Automated Ticket Classification for IT Helpdesks Using NLP

Abstract:
IT helpdesk operations often involve handling large volumes of support tickets, which can be time-consuming and error-prone when done manually. This project proposes an **automated ticket classification system** using **Natural Language Processing (NLP)** to categorize support tickets into relevant categories such as software issues, hardware problems, account-related inquiries, etc.

Methodology:
The system will preprocess the ticket descriptions using NLP techniques like **tokenization**, **stopword removal**, and **stemming**. **Text vectorization** methods such as **TF-IDF** or **word embeddings (Word2Vec, GloVe)** will be used to represent ticket text in a numerical format. The system will be trained on labeled datasets using machine learning models like **Support Vector Machines (SVM)**, **Random Forest**, and **deep learning models (LSTM, BERT)** for multi-class classification. Additionally, the model will incorporate keyword extraction and clustering techniques to improve classification accuracy.

Outcome:
The expected outcome is an automated system that can quickly and accurately classify IT helpdesk tickets, reducing the workload on support staff, minimizing human errors, and speeding up response times. Over time, the system will improve its classification capabilities by learning from new ticket data and evolving helpdesk scenarios.