

Jaswanth Erusu

Data Scientist

+1 571-591-7666 | jaswanth.e@mycvscout.com | [LinkedIn](#) | [GitHub](#) | [Portfolio](#)

SUMMARY

- Data Scientist with 3+ years of experience leveraging Artificial Intelligence (AI), Deep Learning, Machine Learning, Data Mining, Data Visualization, and Natural Language Processing (NLP) to solve complex business problems and drive impactful results.
- Deep learning expertise in implementing Artificial Neural Networks (ANNs), Convolutional Neural Networks (CNNs), Recurrent Neural Networks (RNNs) with LSTMs, and Transformers.
- Skilled in applying various machine learning algorithms such as linear regression, logistic regression, Naive Bayes, support vector machines, decision trees, k-means clustering, and ensemble methods.
- Proficient with various NLP methods for information extraction, topic modeling, sentiment analysis, parsing, and relationship extraction while developing, deploying, and maintaining production NLP models with scalability.
- Extensive experience in Text Analytics, generating data visualizations using R, and Python and creating dashboards using tools like Tableau and Power BI.

EDUCATION

Master of Science in Data Analytics Engineering

George Mason University, Fairfax

TECHNICAL SKILLS

- **Language/ IDE's:** Python, R, SQL, Jupyter Notebook, Google Colab
- **Machine Learning:** Linear, Logistic Regression, Decision Trees, Random Forests, Naive Bayes, SVM, A/B Testing
- **Deep Learning:** ANN, CNN, RNN, Hugging Face Transformers (BERT, GPT-3), LSTM
- **AI Technology:** Generative AI, Natural Language Processing (NLP), Large Language Model (LLM)
- **Cloud/Visualizations:** AWS, Tableau, Power BI, Google Data Studios
- **Packages:** NumPy, Pandas, Matplotlib, Scikit-learn, Seaborn, TensorFlow, Keras, NLTK, XGBoost, PyTorch
- **Database and Tools:** SQL Server, MySQL, PostgreSQL, MongoDB, DynamoDB, Cassandra, SPSS

WORK EXPERIENCE

Northern Trust, CA Data Scientist

Sep 2023 – Present

- Increased data processing efficiency by 80% using LLM-powered automated data cleaning and preprocessing pipelines.
- Established serverless data processing functions using AWS Lambda to automate data cleaning and pre-processing steps, improving workflow efficiency.
- Implemented and deployed various natural language processing (NLP) models using the Hugging Face Transformers library, achieving a 15% improvement in sentiment analysis accuracy compared to baseline models.
- Utilized Scikit-learn's extensive machine-learning library to build, train, and evaluate various models, accelerating the development process by 20%.
- Engineered custom PyTorch Data Loaders to handle complex data formats increasing data ingestion efficiency by 40% and reducing preprocessing time.
- Optimized SQL Server databases for real-time data ingestion and retrieval, supporting high-throughput applications with frequently accessed data and maintained SQL scripts to transform datasets for ML model training.

Kpit Technologies, India Data Scientist

Jul 2020 – Jul 2022

- Implemented XGBoost for advanced statistical modeling in Python, resulting in highly accurate predictions that empowered improved decision-making.
- Accelerated data preparation by 25%, leading to a boost in model accuracy, by designing customized text cleaning and preprocessing tools that effectively handled domain-specific language patterns.
- Enhanced time-series prediction accuracy by 80% through the implementation of Long Short-Term Memory (LSTM) networks.
- Utilized machine learning algorithms and advanced statistical analysis (decision trees, regression models, SVM, clustering) with the scikit-learn package in Python.
- Established deep learning models with Keras for natural language processing (NLP) tasks and improved sentiment analysis accuracy by approximately 20%.
- Leveraged AWS EMR to create highly scalable machine learning pipelines, capable of handling training datasets exceeding terabytes in size and achieving approximately 10% faster model deployment times.
- Built custom SQL queries to analyze customer behavior, leading to a 90% increase in customer segmentation accuracy.
- Collaborated with cross-functional teams to implement automated dashboards and reporting using Tableau, empowering stakeholders with data-driven insights, and a remarkable 25% improvement in workforce performance.
- Executed complex SQL queries to extract and analyze data from various sources, including databases, data warehouses, and API.