ANEESH KRISHNA

aneeshkr@buffalo.edu | Linkedin | Github | Portfolio

EXPERIENCE

Data Science Fellow **Aug 2024 – Present**
Advance2000 Buffalo, NY

- Delivered a RAG-powered MVP product, a helpdesk assistant that boosted ticket resolution speed of support team by 250% by leveraging past solutions to provide relevant response suggestions for new tickets.
- Enhanced model accuracy by 10% through advanced prompt engineering, including chain-of-thought, few-shot prompting, and metadata filtering, optimizing ticket solution relevance.
- Reduced helpdesk response time by 25% by implementing an LLM system using Hugging Face embeddings, Pinecone vector databases, and Llama 3.1 through Groq for real-time ticket analysis.
- Processed solutions for 42,000 open tickets, achieving 98% completion rate within a day by implementing a custom rate-limiting algorithm and parallel processing, optimizing for API token usage and operational efficiency.

Graduate Research Assistant **Aug 2024 – Present**
University at Buffalo Buffalo, NY

- Enhanced AI-driven cybersecurity threat detection through the integration of BERT and GloVe models, achieving 99.98% accuracy and a 1.0 F1 score in context-aware interpretation of network traffic patterns and log anomalies.
- Boosted attack vector classification accuracy by 20% using packet analysis across all 7 OSI layers, extracting advanced features from log sequences for real-time threat detection.
- Achieved 98.55% accuracy in anomaly detection using a Naive Bayes model with TF-IDF Vectorizer, optimizing feature extraction and classification for system log analysis and network security.

Data Migration Analyst **Oct 2020 – Apr 2023**
Infosys Ltd Remote, India

- Designed and automated Power BI dashboards and reports to track data migration progress across 90+ sites, using DAX for calculations and providing real-time analytics to stakeholders for data-driven decisions.
- Streamlined SharePoint site redesign processes by utilizing PowerShell and SQL based scripts, significantly reducing manual intervention by 50% and increasing deployment efficiency for large-scale data migration projects.
- Led a team of 5 engineers in automating complex data workflows and modernizing legacy SharePoint sites, improving data processing efficiency by 60% through the integration of SPFX, Azure Devops and optimized data pipelines.
- Coordinated User Acceptance Testing (UAT) of reports to ensure a seamless transition into production, enhancing reporting accuracy by 40% through advanced data visualization techniques and stakeholder feedback integration.

PROJECTS

YouTube Q&A Bot| Hackathon (~24 hrs) - GitHub *Python, RAG, Flask, LLM*

- Developed a chatbot using Flask, YouTube API, and Pinecone to build a chatbot that answers questions based on YouTube video content and provides timestamps.
- Applied RAG by vectorizing transcripts with Hugging Face embeddings and using GROQ API for LLM-based responses.

Sports Statistics Generator| Published Website (~100 hrs) - Site / GitHub *Python, GenAI, Streamlit, FireStore*

- Engineered a website using Streamlit, Python, Firestore, and Gemini Pro LLM to generate cricket statistics plots based on user queries on a dataset of over 1 million records.
- Collected user queries and feedback on plot accuracy to automatically generate Pandas code, refining the model using Google AI Studio on Google Cloud Platform.

Patient Outcomes Predictor | Team Project (~60 hrs) - GitHub *Python, Optuna, Seaborn, scikit-learn*

- Cleaned and analyzed healthcare data using Polars and Seaborn to reduce dataset size by 90%, identifying trends in diseases and post-operative complications.
- Devised predictive models (Logistic Regression, SVC, Random Forest, Gradient Boosting) to estimate post-operative complications with 82% accuracy after Bayesian optimization.

Business Loan Classifier | Team Project (~40 hrs) - GitHub *Python, Seaborn, PostgreSQL*

- Leveraged PostgreSQL and Python for data preprocessing to maintain integrity and optimize ETL processes; performed EDA to analyze business sectors and default rates, achieving 93% accuracy in loan repayment predictions using Random Forest.

SKILLS AND CERTIFICATIONS

- **Languages**: Python, SQL, R, MATLAB
- **Software:** AWS, GCP, Langchain, TensorFlow, Pytorch, scikit-learn, Keras, Hugging Face Transformers, Power BI
- **Frameworks**: XGBoost, LightGBM, Optuna, Pandas, Numpy, Plotly, SQL, Seaborn, Pinecone, Groq, Huggingface

EDUCATION

University at Buffalo, The State University of New York **Expected Dec 2024**
Master of Science, Majors: Data Science Buffalo, NY
Relevant Coursework: Statistical Learning and Data Mining, Machine Learning, Deep Learning

Visvesvaraya Technological University **Aug 2016 – Aug 2020**
Bachelor of Engineering in Computer Science India
Relevant Coursework: Data Structures, Algorithms, Software Engineering, Database Management