. One of the key methods of evaluating chat bots is human evaluation, where annotators review and rate the quality of chat bot responses based on criteria such as relevance, coherence, and appropriateness. However, this method can be subjective and unreliable, as human annotators may be more likely to give higher ratings to chat bot responses that are similar to their own style of communication.

A defining strength of the methodology lies in its seamless integration into a chat bot interface. By capitalizing on the efficiency of dot product similarity calculation, this system delivers near real-time identification results. Upon image processing, the chat bot seamlessly presents the identified hardware component to the user, ensuring swift and contextually relevant assistance.

### B. Accent-Adaptive Automatic Speech Recognition Model

*1) Dataset :* This study revolves around the development of an Automated Speech Recognition (ASR) model, specifically designed for seamless integration into a chat bot ecosystem. The primary objective is to achieve accurate transcription of user speech, with a special focus on accommodating the distinct characteristics of the Sri Lankan accent. The data set chosen for training and validation purposes is the ”Common Accent” data set [7], a robust collection of audio samples encompassing various South Asian English accents. This data set comprises 100,000 training samples and 450 test samples.

*2) Data Pre Processing:* During the pre processing phase, careful consideration was given to the audio sampling rate. To ensure alignment with the frequency components of human speech, a sampling rate of 8000 Hz was selected. This choice was not only rooted in the characteristics of speech but also influenced by resource constraints and computational efficiency.

To extract meaningful features from the audio data, the log-Mel spectrogram technique was employed [8]. This technique was particularly suitable for objectives as it aids in the identification of distinct vowel sounds. The Whisper feature extractor, provided by the Hugging Face library, played a pivotal role in converting raw audio data into standardized input features suitable for the subsequent stages of the model.

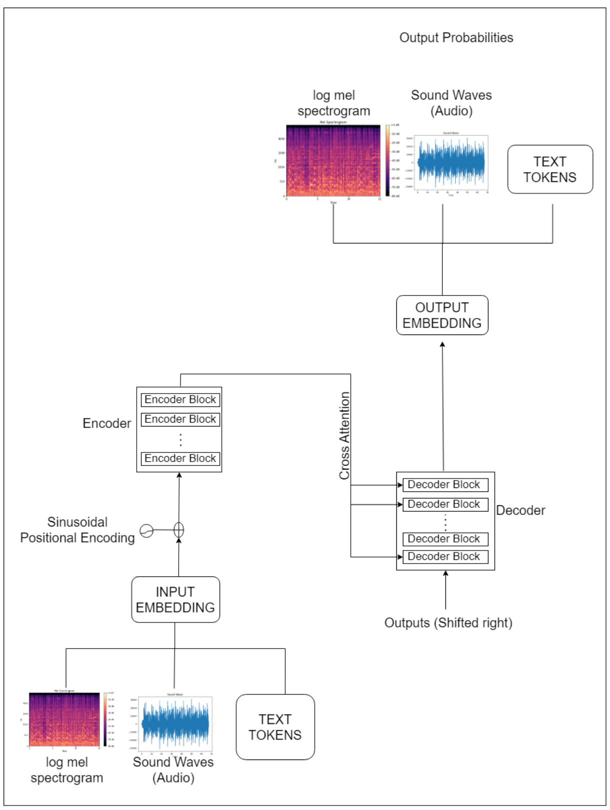


Fig. 2. ASR Model Architecture

*3) Model Architecture:* Figure 2 illustrates the model architecture in more detail, providing a visual representation of its components and their interactions.The model architecture adopted for this ASR system is founded on the Sequence-to-Sequence (Seq2Seq) paradigm, augmented with an encoder-decoder structure, and enriched by cross-attention mechanisms. To achieve optimal performance aligned with the proposed model accent-specific goals, fine-tuning of the Whisper model was carried out [9]. The utilization of Transformers proved invaluable in segmenting lengthy audio data into manageable chunks, thereby streamlining processing while maintaining contextual coherence.

