The Journal of Finance and Data Science 10 (2024) 100137

Contents lists available at ScienceDirect

The Journal of Finance and Data Science

journal homepage: www.keaipublishing.com/en/journals/jfds

Research Article

Reinforcement prompting for financial synthetic data generation

Xiangwu Zuo[a,*], Anxiao (Andrew) Jiang[a], Kaixiong Zhou[b,1]

[a] Computer Science and Engineering Department, Texas A&M University, College Station, TX, USA
[b] Department of Electrical and Computer Engineering, North Carolina State University, Raleigh, NC, USA

ARTICLE INFO

Keywords:
Reinforcement prompting
Reinforcement learning
Synthetic data
Large language model
Financial sentiment analysis
Machine learning

ABSTRACT

The emergence of Large Language Models (LLMs) has unlocked unprecedented potential for comprehending and generating human-like text, fueling advances in the finance domain – a tool that can shape investment strategies and market predictions. Nevertheless, challenges stemming from the necessity for extensive labeled data and the imperative for data privacy remain. The generation of high-quality synthetic data emerges as a promising avenue to circumvent these issues. In this paper, we introduce a novel methodology, named "Reinforcement Prompting", to address these challenges. Our strategy employs a policy network as a Selector to generate prompts, and an LLM as an Executor to produce financial synthetic data. This synthetic data generation process preserves data privacy and mitigates the dependency on real-world labeled datasets. We validate the effectiveness of our approach through experimental evaluations. Our results indicate that models trained on synthetic data generated via our approach exhibit competitive performance when compared to those trained on actual financial data, thereby bridging the performance gap. This research provides a novel solution to the challenges of data privacy and labeled data scarcity in financial sentiment analysis, offering considerable advancement in the field of financial machine learning.

1. Introduction

In recent years, Large Language Models (LLMs) become extremely hot with the advancements in machine learning algorithms and increasing computational capabilities. These models affect a large number of domains, extending beyond the confines of computational linguistics (Hoffmann et al. (2022); Wei et al. (2022)). Trained on a massive corpus of language text, LLMs have exhibited substantially proficiency in generating text that is both contextually appropriate and intricately detailed. For instance, models such as GPT-3 have been shown to generate high-quality text across a variety of contexts and tasks (Brown et al. (2020)). The generated text mirrors the complexity and subtlety of human language, making it virtually indistinguishable from human-written text. This capability allows LLMs to be applied across diverse sectors. In customer service, they power sophisticated chatbots and virtual assistants that handle inquiries with human-like responsiveness. In media and journalism, they assist in generating content that ranges from writing articles to crafting personalized reports. In the field of education, LLMs facilitate the creation of interactive learning tools and personalized learning experiences. Additionally, they play a pivotal role in translating texts, enabling communication across different languages in real-time.

* Corresponding author.
E-mail addresses: dkflame@tamu.edu (X. Zuo), ajiang@cse.tamu.edu (A. Jiang), kzhou22@ncsu.edu (K. Zhou).
Peer review under responsibility of KeAi Communications Co., Ltd.
[1] Present Address: MIT Institute for Medical Engineering & Science, Cambridge, MA, USA.

https://doi.org/10.1016/j.jfds.2024.100137
Received 21 September 2023; Received in revised form 22 June 2024; Accepted 31 July 2024
Available online 3 August 2024

LLMs constitute a specialized subclass within machine learning paradigms, distinguished by their capacity to assimilate and generate text based on vast, context-rich data sets. These models capitalize on architectural innovations like transformers (Vaswani et al. (2017)), and are often instantiated using state-of-the-art frameworks such as GPT-3 (Brown et al. (2020)) and BERT (Devlin et al. (2019)). The essence of LLMs lies in their substantially improved proficiency in capturing and replicating the semantic, syntactic, and pragmatic elements of language. This enables them not only to comprehend contextual subtleties but also to engage effectively with specialized terminologies across various domains. The versatile and robust nature of LLMs makes them indispensable tools for a wide array of natural language processing tasks, and suggests their potential to catalyze further innovation in the realm of linguistic comprehension and generation.

Sentiment analysis, commonly denoted as opinion mining, constitutes a distinct sub-discipline within the broader field of natural language processing. It revolves around the task of deciphering the sentiment or emotional undertones embedded within a piece of text (Awatramani et al. (2021); Mishra et al. (2022)). The financial sector, in particular, employs sentiment analysis as a strategic tool to gauge market sentiment. This tool has the potential to bring significant influence to financial markets. For instance, the dissemination of positive news about a company can trigger a surge in its stock price, while negative news can instigate a downward spiral.

Within the finance domain, Large Language Models (LLMs) hold considerable promise. They have the potential to unlock valuable insights by analyzing extensive quantities of financial data, including but not limited to news articles, financial reports, and social media posts. Despite the potential advantages, the application of LLMs, especially in the realm of financial sentiment analysis, is fraught with challenges. However, a key issue is that these models are typically trained on a broad spectrum of internet text and may not perform optimally on domain-specific text, such as financial news or reports, without further adjustments (Gururangan et al. (2020)).

Moreover, the datasets used for training and fine-tuning for financial tasks may contain various types of sensitive information, such as personal financial details, account numbers, transaction histories, and credit scores (Bryant et al. (2019); Truong et al. (2020)). This introduces a significant element of risk and raises serious concerns about data privacy. The potential for such sensitive information, inadvertently included in the training data, to be exposed or misused represents a critical challenge that must be addressed when applying LLMs in financial sentiment analysis.

Beyond the realm of data privacy, financial sentiment analysis presents a unique set of challenges (Alissa and Alzoubi (2022)). A primary challenge lies in the necessity for large labeled datasets for training sentiment analysis models. The task of labeling data for sentiment analysis is labor-intensive and time-consuming, requiring not only a deep understanding of language but also a comprehensive knowledge of the financial domain. Financial text often includes domain-specific jargon, further complicating the task. Additionally, the subtleties and complexities inherent in financial sentiment make this a particularly challenging task. The scarcity of such labeled data, coupled with the complexities of the task, present significant obstacles in the deployment of LLMs for financial sentiment analysis.

To address the challenges of data privacy and the scarcity of labeled data, this research introduces an innovative methodology that exploits the capabilities of Large Language Models (LLMs) to generate high-quality synthetic data (Efimova et al. (2020); Wu et al. (2022)). Synthetic data refers to artificially produced data that replicates the characteristics of real-world data without containing any confidential or sensitive information. Synthetic data holds the potential to address the issues of data privacy, as it can serve as a substitute for real-world labeled data, which often contain sensitive information. Furthermore, synthetic data can effectively address the challenge of the scarcity of labeled data, as it can be produced in large quantities without the need for manual labeling.

Our proposed methodology is centered on a unique application of Reinforcement Learning (RL) principles within a process we term as *Reinforcement Prompting*. Within this framework, a policy network and a Large Language Model (LLM) are each delegated specific functions. The policy network, termed the “Selector” and mathematically denoted as *Agent*, is entrusted with the selection of keywords that form the basis for prompts generation. On the other side, the LLM, herein referred to as the “Executor” or symbolically represented as *LLM*, undertakes the synthesis of data predicated on the prompts generated by the *Agent*. This entire operation is guided by RL principles, ensuring a robust and adaptive system for synthetic data generation.

Within the *Reinforcement Prompting* framework, the Selector *Agent* curates a selection of keywords from a pre-defined Keywords Vocabulary. These selected keywords are combined with a Prompt Template to generate a context-specific prompt at each iterative step. This formulated prompt guides the Executor *LLM* in the generation of synthetic data. A reward function evaluates the quality of this synthetic data, taking into consideration multiple aspects. This function incorporates several criteria: the performance of a localized model trained on the synthetic data on a held-out validation dataset, the similarity between the synthetic data and real-world financial texts, and the ratio of the synthetic data that contains terms from a Financial Terminology Glossary.

The primary objective of Reinforcement Prompting is to identify the optimal prompt that leads to the generation of high-quality synthetic data. By “high quality”, we refer to synthetic data that excels in two key aspects: 1) accuracy, 2) sentence structure complexity and diversity. Accuracy is measured by the performance of the model on downstream tasks, ensuring that the generated data enables the model to achieve results comparable to or better than those obtained using real data. Sentence structure complexity and diversity are assessed by how closely the generated sentences match the complexity and variety of structures found in true data, avoiding the overly simplistic and uniform structures typically produced by direct LLM generation without our method. Examples of direct LLM generated data and our Reinforcement Prompting generated data are provided in Tables 1 and 7, respectively. This process is iterative, with the Selector *Agent* progressively improving the prompts it generates based on maximizing the total rewards in each training episode. This loop allows the system to learn and adapt, optimizing the quality of the generated synthetic data over time.

Table 1
Samples of direct LLM generated sentences via GPT-3.5-turbo.

Sentence	Label
Company XYZ reported record-breaking profits for the quarter	positive
The CEO's resignation sparked concerns among investors	negative
The market reacted positively to the company's new product launch	positive
Investors remained cautious amid economic uncertainty	neutral
Company GHI faced backlash following a data breach	negative
The company's expansion into new markets was met with skepticism	neutral
The company's cost-cutting measures were well-received by investors	positive
Analysts remained neutral on the company's future prospects	neutral
The company's stock price remained stable despite market volatility	neutral

Table 2
Sampled keyword sets and prompt template.

Sampled Keyword Sets	Prompt Template
Foreign Exchange, Forex, Currency Pair, Exchange Rate, Spot Market, Futures Contract, Option Contract Financial Analysis, Ratio Analysis, Liquidity Ratios, Profitability Ratios, Solvency Ratios, Valuation Investment Strategy, Value Investing, Growth Investing, Passive Investing, Active Investing, Contrarian Investing Options, Futures, Call Option, Put Option, Strike Price, Expiration Date, Option Premium Personal Finance, Budget, Savings, Expenses, Income, Emergency Fund, Credit Card Asset Class, Equity, Fixed Income, Commodities, Real Estate, Cash and Cash Equivalents	Generate 100 financial sentences related to companies or persons that fulfill the following strict criteria: 1) Label each sentence as 'positive', 'negative', or 'neutral'. 2) Format the output as '[sentence] | [label]' and exclude any additional text. 3) The sentences should contain topics related to any of the following keyword sets: "@Selected_KeywordSets"

Table 3
Sample sentences and their respective labels from the Financial PhraseBank dataset.

