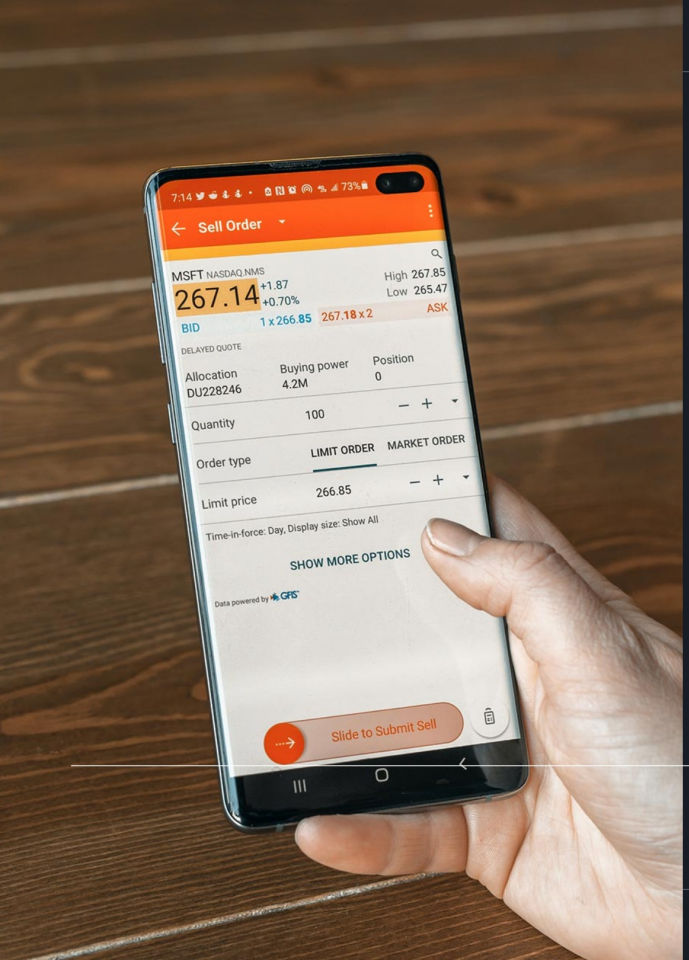



# AI-Powered Service Request Classification

Enhancing efficiency in commercial bank servicing processes through advanced AI-driven email classification and data extraction techniques.

**aininjas**
Presenter

# Challenges in Email Service Requests

Understanding the complexities of email triage

- **High Volume of Service Requests**

  Commercial Bank lending teams receive numerous emails daily, leading to increased workload.

- **Manual Triage Process**

  Emails require manual reading and interpretation by gatekeepers, consuming time and resources.

- **Identification of Intent**

  Gatekeepers must accurately identify the intent and classify requests, which is complex.

- **Classification of Requests**

  Determining 'Request Type' and 'Sub Request Type' adds to the complexity of the process.

- **Data Extraction Needs**

  Key attributes must be extracted for service requests, requiring attention to detail.

- **Assignment Challenges**

  Requests are assigned to teams based on roles and skills, which can lead to misallocations.

- **Time-Consuming Process**

  The manual triage process is inefficient, especially when request volume spikes.

- **Error-Prone Operations**

  High volume increases the risk of errors during the manual triage process.

- **Need for Automation**

  There is a pressing need to automate classification and data extraction using Generative AI.

- **Benefits of Generative AI**

  Automation can improve efficiency, accuracy, and turnaround time in handling service requests.

- **Minimizing Gatekeeping Activities**

  Reducing the need for manual gatekeeping can streamline operations significantly.

# Innovative Email Classification Solution

Overview of a Gen AI-Powered Email Solution

- **Accurate Email Categorization**
  Utilizes Gen AI to categorize emails based on sender intent and context.
- **Context-Based Data Extraction**
  Extracts configurable fields like deal name, amount, and expiration date.
- **Multi-Request Email Handling**
  Detects and categorizes multiple requests within a single email.
- **Priority-Based Extraction Rules**
  Customizable rules prioritize content extraction over attachments.
- **Duplicate Email Detection**
  Identifies duplicate emails to minimize redundancy in requests.

# Comprehensive Solution Architecture

End-to-End AI Pipeline for Email Processing


### Input Sources

The system accepts emails, documents, and predefined rules as inputs for processing.

### LLM Interpretation

Utilizes advanced LLMs like GPT, LLaMA, or Gemini to analyze and interpret the content of emails and attachments.

### Output Generation

Outputs include possible request types, prioritization, confidence scores, and extracted field values.

