**\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_**

**Phase-2 Submission Template**

**\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_**

**Student Name :**R.Monika

**Register Number :** 623323243036

**Institution :** **Vetri Vinayaha** College of Engineering Technology

**Department :** Artificial Intelligence and Data Science

**Date of Submission:** 08.05.2025

**GitHub** **Repository** **Link:**

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

**TITLE:**

**REVOLUTIONIZING CUSTOMER SUPPORT WITH AN INTELLIGENT CHATBOT FOR AUTOMATED ASSISTANCE**

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

**1.PROBLEM STATEMENT**

* Modern customer support systems face challenges such as:
* High response times.
* Limited availability of support staff.
* Inconsistent answers from human agents.
* Increased operational costs.
* Objective: Develop an AI-powered **chatbot** to automate and improve customer support services by providing fast, accurate, and context-aware responses.

**2.PROJECT OBJECTIVES**

* Automate responses to common customer queries using NLP.
* Reduce response time and improve customer satisfaction.
* Handle high volumes of requests without human intervention.
* Learn and adapt based on past interactions and feedback.
* Integrate with existing platforms like websites or mobile apps.

**3.FLOWCHART OF THE PROJECT WORKFLOW**

User Query

|

V

Input Processing (Text Normalization)

|

V

Intent Recognition & Entity Extraction

|

V

Query Classification (via Trained ML/NLP Model)

|

V

Response Generation (Pre-defined / Dynamic)

|

V

**Chatbot** Response to User

|

V

Feedback Collection (Optional)

**4.DATA DESCRIPTION**

* The dataset may include:
* Historical chat logs.
* FAQs and their structured responses.
* Customer support tickets.
* Intent-**labeled** queries.
* Metadata (time, sentiment, satisfaction scores).
* Example fields:
* **User\_query**
* **Response\_text**
* **Intent\_label**
* Timestamp
* **Customer\_sentiment**

**5.DATA PREPROCESSING**

* Key steps:
* Text Cleaning: Remove special characters, numbers, and HTML tags.
* Tokenization: Split sentences into words.
* Lowercasing: Normalize all words.
* **Stopword** Removal: Eliminate common words (e.g., “is”, “the”).
* Lemmatization/Stemming: Reduce words to root form.
* Encoding: Convert words to vectors using techniques like TF-IDF or word **embeddings** (Word2Vec, BERT).

**6.EXPLORATORY DATA ANALYSIS (EDA)**

* Main goals:
* Identify most frequent queries/intents.
* **Analyze** customer sentiment distribution.
* Detect peak support hours.
* Visualize word clouds of queries.
* Understand user **behavior** patterns.
* Tools: Python, pandas, **seaborn, matplotlib.**

**7.FEATURE ENGINEERING**

* Important features:
* Text **embeddings:** Word2Vec, **GloVe,** or BERT-based **embeddings.**
* Intent Labels: Supervised classification targets.
* Conversation Length: Useful for complexity analysis.
* Response Time Metrics: For customer satisfaction **insigTechnique**

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

**8.MODEL BUILDING**

* Techniques:
* Classification Models: To predict intents (e.g., Logistic Regression, SVM, Random Forest, **XGBoost).**
* NLP Models: LSTM, GRU, Transformers (BERT, GPT-based models).
* Seq2Seq Models: For generative response systems.
* Dialogue Management: Rule-based or Reinforcement Learning.
* Frameworks: **scikit**-learn, **TensorFlow, PyTorch, spaCy,** Hugging Face Transformers.

**9.VISUALIZATION OF RESULTS & MODEL INSIGHTS**

* Confusion Matrix: For intent classification accuracy.
* Word Embedding Plots: To understand word relationships.
* Bar Charts: Intent prediction accuracy comparison.
* Response Time Reduction Metrics
* Satisfaction Score **Trendlines**

**10.TOOLS AND TECHNOLOGIES USED**

* Languages: Python
* Libraries: NLTK, **spaCy, scikit-**learn, **TensorFlow/PyTorch,** Transformers
* Visualization: **Matplotlib, seaborn, Plotly**
* Deployment: Flask, **FastAPI**, Docker, **Heroku**/AWS
* Version Control: Git, GitHub
* Collaboration: Trello, Slack, Google Drive

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

**11.TEAM MEMBERS AND CONTRIBUTIONS**

**A. Parkavi (** Data Collection, **Preprocessing** and EDA)

**J. Priyadharshani (**Model selection, Training and Evaluation)

**R. Monika (** Tools Integration**& Final** presentation)