MADIHA (**Madi**) ANSARI

Data Science and Machine Learning Engineer

Mobile: 0423292380 **E-mail**: madihaalansari@gmail.com **LinkedIn**: https://www.linkedin.com/in/madihaalansari/ **GitHub** Portfolio

Experienced data professional helping businesses to achieve data-driven solutions and implementing creative ways to cut costs by leveraging ML/AI-based technologies.

Core Competencies:

Project Management: End-to-end ML project planning, developing, testing, and monitoring, using frameworks like agile and CI/CD for automation and scalability.
Communication: Clear speaking and writing ability to communicate complex data concepts with management and cross-functional teams.
Collaboration: Ability to work closely with diverse stakeholders to identify critical priorities through continuous feedback, ensuring on-time project delivery.
Technological Awareness: Always aware of the business needs. Finding creative ways to implement task-specific technological solutions.
Leadership: Empowering others by encouraging and offering challenging goals with responsibility and appreciation.

Technical Competencies:

Programming Languages: Python, SQL, PowerShell/Bash, C++, VBA.
ML/AI Frameworks: Scikit-learn, Keras, Pytorch, TensorFlow, PySpark, OpenCV, BERT, VADER.
Cloud Technologies: AWS S3, AWS EC2, AWS RDS, Azure Analysis Services.
Data Visualisation Tools: Power BI, Python, and Tableau.
DevOps: CI/CD, APIs, Versioning, Logging, ML workflow, GitHub, Docker. Apache Spark.
Project Management: Knowledge of Agile, and Prince2 frameworks.

Professional Experience

---

1 - (Data Science and ML) Business & Data Analyst, Ventura Motors — Sep 2021 – Present

Data Solution for ML Applications - Electric Bus Performance and Battery Management System (Application)

Total Project worth $10 million

- Implemented End-to-end, **Big data pipeline** in AWS for ingesting MQTT Broker data from CAN Bus.
- Applied advanced ML techniques like **Principle Component Analysis, Moving Averages,** etc. for dimensionality reduction and data normalization to help analyze critical vehicle metrics for **Predictive maintenance**.
- Collaborated with diverse stakeholders, including government, private, and research-based entities.
- ***Improved data-driven decision-making in saving operational costs by 20%.***

MLOPS - Intelligent Accidents Monitoring and Management System (Application)

Estimated annual accident damage $3 million.

- Developed and Deployed **End-to-end ML Application** on **AWS EC2** with bus accident data spanning over 20 years.
- Trained multiple supervised models like **SVM**, **XGBoost**, and **Random Forest** for predictive analytics.
- Trained and used **Silhouette Score** to evaluate performance for **DBSCAN** and **K-Means** clustering models to understand accident patterns.
- Used **GitHub** and **Docker** for CI/CD Implementation to streamline data integration and model monitoring.
- Collaborated with management and cross-functional teams.
- The System helped management mitigate road accident severity, thus by investing in customized driver training programs.
- ***Helped reduce annual damage costs by 15% in better managing third-party accidents and insurance claims.***

* For this project I developed time-series analytics with data from ADAS-fitted vehicles in comparison to non-ADAS vehicles.

Data Solution for ML/AI Applications - Patronage Monitoring System for Public Buses in Southeast Victoria

Business contracts worth 96% of revenue.

- Created **end-to-end data solution** for daily Passenger movement using AWS. I used the **GPR - Gaussian Process Regression** Model, a probabilistic approach to predicting passenger count.
- Preprocessed real-time data using **Scikit-learn.** Analytics based on time-series features like days of the week, peak-off peak, events, and season variations.
- The system helped in understanding diverse customer needs and route demands in the Eastern Region. Thus providing better service and saving operational costs for managing underutilized assets.
- ***Comparison analysis found a 20% discrepancy in passenger counts with Myki ticketing system data***.

Text Analytics using AI modeling - Customer Satisfaction Analysis

Used NLTK Tokenization and Lemmatization techniques. Employed **VADER** for polarity scoring and **BERT** for part-of-speech tagging and named entity recognition(**NER**).

