Being an Emotionally Unaffected Investor: Evidence from Bitcoin

Journal:	*Transactions on Engineering Management*
Manuscript ID	Draft
Manuscript Type:	Special Issue: The Role of Emergent Distributed Ledger Technologies (DLT) including Blockchain and Cryptocurrencies in Achieving Business Resilience and Productivity
Subject Category:	Digital Technologies and Analytics
Keywords:	Large Language Models, BERT, Machine Learning, Textual Analysis, Investor Behavior, Bitcoin, Return Prediction

Being an Emotionally Unaffected Investor: Evidence from Bitcoin

Abstract

As one of the most prominent cryptocurrencies, Bitcoin has been at the forefront of a major revolution in the financial and technological sectors. This study utilizes data from social media to extract the emotional tendencies of investors in the Bitcoin market and analyze differences in investor behavior under various emotional features. We demonstrate that when investors generally exhibit reluctance (such as Sadness and Fear) to buy Bitcoin, it is precisely the opportune moment to invest and achieve returns higher than expected. Conversely, when the emotional tone of investors becomes positive (such as Joy and Love), indicating a tendency to invest, we opt to avoid investments. Our research has also revealed that these emotional cues can assist in better predicting returns in the bitcoin market. Analyzing market emotions contributes to a deeper understanding of market fluctuations and investor behavior. Our findings and insights are expected to help investors recognize the role of subjective emotions in the market and provide them with prudent investment advice: avoid relying excessively on the emotions of others, as this can reduce your investment losses.

Keywords: Large Language Models, BERT, Machine Learning, Textual Analysis, Investor Behavior, Bitcoin, Return Prediction

1 Introduction

In the past, investment was a relatively isolated decision-making process. However, today, thanks to the continuous advancement of internet technology and the deepening development of the digital economy, economic information is becoming increasingly timely and transparent. Online social platforms have become one of the crucial channels for information dissemination, and the emergence of forums has provided investors with an unprecedented opportunity (Wu et al., 2014). More and more investors are flocking to these specific forums, expressing their emotions, and sharing viewpoints with people worldwide.

The current internet is inundated with a wealth of personal emotions. Behavioral finance has demonstrated that investor sentiment can significantly impact stock returns, contradicting the efficient market hypothesis (EMH) (Nassirtoussi et al., 2014). Behavioral finance regards investors, especially retail investors, as ordinary individuals who make irrational decisions based on their own emotions and interests. It posits that emotions can drive investor behavior, thereby influencing their investment decisions in financial markets, providing a better explanation for market anomalies that classical finance cannot explain (Kaplanski and Levy, 2010; Zhou, 2018). Consequently, people have become increasingly eager to explore the potential impact of subjective sentiment in the text on the economy (e.g. Zhou et al., 2023; Jha et al., 2021).

With the continuous advancement of blockchain technology, bitcoin has become a typical representative of digital currency. As a decentralized cryptocurrency, bitcoin's uniqueness lies in its independence from traditional financial institutions and lack of government credit support, resulting in its high price volatility (Sun et al., 2020). From 2013 to 2017, bitcoin's price surged by nearly 2800%, reaching a peak of nearly $\$20,000$, only to plummet to $\$3,000$ in the subsequent months. Despite the recurring bubble behaviors, both academia and investors continue to exhibit a strong interest in bitcoin (Xiong et al., 2020). *Bitcointalk.org* is the first online forum dedicated to discussing and purchasing bitcoin and is considered the world's largest bitcoin forum (Toyoda et al., 2019). This forum was established in 2010 and has seen its user base expand from the initial 3,127 users in 2023 to 3,568,694 users. Hence, we have selected this forum as our source of investor sentiment, measured various emotional states in user posts, and studied the impact of these emotional states on bitcoin market returns.

