Investor sentiment from images: a few-shot learning investigation

Article

Accepted Version

Ren, X., Jiang, W., Sun, X. and Wang, S. ORCID: https://orcid.org/0000-0003-2113-5521 (2025) Investor sentiment from images: a few-shot learning investigation. Journal of Accounting Literature. ISSN 2452-1469 doi: 10.1108/JAL-04-2025-0199 Available at https://centaur.reading.ac.uk/123394/

To link to this article DOI: http://dx.doi.org/10.1108/JAL-04-2025-0199

Publisher: Emerald

Investor Sentiment from Images: A Few-Shot Learning Investigation

Purpose

This research aims to extract emotional features from New York Times news images (2018–2023) using few-shot learning approaches. Leveraging machine learning, it offers a systematic investigation into how image-driven emotions affect investor behavior in the U.S. equity market and contribute to the prediction of market movements.

Design

This study employs the DeepEMD model to extract emotional features from 181,233 news images, constructing a daily sentiment index based on visual media. By defining sentiment thresholds, the study develops differentiated strategies for positive and negative emotional signals. In addition, it integrates four machine learning models—AdaBoost, Support Vector Machine (SVM), ExtraTrees, and Random Forest (RF)—alongside a traditional linear regression model to forecast the prices of various U.S. stock market indices.

Findings This study finds that news image sentiment has a significant impact on financial markets. Positive sentiment strategies applied to serious news topics are associated with higher returns, whereas negative sentiment in entertainment-related content signals potential opportunities for contrarian investment. Moreover, the influence of image-based sentiment on the market exhibits a delayed effect of approximately 2–3 days, with particularly strong predictive power for small-cap stocks. Compared to traditional linear models, machine learning approaches demonstrate superior performance in capturing the nonlinear dynamics between sentiment and market behavior, offering novel analytical tools for behavioral finance research and sentiment-driven anomaly-based investment strategies.

Value: This study integrates visual data analysis into the domain of behavioral finance, highlighting the distinctive role of image-based sentiment in uncovering market anomalies and informing investment strategies.

Keywords: Few-shot Learning, Investor Sentiment, Behavioral Finance, Image Classification, Financial Forecasting

1 Introduction

The existing literature has limited studies on extracting sentiment from images. Instead, the research paradigm on investor sentiment is based on textual data. This is mainly due to the complexity of sentiment extraction from images. Text, as a human-created information carrier, inherently contains sentiment tendencies in each word from the outset (Kozyreva et al., 2020; Quillian, 1967). Through simple methods such as building sentiment dictionaries, we can reliably obtain the sentiment tendency of a piece of text (Jiang et al., 2019; Kearney and Liu, 2014). Images, as a highly intuitive form of information presentation, often divert readers' attention from text when they appear in view (Invernizzi et al., 2022). Psychology has shown that different images can induce varying sentiment in people (Marchewka et al., 2014; Dan-Glauser and Scherer, 2011). When readers encounter a particular image, sentiment signals are already embedded in the image. These sentiment signals may influence their analysis and judgment of everyday matters, causing them to deviate from the path of rational thinking (Pham, 2007; Fenton-O'Creevy et al., 2011). Therefore, studying the impact mechanism of sentiment in images is a valuable complement to the news text itself (Obaid and Pukthuanthong, 2022).

For image data, sentiment stimuli often directly impact viewers visually through factors such as content, color, and composition presented in the image (Zhao et al., 2021; Bhandari et al., 2019). These features are relatively abstract and difficult to quantify directly, making it challenging to identify sentiment polarity from images (Borth et al., 2013). Additionally, a large amount of image data often lacks manual annotation, which increases the difficulty of supervised learning (Moen et al., 2019). To address this issue, few-shot learning (FSL) has become an effective method. As a form of meta learning, FSL aims to classify image categories using only a limited amount of labeled training data to address the challenge of data scarcity (Song et al., 2023; Lu et al., 2023). Moreover, due to the data characteristics of few-shot learning, the algorithm is more efficient, enabling us to obtain model results more quickly.

In this study, we select news images from The New York Times (NYT) as the focus of our research. As a mainstream media outlet with over 150 years of history, NYT covers a wide range of international and domestic news, including politics, social events, and cultural phenomena (Van Belle, 2003; Benson, 2009). Daily news coverage encompasses numerous thematic sections, from which we have selected the most representative and abundant images from nine major sections (including "Arts", "Books", "Business Day", "New York", "Opinion", "Sports", "Technology", "U.S.", and "World") to explore sentiment features across different topics. Considering the lack of annotations for image data, we employ the DeepEMD model proposed by Zhang et al. (2022) for 5-Shot few-shot learning. This method utilizes the Earth Mover's Distance (EMD) as a metric for computing the structural distance between images and determining their relevance. By fine-tuning the pre-trained model of DeepEMD, we have performed sentiment analysis on images under each theme and derived daily sentiment scores for different topics.

When establishing different investment strategies based on changes in sentiment for each topic, we find that the positive sentiment brought by "Business Day," "Technology," and "World" contributes to relatively robust investment returns. These returns surpass the performance of the Dow Jones Industrial Average (DJI) and the Russell 2000 Index (RUT). The news on these three topics involves international economic and political events, thus exhibiting a more serious overall tone (Bianchi et al., 2016). Individuals may feel more distant from serious subjects, which can serve as a buffer reducing the direct impact of sentiment (Zheng et al., 2020). Therefore, the sentiment reflected in these themes can directly influence market prices. In contrast, when formulating investment strategies based on the sentiment conveyed in "Arts" and "Sports", we find that the negative sentiment brought by these topics signals an upward trend in the financial markets. News in these themes tends to be more entertainment-oriented, and entertainment topics are often closer to everyday life (Kross and Ayduk, 2017), making individuals more susceptible to extreme sentiment and resulting in a reverse effect of sentiment. The negative sentiment investors derive from them may lead to an erroneous underestimation of financial market prices, generating irrational investment aversion (Taffler, 2018). Investing at such times, however, may yield excess returns.

To further validate the characteristics of these thematic emotions, we examine their use in predicting stock market prices and observe that they exhibit a certain degree of persistence, potentially leading to temporary pricing errors (Stambaugh and Yuan, 2017). This stickiness implies that the influence of sentiment not only manifests in singular events but may also persistently affect market sentiment and investor decisions over a period, resulting in temporary irrational behaviors in the market (Griffith et al., 2020). To more accurately assess the impact of this sentiment on the stock market, we opt for machine learning models. Compared to traditional linear models, machine learning models can better capture nonlinear factors between variables, hence gaining significant recognition in academia (Dumitrescu et al., 2022; Khandani et al., 2010). Out-of-sample results demonstrate that these machine learning models have better generalization capabilities compared to linear models. Furthermore, we conduct predictions for different lag days and find that in experiments with lag days ranging from 2 to 3 days, some models show greater improvements, further validating the existence of certain lag effects of this sentiment. This stickiness of sentiment provides a deeper explanation for fluctuations in stock market prices. Accurately understanding the lagging period of sentiment helps us better comprehend future market changes.

In robustness tests, we conduct overall statistical analyses of this sentiment. The results indicate that sentiment in the news is more informative for investments in small-cap companies, especially during severe economic crises. Given that small-cap companies are in a growth phase, their asset volatility is considerable, making them susceptible to internal turmoil influenced by policy changes (Shynkevich, 2012). In our investment strategy based on news sentiment, we achieve a cumulative return of 11.95% for the RUT, significantly surpassing the benchmark's 3.79%. However, in investments related to large-cap stocks, our strategy does not diverge significantly from the benchmark. Therefore, by leveraging sentiment analysis to gauge the development status of small-cap companies, we can better mitigate risks.

The main contributions of this study are as follows. First, we develop a framework to employ a few-shot learning (FSL) technique to extract investor sentiment from news images among various news sections, which complements the research paradigm of sentiment analysis based on textual data. Our results demonstrate the interconnections between different news sections, showing that sentiment is often triggered by topics in World and Business Day sections. Second, we examine a simple investment strategy based on image sentiment and find the heterogeneous performance from various news sections. Lastly, we investigate the nonlinear relationship between sentiment and the US stock markets. In particular, sentiment does not immediately impact the financial markets, but with a noticeable delay. In addition, we find that news sentiment has greater relevance for small-cap firms, compared to large-cap firms. The main structure of this paper is illustrated in Fig. 1.

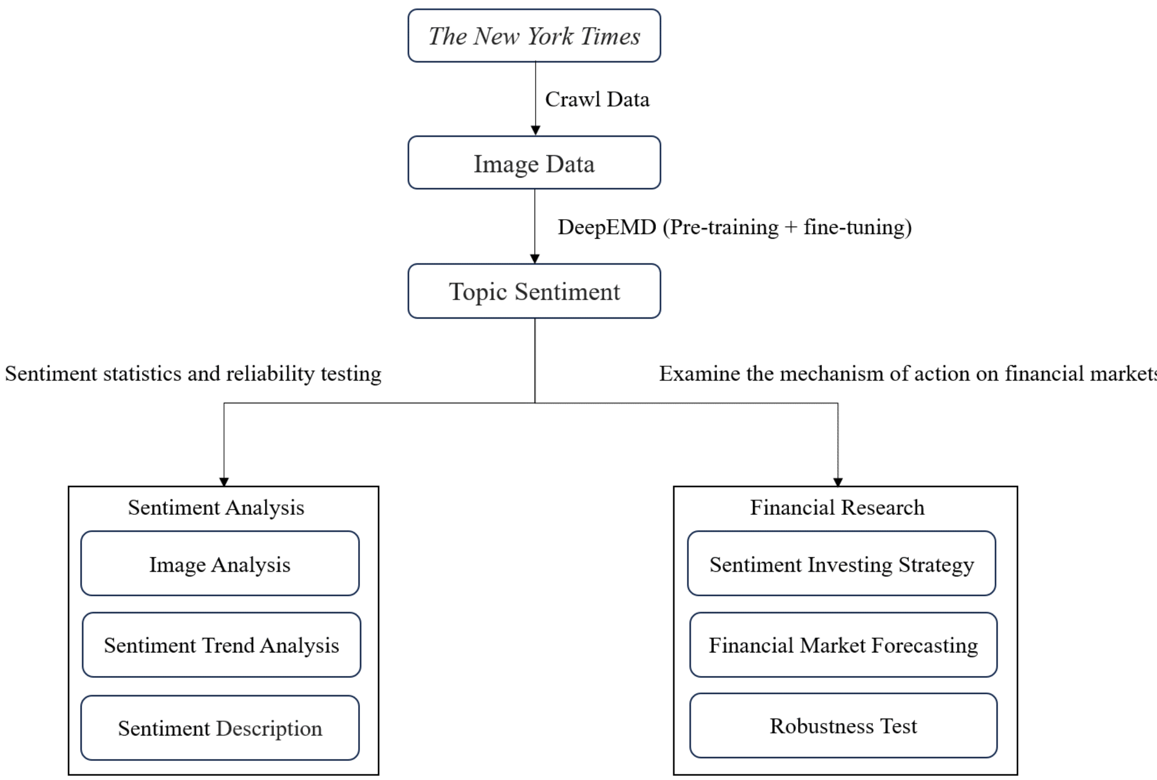

Fig. 1: The Main Technical Research Approach

The structure of this paper is arranged as follows. In Section 2, we review key literature that underpins our research. Section 3 presents the dataset employed in the analysis. Section 4 details the DeepEMD framework alongside the machine learning models utilized for prediction. Section 5 provides a comprehensive analysis of sentiment characteristics across various news topics. Section 6 offers empirical validation of the influence of sentiment on financial markets. Finally, Section 7 summarizes the main findings and closes the discussion.

2 Literature Review

Investor sentiment is no longer a novel topic, as there has been a considerable amount of financial literature devoted to its study. This is attributed to the continuous advancements in natural language processing (NLP) technology, which have made the transformation of textual data into analyzable features much easier (Noh et al., 2015; Phalippou, 2023; Chen et al., 2022a). Consequently, the majority of existing research on financial sentiment is based on text mining. At the corporate level, Chen et al. (2023a) constructed an employee sentiment index by collecting evaluations of their companies from employees of publicly listed firms. This index relies on the difference between the proportions of positive and negative evaluations. Their research found that high levels of employee sentiment predict lower market returns for the company in the future, highlighting the critical role of employee sentiment in the stock market. García-Méndez et al. (2023) developed a Targeted Aspect-Based Emotion Analysis (TABEA) system capable of individually identifying the financial sentiment of different stock market assets within the same tweet. They extracted these sentiments from tweets and accurately identified financial opportunities and precautionary measures. For the selection of news text data, Bai et al. (2022) employed the SeaMNF method, which extracted dynamic sentiment indicators from sparse news headlines. Their text features significantly enhanced the predictive capability for crude oil prices. Li et al. (2019) utilized the Python library, Textblob, to obtain daily sentiment scores from online news, providing real-time high-frequency fundamental data for the crude oil prediction model. Bodilsen and Lunde (2025) found that incorporating macroeconomic news sentiment into regression models significantly enhances the accuracy of long-term volatility forecasts. Doroslovački et al. (2024) proposed a novel market sentiment analysis model that constructs sentiment indicators based on options market data, and found that these indicators exhibit significant predictive power for future spot prices over longer forecast horizons. In the oil market, the study by Cheng et al. (2024) indicated that the interval climate sentiment index constructed using social media can significantly improve the accuracy of interval crude oil price forecasts. Furthermore, an interval trading strategy based on this index not only helps manage market volatility but also contributes to enhancing investment returns.

