

# Sentiment Analysis on Healthcare Reviews

## 1. Introduction

The project aims to analyze sentiments expressed in healthcare reviews. By utilizing Natural Language Processing (NLP) and machine learning techniques, the model classifies reviews as **positive, negative, or neutral**. The results offer insights into customer satisfaction, common healthcare concerns, and trends over time.

## 2. Problem Statement

The primary goal of this project is to **develop a model that classifies sentiments in healthcare reviews**. Analyzing these sentiments helps improve healthcare services by identifying recurring issues and measuring satisfaction.

## 3. Dataset Information

- **Source:** healthcare\_reviews.csv
- **Content:** Textual healthcare reviews with sentiment labels
- **Size:** Multiple reviews categorized under healthcare services/products

## 4. Methodology

The methodology consists of several key steps:

### 4.1 Data Preprocessing

- Load and clean the dataset (CSV format)
- Handle missing data
- Perform text tokenization
- Remove stopwords, special characters, and punctuation
- Convert text data into numerical format using **CountVectorizer**

### 4.2 Model Training

Trained different machine learning models for sentiment classification:

- **Logistic Regression**

- **Random Forest Classifier**
- **K-Nearest Neighbors (KNN)**

### 4.3 Model Evaluation

Performance metrics used:

- **Confusion Matrix**
- **Accuracy Score**
- **Classification Report** (Precision, Recall, F1-score)

## 5. Results and Insights

- **Sentiment Distribution:** Majority of reviews fall under [positive/negative/neutral]
- **Key Themes:** Extracted common phrases and concerns using NLP
- **Trends:** Visualized sentiment variation over time
- **Actionable Insights:** Recommendations for improving healthcare services

## 6. Tools and Technologies Used

- **Programming Language:** Python
- **Libraries:** NLTK, Scikit-learn, Imbalanced-learn, Seaborn, Plotly
- **Machine Learning Models:** Logistic Regression, Random Forest, KNN

## 7. Conclusion

This project implements **sentiment analysis on healthcare reviews**, identifying trends and insights that can improve healthcare service quality.