SRIRANJANI R

au211521243152

PROJECT TITLE

Gen AI-based Handwritten Text to PDF converter

AGENDA

- Problem statement
- Overview of Handwritten Text Recognition
- End Users
- Solution and Proposition
- Modelling
- Results

PROBLEM STATEMENT

In today's digital age, handwritten documents often need to be converted into digital formats for easier storage, sharing, and processing. However, converting handwritten text into a PDF format manually is time-consuming and error-prone. The aim of this project is to develop an automated system that can accurately recognize handwritten text from images and convert it into PDF documents seamlessly. This system should be capable of handling various handwriting styles, languages, and document formats to cater to diverse user needs.

PROJECT OVERVIEW

- Turning Scribbles into Searchable Files: A Handwritten Text to PDF Converter.
- This project tackles the challenge of transforming handwritten documents into easily accessible digital PDFs.
- Here's how it works: Smart Text Detective: We'll build a powerful AI model, like a deep learning system, to decipher your handwriting from images.
- User-Friendly Focus: A smooth user experience is key.

Overall, the goal of the project is to provide users with a reliable and user-friendly solution for digitizing handwritten documents in PDF format, thereby enhancing productivity and facilitating document management in various domains.

WHO ARE THE END USERS?

- ✓ Students and educators
- ✓ Professionals
- ✓ Accountant
- ✓ Customer Service Representative
- ✓ General public

YOUR SOLUTION AND ITS VALUE PROPOSITION

Handwritten to PDF converters use OCR to turn handwritten notes into editable text you can save as PDFs. This lets you organize your notes digitally, edit them easily, and share them with others. While current converters focus on accuracy and editing tools, future applications with generative AI could include improved handwriting recognition, font style conversion within the PDF, and automatic note summarization.

MODELLING

- Data Preparation: Describe the dataset used for training the handwritten text recognition model. Include details such as the size of the dataset, diversity of handwriting styles, and any preprocessing techniques applied.
- Model Architecture: Explain the architecture of the handwritten text recognition model. This could include the use of convolutional neural networks (CNNs), recurrent neural networks (RNNs), or hybrid architectures such as Convolutional Recurrent Neural Networks (CRNNs).
- Training Process: Outline the training process, including hyperparameters, optimization algorithms, and any data augmentation techniques employed.
- Evaluation Metrics: Define the evaluation metrics used to assess the

performance of the handwritten text recognition model. Common metrics include accuracy, precision, recall, and F1-score.

- Text Formatting: Discuss the algorithms and techniques used to format the recognized handwritten text into PDF documents. This may involve considerations such as text alignment, font selection, and page layout.
- PDF Generation: Explain the process of generating PDF files from the formatted text. Describe any libraries or tools used for PDF generation, such as PyPDF2 or ReportLab.
- Quality Assurance: Address any quality assurance measures implemented to ensure the accuracy and fidelity of the generated PDF documents. This could include visual inspections or automated checks for layout consistency.

RESULTS

This AI tool converts handwritten notes to searchable PDFs. We tested its accuracy and how well it preserves layouts. User feedback will help us improve. Overall, it offers a handy way to digitize handwritten content.