

金融予測モデル結果の説明改善のための LLM を用いた文章生成手法の提案

Proposal for a story generation method using LLMs to improve the explanation of financial prediction modelling results

小手川 康太[†] 上田 健太郎[†] 細川 蓮[†] 諏訪 博彦[†] 小川 祐樹[‡]
 Kota Kotegawa Kentaro Ueda Ren Hosokawa Hirohiko Suwa Yuki Ogawa
 梅原 英一[§] 山下 達雄[¶] 坪内 孝太[¶] 安本 慶一[†]
 Eiichi Umehara Tatsuo Yamashita Kota Tsubouchi Keiichi Yasumoto

1. はじめに

将来の金融市場の予測は、投資リスクの管理や獲得利益の最大化のために非常に重要なテーマである。しかし、金融市場に参加する投資家の行動、経済指標、政治的要因など、複雑な要素が絡み合う環境であるため、極めて困難な課題である。近年では、将来の市場の予測に対して過去の価格情報だけでなく、非構造化データの活用が進んでいる。特に、テキストデータは予測に有効な情報源として、活用されている。例えば、ニュースのテキストデータを用いた研究 [1, 2, 3, 4, 5], ソーシャルメディアのテキストデータを用いた研究 [6, 7, 8, 9, 10] など、多様な情報源を対象に行われている。さらに、近年の情報技術の飛躍的な発展により、機械学習モデル、特にディープラーニングの進化が著しい。これらの技術は、高次元のデータから複雑なパターンを検出する性能が高いため、金融市場が生み出す膨大なデータを効果的に処理できることから、金融タスクと相性が良く、優れた精度を示している [11]。

ディープラーニングは高い予測精度を達成できる一方で、その予測プロセスがブラックボックス化しているという問題がある。つまり、モデルの予測がどのような根拠に基づいて予測を行っているのかが不透明であり、これは資産に関わる重要な意思決定を行う金融の領域において大きな問題である。この問題に対処するため、金融タスクにおける説明可能性 (Explainable AI, XAI) の重要性が高まっている [12]。説明可能な AI モデルは、信頼性の向上、倫理的な意思決定の促進など、多くの利点をもたらす。このように金融分野における XAI の重要性が認識される一方で、現状の XAI 技術には依然として多くの課題が存在する [13]。

現在、機械学習分野で広く用いられる XAI 手法である、LIME(Local Interpretable Model-agnostic Explanations)[14]

や SHAP(SHaplay Additive exPlanations)[15] などの手法は、個々の予測結果に対する入力特徴量の重要度を可視化することができる。しかし、このような XAI 手法は、人間が日常的に行う説明方法とは大きな乖離がある [16]。人間が日常的に行う説明方法は、理由や背景を含めて説明をするため、LIME や SHAP のような数値ベースの説明は、直感的に理解しにくい。例えば、「なぜこの株価が上昇したのか」という質問に対して、人間は「新製品が発表されてかつ、業績が好調である。」といった具体的な理由を述べる。一方、LIME や SHAP は、「特徴量 A の寄与度が 0.3, 特徴量 B の寄与度が 0.1」といった数値を提示することのみにとどまる。資産に関わる重要な意思決定を行う金融タスクにおいて、人間が理解しやすい流暢な文章による説明を行うことは重要である。

文章による説明を行う試みとして、近年、大規模言語モデル (Large Language Model, LLM) を活用した XAI の研究が注目を集めている [17]。LLM は自然言語処理タスクにおいて高い性能を示しており、複雑な情報を人間にとって理解しやすいかたちで表現することができる [18]。金融市場においては、説明として「新製品が発表されてかつ、業績が好調である。」のような予測の根拠となる出来事と、それがどれだけ影響しているかを示すことが重要である。

本研究では、この手法を金融タスクに応用し、従来の XAI 手法から取得された予測の根拠となる出来事から人間が自然に解釈可能な文章を生成する手法を提案する。本手法は、具体的に以下の点で先行研究を進展させることで、金融タスクにおける優れた説明の生成を実現する。

