EAS 504 ASSIGNMENT-3

By: - Katyayni Shankar Kaushik

General Information: -

The following Lecture was held by Mr. Ram Narsimhan, he currently holds the position of Principal Data Scientist in GE Digital. He is an alumnus of University at Buffalo and has done his Masters and Phd from UB in Operations Research.

He talks about Industrial Data Science and its use cases, he explained us how GE digital has implemented data science for the growth of the company.

Base Questions: -

1.) What are principal uses of data sciences in this domain?

The principal use of Data Science/ Industrial Data Science in Industrial Operations can be categorized as follows: -

- To make an hourly check of temperature in material sciences domain using Time series plots.
- To calculate the remaining useful life of a machine, this helps in business continuity plan.
- To make anomaly detection analysis for maintenance of machines, in order to increase their lifetime.
- To make supply chain of commodities in day to day life, more optimized and reliable.
- To make earthquake prediction or in broader sense, natural calamity, in order to prevent business interruption.
- To predict machine failure before it actually happens.

2.) How are data and computing related methods used in the organizational workflow?

- Sensor & Signal Processing, Prognostic System & Methods, Control & Optimization, Image Processing & Analysis, Knowledge Discovery & Contextual Analysis, all these are data related techniques that play a major role in making industrial operations more optimized.
- Industrial engineering has evolved with time, a new concept of IIOT i.e. Industrial internet of Things has helped in making industrial operations more optimized and reliable by connecting industrial machines, data & people.
- Time series models are significant of Data Science domain, they play a significant role in monitoring machines, especially temperatures, by generating an hourly trend.

3.) What data science related skills and technologies are commonly used in this sector?

The data science related skills and technologies commonly used in this sector are as follows: -

- Time Series based models (ARIMA, SARIMAX etc) for time-based data modeling.
- Data Visualization tools for generating trend analysis, for e.g. Tableau, Power BI, Qlikview, Qliksense etc
- Python (pandas for time series modeling).
- Application for ensemble methods (Machine Learning) for earthquake prediction.
- Theoretical & Applied Statistics, Image processing & Analysis, Contextual Analysis. Survival Analysis.

4.) What are the primary opportunities for growth?

Data Science is projected to be in huge demand in coming years. Speaker introduced a lot of growth opportunities in use cases that he discussed. Following are they: -

- Oil industry has always been a major player in terms of operations and revenue generation. With the emerging technologies, that make use of data to make industrial operations and supply chain more optimized, Oil companies would certainly like to leverage its vast amount of data being generated on a daily basis.
- Energy and Power Sector has always had industrial operations on a huge scale. Such companies want to ensure business continuity, and therefore plans to use the data in best possible manner for e.g. Anomaly detection, machine monitoring, predicting remaining life etc.
- Speaker also talked about Healthcare domain. He explained how equipment in hospitals can be best used for better monitoring. Data can be leveraged so as to see, which area has what sort problems or diseases that are prevalent there.

<u>Other Questions with respect to this Lecture</u>: -

Ques-1.) Characteristics of Data Science problems discussed in the talk?

Ans Most the Data Science problems in industrial engineering operations pertain to following categories: -

- Anomaly Detection prediction
- Time Series Analysis for hourly temperature monitoring
- Survival Analysis
- Control & optimization
- Remaining useful life analysis and many more

Ques 2.) How industrial data science problems differ qualitatively from other data science problems that you have encountered?

Ans

Most of the industrial data science problems have to deal with Time series analysis, because industrial companies have to monitor the timely trend (dourly, daily, weekly, monthly) trend of the machines being used. While IT giants like Google, Facebook, Linkedin etc work on graph-based models. A major of the risk analysis in Data Science industry is done in Fintech, Industrial, Healthcare and Energy sector. This holds a significance. Industrial Data Science holds some major differences than other data science problems: -

- It involves more physics/ Engineering based models
- It needs much less data
- Such models are powerful, but difficult to maintain and scale.
- Industrial Data Science problems are empirical, and implements heuristic rules & insights.