Sentence	Label
Circulation revenue has increased by 5% in Finland and 4% in Sweden in 2008.	positive
Technopolis plans to develop in stages an area of no less than 100,000 square meters in order to host companies working in computer technologies and telecommunications, the statement said.	neutral
The international electronic industry company Elcoteq has laid off tens of employees from its Tallinn facility; contrary to earlier layoffs the company contracted the ranks of its office workers, the daily Postimees reported.	negative
Both operating profit and turnover for the three-month period increased, respectively from EUR0.9 m and EUR8.3 m, as compared to the corresponding period in 2005.	positive
According to Gran, the company has no plans to move all production to Russia, although that is where the company is growing.	neutral
Operating profit was EUR 11.07 mn, up from EUR 8.65 mn.	positive
Kalmar Espana generated net sales of some 11.3 mln euro $ 14.8 mln in 2005.	neutral
Jan. 6 – Ford is struggling in the face of slowing truck and SUV sales and a surfeit of up-to-date, gotta-have cars.	negative
Rautakesko 's business operations in Norway and Russia, acquired in July 2005, are included in the figures of the comparable period, impacting sales growth starting from August.	neutral

Table 4
Sample terms in the financial terminology glossary.

Investopedia Dictionary	AMEX Exchange	NASDAQ Exchange	NYSE Exchange
SEC Release IA-1092	ASM	CLAY	BA
Accretion	BHB	DYNT	DSU
BAT Stocks	COE	GENQW	FSM
Cash Flow	DPSI	INMB	HAE
Commission	EQX	LIFWZ	JLL
Debt Ratio	FRD	MOBQ	MQY
Exchange Rate	GLU	PAA	NPCT
Forecasting	HNW	SFR	TUYA
Gross Income	INLX	TCBC	YELP

Table 5
Selected keyword sets.

Keywords
Business Valuation, Book Value, Market Value, Liquidation Value, Replacement Value
Market Research, Qualitative Research, Quantitative Research, Market Segmentation, Target Market
Financial Derivatives, Options, Futures, Swaps, Forward Contracts, Counterparty Risk
Annuities, Ordinary Annuity, Annuity Due, Perpetuity, Present Value of Annuity
Personal Finance, Budget, Savings, Expenses, Income, Emergency Fund, Credit Card

Table 6
Generated prompt for executor.

Optimal Prompt
Generate 100 financial sentences related to companies or persons that fulfill the following strict criteria: 1) Label each sentence as 'positive', 'negative', or 'neutral'. 2) Format the output as '[sentence] | [label]' and exclude any additional text. 3) The sentences should contain topics related to any of the following keyword sets:
- Business Valuation, Book Value, Market Value, Liquidation Value, Replacement Value
- Market Research, Qualitative Research, Quantitative Research, Market Segmentation, Target Market
- Financial Derivatives, Options, Futures, Swaps, Forward Contracts, Counterparty Risk
- Annuities, Ordinary Annuity, Annuity Due, Perpetuity, Present Value of Annuity
- Personal Finance, Budget, Savings, Expenses, Income, Emergency Fund, Credit Card

Table 7
Samples of the generated financial synthetic dataset via reinforcement prompting.

Sentence	Label
The Walt Disney Company's acquisition of 21st Century Fox's film and television assets for $71.3 billion was aimed at bolstering its content offerings and competing with streaming services like Netflix.	positive.
Goldman Sachs has been criticized for its involvement in the 1MDB scandal, which involved billions of dollars of fraud and corruption, but the company continues to maintain its reputation as a top investment bank.	negative
The high interest rates associated with credit cards can be a significant financial burden, making it difficult to pay off balances and save money.	negative
The market value of a company's stock reflects the perceived worth of the company by investors and analysts, based on factors such as financial performance and growth potential.	neutral
A recent analysis of the company's product portfolio showed that many were underperforming, indicating a need for greater investment in product development and innovation.	neutral
The CEO announced plans to partner with financial institutions to offer discounted loans and credit to employees who demonstrate strong personal finance habits.	positive
The company's replacement value would be significantly lower than its market value due to substantial investments in research and development that have yet to be reflected in its financial statements.	positive
Morgan Stanley has been accused of engaging in risky trading practices and contributing to the 2008 financial crisis, but the company has also been recognized for its commitment to sustainability and diversity.	neutral
Coca-Cola's introduction of its new flavor Coke Zero Sugar has been successful in attracting younger consumers and bolstering the company's sales.	positive
Intel has faced challenges in diversifying its product offerings and moving beyond its traditional business of manufacturing computer chips, but the company continues to invest in research and development to stay ahead of the competition.	neutral

Reinforcement Prompting presents several distinct advantages over traditional methods of data generation and sentiment analysis.

- Firstly, it addresses the critical issue of data privacy. The generation of synthetic data ensures that sensitive information is neither utilized nor disclosed in the process, offering a particularly crucial benefit in the financial domain, where data often contains sensitive or confidential information.
- Secondly, Reinforcement Prompting mitigates the need for extensive labeled datasets for training sentiment analysis models. Generating high-quality synthetic data offers a practical alternative to real-world labeled data, which can be challenging to obtain and requires substantial manual labeling efforts.
- Thirdly, by capitalizing on the capabilities of LLMs, Reinforcement Prompting can generate synthetic data that is contextually appropriate and closely resembles real-world financial texts. This makes it particularly suitable for training models for financial sentiment analysis, where a deep understanding of domain-specific terms and the ability to capture the sentiment is critical.
- Lastly, Reinforcement Prompting is a flexible and scalable methodology. It can be used with any LLM and can be easily scaled to generate larger volumes of synthetic data as required. This flexibility and scalability make Reinforcement Prompting a versatile tool for financial sentiment analysis, capable of catering to diverse needs and requirements.

The implications of this research are substantial, particularly in the context of financial sentiment analysis. By offering a practical solution to the challenges of data privacy and the scarcity of labeled data, this study contributes to the further development and refinement of sentiment analysis techniques in the financial domain. Moreover, by introducing a novel application of RL principles in the context of LLMs, this research opens up new avenues for exploration in the fields of machine learning and natural language processing. By bridging the gap between LLMs and RL, this study paves the way for a new generation of machine learning methodologies that are both effective and privacy-preserving.

2. Related work

In recent years, the application of machine learning, particularly natural language processing, in finance has been a subject of extensive research. This section provides an overview of the existing literature of Large Language Models (LLMs) and financial data, the challenges associated with data privacy and labeled data, and the application of Reinforcement Learning (RL) in text generation.

The field of research concerning the application of LLMs in finance is rapidly growing (Lu et al. (2023)). LLMs have been used in various financial applications, including market prediction (Wu et al. (2023)), risk assessment (Yang et al. (2023)), and sentiment analysis (Deng et al. (2023)). For instance, Wu et al. (2023) presented BloombergGPT, a large language model specifically designed for financial tasks, demonstrating its effectiveness in market prediction and sentiment analysis. Similarly, Lu et al. (2023) introduced BBT-Fin, a comprehensive pre-trained language model for the Chinese financial domain, aimed at improving performance in financial NLP tasks.

Despite their potential, LLMs face significant challenges when applied in finance, particularly regarding data privacy and the scarcity of labeled data. Real-world financial data often contains sensitive information, such as personal financial details, transaction histories, and confidential corporate information. The use of such data for training models raises serious privacy concerns (Lu et al. (2020)). For example, the leakage of transaction histories or personal financial details can lead to identity theft, financial fraud, and significant financial loss. Additionally, corporate information leaks can have severe consequences for businesses, including loss of competitive advantage and legal ramifications.

Anonymization and Differential Privacy are common approaches to address these issues. Anonymization involves removing sensitive information to prevent the identification of individuals from the data. However, anonymized data can sometimes be re-identified when combined with other data sources, making it an inadequate solution for ensuring complete privacy (Narayanan and Shmatikov (2008)). Differential Privacy, on the other hand, provides a mathematical guarantee of privacy by adding noise to the data or the model's outputs to protect individual data points. While this method enhances privacy protection, it often comes at the cost of reduced model accuracy and utility (Dwork (2006)). In the financial domain, achieving a balance between privacy and data utility is particularly challenging due to the high stakes involved. Solutions like blockchain technology have also been proposed to enhance data privacy in financial applications (Lu et al. (2020)), but these approaches are still evolving and have their limitations.

The scarcity of labeled data poses another significant challenge. Large volumes of labeled data are needed for training LLMs, but obtaining such data is difficult due to the extensive manual labeling required. Existing solutions like semi-supervised learning and active learning help, but they still rely on some labeled data and can be complex to implement (Ji et al. (2022); Weng, Pratama, Za'in, de Carvalho, Appan, Ashfahani and Yee (2022); Presotto et al. (2023)).

To address these challenges, synthetic data generation has emerged as a promising solution. Synthetic data mimics real-world data characteristics without containing sensitive information, helping to maintain data privacy while providing sufficient training data (Sixt et al. (2016); Antoniou et al. (2017); Xie et al. (2018)).

Reinforcement Learning (RL) has been effective in text generation tasks across various domains (Donahue and Rumshisky (2018); Dognin et al. (2021); Upadhyay et al. (2022); Zhang et al. (2022)). However, its potential in financial sentiment analysis using LLMs remains underexplored. Defining suitable reward functions for RL in text generation is challenging due to the subjective nature of text quality.

In summary, the limitations and challenges inherent in the existing literature underscore the pressing need for innovative methodologies. Our study aims to fill this gap through a novel methodology termed *Reinforcement Prompting*, integrating RL principles and LLMs to generate high-quality synthetic financial data.