1. 金融タスク特化型の要因分析

LLM に入力する特徴量とその寄与度を抽出するために従来の XAI 手法を応用して、金融予測に影響を与えた出来事 (例: 業績, 市場動向) を特定し、それらの寄与度を算出する。

2. 重要度を考慮した説明生成

LLM を用いて説明を生成する際、各要因の重要度に基づいて説明の構造と内容を調整する。重要度の高い要因については詳細な説明を行い、重要度

[†] 奈良先端科学技術大学院大学, Nara Institute of Science and Technology

[‡] 東京都市大学, Tokyo City University

[§] 新潟国際情報大学, Niigata University of International and Information Studies

[¶] LINE ヤフー株式会社

の低い要因は省略するか、簡潔な説明を行う。これにより、最も影響力のある要因から順に、論理的かつ簡潔な説明を提供する。

3. 金融専門用語と一般用語の適切な使い分け

金融市場にはプロの投資家から、一般投資家まで様々なユーザが参加する。そのため専門知識レベルに応じて、専門用語と一般用語を適切に使い分け、ユーザに適した説明を生成する。

2. 関連研究

2.1 SHAP (SHapley Additive exPlanations)

SHAP は、機械学習モデルにおいて広く一般的に利用されている XAI 手法で、協力ゲーム理論における Shapley 値の概念を機械学習モデルの解釈に適用した手法である [15]。この手法の目的は、モデルの予測に対する各インスタンスの寄与度を定量化することである。特徴量 i の SHAP 値 ϕ_i は式 (1) で表される。

$$\phi_i = \sum_{S \subseteq F \setminus \{i\}} \frac{|S|!(|F| - |S| - 1)!}{|F|!} [f_x(S \cup \{i\}) - f_x(S)] \quad (1)$$

ここで、 F は全特徴量の集合、 S は特徴量 i を含まない部分集合、 $f_x(S)$ は特徴量集合 S を用いたモデルの予測値を表す。SHAP はモデルに依存しない手法であるため、あらゆる機械学習モデルに利用できるという利点がある。しかし、SHAP の解釈には専門知識が要求される点も課題の一つである。SHAP 値の正確な意味や、それらがどのようにモデルの決定に影響を与えているかを理解するには、手法の背後にある理論的基礎への深い理解が必要となる。これは、非技術者や一般のステークホルダーへの説明を困難にする可能性がある。本研究では SHAP によって得られた知見を LLM によって一般のステークホルダーが解釈可能なかたちで提供することで説明可能性の向上を目指す。

2.2 SHAP を用いた説明文章の生成手法

Zytek ら [19] は、複雑な機械学習モデルの説明を、より自然で人間が読みやすいかたちに変換するために、LLM を活用する研究を行っている。具体的には、SHAP で生成された特徴量と、その寄与度を、LLM を用いて自然言語の説明文に変換するタスクを行った。手法として、GPT-3.5 と GPT-4 を用いたゼロショットでの説明生成を試み、5 種類のプロンプトを設計して比較実験を行った。LLM への入力は、予測に用いた各データセットの中から 3 つのインスタンスを選び、それぞれに対して SHAP を用いて選択された最も寄与度の高い上位 3 つの特徴量を用いている。説明の評価手法として、正確性、

流暢さ、完全性、文脈認識、説明の長さの 5 つの指標を定義した。対象にした予測タスクは、学生の成績を予測するタスクと、住宅価格予測タスクの 2 種類であった。20 人を対象に行ったユーザ調査の結果、従来の XAI 手法による説明と比較して、LLM で生成された説明文の方が説明として好まれる傾向があることが示された。

