

Economic Analysis of Social Network (Part II)

Dr. Vinci Chow
vincichow@cuhk.edu.hk

Motivation

- Age of **Big Data**: data comes in a rate and in a variety of types that exceed economists' ability to analyse it
 - Texts, image, speech, video...
- Businesses in all areas have realized that data is the new gold—the ability to generate insight from data leads to more business opportunities

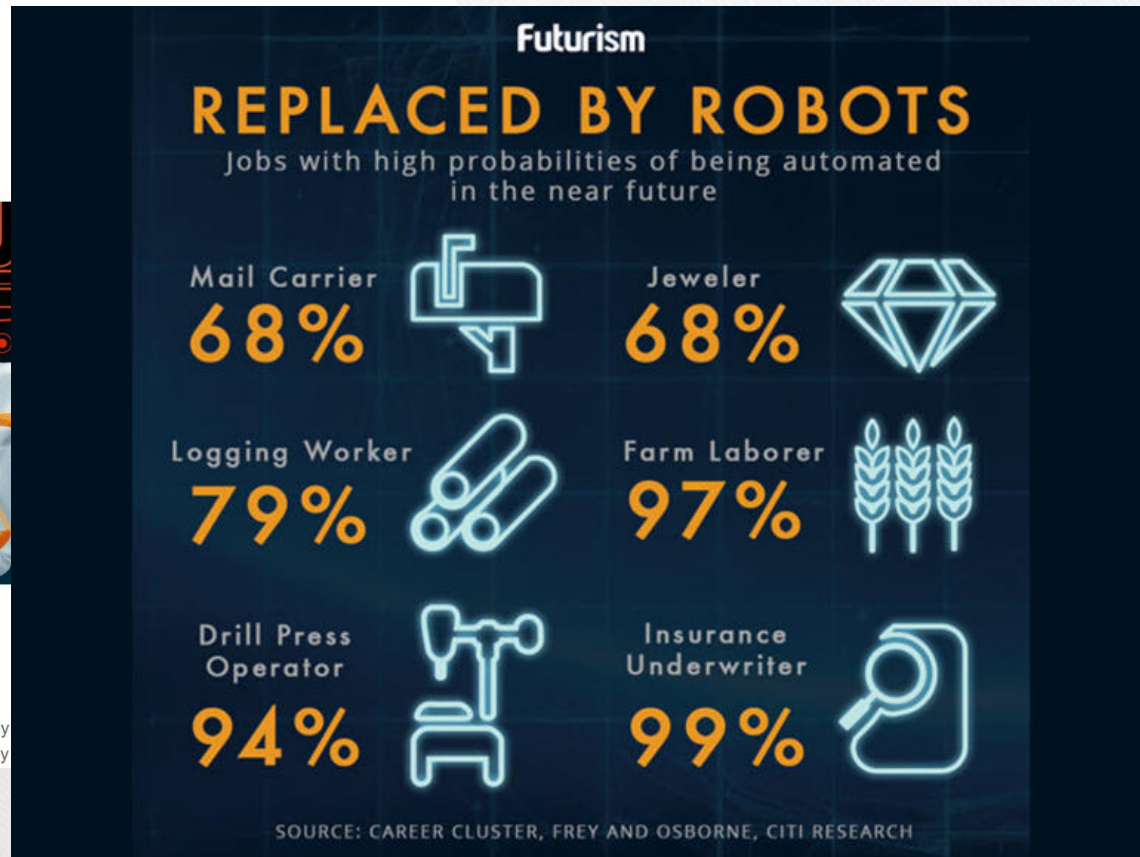
Motivation

An AI just beat top lawyers at their own game



BY MONICA CHIN
3 DAYS AGO

The nation's top lawyers recently
to interpret contracts — and they

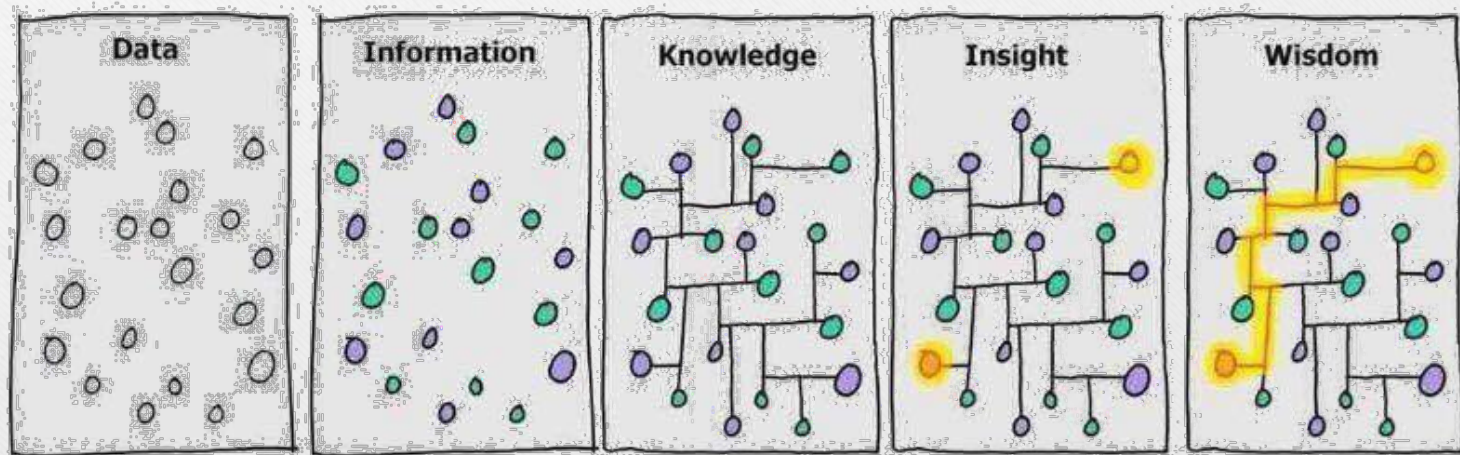


Machine Learning

- Computer scientists have studied data analysis under the umbrella of **machine learning**
 - This includes familiar statistical techniques such as regression and maximum likelihood
 - Less common (to econometricians) such as k-neighbors, regression trees and artificial neural networks
 - To statisticians, machine learning is kind of like econometrics—new names, not necessarily new stuff

Machine Learning

- Objective is usually **accurate prediction**, hypothesis testing is rare



Learning Outcomes

- ✓ Learn Python as a programming language for data analysis
- ✓ Get familiar with basic machine learning ideas and tools
- ✓ Able to conduct data analysis with real data

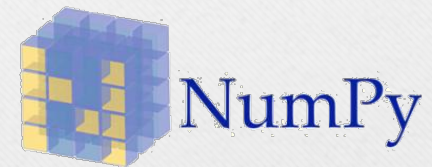
Python

- Python is an open-source, cross-platform programming language
- Very intuition syntax, lots of extensions



Python Libraries Used

- Scikit-learn
 - Python's standard machine learning library
- NumPy
 - High performance array data processing, including matrix and linear algebra
- TensorFlow
 - Google's neural network library



Additional Libraries

- Beautiful Soup 4
 - Web scraper
- Various libraries that help us to load, process, visualize and save data

URL for Course Material

- `econ1101.econ.cuhk.edu.hk:8000`
- Resets every 8 hours. Let me know if you want a persistent version.

What is Your Objective?

- Are you trying to:
 - Let the data tell you what relationship is contained within it?
 - **Unsupervised learning**
 - Trying to predict a continuous variable?
 - Supervised learning - **regression**
 - Trying to predict a discrete variable?
 - Supervised learning - **classification**