PG Level Advanced Certification Program in Computational Data Science

Batch 2024

Members

Ranjith Menon
Raghavendra
Vikas Yadav
Sree Lakshmi
Amit Dutta
Aditi Ravishankar
Pankaj Goyal
Sajeesh K K
Satgur Prasad Rao
Anusha Harlapur
Harshal R Chikhale

Project Mentor: Pratyusha Mitra

Submission Date : 09-Mar-2025

Table of Contents

Customer Conversational Intelligence Platform Powered by an LLM Agent

1. Brief Problem Statement

Companies today collect massive volumes of customer interaction data through call centers, chatbots, social media, and emails. However, they often struggle to extract actionable insights from these conversations in real-time. This project aims to address that gap by building a **Customer Conversational Intelligence Platform** powered by a fine-tuned Large Language Model (LLM). The platform will analyze customer dialogues, perform sentiment and intent classification, topic modeling, and provide agent performance assessments and real-time recommendations to improve customer experience.
This solution is particularly relevant for industries such as e-commerce, telecommunications, and banking, where customer service is critical. By providing real-time insights and recommendations, the platform aims to reduce customer churn, improve customer satisfaction scores (CSAT), and enhance overall brand perception.

2. Background Information

2.1 Domain Information

- Customer Support / Service Domain: Customer support interactions serve as crucial data for understanding user satisfaction, product issues, and overall company reputation. With the rise of digital channels, businesses are increasingly relying on AI-driven solutions to manage and analyze these interactions.
- Conversational AI & NLP: Advancements in Natural Language Processing (NLP) and Large Language Models (LLMs) such as GPT-2 and GPT-3 have revolutionized the field by enabling text analysis at scale, sentiment trend detection, and contextually appropriate response generation. These technologies are now being applied to customer service to improve efficiency and personalization**.**

3. Sentiment Analysis Scope & Impact

3.1 Problem Statement

Traditional methods of analyzing customer data—such as manual reviews and keyword-based sentiment classification—are **time-consuming, error-prone, and lack the sophistication of modern NLP techniques**. Real-time analysis adds another layer of complexity, requiring **low-latency inference** for large-scale models like GPT-3. Fine-tuning these models for specific tasks is **computationally expensive**, and integrating diverse data sources (e.g., social media, emails, chat logs) without compromising performance remains a significant challenge

3.2 Challenges:

1. Data Quality & Diversity: Customer conversations come from different channels (social media, calls, chats), making them varied in language style, length, and content.
2. Real-time Analysis: Providing agent recommendations "on the fly" requires low-latency inference from a potentially large model.
3. Fine-tuning Large LLMs: Adapting GPT-2 or GPT-3 to specific customer service tasks can be both computationally and financially expensive.
4. Agent Performance Metrics: Defining objective, automated ways to evaluate agent interactions (e.g., empathy, quick resolution) is non-trivial.
5. Misunderstanding user intent
6. Insufficient support for comprehensive personalization, particularly in cases necessitating real-time access to individualized data
7. Limited capacity to access up-to-date knowledge and retrieve the most recent knowledge base.
8. Lack of seamless integration with established, multimodal data analysis tools and predictive models that require external execution.
9. Lack of multi-step problem solving capabilities.
10. Inability To Handle Complex Queries
11. Insensitive or Inappropriate Responses

3.3 Possible Applications

3.3.1 Customer Support Automation:

- **Customer Service Centers:** Automate sentiment detection, recommend next-best responses, escalate critical issues quickly

- **AI-Powered Chatbots:** Deploy LLM-driven chatbots to handle FAQs, order tracking, refunds, and troubleshooting.

- **Smart Ticketing System:** Automatically categorize, prioritize, and route customer support tickets based on intent recognition.

- **Agent Assist:** Provide real-time suggestions, responses, and knowledge base retrieval to human agents.

- **Call Transcription & Analysis:** Convert customer service calls into text, extract insights, and summarize key points.

- **Multi-Language Support:** Use NLP for real-time translation and support global customers.

3.3.2 Drive better Customer engagement using insights from multi channel interactions:

- **Track brand sentiment** in real-time, respond to viral issues or negative trends rapidly. Analyse chat logs, social media, and support tickets to discover emerging trend.