A critical aspect of the ASR pipeline involved tokenization, which was facilitated by the Whisper Tokenizer. This process not only enabled the mapping of audio features to their corresponding textual representations but also facilitated the subsequent text-based processing steps. The data collector component was instrumental in preparing preprocessed data in a format compatible with the PyTorch tensors utilized by the model.

*4) Evaluation Metrics:* In assessing the effectiveness of the ASR model, the Word Error Rate (WER) metric was employed [10]. The WER metric comprehensively considers substitution, insertion, and deletion errors, offering valuable insights into the accuracy of the transcription process, particularly at the word level.

## *C. Advanced image processing for PC hardware identification*

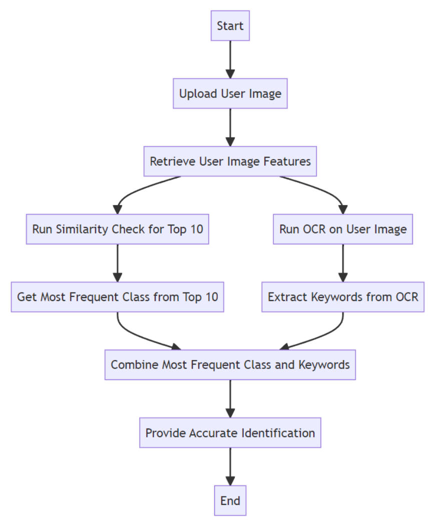


Fig. 3. Overall Component Diagram

This section elucidates the novel methodology employed to achieve rapid and accurate identification of PC hardware components from images provided by users. As shown in Fig. 3, by integrating advanced techniques such as feature extraction, similarity assessment, Optical Character Recognition (OCR), keyword analysis, and utilization of the dot product for near real-time processing, a robust system has been developed, capable of delivering swift and precise recognition of hardware components.

*1) Feature Extraction:* The methodology commences by leveraging the "VGG16" model [11], a well-known Convolutional Neural Network architecture. The extraction of features from the "block5 pool" layer of "VGG16" is prioritized to enhance the capture of intricate features more effectively. This strategic selection facilitates the acquisition of high-level patterns and representations, establishing a robust basis for subsequent analyses.

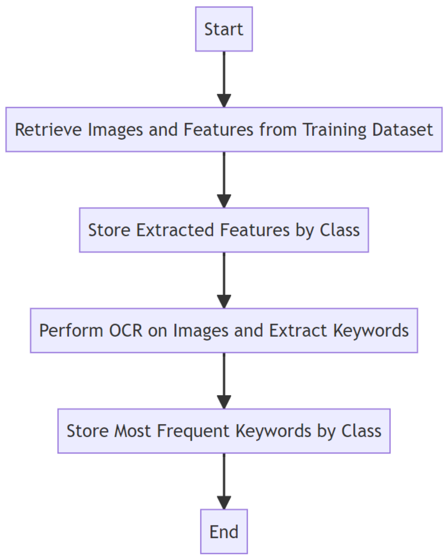


Fig. 4. Feature Extraction and Keyword Extraction

*2) Dot Product for Efficient Similarity Assessment:* A crucial enhancement to the approach involves the adoption of the dot product between feature vectors to assess similarity [12]. This computation exhibits considerably higher speed in contrast to cosine similarity, thereby facilitating near real-time processing capabilities. Through efficient measurement of alignment among feature vectors, the degree of resemblance between the user-uploaded image and the pre-existing dataset is rapidly ascertained.

*3) Optical Character Recognition (OCR) for Text Extraction :* Recognizing the presence of textual information within images, Optical Character Recognition (OCR) technology is harnessed. This process of converting visual text into machine-readable format empowers the system to efficiently extract textual content, thereby enhancing information retrieval capabilities.

*4) Keyword Analysis and Frequency Evaluation:* As illustrated in Fig. 4, from the text extracted through OCR, relevant keywords associated with PC hardware components are systematically identified and extracted. These keywords offer valuable insights into the nature of the depicted hardware component. Furthermore, the analysis extends to evaluating the frequency of these keywords across the data set.

*5) Informed Class Identification through Efficient Processing :* The integration of dot product similarity assessment, keyword analysis, and frequency evaluation culminates in an informed class identification process. This process swiftly selects the most frequent class among the top predicted classes based on dot product similarities and combines this information with identified keywords.