Specifically, this study utilizes text data posted by users on the Bitcointalk.org platform as its research subject and conducts sentiment analysis. With the launch of ChatGPT, large language models (LLMs) have become a focal point of extensive research and discussion across various industries (Saggu and Ante, 2023). These LLMs are designed to handle a multitude of text-related tasks, boasting a vast number of parameters in the billions, which endows them with a high-level comprehension of semantic nuances in text (Leippold, 2023). Given this capability, in this study, we have selected the BERT model developed by Devlin et al. (2018) and the RoBERTa model developed by Liu et al. (2019) as tools for sentiment analysis. Unlike in the past, we have moved beyond simple binary sentiment classification, which categorizes sentiments as positive or negative (e.g. Jegadeesh and Wu, 2013; Airani and Karande, 2022). Instead, we employ a more nuanced six-class sentiment analysis approach. Ekman (1992) has categorized human emotions into six fundamental classes (namely "Love", "Joy", "Surprise", "Anger", "Sadness" and "Fear") and asserted that these basic emotions effectively explain the emotional features of the vast majority of individuals. This multi-emotion classification approach allows us to gain a more accurate insight into investor sentiment fluctuations, ultimately yielding more precise results.

Through LLMs, we obtain distribution data for the six daily emotional states of users. To delve deeper into the influence of different emotions on investment behavior, we devise corresponding investment strategies for bitcoin under various emotional features. Under each emotional feature, we buy bitcoin for half of the days with the highest emotional scores and hold it for one day. We then calculate the returns obtained under these emotional features separately. Surprisingly, the research results far exceeded our expectations. We discover that when investors, particularly retail investors, are generally in a pessimistic mood, especially feeling sadness or fear, our investments exhibit unexpectedly high returns. However, when investors are generally in a positive mood, especially experiencing joy or love, our investments perform below expectations. This counterintuitive correlation piqued our interest, as it seemed to validate Warren Buffett's investment principle: "Be fearful when others are greedy, and be greedy when others are fearful" (Buffett, 2008).

With an in-depth analysis of the textual content under different emotional features, we find that investors with negative emotional features often display a reluctance to engage in investments. At this point, they tend to underestimate the price of bitcoin, providing an opportunity for rational investors to make purchases. Rational investors are less susceptible to short-term emotional disturbances and can acquire assets at lower prices when market sentiment is low, eventually profiting when market sentiment recovers (Shleifer and Summers, 1990). Our research results also aptly demonstrate that such moments often present an excellent opportunity for investment, indirectly confirming that investment behavior is not merely a matter of herd mentality (Dang and Lin, 2016).

Considering the distinct contrast in investment strategies under different emotional features, we hypothesize that these emotional features also have explanatory power for bitcoin returns. To validate this, we apply this emotional distribution data to a predictive model for the bitcoin market. Moreover, to more accurately measure the explanatory value of emotions in market trends, we utilize machine learning methods capable of capturing the nonlinear impacts of these emotional indicators (Gu et al., 2020; Christensen et al., 2022). Empirical results clearly indicate that emotional factors can enhance our ability to forecast the performance of the bitcoin market. This discovery underscores the importance of sentiment analysis in the field of finance. The main technical research approach of this study is illustrated in Fig. 1.

[Fig. 1 about here.]

The main contributions of this paper are as follows. In contrast to previous literature, we conduct a more in-depth exploration of investor emotional characteristics. Unlike the traditional binary emotional classification found in prior literature, we segment investor emotions into six categories, which helps to provide a more comprehensive understanding of the roles investors play in financial markets. This research fills gaps in the fields of behavioral finance and social psychology. Additionally, it also offers a fresh perspective on the field of cryptocurrency research. We shift our research focus to the non-traditional bitcoin market. This is an intriguing endeavor because our results unveiled the distinctions between the bitcoin market and traditional markets like stocks and futures. The prominence of emotional factors in the bitcoin market underscores the limitations of the rational investor assumption in traditional financial frameworks. Moreover, our research provides a clear lesson for investors: do not easily succumb to the influence of others' emotions.

The remainder of this paper is organized as follows. Section 2 outlines important literature relevant to our work. Section 3 presents the data used in this study. Section 4 introduces the emotion analysis and prediction methods. Section 5 extensively explores investment differences under different emotions. We forecast bitcoin returns in Section 6, and Section 7 concludes the paper.