There is no doubt that the rapid advancement of NLP technology owes much to the birth of Transformers (Vaswani et al., 2017). This innovation has made a large language model (LLM) a reality, and Chen et al. (2022b) is a good example. They utilized BERT and RoBERTa models to extract 27 different sentiment category indicators from The Wall Street Journal (WSJ) and used them to construct various emotion-based investment portfolio strategies. This research has opened up new avenues for effectively utilizing text data to formulate investment strategies. Kriebel and Stitz (2022) also made full use of these models, extracting information related to credit default from user-generated text. The research results demonstrate that these textual features significantly improve the predictive performance of credit default. Similarly, Adämmer et al. (2025) found that news data contains valuable information that many economic indicators fail to capture, which can be used to improve tail risk forecasting. In recent years, the emergence of ChatGPT has made the utilization of textual features more flexible and diversified (Zhao et al., 2023; Kaplan et al., 2020). Chen et al. (2023b) validated whether ChatGPT could identify useful news content related to the stock market and macroeconomics. They conducted their research using headlines and alerts from WSJ, and the results showed that the information extracted by ChatGPT was closely related to macroeconomic conditions and had significant market prediction capabilities.

However, in the realm of image visualization in finance, the number of relevant literature has sharply decreased. Currently, only a few articles have explored investor sentiment in images. Obaid and Pukthuanthong (2022) utilized news images to extract sentiment and correlated them with returns, forming a photo sentiment index, namely Photo Pessimism

(PhotoPes). The research findings indicate that the pessimistic sentiment conveyed in news photos, especially in predicting market returns during periods of elevated panic sentiment, is particularly effective. Furthermore, compared to sentiment from textual sources during the same period, the predictive power of image sentiment is superior. Chiah et al. (2022) reutilized the sentiment data from the PhotoPes index and conducted a study on 37 international stock markets. The research findings demonstrate that PhotoPes can effectively predict subsequent market returns and trading volume, indicating the temporary pricing errors caused by sentiment. While previous studies have extensively explored PhotoPes, this sentiment index is based on overall news content, lacking segmentation across different sections and subjectivity in the manual annotation of the training image set (Obaid and Pukthuanthong, 2022). Therefore, our study aims to provide a more refined and objective analysis of image sentiment, exploring the specific mechanisms of news photos on the financial markets.

For classification problems with insufficient annotations, few-shot learning (FSL) has always been a good choice (Yue et al., 2020; Feng and Duarte, 2019). However, in the financial field, this method is extremely rare. Zhou et al. (2021) proposed a novel semi-supervised FSL model, MetaRisk, for operational risk classification. Their model generalizes from a small number of samples and risk type combinations, improving the accuracy of multi-label risk classification. In the sentiment labeling classification task, Liu et al. (2023) combined LLM with FSL to predict the sentiment classification of three major sentiment corpus speeches, achieving an accuracy of over 90%. This study applies FSL to extract investor sentiment from news photos, possibly being the first article to apply FSL to research in financial stock markets.

3 Data Description

3.1 The New York Times

We use the API provided by The New York Times (NYT) to obtain daily news data from January 2018 to December 2023. Additionally, we utilize web scraping techniques in Python to retrieve all the photos accompanying the news during this period. Considering the content differences across various news sections, many sections lack a significant number of images, making it impossible to form a complete time series. Therefore, we decide to focus on nine representative sections for analysis, including “Arts”, “Books”, “Business Day”, “New York”, “Opinion”, “Sports”, “Technology”, “U.S.”, and “World”. These sections encompass a wide range of information sources, including macroeconomic developments, technological transformations, socio-cultural trends and expressions of public opinion (Obaid and Pukthuanthong, 2022; Barbaglia et al., 2023), thereby forming a news dataset with high sectoral representativeness and structural diversity. Table 1 shows the distribution of the number of photos under each section for different years. We obtain a total of $181,233$ news photos, and the differences in the number of images across different sections in different years are not significant. Among them, the “U.S.” section has the highest number of photos, reaching $41,341$. The next largest is the “World” section, with $29,813$ images. Both of these sections tend to cover topical events with higher importance.

Table 1: Distribution of Photos Across Different News Sections

	2018	2019	2020	2021	2022	2023	Sum
Arts	4249	4122	3656	3244	3168	3265	21704
Books	1620	1681	1563	1565	1539	1512	9480
Business Day	3398	3039	3494	3518	3325	2986	19760
New York	2671	2086	2027	2089	1989	2062	12924
Opinion	5788	5203	4572	3426	3279	3193	25461
Sports	3058	2856	2464	3128	2836	1518	15860
Technology	1034	720	978	787	749	622	4890
U.S.	5948	6805	8923	7443	6045	6177	41341
World	5017	4950	4560	5156	5300	4830	29813
Sum	32783	31462	32237	30356	28230	26165	**181233**

Notes: This table displays the number of photos from different news sections for each year, with the last column and the last row representing the corresponding cumulative values.

To explore the temporal changes of these news photos, we aggregate them on a daily basis. Fig. 2 shows the time series of the number of photos over time for each section.

The gray-shaded area in the graph represents the outbreak period of the global COVID-19 pandemic. We can observe that the number of photos in each section universally experienced an upward trend during this period, especially in sections like “World”, “U.S.”, and “Business Day”, where this increase was more pronounced. This indicates that the pandemic became a focal point of global societal and economic attention (Milani, 2021; Verma and Gustafsson, 2020). In the “Sports” section, we identify two significant increases in the number of photos, further analysis reveals that these two points coincide with major international sports events. Additionally, in the overall distribution of photo numbers (the last row in the graph), we observe some level of cyclicality, particularly with noticeable decreases in photo numbers at the end and beginning of each year. We speculate that this may be due to seasonal effects during the holiday period, when news agencies reduce their workload and reporting frequency, leading to a decrease in the number of photos (Fornari et al., 2002).

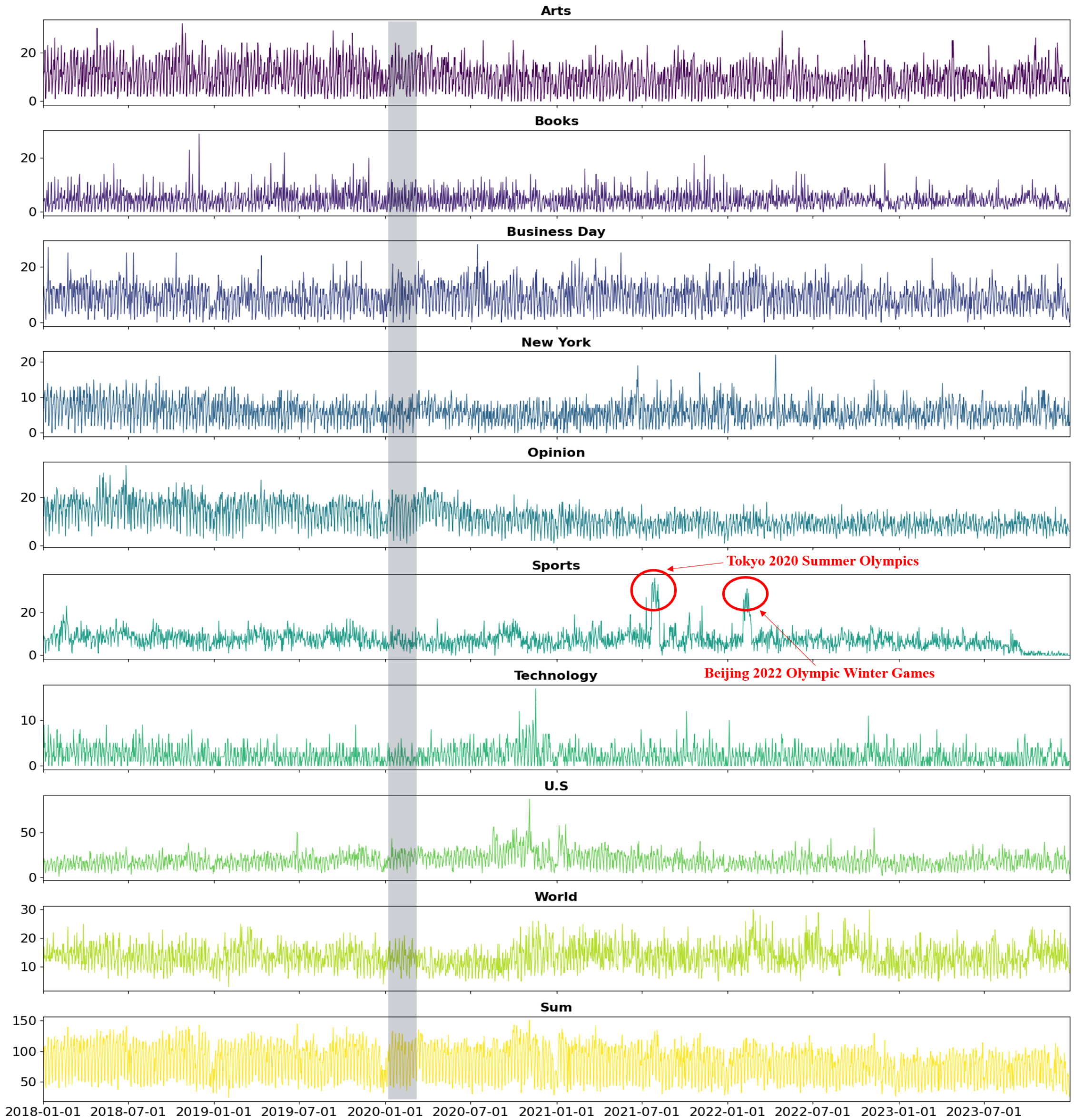

Fig. 2: Daily Time Series of Photos Across Different Sections

3.2 Financial Stock Market Index

To explore the impact of our photo sentiment index on the financial market, based on Degiannakis et al. (2018) and Chang et al. (2015), we select four major stock market indices representing the United States stock market. These indices include the Dow Jones Industrial Average (DJI), the NASDAQ 100 Index (NDX), the Russell 2000 Index (RUT), and the Standard & Poor's 500 Index (SPX). These indices cover stocks of different types and sizes, providing a comprehensive reflection of the overall situation of the US stock market (Petajisto, 2011). Table 2 presents the summary statistics of the prices of these four major indices from January 2018 to December 2023. According to the skewness data, except for the slightly right-skewed RUT index, the skewness of the other indices is close to zero, indicating that their distributions are relatively symmetric. This may be related to the types of stocks selected. The companies covered by the RUT index are typically smaller-cap firms, which often have higher growth potential, potentially resulting in stronger upward trends but also accompanying greater risk and price volatility (Chordia et al., 2011). Additionally, all indices have negative kurtosis values, indicating flatter distributions.

Table 2: Descriptive Statistics of the Stock Indices

	Mean	Median	Std. Dev.	Skewness	Kurtosis
DJI	29893.3769	30015.5100	4167.1383	−0.0687	−1.3925
NDX	11052.4863	11549.6800	3201.2358	0.0187	−1.4375
RUT	1772.4656	1722.3100	283.0561	0.3470	−0.4827
SPX	3587.9087	3677.9500	691.5457	0.0215	−1.4897

Notes: This table displays descriptive statistics of the four most representative stock indices in the United States.

4 Measuring Sentiment in Images

4.1 Few-Shot Learning

In traditional machine learning frameworks, it is often assumed that the training sample size is large enough to cover variations in various categories and features. However, in reality, most data is unlabeled, which means that significant effort is required to annotate large amounts of data by domain experts to support model training (Chen et al., 2013). In such cases, traditional supervised learning becomes time-consuming and labor-intensive, necessitating the use of few-shot learning (FSL) methods for modeling. The goal of FSL is to classify or predict unknown samples using limited data when data availability is insufficient (Wang et al., 2020). This method effectively addresses the problem of insufficient labeled data in supervised learning.

FSL is typically divided into three categories: data enhancement methods, meta-learning methods, and metric learning methods. Data enhancement methods enhance data diversity by applying various transformations and augmentations to the original data (Liu et al., 2020). Meta-learning is a learning approach that adapts to new tasks based on prior knowledge and experience, also known as “learning to learn” (Dong et al., 2025). Metric learning aims to measure the similarity between data points, focusing on learning a feature space where similar samples are closer together and dissimilar samples are farther apart (Jiang et al., 2020). This is also the category of models adopted in this study.