しかし、この研究では金融タスクを対象にして説明文の生成を行っていない。金融市場では、説明として床面積や築年数のような数字では無く、「新製品が発表されてかつ、業績が好調である。」のような具体的な出来事による説明が求められる。そこで本研究では、中間の説明器によって有効的に特徴量を抽出することで、金融タスクに特化した説明の生成を行う手法を検討する。

2.3 LLM による株価変動予測と説明生成手法

Wang ら [20] は、LLM を利用して金融市場の動向予測と説明を、同時に行う手法について研究している。このフレームワークでは、金融テキストデータから株価変動の因子を LLM によって抽出し、抽出された因子と、その他のインスタンスによって株価の変動を予測している。この手法の優れた点は、テキストデータから重要な因子を抽出していることである。金融予測タスクにおいて LLM を用いて、テキストデータからキーワード、センチメントワードを抽出することで予測を行う手法は多数存在する。与えられた金融テキストに対して、キーワードは内容の概略を示す可能性が高いが、株価に直接関連するとは言い切れない。逆に、センチメントフレーズは株価に関連するが、人間が解釈するには情報が不十分で理解しにくい。対して因子は金融テキストの説明可能性、人間の可読性、株価の説明可能性という 3 つの側面ではほかのフレーズの抽出より優れている。

しかし、この手法は予測と説明をひとまとめたフレームワークであるため、このフレームワークより精度の優れたモデルに対して説明を提供することが出来ない。本研究では金融タスクの説明を LLM を用いて生成し、金融タスク自体は通常の機械学習モデルによって行う。現在、金融タスクにおいて、LLM を利用しない手法も数多く存在する。そのため、本研究の提案手法は、既存の LLM を利用しない金融予測モデルの説明可能性を向上させることが期待される。

3. 提案手法

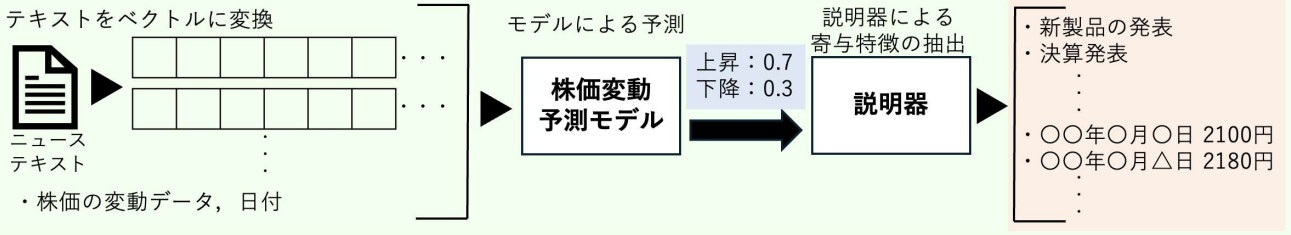
提案手法の概要を図 1 に示す。提案手法は金融予測タスクを行う予測器、予測器が行った予測の根拠となる要因を抽出する説明器、説明器が抽出した要因を人間が理解できる自然なかたちでユーザに提供する LLM によって構成される。