2 Literature Review

At present, the application of text data in financial market analysis is not a new development. An increasing number of studies are focusing on extracting sentiment information from text. Bai et al. (2022) introduced a dynamic sentiment indicator specifically designed for short news and sparse news headlines. This indicator takes into account the cumulative and diminishing effects of sentiment, enhancing the accuracy of price predictions in the crude oil market. On the other hand, Li et al. (2020) extracted investor sentiment from stock comments on a Chinese financial website, categorizing it as positive, neutral, or negative sentiment. The results showed that sentiment extracted from text to some extent contributed to predicting the performance of the Chinese stock market. Cao et al. (2023) went a step further in dissecting emotional dimensions by dividing overall investor sentiment into multiple dimensions, revealing the asymmetric predictability of these multidimensional sentiments in the futures market. This coincides with our approach.

In the macroeconomic domain, Consoli et al. (2021) extracted sentiment states from macroeconomic news in Italy and Spain, demonstrating that this indicator better captured human perception and emotional influences related to investor behavior and decision-making. Their research found that extracted negative sentiment aided in predicting government yield bond spreads during unstable periods. Ren and Wu (2018) approached the study from an investor perspective, employing a sentence-based sentiment analysis method to explore herd behavior among investors. They constructed sentiment indicators to identify group behavior in blue-chip stocks, and the results indicated that herd effects were more pronounced when overall investor sentiment was extremely positive or negative. This underscores the fact that markets are not entirely rational, aligning with our conclusions.

The topic of cryptocurrency has gained rapid popularity in recent years. Given their characteristic of operating without third-party involvement (such as governments and financial institutions), numerous studies have shown that cryptocurrencies are sensitive to investor sentiment. Corbet et al. (2020) analyzed sentiment polarity related to cryptocurrencies, including bitcoin, on social media during the COVID-19 pandemic. They found that negative sentiments significantly impacted the returns of cryptocurrencies. Concurrently, Chen et al. (2022) employed a difference-in-difference (DID) framework to explore the influence of COVID-19-induced uncertainty on bitcoin price differences. They discovered that media hype around the COVID-19 topic and public panic significantly affected bitcoin price differences. These differences were attributed to investors considering bitcoin as an alternative hedge investment in times of uncertainty. These findings provide theoretical support for our research.

Furthermore, Guégan and Renault (2021) explored the role of investor sentiment in the bitcoin market using social media information from the StockTwits platform. Their research, based on high-frequency sentiment, found that during the bitcoin bubble period (from August 2017 to April 2018), irrational sentiments had a statistically significant positive impact on its price returns. Park and Seo (2023) conducted investment simulations using sentiment. They proposed an action recommendation model based on sentiment analysis of text information on Twitter. By adjusting future investment behavior based on the sentiment distribution of tweets, they achieved returns significantly better than the baseline. Poongodi et al. (2021) selected specialized bitcoin forums, specifically *Bitcointalk*, as the source of their literature, which is also the data source for our study. They used a Latent Dirichlet Allocation (LDA) model to obtain daily topic distributions and validated the feasibility of using social media data to predict global cryptocurrency prices. However, despite these studies, bitcoin remains a relatively immature topic compared to other financial markets. These observations suggest that there is a metaphorical dimension to emotions in the context of bitcoin. Prior literature only scratches the surface, and we aim to focus on specific emotions to describe the unique relationship between investors and the bitcoin market.

3 Data Description

By using Python's web scraping techniques, we collect daily user posts and their replies from the *Bitcointalk.org*, spanning from January 2018 to June 2023. *Bitcointalk* is currently the most populous online discussion forum about bitcoin. To facilitate subsequent text sentiment analysis, we removed some web links from the text and discarded data containing fewer than 5 words. Fig. 2 presents the statistical information regarding user posts based on date, month, and the days within each month. In total, we gathered 21,018 posts, comprising 53,093 messages from 897,356 users. Regarding the monthly frequency (as shown in the second figure), we can observe that the messages posted by users exhibit a clear underlying cyclical pattern. There is a significant increase every six months, particularly between April-May and October-November.

[Fig. 2 about here.]