4.2 DeepEMD

Compared to other FSL methods, DeepEMD employs Earth Mover’s Distance (EMD) to measure the similarity between category features and query embeddings (Zhang et al., 2022). EMD measures the distance between two sets of weighted objects. It is based on the fundamental distance between individual objects and the weight of each element, which was originally proposed for image retrieval (Rubner et al., 2000). Given the distances between all units, EMD can find the optimal flow between two distributions, resulting in the minimum cost. Suppose there is a set of suppliers $S = \{s_i|i = 1, 2, ..., n\}$ that need to transport goods to a set of demanders $D = \{d_j|j = 1, 2, ..., m\}$, where s_i represents the i-th supplier unit and d_j represents the j-th demander. The cost of transporting each item from supplier i to demander j is denoted as c_{ij}, and the quantity of units transported is represented by x_{ij}. The objective of this problem is to find the cheapest flow of goods $\tilde{X} = \{\tilde{x}_{ij}|i = 1, ..., n,\ j = 1, ..., m\}$:

$$
\begin{aligned}
&\underset{x_{ij}}{minimize} \sum_{i=1}^{n}\sum_{j=1}^{m} c_{ij}x_{ij} \\
&subject\ to\ x_{ij} \geq,\ i=1,...,n,\ j=1,...,m \\
&\sum_{j=1}^{m} x_{ij} = s_i,\ i=1,...,m \\
&\sum_{i=1}^{n} x_{ij} = d_j,\ j=1,...,n
\end{aligned} \tag{1}
$$

Here, s_i and d_j can also be regarded as weights on each node, used to control the total matching flow generated at each node. Therefore, the objective of EMD is to seek an optimal matching $\tilde{X}$ to minimize the total cost, thus effectively allocating resources between suppliers and demanders. DeepEMD decomposes images into a set of local representations achieved by assigning appropriate weights to the local embeddings of two images. Therefore, all EMD computations are based on feature vectors extracted from network layers (Yuan and Huang, 2020), i.e.,

$$
def\ EMD(weight_1, weight_2, cost\_matrix) \tag{2}
$$

where $weight_1$ represents the weight of input feature S relative to input feature D, $weight_2$ represents the weight of input feature D relative to input feature S, and $cost\_matrix$ denotes the cost of transforming input feature S into input feature D.

Fig. 3 illustrates the overall architecture of DeepEMD. The model consists of two main components. The first part is a fully convolutional neural network (FCN) used for extracting image features. Based on the research by Zhang et al. (2022), we have chosen a 12-layer ResNet (ResNet12) as the backbone model for this part. In this network, we extract feature vectors from the images and compute EMD to measure the distance between embeddings in the embedding space. By weighting the dimensions between the query image and all support set images differently, we can calculate the similarity between each image. The second part is a fully connected layer, which integrates the results obtained from feature extraction and similarity calculation into one-hot encoded labels and outputs them as the final category labels.

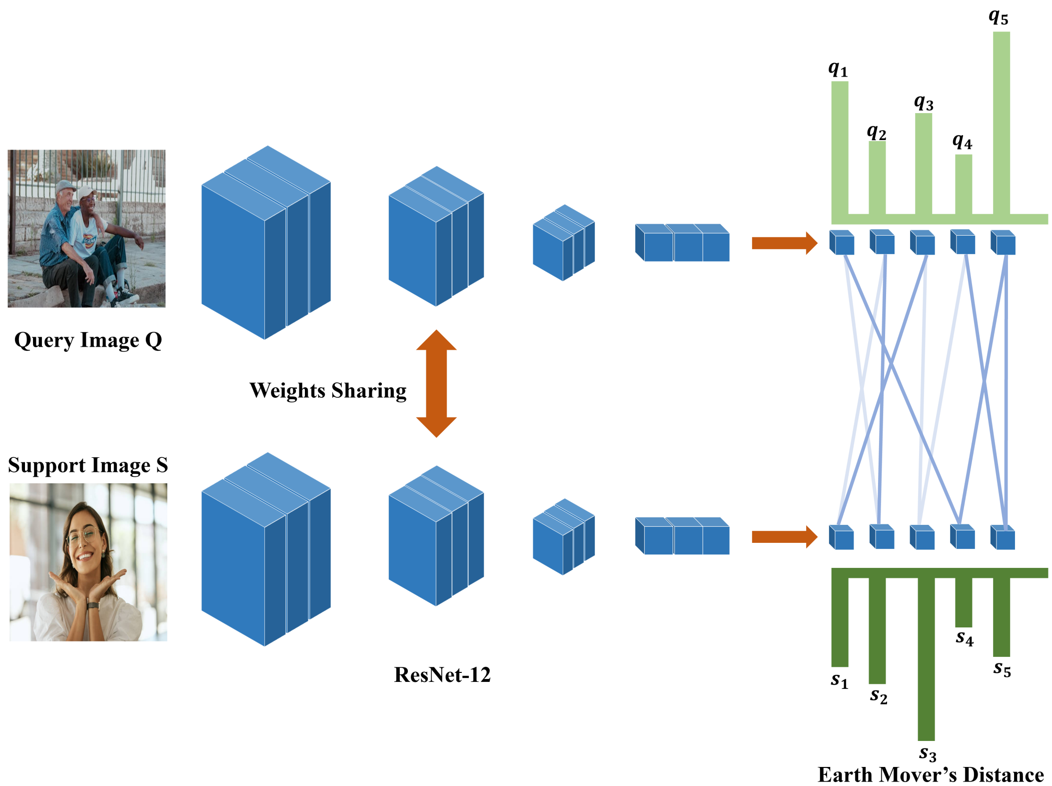

Fig. 3: The Main Architecture of DeepEMD

4.3 Fine-tuning

By conducting prior training on large-scale datasets, pre-trained models can learn feature representations of raw data in advance, enabling parameter sharing and transfer learning. This approach not only accelerates the performance improvement of models with a large number of parameters on various tasks but also reduces the need for a large amount of labeled data. Zhang et al. (2022) provided open-source DeepEMD pre-trained model data in their research. In this study, we select their DeepEMD-FCN pre-trained model trained on the *tieredImageNet* dataset. *TieredImageNet* is a subset of ImageNet, containing 608 categories from 34 superclasses, with a total of 779,165 images (Lee et al., 2019). The pre-trained model achieved an accuracy of 86.03% on the dataset.

To fine-tune the pre-trained model, we use the *artphoto* dataset provided by Machajdik and Hanbury (2010). This dataset leverages theoretical and empirical concepts from psychology and art theory to extract sentiment features from 807 images obtained from an art-sharing website, categorizing them into positive sentiment (amusement, awe, contentment, excitement) and negative sentiment (anger, disgust, fear, sadness). Specifically, we randomly select 5 samples from both positive and negative categories as the support set for each training iteration, with the remaining samples used as the query set, thus constructing a 5-shot FSL model.

5 Validation of the Measurement of Sentiment in Images

In this section, we validate the extracted sentiment from news images. Through the analysis of these thematic sentiments, we verify the transference of sentiment among different news sections.

5.1 Model Evaluation

We additionally utilize the *Emotion*6 dataset provided by Peng et al. (2015) as our out-of-sample validation data. The *Emotion*6 dataset, obtained through a convolutional neural network (CNN), categorizes images into six basic emotions (anger, disgust, fear, joy, sadness, and surprise) according to Ekman (1992). Each category comprises 330 images. Table 3 displays the classification results of fine-tuned DeepEMD on out-of-sample data. Given the higher importance of negative sentiment over positive sentiment (Mikels et al., 2005), we designate negative sentiment samples as positive in the confusion matrix. To assess the performance of the classification model, we select four classification metrics according to Kim et al. (2020) and Jakubik et al. (2023): precision, recall, accuracy, and F1 score:

$$Precision = \frac{TP}{TP + FP} \tag{3}$$

$$Recall = \frac{TP}{TP + FN} \tag{4}$$

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \tag{5}$$

$$F1 = 2 * \frac{Precision * Recall}{Precision + Recall} \tag{6}$$

where TP represents the count of instances where the model correctly predicts positive samples as positive, FP represents the count of instances where the model erroneously predicts negative samples as positive, FN represents the count of instances where the model incorrectly predicts positive samples as negative, and TN represents the count of instances where the model accurately predicts negative samples as negative.

Table 3: Out-of-sample Results of the Fine-tuned DeepEMD

		Actual	
		Positive	Negative
Prediction	Positive	841	153
	Negative	479	507
Precision	84.6076%	Recall	63.7121%
Accuracy	68.0808%	**F1**	**72.6880%**

Notes: This table displays the classification results of the fine-tuned DeepEMD on the validation dataset *Emotion*6. It includes four classification metrics: precision, recall, accuracy and F1 score.

In out-of-sample testing, our fine-tuned model demonstrates good performance. Specifically, we achieve a precision of 84.6%, recall of 63.7%, accuracy of 68.1%, and an F1 score of 72.7%. These results indicate that our model exhibits considerable reliability and accuracy in sentiment extraction. In addition, compared to large-parameter models, DeepEMD has fewer data dependencies and model parameters, thus significantly reducing both training time and operational efficiency of the model (Adadi, 2021). Overall, we believe that this fine-tuned

DeepEMD can serve as an effective tool for extracting sentiment from news images.

5.2 Sentiment Description

Using a fine-tuned DeepEMD model, we successfully conduct sentiment classification on the news images from The New York Times (NYT). The DeepEMD determines the sentiment category of the target image by computing the similarity between images. To further explore the workings of the model, we select several images from each predicted category and conduct a visual analysis of their similarities. Specifically, the network structure of the model divides each image into 25 small blocks and calculates their correlations under different sentiment categories. To present these similarities more intuitively, we add a black filter above the images. We adjust the transparency of the image blocks according to the degree of similarity, aiding in observing which image features had a greater impact on the model's sentiment classification decisions. Representative image samples corresponding to positive and negative emotions are provided in Appendix A.

In images with positive sentiment (refer to Fig. A.1), we can observe that the model focuses on smiling faces. This is consistent with our common knowledge, as smiles are often considered a primary indicator of positive sentiment (Johnson et al., 2010). In contrast, in images with negative sentiment (refer to Fig. A.2), the model pays more attention to elements that may evoke fear, such as flame, weapon, and corpse. Overall, our model demonstrates high credibility in analyzing image similarity. Therefore, we can rely on it to classify sentiment content in news images and construct the sentiment indicators we need. We conduct a daily aggregation of image classification results and calculated daily sentiment scores for different news topics based on various news sources:

$$ImgSent_{d,t} = \frac{1}{n}\sum_{i=0}^{n} Sentiment_{d,t,i} \tag{7}$$

where $ImgSent_{d,t}$ represents the sentiment score of the t-th topic on the d-th day, and $Sentiment_{d,t,i}$ represents the category of the i-th image under the t-th topic on the d-th day. Table 4 displays the descriptive statistics of the daily image sentiment scores under each topic. The average sentiment scores of news images under each topic exceed 0.5, indicating that there are more positive images than negative ones in daily news. Therefore, there is a relatively positive trend in the news content, bringing positive energy to people (Tetlock, 2007), especially in the "World" and "Technology" topics. All sentiment scores show a left-skewed distribution. Following standard practice, we apply standardization to these indicators (Witten and Frank, 2002).

Table 4: Descriptive Statistics of Daily Image Sentiment

TOPIC	Mean	Median	Std. Dev.	Skewness	Kurtosis	Jarque-Bera	Augmented Dicky-Fuller
Arts	0.5957	0.6074	0.1959	-0.5433	1.0999	213.5300***	−46.2448***
Books	0.7030	0.7244	0.2496	-0.8560	0.5589	276.1506***	−44.3030***
Business Day	0.7821	0.7973	0.1526	-0.8949	1.7637	569.6464***	−8.8595***
New York	0.7526	0.7661	0.1915	-0.9782	1.6838	595.4843***	−44.4024***
Opinion	0.6898	0.7005	0.1405	-0.5567	1.0284	20.4909***	−9.5645***
Sports	0.7780	0.7967	0.1760	-1.2111	2.6215	1126.8040***	−7.9962***
Technology	0.7932	0.9433	0.2716	-1.4419	1.4087	748.5096***	−19.2172***
U.S.	0.7469	0.7530	0.1061	-0.3479	0.6115	77.7858***	−17.6862***
World	0.7976	0.8100	0.1051	-0.4304	0.0452	67.6756***	−10.6322***

Notes: This table displays the descriptive statistics of daily news photo sentiment. ***, ** and * denote rejections of the null hypotheses at the 1% , 5% and 10% significance level, respectively.

Next, we inspect the news image sentiments under each topic. Fig. 4 illustrates the daily time series of the sentiment of these topics from 2018 to 2023. To observe the sentiment changes more clearly, we add a 5-day moving average line to each line chart, marked with a red dashed line. Sentiment across different topics may fluctuate due to the influence of relevant events. For instance, at the end of 2023, the emergence of ChatGPT led to a surge in positivity in the "Technology" topic, reflecting a positive attitude (Taecharungroj, 2023). Moreover, sentiment can exhibit a certain degree of propagation across different topics. This indicates that the occurrence of a significant event often triggers sentiment fluctuations in a particular topic first, followed by a ripple effect spreading to other news sections, resulting in a universal sentiment impact. We select two highly representative international events during this period as illustrative cases: the Sino-US trade war in mid-2018 and the COVID-

19 pandemic in early 2020, both of which led to a period of negative trends in news sentiment. The shaded grey areas in the figures indicate the periods when these events began to unfold.