3. Reinforcement Prompting framework

This section delineates the architecture and functionality of the Reinforcement Prompting Framework, a novel methodology explicitly tailored to optimize the generation of financial synthetic data. Although the framework fundamentally aligns with traditional reinforcement learning paradigms, it incorporates unique elements in its structural design and operational implementation. These particular features inspire its naming – Reinforcement Prompting.

3.1. Framework inputs

The framework requires the following inputs:

- A specialized Keyword Vocabulary, denoted by *KV*, and an associated Prompt Template, denoted by *TPL*.
- A Reinforcement Policy Network, designated as the Selector *Agent*, is responsible for the selection of keywords for prompt generation.
- A Large Language Model (LLM), referred to as Executor *LLM*, is assigned the task of generating financial synthetic data based on the prompts formulated by *TPL* and *Agent*.

3.2. Framework outputs

The framework yields the following outputs:

- The optimal policy π^* for the *Agent* in the context of keyword selection.
- The most effective prompt pmt^* synthesized through the combination of *TPL* and *Agent*.
- A dataset $D = \{d_1, d_2, \ldots, d_m\}$ of high-quality synthetic financial data, generated by *LLM* utilizing the prompt pmt^*.

3.3. Optimization objective

The primary objective of our optimization process is to identify an optimal prompt pmt^* synthesized by both *TPL* and *Agent*. Given that *TPL* serves as a static template, the resultant prompt *pmt* is contingent upon the keywords selected by *Agent*. Therefore, the optimization objective is identical to maximizing the cumulative reward accrued by *Agent* in choosing keywords from *KV* that compose an optimal prompt *pmt** for guiding *LLM* to generate high-quality synthetic data. The trained *Agent* can be subsequently utilized to select keywords for generating the most effective prompt *pmt** for *LLM*. The synthetic dataset generated by *LLM* using this optimal prompt can facilitate few-shot training of localized language models while preserving data privacy. Formally, the objective is to identify the optimal policy π^* that maximizes the expected return $J(\theta)$:

$$J(\theta) = \mathbb{E}_{\tau\sim\pi_\theta}[R(\tau)], \tag{1}$$

where:

- θ represents the parameters governing the policy network.
- π_θ is the policy delineated by the parameters θ.
- $\tau = (s_0, a_0, s_1, a_1, \ldots, s_T, a_T)$ denotes a trajectory, representing a series of states and actions from the initial to the terminal state.
- $R(\tau)$ is the cumulative reward corresponding to the entire trajectory.

To update the parameter θ, the gradient of the expected return is employed:

$$\nabla J(\theta) = \mathbb{E}_{\tau\sim\pi_\theta}\left[\sum_{t=0}^{T} \nabla_\theta \log \pi(a_t|s_t;\theta) r_t\right], \tag{2}$$

where:

- $\pi(a|s;\theta)$ denotes the policy function that maps states s to a probability distribution over actions a.
- r_t is the immediate reward at time t, often termed as the return.

The parameter θ is updated through the following gradient ascent, where α signifies the learning rate.

$$\theta \leftarrow \theta + \alpha \sum_{t=0}^{T} \nabla_\theta \log \pi(a_t|s_t;\theta) r_t, \tag{3}$$

The immediate reward function r_t is constructed based on several factors: 1) The agent accrues a reward for taking an action a. 2) A penalty is imposed if action a duplicates any prior actions. 3) At the final step of each episode, the agent receives a reward based on the evaluation of the generated synthetic dataset D_S. This reward includes the performance of a localized BERT model (trained using real dataset D_R) on the generated synthetic dataset set D_S, the level of similarity between the synthetic data D_S and real-world financial texts D_R, and the proportion of the synthetic data D_S that includes terms from a Financial Terminology Glossary *FTG*.

3.4. Environment and reward function

The environment within which the Reinforcement Prompting Framework operates comprises two core components: a policy network agent termed the Selector (*Agent*) and a Large Language Model (LLM) termed the Executor (*LLM*). The Selector is responsible for generating prompts, while the Executor produces synthetic data in response to these prompts.

For the Selector *Agent*, each time step t involves processing the current state s_t (which represents the sets of keywords that have been selected by *Agent*) and an action a_t (the set of keywords to be selected by *Agent*). The system evaluates these inputs to yield a reward r_t and to transition the state to s_{t+1}. For each state except the final state of an episode, this reward r_t is computed as the gain of adopting a new action r_t(action) minus the penalty incurred for duplicating actions r_t(penalty), mathematically expressed as $r_t = r_t(\text{action}) - r_t(\text{penalty})$.

Upon the termination of each training episode, the Executor *LLM* employs the prompt *pmt*, constructed from the Prompt Template *TPL* and the keywords *KV* chosen by *Agent*, to generate a synthetic dataset D_S. The quality of this dataset informs the synthetic data reward r_t(data) for the final time step. Therefore, the reward for the last time step is $r_t = r_t(\text{action}) - r_t(\text{penalty}) + r_t(\text{data})$.

The reward r_t(data) is a composite function of three distinct metrics, formulated as:

$$r_t(\text{data}) = v_1 \times STY(D_S, D_R) + v_2 \times ACR(D_S, \text{BERT}(D_R)) + v_3 \times LXC(D_S, FTG), \tag{4}$$

where $STY(D_S, D_R)$ measures the stylistic similarity between the synthetic data D_S and the real dataset D_R; $ACR(D_S, \text{BERT}(D_R))$ assesses the accuracy of sentiment prediction on D_S using a BERT model trained on D_R; and $LXC(D_S, FTG)$ evaluates the lexical diversity of D_S in conjunction with its resemblance to a Financial Terminology Glossary *FTG*. The weights v_1, v_2, and v_3 are assigned to each metric, subject to the constraint $v_1 + v_2 + v_3 = 1$.

Thus, the cumulative reward function $R(\tau)$ is defined as the sum of discounted future rewards, written as:

$$R(\tau) = r_1 + \gamma \times r_2 + \cdots + \gamma^{T-1} \times r_T, \tag{5}$$

where T denotes the total number of time steps in the episode, and γ is the discount factor between 0 and 1.

4. Methodology

This section provides an in-depth description of the Reinforcement Prompting methodology, with a particular emphasis on synthetic data generation and the evaluation metrics employed to assess data quality. A schematic overview of the complete framework is depicted in Fig. 1.

4.1. Keywords Vocabulary and Prompt Template

The Reinforcement Prompting method integrates a set of curated financial keywords with a structured Prompt Template, both of which are pivotal in the generation of high-quality financial data. Table 2 presents exemplary instances of Keyword Sets and the associated Prompt Template.

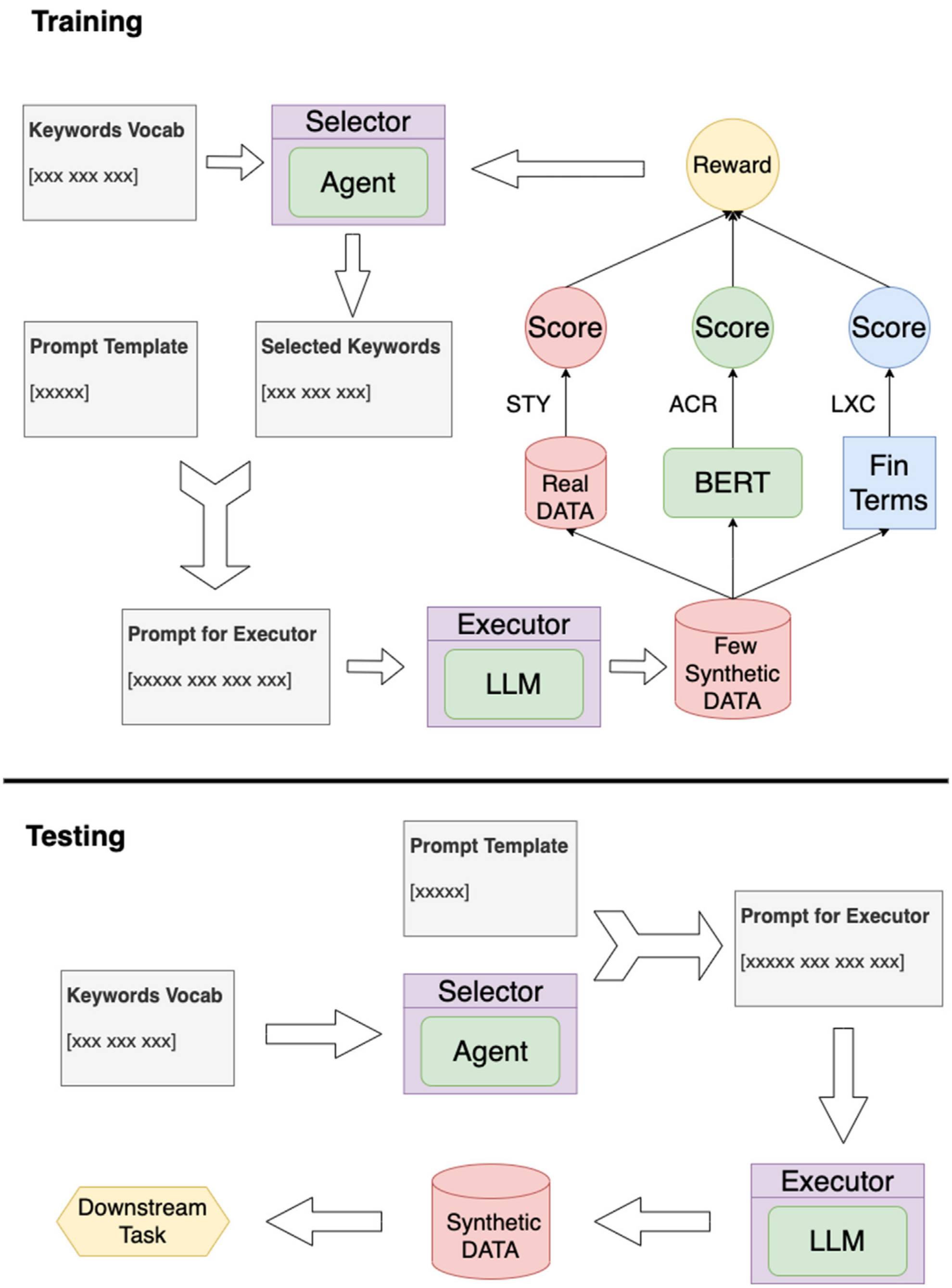

Fig. 1. Framework of financial synthetic data generation via reinforcement prompting.

