Support Ticket Classification & Prioritization

1.Introduction

In modern organizations, customer support teams handle hundreds of support tickets daily. These tickets include product issues, billing complaints, technical errors, and general queries.

Manual sorting of tickets causes:

- Delays in urgent issue handling
- Incorrect categorization
- Increased workload
- Reduced operational efficiency

To solve this problem, this project develops a **Machine Learning-based Support Ticket Classification and Prioritization System** that automatically categorizes and assigns priority levels to support tickets.

This system functions as a decision-support tool to assist support teams in faster and smarter ticket management.

2.Objective

The objective of this project is to build an NLP-based ML system that:

- Reads customer ticket descriptions
- Classifies tickets into appropriate **Ticket Types**
- Predicts **Ticket Priority** (High / Medium / Low)

This helps organizations:

- Route tickets to the correct department
- Handle urgent issues faster
- Reduce manual sorting effort

3.Dataset Used

Dataset: **Customer Support Ticket Dataset**
Source: Kaggle

The dataset includes:

- Ticket Description (text data)
- Ticket Type (category label)
- Ticket Priority (High/Medium/Low)
- Additional customer and resolution information

For this project:

- **Input Feature:** Ticket Description
- **Category Label:** Ticket Type
- **Priority Label:** Ticket Priority

4.Tools & Technologies Used

Development Environment

- Visual Studio Code
- Python

Libraries Used

- NLTK – Stopword removal
- scikit-learn – TF-IDF, Logistic Regression, evaluation
- Pandas – Data handling
- NumPy – Numerical operations

5.Methodology

Step 1: Data Preprocessing

The Ticket Description text was cleaned using:

- Lowercasing
- Punctuation removal
- Stopword removal (NLTK)

This reduces noise and improves model performance.

Step 2: Feature Extraction

Text data cannot be used directly by ML models. Therefore, TF-IDF (Term Frequency – Inverse Document Frequency) was used to convert text into numerical vectors.

TF-IDF helps:

- Identify important words in each ticket
- Reduce impact of common words
- Improve classification quality

Step 3: Model Training

Two separate Logistic Regression models were trained:

1. Ticket Type Classification Model
2. Ticket Priority Prediction Model

The dataset was split into:

- 80% Training Data
- 20% Testing Data

This ensures proper model evaluation.

Step 4: Model Evaluation

The model performance was evaluated using:

- Accuracy
- Precision
- Recall
- F1-Score

These metrics ensure balanced evaluation across all classes.

6.Results

The trained models successfully:

✔ Classified ticket descriptions into correct Ticket Types

✔ Predicted Ticket Priority levels

✔ Achieved reliable accuracy scores

✔ Demonstrated strong precision and recall

This confirms that the system can effectively assist support operations.

7.Business Impact

This ML system improves support workflow by:

- Automatically routing tickets to relevant departments
- Identifying high-priority tickets instantly
- Reducing manual effort
- Improving response time
- Increasing customer satisfaction

Instead of manually reading each ticket, support managers receive intelligent automated predictions.

8.Limitations

- Model performance depends on dataset quality
- Limited to predefined Ticket Types
- Does not yet support real-time deployment

9.Future Improvements

The system can be improved by:

- Using Naive Bayes or Random Forest for better text performance
- Adding Confusion Matrix visualization
- Deploying as a Web Application
- Integrating with live ticketing systems
- Using advanced NLP models (BERT, spaCy pipelines)

10. Conclusion

This project demonstrates a practical application of Machine Learning in business operations.

By applying:

- Text preprocessing
- TF-IDF feature extraction
- Logistic Regression classification

A functional and industry-relevant Support Ticket Classification & Prioritization system was developed.

This project reflects real-world operational AI implementation and showcases applied NLP skills.