Customer Care Call Analytics Powered by ASR & LLM for Indian Languages

DISSERTATION

Submitted in partial fulfillment of the requirements of the

Degree: M.Tech in Artificial Intelligence & Machine Learning

By

Tanmay Jain
2022AA05306

Under the supervision of

Shakti P. Rath
(Principal Architect)

BIRLA INSTITUTE OF TECHNOLOGY AND SCIENCE
Pilani (Rajasthan) INDIA

(June,2024)

BIRLA INSTITUTE OF TECHNOLOGY & SCIENCE, PILANI
SECOND SEMESTER 2023-24

AIMLCZG628T DISSERTATION

Dissertation Title : Customer Care Call Analytics Powered by ASR & LLM for Indian Languages

Name of Supervisor : Shakti P. Rath

Name of Student : Tanmay Jain

ID No. of Student : 2022AA05306

Courses Relevant for the Project & Corresponding Semester :

             1. Deep Learning
             2. Advanced Deep Learning
             3. Natural Language Processing
             4. NLP Applications

Abstract

This dissertation focuses on developing a cost-effective and efficient Automatic Speech Recognition (ASR) system specifically designed for customer care centers serving diverse Indian languages. The primary aim is to address the challenges posed by low-resource languages, noisy environments, and significant variations in speech. This project utilizes custom-trained ASR models, open-source speaker diarization, and an open-source Large Language Model (LLM) to deliver comprehensive call analytics. All components are deployed on-premises, eliminating the need for any paid APIs.

The ASR system will be trained using both open-source and custom data to support low-resource languages like Hindi. This approach ensures higher accuracy and relevance in transcriptions tailored specifically to the domain, which is crucial for the operational needs of customer care centers. The ASR model is designed to utilize advanced architectures, such as transformers and conformers, ensuring the model can learn complex hidden patterns.

A significant aspect of the system is its integration of speaker diarization, which accurately identifies and separates different speakers within a call. This feature enhances the clarity of transcriptions and ensures that dialogues are correctly attributed, facilitating better understanding and analysis of customer interactions. The diarization process is critical in multi-speaker environments typical of customer care settings, where distinguishing between customer and agent speech is essential for accurate data capture and subsequent analysis.

To further enhance the utility of the ASR outputs, a Large Language Model (LLM) is employed for call analytics. The LLM processes the transcriptions to generate actionable insights, such as identifying common customer issues, sentiment analysis, and performance

metrics for customer service representatives. This analytic capability is instrumental in improving service quality, training, and strategic decision-making within customer care operations.

By avoiding the use of paid APIs and utilizing open-source technologies, this dissertation aims to deliver a robust, scalable ASR solution that can be adopted by customer care centers without incurring significant expenses. The focus on on-premise deployment ensures compliance with data privacy regulations, a critical consideration for many organizations.

In summary, this dissertation presents a comprehensive solution for enhancing customer care centers through a tailored ASR system that addresses the operational challenges of Indian languages. By integrating speaker diarization and LLM-based analytics, the proposed system not only improves transcription accuracy but also provides valuable insights to drive service improvements, making it a pivotal tool for modern customer care operations in India.

BITS ID No. 2022AA05306   **Name of Student:** Tanmay Jain

Name of Supervisor: Shakti P. Rath

Designation of Supervisor: Principal Architect

Qualification and Experience: PhD from IIT Madras and 15+ years of experience in ASR

Official E- mail ID of Supervisor: shaktiprasadrath9@gmail.com

Topic of Dissertation: Customer Care Call Analytics Powered by ASR & LLM for Indian Languages

(Signature of Student)   (Signature of Supervisor)

Date: 07 June, 2024   Date: 07 June, 2024

Project Work Title
Customer Care Call Analytics Powered by ASR & LLM for Indian Languages

1.1. Purpose:
The primary purpose of this project is to develop a robust ASR (Automatic Speech Recognition) and LLM (Large Language Model) powered solution for call analytics in customer care centers catering to Indian languages. This system aims to improve the efficiency and effectiveness of customer support by accurately transcribing and analyzing calls.

1.2. Expected Outcome:
- A comprehensive ASR system tailored to Indian languages.
- Enhanced call analytics through integration with LLMs.
- Improved customer service through accurate transcriptions and insights.
- Generate revenue by leveraging conversational insights to analyze customer feedback, improve product design, and address pain points effectively.
- Cost-effective and on-premise solution ensuring data privacy.

