THE UNIVERSITY *of* EDINBURGH

Stock AI Assistant

2nd Hons Meeting

Finn Zhan Chen - 28th Novermber 2018

Project Overview

Objectives

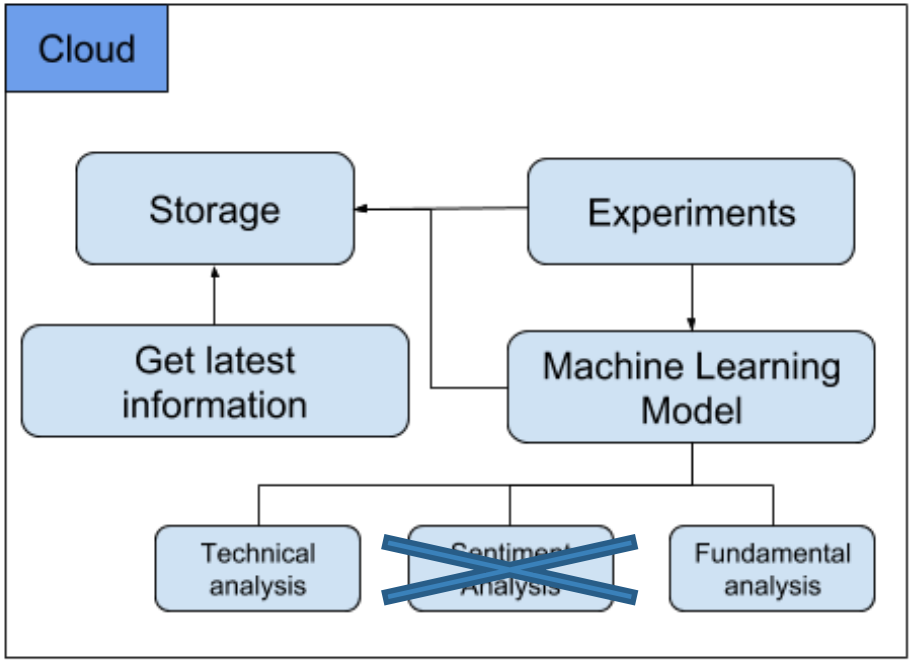
- Discover undervalued stocks (fundamental analysis)

- Discover high performing stocks (technical analysis)

- ~~Watch selected stocks (sentiment analysis)~~

- Overall score value

- Price prediction at certain time interval



Current Progress: System Design

Dataset ready to be used

	DATE	AMZN	AMZN.1	AMZN.2	AMZN.3	AMZN.4	AMZN.5	AMZN.6	AMZN.7	AMZN.8	...	AMZN.49	AMZN.50	AMZN.51	AMZN.52	AMZN.53
0	01/01/2017	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	...	NaN	NaN	NaN	NaN	NaN
1	02/01/2017	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	...	NaN	NaN	NaN	NaN	NaN
2	03/01/2017	753.67	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	...	NaN	NaN	NaN	NaN	NaN
3	04/01/2017	757.18	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	...	NaN	NaN	NaN	NaN	NaN
4	05/01/2017	780.45	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	...	NaN	NaN	NaN	NaN	NaN
5	06/01/2017	795.99	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	...	NaN	NaN	NaN	NaN	NaN
6	09/01/2017	796.92	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	...	NaN	NaN	NaN	NaN	NaN
7	10/01/2017	795.90	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	...	NaN	NaN	NaN	NaN	NaN
8	11/01/2017	799.02	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	...	NaN	NaN	NaN	NaN	NaN
9	12/01/2017	813.64	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	...	NaN	NaN	NaN	NaN	NaN
10	13/01/2017	817.14	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN		NaN	NaN	NaN	NaN	NaN

Current Progress: System Design

58 indicator explained (subset shown below)

Share Price	EBITDA	Current Assets	Revenues	Net Income from Discontinued Op.	Total Noncurrent Assets	Total Noncurrent Liabilities	Receivables	Current Liabilities	Treasury Stock	Cash From Financing Activities
Common Shares Outstanding	Interest expense, net	Net PP&E	COGS	Net Profit	Total Assets	Total Liabilities	Cash From Operating Activities	Total Equity	Retained Earnings	Net Change in Cash
Avg. Basic Shares Outstanding	Abnormal Gains/Losses	Intangible Assets	SG&A	Dividends	Short term debt	Preferred Equity	Net Change in PP&E & Intangibles	Depreciation & Amortisation	Equity Before Minorities	Free Cash Flow
Avg. Diluted Shares Outstanding	Income Taxes	Goodwill	R&D	Cash and Cash Equivalents	Accounts Payable	Share Capital	Cash From Investing Activities	Change in Working Capital	Minorities	Gross Margin

