RUPANSHU KAPOOR

Jaipur, India • +91-9983810448 • rupanshukapoor@outlook.com
Portfolio GitHub LinkedIn

SKILLS

Data Analysis: Excel, MySQL, Python (Pandas, NumPy, SciPy), Matplotlib, Seaborn, PowerBI, Tableau
Machine Learning: Statistical Analysis, scikit-learn, Keras, PyTorch, ANN, CNN, NLP, OpenCV
Generative AI: LangChain, HugginFace, LLM, RAG, Pinecone, ChromaDB, FAISS
Big Data Technologies: Azure Data Factory, PySpark, Hadoop, ETL
Additional Tools: Git, Data Version Control(DVC), JIRA, Flask, GCP, Streamlit

EDUCATION

Post Graduation Program- Data Science and Machine Learning **Nov 2023 - Aug 2024**
Imarticus Learning

Bachelor of Engineering - Electronics and Computer **Aug 2015 - July 2019**
M.B.M Engineering College, Jodhpur

WORK EXPERIENCE

Data Science Intern, Imarticus **May 2024 - Present**

- Developed an advanced resume parsing tool using NLP for text extraction and LLMs for context-aware information extraction, enhancing candidate profile accuracy by 30%.
- Implemented grammar and spelling error detection with intelligent suggestions, improving the quality and professionalism of resumes by 40%.

Data Engineer, Pratham Software **Jul 2019 - Dec 2020**

- Designed and developed a CPQ (Configure Price Quote) tool tailored to specific customer requirements, enabling efficient and customized pricing strategies.
- Created automated ETL pipeline using Pyspark on Azure Data Factory(ADF), reducing delivery times by 30%.

PROJECTS

SnapText: AI Image Chatbot

- Designed a Streamlit app, enabling user authentication via Firebase maintaining user security and accessibility.
- Implemented advanced text extraction from images and PDFs using Google Document AI, achieving an 85% accuracy rate in text recognition .
- Integrated a RAG pipeline for creating chroma embeddings of extracted text, facilitating efficient text search and query performance, resulting in a 40% reduction in query response time.

Technologies Used: Python, Streamlit, Firebase, Google Document AI, RAG, LangChain, ChromaDB
GitHub Repository: SnapText

Face Mask Detection

- Created a face mask detection application using a custom Convolutional Neural Network (CNN) model, achieving an accuracy rate of 99% .
- Enhanced model performance and robustness by augmenting a dataset of 12000 images, improving detection accuracy under various lighting and angle conditions.
- Deployed the application using Streamlit, creating a user-friendly interface for seamless interaction and easy integration into public safety systems during the pandemic like COVID 19.

Technologies Used: Python, Keras, OpenCV, Numpy, Pandas, Seaborn, Image Generators, Streamlit
GitHub Repository: Face Mask Detection

ADDITIONAL INFORMATION

- **Certifications**:
  - Stanford | Deep Learning AI- Machine Learning Specialization
  - IBM - Machine Learning
- **Won** Imarticus Data Science Hackathon - 2024