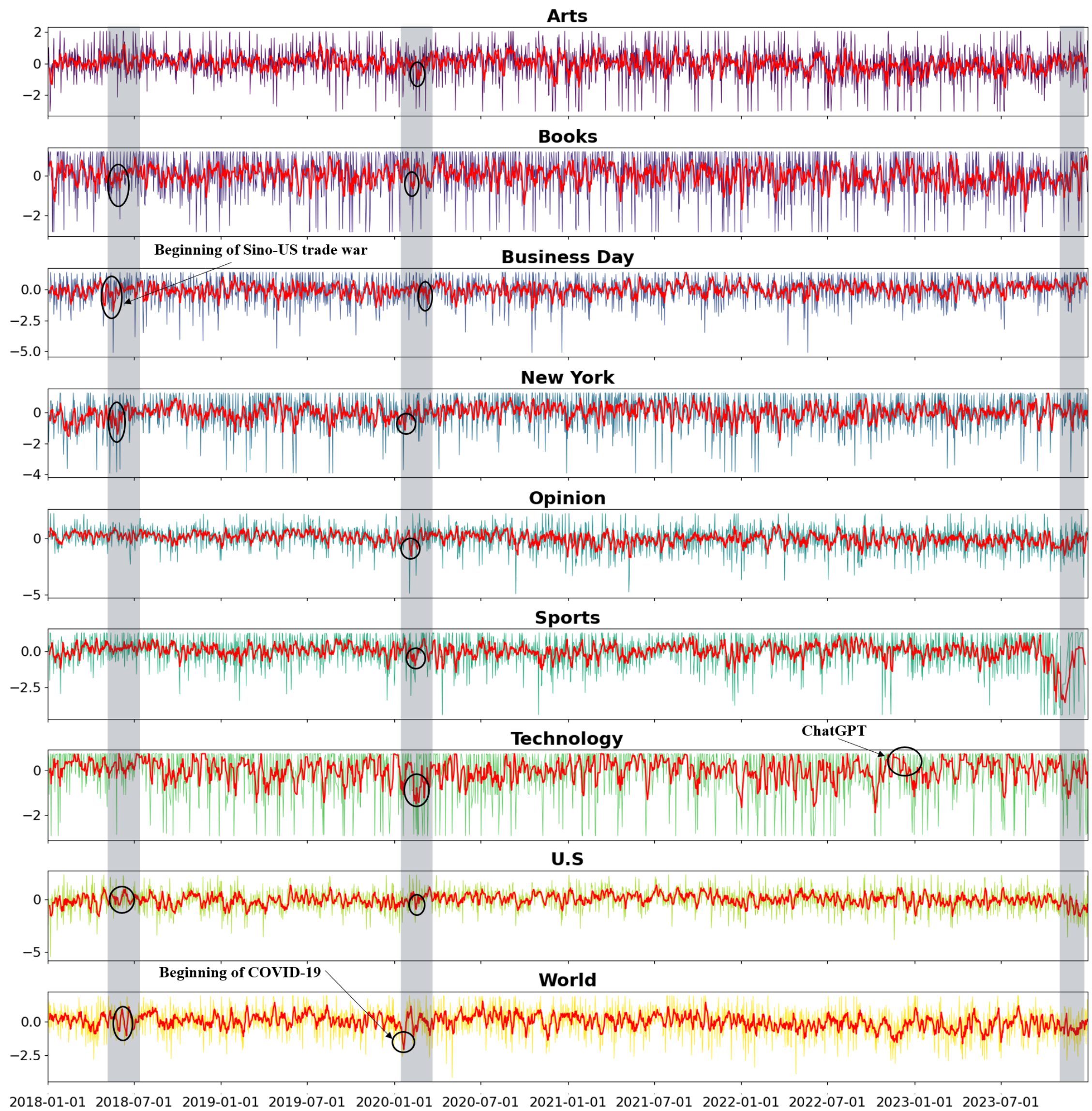

Fig. 4: Daily Trends of News Image Sentiment Scores

At the outset of the Sino-US trade war, media outlets began expressing concerns about its global economic impact (Chen and Wang, 2022), leading to negative sentiment initially surfacing in the "Business Day" topic. Subsequently, these sentiments spread to other topics such as "New York", "U.S.", and "World", which typically cover international political and societal news. Surprisingly, this negative sentiment even extended to the "Books" topic, indicating its influence on the overall operations of the news media. During the outbreak of the COVID-19 pandemic, this pattern became particularly pronounced. As a global infectious disease event, it initially garnered attention in news reports under the "World" topic, leading to negative sentiment being first observed on this topic. Subsequently, these sentiments were swiftly disseminated to almost all other news topics. The pandemic profoundly disrupted normal operations worldwide (Søreide et al., 2020), eliciting negative sentiment among people regarding the economy ("Business Day"), society ("U.S." and "New York"), technology ("Technology"), as well as entertainment ("Arts", "Books" and "Sports"). Thus, as the previous study of Baker et al. (2012) has indicated, the impact of sentiment is not confined solely to specific events but rather exhibits broad contagion. The occurrence of a significant event can influence the sentiment atmosphere across the entire news media. This provides robust initial evidence for exploring the mechanism of sentiment influence on financial markets.

6 Application to Investment Strategy

To validate whether news sentiment has profit potential for investors, we conduct financial market investment simulations based on different topic sentiments in this section. We select the Dow Jones Industrial Average (DJI), representing large enterprises (Donaldson and Kim, 1993), and the Russell 2000 Index (RUT), representing small enterprises (Cremers et al., 2020), to comprehensively reflect the characteristics of the stock market. We formulate investment strategies based on the relationship between the prices of these two indices and different topic sentiments. Specifically, we use the median sentiment value of each topic as a threshold and update this threshold on a rolling basis every six months based on historical data (the initial threshold is calculated using data from January 2018 to June 2018). We adopt two strategies: positive sentiment-oriented and negative sentiment-oriented. The positive sentiment-oriented strategy assumes that positive sentiment predicts future stock price increases. Therefore, when sentiment is above the threshold, we buy before the close of the day and sell before the close of the next day; otherwise, we choose to short-sell. Conversely, the negative sentiment-oriented strategy assumes that stock prices will rise when sentiment is below the threshold, so it takes the opposite trading actions. To evaluate the effectiveness of these strategies, we devise a benchmark strategy, which involves buying the index every day and holding it for one day before selling. Table 5 displays the return results obtained from different sentiment-oriented strategies under these topics. To evaluate the performance of the investment strategy, we use return, Sharpe ratio, and maximum drawdown as key evaluation metrics (Han et al., 2023).

Table 5: The Return Results of Different Investment Strategies

	DJI						RUT					
	Return		Sharpe Ratio		Maximum Drawdown		Return		Sharpe Ratio		Maximum Drawdown	
	Negative	Positive	Negative	Positive	Negative	Positive	Negative	Positive	Negative	Positive	Negative	Positive
Arts	47.44%	-46.36%	0.30	-0.59	**0.36**	0.54	**43.67%**	-53.44%	**0.27**	-0.50	**0.37**	0.63
Books	-39.92%	33.93%	-0.49	0.21	0.50	**0.21**	-15.87%	-21.10%	-0.09	-0.14	0.50	**0.38**
Business Day	-62.49%	**110.86%**	-0.91	**0.62**	0.68	**0.24**	-71.57%	**137.19%**	-0.83	**0.61**	0.79	**0.27**
New York	-2.93%	-18.52%	-0.07	-0.22	**0.32**	0.50	-3.23%	-32.08%	0.00	-0.24	**0.36**	0.54
Opinion	15.12%	-31.07%	0.08	-0.37	**0.29**	0.45	-10.45%	-23.70%	-0.05	-0.16	0.52	0.48
Sports	31.70%	-44.97%	0.20	-0.57	0.50	0.60	-2.54%	-36.55%	0.01	-0.29	0.61	0.62
Technology	-47.20%	**64.09%**	-0.61	**0.39**	0.58	0.38	-53.01%	**62.81%**	-0.49	**0.35**	0.62	**0.36**
U.S.	-40.41%	32.73%	-0.50	0.21	0.54	**0.28**	-18.33%	-16.32%	-0.12	-0.10	0.60	0.48
World	-57.85%	**87.65%**	-0.80	**0.51**	0.72	0.38	-54.13%	**45.86%**	-0.51	**0.27**	0.75	0.56
BaseLine	**55.28%**		**0.35**		**0.37**		**23.37%**		**0.16**		**0.43**	

Notes: This table displays the final return results of sentiment-based strategies across different topics under the Dow Jones Industrial Average (DJI) and the Russell 2000 Index (RUT). This table lists the return rate, Sharpe ratio and maximum drawdown under positive and negative sentiment strategies respectively.

In evaluating the returns of the investment strategy, we observe that the sentiment of different topics exhibits varying emotional orientations. Among them, the "Business Day", "Technology" and "World" topics are particularly instrumental in achieving the highest returns. In these topics, positive sentiment often serves as a reliable indicator of upward market trends. These three themes encompass key political and economic information that significantly influences societal development. News related to these topics can directly reflect market shifts (Calomiris and Mamaysky, 2019). Within these sentiment-driven investment strategies, the "Business Day" theme demonstrate the highest return, achieving 110.86% in the DJI and 137.19% in the RUT, significantly outperforming their respective benchmark returns (55.28% in DJI and 23.37% in RUT). This outcome aligns with our understanding, as the "Business Day" reports are closely tied to economic factors, making market reactions to sentiment more pronounced. Moreover, the positive sentiment-driven strategy associated with this theme also yielded the highest Sharpe ratio and the lowest maximum drawdown, suggesting the relative robustness of the investment strategy.

In the sentiment-driven investment strategies oriented toward negative sentiment, almost no topic manages to outperform the benchmark. However, it is worth noting that in investments involving the DJI index, strategies based on the "Arts" (47.44%) and "Sports" (31.70%) themes both yield positive returns. In the case of investments involving the RUT index, the strategy based on the "Arts" theme even outperforms the baseline, achieving 43.67%. This phenomenon can be partially attributed to investor behavioral psychology. News reports within these themes are often of an entertainment nature, and people typically read them for relaxation purposes. However, when these reports are coupled with negative sentiment, they may disrupt people's expectations of a relaxed atmosphere, leading to a significant psychological gap. As Lee et al. (2002) pointed out, extreme emotional fluctuations are more likely to influence investor behavior. Therefore, this negative sentiment often causes investors to underestimate future market trends, generating aversion and leading to irrational decisions. As mentioned by Barberis and Thaler (2003), this irrationality has a contrarian effect on prices. Choosing to invest at such times can yield greater excess returns.

When comparing the returns of different stock indices, we observe that the cumulative returns of the sentiment strategy for the RUT are generally higher than those for the DJI. This indicates that news sentiment may have greater relevance in the investment of small-cap companies, thus highlighting the applicability of sentiment strategies. Small-cap companies, due to their high volatility, are more susceptible to external events (Levis, 2002). Our news sentiment reflects the dynamic changes in external information.

Additionally, we conduct a detailed analysis of the daily investment returns for each strategy. Fig. 5 illustrates the cumulative return trends of these sentiment-based strategies. The upper graph represents the investment performance for the DJI index, while the lower one for the RUT index. Solid lines indicate strategies based on positive sentiment, while dashed lines represent those based on negative sentiment. Furthermore, we use thick black lines to denote the performance of the benchmark. Comparing the two graphs, we observe that the returns for the RUT index are more volatile than those for the DJI index, which aligns with our previous findings. Among the performance of the two indices, the "Business Day", "Technology" and "World" sentiment based on positive orientation exhibit the best investment returns. It is noteworthy that the investment discrepancies between these strategies began to widen gradually from the beginning of 2020. This timeframe coincides with the global outbreak of the COVID-19 pandemic. This indirectly demonstrates that such news sentiment can help investors mitigate potential risks or seize investment opportunities by quantifying the impact of significant international events.