2. Literature Review:
[1] Making a Case for Speech Analytics to Improve Customer Service Quality: Vision, Implementation, and Evaluation
  - This paper presents a strategic initiative that examines the use of speech analytics to improve customer service quality at call centers. It discusses the vision, implementation, and evaluation of speech analytics in enhancing customer service quality.

[2] The Impact of Call Center Employees' Customer Orientation Behaviors on Service Quality
  - This study investigates the impact of call center employees' customer orientation behaviors on service quality. It highlights the importance of speech analytics in understanding customer interactions and improving service quality.

[3] Text Summarization for Call Center Transcripts
  - Text summarization of call center transcripts is essential for detailed analysis but is challenging due to context switching, cross talk, and transcription errors from ASR systems. This work develops a summarization model for on-premise deployment at call centers by fine-tuning pre-trained open-source large language models, using reference summaries generated by GPT-3.

[4] Leveraging AI via speech-to-text and LLM integration for improved healthcare decision-making in primary care
- Recent research highlights the potential of AI, particularly ASR and large language models, to alleviate workloads and streamline documentation in healthcare. A proposed framework for general practices includes ASR for transcriptions, decision support systems, and automated prescription email generation. Qualitative and quantitative analyses demonstrate improvements in reducing administrative burdens and enhancing medic-patient relationships, potentially improving diagnostic outcomes. This approach can be adapted for

customer care call analysis to automate call summaries, improve documentation accuracy, support real-time decision-making, and enhance customer interactions.

3.1. Existing Process:
Current customer care centers employ one of two methods:
1. A manual process where a designated person reviews a selection of recordings from customer care agents by listening them manually, providing feedback to both the agents and the company.
2. Manual transcription or commercial ASR services for call analysis, which are often costly, less accurate for Indian languages, and ineffective in noisy environments. These services typically rely on cloud-based solutions, raising concerns about data privacy and latency.

3.2. Limitations:
- The manual process prevents agents from reviewing all recordings, leading to inaccurate insights, improper feedback for both the agent and the company, potential bias, and high costs.
- Commercial ASR services are associated with high costs and limited accuracy in transcribing Indian languages and handling dialects/accents. They are also ineffective in noisy environments and pose data privacy issues due to their reliance on cloud-based solutions.

4. Justification for Methodology:
- Custom Data Training: Ensures the ASR system is finely tuned to Indian languages and their specific challenges.
- On-Premise Deployment: Addresses data privacy concerns and reduces operational costs.
- Noise Robustness: Essential for real-world customer care environments which are often noisy.
- Integration with LLMs: Provides advanced call analytics capabilities, enhancing the overall functionality of the system.

5. Project Work Methodology:
- Phase 1: Data Collection - Explore, gather, and annotate speech data in various Indian languages.
- Phase 2: ASR Model Development - Develop and train ASR models with noise robustness.
- Phase 3: Speaker Diarization - Explore open-source speaker diarization models/ APIs and integrate them for accurate speaker identification.
- Phase 4: LLM Integration - Explore open-source LLMs for analyzing transcriptions to extract insights and perform sentiment analysis.
- Phase 5: Evaluation and Optimization - Test the system in real-world conditions and optimize for accuracy and efficiency.
- Phase 6: Deployment - Deploy the system on-premises in customer care centers.

6. Benefits Derivable from the Work:
Benefits:
- Enhanced Accuracy: Significantly improved transcription accuracy for Indian languages.
- Cost Efficiency: Drastic reduction in costs by eliminating the need for manual review and avoiding expensive commercial ASR services.
- Real-Time Analysis: Capability to conduct real-time call analysis and deliver immediate insights.

- Data Privacy: Guaranteed through on-premise deployment, ensuring the confidentiality of sensitive information.
- Improved Customer Service: Enhanced customer experience through actionable insights and faster response times.

7. Additional Details:
- Scalability: The system is designed to scale across different customer care centers and can be adapted for additional languages and dialects as needed.
- User Training: Provision of training modules for customer care agents to effectively use and interpret the analytics provided by the system.
- Future Enhancements: Potential for future enhancements includes integrating with CRM systems for seamless customer interaction and adding multilingual support, as well as code-switching capabilities for broader applicability.

