GRADUATE CERTIFICATE: Intelligent Reasoning Systems (IRS)
PRACTICE MODULE: Project Proposal

Date of proposal:

16 March 2024

Project Title:

Intelligent Sentiment Trading Recommender

Sponsor/Client: *(Name, Address, Telephone No. and Contact Name)*

Institute of Systems Science (ISS) at 25 Heng Mui Keng Terrace, Singapore
NATIONAL UNIVERSITY OF SINGAPORE (NUS)
Contact: Mr. GU ZHAN / Lecturer & Consultant
Telephone No.: 65-6516 8021
Email: zhan.gu@nus.edu.sg

Background/Aims/Objectives:

1. *Background*

- In 2023, the Fed had implemented a series of interest rate hikes throughout the year to slow down inflation which had risen well above their goal of 2%. **With the rising cost of living caused by inflation, it becomes inevitable that a more efficient trading analysis tool is required to help investors**. Whilst traditional analysis of stocks focuses on hard numbers, markets are also driven by emotions. Investor sentiment including fear, greed, and optimism, significantly affects stock prices. **Sentiment analysis helps to gauge these emotions by analyzing news articles, social media posts, and other financial data to identify shifts in investor confidence before they are reflected in the actual price movements**. This allows investors to potentially make informed decisions ahead of the curve. Sentiment analysis helps to quantify the "softer" side of the market, providing a more comprehensive view.

- There are multiple available pre-trained Natural Language Processing (NLP) models in the market that could facilitate sentiment tagging. Some examples are **FinBERT** (ProsusAI, 2022)[1], which focuses solely on financial news articles and **X-roBERTa-base for Sentiment Analysis** (X formerly known as Twitter) which looks into the social media tweets analysis (Fan et al., 2019)[2]. NLP libraries like spaCy, NLTK have made sentiment analysis more possible and accurate.

- While there are plenty of common sentiment analysis models, **these models tend to face difficulty in negation handling, domain dependence, spam detection, and ambiguity in the form of abbreviations or sarcasm.** Improved tools like **PyFin-sentiment** appear to

better reflect “market sentiments” rather than “general sentiments”[3] (but fail to work as issues were not debugged by the code owner).

- On the financial side, there are alot of algorithm-based trading bots and **these bots execute trades based on data trends and a set of predefined rules**. **They do not consider the sentiments of the market, as well as any social, economic, geopolitical events**, etc (Medha et al.,2021)[4].

- **From a global point of view, in 2024, the emotion recognition and sentiment analysis software market size is projected to increase by USD 797.17 million, at a CAGR of 14.15% between between 2023 and 2028**[5]. Specific to the finance applications, the team has identified that there are increasing demands in finance markets for market insights and sentiment extractions. **The NLP in global finance markets is estimated to be growing at a CAGR of 27.6%, worth USD18.8 billion by 2028.** (Marketsandmarkets, 2023)[6]. Key players in the market include Microsoft, Google, IBM, AWS, Oracle, etc.

- On the social media front, Reddit, out of top 100 groups, 3 of which are related to finance/crypto with a total of 29M subscribers[7]. Of which, **r/WallStreetsBets was popularized from the Gamestop (GME) short squeeze saga in 2021 and even the traditional media recognizes the community effect on sentiment trading[8].**

- This project concept continues to be attractive for businesses to retail consumers as both groups are always keen on profit-generating ideas. Online bots have poor reviews[9] and there seems to be a lack of widely-discussed existing recommenders out in the market.

2. *Final Goal:*

- **To design an intelligent sentiment trading recommender that aims to provide market entry suggestions to the user.**
- The software will push notifications based on the user's subscribed company, industry and favorites.
- To push a notification signal for a recommended action (e.g strong buy), the underlying model will label news articles when they are first released. If it hits the right combination of articles and sentiment polarities with the right weightage between traditional finance (tradfi) and social finance (socialfi), a notification will be sent.
- The notification will consist of information (see *Roadmap*) to educate the user on the market sentiments before making any informed trading/investment decisions.

3. *Milestones/Objectives:*

- To leverage on **pre-trained NLP models with financial news' texts/social media inputs and categorize into different polarities of sentiments and company name/industry.**
- **To clean and prepare financial data from stock prices sources, converting into h-o-h or d-o-d %change time series**.
- **To train a model using supervised classification machine learning techniques using pre-trained NLP sentiment outputs and pre-processed financial data.** The model will be trained by breaking up different periods of trading years as train, validate and test dataset. The output will provide **recommendations to users to make investment decisions.**

4. *Roadmap:*

- **Phase 1 MVP:**
  A prediction software that is able to take user's inputs, read market sentiments and provide suggestions on trading actionables. The outputs from the software contain:
  a. Types of recommended actions (Strong Buy, Buy, Neutral, Sell, Strong Sell)
  b. Company/Sector
  c. Expected profit taking period (e.g Day 0, Day +1)
  d. A link to the article related to the stock.

  Phase 2 (Not included in this phase)
- Provide a confidence level for each recommendation to the users.
- Allow the user to rate the sentiments on the articles and feedback to the model to self-learn.
- To include Technical Analysis features in model training.
- To use image analysis machine learning software to scan for chart shapes and patterns to suggest entry points.