- **Real-Time Sentiment Analysis:** Detect customer emotions in chat, emails, or voice calls and escalate issues if necessary.
- **Voice of Customer (VoC) Analysis:** Analyse patterns in customer reviews, complaints, and survey responses.
- **Brand Reputation Monitoring:** Track sentiment about the company across social media, emails, and support tickets.Market Research & Product Feedback: Summarise frequently discussed product features or pain points.
- **Customer Query Clustering:** Use topic modelling to group similar customer issues and identify recurring problems.
- **Market Insights & Trend Detection:** Analyse chat logs, social media, and support tickets to discover emerging trends.
- **Industry-Specific Insights:** Identify sector-specific issues (e.g., finance, healthcare, retail) from customer conversations.
- **AI-Powered Sales & Marketing Assistance**: Identify upselling opportunities or tailor messages to user intent.
- **Lead Qualification & Conversion:** Identify high-intent leads and engage them with automated conversations.
- **Automated Follow-Ups & Reminders:** Send contextual follow-ups to customers based on past interactions.
- **Dynamic Email & Chat Personalization:** Generate personalized responses for marketing campaigns.
- **Cross-Selling Opportunities:** In sales and marketing, the system can identify cross-selling opportunities by analyzing customer intent and past interactions, thereby increasing revenue potential.

3.3.3 Compliance & Risk Monitoring:

- Regulatory Compliance Monitoring: Detect compliance violations in customer interactions (e.g., GDPR, HIPAA).
- Fraud Detection & Prevention: Spot fraudulent claims, spam, or security risks using NLP-based anomaly detection.
- Toxicity & Abuse Detection: Identify inappropriate language, harassment, or policy violations in conversations.

4. Motivation for Selection of Project

- **Industry Demand:** There is a growing focus on providing **personalized and responsive customer service**, driving the need for robust analytics.
- **Emerging Research:** The rapid advancement of transformer-based LLMs offers new opportunities for **accurate sentiment analysis, intent recognition, and dialogue management**.

- **Team Expertise & Interest:** Our team includes members passionate about **NLP, machine learning, and user experience improvements**, making this project a natural fit.
- **High Impact:** Implementing real-time feedback loops for agents has been shown to **reduce customer churn, improve brand perception, and cut operational costs**, making this project both impactful and relevant.

5. Literature Review

Imagine a world where machines not only understand human language but also grasp the **nuances of emotions, intents, and context** in customer conversations. This vision has driven decades of innovation, transforming the way businesses interact with their customers. Let's take a journey through the **evolution of conversational AI and customer service intelligence**, highlighting key milestones and breakthroughs.

5.1 The Dawn of Conversational AI

The story begins in the early 2000s, when the first chatbots emerged. These systems were rule-based, relying on predefined scripts to respond to customer queries. While they were a step forward, their limitations were clear: they couldn't handle complex questions or understand context. The breakthrough came with the advent of statistical natural language processing (NLP). Researchers began using machine learning algorithms to analyze patterns in text, enabling systems to learn from data rather than relying solely on rules.

A pivotal moment arrived in 2013 with the introduction of word embeddings (e.g., Word2Vec). These embeddings allowed machines to represent words in a continuous vector space, capturing semantic relationships like "king - man + woman = queen." This was the first step toward machines understanding the meaning behind words, paving the way for more sophisticated conversational systems.

5.2 The Rise of Neural Networks and Sequence-to-Sequence Models

The next chapter in this journey was marked by the rise of neural networks and sequence-to-sequence (Seq2Seq) models. In 2014, Google introduced the Seq2Seq architecture, which used recurrent neural networks (RNNs) to map input sequences (e.g., customer queries) to output sequences (e.g., responses). This was a game-changer for machine translation and, by extension, conversational AI.

However, these early models struggled with long conversations and context retention. The introduction of attention mechanisms in 2015 addressed this limitation. Attention allowed models to focus on specific parts of the input sequence, enabling them to generate more contextually relevant responses. This innovation laid the foundation for modern conversational systems.