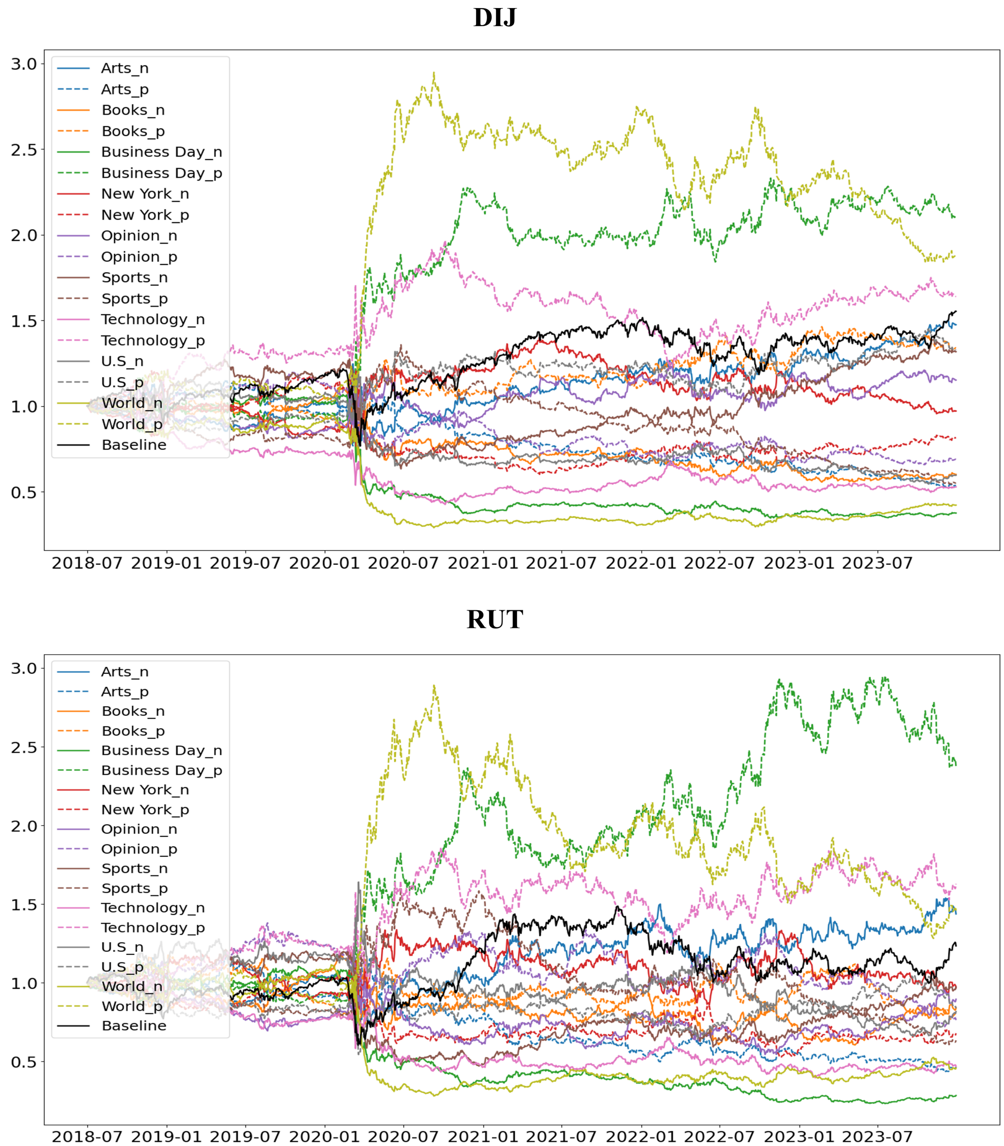

Fig. 5: The Cumulative Return Trends under Different Investment Strategies

7 Application to Financial Market Forecasting

7.1 Forecasting Models

We opt for machine learning models as the primary predictive tool to comprehensively understand the mechanism of impact of these indicators on the financial market. Compared to traditional econometric methods, machine learning models can more flexibly capture nonlinear relationships, thus predicting market price trends more accurately (Alshater et al., 2022). Therefore, based on the research of Lu et al. (2022) and Goodell et al. (2023), we select four typical machine learning models for stock index price prediction: Adaptive Boosting (AdaBoost), Support Vector Regression (SVR), Extremely Randomized Trees (ExtraTrees), and Random Forest (RF). To accurately assess the generalization ability of these nonlinear models, we choose Ordinary Least Squares (OLS) as the baseline model to compare the performance differences between linear and nonlinear models (Bali et al., 2021).

AdaBoost is an ensemble learning method that aims to combine multiple weak learners, which are models with modest performance, to create a stronger learner with improved predictive ability (Freund and Schapire, 1997). In AdaBoost regression tasks, each weak learner adjusts the weights of samples based on prediction errors, giving more attention to those difficult-to-predict samples in the next round of training. By iteratively training multiple weak learners and combining them, AdaBoost regression can better adapt to complex data structures and nonlinear relationships, particularly when there is noise in the dataset or sample distribution is uneven.

Support Vector Regression (SVR), an adaptation of the Support Vector Machine (SVM) for regression tasks, is introduced in (Smola and Schölkopf, 2004). Like SVM, SVR seeks a function that not only fits the training set well but also generalizes effectively to unseen data. The fundamental principle involves constructing a regression function with the widest possible margin, ensuring that deviations from actual targets remain within a defined threshold. The model's support vectors—training instances lying near the margin—are pivotal in shaping the regression function and significantly influence its predictive capability.

Random Forest (RF), introduced by Breiman (2001), is an ensemble-based approach that aggregates the outputs of numerous decision trees to conduct both classification and regression. By combining predictions from multiple trees, RF enhances model robustness and generalization. Moreover, the use of random sampling and randomized feature subsets contributes to minimizing the risk of overfitting.

ExtraTrees is also an ensemble learning algorithm that introduces additional randomness during the construction of decision trees (Geurts et al., 2006). Therefore, compared to RF, ExtraTrees is more randomized. It not only randomly selects the splitting features at each node during tree construction but also randomly selects the splitting thresholds. This randomness increases the diversity of the models and reduces the risk of overfitting. In extreme cases, ExtraTrees may build completely random trees that are independent of the output values of the training samples. Due to its simplicity, efficiency, and insensitivity to hyperparameters, ExtraTrees are widely used in regression problems with high data noise and feature dimensions.

7.2 Forecasting Results

In the above discussion, we find that relying on sentiment from news images could potentially generate additional investment returns. To validate these returns in economic terms, we will utilize this sentiment to predict stock prices. Here, we have introduced two additional prominent US stock indices to enhance the validity of our predictive conclusions. These indices are the NASDAQ 100 Index (NDX), representing many high-growth industries in the US (Pástor and Veronesi, 2006; Chen et al., 2022a), and the Standard & Poor's 500 Index (SPX), representing large-cap companies in the US stock market (Baral and Pokharel, 2017). The price changes of these two indices not only reflect the overall performance of the US stock market but also influence global financial markets (Buncic and Gisler, 2016). Considering that sentiment may exhibit some degree of persistence, we have chosen a lag period of four days, using the sentiment of each topic from the current day and the previous three days to predict the closing price of the indices for the current day (Batchelor et al., 2007). Table 6 displays the in-sample results of linear fitting for the four stock indices.

Consistent with the results of the investment simulation, the in-sample fitting results show significant correlations between the sentiment from the “Arts”, “Business Day”, “New York”, “Opinion”, “U.S.” and “World” topics and stock index prices, validating the predictability of sentiment for the market. Particularly, sentiment from the “Opinion” exhibits the strongest negative linear relationship with market prices, with its four lagged values performing well in the fitting results across all markets. Moreover, the sentiment from the “Business Day”, which best reflects economic conditions, and the sentiment from the “World”, which best reflects international situations, also hold considerable predictive value.

Furthermore, consistent with Liu (2015), we observe a lag effect of these sentiment, indicating that the sentiment of the current day does not directly impact the market prices on the same day but requires some time to influence investor behavior, which is then reflected in market prices. In the fitting results, sentiment from the “Books” and “Technology” confirms this assertion. For instance, in the “Books” sentiment, only the sentiment from two days ago significantly affects the stock index prices, and this effect is manifested in the prices of three indices (DJI, RUT, and SPX). Among the four lagged sentiments provided by the “U.S.” theme, only the sentiment of the current day does not exhibit significance, further confirming the non-immediacy of sentiment.

Table 6: In-sample Results of the Linear Regression for Stock Indices

	DJI	NDX	RUT	SPX
Arts_0	**−0.0130****	−0.0137	−0.0084	**−0.0150****
Arts_1	**−0.0139****	−0.0156	−0.0095	**−0.0167****
Arts_2	**−0.0151*****	−0.0189	**−0.0109***	**−0.0186****
Arts_3	**−0.0133****	−0.0162	−0.0080	**−0.0164****
Books_0	0.0077	0.0152	0.0077	0.0106
Books_1	0.0076	0.0138	0.0094	0.0100
Books_2	**0.0089***	0.0150	**0.0096***	**0.0115***
Books_3	0.0072	0.0134	0.0075	0.0095
Business Day_0	**0.0111****	**0.0227****	0.0094	**0.0160****
Business Day_1	**0.0104****	**0.0222****	0.0086	**0.0154****
Business Day_2	**0.0125****	**0.0255****	**0.0124****	**0.0178****
Business Day_3	**0.0105****	**0.0210***	**0.0116***	**0.0148****
New York_0	0.0034	**0.0270****	0.0021	0.0097
New York_1	0.0011	0.0176	0.0016	0.0053
New York_2	0.0042	**0.0234****	0.0057	0.0088
New York_3	0.0055	**0.0265****	0.0063	0.0112
Opinion_0	**−0.0199*****	**−0.0452*****	**−0.0270*****	**−0.0276*****
Opinion_1	**−0.0195*****	**−0.0431*****	**−0.0282*****	**−0.0264*****
Opinion_2	**−0.0173*****	**−0.0359*****	**−0.0253*****	**−0.0227*****
Opinion_3	**−0.0154*****	**−0.0328*****	**−0.0239*****	**−0.0203*****
Sports_0	0.0021	-0.0027	0.0069	0.0020
Sports_1	0.0002	-0.0050	0.0020	-0.0002
Sports_2	-0.0030	-0.0139	-0.0001	-0.0056
Sports_3	-0.0043	-0.0132	-0.0030	-0.0063
Technology_0	-0.0029	-0.0090	0.0053	-0.0056
Technology_1	-0.0026	-0.0070	0.0037	-0.0045
Technology_2	-0.0005	0.0001	0.0068	-0.0002
Technology_3	0.0023	0.0029	**0.0107****	−0.0019
U.S._0	0.0047	0.0156	0.0071	0.0080
U.S._1	**0.0085***	**0.0238****	**0.0103***	**0.0134****
U.S._2	0.0051	**0.0182***	0.0070	0.0090
U.S._3	**0.0086***	**0.0260****	**0.0101***	**0.0140****
World_0	**−0.0120*****	**−0.0199****	**−0.0094***	**−0.0157*****
World_1	**−0.0101****	−0.0145	**−0.0087***	**−0.0128****
World_2	**−0.0086****	−0.0125	−0.0067	**−0.0111***
World_3	**−0.0105****	**−0.0160***	−0.0074	**−0.0134****
const	**10.2914*****	**9.2514*****	**7.4772*****	**8.1587*****
Adj. R^2	12.00%	11.60%	11.40%	11.60%

Notes: This table displays the in-sample coefficient results for fitting the prices of four stock indices using sentiment lagged by four days. ***, ** and * denote rejections of the null hypotheses at the 1% , 5% and 10% significance level, respectively.

To delve deeper into the predictive effect of sentiment on stock market prices, we conduct out-of-sample prediction tests. Given the presence of numerous non-linear relationships in financial markets, traditional linear models struggle to capture these relationships effectively. Therefore, we opt for machine learning models as predictive models, which have been widely used in current financial literature (e.g. Aziz et al., 2022; Ghoddusi et al., 2019; Christensen et al., 2023). We select four typical machine learning models: Adaptive Boosting (AdaBoost), Support Vector Regression (SVR), Extremely Randomized Trees (ExtraTrees), and Random Forest (RF). Additionally, to further validate the lag effect of sentiment, we utilize the sentiment of all topics within a day to predict the stock market prices for the next five days. Specifically, we use data from 2018 to 2021 as the training set for the models and data from 2022 to 2023 as the out-of-sample testing set. Table 7 presents a detailed comparison of the out-of-sample results, with OLS chosen as the baseline model to evaluate the non-linear capabilities of the machine learning models.

In the comparison of out-of-sample results, we use two primary evaluation metrics: Mean Absolute Error (MAE) and Root Mean Squared Error (RMSE), defined as follows:

$$\text{MAE} = \frac{1}{n}\sum_{t=1}^{n}|p_t - \hat{p}_t| \tag{8}$$

$$\text{RMSE} = \sqrt{\frac{1}{n}\sum_{t=1}^{n}(p_t - \hat{p}_t)^2} \tag{9}$$

Where n is the number of observed data points during the testing period, p_t denotes the actual value of price on day t, and $\hat{p}_t$ is the price forecast obtained using the forecasting model. According to Yan et al. (2020) and Chen et al. (2021), these two metrics are widely used in machine learning prediction problems as they provide a comprehensive assessment of overall error and large errors, aiding in evaluating model performance.

In the table, we compute the loss ratio of each prediction model relative to the baseline model and provide the specific values in parentheses. Firstly, regardless of the context, the machine learning prediction performance surpasses that of the baseline linear model, further confirming that the relationship between sentiment and financial markets exhibits more nonlinear characteristics (Ni et al., 2015). Especially, AdaBoost and SVM perform the best, being the top two nonlinear models in overall performance. Secondly, as the prediction days increase, the performance of each prediction model declines, and the loss correspondingly increases. However, we find that the upward trend in model loss is not consistent. AdaBoost model exhibits the most prominent features. The second-day prediction shows the best performance, with an MSE loss of 0.1317 (0.7692 of the baseline) and an RMSE loss of 0.1401 (0.7519 of the baseline). Except for the fifth-day prediction (H=5), the AdaBoost model's performance is better at other time points than its out-of-sample performance on the first day (H=1). This again confirms our hypothesis that sentiment has a lag effect. This viewpoint aligns with the conclusion of Edmans et al. (2022) that sentiment influences investor behavior, leading to temporary pricing errors. This situation is also evident in ExtraTrees (the third-day prediction is better than the second day) and RF (the fourth-day prediction is better than the third day).

