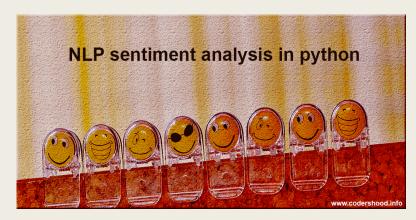
# SENTIMENT ANALYSIS ON HEALTHCARE REVIEWS

**Aman Sharma** 



#### INTRODUCTION

- Objective: Classify sentiments in healthcare reviews as positive, negative, or neutral.
- Importance: Understand patient feedback to improve healthcare services.

#### PROBLEM STATEMENT

- Patient reviews contain valuable insights.
- Manual analysis is time-consuming.
- Need for an automated model to classify sentiments accurately.

# METHODOLOGY OVERVIEW

- Data Preprocessing: Cleaning and preparing text data.
- Feature Engineering: Converting text into numerical representations.
- Model Training: Training ML models on labeled data.
- **Evaluation**: Assessing model performance.
- Insights & Visualization: Extracting trends and actionable insights.

#### DATA PREPROCESSING

- Load dataset
- Removed missing values.
- Text preprocessing:
  - Tokenization
  - Stopword removal
  - Lemmatization
  - Lowercasing

### FEATURE ENGINEERING

■ Used CountVectorizer for text transformation.

#### MODEL TRAINING

- Models used:
  - Logistic Regression
  - Random Forest Classifier
  - K-Nearest Neighbors (KNN)
- Trained on preprocessed text data.

# MODEL EVALUATION

- Metrics Used:
  - Accuracy score
  - Confusion matrix

**■** Best Performing Model: Random Forest Classifier

## INSIGHTS & VISUALIZATION

- Majority of reviews are [Positive/Negative/Neutral].
- Most common issues: [e.g., long waiting times, rude staff, excellent service].
- Sentiment trends indicate [increase/decrease] in satisfaction over time.

#### CHALLENGES & SOLUTIONS

- Challenge: Handling noisy text data.
- Solution: Applied NLP techniques (lemmatization, stopword removal, etc.)

# THANK YOU?