タスクの定義

モデルへの入力：ニュースのテキストデータ，株価変動データ，日付 ▶ モデルの出力：翌日の調整後終値が上がるか下がるか

システムの概要

STEP1：予測に寄与した特徴の抽出



STEP2：LLMによる文章説明の生成

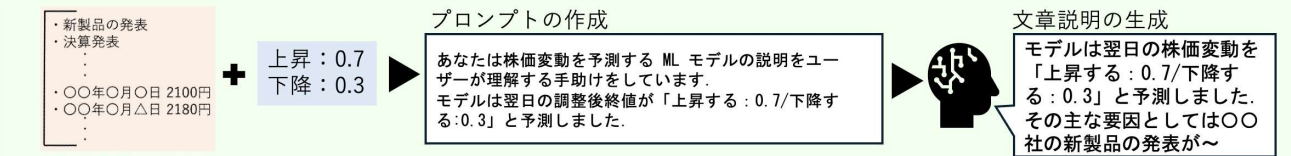


図 1: 提案手法の概要

3.1 タスクの定義

金融タスクには、株価の予測やポートフォリオ最適化など様々なタスクが存在するが、本研究では、株式市場のとある銘柄について、過去 N 日分のニューステキスト、株価変動、日付から翌日上がるか下がるかを予測するタスクを設定する。ニューステキストは、BERT[21]等の事前学習済み言語モデルで埋め込みベクトルに変換したものを利用する。予測器には、LSTM(Long Short-Term Memory)[22]のような、時系列データを扱う金融予測で広く用いられるモデルを想定している。モデルの出力は、翌日の株価が上昇する確率を 0 から 1 の間の値で示すものとし、0.5 以上であれば株価上昇、0.5 未満であれば株価下落と予測する。

3.2 SHAP による寄与度の抽出

予測に寄与した重要な情報を抽出するため、SHAP を用いて予測モデルの予測値をもとに、各インスタンスの寄与度を定量化する。これにより、説明に含めるべき情報と、その情報がどれほど重要かを算出する。予測器に入力したニューステキスト、株価変動、日付情報のうちニューステキストと株価変動データに対して、SHAP の説明を生成する。その中から、各インスタンスにおける寄与度の最も高かった特徴量を、ニューステキストからは 10 件、株価変動データからは 3 件取得し、次のステップの LLM への入力とする。

3.3 LLM による文章説明の生成

SHAP によって得られたそれぞれの特徴と、その寄与度を入力として、LLM によって流暢な文章による説明を生成する。このとき、プロンプトに説明の粒度に関する

要求や、説明を行う対象の情報を追加することで、ユーザに合わせた説明の提供を実現する。具体的には、投資初心者に対して説明を生成する場合、プロンプトに「回答では投資初心者にもわかるような簡単な言葉を利用し、必要に応じて専門用語の説明や言い換えを行ってください」のような文言を追加する。

4. 実験

本節では、金融タスクにおける、LLM によって流暢な文章に変換された説明の有効性を検証するための実験と、その評価方法について検討する。過去の株価と、それに関するテキストデータから、価格変動予測モデルを作成し、そのモデルの予測をもとに表 1 のプロンプトを入力することで文章を生成し、その文章の説明可能性を評価する。プロンプトの作成は、Zytek ら [19] を参考に行った。P3, P4, P5 は説明の詳細さを、P6 はユーザに合わせた説明を提供できるか検証する目的で作成している。

4.1 株価変動予測モデルの設計

4.1.1 データセット

金融予測を行うモデルの訓練に用いるデータセットには、DJAI[23]を利用する。DJIA は、2008 年 6 月 8 日から 2016 年 7 月 1 日までのダウ平均のニュースと、株価の情報を含んでいる。

4.1.2 使用する LLM の検討

本手法は、利用する LLM の性能によって、出力が大きく変わることが予想される。それを検証するために本実験では、説明を生成する LLM として GPT-3.5 と

表 1: LLM への入力プロンプト

コード	繰り返し使う文章
[イントロ]	あなたは株価変動を予測する ML モデルの説明をユーザが理解する手助けをしています。
[予測結果]	モデルは翌日の調整後終値が「上昇する:0.7/下がる:0.3」と予測しました。
[タスク]	SHAP を使って生成した寄与した特徴の説明を、(feature, feature.value, contribution, average.feature.value) 形式でお渡しします。説明を組み合わせる自然な説明に変換してください。
[SHAP 説明]	(feature, feature.value, contribution, average.feature.value) ...
コード	プロンプト
P1	[イントロ][予測結果][タスク][SHAP 説明]
P2	あなたは、ML を扱った経験のないユーザが ML モデルの予測を理解する手助けをしています。 [予測結果][タスク]. あなたの答えは、できるだけ自然で、会話をしているような文章にしてください。 [SHAP 説明].
P3	[イントロ][予測結果][タスク] 特定のインスタンスと平均との比較に関する文脈を含め、全体的な予測に対する各特徴の貢献の重要性を強調した説明を行ってください。 [EXP].
P4	[イントロ][予測結果][タスク] 回答の中で、特徴や寄与度について明確に言及しないこと。 [SHAP 説明].
P5	[イントロ][予測結果][タスク] 回答ではすべての特徴と寄与度について明確に言及するようにしてください [SHAP 説明].
P6	[イントロ][予測結果][タスク] 回答では投資初心者にもわかるような簡単な言葉を利用し、必要に応じて専門用語の説明や言い換えを行ってください [SHAP 説明].