Table 7: Comparative Analysis of Model Predictive Performance

		OLS	AdaBoost	SVM	ExtraTrees	RF
H=1	MAE	0.1703	0.1418	0.1319	0.1564	0.1527
		(1.0000)	(0.8328)	(0.7746)	(0.9184)	(0.8972)
	RMSE	0.1856	0.1513	0.1468	0.1682	0.1663
		(1.0000)	(0.8151)	(0.7909)	(0.9062)	(0.8958)
H=2	MAE	0.1713	**0.1317**	0.1355	0.1574	0.1574
		(1.0000)	**(0.7692)**	(0.7914)	(0.9192)	(0.9192)
	RMSE	0.1864	**0.1401**	0.1507	0.1693	0.1698
		(1.0000)	**(0.7519)**	(0.8084)	(0.9084)	(0.9111)
H=3	MAE	0.1714	0.1345	0.1364	0.1571	0.1582
		(1.0000)	(0.7844)	(0.7959)	(0.9166)	(0.9228)
	RMSE	0.1864	0.1430	0.1518	0.1694	0.1701
		(1.0000)	(0.7670)	(0.8145)	(0.9089)	(0.9126)
H=4	MAE	0.1724	0.1336	0.1385	0.1601	0.1579
		(1.0000)	(0.7751)	(0.8035)	(0.9287)	(0.9158)
	RMSE	0.1873	0.1425	0.1536	0.1713	0.1702
		(1.0000)	(0.7607)	(0.8204)	(0.9149)	(0.9089)
H=5	MAE	0.1729	0.1542	0.1387	0.1623	0.1590
		(1.0000)	(0.8922)	(0.8025)	(0.9387)	(0.9199)
	RMSE	0.1879	0.1633	0.1540	0.1733	0.1713
		(1.0000)	(0.8689)	(0.8192)	(0.9223)	(0.9112)

Notes: This table displays the predictive performance of the forecasting model at different lag days. The values within parentheses indicate the ratios of these non-linear models compared to the baseline.

7.3 Robustness Test

Finally, in this section, we conduct a robustness test of sentiment indicators to validate the precise impact of news image sentiment on the financial markets. To achieve this goal, we adopt an approach similar to the research methodology of Chiah et al. (2022), which involves aggregating sentiment indicators across all topics to calculate a composite indicator reflecting overall news media sentiment. The formula for calculating this composite indicator is as follows:

$$OImgSent_d = \frac{1}{n}\sum_{i=0}^{n} Sentiment_{d,i} \tag{10}$$

Where $OImgSent_d$ represents the overall news image sentiment for the d-th day, and $Sentiment_{d,i}$ represents the category of the i-th image on the d-th day. Fig. 6 depicts the daily trend of overall news image sentiment. We also plot a 5-day moving average line in red. Similar to the volatility observed in financial market prices, news sentiment exhibits significant fluctuations. At the beginning of 2020, overall news sentiment remained highly negative, showing a noticeable decline.

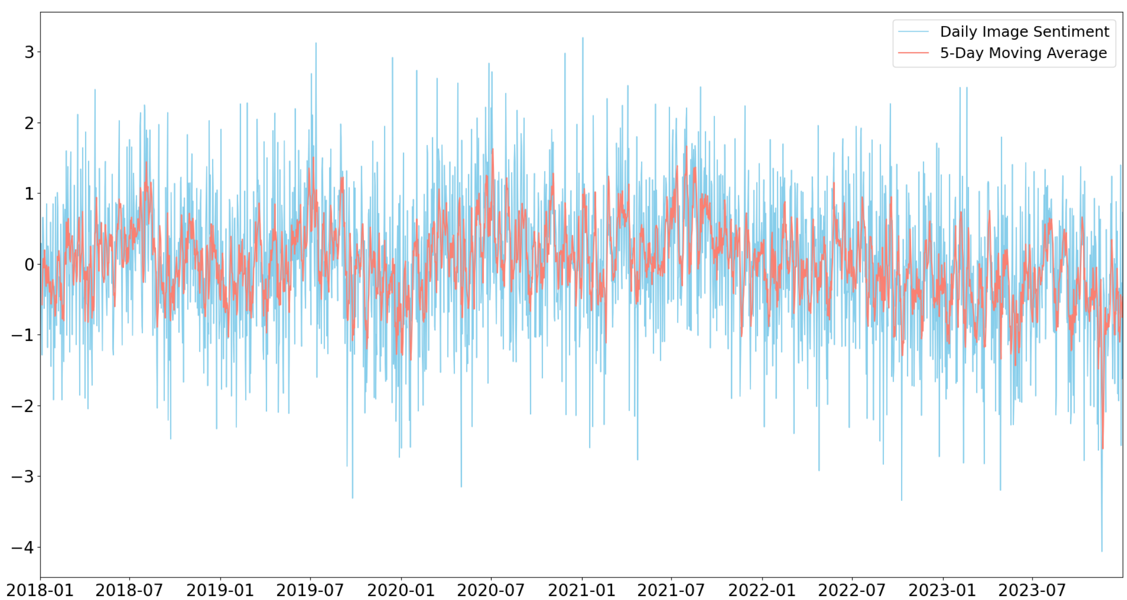

Fig. 6: Daily Trends of Overall News Photos Sentiment

Continuing with the approach used in the previous investment simulations, we devise positive sentiment-oriented and negative sentiment-oriented strategies accordingly. We still target DJI and RUT as the market indices of interest. Table 8 presents the final performance of our investment strategies. Positive sentiment aligns with the direction of future market uptrends.

Under the positive sentiment-oriented strategy, our investments yielded a total return of 4.24% on the DJI index and 11.95% on the RUT index, both exceeding the benchmarks for these indices (4.17% and 3.79%). The performance of the RUT index significantly outperformed that of the DJI index, suggesting that news sentiment might be more suitable for investments in small-cap companies in growth phases. These smaller companies are more susceptible to public opinion and policy influences, thus exhibiting greater asset volatility (Michaelas et al., 1999; Crouzet and Mehrotra, 2020). Therefore, leveraging news sentiment as an investment guide may offer better operational and investment guidance for these companies and their investors.

Table 8: The Return Results of Overall Sentiment Investment Strategies

	DJI		RUT	
	Negative	Positive	Negative	Positive
Sentiment Strategy	-4.30	**4.24%**	-11.36	**11.95%**
BaseLine	**4.17%**		**3.79%**	

Notes: This table displays the final return results of sentiment-based strategies under the Dow Jones Industrial Average (DJI) and the Russell 2000 Index (RUT). The second and fourth columns represent investment strategies oriented toward negative sentiment, while the third and fifth columns represent strategies oriented toward positive sentiment.

We opt for the sentiment-oriented strategy with higher returns and depict the investment performance of these strategies on the stock indices in Fig. 7. We observe significant discrepancies between the benchmark strategy and our sentiment-oriented strategy during the COVID-19 pandemic. This global crisis led to an overall contraction trend in financial markets, triggering a decline in stock index prices (Berkman and Malloch, 2023). However, our strategy successfully mitigated the risk of negative returns based on the performance of news sentiment, and opted for short selling, which resulted in substantial profit increases. This also demonstrates that news image sentiment helps predict the impact of major events on financial markets to some extent, aiding us in risk avoidance. However, in future investment trends, DJI never further expanded its returns, eventually slightly outperforming the benchmark. On the other hand, our strategy consistently widened the gap with the benchmark on RUT, reaffirming the superior reference value of sentiment in investments in small companies.

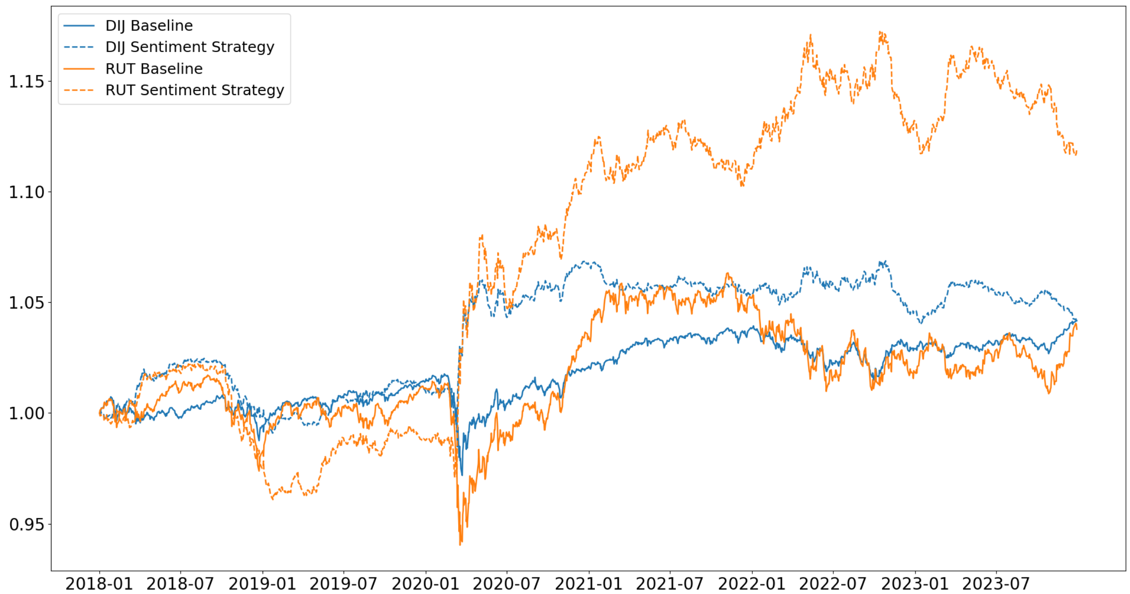

Fig. 7: The Cumulative Return Trend of Overall News Sentiment Strategy

8 Conclusion

This study employs a few-shot learning (FSL) technique to extract investor sentiment from news images across various prominent sections of The New York Times (NYT). By segmenting the sentiment of news topics, we expand the application of few-shot learning in the financial domain and validated the role of different news topics in the variability of investor decision-making. Firstly, our results demonstrate the intricate interconnections among different news topics. We have studied specific event time points (such as the Sino-US trade war and the COVID-19 pandemic) and found that the sentiment triggered by such events often resonates from news topics in “World” or “Business Day” and has chain effects on almost all other news topics.

In terms of investment strategies, we validate the effectiveness of sentiment associated with various topics. Specifically, we find that serious news related to economics and politics often contains substantial information that can positively reflect financial market dynamics. Therefore, the sentiment from such news has a positive guiding effect on economic development. In contrast, some entertainment-oriented negative news can influence investor decision-making, leading to incorrect assessments of market prices. The emotional differences brought about by these different news topics help us gain a deeper understanding of how investor behavior impacts market fluctuations.

Furthermore, our empirical results validate that these sentiments captured from news images can predict changes in stock index prices. Financial markets entail numerous linear and nonlinear relationships. By delving into the nonlinear relationship between sentiment and the US stock market, we can better understand the mechanisms of market operation. Therefore, machine learning models with stronger capabilities to capture nonlinear relationships have immense potential in the financial domain.

Additionally, we observe a lag effect of sentiment. This implies that the news sentiment does not immediately impact the financial markets. Instead, there is a noticeable delay period during which it gradually influences investor behavior, ultimately reflecting changes in market prices, leading to temporary pricing errors. Therefore, paying attention to past sentiment trends can better help us understand the dynamics of the current financial market.

Finally, news sentiment holds greater relevance for investing in small-cap companies. Small-cap firms are more susceptible to external events, thus exhibiting higher volatility. Our investment strategy simulations validate the stronger predictive ability of these sentiments for small-cap stock markets. Particularly during extreme events, our sentiment indicators demonstrate stronger predictive capabilities. Our sentiment-based strategies successfully mitigated the market downturn risk caused by the COVID-19 pandemic. Hence, leveraging news sentiment aids investors in better risk management.

In today's era of advancing technology, media coverage is increasingly expansive. Transforming this reliable information into profit opportunities has always been the relentless pursuit of investors and scholars. This study is just a preliminary exploration of this field. Future research could extend the application of sentiment indicators to more types of markets to explore their applicability in different domains. Additionally, extracting more valuable information from these non-traditional data sources is also an important direction for future research. Financial technology has already opened new doors for us, and we eagerly await further developments in the future.

Acknowledgements

We acknowledge the financial support from by the National Natural Science Foundation of China (72403258, 72171234), the Natural Science Foundation of Hunan Province (2022JJ40647), Fundamental Research Funds for the Central Universities (2722024EJ011), and the Innovation and Talent Base for Digital Technology and Finance (B21038).

References

Adadi, A. (2021). A survey on data-efficient algorithms in big data era. *Journal of Big Data*, 8(1):24.

Adämmer, P., Prüser, J., and Schüssler, R. A. (2025). Forecasting macroeconomic tail risk in real time: Do textual data add value? *International Journal of Forecasting*, 41(1):307–320.

Alshater, M. M., Kampouris, I., Marashdeh, H., Atayah, O. F., and Banna, H. (2022). Early warning system to predict energy prices: The role of artificial intelligence and machine learning. *Annals of Operations Research*, 1–37.

Aziz, S., Dowling, M., Hammami, H., and Piepenbrink, A. (2022). Machine learning in finance: A topic modeling approach. *European Financial Management*, 28(3):744–770.

Bai, Y., Li, X., Yu, H., and Jia, S. (2022). Crude oil price forecasting incorporating news text. *International Journal of Forecasting*, 38(1):367–383.

Baker, M., Wurgler, J., and Yuan, Y. (2012). Global, local, and contagious investor sentiment. *Journal of Financial Economics*, 104(2):272–287.

Bali, T. G., Goyal, A., Huang, D., Jiang, F., and Wen, Q. (2021). Different strokes: Return predictability across stocks and bonds with machine learning and big data. *Swiss Finance Institute, Research Paper Series*, (20-110).

Baral, N. and Pokharel, M. P. (2017). How sustainability is reflected in the s&p 500 companies' strategic documents. *Organization and Environment*, 30(2):122–141.

Barbaglia, L., Consoli, S., and Manzan, S. (2023). Forecasting with economic news. *Journal of Business and Economic Statistics*, 41(3):708–719.

