write the objective, implementation and applications of your idea. you must include the details about the final result of your idea and how you plan on building upon your idea after the campaign is over

Description:

Objective:

The aim of our project is to create a service for an investor, in order to track the news pertaining to the assets in their portfolio and notify them about its effects, quickly and effectively.

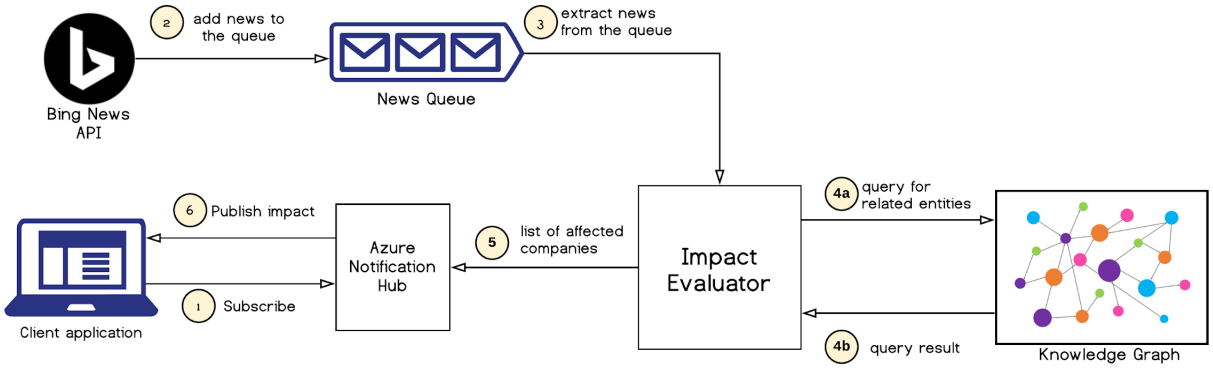
An investor invests in multiple stocks for better returns. However, it is very difficult for them to track the effect of some news on the stock prices of a company, it's products, its subsidiaries and other companies it carries out trade with.

We aim to find the effect of a news article on the stock price of a company to which the news is related. We will also try to give a prediction on which subsidiary, competitor and other dependent companies will stand to gain or lose from this event.

This application will help the investor get better insights from news articles and hence do a better trade.

Implementation:

We are going to use four sub-modules for implementing this project. We have further explained them in detail as listed below:



1. **Fetching data:** Fetching latest new articles using Azure's 'Bing news search API' at regular intervals, and pushing them into Azure's service bus.

2. **Knowledge Graph:** Knowledge Graphs store facts in the form of relations between different entities. In our case, the entities will be the companies and their subsidiaries. We will maintain two knowledge graphs,

   a. (Organisation A → Organisation B)

   b. (Organisation A → Commodity X)

   Note: A commodity influences an organisation's ability to generate profit or loss.
   Eg: If OPEC updates oil prices, then it will affect other petroleum based industries

3. **Impact evaluator:** Impact evaluator contains a NLP tool which will process the received news article using sentiment analysis. The NLP tool will output the organisation, related commodities and the relevant sentiment of the news article(either positive or negative) and output a score depending on the sentiment of the news article. When we analyse a news article we will,
    a. We use knowledge graph to find its effect on related companies
    b. Identify which companies are to be affected
    c. Normalize impact score of companies between a fixed range
    d. If the normalized output value crosses a threshold, it means that the impact on the company is substantial
4. **Client application:** The user will subscribe to the companies which he/she is interested in. Whenever any message related to the company is published, the user is notified about it.

Idea application

1. Keep investors up-to-date on the events affecting their portfolio.
2. Investors will not need to analyse each and every news article available on the internet
3. It will help budding investors to analyse and understand the market better.

Result

Investors will now get analysed results related to their subscribed companies instantaneously. This will enable investors to take decisions related to put or call options on his stocks, thus reducing their loss margin if any.

Future Scope:

1. We can incorporate tweets as a news source and hence broaden the domain of the service.
2. We can also incorporate the quarterly financial reports of companies as a news source.

Technology stack:

1. **Bing News Search:** The Bing news Search API will be used to obtain financial news at regular intervals to facilitate the prediction of impacts on establishments
2. **Azure Machine Learning;** Azure Machine Learning will be used to host machine learning pre-trained models that will be used to calculate which organisations a particular news will affect and how it will affect the organisations.
3. **Azure notification hub:** Azure notification will be used to provide notification about subscribed companies to the users
4. **Azure Computing(Virtual Machines):** A linux Virtual Machine will be used to host a front end of the project https://docs.microsoft.com/en-us/azure/app-service/app-service-web-get-started-nodejs

5. **Service Bus:** To coordinate the activities between all other Azure services
6. **Azure Storage:** To store data necessary for the application

Objective:

The aim of our project is to create a service for an investor, in order to track the news pertaining to the assets in their portfolio and notify them about its effects, quickly and effectively.