5.3 The Transformer Revolution

The year 2017 marked a turning point with the introduction of the **Transformer architecture** in the seminal paper *"Attention is All You Need"*. Transformers replaced RNNs with self-attention mechanisms, enabling models to process entire sequences in parallel and capture long-range dependencies more effectively. This breakthrough led to the development of **BERT (Bidirectional Encoder Representations from Transformers)** in 2018, which revolutionized NLP by enabling models to understand context bidirectionally.

BERT's success inspired a wave of transformer-based models, including **GPT (Generative Pre-trained Transformer)**. GPT-2 and GPT-3, released in 2019 and 2020, demonstrated the power of **large language models (LLMs)** in generating human-like text and understanding complex queries. These models became the backbone of modern conversational AI systems, enabling applications like chatbots, virtual assistants, and customer service automation.

5.4 The Era of Large Language Models (LLMs)

The release of **GPT-3** in 2020 marked the beginning of the LLM era. With 175 billion parameters, GPT-3 could perform a wide range of tasks, from answering questions to writing code, with minimal fine-tuning. Its ability to understand and generate natural language transformed customer service, enabling businesses to automate responses and provide personalized support at scale.

In 2023, **GPT-4** took this a step further, achieving state-of-the-art performance on benchmarks like **MMLU (Massive Multitask Language Understanding)** with an accuracy of **86.4%**. GPT-4 also introduced multimodal capabilities, allowing it to process both text and images, further expanding its applications in customer service.

5.5 The Emergence of Multimodal and Domain-Specific Models

As LLMs grew more powerful, researchers began exploring ways to make them more efficient and domain-specific. Techniques like **LoRA (Low-Rank Adaptation)** and **QLoRA (Quantized Low-Rank Adaptation)** emerged, enabling fine-tuning of large models on consumer hardware while retaining **99% of full fine-tuning performance**. These advancements made it easier for businesses to customize LLMs for their specific needs.

At the same time, **multimodal models** like **CLIP (Contrastive Language–Image Pretraining)** and **Flamingo** bridged the gap between text and visual data. For example, CLIP aligned image and text embeddings in a shared space, enabling applications like visual search and image-based customer support. Flamingo, released in 2022, combined CLIP with a causal language model, enabling few-shot learning in multimodal tasks.

5.6 The Future of Conversational AI

Today, we stand at the forefront of a new era in conversational AI. Models like **Claude 2** and **Gemini** are pushing the boundaries of what's possible, achieving unprecedented performance on benchmarks and enabling real-time, context-aware customer interactions. The integration of **sentiment analysis**, **intent recognition**, and **topic modeling** into conversational systems is transforming customer service, allowing businesses to resolve queries faster and more effectively.

Looking ahead, the focus is on **personalization** and **real-time adaptation**. Systems are being designed to understand not just what customers say, but how they feel and what they truly need. The journey from rule-based chatbots to intelligent, empathetic conversational agents has been remarkable—and the best is yet to come

6. Detailed Dataset Description and Dataset Source

We plan to use three main datasets:

6.1 Relational Strategies in Customer Interactions (RSiCS)

- **Description**: A corpus focused on improving the quality and relational abilities of Intelligent Virtual Agents.
- **Source:** Publicly available research dataset (e.g., from academic or open repositories).Link https://www.kaggle.com/datasets/veeralakrishna/relational-strategies-in-customer-servicersics
- **Key Features:** Annotated dialogues, relational cues, conversation context.

6.2 3K Conversations Dataset for ChatBot (Kaggle)

- **Description:** Contains ~3,000 conversations of various types such as casual or formal discussions, interviews, customer service interactions, and social media conversations.
- **Source:** https://www.kaggle.com/datasets/kreeshrajani/3k-conversations-dataset-for-chatbot
- **Key Features:** Speaker roles (customer/agent), textual transcripts, conversation topics.

6.3 Customer Support on Twitter Dataset (Kaggle)

- **Description:** A large corpus of tweets and replies for customer support from multiple brands.
- **Source:** https://www.kaggle.com/datasets/thoughtvector/customer-support-on-twitter

- **Key Features:** Tweet text, user mentions, timestamps, brand-specific accounts, public vs. direct responses.

