Vikram Kalister

● 224-733-2591 ● vikram.kalister@gmail.com ● LinkedIn

WORK EXPERIENCE

Antline LLC Jun 2023 - Present

Data Scientist | AWS, Azure, Python, Pandas, Scikit-learn, Seaborn, Docker, R, PowerBI

- Developed a Tenant Screener tool using classification models (Logistic Regression, Random Forest, XGBoost) to evaluate applicant reliability, achieving 87% accuracy and integrating with internal systems for real-time screening.
- Created a House Price Prediction system to support property investment decisions, using regression models (Gradient Boosting, Random Forest) and feature engineering on real estate and market data; improved valuation accuracy and reduced manual analysis time by 60%.
- Conducted customer segmentation using K-means clustering on demographic and behavioral data, identifying four target personas and enabling marketing strategies that improved engagement by 12% and conversions by 9%.

Wintrust Financial Jun 2021 - Aug 2021

IT Intern | PowerBI, SQL

- Enhanced data accuracy by 20% through detailed analysis of credit reports, eliminating discrepancies in a database of over 1 million customers.
- Improved reporting efficiency by 50% using Power BI to showcase trends based on various metrics including date, account number, and location.
- Developed over 10 interactive dashboards in Power BI to analyze call volume and patterns for Microsoft Teams, driving insights into communication efficiency.

TECHNICAL SKILLS

Certifications: AWS Certified Cloud Practitioner (Dec 2022), Oracle Certified Foundations Associate (Oct 2023)
Languages: Python, R, Java, SQL, JavaScript, HTML, CSS
Tools: scikit-learn, ARIMA, Google Colaboratory, RStudio, Power BI, Databricks, Django
Cloud/Databases: AWS (Lambda, EC2, Glue, EMR, S3, QuickSight, SNS, SQS, CloudWatch), Azure, Oracle, MS SQL Server

PROJECTS

Web Traffic Analysis Pipeline | AWS Glue, QuickSight, S3, EMR, Spark, Hive Jun 2023 - Aug 2023

- Built a cloud-based log analytics solution using AWS Glue and EMR for a 1.2 million record dataset.
- Automated data extraction from S3 using AWS Glue, streamlining data transformation and improving performance by 200%.
- Configured EMR clusters for SQL analytics on web traffic logs; visualized data in QuickSight, for 16% year-over-year increase in website traffic.

Wind Power Output Forecasting | Python, Pandas, Scikit-learn, Seaborn, Google Colab Jan 2023 - Feb 2023

- Developed an ARIMA model in Python to predict wind power output, providing insights to optimize wind energy generation.
- Analyzed 5 years of wind power data, allowing for a 5% increased optimization in wind turbine energy production.
- Preprocessed and visualized time series data, achieving forecasts with a mean absolute percentage error below 6%.

Heart Disease Prediction | Python, Pandas, Scikit-learn, Imbalanced-learn, Seaborn, Google Colab, GitHub Mar 2022 - Jun 2022

- Built machine learning models to predict heart disease likelihood from a dataset of 100,000+ patient records.
- Addressed data imbalances utilizing SMOTE, increasing model effectiveness by 15%.
- Improved model accuracy by 7% using Support Vector Machines, leading to earlier detection of heart disease in 3% of patients.
- Performed hyperparameter tuning with Grid Search, optimizing model performance by an added 2%.

EDUCATION

DePaul University | BS in Data Science Sept 2018 - Jun 2022

- **Relevant Coursework:** Advanced Data Analysis, Data Visualization, Advanced Machine Learning, Foundations of Data Science, Introduction to Big Data Processing

Linköping University & INSEEC Business School | BS in Business Administration Aug 2020 - Jun 2021

- **Relevant Coursework:** IT Project Management, Business and Professional Communication, Business and Economics, Research Methodology