Harshal Honde

📞 +91 9021664306 ✉ Harshalhonde50@gmail.com 🔗 HarshalHonde 🔗 Harshalhonde 🌐 Portfolio </> Leetcode

Experience

Nroad Apr 2024 – Sep 2024
Data Science Intern *Pune, IN*

- Developed and executed full OCR workflows for financial reports using PyTesseract and PyMuPDF, transforming scanned documents into structured, editable formats for data pipelines.
- Designed and deployed NER models using PyTorch and TensorFlow, achieving a 30% increase in entity extraction accuracy for financial documents.
- Automated OCR and NER tasks, reducing manual data processing time by 50% and increasing opera-tional efficiency in financial analysis workflows

Projects

Precision Summarizer with AI | *Python, NLP, NLTK, Hugging-face-datasets, flask*

- Built a text summarization tool using LSTM and BERT, reducing text analysis time by 50%.
- Unified over 1000 summarized documents, leading to a 50% reduction in manual report preparation time and improving decision-making efficiency for business users.
- Enhanced summarization accuracy by 30% through fine-tuning models on domain-specific datasets, ensuring relevance and precision in extracted summaries.

PDF Chunk Extractor | *PyPDF2, PyTesseract, PostgreSQL*

- Developed a document processing pipeline featuring an intuitive interface for seamless PDF uploads; integrated PyTesseract and PyPDF2, which enabled rapid text extraction from 500+ scanned and non-scanned documents.
- Reduced manual document processing time by 60%, enhancing efficiency for document-heavy workflows.
- Incorporated error handling mechanisms to manage corrupted or non-standard PDFs, increasing system reliability by 40%.

ConversationFilterAI | *Jupyter notebook, python, Hugging Face Transformers, PyTorch*

- Designed a real-time voice conversation filtering system that reduced instances of abusive language in live customer-executive interactions by 50%, improving communication quality and professionalism.
- Leveraged Hugging Face Transformers to build NLP models, achieving 95% transcription accuracy and enabling real-time analysis of voice data.
- mplemented a speech-to-text pipeline using PyTorch, processing live audio streams with a 30% faster detection rate compared to existing solutions for identifying inappropriate language.

Cryptocurrency Data Scraping | *Python, Requests, BeautifulSoup, Selenium, streamlit*

- Extracted cryptocurrency market data from CoinMarketCap using web scraping tools such as Requests and Pandas, ensuring accurate and up-to-date information for analysis.
- Analyze the performance metrics of various cryptocurrencies to uncover key market trends and patterns critical for informed investment decisions.
- Generated actionable insights based on data analysis, enabling the development of investment strategies to capitalize on emerging trends in the cryptocurrency market.

Technical Skills

Languages: Python, C++, Golang, SQL
Technologies:TensorFlow, PyTorch, Flask, Docker, FastAPI, AWS
Concepts: NLP, OCR, XGBoost, Transformers, CatBoost, Cache Memory, Artificial Intelligence, Machine Learning, Neural Networks, API, Database Normalization, RNNs, Word2Vec, Vector Databases, Gradient Descent

Education

Savitribai Phule Pune University Aug 2019 - July 2023
Bachelor of Engineering in Electrical , Minor: Computer Science (GPA: 3.80 / 4.00) *Pune, IN*