Analysis of stock market recommendations using computer vision

In partial fulfillment of the requirements of the degree of
M.Tech. in Data Science

A. B. Arthur[1]   J. Doe[2]

[1]Faculty of Physics
Very Famous University

[2]Faculty of Chemistry
Very Famous University

April 27, 2022

What are we going to learn?

Part I - Data science: the industrial perspective

Need of data science in industries
Data Science - Industries vs Consumer driven
Enabling technologies for data science

Need of data science in industries

"Why do we need data science when most of our modern devices are computer controlled and heavily automated?"

Part I - Data science: the industrial perspective

Need of data science in industries
Data Science - Industries vs Consumer driven
Enabling technologies for data science

Need of data science in industries

"Why do we need data science when most of our modern devices are computer controlled and heavily automated?"

1. Voluminous data generated from user metrics helps us optimise various processes
2. Helps personalise Human Machine Interaction (HMI)
3. Helps understand data whose volume, velocity and variety etc. can't be processed and understood easily in terms of patterns and trends

Part I - Data science: the industrial perspective

Need of data science in industries
Data Science - Industries vs Consumer driven
Enabling technologies for data science

Data Science - Industries vs Consumer driven

"Which of the data science domains: industries OR consumer driven applications deserves more focus"

Part I - Data science: the industrial perspective

Need of data science in industries
Data Science - Industries vs Consumer driven
Enabling technologies for data science

Data Science - Industries vs Consumer driven

"Which of the data science domains: industries OR consumer driven applications deserves more focus"

1. Large industries produce thousands of terabytes of data as opposed to few gigabytes of data for small industries or home based solutions over a year.

Part I - Data science: the industrial perspective

Need of data science in industries
Data Science - Industries vs Consumer driven
Enabling technologies for data science

Data Science - Industries vs Consumer driven

"Which of the data science domains: industries OR consumer driven applications deserves more focus"

1. Large industries produce thousands of terabytes of data as opposed to few gigabytes of data for small industries or home based solutions over a year.

Fun fact!!

A smart home generates about 1 GB of data per week ($\approx$ 200 MB/day) while a simple processing rack of a manufacturing plant may use up to 1 TB of memory.

Part I - Data science: the industrial perspective

Need of data science in industries
Data Science - Industries vs Consumer driven
Enabling technologies for data science

Data Science - Industries vs Consumer driven

"Which of the data science domains: industries OR consumer driven applications deserves more focus"

1. Large industries produce thousands of terabytes of data as opposed to few gigabytes of data for small industries or home based solutions over a year.

Fun fact!!

A smart home generates about 1 GB of data per week ($\approx$ 200 MB/day) while a simple processing rack of a manufacturing plant may use up to 1 TB of memory.

2. Industries use wide varied ML algorithms for analysing and acting on collected data as opposed to dedicated algorithms and data collecting regimes for small industries

3. Industrial applications of data science have grown exponentially with the advent of Industry 4.0 (IoT - connected devices and machines)

Part I - Data science: the industrial perspective

Need of data science in industries
Data Science - Industries vs Consumer driven
Enabling technologies for data science

Data Science - Industries vs Consumer driven (contd.)

4. Applications in industries are delay sensitive, while in homes or small consumer driven solutions they are not.

5. Harsh environments are prominent in factories and industries while the same may not be true for home applications.

6. Industrial applications are mission critical while small scale data science solutions may not pertain to some mission.

Part I - Data science: the industrial perspective

Need of data science in industries
Data Science - Industries vs Consumer driven
Enabling technologies for data science

Data acquisition

A typical factory or industry employs hundreds of sensors at various places.
The data from these sensors need to be efficiently communicated to processing
or storage hubs.[**Dai˙2019**]
In general (and unfortunately), increased data rates usually correspond to
decreased data range.

Part I - Data science: the industrial perspective

Need of data science in industries
Data Science - Industries vs Consumer driven
Enabling technologies for data science

Data acquisition

A typical factory or industry employs hundreds of sensors at various places.
The data from these sensors need to be efficiently communicated to processing
or storage hubs.[**Dai˙2019**]
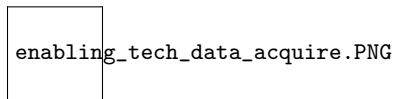In general (and unfortunately), increased data rates usually correspond to
decreased data range.

enabling_tech_data_acquire.PNG

Figure: Bandwidth vs Coverage for a variety of technologies

Part I - Data science: the industrial perspective

Need of data science in industries
Data Science - Industries vs Consumer driven
Enabling technologies for data science

Data preprocessing, storage and communications

- Data preprocessing - Data consists of heterogeneous data points which could be erroneous and noisy as well as redundant.

Part I - Data science: the industrial perspective

Need of data science in industries
Data Science - Industries vs Consumer driven
Enabling technologies for data science

Data preprocessing, storage and communications

- Data preprocessing - Data consists of heterogeneous data points which could be erroneous and noisy as well as redundant.
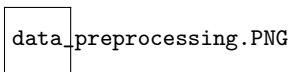
data_preprocessing.PNG

Figure: Data preprocessing techniques

Part I - Data science: the industrial perspective

Need of data science in industries
Data Science - Industries vs Consumer driven
Enabling technologies for data science

Data preprocessing, storage and communications

- Data preprocessing - Data consists of heterogeneous data points which could be erroneous and noisy as well as redundant.
- Storage - It consists of a *network* of interconnected storage devices or platforms. Almost all common DBMS software is an ideal choice for data storage.

Part I - Data science: the industrial perspective

Need of data science in industries
Data Science - Industries vs Consumer driven
Enabling technologies for data science

Data preprocessing, storage and communications

- Data preprocessing - Data consists of heterogeneous data points which could be erroneous and noisy as well as redundant.
- Storage - It consists of a *network* of interconnected storage devices or platforms. Almost all common DBMS software is an ideal choice for data storage.

database_logos.PNG

Figure: Commonly used DBMS software[**dbms˙pic**]

Part I - Data science: the industrial perspective

Need of data science in industries
Data Science - Industries vs Consumer driven
Enabling technologies for data science

Data analytics

- Data analytics - To analyse processed data (so as to take subsequent actions on the same) using any of the methods - 1) Statistical modeling schemes, 2) Data mining schemes, 3) Machine learning schemes and 4) Data visualization.

Part I - Data science: the industrial perspective

Need of data science in industries
Data Science - Industries vs Consumer driven
Enabling technologies for data science

Data analytics

- Data analytics - To analyse processed data (so as to take subsequent actions on the same) using any of the methods - 1) Statistical modeling schemes, 2) Data mining schemes, 3) Machine learning schemes and 4) Data visualization.
  Luckily, business value increases with increasing complexity of an approach.

Part I - Data science: the industrial perspective

Need of data science in industries
Data Science - Industries vs Consumer driven
Enabling technologies for data science

Data analytics

- Data analytics - To analyse processed data (so as to take subsequent actions on the same) using any of the methods - 1) Statistical modeling schemes, 2) Data mining schemes, 3) Machine learning schemes and 4) Data visualization.
  Luckily, business value increases with increasing complexity of an approach.

enabling_tech_data_anal.PNG

Figure: Business value vs complexity for various data analysis approaches

Part I - Data science: the industrial perspective

Need of data science in industries
Data Science - Industries vs Consumer driven
Enabling technologies for data science

References I

Part I - Data science: the industrial perspective

Need of data science in industries
Data Science - Industries vs Consumer driven
Enabling technologies for data science

Thank You