RANJEET NAGARKAR

MACHINE LEARNING ENGINEER

+1 (415)-579-6333 | ranjeet.sn96@gmail.com | San Antonio

SUMMARY

• Machine Learning Engineer with 5 years of experience in Artificial Intelligence, DeepLearning, Machine Learning, and Data Mining. • Experience in building production quality and large-scale deployment of applications related to natural language processing (NLP) and machine learning algorithms.
• Experience in developing and deploying machine-learning models, specializing in Large Language Models (LLM) andDatabricks.
• Excellent at using Python to manipulate data for data loading and extraction and worked with Python libraries like NumPy, Pandas, Matplotlib, SciPy, Scikit-learn, Seaborn, TensorFlow, Ggplot2, OpenCV, PyTorch and NLTK.
• Proficient in managing AWS cloud resources like EC2, S3, Lambda, EBS, EMR, DynamoDB, SQS, SNS, and CloudWatch.

WORK EXPERIENCE

Fractal Analytics, USA | Data Scientist ***Sept 2024 – Present***

• Developed graph analytics pipelines to detect anomalous customer associations, leveraging isolation forests, community detection algorithms, and graph traversal techniques, preventing an estimated $2 million in potential fraud losses.

• Designed unsupervised clustering models (k-means, DBSCAN) to group customer emails and addresses

• Built anomaly detection framework integrating customer behavior data, CAP (Customer Associated Persons) linkages, and transactional features to uncover high-risk entities, improving investigative efficiency by 20%.

• Collaborated with cross-functional teams to automate anomaly review dashboards using Python, SQL, and Tableau

• Performed causal inference analysis to quantify the impact of CAP rule changes on anomaly rates, presenting findings to senior leadership and influencing future fraud mitigation strategies.

Data Knobs, USA | MACHINE LEARNING ENGINEER ***Oct 2023 – July 2024***

• Designed a scalable chatbot framework using DAG architecture and langchain agents library with Voyager + ReACT prompting

• Estimated cost savings of $300,000 with enhanced dialogue management for ReAct chatbot products

• Fine Tuned LLM using PEFT and LoRA techniques for Customer Support Automation task improving KPI by 8%, leveraging Hugging Face Transformers (BERT, GPT, ALBERT, CLIP, BART) for various NLP tasks extensively utilizing Python.

• Deployed the LLM application on cloud platforms including AWS (SageMaker, EC2, EMR, S3, Lambda serverless), and GCP.

• Conducted A/B testing and causal inference experiments to optimize web applications.

Gyandata, India | Data Scientist ***Apr 2021 – Mar 2023***

• Analyzed~20M rows of Data on Snowflake using Apache Spark in Databricks to extract insights on daily equipment procurement

• Employed K-means clustering to group equipment with similar consumption patterns reducing maintenance cost by 7-8% in Python

• Deployed Gradient Boosted Regressor, engineered 20+ lagged variables, time series forecasting reducing procurement by ~5%

• Accomplished the integration of AWS services with on-premises resources, creating a hybrid cloud environment that increased flexibility and compact operational costs by 29%.

• Performed Sentimental analysis in NLP on the email feedback of the customers to determine the tone behind the series of words by Neural network techniques like Long-Short Term Memory (LSTM) cells in Recurrent Neural Networks (RNN) using Pytorch Python

Robert Bosch Center for Data Science and Artificial Intelligence, India | Data Scientist ***Feb 2020 – Apr 2021***

• Built MATLAB Application for DIPCA early fault detection and process monitoring with savings of 9% in operating costs

• Contributed to the development of an automated customer response application serving an estimated 50+ interactions daily

• Improved baseline Linear Regression model to Gradient Boosting using GridSearchCV improving prediction by 25%

• Integrated Hugging Face Transformers to optimise text processing, increasing efficiency by 10%

SKILLS

• **Language:** Python, R, Java, SQL, MATLAB
• **IDEs:** Visual Studio Code, Anaconda, JupyterLab, Jupyter Notebook
• **ML Algorithm:** Linear Regression, Logistic Regression, Decision Trees, SVM, Random Forests, Naive Bayes, K Means, Supervised, Unsupervised, Clustering, NLP, Large Language Models (LLM), LSTM, NLTK
• **Deep Learning:** ANN, CNN, RNN, Hugging Face Transformers (BERT, GPT, ALBERT, CLIP, BART), Stable Diffusion •
Packages: NumPy, Pandas, Matplotlib, SciPy, Scikit-learn, Seaborn, TensorFlow, PyTorch, NLTK, Ggplot2, OpenCV, Plotly •
Cloud Technology: Azure, AWS (SageMaker, EC2, EMR, S3), GCP(Pub/Sub, Big Query, Composer),
• **Web Experimentation:** Causal Inference, A/B Testing, Hypothesis testing, Statistical Modeling, Optimization, Experimental Design, Docker, Flask, CI/CD pipelines
• **Big Data/Database:** SQL Server, MySQL, NoSQL, MongoDB, PostgreSQL, PySpark, SparkSQL, Spark, ETL Data Pipeline, Hadoop, SVN, Snowflake, Databricks, Hive
• **Other Tools & Skills:** GitHub, Jira, Selenium, FastAPI, Airflow, Looker, Excel, PowerPoint,
• **Operating System:** Windows, Linux, MacOS

EDUCATION

Master of Science in Data Science | University of San Francisco, San Francisco, CA ***July 2023 - July 2024***

Bachelor of Technology | Indian Institute of Technology Madras, Chennai, India ***August 2015 - May 2019***