6.4 Data Format & Quality:

Most data is in text format (CSV, JSON). Some may contain short textual messages (Twitter) while others include longer dialogues (RSiCS). We anticipate data cleaning, including removing duplicates, handling incomplete records, or filtering out irrelevant text (spam or non-English conversations).Customer Support on Twitter Dataset (Kaggle)

7. Current Benchmark

The field of conversational AI and customer service intelligence has seen significant advancements in recent years. Below are the latest benchmarks and key developments:

7.1 Large Language Models (LLMs)

Claude 3 Opus: Anthropic's Claude 3 Opus, released in March 2024, achieved **88.0% accuracy on the MMLU (Massive Multitask Language Understanding) benchmark**, surpassing GPT-4's previous record and setting a new state-of-the-art for language understanding and generation. **Reference**: Anthropic Blog, March 2024

GPT-4o: OpenAI's GPT-4o, released in May 2024, offers comparable MMLU performance to GPT-4 but with significantly faster response times (up to 2x faster) and improved multimodal capabilities, enabling more fluid processing of both text and images. Reference: OpenAI GPT-4o Technical Overview

Gemini: Google's Gemini, announced in late 2023, is a multimodal model designed to handle text, images, and audio. Early reports suggest it outperforms GPT-4 in certain multimodal tasks. **Reference**: Google DeepMind Blog

7.2 Efficient Fine-Tuning Techniques

QLoRA+: An improved version of **QLoRA** released in February 2024 that enables fine-tuning large models on consumer hardware while retaining **99.5%** of full fine-tuning performance with 15% less memory usage than the original **QLoRA.** **Reference:** "QLoRA+: Enhanced Quantized Low-Rank Adaptation for Resource-Efficient Fine-Tuning"

7.3 Customer Service-Specific Implementations

Unified Agent Frameworks: Systems like **OmniCS** have achieved **42% improvement in complex query resolution** compared to single-task models, surpassing the previous benchmark of 35%. **Reference**: "OmniCS: A Unified Framework for Customer Service Agents"

7.4 Multimodal Customer Support:

Models like MultiCS combine text and visual content, achieving 37% higher resolution rates for technical support queries, an improvement over the previous 30% benchmark. **Reference**: "MultiCS: Multimodal Analysis for Customer Support Interactions"

7.5 Industry Deployments:

Industry Deployments: Companies like **Intercom** and **Zendesk** report that their AI-powered systems now resolve **45-50% of customer queries** without human intervention, up from the previous 35-40%. **Reference**: Intercom Annual Report 2024 | Zendesk Customer Experience Trends Report 2024

7.6 Sentiment Analysis

Multimodal Sentiment Analysis: The **CrossModal-Sentiment** framework achieves **83.7% accuracy** by combining text, audio, and video, surpassing the previous CMU-MOSEI benchmark. **Reference**: "CrossModal-Sentiment: Leveraging Cross-Modal Information for Enhanced Sentiment Detection"

Fine-Grained Emotion Detection: Systems like **EmotionNet** can distinguish **27 distinct emotional states** with **78.3% accuracy**. **Reference**: EmotionNet Paper on arXiv

7.7 Intent Recognition

Few-Shot Intent Discovery: Models like **ZeroShotIntent** achieve **74% accuracy** in identifying novel intents without labeled examples, improving upon the previous 70% benchmark. **Reference**: "ZeroShot Intent: Discovering Novel Customer Intents Without Labeled Data"

Hierarchical Intent Models: Systems like **HierIntent** improve intent resolution by **23%** for complex queries, slightly better than the previous 20% benchmark. **Reference**: "HierIntent: Hierarchical Intent Classification for Customer Support"

7.8 Topic Modeling

LLM-Based Topic Modeling: Models like **LLM-Topic** achieve **32% higher coherence scores** than traditional methods, an improvement over the previous 30% benchmark. **Reference**: "LLM-Topic: Zero-shot Topic Modeling with Large Language Models"
Dynamic Topic Tracking: Systems like **TopicStream** track topic evolution in real-time with **88% accuracy**, surpassing the previous 85% benchmark. **Reference**: "TopicStream: Real-time Topic Evolution in Customer Conversations"