Barberis, N. and Thaler, R. (2003). A survey of behavioral finance. *Handbook of the Economics of Finance*, 1:1053–1128.

Batchelor, R., Alizadeh, A., and Visvikis, I. (2007). Forecasting spot and forward prices in the international freight market. *International Journal of Forecasting*, 23(1):101–114.

Benson, R. (2009). What makes news more multiperspectival? A field analysis. *Poetics*, 37(5-6):402–418.

Berkman, H. and Malloch, H. (2023). Stock valuation during the covid-19 pandemic: An explanation using option-based discount rates. *Journal of Banking and Finance*, 147:106386.

Bhandari, U., Chang, K., and Neben, T. (2019). Understanding the impact of perceived visual aesthetics on user evaluations: An emotional perspective. *Information and Management*, 56(1):85–93.

Bianchi, J., Liu, C., and Mendoza, E. G. (2016). Fundamentals news, global liquidity and macroprudential policy. *Journal of International Economics*, 99:S2–S15.

Bodilsen, S. T. and Lunde, A. (2025). Exploiting news analytics for volatility forecasting. *Journal of Applied Econometrics*, 40(1):18–36.

Borth, D., Ji, R., Chen, T., Breuel, T., and Chang, S.-F. (2013). Large-scale visual sentiment ontology and detectors using adjective noun pairs. In *Proceedings of the 21st ACM International Conference on Multimedia*, pages 223–232.

Breiman, L. (2001). Random forests. *Machine Learning*, 45:5–32.

Buncic, D. and Gisler, K. I. (2016). Global equity market volatility spillovers: A broader role for the united states. *International Journal of Forecasting*, 32(4):1317–1339.

Calomiris, C. W. and Mamaysky, H. (2019). How news and its context drive risk and returns around the world. *Journal of Financial Economics*, 133(2):299–336.

Chang, Y.-C., Hong, H., and Liskovich, I. (2015). Regression discontinuity and the price effects of stock market indexing. *Review of Financial Studies*, 28(1):212–246.

Chen, C. Y.-H., Fengler, M. R., Härdle, W. K., and Liu, Y. (2022a). Media-expressed tone, option characteristics, and stock return predictability. *Journal of Economic Dynamics and Control*, 134:104290.

Chen, F. and Wang, G. (2022). A war or merely friction? Examining news reports on the current sino-us trade dispute in the new york times and china daily. *Critical Discourse Studies*, 19(1):1–18.

Chen, J., Tang, G., Yao, J., and Zhou, G. (2023a). Employee sentiment and stock returns. *Journal of Economic Dynamics and Control*, 149:104636.

Chen, J., Tang, G., Zhou, G., and Zhu, W. (2023b). Chatgpt, stock market predictability and links to the macroeconomy. *Available at SSRN 4660148*.

Chen, W., Zhang, H., Mehlawat, M. K., and Jia, L. (2021). Mean–variance portfolio optimization using machine learning-based stock price prediction. *Applied Soft Computing*, 100:106943.

Chen, Y., Carroll, R. J., Hinz, E. R. M., Shah, A., Eyler, A. E., Denny, J. C., and Xu, H. (2013). Applying active learning to high-throughput phenotyping algorithms for electronic health records data. *Journal of the American Medical Informatics Association*, 20(e2):e253–e259.

Chen, Y., Kelly, B. T., and Xiu, D. (2022b). Expected returns and large language models. *Available at SSRN 4416687*.

Cheng, Z., Li, M., Sun, Y., Hong, Y., and Wang, S. (2024). Climate change and crude oil prices: An interval forecast model with interval-valued textual data. *Energy Economics*, 134:107612.

Chiah, M., Hu, X., and Zhong, A. (2022). Photo sentiment and stock returns around the world. *Finance Research Letters*, 46:102417.

Chordia, T., Roll, R., and Subrahmanyam, A. (2011). Recent trends in trading activity and market quality. *Journal of Financial Economics*, 101(2):243–263.

Christensen, K., Siggaard, M., and Veliyev, B. (2023). A machine learning approach to volatility forecasting. *Journal of Financial Econometrics*, 21(5):1680–1727.

Cremers, M., Pareek, A., and Sautner, Z. (2020). Short-term investors, long-term investments, and firm value: Evidence from russell 2000 index inclusions. *Management Science*, 66(10):4535–4551.

Crouzet, N. and Mehrotra, N. R. (2020). Small and large firms over the business cycle. *American Economic Review*, 110(11):3549–3601.

Dan-Glauser, E. S. and Scherer, K. R. (2011). The Geneva affective picture database (GAPED): A new 730-picture database focusing on valence and normative significance. *Behavior Research Methods*, 43:468–477.

Degiannakis, S., Filis, G., and Hassani, H. (2018). Forecasting global stock market implied volatility indices. *Journal of Empirical Finance*, 46:111–129.

Donaldson, R. G. and Kim, H. Y. (1993). Price barriers in the dow jones industrial average. *Journal of Financial and Quantitative Analysis*, 28(3):313–330.

Dong, J., Wang, Y., Xie, X., Lai, J., and Ong, Y.-S. (2025). Generalizable and discriminative representations for adversarially robust few-shot learning. *IEEE Transactions on Neural Networks and Learning Systems*, 36(3):5480–5493.

Doroslovački, K., Gradojevic, N., and Tarnaud, A. C. (2024). A novel market sentiment analysis model for forecasting stock and cryptocurrency returns. *IEEE Transactions on Systems, Man, and Cybernetics: Systems.*

Dumitrescu, E., Hué, S., Hurlin, C., and Tokpavi, S. (2022). Machine learning for credit scoring: Improving logistic regression with non-linear decision-tree effects. *European Journal of Operational Research*, 297(3):1178–1192.

Edmans, A., Fernandez-Perez, A., Garel, A., and Indriawan, I. (2022). Music sentiment and stock returns around the world. *Journal of Financial Economics*, 145(2):234–254.

Ekman, P. (1992). An argument for basic emotions. *Cognition and Emotion*, 6(3-4):169–200.

Feng, S. and Duarte, M. F. (2019). Few-shot learning-based human activity recognition. *Expert Systems with Applications*, 138:112782.

Fenton-O'Creevy, M., Soane, E., Nicholson, N., and Willman, P. (2011). Thinking, feeling and deciding: The influence of emotions on the decision making and performance of traders. *Journal of Organizational Behavior*, 32(8):1044–1061.

Fornari, F., Monticelli, C., Pericoli, M., and Tivegna, M. (2002). The impact of news on the exchange rate of the lira and long-term interest rates. *Economic Modelling*, 19(4):611–639.

Freund, Y. and Schapire, R. E. (1997). A decision-theoretic generalization of on-line learning and an application to boosting. *Journal of Computer and System Sciences*, 55(1):119–139.

García-Méndez, S., de Arriba-Pérez, F., Barros-Vila, A., and González-Castaño, F. J. (2023). Targeted aspect-based emotion analysis to detect opportunities and precaution in financial twitter messages. *Expert Systems with Applications*, 218:119611.

Geurts, P., Ernst, D., and Wehenkel, L. (2006). Extremely randomized trees. *Machine Learning*, 63:3–42.

Ghoddusi, H., Creamer, G. G., and Rafizadeh, N. (2019). Machine learning in energy economics and finance: A review. *Energy Economics*, 81:709–727.

Goodell, J. W., Jabeur, S. B., Saâdaoui, F., and Nasir, M. A. (2023). Explainable artificial intelligence modeling to forecast bitcoin prices. *International Review of Financial Analysis*, 88:102702.

Griffith, J., Najand, M., and Shen, J. (2020). Emotions in the stock market. *Journal of Behavioral Finance*, 21(1):42–56.

Han, C., He, Z., and Toh, A. J. W. (2023). Pairs trading via unsupervised learning. *European Journal of Operational Research*, 307(2):929–947.

Invernizzi, A. C., Bellucci, M., Acuti, D., and Manetti, G. (2022). Form and substance: Visual content in csr reports and investors' perceptions. *Psychology and Marketing*, 39(5):974–989.

Jakubik, J., Nazemi, A., Geyer-Schulz, A., and Fabozzi, F. J. (2023). Incorporating financial news for forecasting bitcoin prices based on long short-term memory networks. *Quantitative Finance*, 23(2):335–349.

Jiang, F., Lee, J., Martin, X., and Zhou, G. (2019). Manager sentiment and stock returns. *Journal of Financial Economics*, 132(1):126–149.

Jiang, W., Huang, K., Geng, J., and Deng, X. (2020). Multi-scale metric learning for few-shot learning. *IEEE Transactions on Circuits and Systems for Video Technology*, 31(3):1091–1102.

Johnson, K. J., Waugh, C. E., and Fredrickson, B. L. (2010). Smile to see the forest: Facially expressed positive emotions broaden cognition. *Cognition and Emotion*, 24(2):299–321.

Kaplan, J., McCandlish, S., Henighan, T., Brown, T. B., Chess, B., Child, R., Gray, S., Radford, A., Wu, J., and Amodei, D. (2020). Scaling laws for neural language models. *arXiv preprint arXiv:2001.08361*.

Kearney, C. and Liu, S. (2014). Textual sentiment in finance: A survey of methods and models. *International Review of Financial Analysis*, 33:171–185.

Khandani, A. E., Kim, A. J., and Lo, A. W. (2010). Consumer credit-risk models via machine-learning algorithms. *Journal of Banking and Finance*, 34(11):2767–2787.

Kim, A., Yang, Y., Lessmann, S., Ma, T., Sung, M.-C., and Johnson, J. E. (2020). Can deep learning predict risky retail investors? A case study in financial risk behavior forecasting. *European Journal of Operational Research*, 283(1):217–234.

Kozyreva, A., Lewandowsky, S., and Hertwig, R. (2020). Citizens versus the internet: Confronting digital challenges with cognitive tools. *Psychological Science in the Public Interest*, 21(3):103–156.

Kriebel, J. and Stitz, L. (2022). Credit default prediction from user-generated text in peer-to-peer lending using deep learning. *European Journal of Operational Research*, 302(1):309–323.

Kross, E. and Ayduk, O. (2017). Self-distancing: Theory, research, and current directions. In *Advances in Experimental Social Psychology*, volume 55, pages 81–136.

Lee, K., Maji, S., Ravichandran, A., and Soatto, S. (2019). Meta-learning with differentiable convex optimization. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 10657–10665.

Lee, W. Y., Jiang, C. X., and Indro, D. C. (2002). Stock market volatility, excess returns, and the role of investor sentiment. *Journal of Banking and Finance*, 26(12):2277–2299.

Levis, M. (2002). The record on small companies: A review of the evidence. *Journal of Asset Management*, 2:368–397.

Li, X., Shang, W., and Wang, S. (2019). Text-based crude oil price forecasting: A deep learning approach. *International Journal of Forecasting*, 35(4):1548–1560.

Liu, S. (2015). Investor sentiment and stock market liquidity. *Journal of Behavioral Finance*, 16(1):51–67.

Liu, T., Chen, Z., Yang, Y., Wu, Z., and Li, H. (2020). Lane detection in low-light conditions using an efficient data enhancement: Light conditions style transfer. In *2020 IEEE Intelligent Vehicles Symposium (IV)*, pages 1394–1399.

Liu, Z.-T., Wu, B.-H., Han, M.-T., Cao, W.-H., and Wu, M. (2023). Speech emotion recognition based on meta-transfer learning with domain adaption. *Applied Soft Computing*, 147:110766.

Lu, J., Gong, P., Ye, J., Zhang, J., and Zhang, C. (2023). A survey on machine learning from few samples. *Pattern Recognition*, 139:109480.

Lu, X., Ma, F., Xu, J., and Zhang, Z. (2022). Oil futures volatility predictability: New evidence based on machine learning models. *International Review of Financial Analysis*, 83:102299.

Machajdik, J. and Hanbury, A. (2010). Affective image classification using features inspired by psychology and art theory. In *Proceedings of the 18th ACM International Conference on Multimedia*, pages 83–92.

Marchewka, A., Żurawski, Ł., Jednoróg, K., and Grabowska, A. (2014). The nencki affective picture system (naps): Introduction to a novel, standardized, wide-range, high-quality, realistic picture database. *Behavior Sesearch Methods*, 46:596–610.

Michaelas, N., Chittenden, F., and Poutziouris, P. (1999). Financial policy and capital structure choice in uk smes: Empirical evidence from company panel data. *Small Business Economics*, 12:113–130.

Mikels, J. A., Fredrickson, B. L., Larkin, G. R., Lindberg, C. M., Maglio, S. J., and Reuter-Lorenz, P. A. (2005). Emotional category data on images from the international affective picture system. *Behavior Sesearch Methods*, 37:626–630.

Milani, F. (2021). Covid-19 outbreak, social response, and early economic effects: A global var analysis of cross-country interdependencies. *Journal of Population Economics*, 34(1):223–252.

Moen, E., Bannon, D., Kudo, T., Graf, W., Covert, M., and Van Valen, D. (2019). Deep learning for cellular image analysis. *Nature Methods*, 16(12):1233–1246.

Ni, Z.-X., Wang, D.-Z., and Xue, W.-J. (2015). Investor sentiment and its nonlinear effect on stock returns—New evidence from the chinese stock market based on panel quantile regression model. *Economic Modelling*, 50:266–274.