Subsequently, we collect the corresponding bitcoin price data for the mentioned time period, as shown in Fig. 3, depicting the bitcoin price trends from January 2018 to June 2023. From the figure, it is evident that the price of bitcoin exhibits extreme volatility, especially between October 2020 and April 2021, where the price surged from $\$10,000$ to over $\$60,000$, followed by a rapid decline in May 2021. This volatility reflects bitcoin's status as a financial asset free from government and traditional financial institution regulations and its susceptibility to more irrational factors compared to traditional markets Güler (2023). This is a topic that this paper seeks to explore in depth.

[Fig. 3 about here.]

4 Methodology

4.1 Sentiment Analysis Methods

4.1.1 Large Language Model

With Google's introduction of the Attention mechanism and the Transformer architecture based on this mechanism (Vaswani et al., 2017), the performance of Natural Language Processing (NLP) tasks has significantly improved, officially leading the trend in NLP. The Transformer represents an innovative deep neural network encoder with its built-in self-attention mechanism, giving the model the ability to establish connections between various positions in a sequence. It can handle long-range dependencies, surpassing the limitations of traditional Recurrent Neural Networks (RNN) and Convolutional Neural Networks (CNN) in handling long-distance dependencies. Furthermore, the Transformer overcomes constraints related to parallelization and memory, thereby enhancing the model's scalability. Large language models (LLMs), specifically Transformer models trained on massive text data, possess massive parameter sizes (Stevenson et al., 2021). Thus, today, these LLMs can capture rich language knowledge and representations, excelling in numerous text tasks, including but not limited to text classification, machine translation, question-answering systems, and more.

In this study, we leverage LLMs based on bidirectional Transformer encoders, namely BERT and RoBERTa, to perform sentiment classification tasks. These models have been widely proven to have strong text processing capabilities (Tian et al., 2022).

4.1.2 Pre-training and Fine-tuning

BERT is a language model created by Devlin et al. (2018), and it is built upon the concept of transfer learning. Therefore, this model undergoes a pretraining process. Pretraining involves initially learning general knowledge on a large-scale dataset and then fine-tuning specific tasks to meet particular requirements (Wang et al., 2018). This approach accelerates the model's performance on various tasks while reducing the need for extensive labeled data. For BERT, its pre-training was conducted on a corpus of over 3 billion words from the BookCorpus and English Wikipedia. The model comprises 300,000 billion parameters to better understand syntax and semantics. During the pre-training process, BERT involves two parallel unsupervised tasks: the Masked Language Model (MLM) and Next Sentence Prediction (NSP). To facilitate the training of the BERT model, the data encoding process is as follows:

a) Use the BERT tokenizer to tokenize the comment text.

b) Attach a special token [CLS] at the beginning to indicate the classification task and attach another special token [SEP] at the end.

c) Choose a sequence length (in this paper, we select 512), and truncate or pad all comments to a single fixed length.

d) Create an attention mask to distinguish between real tokens and padding tokens.

For better comprehension, the following example illustrates these steps. Due to space constraints, we have chosen a sequence length of 64.

- Input Text: "there we go again first russia and now the uk there be no point in having cryptocurrency if it run by central authority that defeats the whole purpose and the appeal of bitcoin"

- Tokenized: ['there', 'we', 'go', 'again', 'first', 'russia', 'and', 'now', 'the', 'uk', 'there', 'be', 'no', 'point', 'in', 'having', 'crypt', '##oc', '##ur', '##ren', '##cy', 'if', 'it', 'run', 'by', 'central', 'authority', 'that', 'defeats', 'the', 'whole', 'purpose', 'and', 'the', 'appeal', 'of', 'bit', '##co', '##in']

- Special Tokens: ['[CLS]', there', 'we', 'go', 'again', 'first', 'russia', 'and', 'now', 'the', 'uk', 'there', 'be', 'no', 'point', 'in', 'having', 'crypt', '##oc', '##ur', '##ren', '##cy', 'if', 'it', 'run', 'by', 'central', 'authority', 'that', 'defeats', 'the', 'whole', 'purpose', 'and', 'the', 'appeal', 'of', 'bit', '##co', '##in', '[SEP]']

