# PHASE-1

**STUDENT NAME :**

1. Kavitha  
   **REGISTER NUMBER:**

422223243029

Surya Group Of Institutions

**DEPARTMENT:**

Department of Artificial Intelligence and Data Science(B.Tech)

**DATE OF SUBMISSION:**

19th April 2025(19/04/2025)

**PROBLEM STATEMENT:**

The topic is **“Revolutionizing customer support with an intelligent chatbox for automated assistance”.** This topic is based on a real world problem. Many products has customer service,if the customer has a query or a problem with the product the customer can directly contact the customer service through calls,emails etc…The person working in the product has to reply to the customer and take actions according to it. This project recieves the customer reviews and produce automated assistance based on the customer feedback through an intelligent chatbox.This problem is important because it saves the time of the product producer. It is worth solving because it saves time for both the customer and the producer,when the customer gives the review or the feedback the next moment the automated assistance is provided by the chatbox.

**OBJECTIVE OF THE PROJECT:**

The main objective of the project is to satisfiy the customer needs and make relevant measures to it. It also focuses on the time saving of both the customer and the producer. Based one the feedback of one customer the chatbox involves in machine learning and thus automatically solves

**SCOPE OF THE PROJECT:**

There are some features of the project such as:

**\***It includes data sets - The customer reviews and feedbacks are converted to a dataframe for futher uses.

**\***Basic information- Basic requirement of the product should be needed (example:date of manufacturing, date of expire,ingredients used to make the product etc…)

**\***Chatbot- An automated chatbot is required to respond to the customer.

Constraints of the project:

A specific datset such as the customer related is needed.The machine learning model can only be used. Data related to the product should be provided to the chatbot. Algorithm baded on the customer input the results only similar to the data should be shown hence related keywords should be used.

**DATA SOURCES:**

The dataset required for this project is “Customer service dialogue datasets.” This dataset can be downloaded from the websites such as Kaggle,UCI,APIs,synthetic etc.The dataset is a public dataset.The dataset is based on both static and dynamic.

**HIGH-LEVEL METHODOLOGY:**

**Data Collection:** The data set is collected from the Kaggle websites. The user can also create their own datsets.

**Data Cleaning:** The dataset may contain potential issues such as missing values,duplicates or inconsistent formats. To identify the missing value is.na() function can be used. To remove the column or a specified row which has a missing value na.omit(row or column) can be used. To fill the missing values fill.na() function is used. This function uses the mean method to fill the missing values.To remove the duplicates drop\_duplicates() can be used. Even unique(x) function is used to remove the duplicates.

**Exploratory Data Analysis(EDA):** Libraries such as Matplotlib, Numpy, Pandas, Tensorflow, Pytorch, Scikit-learn are used. Tools such as Rasa,Dialog flow, Microsoft bot framework are used.

**Feature Engineering:** The chatbot for providing autoamted assistance are already in current use. Hence, this project focuses on the feature of transforming existing ones to improve model performance.

**Model Building:** The algorithm used in this project are Natural Language Understanding(NLU), Deep Learning, Dialogue management, Natural Language Generation(NLG).

**Model Evaluation:** Accuracies used such as Intent Recognition Accuracy, Entity Recognition Accuracy, Response Accuracy and precision. Effectiveness such as Goal completion rate, Self-service rate, First contact resolution are also used.

**Visualization and Interpretation:** The insights should have a perfect format. The audience which is the customer should be well defined.Predictions should be based on machine learning model. Bar charts and LIne charts are commonly used for visualizing these type of datasets.

**Deployment:** The deployment is absolutely needed for this project for realizing the benefits. The colab notebook will be best to use since, it is a group project and it also contains visualizing tools.

**TOOLS AND TECHNOLOGIES:**

**Programming Lnaguage:** The main language used is python since it is easy to understand and easy and efficient to code.

**Notebook/IDE:** The platform used is Google colab. As mentioned above, google colab is best for group projects.

**Libraries:** The libaries used are Matplotlib, Numpy, Pandas, TensorFlow, PyTorch, Scikit-learn.

**Optional Tools for Deployment:** The tools used for deployment are Streamlit, Flask, Django, Plotydash.

**TEAM MEMBERS AND ROLES:**

**P.Kavitha-** She will be responsible for collection of datasets,cleaning processing and model building.

**D.Swathi-** She will be responsible for data analyzing and model evaulation.

**X.Mary Prakasam-** She will be responsible for monitoring the sources of data and feature engineering.

**A.Jasmine Joicy-** She will responsible for taking the lead in monitoring the group. Her role is to Visualizing and Interpretation part, building the chatbots, deployment and guiding the overall project completion process.