As previous posts mentioned that data mining could be widely used, I'm going to talk about data mining applying in the Banking Industry.

The banking system has gathered daily transactions and operations for years. This valuable information could be used by data mining in many different ways. Most commonly, we could use them in descriptive statistics by clustering, or predictive statistics by classification.

Clustering helps to identify similar classes of objects. This is the technique to combine transactions with similar behaviors or the same set of queries into one group. For example, people living in a certain area and of particular age group demand a particular set of services, like customers who are 45 years older always a demand for value investment style products rather than growth, likewise the same age of people who lives in different areas have preferences for some particular wealth management companies. All of these information would help banking sales representatives to sell their products more effectively and easier to target their 20% customers who will provide their 80% profits.

Classification, on the other hand, is the most commonly used technology on fraud detection and credit risk management. For example, in fraud detection, the part of current data (training data) would be analyzed by the decision tree algorithm. Then apply the other part of data (test data) to estimate the accuracy of the classification rules. If the accuracy is acceptable, then the rules can be applied to supervise future transactions.

In general, clustering could help banks to learn from the previous data and make profits more effectively. Classification could help banks and customers to prevent future risks.

Reference: Farooqi, R., Iqbal, N. (2017). *Effectiveness of Data mining in Banking Industry*. Retrieved from https://www.researchgate.net/publication/329518897\_Effectiveness\_of\_Data\_mining\_in\_Banking\_Industry\_An\_empirical\_study