- Token IDs: [101, 2045, 2057, 2175, 2153, 2034, 3607, 1998, 2085, 1996, 2866, 2045, 2022, 2053, 2391, 1999, 2383, 19888, 10085, 3126, 7389, 5666, 2065, 2009, 2448, 2011, 2430, 3691, 2008, 14222, 1996, 2878, 3800, 1998, 1996, 5574, 1997, 2978, 3597, 2378, 102, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0]

For vocabulary words not found in the tokenizer's vocabulary, we break them down into existing words and connect them with a pound sign "##" symbol. As shown in the example above, the word "cryptocurrency" is broken down into five tokens: "crypt", "##oc", "##ur", "##ren", "##cy". These token IDs, with the attached attention mask, can then be used to fine-tune the BERT base model for classification tasks.

RoBERTa is an improved version of BERT proposed by Liu et al. (2019). Compared to BERT, RoBERTa employs a larger training dataset, larger batch sizes, and longer pre-training steps in its training methodology. It further enhances the quality of pre-training compared to BERT by omitting the NSP task and adjusting hyperparameters. Additionally, RoBERTa utilizes a dynamic masking strategy, generating a new masking pattern for each input sequence during model training. This enhancement allows RoBERTa to support more training steps and larger datasets.

Fine-tuning is the process of training a pre-trained model, which has undergone extensive training, for specific task adaptations. This phase typically requires relatively small amounts of data, as the model already possesses a broad understanding of language. According to Alain and Bengio (2016), we use feature extraction to directly utilize the pre-trained model's parameters for generating text features relevant to downstream tasks. Specifically, we extract features directly from the pre-trained model for the target text, without updating any parameters generated during training. This simplified approach significantly reduces the computational burden, and previous research has indicated that pre-trained model parameters are sufficient for understanding text semantics (Jiang et al., 2022).

Based on the study of Kriebel and Stitz (2022), we utilize the pre-trained bert-base-uncased for BERT and pre-trained roberta-base for RoBERTa. Saravia et al. (2018) applied graph-based algorithms for six-class (Joy, Sadness, Fear, Surprise, Love, and Anger) sentiment analysis of user text collected from Twitter. They annotated tweets and constructed the *Emotion* dataset. As a result, we obtained their provided dataset through *HuggingFace* and built a sentiment classification model based on LLMs. To achieve the final six-class output, we feed the text features extracted from the pre-trained model into a fully connected layer to generate a result with six elements. To ensure that the prediction results form a probability distribution with a sum of one, we also add a Softmax function to the output layer, which is mathematically expressed as follows:

$$Softmax(z_i) = \frac{exp(z_i)}{\sum_{j=1}^{6} exp(z_j)} \tag{1}$$

Where z is the emotion score, z_i and z_j represent one of the emotions. Finally, we apply the emotion classifier based on LLMs to the bitcoin forum texts we've collected and calculate daily emotion distributions using the following formula:

$$Emotion_{d,t} = \frac{1}{n} \sum_{j=0}^{n} Emotion_{t,d,i} \tag{2}$$

Where $e_{t,d}$ represents the probability of the t-th emotion on the d-th day, $e_{d,t,i}$ represents the probability of the t-th emotion of the i-th text message on the d-th day, and n represents the number of text messages posted on that day. The overall training framework of the model is as shown in Fig. 4.

[Fig. 4 about here.]

4.2 Forecasting Model

To investigate the predictive effect of emotion indicators on the bitcoin market, we construct a prediction model for bitcoin returns. To emphasize the importance of emotional factors, we employ machine learning models. Compared to traditional econometric methods, machine learning models have a powerful capability to capture nonlinear relationships, allowing for more accurate predictions and analyses (Xiao et al., 2022). Furthermore, existing research suggests that financial markets contain many nonlinear factors that significantly influence market trends (Bukhari et al., 2020). Therefore, based on Li et al. (2019), we select support vector regression (SVR) and random forest (RF) as our prediction models.

SVR is a variation of support vector machine (SVM) used for regression tasks (Smola and Schölkopf, 2004) Similar to SVM, the goal of SVR is to find a regression function that can fit the training data and make predictions on test data. The crucial idea is to maximize the upper and lower margins of model predictions, which ensures the minimization of prediction errors. The key components of the model are the support vectors, which are training data points closest to the margin boundaries and are essential for the model's performance.