References:

1. ProsusAI. (2022). ProsusAI/finBERT: Financial sentiment analysis with bert. GitHub. https://github.com/ProsusAI/finBERT
2. Fan, R., Talavera, O., & Tran, V. (2019). *(PDF) Social Media Bots and stock markets*. Social media bots and stock markets. https://www.researchgate.net/publication/331639758_Social_media_bots_and_stock_markets
3. Wilksch, M., & Abramova, O. (2023a). Pyfin-sentiment: Towards a machine-learning-based model for deriving sentiment from financial tweets. International Journal of Information Management Data Insights, 3(1), 100171. https://doi.org/10.1016/j.jjimei.2023.100171
4. Medha Mathur, M., Mhadalekar, S., Mhatre, S. S., & Mane, V. (2021, August 1). Algorithmic Trading Bot. Researchgate. Retrieved August 3, 2024, from https://www.researchgate.net/publication/353770459_Algorithmic_Trading_Bot
5. Emotion Recognition and Sentiment Analysis Software Market Analysis North America, Europe, APAC, South America, Middle East and Africa - US, China, Japan, UK, Germany - Size and Forecast 2024-2028 https://www.technavio.com/report/emotion-recognition-and-sentiment-analysis-software-market-industry-analysis
6.MarketsandMarkets. (2023, April). *NLP in Finance Market Size & Share Analysis - Industry Research Report - Growth Trends*. MarketsandMarkets. https://www.marketsandmarkets.com/Market-Reports/nlp-in-finance-market-21737879.html
7. Top communities on reddit - page 1. (n.d.-c). https://www.reddit.com/best/communities/1/
8. The GameStop Episode: What Happened and What Does It Mean?. Cato.org. (n.d.). https://www.cato.org/cato-journal/fall-2021/gamestop-episode-what-happened-what-does-it-mean#game-stop-goes-crazy-in-an-interesting-way

8. 6 live sentiment analysis trading bots using Python | Udemy. (n.d.-a). https://www.udemy.com/course/sentiment-trading-python/

Requirements Overview:

- **Research ability**: To extract/interpret financial market sentiments based on financial/social media/event-driven news outlets.
- **Programming ability**: To perform sentiment analysis of stock's traditional and social media data.
- **System integration ability**: To combine financial news' sentiment and stock price actions to give user overall advices to make informed trading/investing decisions

Resource Requirements (please list Hardware, Software and any other resources)

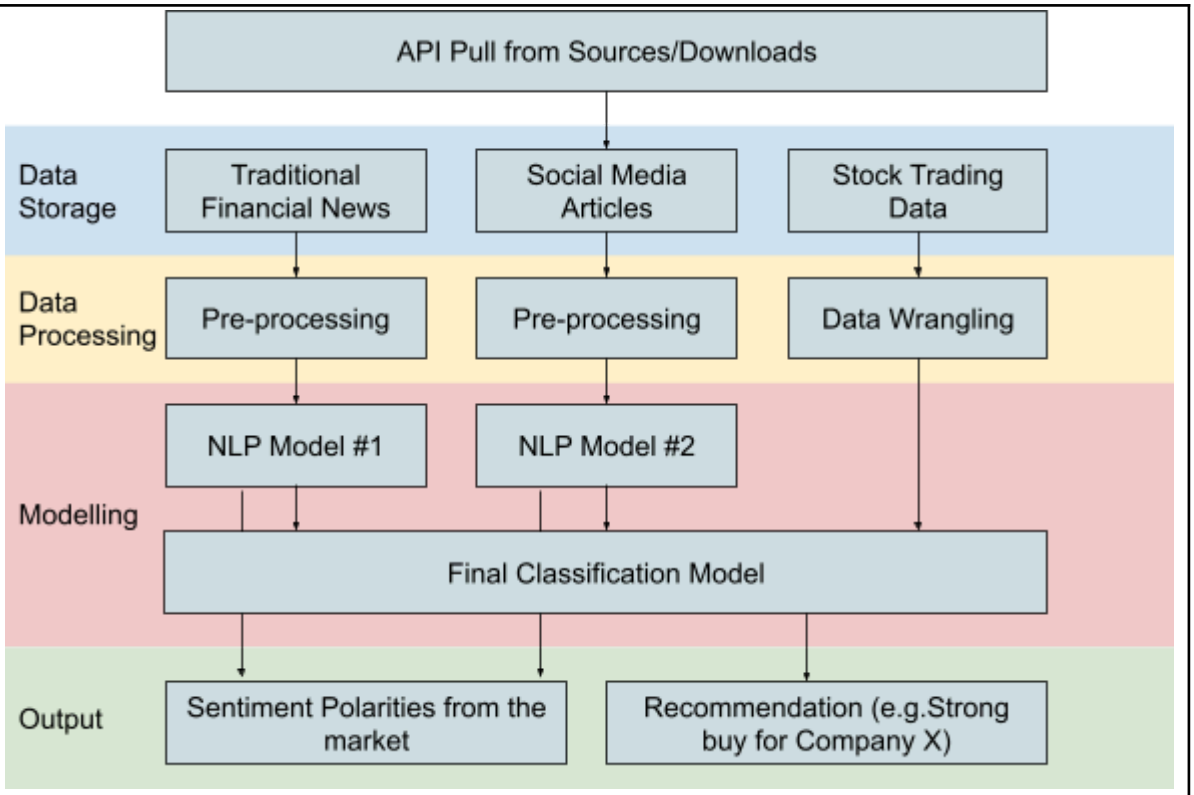
1. **Data Collection/Preparation**
    a. **News**
    - Open News API, PRAW, python-twitter
    - Social Media API (Twitter / X , Facebook / Meta, Reddit, StockTwits, LinkedIn)
    - Reuters, CNBC, Yahoo Finance News, The Business Times, Bloomberg

