For over 90 years, MFS has maintained its leading position in the financial services marketplace by being a customer-centric company. To maintain its competitive position, MFS has enacted a policy that sales staff regularly interact with their clients on a personal basis, by regularly emailing, calling and meeting their clients. After each of these interactions, the data generated is recorded in MFS’s internal system. In the age of big data, how to utilize these data to help MFS expand its market share becomes a new challenge for MFS faces. In this project, we assist MFS in extracting valuable information from unstructured data on activities with one of the 300+ partner financial advising companies of MFS. This allows MFS to understand the influence of different activities on client relationships, and manage sales activities in a more efficient manner.

**ABSTRACT**

Better Understand Activity Influence

**Methodology**

The prediction model gives a better understanding of how different frequency and different types of activities influence future purchase amount. For example, we examine by adding a specific activity, what change will bring compared to baseline which includes 30K financial advisors from 30K+ other 300+ partner financial advising companies. It is the first time unstructured data has played an important role in data science project in MFS.

**Prediction Model Performance**

It provides a new perspective on how unstructured data could be stored and managed in the future. By using the pipeline for processing unstructured data we developed, this project could be expanded to cover activities with the other 300+ partner financial advising companies.

**Influence on Future Data Management**

- It is the first time unstructured data has played an important role in data science project in MFS.
- It provides a new perspective on how unstructured data could be stored and managed in the future.

**DATA SETS**

Records of interactive activities between sales specialists and 300 financial advisors from 2015 to 2018.

**Activities Categorized Into 4 Types Based On Content**

- Non Business Topic (6.8%)
- Life (26.8%)
- Product Sale (15.0%)
- Market Fact (51.4%)

**Graph 1:** Pie Chart on the distribution of the four activity content we identified in training set, the model has an over 70% accuracy on the test set.

**Graph 2:** The performance of different methods in predicting different future transactions evaluation metrics measured in accuracy.