GPT-4 の二つを用意し、生成された説明を比較検討する。GPT-4 は、GPT-3.5 と比較して文章生成の性能が高いことが知られているため、GPT-4 の方が、自然で優れた説明を生成できることが予想される。

4.2 評価指標

4.2.1 株価変動予測モデルの評価

金融予測モデルの予測を精度 (ACC) と、マシューズ相関係数 (MCC) によって評価する。ACC は全ての予測のうち、正しく予測することができたものの割合を表し、式 (2) と表す。ここで TP は上昇を正しく予測できた数、FN は下降を正しく予測できた数、FP は下降を上昇と誤って予測した数、FN は上昇を下降と誤って予測した数を表している。

$$ACC = \frac{TP + TN}{TP + TN + FP + FN} \quad (2)$$

MCC は式 (3) のように表され、予測モデルの質をより総合的に評価する指標である。-1 から 1 の間の値を取り、1 に近いほど予測が優れていることを示す。MCC は不均衡なデータセットでも有効な指標であり、金融市場のような上昇・下降が不均衡な状況でも適切に評価できる。

$$MCC = \frac{TP \times TN - FP \times FN}{\sqrt{(TP + FP)(TP + FN)(TN + FP)(TN + FN)}} \quad (3)$$

4.2.2 LLM による説明可能性の評価

Koa ら [24] をもとに、生成された説明の LLM による客観的な評価手法を検討する。生成された文章を GPT-4 に入力し、次の 10 個の項目について 1 から 7 までの値で評価させることで実現する。

5. 現状の課題

5.1 SHAP による寄与特徴の抽出

提案手法では、ニューステキストをベクトル化するために、BERT を用いることを想定しているが、BERT が生成するベクトルは、入力テキストの文脈情報を高次元空間で表現するため、非常に抽象的である。そのため、SHAP によって特徴量を抽出しても、各ベクトル成分が具体的にどの単語やフレーズに対応しているのかを直接的に理解することが難しい。適切に情報を抽出できなければ LLM による自然な説明も期待できないため、ニューステキストから予測に寄与した出来事を、意味が解釈できるかたちで抽出する手法を検討する必要がある。

ベクトル化されたニューステキストを、より解釈性の高いかたちで利用する方法はいくつか存在する。BERTopic[25] は BERT でベクトル化されたテキストデータから特徴的なトピックをベクトル化し、階層的クラスタリングで分類する手法である。この手法で取得したトピックベクトルを予測の特徴として利用すれば、予測に寄与したトピックベクトルから、株価変動の要因として重要な出来

表 2: 評価項目とその意味

評価項目	評価項目の説明
株価の動きと関連性	株価の動きに直接関連するような要因に重点を置いて説明されているか。
財務指標	業績予測や時価総額など、関連する財務指標を含んで説明がされているか。
世界と業界に関する要因	広範な経済情勢が考慮されているか。
企業の発展	関連する企業の具体的な進展に関する内容が含まれているか。
時間的認識	出来事の順序を考慮した説明になっているか。
プラス要因とマイナス要因のバランス	予測の肯定的な内容と否定的な内容をバランスよく含んでいるか。
文脈理解	自然な文章で行われているか。
明快さと一貫性	明確で理解しやすく、要因を論理的に結びつけた守備一貫した内容になっているか。
情報との整合性	事実と一致しているか。また、情報に不正確さや矛盾はないか。
更新への感受性	銘柄に影響を与えうる将来の可能性に対して触れられているか。

