

Project Description for Infosys Interns

Project Title: AI Powered Smart Email Classifier for Enterprises

Project Statement:

Enterprises receive thousands of customer support emails daily, covering complaints, requests, feedback, and spam. Manually triaging these emails is inefficient and delays response times. This project focuses on building an AI-powered email classification system that automatically categorizes incoming emails and assigns urgency levels. By leveraging Natural Language Processing (NLP) and machine learning, the solution will enhance customer service efficiency and ensure critical issues are prioritized.

Outcomes

- **Automated Email Categorization:** Classify emails into categories (complaints, requests, feedback, spam).
- **Urgency Detection:** Assign priority levels (high, medium, low) to ensure timely resolution.
- **Dashboard Interface:** Provide a visualization dashboard for customer service teams.
- **Productivity Boost:** Reduce manual sorting time and improve customer satisfaction.

Modules to be Implemented

Module 1: Email Data Collection & Preprocessing

- Collect historical email datasets (or use publicly available datasets for training).
- Preprocess emails (remove signatures, stopwords, normalize text).
- Label sample dataset with categories and urgency tags.

Module 2: Email Categorization Engine

- Train NLP-based classification models (BERT, DistilBERT, or Scikit-learn classifiers).
- Implement multi-class classification (complaints, requests, feedback, spam).
- Evaluate with precision, recall, and F1 scores.

Module 3: Urgency Detection Module

- Train a model to predict urgency (high/medium/low).
- Extract urgency signals (keywords like “urgent”, “not working”, “asap”).
- Combine rule-based and ML-based approaches for higher accuracy.

Module 4: Dashboard & Visualization Layer

- Build a dashboard displaying classified emails in real time.

Project Description for Infosys Interns

- Provide filters (by category, urgency, date).
- Show analytics (volume trends, most common complaint types).

Module 5: Deployment & Integration

- Deploy system on cloud (Azure/AWS/GCP).
- Integrate APIs for enterprise email systems.
- Provide endpoints for real-time classification.

Milestones & Timeline (8 Weeks Total)

Milestone 1 (Weeks 1–2): Data Collection & Preprocessing

- Objective: Prepare labeled dataset for training.
- Tasks:
 - Collect and clean email datasets.
 - Remove noise (HTML, signatures).
 - Label emails with categories and urgency levels.

Milestone 2 (Weeks 3–4): Email Categorization Engine

- Objective: Develop an NLP-based classification system.
- Tasks:
 - Train baseline classifiers (Logistic Regression, Naive Bayes).
 - Fine-tune transformer models (BERT/DistilBERT).
 - Evaluate classification accuracy.

Milestone 3 (Weeks 5–6): Urgency Detection & Scoring

- Objective: Implement urgency prediction.
- Tasks:
 - Train urgency classification model.
 - Combine ML + keyword-based detection.
 - Validate results with confusion matrix & F1 score.

Milestone 4 (Weeks 7–8): Dashboard & Deployment

- Objective: Deliver enterprise-ready solution.
- Tasks:
 - Build an interactive dashboard with filters & charts.
 - Deploy system on Azure/AWS/GCP.
 - Test with simulated email streams.