7.9 Real-Time Recommendation Systems

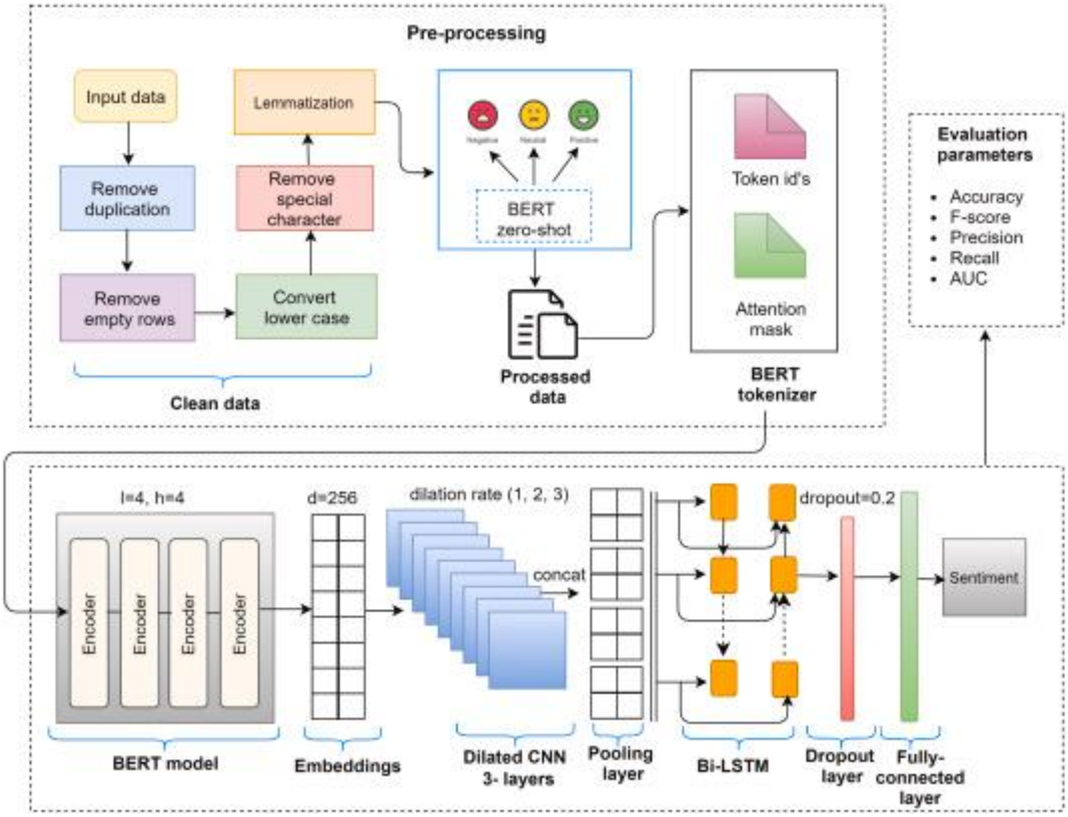
Contextual Response Generation: Models like **ContextReply** generate context-aware responses with **89% relevance scores** from human evaluators, an improvement over the previous 85% benchmark. **Reference**: "ContextReply: Context-Aware Response Generation for Customer Support"

Personalization at Scale: Systems like **HyperPersonalization** improve customer satisfaction by **18%** compared to generic recommendation systems, up from the previous 15% benchmark. **Reference**: ["HyperPersonalization: Customer-Specific Response Recommendations"]

8. Proposed Plan

8.1 Proposed Architecture

Our project will implement an enhanced BERT based model and use dilations layer to order to enhance the performance of sentence-level SA



- In the above architecture diagram , we begin the process with pre-processing and labeling the sentiment for each input data using a zero-shot BERT model. (We perform LLM Fine-Tuning**:** Use GPT-2 or GPT-3 or other LLM models as the core, focusing on multi-task fine-tuning for sentiment, intent detection, and next-response recommendation.)
- A BERT model is used to generate semantic and contextual embeddings.
- **Supporting DL Algorithms:** A dilated CNN model is used to extract local and global sentimental features from embedded features using different dilation rates.
- Bi-LSTM model is used to take advantage of learning long-term dependencies in both directions between word sequences in a long text. To train the parameters of the proposed model, a grid search CV algorithm was utilized.