事に関連する単語を抽出することが出来、BERT でベクトル化されたテキストから抽出した特徴量より、解釈性の高い特徴量が得られることが期待される。

5.2 説明可能性の評価

提案手法によって生成された文章の有効性を評価する手法は、再検討の余地がある。本研究では、先行研究 [24] に倣い、GPT を用いた評価を検討している。しかし、生成 AI を利用した評価手法であるため、高い評価結果が得られたとしても、それは提案手法により生成された説明が、金融市場を正確に説明していることを保証するものではない。別の評価手法としてユーザアンケートによる評価があるが、評価が主観的になってしまう問題がある [26]。

6. おわりに

本研究では、金融市場における機械学習モデルの説明の改善を目的として、従来の XAI 手法で生成された説明を LLM を活用して人間が自然に読める文章に変換する手法を提案した。この手法は、SHAP などの XAI 手法を用いて抽出した重要特徴量を LLM に入力し、自然言語による説明文を生成するものである。現状の課題としては、SHAP による寄与特徴量の抽出方法の改善や、生成された説明の評価方法の精緻化が挙げられる。今後は実際に提案手法の実装を行い、金融価格予測とその説明の生成を行うことで、手法の問題点を明らかにするとともに、有効性の検証を行う予定である。

参考文献

- [1] P. C. Tetlock, M. Saar-Tsechansky, and S. Macskassy. More than words: Quantifying language to measure firms' fundamentals. *The Journal of Finance*, Vol. 63, No. 3, pp. 1437–1467, 2008.
- [2] Robert P. Schumaker and Hsinchun Chen. Textual analysis of stock market prediction using breaking financial news: The azfin text system. *ACM Trans. Inf. Syst.*, Vol. 27, No. 2, 2009.
- [3] Saloni Mohan, Sahitya Mullapudi, Sudheer Sammeta, Parag Vijayvergia, and David C. Anastasiu. Stock price prediction using news sentiment analysis. In *2019 IEEE Fifth International Conference on Big Data Computing Service and Applications (BigDataService)*, pp. 205–208, 2019.
- [4] Nohyoong Seong and Kihwan Nam. Predicting stock movements based on financial news with segmentation. *Expert Systems with Applications*, Vol. 164, p. 113988, 2021.
- [5] Matin N. Ashtiani and Bijan Raahemi. News-based intelligent prediction of financial markets using text mining and machine learning: A systematic literature review. *Expert Systems with Applications*, Vol. 217, p. 119509, 2023.
- [6] J. Bollen, H. Mao, and X. Zeng. Twitter mood predicts the stock market. *Journal of Computational Science*, Vol. 2, No. 1, pp. 1–8, 2011.
- [7] B. Li, K. C. C. Chan, C. Ou, and R. Sun. Discovering public sentiment in social media for predicting stock movement of publicly listed companies. *Information Systems*, Vol. 69, pp. 81–92, 2017.
- [8] G. Chen and A. Kurov. Informational role of social media: Evidence from twitter sentiment. *Journal of Banking and Finance*, Vol. 121, , 2020.
- [9] Xuan Ji, Jiachen Wang, and Zhijun Yan. A stock price prediction method based on deep learning technology. *International Journal of Crowd Science*, Vol. 5, No. 1, pp. 55–72, 2021.

