ANEESH KRISHNA

+1(716) 303-1623 | aneeshkr@buffalo.edu | Linkedin | Github | Portfolio

EDUCATION

University at Buffalo, The State University of New York **Buffalo, NY**
Master of Science in Data Science **Dec 2024**
Key Courses: Statistical Learning and Data Mining, Machine Learning, Deep Learning, Database Management

Visvesvaraya Technological University **Mangalore, INDIA**
Bachelor of Engineering in Computer Science **July 2020**
Key Courses: Algorithm Analysis, Big Data Analytics, Cloud Computing, Data Structures and Applications

SKILLS AND CERTIFICATIONS

- **Programming Languages & Tools**: Python, C#, SQL, R, MATLAB, Microsoft Excel.
- **Data Visualization & Business Intelligence:** matplotlib, seaborn, Plotly, Tableau, Power BI, Streamlit
- **Data Science & Machine Learning**: pandas, NumPy, scikit-learn, TensorFlow, PyTorch, Keras, Deep Learning, NLP, Transformers, CNN, LLM, Time Series Analysis, Langchain, Llama, GenAI, Dask, Optuna
- **Cloud & Big Data Technologies:** AWS, GCP, Azure, Hadoop, Spark, ETL, Cloud Firestore
- **Certifications and Specializations:** Google Data Analytics Specialization, Data Science and Business Analytics, Deep Learning Specialization.

EXPERIENCE

Graduate Research Assistant, University at Buffalo, NY: **July 2024 - Present**

- Accomplished a 30% increase in anomaly detection accuracy by designing and implementing a hybrid model pipeline using supervised machine learning and behavioral analysis on system log data.
- Achieved a 20% improvement in real-time attack vector classification accuracy by developing a method for packet analysis across all 7 OSI layers and extracting advanced features from log sequences.
- Enhanced threat detection by 25% through the integration of NLP models, which enabled context-aware interpretation of network traffic patterns and log anomalies.
- Reduced false-positive rates by 15% and improved incident response times by 40% through leading the creation of a comprehensive evaluation framework for intrusion detection systems, utilizing both historical and real-time security log data.

Senior System Engineer, Infosys Ltd, Mysore, India: **Oct 2020 - Apr 2023**

- Modernized 15+ outdated web components using SPFX, reducing loading time by 30% and enhancing user experience through improved performance and streamlined functionality.
- Monitored migration of over 10 million files and metadata from 67 sites using the Microsoft ShareGate tool to the latest site, meeting stakeholder expectations.
- Developed PowerBI reports and dashboards, providing critical insights and aiding strategic planning.

PROJECTS

Cricket Statistics LLM Chatbot click here *LLM, GenAI, Streamlit, FireStore*

- Engineered an interactive Streamlit-based chatbot utilizing Google's Gemini AI model to generate and execute pandas code for cricket statistics analysis, enhancing user engagement with complex data.
- Improved AI model accuracy by 25% as measured by user feedback scores by implementing a comprehensive feedback system using Google Firestore and creating a custom fine-tuning pipeline for the Gemini AI model.
- Expanded the cricket statistics database by 30% through ESPNCricinfo web scraping, enriching player information for comprehensive cricket analysis.
- Reduced application deployment time by 50% as measured by time-to-market metrics by leveraging Google Cloud Platform services to seamlessly integrate AI, database, and web technologies in a cloud-based environment.

Optimizing Healthcare Facilities Through EDA & Machine Learning click here *Pandas, Optuna, Seaborn, scikit-learn*

- Accomplished 90% size reduction by data-cleaning, encoding (to JSON), and performing feature selection on a large dataset, leveraging libraries such as polars.
- Extracted healthcare insights, spotlighting gender and age-specific trends in diseases, post-operative complications, and hospital length of stay via Plotly, Matplotlib and Seaborn libraries.
- Built predictive models for patient length of stay, attaining an 82% accuracy through Gradient Boosted Trees after Bayesian hyperparameter optimization.

Sports Data Analytics of Cricket World Cup click here *Pandas, matplotlib, Seaborn, Tableau*

- Cleaned and preprocessed over 1 million records from ESPNcricinfo of all T20 matches (2015-2023) using Python and Pandas, enhancing data quality for analysis.
- Extracted and analyzed stadium statistics and team performance data for the 2024 T20 World Cup using SQL, Pandas, Tableau, Matplotlib, and Seaborn, providing strategic insights and detailed team profiles.

British Airways Review Dashboard click here *KPI, ETL, Tableau*

- Developed an interactive Tableau dashboard analyzing 300+ reviews of British Airways' services, focusing on cabin staff performance, aircraft efficiency, and traveler feedback to provide insights into customer satisfaction.
- Utilized advanced filtering, parameter control, and data visualization techniques to explore service ratings by traveler type, seat type, and geographical trends, enabling detailed analysis for stakeholders.

Adidas USA Sales Dashboard **click here** *KPI, ETL, Tableau*

- Designed a dynamic Tableau dashboard to track Adidas USA sales performance, focusing on month-over-month growth, retailer contributions, and regional sales, driving actionable insights for strategic decision-making.
- Enhanced business forecasting by visualizing sales trends across products and regions, leveraging custom calculations, interactive maps, and real-time data updates for more precise insights.

Loan Risk Prediction Model for Small Businesses **click here** *Python, matplotlib, Seaborn, PostgreSQL*

- Performed preprocessing on initial Excel data, employing NoSQL to normalize the data and establish a relational database.
- Conducted EDA by querying data with SQLite, producing informative visualizations on business sector distributions and default rates using seaborn and plotly libraries.
- Implemented an ML model for classifying businesses' loan repayment likelihood, achieving a 93% accuracy rate with Random Forest algorithm.