NAVYA MOHAN K.K.

DATA SCIENTIST/MACHINE LEARNING SPECIALIST

SUMMARY:

- 10+ years of experience in **technical research & development**, mainly in **Machine learning** and **Data science.** Worked in automotive industry research and development for 8 years.
- 5+ years of experience in **Python.**
- Experienced in Python to manipulate data for data loading and extraction and worked with python libraries like Matplotlib, Scipy, Numpy and Pandas for data analysis.
- Involved in the entire **data science** project life cycle and actively involved in all the phases including data preprocessing, EDA (exploratory data analysis), feature extraction, ML model creation, data visualization with large data and performance evaluation for accuracy.
- Skilled in performing **data analytics** - data extraction, data ingestion, data preparation, data mining, data parsing, data manipulation.
- Excellent knowledge on Supervised and Unsupervised **Machine Learning** Algorithms – Classification and Regression analysis.
- **Data engineering** experience along with **synthetic data** generation and handling of pii fields.
- Deep knowledge in **ML frameworks** and libraries.
- Familiar in **Big Data** concepts with Hadoop, **Spark** and **Scala**.
- Knowledge on **Azure** Pipelines and Dataflow, SQL Pools, Azure Data Lake, Azure Synapse Analytics, Azure Data bricks, Azure DevOps and MLOps.
- Knowledge of **SQL and NoSQL** databases.
- Experience in development and testing for VIL (Vehicle in Loop).
- Knowledge on the concepts of Neural networks, Deep learning and Artificial intelligence.
- Experience in **automation** scripts.
- Experienced the full software life cycle in SDLC, **Agile** Methodology, DevOps and Scrum methodologies including creating requirements, test plans and JIRA.
- Excellent **communication skills** with leadership abilities along with analytical and **problem-solving** skills.
- Experience in **automotive** and **banking** domains.
- Dealt with sensitive data (financial) that have security concerns.

EDUCATION:

- M.Tech from Kerala University, Trivandrum 2012 (88.00%)
- B. Tech Computer Science and Engineering, CUSAT 2010 (76.67%)

PUBLICATION:

- Presented a paper based on Image Processing in IEEE conference which got published in IEEE xplore website - https://ieeexplore.ieee.org/document/6497939

TECHNICAL SKILLS:

Software	Matlab, Simulink, Rational DOORS, Enovia, Mindmap, Visio, Apache NiFi
Languages	Python, Matlab, R, SQL, Open CV, C++, Scala

Tools/Platform	Azure Synapse Analytics, ML Studio, Apache Spark, dSpace Automation Desk, Vehicle Spy, EPIC sensors, Data loggers
ML Libraries	Numpy, Pandas, Scipy, Scikit-Learn, TensorFlow, Matplotlib, SDV, Keras
Data Visualization	Tableau, Power BI, Matplotlib, Seaborn, Matlab,
Requirement Management tools	JIRA, Azure DevOps, Rational DOORS, Enovia
IDEs	Jupyter Notebook, Spyder, Visual Studio, Pycharm
Methodologies	Agile, Scrum, Waterfall
Version control	Git, SVN

PROFESSIONAL EXPERIENCE:

Accenture
Data Science /Machine Learning Specialist
May 2022 – June 2023
URL: https://www.accenture.com/in-en/
A banking domain-based project. The aim of the project is to create synthetic data for the bank customer database due to security reasons.
Roles & Responsibilities:

- Pre and post processing of data –Automate the preprocessing of data for data cleaning and based on user specific requirements. Post processing of data to revert the custom changes.
- Data Analysis and Parsing – Detailed data analysis for interpretation of categorical and numerical data characteristics. Meta data creation based on data characteristics.
- Synthetic data creation – Collaborate with Datacebo for synthetic data creation. Use synthetic data vault library for evaluating synthetic data.
- Metrics and performance evaluation – Perform tests such as Chi-squared and KS metrics for performance comparison between synthetic data and real data.

Environment: Python, Data Analytics Platform, SDV

Additional Project Responsibility:
June 2023 – Till Date
Amazech Solutions, LLC

- Onsked Scheduling Solution - A seamless centralized schedule management solution which will significantly reduce the complexities involved in tracking appointments to individuals, families and businesses.
- The project is divided into 4 parts – Scheduling OS, Scheduling App, Schedule predictor (AI/ML), Scheduling Application Data, Metadata.
- Supported in the AI/ML part to identify potential models or use Azure ML services to identified the best fit Model.

Environment: Azure Synapse Analytics

Tata Elxsi, Trivandrum
Specialist Software Engineer
December 2012 to May 2022
URL: https://www.tataelxsi.com/

- Software-over-the-Air (SOTA) automation process - Development and testing for vehicle in loop (VIL)
- Noise Factor Development and Testing (NFT) – Identification of potential failures during software update and the method to tackle the same.
- **Data Science** for experience enhancement - The project involves identifying the techniques and feature that may improve the experience of the occupant and implementation of feature using machine learning techniques.
- Participated in all phases of data mining, data collection, data cleaning, developing models, validation, and visualization to deliver data science solutions.
- Machine learning algorithm implementation - based on the random forest and SVM.
- Data logger configuration (both in bench and remotely) and linking logger to server.
- Script creation for loggers based on vehicle type and model year.
- Installation of data logger in vehicle. Perform different verification checks such as pre installation checks on the bench, checks with installation on a test vehicle, tests after deployment of logger.
- Investigation on Heart Rate Variability for Driver Stress Analysis - The project aims to investigate the capability of biosensors to detect variability in heart rate.
- Machine learning algorithm was used for the implementation. Involved in problem analysis & Identification, Information gathering, Database collection.
- Design, Matlab Coding, Debugging, Implementation and Testing
- Self-learning car project - The aim of the Self Learning Car project is to develop an intelligent self-learning vehicle that will offer a completely personalized driving experience and help prevent accidents by reducing driver distraction.
- Feature feasibility study, detailed literature survey and benchmarking for different self learning technologies for JLR cars.
- Feature related signal identification. Matlab coding for analyzing data.
- Pattern recognition and machine learning to find similar patterns and repeating scenarios.

Environment: Python, SQL, dSpace Automation Desk, Matlab, Jupyter, Simulink, Spark, Scala, Enovia, Rational DOORS, Vehicle Spy Software, Data logger, Vector tools, EPIC Software, ECG detection sensors, Open CV, Visio, Mindmap, MS Excel, MS Project, Apace NiFi,SVN.