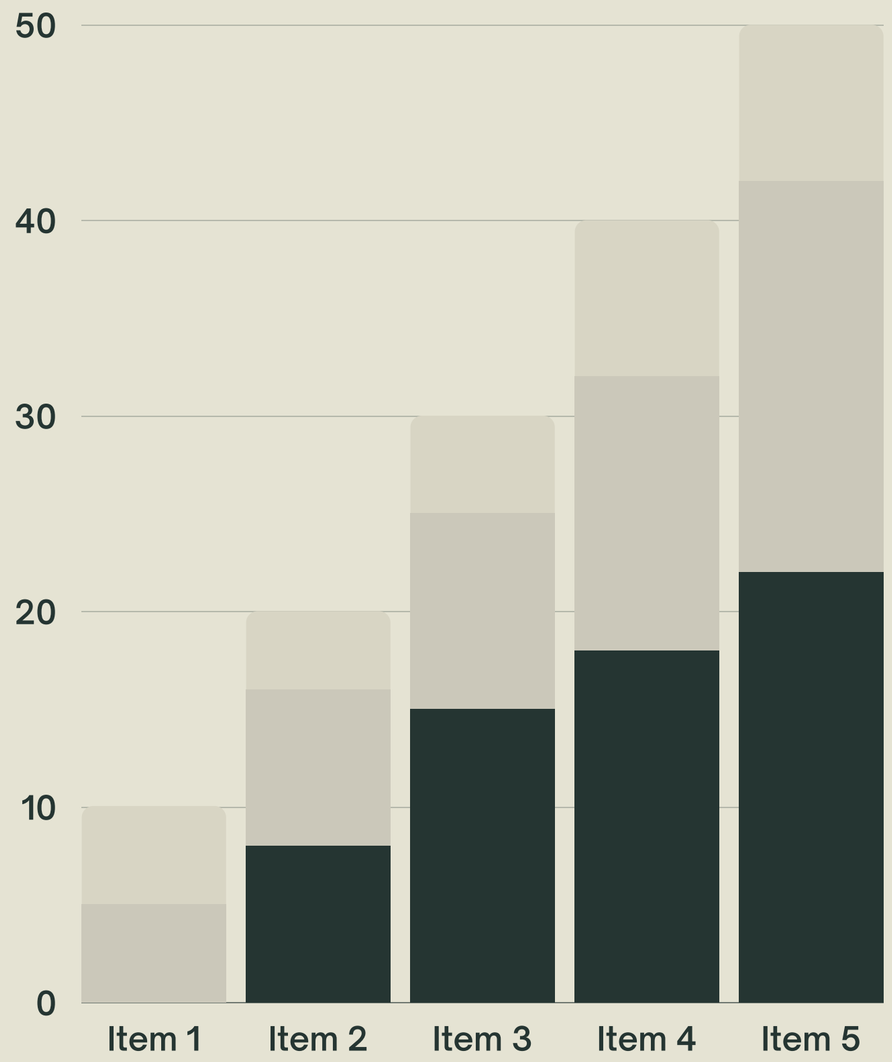
IMDb

# REVIEW RATINGS PREDICTOR

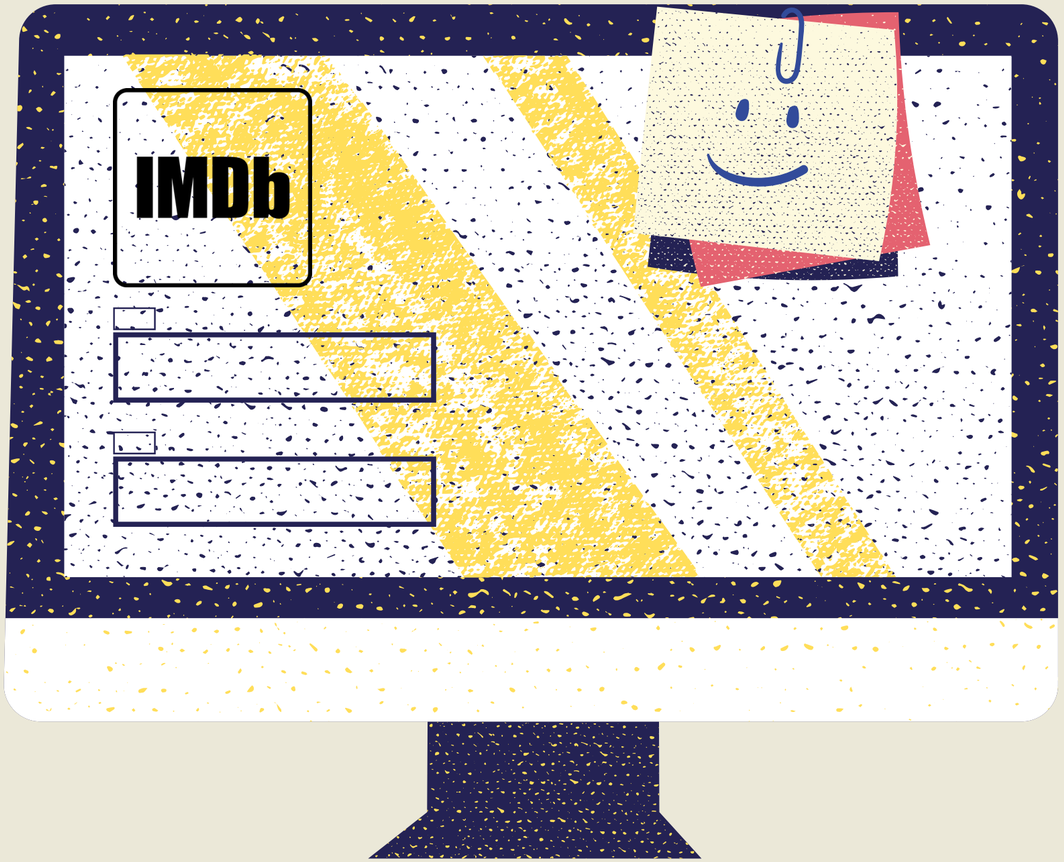



**Hung Pham**
Data Science Student - BrainStation

# Why?

# Getting data

- Hand scraping data from IMDb.
- 80k reviews from 150 movies

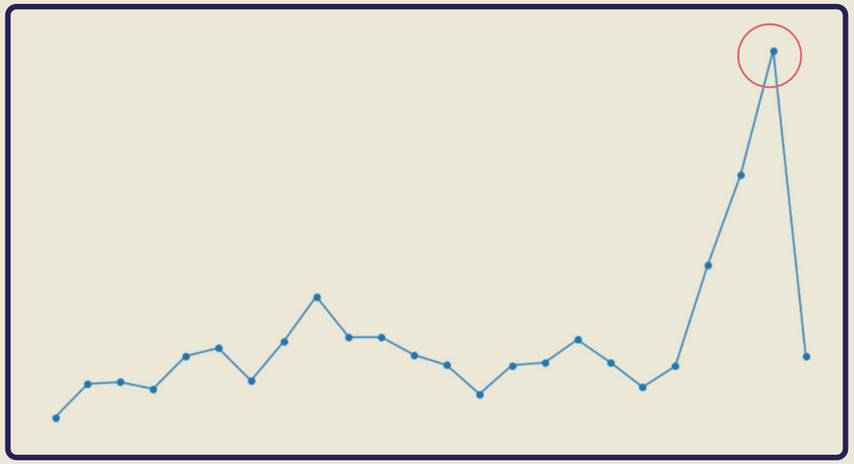



Se

BeautifulSoup

# Preparing data

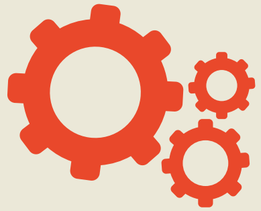
- Cleaning, removing duplicates, empty values...
- Exploring
- Transform text to numerical data



?

pandas

seaborn



# Modeling

- Logistic regression
- Random Forest
- Naive Bayer

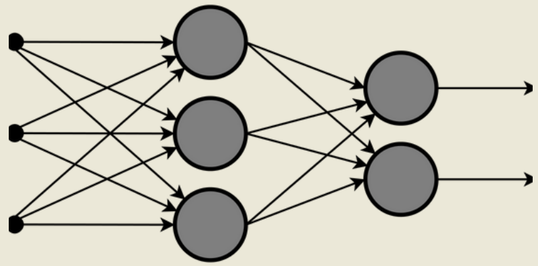
- Feedforward Neural Network

**65%**

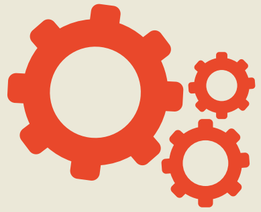
within error range of 1

**85%**

within error range of 2

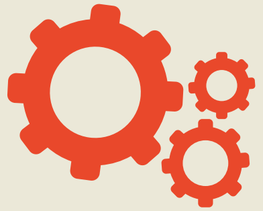


# Web application



# Contact


linkedin.com/in/hungpham89

github.com/hungpham89

ngochung3989@gmail.com