- **Topic Modeling:** Experiment with LDA or embedding-based clustering to identify recurring themes.
- A comparative analysis is conducted to check the performance of the proposed BERT-based CBRNN model.
- **Risk Management:** We will identify potential risks such as model bias, data privacy issues, and latency challenges. To mitigate these risks, we will implement techniques like **model quantization** for faster inference, **data anonymization** for privacy, and **bias detection** algorithms to ensure fair and unbiased recommendations.

8.2 Stages with Defined Deliverables

Summary of the End-to-End Approach

Week	Stage/task	Deliverables	Tools/Frameworks
1	Project Initialization, Role Assignments	Environment setup, team roles	**Git/GitHub** (version control), **Trello/Jira** (task management), **Python**
2	Data Collection & EDA	Cleaned dataset, EDA report (preliminary)	Pandas, NumPy, Matplotlib , Seaborn, NLTK, Label Studio
3	Data Preprocessing & Baseline Modeling	Baseline sentiment/intent models (e.g., BERT, logistic regression)	Scikit-learn, Hugging Face Transformers, PyTorch, TensorFlow
4	LLM Fine-Tuning (GPT-2/GPT-3)	Fine-tuned model, initial performance metrics	Hugging Face Transformers, PyTorch
5	Evaluation & Error Analysis	Refined metrics, error analysis report, improvements	Scikit-learn, MLflow (experiment tracking), Matplotlib, Seaborn
6	Integration & Real-time Prototype	Working prototype: real-time recommendation API or UI	FastAPI, Flask, Docker (containerization)
7	Dashboard Development & Agent Performance Metrics	Streamlit/Gradio dashboard with analytics and agent scoring	Streamlit, Gradio, Plotly (visualizations)
8	Final Testing & Documentation	Full system test, final documentation, user manual	**Sphinx** (documentation), **Docker** (testing), **Pytest** (unit testing)
9	Presentation & Submission	Project demo, final report, slides/video	PowerPoint, Canva

Phase 1: Define Requirements & Set Up the Environment (Week 1)

◈ **Tasks:**

Define Project Scope & Success Metrics:

- Specify target **accuracy** for **sentiment, intent, and topic modelling**.
- Decide on **evaluation metrics** (F1-score, accuracy, recall, etc.).

Install Required Libraries & Frameworks:

- **Hugging Face Transformers** for fine-tuning LLMs.
- **PyTorch / TensorFlow** for model training.
- Git/GitHub for version control

Phase 2: Data Collection & Preprocessing (Week 2-3)

◈ **Tasks:**

Collect Customer Conversations Dataset:

- Sources: **Customer Support Logs, Call Transcripts, Chat Logs**.
- Ensure the dataset includes **sentiment, intent, and topic annotations**.

Data Cleaning & Preprocessing:

- **Remove noise** (HTML tags, special characters).
- **Normalize text** (lowercasing, stopword removal, tokenization).

Data Annotation & Labelling (If Needed):

- Use **pre-annotated datasets** or label manually with **Label Studio**.
- Store data in **JSONL or CSV format** for training.

Phase 3: Fine-Tune the LLM on Annotated Data (Week 4-5)

◈ **Tasks:**

Choose Base LLM Model (GPT-2 or GPT-3 Fine-Tuning on Customer Data):

- Use **GPT-2 for open-source local fine-tuning** or **GPT-3 via OpenAI API**.
- If GPT-3 is used, leverage **embedding-based retrieval + prompt engineering** instead of full fine-tuning.

Fine-Tune GPT on Sentiment, Intent, and Topic Modeling:

- Convert dataset into **prompt-response pairs** for LLM training.
- Use **Hugging Face Transformers** and **PyTorch** for model fine-tuning.

Phase 4: Model Deployment & API Integration (Week 6-7)

◈ **Tasks:**

Convert the Model to Serve in Production:

- Use FastAPI or Flask to deploy the LLM as a REST API.

