Assignment 1

EAS 504

Anish Shah

Base Questions:

- What are the principal uses of data sciences in this domain?
    1. Rapid Analytics for on-time delivery of products and services.
    2. Aftermarket demand forecasting.
    3. Developing new services and products that will utilize big data.
    4. Repositioning existing services and products to utilize the data science models.

- How are data and computing related methods used in the organizational workflow?
    A. MOOG's manufacturing and procurement workflows were reviewed and business data related to Shipment/Order information, Part Information, Manufacturing/Work Orders, Inventory, Vendor/Purchase Orders was identified and mapped to support the project objectives.

- What data science related skills and technologies are commonly used in this sector?
    A. Skills such as Exploratory Data Analysis, Compromise Programming, Bayesian Machine Learning, Gaussian Process Learning and Extrapolation, Bayesian Network, Text Mining for finding Word Co-Occurrences and Topical analysis using latent dirichlet allocation are used.
    Technologies such as MATLAB for mathematical calculations and model development, SQL for understanding and preparing the data for analysis using SQL queries and data aggregation is used.

- What are the primary opportunities for growth?
    A. Primary opportunities for growth in this sector would be collecting more data as the current models were just trained on the 17 months data and training the models on the larger dataset so that we can see how the model accuracy changes from the previous version and therefore making changing and keep developing new models to improve accuracy and insights.

Organizational values:
- Speakers Andrew Loeb and Dan Rubach emphasized on the 12 organizational values.

Our Values

TRUST IS A MUST
COMPETENCE IS KING
WE TRY HARDER
WE'RE ALL IN THIS TOGETHER
IT'S MY JOB
WE LOOK FOR SOLUTIONS - NOT SCAPEGOATS
COMMUNICATION IS CRUCIAL
FORMALITY DOESN'T HELP
WE HAVE TO BE ADAPTABLE AND READY TO CHANGE
COMMITMENT SHOULD BE REWARDED
WORK SHOULD BE AN ENJOYABLE EXPERIENCE
YOUR PERSONAL LIFE IS IMPORTANT

- We got to know how following this values by each person in company creates a good work environment and due to being good place to work how there are many people who are in the company for a substantial amount of time.
- These values just not being essential for their company but wherever we work and some of them are very important to follow in our day-to-day life also as they help us become good human beings which is a must in today's age.

Case Studies:

1. Bayesian Network and Text Mining:
   - 11 nodes were used in developing an Averaged Bayesian Network. They were of three different categories Quality, Delivery and Cost. An Intuitive framework was obtained using this analysis. All the parent nodes were mainly quality statistics.
   - For Text Mining non-conformance documents gathered from past two years were used and Word Frequency Cloud was developed. Topic Analysis was also performed using latent Dirichlet Allocation technique which uses Bayesian sampling technique to discover themes from document corpus and word frequencies. Different potential systemic topics were identified,
   - Focused effort was used as that will allow the greatest positive impact on the whole network.

2. Aftermarket Demand Forecasting:
   - First step in developing a predictive model for aftermarket demand forecasting is establishing demand. We also got to know that forecast is always wrong.
   - Time Series for Demand was created using the actual customer contracts data.
   - For building the predictive model Bayesian Machine Learning techniques such as Gaussian process was used. Using this algorithm, a structure was learnt from the data.
   - The posterior sample obtained using the model seems to predict identical results to historical data which tells us about consistency.

Conclusion:
- We learnt many things such as how we should identify a business case before deciding on solution.
- We should try different tools, understand their limitations.
- We should let data tell its own story and how machine learning methods are useful for learning structure.
- Interpreting and communicating the structure is the final step in the process.