    b. **Stock screener**
    - FinViz
    - Yahoo Finance / Yahoo Finance API
    - Other (Alpha Vantage, Quandl, IEXCloud, Tiingo)

2. **Verification bots/Teaching assistant**
    - ChatGPT/Prof guidance

3. **System Design**
    a. **Hardware considerations:**
    - GPU, Laptop

    b. **Software considerations:**
    - spaCy, NLTK, Gensim
    - BERT, FinBERT, roBERTa, DistilBERT, LSTM (long short-term memory networks), VADER
    - sklearn, Decision Tree, KNN, SVM, RandomForest, Fuzzy Logic
    - Backtesting, Ta-lib, bt
    - Flask, Django, Bootstrap, jQuery, HTML, CSS, JS



Number of Learner Interns required: (Please specify their tasks if possible)

4 Learner Interns:

Learners	**Tasks**
Kenn Ng Ying Wee	Team Leader, NLP modelling (socialfi), Trading data wrangling, UI/UX, Video Editing, Report
Tan Kok Tong	NLP modelling (tradfi), Final Model compilation, UAT, Report
Lai Kah Hoe	NLP modelling (socialfi), UI/UX, UAT, Report
Yeo Li Ying	NLP modelling (tradfi), Trading data wrangling, Final Model compilation, Report

Methods and Standards:

Procedures	Objective	Key Activities
Requirement Gathering and Analysis	The team should meet with ISS to scope the details of project and ensure the achievement of business objectives.	1. Gather & Analyze Requirements 2. Define application design 3. Prioritize & Consolidate Requirements 4. Establish Functional Baseline
Technical Construction	· To develop the source code in accordance to the design. · To perform unit testing to ensure the quality before the components are integrated as a whole project	1. Setup Development Environment 2. Understand the System Context, Design 3. Perform Coding 4. Conduct Unit Testing (NLP, final classification model)
Integration Testing and acceptance testing	To ensure interface compatibility and confirm that the integrated system hardware and system software meets requirements and is ready for acceptance testing.	1. Prepare System Test Specifications 2. Prepare for Test Execution 3. Conduct System Integration Testing 4. Evaluate Testing 5. Establish Product Baseline
Acceptance Testing	To obtain ISS user acceptance that the system meets the requirements.	1. Plan for user acceptance testing 2. Conduct training for acceptance Testing 3. Prepare for Acceptance Test Execution 4. ISS Evaluate Testing 5. Obtain Customer Acceptance Sign-off
Delivery	To deploy the system (ISS standalone server) environment.	1. Software must be packed by following ISS's standard 2. Deployment guideline must be provided in ISS production (ISS standalone server) format 3. Production (ISS standalone server) support and troubleshooting process must be defined.

Team Formation & Registration

Team Name: SenTEAMent
Project Title (repeated): Intelligent Sentiment Trading Recommender

System Name (if decided):
Team Member 1 Name: Kenn Ng Ying Wee
Team Member 1 Matriculation Number: A0111692L
Team Member 1 Contact (Mobile/Email): +65 86864252 / ngyingwee92@hotmail.com
Team Member 2 Name: Yeo Li Ying
Team Member 2 Matriculation Number: A0292287R
Team Member 2 Contact (Mobile/Email): +65-94359687/yeoliying@gmail.com
Team Member 3 Name: Lai Kah Hoe
Team Member 3 Matriculation Number: A0292131N
Team Member 3 Contact (Mobile/Email): +65 98655569 / laikahhoe@gmail.com
Team Member 4 Name: Tan Kok Tong
Team Member 4 Matriculation Number:

A0155101Y
Team Member 4 Contact (Mobile/Email): +65 81610164 / tan1995_@hotmail.com
Team Member 5 Name:
Team Member 5 Matriculation Number:
Team Member 5 Contact (Mobile/Email):
Team Member 6 Name:
Team Member 6 Matriculation Number:
Team Member 6 Contact (Mobile/Email):

For ISS Use Only		
Programme Name:	**Project No:**	**Learner Batch:**
Accepted/Rejected/KIV:		
Learners Assigned:		
Advisor Assigned: Contact: Mr. GU ZHAN / Lecturer & Consultant Telephone No.: 65-6516 8021 Email: zhan.gu@nus.edu.sg		