Sanath Vijay Haritsa

Mannheim | +4917648994438 | sanath.vh@gmail.com

Experience

Nichesolv Private Limited **June 2020 - September 2023**
Data Scientist Bangalore, India
https://www.nichesolv.com/

Ace

- Developed an application for the analysis of a Tennis match - provide performance statistics of a player using computer vision for stroke recognition, ball tracking, detecting court lines, etc.
- Tools - Python, Tensorflow, OpenCV, YOLO, transfer learning using VGG16 and ResNet.

Cavitation Prediction

- Performed time series analysis to predict the cavitation in pumps before it reaches a critical level.
- Plotted the trend in the calculated variables to get an understanding of when the cavitation will occur.
- Tools - Python, ARIMA, matplotlib.

Quality Chain Committee analysis

- Detected the causal factor for the failure of a certain semiconductor chip.
- Analyzed the inspection data of the chip to check if there is any error in the inspection process.
- Analyzed the manufacturing data using control charts to find the main cause of defective chips.
- Tools - SVM, XGBoost, Random Forest, KDE plots, scatter plots and box plots.

Explainable AI

- Conducted literature survey on methods for obtaining local interpretations of black box models.
- Methods investigated - LIME, SHAP, Anchors, DiCE.
- Implemented these methods on multiple open source datasets, as well as industry data.
- Implemented evaluation metrics to compare different methods.
- We found that LIME was significantly faster in execution but SHAP had considerably better performance.

Dashboard web application

- Created a PoC web application to monitor the Health Index of a ship engine.
- Used a pre-trained SVM model to predict health index based on data coming from sensors.
- Used SHAP along with scatter, line and bar graphs to convey which factors are causing the engine to fail.
- Tools - Python Flask and Dash.

Clustering

- Used python to develop executable files to be used in a No-code ML tool.
- The executables are capable of performing data processing (scaling, normalization, label encoding), build clustering models (KMeans, XMeans) on the data, as well as create clusters and provide visualization.