- [10] Kentaro Ueda, Hirohiko Suwa, Masaki Yamada, Yuki Ogawa, Eiichi Umehara, Tatsuo Yamashita, Kota Tsubouchi, and Keiichi Yasumoto. Sscdv: Social media document embedding with sentiment and topics for financial market forecasting. *Expert Systems with Applications*, Vol. 245, p. 122988, 2024.
- [11] Latrisha N. Mintarya, Jeta N. M. Halim, Callista Angie, and Said Achmad. Machine learning approaches in stock market prediction: A systematic literature review. In *Procedia Computer Science*, volume 216, pp. 96–102, 2023.
- [12] Christoph Molnar. *Interpretable Machine Learning*. Lulu.com, 2020.
- [13] Umang Bhatt, Alice Xiang, Shubham Sharma, Adrian Weller, Ankur Taly, Yunhan Jia, Joydeep Ghosh, Ruchir Puri, José M. F. Moura, and Peter Eckersley. Explainable machine learning in deployment. Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency (FAT* ’20), pp. 648–657, 2020.
- [14] Marco Tulio Ribeiro, Sameer Singh, and Carlos Guestrin. ”why should i trust you?”: Explaining the predictions of any classifier. In *KDD ’16: Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pp. 1135–1144, 2016.
- [15] Scott M. Lundberg and Su-In Lee. A unified approach to interpreting model predictions. In *NIPS’17: Proceedings of the 31st International Conference on Neural Information Processing Systems*, pp. 4768–4777, 2017.
- [16] Waddah Saeed and Christian Omlin. Explainable ai (xai): A systematic meta-survey of current challenges and future opportunities. In *Knowledge-Based Systems*, volume 263, 2023.
- [17] Yupeng Chang, Xu Wang, Jindong Wang, Yunan Wu, Linyi Yang, Kaijie Zhu, Hao Chen, Xiaoyuan Yi, Cunxiang Wang, Yidong Wang, Wei Ye, Yue Zhang, Yi Chang, Philip S Yu, Qiang Yang, and Xing Xie. A survey on evaluation of large language models. *ACM Transactions on Intelligent Systems and Technology*, Vol. 15, No. 39, pp. 1–45, 2024.
- [18] David Martens, James Hinns, Camille Dams, Mark Vergouwen, and Theodoros Evgeniou. Tell me a story! narrative-driven xai with large language models. May 2024.
- [19] Alexandra Zyteck, Sara Pido, and Kalyan Veeramachaneni. Llms for xai: Future directions for explaining explanations. In *ACM CHI 2024, HCXAI. Workshop on Human-Centered Explainable AI*, 2024.
- [20] Meiyun Wang, Kiyoshi Izumi, and Hiroki Sakaji. Llmfactor: Extracting profitable factors through prompts for explainable stock movement prediction. June 2024.
- [21] Jacob Devlin, Ming-Wei Chang, Kenton Lee, and Kristina Toutanova. Bert: Pre-training of deep bidirectional transformers for language understanding. In *Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long and Short Papers)*, pp. 4171–4186, 2019.
- [22] Sepp Hochreiter and Jürgen Schmidhuber. Long short-term memory. *Neural computation*, Vol. 9, No. 8, pp. 1735–1780, 1997.
- [23] aaron7sun. Daily news for stock market prediction. <https://www.kaggle.com/datasets/aaron7sun/stocknews>. アクセス日: 2024-07-21.
- [24] Kelvin J. L. Koa, Yunshan Ma, Ritchine Ng, and Tat-Seng Chua. Learning to generate explainable stock predictions using self-reflective large language models. In *WWW ’24: Proceedings of the ACM on Web Conference 2024*, pp. 4304–4315, May 2024.
- [25] M.Grootendorst. Bertopic: Neural topic modeling with a class-based tf-idf procedure. March 2022.
- [26] Meike Nauta, Jan Trienes, Shreyasi Pathak, Elisa Nguyen, Michelle Peters, Yasmin Schmitt, Jörg Schlötter, Maurice van Keulen, and Christin Seifert. From anecdotal evidence to quantitative evaluation methods: A systematic review on evaluating explainable ai. In *ACM Computing Surveys*, Volume 55, Issue 13s, pp. 1–42, 2023.