1. **Broad Area of Work:** Automatic Speech Recognition (ASR)

2. **Objectives**

The objectives of my project are as follows:

- Develop a robust ASR system optimized for Indian languages to accurately transcribe customer calls.
- Integrate speaker diarization to identify and differentiate between speakers during calls.
- Implement noise robustness techniques to enhance ASR performance in noisy environments.
- Incorporate a Large Language Model (LLM) for call analytics, extracting actionable insights from transcriptions.
- Deploy the ASR system on-premise to ensure data privacy and reduce operational costs.
- Provide a cost-effective and efficient solution for customer care centers catering to Indian languages.

3. Scope of Work

- Development of Custom ASR Models: Design and train ASR models tailored specifically for Indian languages.

- Integration of Speaker Diarization: Explore open-source speaker diarization models/ API's to accurately identify and differentiate between speakers during customer calls, enhancing the clarity and relevance of transcriptions.

- Noise Robustness Enhancement: Develop and integrate noise robustness techniques to improve ASR performance in noisy environments commonly encountered in customer care settings.

- Incorporation of Large Language Models (LLMs): Integrate LLMs into the ASR system for advanced call analytics, including sentiment analysis, topic extraction, and customer behavior insights.

- On-Premise Deployment: Deploy the ASR system on-premise within customer care centers to ensure data privacy and security, while also reducing reliance on external services and minimizing latency.

- Cost-Effectiveness and Efficiency: Focus on delivering a cost-effective and efficient solution that meets the operational needs of customer care centers catering to Indian languages, without compromising on performance or accuracy.

- Scalability and Adaptability: Design the ASR system to be scalable and adaptable, allowing for easy integration with existing infrastructure and potential expansion to support additional languages or dialects in the future.

- Evaluation and Optimization: Conduct thorough evaluation and optimization of the ASR system to ensure high accuracy, reliability, and usability in real-world customer care environments.

4. Detailed Plan of Work

Sno.	Tasks/Subtasks	Start Date - End Date	Planned Duration (weeks)	Specific Deliverable
1	Marker Research, Data Collection	Week 1 - Week 2	2	Market Research review report
2	Data Collection, & Processing	Week 3 - Week 4	2	Annotated speech corpus
3	ASR Model Development	Week 5 - Week 7	3	Initial ASR models
4	Speaker Diarization Implementation	Week 8 - Week 9	2	Integrated speaker diarization
5	Noise Robustness Enhancement	Week 10 - Week 11	2	Improved ASR performance in noisy environments
6	LLM Integration	Week 12 - Week 13	2	Integrated LLM for call analytics
7	On-Premise Deployment	Week 14 - Week 15	2	Deployed ASR system on-premise
8	Evaluation and Optimization	Week 16	1	Evaluated and optimized ASR system

5. Literature References:

The following are referred journals from the preliminary literature review.

[1]: Scheidt, S., & Chung, Q. B. (2019). Making a case for speech analytics to improve customer service quality: Vision, implementation, and evaluation. Library and Information Sciences, Computer Networks and Communications, Information Systems, 1–12. DOI: https://doi.org/10.1016/j.ijinfomgt.2018.01.002

[2]: Rafaeli, A., Ziklik, L., & Doucet, L. (2008). The Impact of Call Center Employees' Customer Orientation Behaviors on Service Quality. Journal of Service Research, 10(3), 239-255. https://doi.org/10.1177/1094670507306685

[3]: Ahmed, I., Zhou, Y., Sharma, N., Hosier, J. (2024). Text Summarization for Call Center Transcripts. In: Arai, K. (eds) Intelligent Systems and Applications. IntelliSys 2023. Lecture Notes in Networks and Systems, vol 822. Springer, Cham. https://doi.org/10.1007/978-3-031-47721-8_36

[4]: https://hdl.handle.net/10589/218053

Supervisor's Rating of the Technical Quality of this Dissertation Outline

EXCELLENT / GOOD / FAIR/ POOR (Please specify): EXCELLENT

Date: 07 June, 2024

(Signature of Supervisor)

Name of the supervisor: Shakti P. Rath

Email Id of Supervisor: shaktiprasadrath9@gmail.com

Mob # of supervisor: 9027696989