Daneshwari Savadkar

Bangalore, Karnataka
+91-9535792581 daneshwari924@gmail.com LinkedIn

SUMMARY

Jr. Data Scientist with 1.3 year of hands-on experience in developing AI-powered solutions for customer lifetime value (CLTV) prediction, invoice data extraction, and image processing. Skilled in machine learning algorithms, deep learning architectures, and natural language processing (NLP). Proficient in utilizing Python, TensorFlow, PyTorch, and advanced AI frameworks to build and deploy scalable models. Demonstrated expertise in developing custom NER models and optimizing deep learning solutions for multilingual document processing. Fine-tuned large language models using QLoRA and LoRA techniques to enhance model efficiency and scalability. Strong background in data preprocessing, feature engineering, and model validation, with certifications in Databricks Machine Learning and Data Engineering.

TECHNICAL SKILLS

Languages	Python, C, Java, SQL
Data Science	Machine Learning, Deep Learning, NLP, Data Preprocessing, Feature Engineering, Model Validation ,Agentic RAG
Algorithms	Pandas, Numpy, Scikit-Learn, XGBoost, Logistic Regression, Decision Trees, Random Forests, Boosting Algorithms, Neural Networks
AI/ML Frameworks	Hugging Face, Lang Chain, Spark ML, PyTorch, TensorFlow, Scikit-learn, Auto ML, NLP
Big Data Tools	Azure Databricks, Lakehouse, Apache Spark, PySpark
Technologies/Frameworks	HTML5, CSS3, JavaScript, MySQL, Pinecone
Developer Tools	VS Code, Jupyter, PyCharm, Eclipse, Google Colab, Android Studio
Cloud Tools	AWS, Azure
API	FastAPI, RestAPI
LLM	LLM Model Training, Evaluation
Container	Docker

WORK EXPERIENCE

Element Technologies INC | Jr. Data Scientist **Feb 2024 – Present**

Customer Lifetime Value (CLTV) Prediction and Analysis |
Python, TensorFlow, Scikit-learn, MongoDB, FastAPI, Uvicorn

- Designed and implemented an API for real-time CLTV predictions using Python, TensorFlow, Scikit-learn, and MongoDB.
- Conducted data extraction, preprocessing (handling missing values, outliers), and data integrity checks on large datasets.
- Performed exploratory data analysis (EDA) to derive preliminary insights and inform feature engineering.
- Deployed predictive models using FastAPI and Uvicorn, achieving scalability and reliability for production use.
- Monitored model performance post-deployment, adjusted models to address overfitting, and ensured compliance with model validation protocols.

Entity Recognition (ER) for Invoice Data Extraction | LLaMA 3.1 (Open Source) Finetuning

- Designed and fine-tuned **open-source NER models** using **LLaMA 3.1** for high-accuracy entity extraction (e.g., vendor, invoice date, amounts) from diverse and multilingual invoice formats.
- **F**inetuned the model by applying **QLoRA (Quantization + LoRA)** to improve model efficiency, along with **GPTQ and GGML** for quantization. Leveraged **ExLlama** for fast inference, achieving a 25% increase in processing speed and an F1-score of 0.92, outperforming baseline models by 15/multilingual contexts.
- Optimized for **cloud deployment**, reducing resource consumption by 20% and enabling real-time digitization for large-scale invoice processing.

AI-Powered Resume Screening Platform | Advanced AI and Semantic Search

- Pioneered the development of an **Advanced AI tool** leveraging a **Retrieval-Augmented Generation (RAG)** methodology, achieving a 98% accuracy rate in matching resumes to job descriptions, significantly enhancing screening precision.
- Transformed Job Descriptions (JDs) and resumes into high-dimensional semantic embeddings using **Large Language Models**, and stored them in a **Pinecone vector database** for precision-driven similarity analysis.
- Architected a robust system to intelligently rank resumes based on **cosine similarity**, delivering a prioritized list of top candidates with percentage match scores, effectively aligning with role requirements.
- Enabled rapid resume scanning, reducing manual screening time for HR teams to under 25 seconds for top matching resumes, while maintaining a high level of accuracy.

PROJECTS

Nephropathy Prediction using Deep Learning | ANN, CNN, MobileNet **2023-2024**

∗ Implemented Artificial Neural Networks (ANN) for blood test data and Convolutional Neural Networks (CNN) for CT scan image analysis.
∗ Achieved a remarkable training accuracy of 99.0% and a validation accuracy of 99.0% using MobileNet architecture for CT scan image analysis.
∗ Utilized a dataset comprising 5,077 normal class and 2,283 tumor class images for CNN and a blood test dataset of 400 records, featuring 26 essential attributes for ANN.

Car Number Plate Detection Using OpenCV: A Contour-Based Approach | OpenCV **2023**

∗ Developed a car number plate detection system using OpenCV, employing a contour-based approach for efficient and accurate recognition of license plates.

CERTIFICATIONS

- **Databricks Certified Machine Learning Associate**
- **Databricks Certified Data Engineer Associate**

EXTRACURRICULAR

∗ **Research Paper:** "A NOVEL HEALTH MONITORING SYSTEM FOR MULTIPARAMETERS: BLOOD GLUCOSE, TEMPERATURE AND EC" - Framework using wireless technology and geolocation-based monitoring for patient data accuracy.
∗ **Awards:**
∗ 1st prize in oral presentation at ICETEST 2023 Conference.
∗ Presented a research paper at the International Conference ICWITE 2022 held at IISc Bangalore.
∗ Participated in a poster presentation at IEEE NKCon 2022 Conference for the project "Digital Data Transmission System Of KSRTC."

EDUCATION

BLDEA College of Engineering and Technology, Vijayapur **2020 – 2024**
B.E - CSE (Artificial Intelligence and Machine Learning Engineering) ***CGPA - 8.00***