### *D. Device Recommendation and Repair Center Suggestions Model*

relevant user inputs, including information about the user's type, location, and relevant keywords obtained via the use of NLP, speech recognition, or image processing algorithms. The laptops in the data set will be rigorously divided into different user typologies based on pre-established criteria depending on the intended purpose of use during the pre processing stage. [13]

Whenever a user expresses a need for the purchase of a laptop, a combined framework made up of collaborative and content-based recommendation models will be used. This involves creating a TFIDF vector based on the identified keywords and comparing it to the entire data set of computers. The laptop with the highest cosine value will be identified after computing the cosine similarity between the user's query and the data set. This system will then provide recommendations for laptops with similar specs, supported by a content-based methodology. In addition, the collaborative-based recommender mechanism must be used to provide recommendations that are in harmony with those that users of similar typologies have endorsed. Through the user interface, the combined recommendations resulting from the content-based and collaborative-based paradigms will be channelled. [14]

As an alternative, the system will provide a list of repair facilities within a 10-kilometre radius of the user's current location if the user's required requirement requires the identification of a repair centre. The evaluation of Google reviews will be used to construct this recommendation, allowing users to effortlessly identify better places and the users will be able to navigate to the chosen repair centre easily with navigational applications of their choice directly from the chat bot interface.

Laptop recommendation model and the computer part identification model it will returns an output of laptops and computer parts including variety of elements, such as cost comparisons, embedded hyperlinks that facilitate optimal pricing, technical specifications, customer reviews and purchase options across various retail channels.

The system must initiate a method wherein the cosine similarity between the specified keywords and the corpus of computer parts is evaluated when the user's expressed need is for the acquisition of particular computer part. The result will be a list of well-chosen outcomes that indicate compatibility and relevance. And for the repair centre recommendation it will return an output that includes contact numbers to repair centres, their locations, and rating for them computed through reviews.

## V. Results And Discussion

### A. Chat - Bot Component and User Interface

In the evaluation process, A/B testing was used. A/B testing involves testing different versions of a chat bot to see which one performs better. This can be done by changing the prompts, the conversational flow, or the features of the chat bot. A/B testing can be a powerful tool for improving the performance of chat bots, as it allows us to test different hypotheses and see which one leads to the best results. A/B testing stands a high importance since prompting the chat bot in different ways yield different results. Thus, prompting accordingly to get the desired results is our objective.

As for this study several prompting strategies was used to yield desired output. In summary three trials of prompting was done to yield the desired output.

In the initial strategy, the chat bot was capable of gathering information about laptop specifications despite of the users' personal preferences. Since the goal of this novel approach is the capture a vast amount of customer base, users with or without a particular knowledge regarding laptops should be able to find their desired device. Thus, it is important to capture users' personal preferences, their job roles and such information. Since, the initial prompting strategy wasn't capable of capturing such information, it was decided to change the prompting strategy.

In the next phase, the chat bot was able to capture the relevant information according to a different prompting strategy. Yet, the chat bot wasn't able to clearly distinguish the user whether the user has a IT knowledge base regarding devices or not. Thus, this strategy fails since this short list our user base.

Hence, as the final strategy, prompting was done to initially distinguish such users IT knowledge base about devices. Thus, if they do not, the chat bot assess such users' further personal preferences and desires to use devices and then recommend a device accordingly. On the other hand, if the user does have a knowledge of devices, then the chat bot would assess for device specifications specifically to recommend the best device. Thus, in this particular case A/B testing incorporated with evaluation was used to assess the chat bot. This particular method can also be mentioned as 'Iterative testing and refining', as mentioned in 'Prompt Engineering: A Quick Guide To Techniques, Tips, And Best Practices'. [15]

As summary the following accuracy figures were obtained by passing a similar input to all three prompt strategies simultaneously as shown in the table II. The desired output is defined such that, the chat bot is capable of identifying whether the user has a IT knowledge base or not and gathering information, such as their personnel preferences, job role and last but not least specifications of devices.