As illustrated in the gray box labeled “Keywords Vocab” in Fig. 1, the Keyword Sets act as the action space for the Reinforcement Prompting process. They provide the Selector *Agent* with a pool of essential keywords for selection. These keywords are primarily sourced from the Financial Terminology Glossary available at https://www.investopedia.com/. To mitigate the dimensionality of the action space, we implement GPT-3.5-turbo to categorize the keywords into various groups based on their co-occurrence frequencies in the articles of GPT-3.5-turbo's training corpus. The top 100 keyword groups, ranked by appearance frequency, are picked and built into the Keywords Vocabulary.

The gray box labeled “Prompt Template” in Fig. 1 is the template sentence structure for generated prompts. This structure is elaborated upon in Table 2. Our empirical studies reveal that the maximum number of sample sentences GPT-3.5-turbo can produce in a single iteration is 100; therefore, our Prompt Template is designed to solicit the generation of 100 financial sentences at each episode.

Combining the selected keywords from each training episode with the Prompt Template, the Reinforcement Prompting methodology is able to produce prompts. These prompts subsequently guide the Executor *LLM* in the generation of financial synthetic data.

4.2. Selector-Executor Network

Our Reinforcement Prompting procedure is performed via a Selector-Executor Network. During each prompting step, the Selector, denoted as *Agent*, chooses a set of keywords (represented as action a) based on the currently selected keyword sets (represented as state s). If the chosen set of keywords has not been previously selected, a minor reward is granted; otherwise, a penalty is incurred for action duplication. The episode terminates either when the *Agent* selects the “<end>” token as its action or when the maximum number of allowable steps is reached. The cumulatively selected keyword sets are then combined with the Prompt Template to compose the “Prompt for Executor”.

Upon receiving this composite prompt, the Executor *LLM* generates a collection of data points, termed as the “Synthetic DATA”. The quality of this synthetic dataset is subsequently evaluated in comparison to a real financial dataset (“Real DATA”), a specialized BERT model fine-tuned for financial contexts, and a comprehensive Financial Terminology Glossary (“Fin Terms”). The evaluation metrics are aggregated to form a composite reward, which is then passed to the Selecter “Agent”. This accumulated reward serves as the final step's reward for each episode.

Utilizing the Reinforce Algorithm, we employ gradient ascent to optimize the policy of the network. Consequently, the *Agent* becomes proficient at selecting financial keywords to construct effective prompts.

4.3. Evaluation metrics

We adopt three distinctive metrics to evaluate the quality of synthetic data yielded by *LLM* at each prompting step. These metrics, denoted as *STY*, *ACR*, and *LXC*, are consolidated via Eq. (4) to compute the reward for our Reinforcement Prompting.

4.3.1. STY metric

The *STY* metric gauges the stylistic resemblance between the synthetically generated data D_S and the real financial data D_R. It comprises two elements.

1) **STY_1: Sentiment Probability Distribution Similarity Between Real Data and Synthetic Data.**

The objective is to generate a dataset whose sentiment labels, namely “positive”, “negative” and “neutral”, mirror the distribution observed in the real financial dataset. To quantify this congruity, we compute the Jensen-Shannon Distance (JSD) between the sentiment distributions in D_S and D_R.

JSD is a symmetric measure that quantifies the similarity between two probability distributions. It is the mean of two Kullback-Leibler (KL) divergences, where KL divergence measures the information loss when one distribution is utilized to approximate another. The mathematical formulation of JSD is as follows:

$$JSD(P,Q) = \frac{1}{2}D(P\|M) + \frac{1}{2}D(Q\|M) \tag{6}$$

where $D(P\|M)$ and $D(Q\|M)$ represent the KL divergences of P and Q from M, respectively, and M is the average of P and Q:

$$M = \frac{1}{2}(P+Q) \tag{7}$$

The KL divergence, or relative entropy, between two probability distributions quantifies the information loss when one distribution is used to approximate another. For discrete distributions, the KL divergence of P from Q is defined as:

$$D(P\|Q) = \sum P(i)\log\left(\frac{P(i)}{Q(i)}\right) \tag{8}$$

where the summation is over all possible events i, $P(i)$ is the probability of event i according to distribution P, and $Q(i)$ is the probability of event i according to distribution Q. For continuous distributions, the summation in the above equation is replaced by an integral.

While KL divergence is not symmetric, i.e., $D(P\|Q) \neq D(Q\|P)$, the JSD measure, by averaging the two KL divergences, is symmetric. Furthermore, it is bounded between 0 and 1, where 0 indicates identical distributions P and Q, and 1 signifies complete dissimilarity.

Therefore, representing the distribution of “positive”, “negative”, and “neutral” sentiment sentences in D_S and D_R as P and Q respectively, a JSD value is calculated and confined between 0 and 1. Given that STY_1 contributes to the calculation of the STY metric and subsequently the reward r, larger STY_1 values are preferred. Hence, we define STY_1 as:

$$STY_1 = 1 - JSD \tag{9}$$

2) **STY_2: Unique Word Appearance Ratio of Synthetic Data in Real Data.**

The STY_2 sub-metric operates under the assumption that a high incidence of terms from D_S appearing in the real financial dataset D_R is indicative of a stylistic alignment between D_S and D_R. This reasoning forms the basis for the inclusion of STY_2 in our comprehensive STY metric.

During each prompting step, the synthetic dataset D_S is processed to ensure uniformity by converting all words to lowercase. Subsequently, all unique terms are extracted to constitute a list denoted as $Term_S$. The total count of words in $Term_S$ that also appear in D_R is computed and represented as $N(SnR)$. Concurrently, the total word count of D_R is ascertained, represented as $N(D_R)$. These quantities facilitate the computation of the STY_2 value as:

$$STY_2 = \frac{N(SnR)}{N(D_R)} \tag{10}$$

The derived STY_2 is a bounded value within the range of 0 to 1. A larger STY_2 value implies a higher proportion of words in D_S that also appear in D_R, thereby indicating a greater stylistic similarity.

The final STY metric is then computed as a weighted linear combination of STY_1 and STY_2, where w_1 and w_2 are the respective weights and obey the constraint $w_1 + w_2 = 1$. This facilitates the balancing of the importance of sentiment distribution similarity and unique word appearance ratio in the synthetic data:

$$STY = w_1 \times STY_1 + w_2 \times STY_2 \tag{11}$$

4.3.2. ACR metric

The ACR metric serves to evaluate the quality of the synthetic data D_S through the lens of its utility in a downstream application. In our context, the downstream application involves sentiment classification of sentences within D_S, which is performed using a BERT model that has been fine-tuned on the real financial data D_R.

More specifically, the BERT model undergoes a fine-tuning process on D_R, learning to classify sentiments into categories such as positive”, negative”, and “neutral”. Once fine-tuned, this BERT model, denoted as $BERT(D_R)$, is used to predict sentiment labels for the sentences in the synthetic dataset D_S. The quality of D_S is then quantitatively assessed by comparing the accuracy of these predictions to the actual sentiment labels of D_S. High prediction accuracy signifies that D_S closely resembles D_R, indicating the high quality of the synthetic data D_S.

The ACR metric is mathematically defined as:

$$ACR = \frac{N_{correct}(D_S)}{N_{total}(D_S)}, \tag{12}$$

where $N_{correct}(D_S)$ represents the number of sentences in D_S where the sentiment label predicted by $BERT(D_R)$ matches the actual sentiment label, and $N_{total}(D_S)$ represents the total number of sentences in D_S. The ACR metric yields a value that falls within the range of 0 to 1.

In summary, the ACR metric offers a performance-oriented evaluation of the synthetic data, thereby adding a practical perspective to the appraisal of the data generation process. This metric not only gauges the ability of the synthetic data to mirror the sentiment distribution of the real data, but it also provides insights into the utility of the synthetic data for downstream applications, such as sentiment analysis.

4.3.3. LXC metric

The LXC metric has been designed as a composite measure to evaluate the richness and relevance of the synthetic data D_S to the domain of finance. Specifically, it considers the lexical diversity of D_S and the inclusion of terms from a Financial Terminology Glossary (FTG). This metric is composed of two distinct components.

1) **LXC_1: Lexical Diversity of Synthetic Data**.

The lexical diversity of synthetic data D_S is measured as the ratio of unique words to the total word count in D_S. This measure of diversity reflects the range of vocabulary utilized in the synthetic data. A higher lexical diversity signifies a more extensive use of vocabulary, thereby enhancing the richness and complexity of the generated text, which is desirable when attempting to mimic real-world financial text.

The mathematical computation of LXC_1 is:

$$LXC_1 = \frac{N_{unique}(D_S)}{N_{total}(D_S)}, \tag{13}$$

where $N_{unique}(D_S)$ represents the count of unique words in D_S, and $N_{total}(D_S)$ denotes the total count of words in D_S. The value of LXC_1 falls within the range of 0 to 1.

2) **LXC_2: Appearance Rate of Financial Terminology Glossary in Synthetic Data.**

The frequency of terms from the Financial Terminology Glossary (*FTG*) appearing within the synthetic data (D_S) offers a valuable metric for evaluating the quality of the generated text. A higher frequency of *FTG* terms in D_S indicates a significant presence of financial terms, suggesting a closer approximation to the language and style of actual financial text.

LXC_2 is mathematically computed as:

$$LXC_2 = \frac{N_{FTG}(D_S)}{N_{total}(D_S)}, \tag{14}$$

where $N_{FTG}(D_S)$ represents the count of words in D_S that also appear in *FTG*, and $N_{total}(D_S)$ denotes the total count of words in D_S. The value of LXC_2 is bounded between 0 and 1.

