

Data Science: extract knowledge from large scale data and use it for prediction, decision making, etc

Important field because:

- ① new sources of data
- ② new capabilities to acquire, store, process data
- ③ new algos & methods to analyze data

this course covers 3 main topics:

- ① Introduction & Applications of data science
- ② Machine learning & data analytics
- ③ Big data processing & manipulation

Ex Netflix has \$1x10⁶ in competition for algo for best user ratings prediction

- ↳ massive data (100x10⁶ ratings, 5x10⁵ users, etc)
- ↳ High dimensionality: complicated set of factors
- ↳ missing data: 99% of data missing

Ex - Amazon recommender system

- speech recognition (Siri, Alexa, etc...)

- stock market prediction

- real estate prediction

- robotics

- spam detection

- object / face recognition

- self-driving cars

Ingredients for data science:

- data

- machine learning

- big data manipulation techniques

① Video:

Difference between data & knowledge?

- knowledge is useful part of data for your application
- data is all bits & bytes

②