## Project roadmap

Our report is generated from Latex. The .docx file is a converted copy. please see source codes in: Group7\_Report[Latex].zip

We are proposing two product lines **Hawkeye** and **Proofreader** in part 3 and 4, both using different models and datasets from part 1 and 2. Thus our notebook will be mainly divided into 3 blocks.

1. **\*\*Section part 1 and part 2\*\***

Contains the code for part 1 and 2, where we follow the project requirements and test self-trained (part 1) and pre-trained models on various labelled and un-labelled datasets.

2. **\*\*Section part 3&4: Hawkeye\*\***

\*Contains the code for the product: Hawkeye\*

a. Use entity-based sentiment analysis model on Twitter data and plot changes in sentiment over time.​

b. For days with large sentiment changes, perform topic modelling to identify key drivers of sentiment.

3. **\*\*Section part 3&4: Proofreader\*\***

\*Contains the code for the product: Proofreader\*

a. Calculate sentiment response scores based on the ratio of positive to negative replies. ​

b.Link the sentiment of replies to the contents of the main tweet via a regression model.

## Folder content

/content/drive/MyDrive/UofT\_MEng/MIE1624/Project/submit/

├── config

├── data

│ ├── part\_1

│ ├── part\_2

│ └── part\_3

│ ├── twitter\_all

│ │ └── part3\_twitter\_1000sample\_per\_day

│ ├── twitter\_elon\_musk

│ │ └── part3\_elon\_twitter\_scores\_1000\_sample\_replies

│ └── twitter\_Ukraine\_Official

└── models