The composite *LXC* metric is formulated as a weighted sum of LXC_1 and LXC_2, yielding a final result that falls within the range of 0 to 1:

$$LXC = u_1 \times LXC_1 + u_2 \times LXC_2, \tag{15}$$

where u_1 and u_2 are the weights allocated to LXC_1 and LXC_2, respectively, and are subject to the constraint $u_1 + u_2 = 1$. These weights can be flexibly adjusted to cater to the specific requirements of the task at hand.

4.4. Optimal prompt generation and evaluation

The process of prompt generation operates iteratively over a pre-specified number of cycles, denoted as *L*. In each cycle, a designated prompt, symbolized as *pmt*, instructs our Language Likelihood Model (LLM) to generate a corresponding synthetic dataset, denoted D_S. Upon the conclusion of these iterative cycles, the optimal policy π^* governing the Selector *Agent* is determined.

Once π^* is identified, the Selector *Agent* is applied to the Keywords Vocabulary, as illustrated in the "Testing" phase in Fig. 1. The optimal prompt *pmt** is synthesized by combining the selected keywords with the Prompt Template. This assembled prompt is applied to the Executor *LLM* multiple times to generate numerous synthetic data, collectively represented as D_S. The volume of D_S is sufficiently expansive to fulfill the requirements of subsequent sentiment analysis tasks.

For the empirical evaluation of the synthetic data, we construct an experimental setup involving the training of two BERT models: one utilizes the synthetic dataset D_S, and the other employs a subset of the real dataset D_R, equated in size to D_S. The efficacy of these models is then assessed through their prediction accuracy on two intact real financial datasets. This experimental design furnishes a quantitative methodology for evaluating the quality and utility of the synthetic data, specifically in relation to sentiment analysis applications. The results thereby provide valuable insights into the efficacy of the synthetic data generation process.

5. Synthetic data generation task

In this section, we delineate the experiment of generating financial synthetic data, detail the experimental setup, and discuss the results obtained from the Reinforcement Prompting.

5.1. Real financial dataset for training

The primary dataset employed in our study is the Financial PhraseBank Dataset (Malo et al. (2014)). This dataset comprises sentences extracted from financial news, meticulously curated from the LexisNexis database, and reflects a substantial investment of domain-specific expertise and manual labor. A total of 4, 846 English sentences were randomly sampled and annotated by a cohort of 16 annotators possessing substantial experience in the fields of finance and business.

To mitigate the potential for cognitive overload, the dataset was partitioned into manageable segments, which were then distributed over the duration of the annotation process. On average, each annotator was responsible for annotating approximately 1, 500 sentences. This distributed approach ensured that each sentence received evaluations from between 5 and 8 annotators.

Furthermore, the dataset incorporates metadata on the levels of inter-annotator agreement, offering valuable insights into the reliability and consistency of the annotation process. Of the 4, 846 annotated sentences, 2, 262 achieved unanimous agreement among the annotators, while 3, 453 achieved a consensus rate exceeding 75% for their corresponding sentiment labels. Representative samples from this dataset are showcased in Table 3. The rigor and comprehensiveness of this data collection and annotation process confer a high degree of robustness and reliability upon the Financial PhraseBank Dataset, rendering it an exemplary resource for our experimental endeavors.

5.2. Financial Terminology Glossary

In addition to deriving a portion of this glossary from Investopedia's financial dictionary, we enrich it with terms from a range of databases to build our comprehensive Financial Terminology Glossary.

We incorporate stock symbols into the glossary from three leading U.S. stock exchanges: the American Stock Exchange (AMEX), the NASDAQ Exchange, and the New York Stock Exchange (NYSE). The integration of these stock symbols enhances the practical applicability of our glossary.

Consequently, the glossary serves as a thorough compendium of financial language, encompassing both general financial terms and particular stock symbols. It thus plays a pivotal role in the Reinforcement Prompting methodology.

Table 4 presents a snapshot of the Financial Terminology Glossary, illustrating a selection of terms sourced from different databases. Please note that the actual glossary is far more comprehensive, accommodating an extensive collection of financial terms and stock symbols. This broad-spectrum glossary is aggregated into a single file, thereby enabling its ready deployment in the generation of high-quality synthetic financial texts via the Reinforcement Prompting process.

5.3. Selection of language models

The architecture of our Reinforcement Prompting method allows for the potential utilization of a variety of large language models (LLMs) to instantiate the Executor *LLM*. Among the considered alternatives are OpenAI's ChatGPT, Anthropic's Claude, and Google's Bard, each offering distinct advantages in terms of linguistic competence and ease of integration.

After conducting a rigorous evaluation, we elect to use OpenAI's ChatGPT as the Executor *LLM*. Several factors influence this decision. Primarily, ChatGPT provides API access to its gpt-3.5-turbo architecture, thereby facilitating the integration and execution within the framework of our Reinforcement Prompting methodology. Furthermore, the linguistic capabilities of ChatGPT are well-aligned with our objectives for generating high-quality synthetic financial data.

In parallel, we also identify appropriate models for the supplementary components of our methodology. Specifically, for computing the *ACR* reward metric in each iteration, we utilize a BERT model pre-trained on the Financial PhraseBank dataset, commonly referred to as FinBERT (Araci (2019)). Employing FinBERT obviates the necessity of training a BERT model from scratch, thereby saving computational resources and simplifying the overall workflow.

For the downstream task of sentiment analysis, applied to the synthetic data produced by our methodology, we select the untrained bert-base-uncased model. This model is then trained on our synthetic dataset and subsequently deployed to assess the quality of the generated data, with respect to the specific task of sentiment analysis.

5.4. Results of Reinforcement Prompting implementation

We implement our Reinforcement Prompting methodology with the *Agent* initially equipped with a curated vocabulary of high-frequency keyword sets. These keywords are selected based on financial terminology from Investopedia and the pretraining data of ChatGPT.

During the iterative Reinforcement Prompting process, the *Agent* selects sets of keywords in each iteration and naturally combines them with a pre-defined Prompt Template. These prompts generated via this process then guide the Executor *LLM* to produce synthetic data samples. For the training phase, we employ a low-temperature setting of GPT-3.5-turbo to ensure stable rewards at the final step of each iteration.

We set the total number of episodes at 1, 000, aiming to achieve a balance between extensive exploration and optimization, while also considering computational constraints. Fig. 2 depicts the cumulative rewards accrued for each episode, indicating that the training process stabilizes after approximately 650 epochs.

Table 5 presents the keywords selected through the Reinforcement Prompting process. These chosen keywords offer insights into the quality of the generated prompt, serving as additional evidence of the methodology's effectiveness.

A comprehensive prompt, as displayed in Table 6 and is subsequently employed to instruct GPT-3.5-turbo multiple times. Contrary to the training phase, here we use a high-temperature setting to encourage the generation of diverse financial synthetic data points. The resulting synthetic data samples can be further utilized for downstream tasks.

6. Sentiment analysis: a downstream task

In the subsequent phase of our research, we deploy the optimized prompt, obtained through the Reinforcement Prompting methodology, to instruct the Executor *LLM* in generating 2, 300 synthetic financial data samples. With those synthetic data, we conduct a downstream sentiment analysis task, comparing the performance of BERT models trained on the synthetic data and the real Financial Phrasebank data, respectively. The comparisons are conducted using the Twitter Financial Dataset and the Fin-news Financial Dataset. Due to the character output limitations enforced by OpenAI's API, we conduct multiple prompting iterations to accumulate enough volume of synthetic data. A subset of these generated samples is presented in Table 7.

To provide better insight into the nature of the synthetic samples being generated, we first analyze the characteristics of the real dataset. The Financial PhraseBank dataset, as shown in Table 3, exhibits several key qualities:

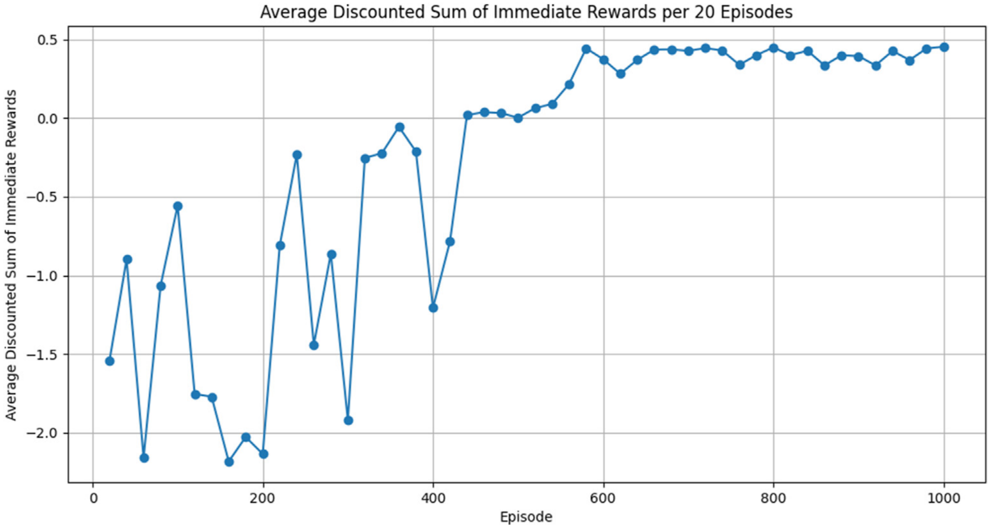

Fig. 2. Reinforcement prompting cumulative rewards for each episode.

- Complexity of Text Instances: Sentences vary in structure, from simple statements to more complex descriptions involving multiple clauses.
- Lexical Diversity: The dataset includes specific financial terminology and diverse vocabulary.
- Varied Sentiment Labels: Sentences are labeled with positive, neutral, or negative sentiments, reflecting a range of opinions and tones.

Our generated synthetic dataset, presented in Table 7, mirrors these characteristics. The synthetic sentences demonstrate similar sentence structure complexity and lexical diversity, incorporating specific financial contexts and terminology, and varying sentiment labels. This alignment ensures that the synthetic data maintains the utility and richness of the real data, making it suitable for downstream tasks.

