

Data Analysis Overview

Motivation

- Science of examine raw data with purpose of deriving conclusions about information
- Applying algorithms to derive insights , ex: running through different data sets to look for meaningful correlations between each other
- It is used in a number of industries to allow the organizations and companies to make better decisions as well as verify and disprove existing theories or models.
- The focus of Data Analytics lies in inference, which is the process of deriving conclusions that are solely based on what the researcher already knows.

Applications

Healthcare: The main challenge for hospitals with cost pressures tightens is to treat as many patients as they can efficiently, keeping in mind the improvement of the quality of care. Instrument and machine data is being used increasingly to track as well as optimize patient flow, treatment, and equipment used in the hospitals. It is estimated that there will be a 1% efficiency gain that could yield more than \$63 billion in the global healthcare savings.

Travel: Data analytics is able to optimize the buying experience through the mobile/ weblog and the social media data analysis. Travel sights can gain insights into the customer's desires and preferences. Products can be up-sold by correlating the current sales to the subsequent browsing increase browse-to-buy conversions via customized packages and offers. Personalized travel recommendations can also be delivered by data analytics based on social media data.

Gaming: Data Analytics helps in collecting data to optimize and spend within as well as across games. Game companies gain insight into the dislikes, the relationships, and the likes of the users.

Energy Management: Most firms are using data analytics for energy management, including smart-grid management, energy optimization, energy distribution, and building automation in utility companies. The application here is centered on the controlling and monitoring of network devices, dispatch crews, and manage service outages. Utilities are given the ability to integrate millions of data points in the network performance and lets the engineers use the analytics to monitor the network.

Skills required

- Basic programming skills in python
 - Simple usage instructions will work
- Statistical skills and mathematics: Descriptive and inferential statistics and experimental designs are a must for data scientists.
- Machine learning skills
- Data wrangling skills: The ability to map raw data and convert it into another format that allows for a more convenient consumption of the data.
- Data Intuition: it is extremely important for professional to be able to think like a data analyst.

Python Packages

- Numpy
- Pandas
- Machine learning libraries
- Matplotlib for visualization