Thus, from the above accuracy figures it was concluded that it is best to move on forward with the prompt Strategy No.3

TABLE II
Accuracy Figures on Prompt Strategies

| Prompt Trial No. | Delivering the Desired Output | Accuracy |
|---|---|---|
| Prompt 1 | 5/20 | 25% |
| Prompt 2 | 13/20 | 65% |
| Prompt 3 | 18/20 | 90% |

### B. Accent-Adaptive Automatic Speech Recognition Model

This section analyzes the results of the automatic speech recognition (ASR) experiment and discusses the implications of those results.

*1) Metrics for Performance :* As illustrated in the table III quantifiable metrics that reveal information about the accuracy and effectiveness of the ASR model were used to assess its performance. The loss value of 0.4234 shows the model's ability to reduce discrepancies between predicted and actual transcriptions during training.

The Word Error Rate (WER), which was calculated at 13.0605, proves how well the model can translate spoken language into written text. The percentage of words in the output of the transcription that differ from the actual text is measured by this metric.

Similarly, the model's capacity to precisely detect and correct word-level variations during transcription is demonstrated by the orthographic word error rate (WER Ortho) of 17.9229.

The WER is subtracted from 100 to get the Word Accuracy (WAC), which gives a value of 86.9395%. This value indicates how well the model can accurately translate about 86.94% of spoken language into text.

TABLE III
Automatic Speech Recognition Performance Metrics

| Epoch | Step | Training Loss | Validation Loss | WER Ortho | WER |
|---|---|---|---|---|---|
| 6 | 2500 | 0.006 | 0.423 | 17.923% | 13.061% |

*2) model for allocating resources and training :* A high-performance Collab A100 GPU was used for training the Whisper architecture, which forms the basis of the ASR model. About 27GB of GPU memory were used during the training phase, which lasted for about 3 hours and 34 minutes. The effective convergence of the model was made possible by the efficient acceleration provided by the GPU integration.

The model showed a tendency to over fit the training audio clips at a discernible point around the 3.42 epoch mark, according to experimentation observations. In response, the final model's weights were chosen and kept that corresponded to the best validation performance. The hugging face hub then received these improved weights for wider distribution.

*3) Accentuation and Model Generalization :* Pre-training and fine-tuning were both included in a two-stage process that increased the effectiveness of the ASR model. The Whisper model underwent subsequent fine-tuning using a dataset representative of South Asian accents after pre-training on a varied dataset. This method skillfully captured the distinctive nuances present in South Asian speech patterns while facilitating generalization across a wide range of accents.

The achieved WER of 13.0605 confirms the effectiveness of this methodology when placed within the complex phonological and intonational characteristics of South Asian accents. However, there are still opportunities for improvement and additional research. Through the incorporation of methods like data augmentation, accent-specific fine-tuning, and the incorporation of a larger and more complete training dataset, the model's performance may be improved.

### C. Advanced image processing for PC hardware identification

**Overall Accuracy:** The overall accuracy metric measures the proportion of correctly classified hardware components in the entire dataset. The formula for overall accuracy is as follows:

$$Overall\ Accuracy = \frac{Number\ of\ Correct\ Predictions}{Total\ Number\ of\ Predictions} \times 100\%$$

The research achieved an overall accuracy of 89.1%, showcasing the robustness of their methodology in accurately identifying hardware components across the entire dataset.

**Class-wise Accuracy:** Class-wise accuracy is a set of metrics that assess the model's performance for each individual hardware component class.

The class-wise accuracies ranged from 81.81% to 94.55%, highlighting the model's proficiency in distinguishing between different hardware component classes.

TABLE IV
IMAGE PROCESSING COMPONENT CLASS-WISE ACCURACY

| Class | Accuracy (%) |
|---|---|
| GPU | 87.27 |
| HDD | 94.55 |
| CPU | 89.09 |
| RAM | 81.81 |
| SSD | 92.73 |

## VI. CONCLUSION

In conclusion, the E-commerce Chat Bot Recommender System for laptops is a creative solution that makes use of cutting-edge technologies to completely transform the online shopping experience for Sri Lankan customers of laptops. The system customizes the shopping experience by integrating chat bot and image recognition capabilities, understanding individual preferences and requirements to offer the best results while reducing time and effort. It is inclusive for customers of various financial capacities thanks to the inclusion of features like web-scraping and repair center recommendations.