Our experimental design entails the training of two separate BERT models on incrementally larger datasets, both synthetic datasets and real Financial Phrasebank datasets. These datasets contain 1, 000, 1, 100, 1, 200, 1, 300, 1, 400, 1, 500, 1, 600, 1, 700, and 1, 812 samples, respectively. Each training dataset is carefully curated to maintain a balanced distribution of positive, negative, and neutral sentiments. Due to the Financial Phrasebank being a pre-existing dataset, the maximum number of balanced sentences obtainable is 1, 812, thereby setting the upper limit for training samples. To enhance model performance, 10% of each training dataset is allocated for validation. For testing, we employ the Twitter Financial Dataset and the Fin-news Financial Dataset, both accessible at https://www.huggingface.com.

6.1. Analysis on Twitter Financial Dataset

We evaluate the performance of BERT models trained on synthetic data (BERT-Syn) and the Financial Phrasebank dataset (BERT-Real) using the Twitter Financial Dataset as the test dataset. This dataset is trimmed to 5,064 data points from 11, 882 samples to maintain a balanced distribution of positive, negative, and neutral sentiments. Our evaluation metrics encompass accuracy, precision, and recall, and we consider varying training sample sizes ranging from 1, 000 to 1, 812. These metrics for both BERT models are shown in Tables 8–10. This analysis aims to provide a thorough assessment of the utility and reliability of synthetic and real training data for financial sentiment analysis in social media contexts.

The summary of overall accuracy is presented in Table 8. Observations from the table indicate that BERT-Syn consistently outperforms BERT-Real across various training set sizes when tested against the Twitter Financial Dataset. Specifically, BERT-Syn attains its highest accuracy of 0.67 with a training set size of 1, 500 data points, while BERT-Real peaks at 0.64 with 1, 600 data points. Interestingly, the BERT-Syn model demonstrates relatively stable performance, with accuracy ranging from 0.59 to 0.67. In contrast, BERT-Real exhibits a bit more variance, ranging from 0.53 to 0.64. This suggests that the synthetic data may provide a more consistent training signal for the model.

Table 8
Sentiment analysis accuracy comparison: BERT Models Trained on Generated Synthetic Data (BERT-Syn) vs. Financial Phrasebank Real Data (BERT-Real), tested on a balanced twitter financial dataset comprising 5064 data points, equally distributed across sentiment labels.

Accuracy	1000	1100	1200	1300	1400	1500	1600	1700	1812
BERT-Syn	0.65	0.64	0.59	0.66	0.66	0.67	0.63	0.65	0.64
BERT-Real	0.56	0.57	0.53	0.56	0.55	0.53	0.64	0.53	0.57

Table 9
Precision scores of different sentiments comparison: BERT Models Trained on Generated Synthetic Data (BERT-Syn) vs. Financial Phrasebank Real Data (BERT-Real), tested on a balanced twitter financial dataset comprising 5064 data points, equally distributed across sentiment labels.

Precision	1000	1100	1200	1300	1400	1500	1600	1700	1812
BERT-Syn “Positive”	0.69	0.69	0.81	0.57	0.69	0.62	0.83	0.68	0.56
BERT-Real “Positive”	0.69	0.76	0.63	0.65	0.48	0.76	0.73	0.59	0.72
BERT-Syn “Negative”	0.78	0.83	0.86	0.74	0.87	0.81	0.85	0.83	0.83
BERT-Real “Negative”	0.80	0.79	0.86	0.78	0.81	0.84	0.73	0.86	0.78
BERT-Syn “Neutral”	0.54	0.53	0.46	0.78	0.55	0.64	0.49	0.54	0.64
BERT-Real “Neutral”	0.45	0.45	0.43	0.45	0.52	0.42	0.52	0.44	0.45

Table 10
Recall scores of different sentiments comparison: BERT Models Trained on Generated Synthetic Data (BERT-Syn) vs. Financial Phrasebank Real Data (BERT-Real), tested on a balanced twitter financial dataset comprising 5064 data points, equally distributed across sentiment labels.

Recall	1000	1100	1200	1300	1400	1500	1600	1700	1812
BERT-Syn “Positive”	0.64	0.65	0.45	0.88	0.67	0.77	0.46	0.63	0.78
BERT-Real “Positive”	0.46	0.45	0.48	0.41	0.68	0.30	0.51	0.55	0.41
BERT-Syn “Negative”	0.59	0.46	0.40	0.66	0.50	0.59	0.52	0.52	0.51
BERT-Real “Negative”	0.36	0.40	0.31	0.47	0.41	0.39	0.68	0.30	0.47
BERT-Syn “Neutral”	0.72	0.80	0.91	0.44	0.81	0.67	0.90	0.79	0.65
BERT-Real “Neutral”	0.85	0.86	0.81	0.80	0.57	0.91	0.72	0.74	0.82

The BERT models are fine-tuned using a limited number of samples (ranging from 1, 000 to 1, 812) of either synthetic data or real data from the Financial Phrasebank. Their performances are evaluated on the Twitter Financial Dataset, which has not been seen by the fine-tuned models before. This study's objectives differ from previous works, such as those involving FinBERT for financial sentiment analysis (Araci (2019); Huang et al. (2022)), where the models are typically pre-finetuned on a large corpus of financial data and then trained and tested within the same domain. Consequently, the performance of our results (Table 8)—where models are trained and tested on two different datasets—is naturally lower than that of the FinBERT models trained and tested on the same dataset (the Twitter Financial Dataset). Our primary aim is to evaluate the effectiveness of synthetic data generated through Reinforcement Prompting for training BERT models in financial sentiment analysis. To achieve this, we design our experiments to test the models' ability to perform across different financial text corpora (testing datasets in different distributions). Unlike the FinBERT models, our BERT models are not pre-finetuned on any financial data before being fine-tuned on the synthetic or real data samples. This approach provides a more realistic and challenging training scenario. Moreover, we fine-tune the BERT models using a small amount of data to simulate real-world conditions where financial data samples are often scarce due to privacy concerns and accessibility issues. These requirements naturally lead to lower accuracy results. However, this lower accuracy is expected and necessary, as the BERT trained on the comparable real Financial Phrasebank dataset (BERT-Real) has a similarly low accuracy performance. By designing the experiments in this way, we aim to ensure that the synthetic datasets are useful across multiple different datasets, making them valuable in various financial contexts where real data is limited or difficult to obtain. This demonstrates that our approach provides a practical and realistic evaluation of synthetic data's effectiveness, emphasizing its potential for broader applicability rather than just optimizing for a single dataset.

Table 9 details the precision scores for different sentiment categories when evaluated on the Twitter Financial Dataset. Both BERT-Syn and BERT-Real models exhibit similar performance in identifying “Positive” sentiments. However, BERT-Syn generally outperforms BERT-Real in capturing “Negative” sentiments, achieving a peak precision score of 0.87 with a training set of 1, 400 data points. For “Neutral” sentiments, both models exhibit relatively lower precision scores, with BERT-Syn slightly outperforming BERT-Real.

Table 10 illustrates the recall scores for the Twitter Financial Dataset. In the cases of “Positive” and “Negative” sentiments, BERT-Syn's recall metrics are generally comparable to or better than those of BERT-Real. However, for “Neutral” sentiments, BERT-Real exhibits a higher recall, peaking at 0.91 with a training set of 1, 500 data points. This observation suggests that BERT-Real may be more sensitive to neutral sentiments within the dataset.

6.2. Analysis on Fin-news financial dataset

We extend our analysis to the performance of BERT models trained on synthetic financial data (BERT-Syn) and Financial Phrasebank real data (BERT-Real) when tested on the Fin-news Financial Dataset. The dataset comprises 2, 103 datapoints from 8, 675 samples, with a balanced distribution across the three sentiment classes: positive, negative, and neutral. The metrics considered include accuracy, precision, and recall across varying numbers of training samples (1, 000, 1, 100, 1, 200, 1, 300, 1, 400, 1, 500, 1, 600, 1, 700, 1, 812). The results are depicted in Tables 11–13.

The accuracy results on the Fin-news Financial Dataset, as summarized in Table 11, reveal distinct patterns in comparison to those observed on the Twitter Financial Dataset. Both BERT-Syn and BERT-Real exhibit fluctuating performance with varying training set sizes. BERT-Real achieves its highest accuracy of 0.90 with training sets of 1, 600 and 1, 812 data points, while BERT-Syn peaks at 0.84 with 1, 300 data points.

Table 11
Sentiment analysis accuracy comparison: BERT Models Trained on Generated Synthetic Data (BERT-Syn) vs. Financial Phrasebank Real Data (BERT-Real), tested on a balanced fin-news financial dataset comprising 2103 data points, equally distributed across sentiment labels.

Accuracy	1000	1100	1200	1300	1400	1500	1600	1700	1812
BERT-Syn	0.82	0.64	0.57	0.84	0.73	0.78	0.79	0.75	0.74
BERT-Real	0.72	0.72	0.70	0.75	0.81	0.86	0.90	0.60	0.90

Table 12
Precision scores of different sentiments comparison: BERT Models Trained on Generated Synthetic Data (BERT-Syn) vs. Financial Phrasebank Real Data (BERT-Real), tested on a balanced fin-news financial dataset comprising 2103 data points, equally distributed across sentiment labels.

Precision	1000	1100	1200	1300	1400	1500	1600	1700	1812
BERT-Syn "Positive"	0.80	0.90	0.90	0.74	0.90	0.91	0.93	0.84	0.63
BERT-Real "Positive"	0.91	0.91	0.61	0.60	0.67	0.96	0.91	0.50	0.92
BERT-Syn "Negative"	0.92	0.98	0.94	0.92	0.97	0.95	0.96	0.97	0.98
BERT-Real "Negative"	0.98	0.96	0.98	0.97	0.99	0.98	0.95	1.00	0.95
BERT-Syn "Neutral"	0.76	0.48	0.44	0.91	0.57	0.63	0.63	0.60	0.76
BERT-Real "Neutral"	0.56	0.55	0.63	0.78	0.88	0.72	0.86	0.66	0.84

Table 13
Recall scores of different sentiments comparison: BERT Models Trained on Generated Synthetic Data (BERT-Syn) vs. Financial Phrasebank Real Data (BERT-Real), tested on a balanced fin-news financial dataset comprising 2103 data points, equally distributed across sentiment labels.