Current Progress: System Design

Object oriented approach to data

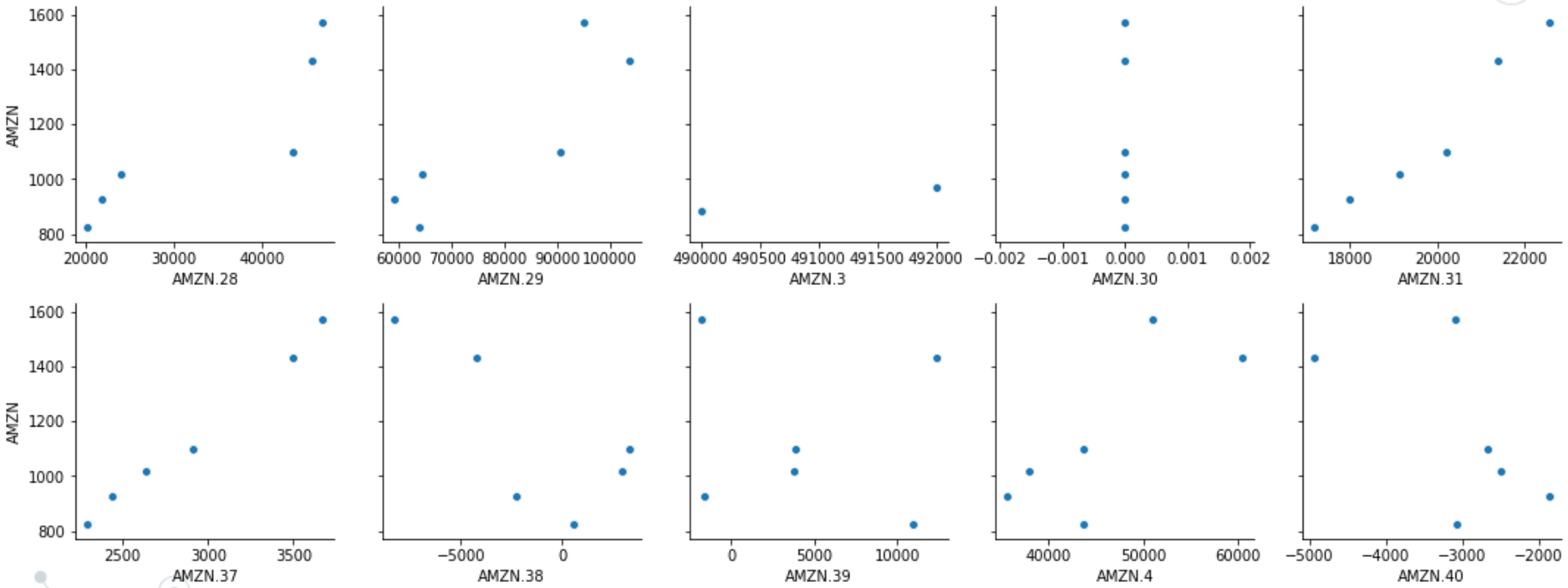
```
class Company(object):
    def __init__(self, name, pdframe):
        self.name = name
        self.share_prices = pdframe[name]
        self.converted_dates = self.preprocess_dates_raw(pdframe["DATE"])
        self.indicator_names_dict = {
            name + ".1" : "Common Shares Outstanding",
            name + ".2" : "Avg. Basic Shares Outstanding",
            name + ".3" : "Avg. Diluted Shares Outstanding",
            name + ".4" : "Revenues",
```

Current Progress: System Design

Share price visualisation

Current Progress: System Design

Indicator's correlation to share price (indicator are from quarterly results, hence small datapoints)



Current Progress: System Design

Indicator's correlation to share price in numbers

1.000 AMZN.57: Market Capitalisation
0.999 AMZN.58: Enterprise Value
0.995 AMZN.37: Depreciation & Amortisation
0.986 AMZN.36: Total Equity
0.976 AMZN.31: Share Capital
0.967 AMZN.33: Retained Earnings
0.949 AMZN.19: Net PP&E
0.936 AMZN.10: Interest expense, net
0.925 AMZN.22: Total Noncurrent Assets
0.915 AMZN.6: SG&A
0.913 AMZN.23: Total Assets
0.904 AMZN.9: EBITDA
0.892 AMZN.17: Receivables
0.884 AMZN.28: Total Noncurrent Liabilities
0.866 AMZN.29: Total Liabilities

0.568 AMZN.25: Accounts Payable
0.508 AMZN.16: Cash and Cash Equivalents
0.497 AMZN.46: Operating Margin
0.269 AMZN.50: Current Ratio
0.264 AMZN.51: Liabilities to Equity Ratio
0.235 AMZN.41: Cash From Investing Activities
0.196 AMZN.53: EV / EBITDA
0.129 AMZN.56: Operating Income / EV
-0.033 AMZN.43: Net Change in Cash
-0.119 AMZN.39: Cash From Operating Activities
-0.183 AMZN.42: Cash From Financing Activities

Current Progress: Next Steps

LSTM Data Preparation

1. Transform the time series into a supervised learning problem
2. Transform the time series data so that it is stationary.
3. Transform the observations to have a specific scale.

LSTM Stock Prediction

1. Parameter optimisation
2. Sliding Window vs Vector Auto Regression
3. One-step vs Multi-step forecasting
4. Univariate (Share price only) vs Multivariate Time series (Share price and other indicators)

Current Progress: Next Steps

Evaluation

1. Compare to baseline models like Persistence Model Forecast
2. Training, validation, and testing set

Industry specific performance

1. 100+ companies with from 13 different industries such as Utilities, Construction, and Manufacturing
2. Use this to help predicting fundamental and technical performance

Timeline

Gantt Chart

Semester 1:

- Focus on full-time job hunt
- Focus on planning Hons project
- Focus on Masters application

Semester 2:

- Minimal work from other courses, can devote almost full-time for Hons project