Noh, H., Jo, Y., and Lee, S. (2015). Keyword selection and processing strategy for applying text mining to patent analysis. *Expert Systems with Applications*, 42(9):4348–4360.

Obaid, K. and Pukthuanthong, K. (2022). A picture is worth a thousand words: Measuring investor sentiment by combining machine learning and photos from news. *Journal of Financial Economics*, 144(1):273–297.

Pástor, L. and Veronesi, P. (2006). Was there a nasdaq bubble in the late 1990s? *Journal of Financial Economics*, 81(1):61–100.

Peng, K.-C., Chen, T., Sadovnik, A., and Gallagher, A. C. (2015). A mixed bag of emotions: Model, predict, and transfer emotion distributions. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 860–868.

Petajisto, A. (2011). The index premium and its hidden cost for index funds. *Journal of Empirical Finance*, 18(2):271–288.

Phalippou, L. (2023). Thematic investing with big data: The case of private equity. *Financial Analysts Journal*, 79(4):30–40.

Pham, M. T. (2007). Emotion and rationality: A critical review and interpretation of empirical evidence. *Review of General Psychology*, 11(2):155–178.

Quillian, M. R. (1967). Word concepts: A theory and simulation of some basic semantic capabilities. *Behavioral Science*, 12(5):410–430.

Rubner, Y., Tomasi, C., and Guibas, L. J. (2000). The earth mover's distance as a metric for image retrieval. *International Journal of Computer Vision*, 40:99–121.

Shynkevich, A. (2012). Performance of technical analysis in growth and small cap segments of the us equity market. *Journal of Banking and Finance*, 36(1):193–208.

Smola, A. J. and Schölkopf, B. (2004). A tutorial on support vector regression. *Statistics and Computing*, 14:199–222.

Song, Y., Wang, T., Cai, P., Mondal, S. K., and Sahoo, J. P. (2023). A comprehensive survey of few-shot learning: Evolution, applications, challenges, and opportunities. *ACM Computing Surveys*, 55(13s):1–40.

Søreide, K., Hallet, J., Matthews, J. B., Schnitzbauer, A. A., Line, P. D., Lai, P. B., Otero, J., Callegaro, D., Warner, S. G., Baxter, N. N., et al. (2020). Immediate and long-term impact of the covid-19 pandemic on delivery of surgical services. *Journal of British Surgery*, 107(10):1250–1261.

Stambaugh, R. F. and Yuan, Y. (2017). Mispricing factors. *Review of Financial Studies*, 30(4):1270–1315.

Taecharungroj, V. (2023). "What can chatgpt do?" Analyzing early reactions to the innovative ai chatbot on twitter. *Big Data and Cognitive Computing*, 7(1):35.

Taffler, R. (2018). Emotional finance: Investment and the unconscious. *European Journal of Finance*, 24(7-8):630–653.

Tetlock, P. C. (2007). Giving content to investor sentiment: The role of media in the stock market. *Journal of Finance*, 62(3):1139–1168.

Van Belle, D. A. (2003). Bureaucratic responsiveness to the news media: Comparing the influence of the new york times and network television news coverage on us foreign aid allocations. *Political Communication*, 20(3):263–285.

Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, Ł., and Polosukhin, I. (2017). Attention is all you need. *Advances in Neural Information Processing Systems*, 30.

Verma, S. and Gustafsson, A. (2020). Investigating the emerging covid-19 research trends in the field of business and management: A bibliometric analysis approach. *Journal of Business Research*, 118:253–261.

Wang, Y., Yao, Q., Kwok, J. T., and Ni, L. M. (2020). Generalizing from a few examples: A survey on few-shot learning. *ACM Computing Surveys (CSUR)*, 53(3):1–34.

Witten, I. H. and Frank, E. (2002). Data mining: Practical machine learning tools and techniques with java implementations. *Acm Sigmod Record*, 31(1):76–77.

Yan, B., Aasma, M., et al. (2020). A novel deep learning framework: Prediction and analysis of financial time series using ceemd and lstm. *Expert Systems with Applications*, 159:113609.

Yuan, Z. and Huang, W. (2020). Multi-attention deepemd for few-shot learning in remote sensing. In *2020 IEEE 9th Joint International Information Technology and Artificial Intelligence Conference (ITAIC)*, volume 9, pages 1097–1102.

Yue, Z., Zhang, H., Sun, Q., and Hua, X.-S. (2020). Interventional few-shot learning. *Advances in Neural Information Processing Systems*, 33:2734–2746.

Zhang, C., Cai, Y., Lin, G., and Shen, C. (2022). Deepemd: Differentiable earth mover's distance for few-shot learning. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 45(5):5632–5648.

Zhao, S., Yao, X., Yang, J., Jia, G., Ding, G., Chua, T.-S., Schuller, B. W., and Keutzer, K.

(2021). Affective image content analysis: Two decades review and new perspectives. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 44(10):6729–6751.

Zhao, W. X., Zhou, K., Li, J., Tang, T., Wang, X., Hou, Y., Min, Y., Zhang, B., Zhang, J., Dong, Z., et al. (2023). A survey of large language models. *arXiv preprint arXiv:2303.18223*.

Zheng, L., Miao, M., and Gan, Y. (2020). Perceived control buffers the effects of the covid-19 pandemic on general health and life satisfaction: The mediating role of psychological distance. *Applied Psychology: Health and Well-Being*, 12(4):1095–1114.

Zhou, F., Qi, X., Xiao, C., and Wang, J. (2021). Metarisk: Semi-supervised few-shot operational risk classification in banking industry. *Information Sciences*, 552:1–16.

A Additional Figures

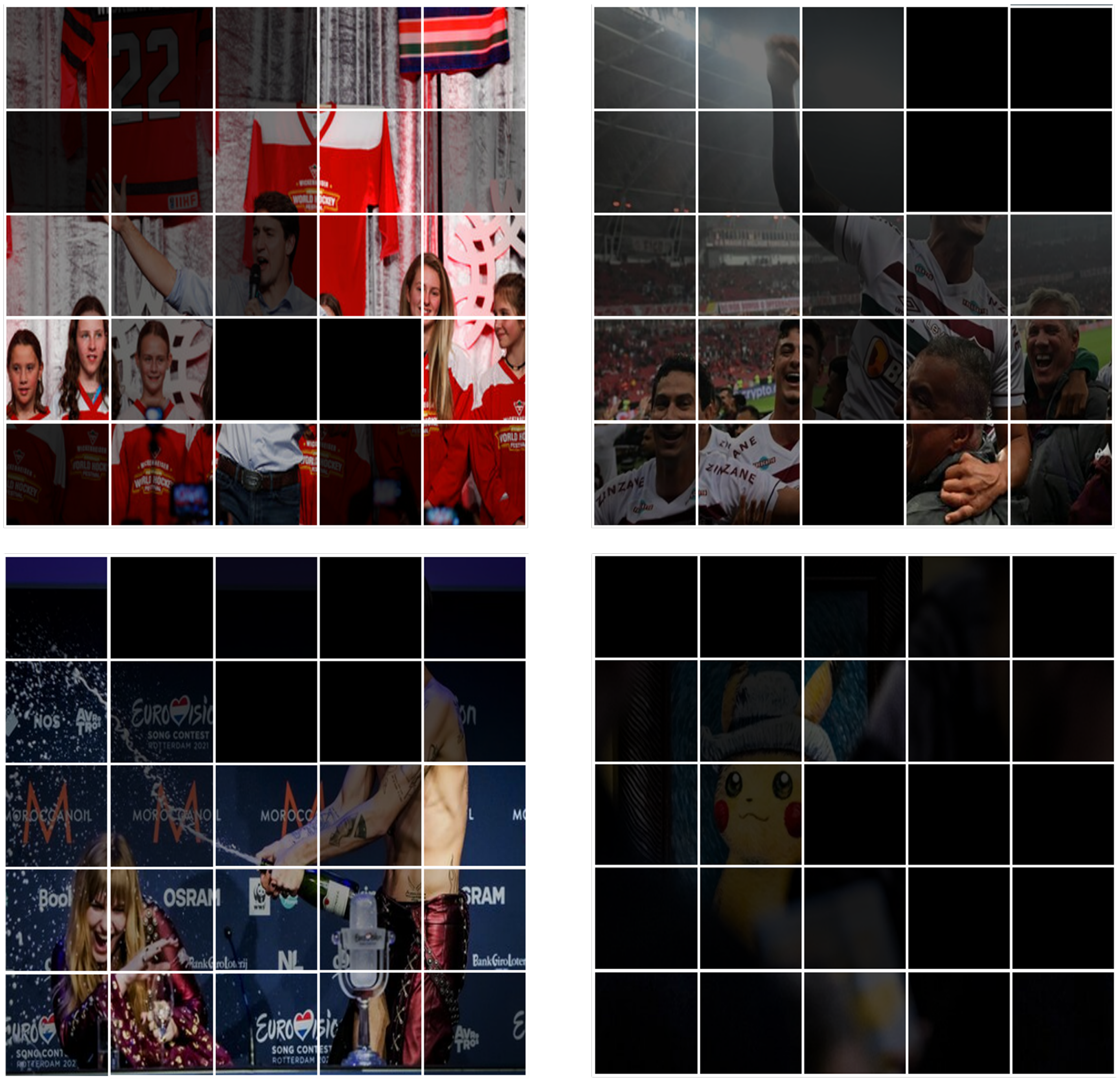

Fig. A.1: Image Similarity for Positive Sentiment

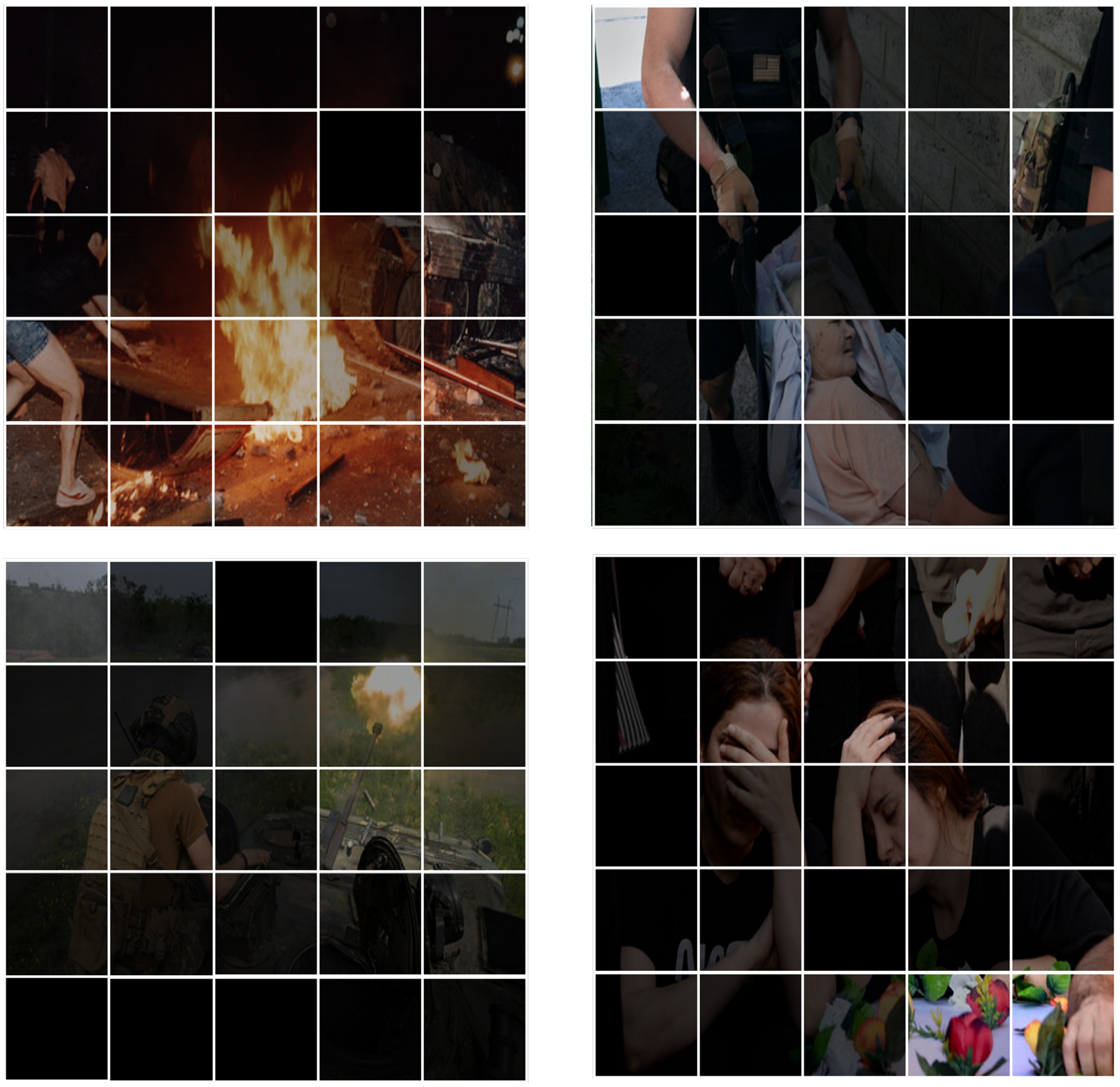

Fig. A.2: Image Similarity for Negative Sentiment