Recall	1000	1100	1200	1300	1400	1500	1600	1700	1812
BERT-Syn "Positive"	0.84	0.51	0.50	0.93	0.58	0.60	0.64	0.78	0.91
BERT-Real "Positive"	0.64	0.73	0.88	0.90	0.95	0.78	0.86	0.94	0.85
BERT-Syn "Negative"	0.83	0.45	0.27	0.94	0.67	0.79	0.77	0.58	0.61
BERT-Real "Negative"	0.59	0.47	0.67	0.87	0.85	0.82	0.96	0.42	0.95
BERT-Syn "Neutral"	0.80	0.96	0.95	0.67	0.95	0.94	0.95	0.88	0.70
BERT-Real "Neutral"	0.94	0.94	0.55	0.47	0.64	0.98	0.90	0.46	0.92

Interestingly, BERT-Real outperforms BERT-Syn when trained on larger data sets (1, 500, 1, 600, and 1, 812), implying that real financial data may offer a more nuanced feature representation for this specific testing dataset.

Table 12 details the precision scores for different sentiments on Fin-news Financial Dataset. For "Positive" sentiments, both models show competitive performance, but BERT-Syn excels in the medium training set sizes (1, 300-1, 500). In the case of "Negative" sentiments, both models perform exceptionally well, but BERT-Real slightly outperforms BERT-Syn. For "Neutral" sentiments, BERT-Real generally shows higher precision, particularly in the larger training set sizes.

Recall scores, as presented in Table 13, offer additional insights for the testing result on Fin-news Financial Dataset. BERT-Syn and BERT-Real both show high recall scores for "Positive" and "Neutral" sentiments, but BERT-Real generally outperforms BERT-Syn for "Negative" sentiments, especially in the larger training set sizes.

The difference in performance between this dataset and the Twitter Financial Dataset can be attributed to the following factors:

- Dataset Characteristics: The Financial Phrasebank data, used for training the models, shares a more similar structure and distribution with the Fin-news Financial Dataset, compared to Twitter Financial Dataset. This similarity allows models trained on the Financial Phrasebank to perform better when tested on the Fin-news Financial Dataset.
- Nature of Twitter Data: In contrast, the Twitter Financial Dataset consists of tweets that are generally shorter, less structured, and contain more informal language. The distribution of sentiments and the style of text in Twitter data differ significantly from the structured format of news articles. As a result, models trained on the Financial Phrasebank or synthetic dataset may struggle to generalize well to the Twitter Financial Dataset, leading to lower performance than testing on Fin-news Financial Dataset.

6.3. Analysis of the variance

For both accuracy results of the Twitter Financial Dataset (Table 8) and the Fin-news Financial Dataset (Table 11), the variances between some of the results tend to be large for different numbers of training samples. For example, in Table 8, the BERT-Real achieves 0.53 in accuracy when fine-tuned with 1, 500 data samples, achieves 0.64 in accuracy when fine-tuned with 1, 600 data samples, then drops to 0.53 again with 1, 700 data samples. To analyze the cause of the variance, we repeat the experiments on fine-tuning the BERT models with different numbers of training samples, and we observe this variance present in every run of the experiment. Based on these observations, we can draw the following conclusions:

- **Impact of Sample Size:** The primary cause of this variance is the small sample size used for training. In scenarios with limited data, the model's performance can be significantly impacted by the specific characteristics of the training samples, leading to higher variability. This is a well recognized phenomenon in machine learning research, where smaller datasets often result in less stable models.
- **Variance in Real Data:** It is important to note that this variance is not exclusive to synthetic data; it is also observed in models trained on real data. To illustrate this point, we conducted multiple experiments using both the generated synthetic data and real data from the Financial Phrasebank Dataset, and then tested them on the Twitter Financial Dataset. Table 14 presents the accuracy results from three different runs of the experiment, highlighting the inherent variability.
- **Averaged Performance:** To mitigate the impact of this variance, we averaged the results of 30 runs for each training sample size for the experiments on the Twitter Financial Dataset. Table 15 presents the averaged accuracy results, showing that the performance of models trained on synthetic data closely matches those trained on real data. This demonstrates the high quality of the synthetic data generated through the Reinforcement Prompting methodology.

These observations suggest that the variance is primarily due to the small sample sizes used for training rather than the nature of the synthetic data itself. By averaging the results across multiple runs, we can obtain a more reliable assessment of the model performance, thereby validating the effectiveness of our synthetic data.

6.4. Case study: human evaluation

We conducted a case study comparing real and synthetic datasets in terms of syntactic and semantic quality. This evaluation aims to assess the reliability and effectiveness of the generated synthetic data.

6.4.1. Human evaluation setup

We select 1, 000 sample sentences from both the real Financial Phrasebank dataset and the generated synthetic dataset. A group of three human annotators with expertise in financial texts is asked to evaluate these samples based on the following metrics:

- Semantic Coherence: The logical consistency and meaningfulness of the sentences.
- Domain Relevance: The applicability and relevance of the sentences to the financial domain.

Annotators are instructed to rate each sentence on a scale from 1 to 5 for each criterion, where 1 indicates poor quality and 5 indicates excellent quality. All annotators who participated in the data processing for this study were informed about the nature of the research and provided their consent for the use of their annotations. The process was confirmed exempt from ethical approval, and their participation was entirely voluntary.

6.4.2. Annotator agreement and results

To quantify the agreement among annotators, we calculat the average scores for each criterion. The results are summarized in Table 16.

Table 14
Accuracy results from multiple runs on twitter financial dataset.

Accuracy	1000	1100	1200	1300	1400	1500	1600	1700	1812
BERT-Syn Run 1	0.65	0.64	0.59	0.66	0.66	0.67	0.63	0.65	0.64
BERT-Syn Run 2	0.66	0.64	0.61	0.67	0.59	0.68	0.64	0.66	0.65
BERT-Syn Run 3	0.64	0.67	0.58	0.69	0.65	0.66	0.62	0.64	0.63
BERT-Real Run 1	0.56	0.57	0.53	0.56	0.55	0.53	0.64	0.53	0.57
BERT-Real Run 2	0.63	0.56	0.56	0.59	0.67	0.57	0.63	0.64	0.68
BERT-Real Run 3	0.58	0.62	0.59	0.57	0.54	0.69	0.52	0.67	0.55

Table 15
Average sentiment analysis accuracy comparison: BERT Models Trained on Generated Synthetic Data (BERT-Syn) vs. Financial Phrasebank Real Data (BERT-Real), tested on a balanced twitter financial dataset comprising 5064 data points, equally distributed across sentiment labels.

Accuracy	1000	1100	1200	1300	1400	1500	1600	1700	1812
BERT-Syn	0.66	0.65	0.60	0.64	0.66	0.68	0.66	0.67	0.67
BERT-Real	0.57	0.61	0.58	0.59	0.61	0.59	0.62	0.60	0.61

Table 16
Annotator agreement on real and synthetic datasets.

Metric	Synthetic Data	Real Data
Semantic Coherence	4.3	4.5
Domain Relevance	4.1	4.2

The scores indicate the synthetic data closely matches the quality of the real data in terms of semantic coherence, and domain relevance.

6.5. Synthetic data versus real data: a comparative summary

We conduct a comprehensive analysis of BERT models trained on synthetic data (BERT-Syn) and real financial data from the Financial Phrasebank (BERT-Real), benchmarking their performance against two distinct testing datasets: the Twitter Financial Dataset and the Fin-news Financial Dataset. Our experiments reveal that BERT models trained on synthetic data demonstrate performance metrics that are not only comparable but occasionally even surpass those of models trained on real financial data in the context of sentiment analysis tasks (see Tables 8 and 11).

Our empirical observations underscore several advantages of employing synthetic data:

- **Consistency**: As evidenced by the Fin-news Financial Dataset, BERT-Syn exhibits less variability in accuracy metrics, suggesting a more stable training signal.
- **Flexibility**: Our experiments demonstrate that synthetic data allows for extensive experimentation with balanced samples, which is crucial when the real data is scarce. Specifically in our experiments, we generate multiple synthetic datasets and select balanced samples for training, which help achieve more reliable and fair performance evaluations.
- **Cost-Effectiveness**: The generation of synthetic data often proves to be more economically viable than the collection and annotation of real-world data, particularly in specialized sectors such as finance.

However, it is also important to acknowledge the advantages of using real data:

- **Authenticity**: Real data inherently reflects the true distribution and complexity of real-world scenarios, ensuring that models trained on such data are well-tuned to actual events.
- **Diversity**: Real data encompasses the full range of variability found in actual financial texts, capturing rare events and edge cases that synthetic data might miss.
- **Legal and Compliance Considerations**: Using real data helps ensure that models adhere to regulatory standards and compliance requirements, which is crucial in sensitive fields like finance.
- **Trust and Acceptance**: Models trained on real data may be perceived as more trustworthy and acceptable by stakeholders, as they rely on actual, verifiable information.

In summary, while synthetic data offers notable benefits in terms of consistency, flexibility, and cost-effectiveness, real data remains crucial for its authenticity, diversity, compliance, and stakeholder trust. A balanced approach that leverages both synthetic and real data can maximize the strengths of each, leading to more robust and versatile models for sentiment analysis tasks in the financial domain.

7. Conclusion

This work presents a novel approach—*Reinforcement Prompting*—for the generation of high-quality financial synthetic data for sentiment analysis tasks. By integrating the advantages of Large Language Models (LLMs) with the optimization techniques of Reinforcement Learning, the methodology provides a solution to the challenges associated with data privacy and the scarcity of labeled data.

The methodology distinguishes itself by deploying a neural network as a Selector for prompt generation and an LLM as an Executor for synthetic data generation. The generation of prompts is guided by a reward function that considers the quality of the synthetic data, with the objective of identifying the prompt that engenders the highest quality synthetic data.

