EDUCATION

Virginia Tech

Blacksburg, USA

Aug 2024 - Present

Masters in Computer Engineering

Coursework: Advanced Machine Learning, Large Scale Software Development

Sri Sivasubramaniya Nadar (SSN) College of Engineering

Chennai, India

Bachelor Engineering in Computer Science Engineering | GPA: 8.5/10.0

Aug 2018 – May 2022

Coursework: Data Structures, Design Analysis of Algorithms, Probability and Statistics, Artificial Intelligence, Machine Learning, Data Analytics

WORK EXPERIENCE

Fidelity Investments

Chennai, India

Data Engineer

Aug 2022 - Jul 2024

- Developed an ML automation model using Python, TensorFlow, and SQL to interlink multiple datasets, ensuring real-time synchronization across use cases, improving data consistency by 95%, and accelerating development cycles by 25%, saving one week per cycle while enabling faster company-wide decision-making
- Analyzed and resolved data issues related to time overlaps and duplicate records in the data lake, implementing
 validation checks and deduplication processes that improved data accuracy by 98% and enhanced overall process
 efficiency.
- Led the migration of ETL pipelines from Concourse to Jenkins, developing automation scripts to convert Concourse YAML files into Jenkins declarative pipelines, reducing manual effort by 70%. Completed the migration ahead of schedule, saving 15% in costs by decommissioning legacy systems early.
- Migrated data from legacy systems to the cloud using AWS, Airflow, and Snowflake, enabling seamless transitions for 800 clients and improving scalability and system performance.

Software Engineering Intern

Jan 2022 - Jul 2022

• Developed an automation framework on Airflow using Python, AWS, Snowflake, SQL, and Django to track daily job runs, providing real-time visibility into task statuses, reducing issue response time by 40%, minimizing downtime, and decreasing manual efforts by 75% across the team

Vellore Institute of Technology (VIT)

Chennai, India

Research Intern

August 2023 - July 2024

- Utilized one-shot learning and MLP neural networks with SMOTEENN and PowerTransformer for skewness correction on a UCI dataset (11 features, 1,599 samples), improving predictions on a small, imbalanced dataset.
- Achieved 98.36% accuracy, outperforming previous models, with 98% precision and recall, using an ensemble of Random Forest, SVM, Gradient Boosting, and XGBoost.
- Developed a stacked ensemble model that was distinctive in handling small datasets while maintaining high accuracy, improving upon existing literature in wine quality prediction.

PROJECTS

Gaussian Processes for Automating Model Selection (paper)

Developed a two-tiered system for model selection and hyperparameter optimization using a multi-armed Gaussian Bandit and Bayesian optimization, achieving an 8-12% accuracy improvement, 30% faster convergence, and reducing computational costs by 15% on large-scale datasets.

NAS for Automated ML Deployment on Extreme Edge Devices (paper)

Researched and optimized Neural Architecture Search (NAS) models for extreme edge devices, achieving up to 4.8% accuracy improvement, 4-13x memory reduction, and 1.7-3.3x faster inference times

A Computational Analysis of Climate Change Sentiment on Social Media (paper)

Conducted research analyzing climate change sentiment using Word2Vec and BERT models on 600M+ tweets, achieving high-accuracy sentiment classification and uncovering a 15% increase in urgency-related discussions over three years

Monitoring Mental Wellness through Sentiment Analysis (code)

Devised a mental health analysis system using text mining, text analysis, bidirectional LSTM and word embedding to interpret contextual emotions from text only narratives and monitor the user's mental state with 90% accuracy

SKILLS

Programming Languages: Python, R, Java, JavaScript, TypeScript, C++, SQL, HTML, CSS, Linux scripting Tools/Technologies: Git, Docker, Airflow, AWS, Jupyter, Pandas, Seaborn, Scikit-Learn, Kubernetes, CI/CD, React, REST