Eckertstrasse 30C, Graz, Styria, 8020, Austria

Oct 1992

mnv.harisyam@gmail.com

+436605607756

https://harisyammnv.github.io/

harisyammnv

SKILLS

Python

PyData Stack, PyTorch, PyTorch-Lightning

PostgreSQL

AWS Cloud

Serverless Architechtures

Docker, Kubernetes, Git (CI/CD)

Tableau

Prefect, Spark and Dask

Scalable Data pipelines

Atlassian Stack (JIRA, Confluence)

Expert

LANGUAGES

English

Expert

German

Intermediate (Niveau B1)

Harisyam Manda

Data Scientist

A passionate Data Scientist with **4 years** of Industry Experience and a Certified Professional **Scrum Master** with excellent communication skills

SUMMARY

- Extensive knowledge in **Python** with libraries such as **scikit-learn**, **Pandas**, **PyTorch**, matplotlib, plotly and BI-Tools like **Tableau**
- Experience in building large, scalable ML apps using **AWS** Cloud
- Working on streaming analytics with **Kafka** and **Spark-Streaming**

WORK EXPERIENCE

AVL List GmbH (Jul 2021 - Present)

Data Scientist

Customer Project: Development of a scalable analytics pipeline for classifying vehicles with risk of catalyst failure

Customer: FORD Motors, US

Roles & Responsibilities:

- Analyzed **Timeseries** sensor data collected from **10K** vehicles located in North America which was sent by the customer in batches
- Optimized and pre-processed the raw data using **Dask** and created partitioned parquet files (**Prefect** Data pipeline)
- Performed **aggregations** on the data at scale
- Calculated custom KPIs using domain knowledge from Engineering
- Built and Trained **XGBoost** model for risk classification
- Observed model robustness through **feature engineering** and analyzed performance using metrics such as Precision, Recall
- Tuned the model using **skopt** bayesian hyperparameter search
- Wrapped the model using **FastAPI** REST endpoints for inference
- Deployed the optimized model in an **AWS Lambda** function for active inference and feedback
- Developed **Tableau** Dashboards for displaying the results and KPIs to key stakeholders

Key Achievements: Deployed an end to end ML service which helped in monitoring the fleet and thereby reducing vehicle downtime and increased customer satisfaction

R&D Project: Reinforcement Learning based Thermal Management for Cabin Heating Mode Selection **Funded by: ECSEL FRACTAL**, EU Commission

- Implementation of **reinforcement learning** based models aimed at improving energy efficiency and reduction of environmental pollutants by having effective heating mode selection for Cabin Thermal Management
- Leading a team of 3 junior data scientists and assigned research targets through scrum framework

NANODEGREES

Data Engineering Nanodegree (Jun 2020)
Udacity
https://bit.ly/3kO0fy0

Machine Learning Nanodegree (Mar 2018)
Udacity
https://bit.ly/3LW1JSO

CERTIFICATIONS

Professional Scrum Master (Jan 2020)
Scrum.org
https://bit.ly/3kOpxMj

PUBLICATIONS

Next Service Date Forecasting for Commercial Vehicles (Jun 2022)
JSAE

Development of a ML pipeline for proactive scheduling of maintenance for Commercial Vehicles based on their Telematics data

INTERESTS

Bicycling

VolleyBall

INNO Power (previously GE Power) (Aug 2018 - Jul 2021)
Data Scientist

Project Goals: Development of analytical application for pro-active maintenance of spark plugs in gas engines

Role and Responsibilities:

- Managed the scrum team as a **scrum master**
- Analyzed task requirements as per daily scrum calls
- Responsible for creating an in-house **time-series** data processing **library** that can process data at scale
- Analyzed and investigated different Root Causes (RCA's) for spark plug failures in **Tableau** using KPIs calculated with the in-house library
- Pre-Processed timeseries data coming from **spark-ignition** system
- Segmented different categories of spark plugs using **K-Means** clustering
- Developed and trained multiple deep learning models using **LSTMs** to forecast end of the life of spark plugs
- Deployed the model in **AWS Fargate** as a cronjob for **batch-predictions**
- Used **AWS SNS** for sending notifications to **maintenance** team for pro-active maintenance of spark plugs which helped in increasing the usage life and reduced downtime of engines

Key Achievements: Developed and deployed ML-based services in **AWS cloud** which helped in saving warranty costs and increased uptime of gas engines

EDUCATION

RWTH Aachen University (Oct 2015 - Mar 2018)
M.Sc. Computational Science
1.5

BITS Pilani (Aug 2010 - Jul 2015)
M.Sc. Chemistry and B.E. Mechanical Engineering
9.3/10.0

PROJECTS

Data Science Trainee (Jun 2017 - Apr 2018)
DAIMLER AG, Stuttgart

Developed a robust Machine Learning pipeline for dealing with noise and skewness in very large databases. Implemented a custom Deep Neural Network model comprising of denoising autoencoders in PyTorch with python which can deal with annotation and feature noise and skewness in very large data sets

Key Achievements: The developed ML pipeline is used by the DAIMLER Trucks Big Data Team for analyzing the root causes of truck damages when used in semi-autonomous mode

https://bit.ly/3w9I9vz