Rajesh Babu Pasupuleti Data Scientist

rajeshbabu@mywebemails.com | +1 (937) 580-8681 | Dayton, OH | LinkedIn | GitHub | Portfolio

Summary

Data Scientist with 2+ years of experience in applying machine learning and deep learning techniques using Python and R. Proficient in building neural networks with TensorFlow and Keras and performing big data processing with PySpark and Apache Kafka. Skilled in SQL, Snowflake, and cloud platforms AWS and Azure, with expertise in data visualization using Power BI and Tableau.

Technical Skills

- **Programming :** Python, R, SQL, NoSQL
- **Data Processing & Streaming:** PySpark, Apache Kafka, Pandas, NumPy, Apache Airflow, Snowflake, ETL, MS Excel, Matplotlib, Seaborn, Plotly, scikit-learn
- **Machine Learning Algorithms:** Random Forest, Linear Regression, PyTorch, SVM, Decision Tree, ARIMA, Xgboost, TensorFlow, Keras, Scikit-learn, Statsmodels, Deep Learning Algorithms, LSTM, SciPy, NLTK, RestNet 50, VGG16, FFNN, RNN, CNN, Siamese Network, GAN, LSTM
- **Frameworks & Others:** Flask, MS SQL, MySQL, PostgreSQL, MongoDB, Git, GitHub, Docker, Statistics, AWS, Azure, Power BI, Tableau, Kubernetes, FlaskAPI, Eclipse, Google Colab, A/B Testing

Professional Experience

Data Scientist, *NVIDIA*　　　　　　　　　　　　　　　　　　　06/2023 – 05/2024 | Remote, USA

- Collected and annotated 50,000 historical customer support chat logs and 10,000 FAQs using Python, employing libraries like NLTK for text processing and SpaCy for named entity recognition. Annotated data with 25,000 unique intents and entities to create a robust training dataset.
- Implemented text preprocessing using Python libraries like NLTK and SpaCy. Processed over 200,000 text samples, applying tokenization, lemmatization, and stop-word removal to transform raw text into a clean, structured format suitable for training NLP models.
- Developed intent recognition models using machine learning algorithms like Support Vector Machines (SVM) and Random Forest with Scikit-learn. Achieved model accuracy of 92% and improved F1-score by 15% through feature extraction and iterative evaluation.
- Utilized NLP models such as BERT for intent recognition and entity extraction. Fine-tuned BERT using Hugging Face's Transformers library with a dataset of 50,000 annotated examples, achieving a 90% accuracy in understanding and categorizing user inputs.
- Built response generation capabilities with deep learning models like GPT (Generative Pre-trained Transformer) using TensorFlow and PyTorch. Applied sequence-to-sequence architecture to generate responses, achieving a 95% relevance rate in contextually appropriate replies.
- Deployed the chatbot on AWS using AWS Lambda for serverless computing and AWS Elastic Beanstalk for application management. Monitored and logged over 1 million interactions using AWS CloudWatch, ensuring a 99.9% uptime and continuous system improvement.

Data Scientist, *Wells Fargo*　　　　　　　　　　　　　　　　　　04/2021 – 07/2022 | Chennai, India

- Developed a sales forecasting model using Python, Pandas, and NumPy to analyze historical sales data. Achieved a 20% improvement in forecast accuracy, predicting quarterly sales with an average error rate of less than 5%.
- Implemented a real-time data processing pipeline using Apache Kafka and PySpark for streaming analytics. Processed over 100,000 events per minute, reducing data latency by 30% compared to batch processing.
- Designed and automated ETL processes using Airflow and Snowflake, integrating data from MySQL and PostgreSQL databases. Reduced data integration time by 40% and improved data freshness by implementing incremental updates.
- Built a customer churn prediction model using TensorFlow and Scikit-Learn, analyzing customer behavior data. Achieved a predictive accuracy of 85%, reducing customer churn by 15% through targeted retention strategies.
- Developed a web-based dashboard using Flask and Power BI to visualize machine learning model outputs. Enabled stakeholders to interactively explore sales trends and forecasted data, improving decision-making processes

Education

Master of Science, *University of Dayton*　　　　　　　　　　　　08/2022 - 05/2024 | Dayton, OH
Computer Science

Bachelor of Technology, *SRM Institute of Science and Technology*　　　06/2018 - 05/2022 | Chennai, India
Computer Science and Engineering