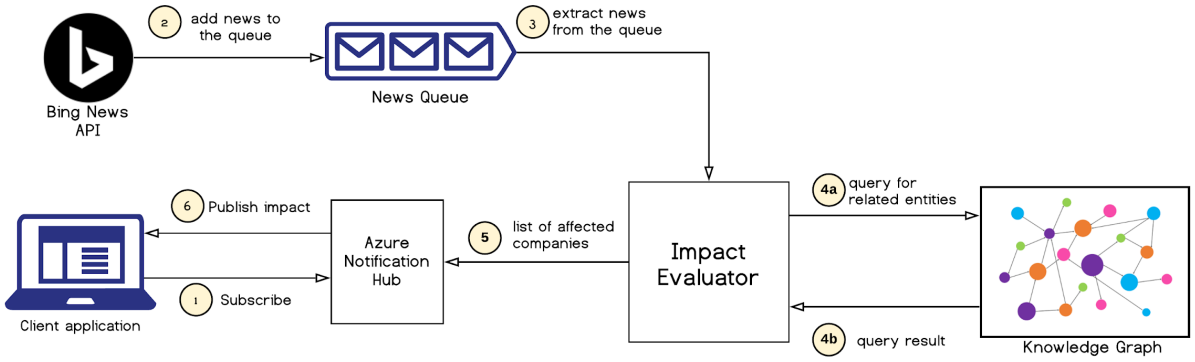
An investor invests in multiple stocks for better returns. However, it is very difficult for them to track the effect of some news on the stock prices of a company, it's products, its subsidiaries and other companies it carries out trade with.

By using Machine learning we aim to find the sentiment of a news article and try to predict its effect on the stock price of a company to which the news is related. We will also try to give a prediction on which subsidiary will either gain or lose from this event.

This application will help the investor get better insights from news articles and hence do a better trade.

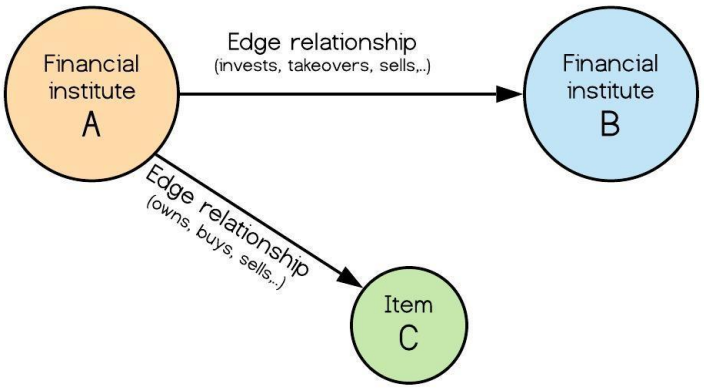
Implementation:

We are going to use four sub-modules for implementing this project. We have further explained them in detail as listed below:



1. Fetching data: Fetching latest new articles from publicly available api at regular intervals, and pushing into the news articles queue.

2. Knowledge Graph: A knowledge base is any collection of information. A Knowledge Graph is a structured Knowledge Base. Knowledge Graphs store facts in the form of relations between different entities.



As we can see from the above example figure,
We will store relationships between different Organisations as seen above

`(Organisation A → Organisation B)`

We also store relationships between different institutes and commodities or items

`(Organisation A → Commodities C)`

Note: A commodity influences an organisation's ability to generate profit or loss.
Eg: If OPEC updates oil prices, then it will affect other petroleum based industries
These relationships are extracted by using Natural Language Processing (NLP) which is applied over news articles to find out all the subject-predicate-object relationships.

By extracting facts from a knowledge base and representing these facts in the form of entities and relations, a knowledge graph claims to have an understanding of the information. Many knowledge graphs currently represent extracted facts in the form of **Subject-Predicate-Object (SPO)** triples which are in line with the standard prescribed by RDF (Resource Description Framework).

Here are the steps for creating a Knowledge Graph:

1. **Knowledge Extraction:**
    a. **Extraction of SPO triples** (facts) from text.
        i. Uses **Natural Language Processing (NLP)** techniques like **dependency parsing**. NLP is the backbone of forming a good knowledge graph from textual information.
    b. **Entity Recognition & Linking:**
        i. The extraction of a single word entity from a sentence is not a tough task. We can easily do this with the help of **parts of speech (POS)** tags. The **nouns** and the **proper nouns** would be our **entities**.

ii. However, when an entity spans across multiple words, then **POS tags** alone are not sufficient. We need to parse the **dependency tree** of the sentence.
iii. We can do this using the **spaCy** library.

2. **Graph Construction:**
    a. Removing ambiguities and storing the SPO triples on a Graph Database. Here, the fact represented as an SPO conveys that the Subject is related to the Object through the relationship described by the Predicate.
    b. Another step here would be to process the Graph to achieve things like filling missing links, clustering entities, etc.

3. **Impact evaluator:** Impact evaluator contains a NLP tool which will process the received news article using sentiment analysis. The NLP tool will output the organisation, related commodities and the relevant sentiment of the news article(either positive or negative) and output a score depending on the sentiment of the news article. When we analyse a news article we will,
    a. We use knowledge graph to find its effect on related companies
    b. Identify which companies are to be affected
    c. Normalize impact score of companies between a fixed range
    d. If the normalized output value crosses a threshold, it means that the impact on the company is substantial
4. **Client application:** The user will subscribe to the companies which he/she is interested in. Whenever any message related to the company is published, the user is notified about it.

Technology stack:

1. **Bing News Search:** The Bing news Search API will be used to obtain financial news at regular intervals to facilitate the prediction of impacts on establishments
2. **Azure Machine Learning;** Azure Machine Learning will be used to host machine learning pre-trained models that will be used to calculate which organisations a particular news will affect and how it will affect the organisations.
3. **Azure notification:** Azure notification will be used to provide notification about subscribed companies to the users
4. **Azure Computing(Virtual Machines):** A linux Virtual Machine will be used to host a front end of the project
5. **Service Bus:** To coordinate the activities between all other Azure services
6. **Azure Storage:** To store data necessary for the application.

Result

Investors will now get analysed results related to their subscribed companies instantaneously. This will enable investors to take decisions related to put or call options on his stocks, thus reducing their loss margin if any.