The implementation of Reinforcement Prompting not only alleviates the dependency on real-world labeled data but also furnishes a strategy that is inherently privacy-preserving. This improvement renders the methodology a valuable tool for financial sentiment analysis, providing a mechanism to extract valuable insights from large volumes of unstructured data, without compromising data privacy or necessitating extensive labeled data.

Therefore, this work signifies a considerable advancement in the utilization of Large Language Models within the financial domain, paving the way for novel research opportunities and applications.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

This work was supported in part by NSF Project CCF-2416361. We would like to thank the reviewers and the editor-in-chief for their valuable comments and suggestions that have improved the quality of this paper.

References

Alissa, K., Alzoubi, O., 2022. Financial sentiment analysis based on transformers and majority voting. In: 2022 IEEE/ACS 19th International Conference on Computer Systems and Applications (AICCSA). IEEE.

Antoniou, A., Storkey, A., Edwards, H., 2017. Data augmentation generative adversarial networks. arXiv preprint arXiv:1711.04340.

Araci, D., 2019. Finbert: Financial Sentiment Analysis with Pre-trained Language Models, 10063 arXiv:1908.

Awatramani, P., Daware, R., Chouhan, H., Vaswani, A., Khedkar, S., 2021. Sentiment analysis of mixed-case language using natural language processing. In: 2021 Third International Conference on Inventive Research in Computing Applications (ICIRCA). IEEE.

Brown, T., Mann, B., Ryder, N., Subbiah, M., Kaplan, J.D., Dhariwal, P., Neelakantan, A., Shyam, P., Sastry, G., Askell, A., Agarwal, S., Herbert-Voss, A., Krueger, G., Henighan, T., Child, R., Ramesh, A., Ziegler, D., Wu, J., Winter, C., Hesse, C., Chen, M., Sigler, E., Litwin, M., Gray, S., Chess, B., Clark, J., Berner, C., McCandlish, S., Radford, A., Sutskever, I., Amodei, D., 2020. Language models are few-shot learners. In: Larochelle, H., Ranzato, M., Hadsell, R., Balcan, M., Lin, H. (Eds.), Advances in Neural Information Processing Systems. Curran Associates, Inc., pp. 1877–1901. URL: https://proceedings.neurips.cc/paper_files/paper/2020/file/1457c0d6bfcb4967418bfb8ac142f64a-Paper.pdf

Bryant, R., Cintas, C., Wambugu, I., Kinai, A., Weldemariam, K., 2019. Analyzing bias in sensitive personal information used to train financial models. arXiv preprint arXiv:1911.03623.

Deng, X., Bashlovkina, V., Han, F., Baumgartner, S., Bendersky, M., 2023. What do llms know about financial markets? a case study on reddit market sentiment analysis. In: Companion Proceedings of the ACM Web Conference 2023. https://doi.org/10.1145/3543873.3587324.

Devlin, J., Chang, M.W., Lee, K., Toutanova, K., 2019. 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). Association for Computational Linguistics, Minneapolis, Minnesota, pp. 4171–4186. https://doi.org/10.18653/v1/N19-1423. URL: https://aclanthology.org/N19-1423.

Dognin, P., Padhi, I., Melnyk, I., Das, P., 2021. Regen: reinforcement learning for text and knowledge base generation using pretrained language models. In: Proceedings of the 2021 Conference on Empirical Methods in Natural Language Processing. https://doi.org/10.18653/v1/2021.emnlp-main.83.

Donahue, D., Rumshisky, A., 2018. Adversarial text generation without reinforcement learning. arXiv preprint arXiv:1810.06640.

Dwork, C., 2006. Differential privacy. In: Bugliesi, M., Preneel, B., Sassone, V., Wegener, I. (Eds.), Automata, Languages and Programming. Springer Berlin Heidelberg, Berlin, Heidelberg, pp. 1–12.

Efimova, V., Shalamov, V., Filchenkov, A., 2020. Synthetic dataset generation for text recognition with generative adversarial networks. In: Osten, W., Nikolaev, D.P. (Eds.), Twelfth International Conference on Machine Vision (ICMV 2019). SPIE.

Gururangan, S., Marasović, A., Swayamdipta, S., Lo, K., Beltagy, I., Downey, D., Smith, N.A., 2020. Don't stop pretraining: adapt language models to domains and tasks. In: Jurafsky, D., Chai, J., Schluter, N., Tetreault, J. (Eds.), Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics. Association for Computational Linguistics, Online, pp. 8342–8360. https://doi.org/10.18653/v1/2020.acl-main.740. URL: https://aclanthology.org/2020.acl-main.740.

Hoffmann, J., Borgeaud, S., Mensch, A., Buchatskaya, E., Cai, T., Rutherford, E., Casas, D.d.L., Hendricks, L.A., Welbl, J., Clark, A., et al., 2022. Training compute-optimal large language models. arXiv preprint arXiv:2203.15556.

Huang, A., Wang, H., Yang, Y., 2022. Finbert: a large language model for extracting information from financial text†. Contemporary Accounting Research 40. https://doi.org/10.1111/1911-3846.12832.

Ji, Y., Zhang, Z., Tang, X., Shen, J., Zhang, X., Yang, G., 2022. Detecting cash-out users via dense subgraphs. In: Proceedings of the 28th ACM SIGKDD Conference on Knowledge Discovery and Data Mining. https://doi.org/10.1145/3534678.3539252.

Lu, D., Wu, H., Liang, J., Xu, Y., He, Q., Geng, Y., Han, M., Xin, Y., Xiao, Y., 2023. Bbt-fin: comprehensive construction of Chinese financial domain pre-trained language model, corpus and benchmark. https://doi.org/10.48550/ARXIV.2302.09432.

Lu, Y., Huang, X., Dai, Y., Maharjan, S., Zhang, Y., 2020. Blockchain and federated learning for privacy-preserved data sharing in industrial iot. IEEE Transactions on Industrial Informatics 16, 4177–4186. https://doi.org/10.1109/tii.2019.2942190.

Malo, P., Sinha, A., Korhonen, P., Wallenius, J., Takala, P., 2014. Good debt or bad debt: detecting semantic orientations in economic texts. Journal of the Association for Information Science & Technology 65, 782–796. URL: https://ideas.repec.org/a/bla/jinfst/v65y2014i4p782-796.html.

Mishra, S., Choubey, S., Choubey, A., Yogeesh, N., Durga Prasad Rao, J., William, P., 2022. Data extraction approach using natural language processing for sentiment analysis. In: 2022 International Conference on Automation, Computing and Renewable Systems (ICACRS). IEEE.

Narayanan, A., Shmatikov, V., 2008. Robust de-anonymization of large sparse datasets. In: 2008 IEEE Symposium on Security and Privacy (Sp 2008), pp. 111–125. https://doi.org/10.1109/SP.2008.33.

Presotto, R., Ek, S., Civitarese, G., Portet, F., Lalanda, P., Bettini, C., 2023. Combining Public Human Activity Recognition Datasets to Mitigate Labeled Data Scarcity. https://doi.org/10.48550/ARXIV.2306.13735.

Sixt, L., Wild, B., Landgraf, T., 2016. Rendergan: generating realistic labeled data. Frontiers in Robotics and AI 5, URL. https://api.semanticscholar.org/CorpusID:2483823.

Truong, A., Walters, A., Goodsitt, J., 2020. Sensitive data detection with high-throughput neural network models for financial institutions. arXiv preprint arXiv:2012.09597.

Upadhyay, B., Sudhakar, A., Maheswaran, A., 2022. Efficient Reinforcement Learning for Unsupervised Controlled Text Generation. https://doi.org/10.48550/ARXIV.2204.07696.

Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A.N., Kaiser, L.u., Polosukhin, I., 2017. Attention is all you need. In: Guyon, I., Luxburg, U.V., Bengio, S., Wallach, H., Fergus, R., Vishwanathan, S., Garnett, R. (Eds.), Advances in Neural Information Processing Systems. Curran Associates, Inc. URL. In: https://proceedings.neurips.cc/paper_files/paper/2017/file/3f5ee243547dee91fbd053c1c4a845aa-Paper.pdf.

Wei, J., Wang, X., Schuurmans, D., Bosma, M., Ichter, B., Xia, F., Chi, E., Le, Q., Zhou, D., 2022. Chain-of-thought prompting elicits reasoning in large language models. arXiv preprint arXiv:2201.11903.

Weng, W., Pratama, M., Za'in, C., de Carvalho, M., Appan, R., Ashfahani, A., Yee, E.Y.K., 2022. Autonomous cross domain adaptation under extreme label scarcity. IEEE Transactions on Neural Networks and Learning Systems PP, 1–12. https://doi.org/10.1109/TNNLS.2022.3183356.

Wu, S., Irsoy, O., Lu, S., Dabravolski, V., Dredze, M., Gehrmann, S., Kambadur, P., Rosenberg, D., Mann, G., 2023. Bloomberggpt: A Large Language Model for Finance. https://doi.org/10.48550/ARXIV.2303.17564.

Wu, Z., Wang, L., Wang, W., Shi, T., Chen, C., Hao, A., Li, S., 2022. Synthetic data supervised salient object detection. In: Proceedings of the 30th ACM International Conference on Multimedia. ACM.

Xie, L., Lin, K., Wang, S., Wang, F., Zhou, J., 2018. Differentially private generative adversarial network. CoRR abs/1802.06739. URL: http://arxiv.org/abs/1802.06739. arXiv:1802.06739.

Yang, H., Liu, X.Y., Wang, C.D., 2023. Fingpt: Open-Source Financial Large Language Models. https://doi.org/10.48550/ARXIV.2306.06031.

Zhang, M., Tian, G., Gao, H., Zhang, Y., 2022. Autonomous generation of service strategy for household tasks: a progressive learning method with a priori knowledge and reinforcement learning. IEEE Transactions on Circuits and Systems for Video Technology 32, 7473–7488. https://doi.org/10.1109/tcsvt.2022.3189357.