Additionally, the system's image processing methodology introduces a ground-breaking method for real-time hardware component identification that is enhanced with dot product similarity calculation. This feature-rich system seamlessly integrates text recognition, feature extraction, efficient similarity assessment, and chat bot interaction to deliver prompt and accurate support for hardware-related questions.

Changing the attention to the Whisper architecture-based ASR model, it shows promising results in converting spoken language to written text, especially when working with a variety of South Asian accents. Word error rate (WER) has significantly decreased as a result of the effective integration of pre-training and fine-tuning methodologies, which is a positive development. However, more research is still needed to improve accuracy, particularly in cases where there are complex linguistic structures and accents. Future directions for improvement could include investigating advanced architectural refinements, incorporating large and diverse datasets, and investigating novel training techniques.

Together, these developments demonstrate how technology can improve customer experiences and speed up decision-making processes.

## REFERENCES


[1] A. R. Asadi and R. Hemadi, *Design and Implementation of a chatbot for e-commerce*, 2018.

[2] A. Olomo and V. Oguntosin, "Development of an e-commerce chatbot for a university shopping mall," *Applied Computational Intelligence and Soft Computing*, 2021.

[3] I. Islek and S. G. Oguducu, "A hierarchical recommendation system for e-commerce using online user reviews," *Electronic Commerce Research and Applications*, vol. 52, 2022.

[4] J. B. Idoko, M. Mohammed, and A. U. Mohammed, "Machine learning based cardless atm using voice recognition techniques," in *Machine Learning and the Internet of Things in Education: Models and Applications*. Springer, 2023, pp. 75–84.

[5] W. Chandrasena, A. Gunesekera, and I. Uwanthika, *Sinhala Chatbot with Recommendation System for Sri Lankan Traditional Dancers*. KDU IRC, 2021.

[6] geeksforgeeks, "Prompt engineering for chatbot," Available at https://www.geeksforgeeks.org/prompt-engineering-for-chatbot/ (Accessed Jul. 30, 2023).

[7] Huggingface, "Dtu54dl common-accent," Available at https://huggingface.co/datasets/DTU54DL/common-accent (Accessed Jul. 30, 2023).

[8] A. Meghanani, A. C. S., and A. G. Ramakrishnan, "An exploration of log-mel spectrogram and mfcc features for alzheimer's dementia recognition from spontaneous speech," in *2021 IEEE Spoken Language Technology Workshop (SLT)*, 2021, pp. 670–677.

[9] A. Radford, J. W. Kim, T. Xu, G. Brockman, C. McLeavey, and I. Sutskever, *Robust speech recognition via large-scale weak supervision*, 2023.

[10] Wikikpedia, "Word error rate," Available at https://en.wikipedia.org/wiki/Word_error_rate (Accessed Jul. 30, 2023).

[11] S. Mascarenhas and M. Agarwal, "A comparison between vgg16, vgg19 and resnet50 architecture frameworks for image classification," in *2021 International Conference on Disruptive Technologies for Multi-Disciplinary Research and Applications (CENTCON)*, vol. 1, 2021.

[12] S. Gupta, "Building the gen ai applications and choosing the right similarity metrics," Available at https://www.linkedin.com/pulse/building-gen-ai-applications-choosing-right-similarity-sharad-gupta/ (Accessed Jul. 30, 2023).

[13] S. Ekin, *Prompt Engineering :A Quick Guide To Techniques, Tips, And Best Practices*, 2023.

[14] P. Govindarajan and S. S. Devi, "Hybrid recommendation system based on collaborative and content-based filtering," *Cybernetics and Systems*, vol. 54, 2021.

[15] Y. Hou and S. S. Yang, "An e-commerce personalized recommendation algorithm based on fuzzy clustering," *International Journal of Economics and Statistics*, vol. 10, 2022.