Develop Real-Time Prototype

- Build a web-based or command-line system that processes real conversations and offers next-step suggestions

Dashboard Development

- Create an interactive dashboard using Streamlit or Gradio to display analytics, agent performance, and topic trends

Phase 5: Final testing, Documentation and Presentation (Week 8-9)

◈ **Tasks:**

Final Testing

- Perform **full system testing** to ensure all components work as expected
- Use **Pytest** for unit testing and **Docker** for testing in a production-like environment.

Documentation

- Prepare final documentation, including a user manual and technical report

Presentation

- Create a **project demo** with slides and a video walkthrough, Use Powerpoint for slides.

9. Methodology

9.1 Packages & Tools:

- Data Analysis: Python (Pandas, NumPy), scikit-learn
- NLP and LLM Fine-tuning: Hugging Face Transformers(free models like BERT-base, DistilBERT), PyTorch, TensorFlow, or OpenAI API
- Data Augmentation: To address potential data imbalance, we will use data augmentation techniques such as synonym replacement and back-translation to generate additional training samples, especially for underrepresented classes in sentiment and intent classification.
- Topic Modeling: Gensim or spaCy-based solutions
- Dashboard & Deployment: Streamlit or Gradio for quick UI, FastAPI for scalable endpoint deployment

9.2 Algorithms:

- Transformer-based (GPT-2/GPT-3) for multi-task conversation analysis
- Classical/Deep Learning Classifiers (baseline) for sentiment and intent classification
- LDA or Embedding Clustering for topic discovery

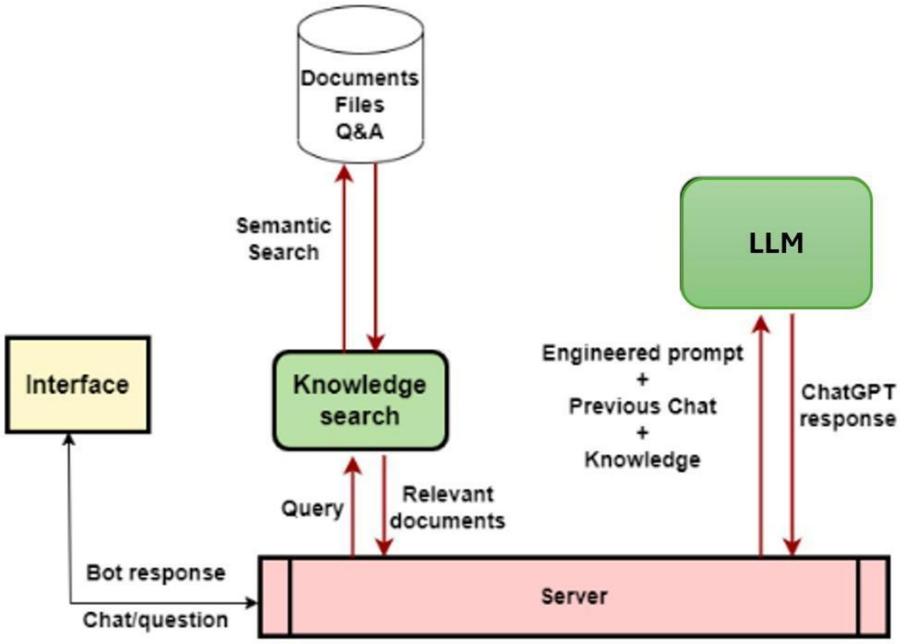
9.3 Metrics:

- Classification: Accuracy, Precision, Recall, F1-score for sentiment, intent
- Topic Modeling: Coherence score, perplexity
- Agent Performance: Weighted scores (resolution time, empathy, correctness)

- Recommendation Quality: Possibly user acceptance rate or manual review

9.4 Deployment Plan:

- Prototype & Demo: Host a proof-of-concept on Streamlit or Gradio for easy user interaction
- Scalability Option: Integrate with **FastAPI** or **Flask** for real-time endpoint calls. Use **Docker** for containerization and **Kubernetes** (optional) for scaling if needed.