### Duplicate Detection

Includes a field to detect duplicates and classify reasons for flagging.

### Scalability and Efficiency

Designed for scalability, explainability, and efficiency, ensuring smooth integration into loan servicing workflows.

# Comprehensive Tech Stack Overview

Explore the key elements of our technology stack

01

### Frontend Framework

Utilizes Flask and HTML-CSS for building user interfaces.

02

### Backend Technology

Powered by Python FastAPI, facilitating seamless API integration.

03

### AI Model Selection

Incorporates fine-tuned models like bert-base-uncased for classification tasks.

04

### Embedding Techniques

Employs SentenceTransformer MODEL all-MiniLM-L6-v2 for efficient text embedding.

05

### Database Management

ChromaDB is used as a vector database to eliminate duplicate applications.

06

### Hosting Platform

Application is hosted on Hugging Face Spaces for optimal performance.

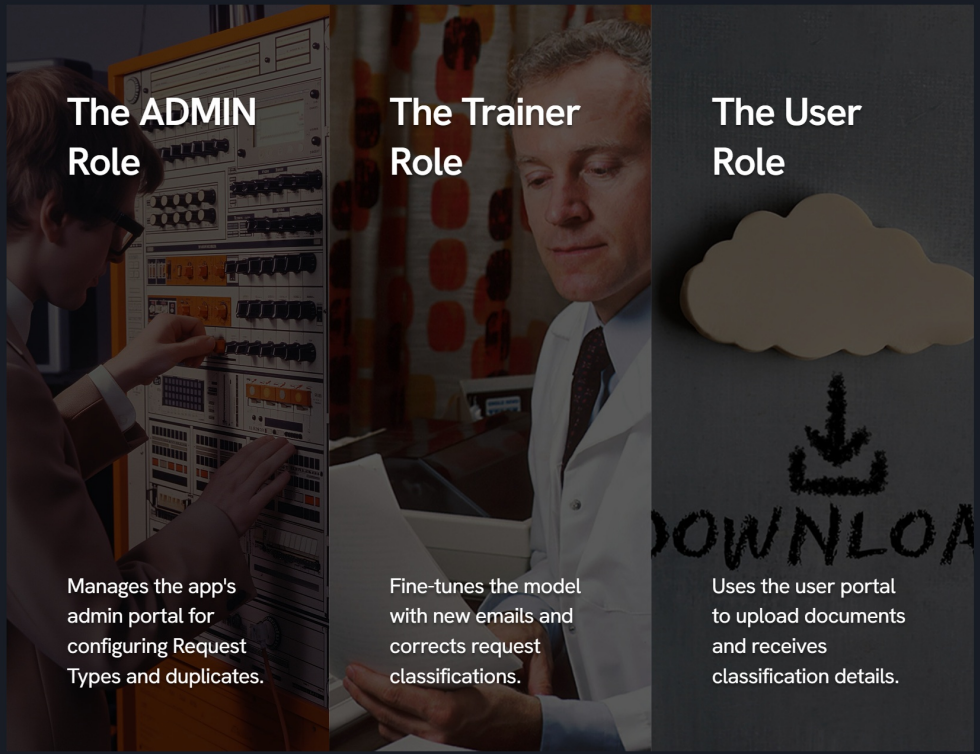
07

### Data Processing Format

Utilizes JSON-based processing for data handling and exchange.

# Understanding User Personas in Application

Explore the key roles in our application system



# Key Considerations for Implementation

Key aspects of solution implementation

- Microservice Architecture Overview
  - Three microservices are hosted on Hugging Face Space, ensuring scalability and flexibility.
- User Interface Micro Frontend
  - A UI Micro Frontend (MFE) is developed to cater to various user personas, enhancing user experience.
- Utility Service with FASTAPI
  - A domain-based utility service is implemented using the FASTAPI framework for efficient performance.
- Training and Prediction APIs
  - A dedicated service exposes training and prediction APIs through FASTAPI, streamlining model interaction.
- AI Models Utilization
  - Utilizes the bert-base-uncased model for general language support and prosusai/finbert for financial classification.
- Multiclassification Approach
  - Employs multiclassification to enable multiple predictions based on request and sub-request types.
- ChromaDB for Semantic Search
  - ChromaDB vector database is utilized for storing high-dimensional embeddings, allowing semantic search for duplicates.

# Unlock Advanced AI Solutions Today

Transform your email servicing process with cutting-edge AI technology that enhances efficiency and customer satisfaction. Reach out now at 9740699078