Image Processing using OpenCV - Paperless Invoice data capture

- Automated text extraction from encrypted financial remittances in **PDF** files and other invoices in **image** formats**.**
- Utilized **OpenCV** and **PIL** to preprocess noise in images like **grayscale conversion**, **thresholding,** etc.
- Tuned the **Tesseract-OCR** engine to optimize performance for documents with complex layouts, like columns and tables.
- Finally, used **Regex** pattern matching for the targeted extraction of invoice numbers, dates, payment amounts, and account details.
- ***OCR-enabled solution helped in time and cost efficiency, by cutting time from 10 mins to 1 min to go through 1 invoice. It saved the company monthly $2,500 on about 1000 docs per month.***

Performance Monitoring Analytics - Vehicle Safety Inspection & Vehicle Off-Road Tracking System

- Developed an automated data pipeline for the **Power BI** application.
- Used various **Python** libraries to preprocess data for a traffic light predictive safety maintenance dashboard.
- ***It Improved asset management significantly by achieving 100% safety compliance on 1000 vehicles thus saving $1,500 per vehicle if they default. Previously 5 vehicles were defaulting each month.***

Information Protection - Risk Management for Data & Information Security.

Worked on various data security frameworks like ISO 27001 & Essential 8 to define an Industry-specific Risk Register. Collaborated with government and non-government agencies to improve sensitive data security (resilience, response & recovery).

Data Strategy for ESG - Implementing Sustainability Standards Framework - Reporting for Corporate Carbon Footprint

Developed strategic data required for tracking corporate emissions and reporting for compliance frameworks. Pioneered **Ventura's Corporate ESG report in 2023**.

2 - Data Scientist, Miepol (Pty) Ltd. Oct 2019 – Sep 2021

- **Data Science & Engineering: Bus Stop Management & Predictive Maintenance System.** *It helped the business save about **$50,000 annually** on purchasing complex data solutions and tailoring them to the stakeholder's needs.*
- **Data Engineering:** Data-driven **Wayfinding for passengers** to identify bus stops that require route change. Automated **ETL** pipeline using **Power BI**. Helped the business save **hours** on manual data processing

3 - Academia – Mathematics, STEM, and Programming in Python /C++. Aug 2009 - Oct 2019

A STEM-focused educator with project-based programming for more than 10 years.

4 - System Research Analyst, LMKR. Jan 2005 – Nov 2007

- **Analytics for Reservoir Performance for Oil and Gas Production.**
- **Seismic Activity Monitoring and Anomaly Detection.**

5 - Database Coordinator, Petrosin Engineering. Sep 2003 – Dec 2004

Data-driven Procurement Optimization for LPG Cylinder Supply. Created **SQL database** for inventory stocks, procurement costs, and client information. Cross-application automation using **VBA**. Advanced **Excel** for cleaning and forecasting analysis.

Education and Certifications

- **MLOps (AWS): Data Pipeline Automation & Optimization using Amazon Web Services.**
- **MLOps (Azure): Data Pipeline Automation & Optimization using Microsoft Azure Machine Learning.**
- **Apache Spark for Data Engineering and Machine Learning.**
- **Cyber Security Risk Management.**
- **Analyzing and Visualizing Data with Microsoft Power BI.**

- **Master of Data Science. University of San Diego, CA. Relevant Coursework:** SQL for Data Science, Python for Data Science, Machine Learning, Probability and Statistics in Python, Big Data Analytics Using Spark, Data Protection & Security.

- **Bachelor of Computer Engineering (Hons). Comsats University. Relevant Coursework:** Computer programming, Embedded Systems and Design, AI and Machine Learning, Computational Algorithms, Signals and Image Processing, Calculus and Discrete Mathematics, Linear Algebra, Systems Engineering, Database Design and Data Structures.

Community Engagement and Volunteering

- An active member of the Glen Eira Strategic Transport Advisory Committee.
- Active member aimed at supporting govt's mission: 'Go Electric Plan'.
- Project Engagement with the City of Glen Eira Sustainability program called 'Energy Smart'.
- Training educators for makeshift SOS schools in disaster-stricken areas.
- Helped SAYA, an NFP Organization for orphaned and abused children.

***References available on demand**