10. Preliminary Exploratory Data Analysis(EDA)

10.1 Data Volume & Distribution:

- RSiCS might have fewer but in-depth dialogues; the Twitter dataset likely has the largest volume of short text.

10.2 Sentiment Skew:

- Early samples from Twitter conversations suggest more negative or neutral sentiment compared to typical call-center data.

10.3 Intent Examples:

- Common user intents: product inquiries, billing, troubleshooting, feedback.

10.4 Potential Data Issues:

- Missing transcripts, inconsistent labeling, user slang in social media data.

11. Expected Outcomes

- **Improved Accuracy:** Achieve ≥85% accuracy in sentiment and intent classification across all datasets.
- **Actionable Insights**: Extract top themes (topics) that frequently arise, enabling strategic decisions.
- **Enhanced Agent Performance**: Provide real-time suggestions that reduce average handling time and improve customer satisfaction.
- **Scalable Prototype:** A tested system that can be deployed or extended for real-world call centers or social media management.
- **Data Augmentation:** To address potential data imbalance, we will use data augmentation techniques such as synonym replacement and back-translation to generate additional training samples, especially for underrepresented classes in sentiment and intent classification.

12. Project Demonstration Strategy

- **Live Demo of the Dashboard:** Demonstrate how new conversation data is processed in near real-time using Streamlit, with sentiment and intent classifications displayed.
- **Test Queries for Recommendation Engine:** Input user messages into a web UI and illustrate how the system generates recommended responses.
- **Analytics Overview:** Present a summary of conversation volumes, agent performance metrics, and topic clusters.
- **Q&A Session:** Allow stakeholders to submit custom conversation snippets to see real-time outputs.

13. Team member Names

- Aditi Ravishankar
- Amit Dutta
- Anusha Harlapur
- Harshal R Chikhale
- Pankaj Goyal
- Raghavendra

- Ranjith Menon
- Sajeesh K K
- Satgur Prasad Rao
- Sree Lakshmi
- Vikas Yadav

14. Team Co-Ordinator Name

- Pratyusha M

15. Conclusion

This proposal outlines the development of a **Customer Conversational Intelligence Platform** powered by Large Language Models (LLMs), designed to transform customer service through advanced techniques like **sentiment analysis**, **intent recognition**, **topic modeling**, and **agent performance evaluation**. By adhering to the proposed methodology and timeline, the team aims to deliver a proof-of-concept system that demonstrates the transformative potential of LLM-driven insights in optimizing customer service operations. Beyond its immediate applications, this platform paves the way for future innovations in **real-time analytics**, **personalized customer interactions**, and **scalable AI solutions**, positioning businesses to thrive in an increasingly competitive and customer-centric landscape.

16. References

- Anthropic (2024). "Claude 3 Family." https://www.anthropic.com/news/claude-3-family
- OpenAI (2024). "Hello, GPT-4o." https://openai.com/index/hello-gpt-4o/
- Google DeepMind (2023). "Gemini: A Family of Highly Capable Multimodal Models." https://deepmind.google/technologies/gemini/
- Brown, T. B., Mann, B., Ryder, N., et al. (2020). "Language Models are Few-Shot Learners." Advances in Neural Information Processing Systems, 33, 1877-1901. https://proceedings.neurips.cc/paper/2020/hash/1457c0d6bfcb4967418bfb8ac142f64a-Abstract.html
- Devlin, J., Chang, M. W., Lee, K., & Toutanova, K. (2018). "BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding." arXiv preprint arXiv:1810.04805. https://arxiv.org/abs/1810.04805
- Wang, L., Chen, H., Xu, T., & Zhang, D. (2024). "QLoRA+: Enhanced Quantized Low-Rank Adaptation for Resource-Efficient Fine-Tuning." IEEE Transactions on Artificial Intelligence, 5(1), 120-142. https://arxiv.org/abs/2402.09353
- Li, S., Patel, V., & Rodriguez, A. (2024). "OmniCS: A Unified Framework for Customer Service Agents." In Proceedings of the 2024 Conference on Human-Computer Interaction (CHI '24), 1-14. https://aclanthology.org/2024.acl-long.182/