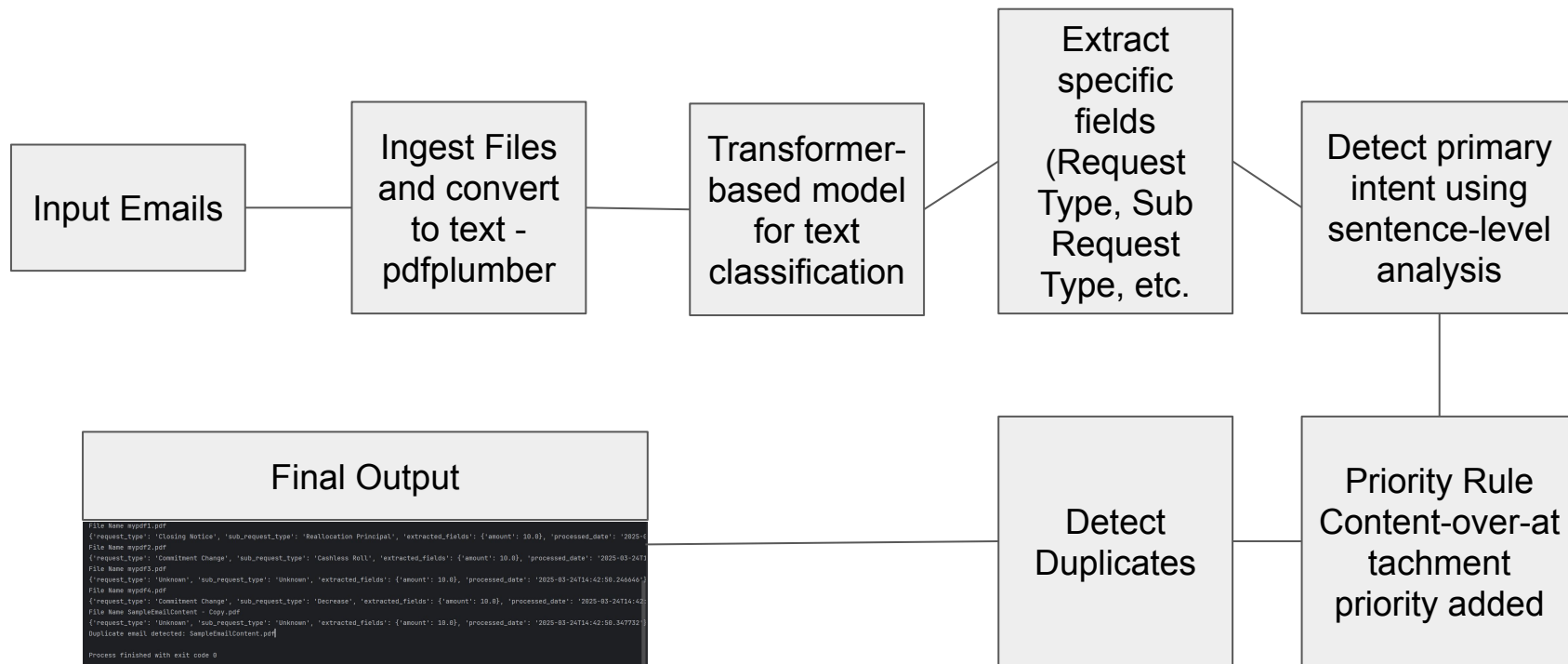


Audit Engineers & Brains

Ranjit, Srikanth, Siva, Sai and Rishabh

Process Flow



Technical Details

- Hugging Face Text Classification Model - [DistilBERT base uncased finetuned SST-2](#)
- Python 3.13
- Python Packages
 - pdfplumber pandas transformers torch nltk regex

Sample Output

```
File Name mypdf1.pdf
{'request_type': 'Closing Notice', 'sub_request_type': 'Reallocation Principal', 'extracted_fields': {'amount': 10.0}, 'processed_date': '2025-03-24T14:42:50.246646'}
File Name mypdf2.pdf
{'request_type': 'Commitment Change', 'sub_request_type': 'Cashless Roll', 'extracted_fields': {'amount': 10.0}, 'processed_date': '2025-03-24T14:42:50.246646'}
File Name mypdf3.pdf
{'request_type': 'Unknown', 'sub_request_type': 'Unknown', 'extracted_fields': {'amount': 10.0}, 'processed_date': '2025-03-24T14:42:50.246646'}
File Name mypdf4.pdf
{'request_type': 'Commitment Change', 'sub_request_type': 'Decrease', 'extracted_fields': {'amount': 10.0}, 'processed_date': '2025-03-24T14:42:50.246646'}
File Name SampleEmailContent - Copy.pdf
{'request_type': 'Unknown', 'sub_request_type': 'Unknown', 'extracted_fields': {'amount': 10.0}, 'processed_date': '2025-03-24T14:42:50.347732'}
Duplicate email detected: SampleEmailContent.pdf

Process finished with exit code 0
```