RF is an ensemble learning method based on decision trees. It works by constructing multiple decision trees and combining their predictions for classification and regression tasks (Breiman, 2001). RF has strong generalization capabilities, making it suitable for high-dimensional and large-scale datasets. It can automatically handle feature selection, reducing the need for extensive feature preprocessing. Additionally, through random sampling and feature selection, it effectively mitigates the problem of overfitting.

To prevent overfitting, machine learning models typically require data splitting (Lu et al., 2022). Therefore, we choose to use data from January 2018 to June 2021 as the training set to predict bitcoin returns in the test set (from July 2021 to June 2023, spanning 2 years). Following the recommendations from previous research (Li et al., 2022), we also consider lag effects by incorporating the returns from the past five days. We use a rolling window approach (with a prediction window matching the length of the training set, which is 1261 days) to predict bitcoin returns for the following day. While obtaining the best prediction model might not be guaranteed, we keep the hyperparameters of each machine learning model constant to better compare the contributions of different indicators to improving prediction performance.

5 Sentiment Analysis

In this section, we specifically investigate the differences that exist among the six different emotions in the text. By contrasting various texts, we analyze the behaviors of investors under different emotions. After formulating distinct investment strategies based on different emotions, we have reached some unique conclusions.

5.1 Emotional Description

We use the fine-tuned emotion classifier to perform sentiment analysis on the forum text, resulting in daily sentiment distributions. We plot trend graphs for the six emotion scores along the time axis. Additionally, to compare the relationship between bitcoin price movements and sentiment, we include daily return rate changes for bitcoin during the same period. Fig. 5 presents these results. The sentiment trends in the upper part of the graph are based on the BART pre-trained model, while the lower part is based on the RoBERTa model.

[Fig. 5 about here.]

In the graphs, we can observe distinct fluctuations in different sentiments, which resemble the trends in bitcoin returns. We've marked three-time points in the graph: the downturn in March 2020, the sustained uptrend in April 2021, and the downturn in May 2022. During these three periods, bitcoin experienced significant price fluctuations. Through comparison, we can see that when bitcoin prices decline, positive sentiments (such as Joy, Love, and Surprise) significantly decrease, while negative sentiments (Sadness, Fear, and Anger) show an upward trend. Conversely, sentiment changes move in the opposite direction to price changes. This aligns with real-world observations, as investors tend to feel happier when prices rise and experience lower moods when they incur losses (Garcia, 2013).

Furthermore, it's worth noting that during these three marked time points, sentiment changes sometimes precede price changes. In other words, when prices are about to decline, investor sentiment may turn negative first. This may imply that investor sentiment plays a certain leading role in price changes (Xu and Zhao, 2022). However, it's worth considering that when investors are extremely negative, it doesn't mean that a bear market has already arrived but may indicate that the bear market is imminent (after some time). We will validate this assumption below.

5.2 Interpreting Textual Analysis

Then, we will delve into the differences between different emotions. To do this, we perform sentiment classification on each user's posted text, labeling them with the emotion category having the highest probability. As a result, we categorize different texts into six classes representing their most likely emotional attributes. We conduct word frequency analysis for each category to demonstrate the frequency of occurrence of different words within each category. To emphasize the differences between the categories, we remove common words that appeared in each category with a frequency of more than 25%, such as "you," "be," "the," "we," and so on.

We generate emotion word clouds using the processed word frequency data, and Fig. 6 displays the word clouds for each emotion. In these word clouds, the font size is proportional to the word's frequency within the category. In the word cloud under the "Joy" emotion label, we can observe positive words like "kudos," "salute," and "celebration," indicating that investors on the forum are in a cheerful mood during this emotional feature. Similarly, from the "Love" emotion, we can see positive words like "passion" and "blessings," clearly showing positive emotions. However, in the word cloud under the "Surprise" category, the vocabulary is mainly neutral, without distinct positive features. It's important to note that people in a "Surprise" emotional feature don't necessarily feel happy. Surprise has the unclear valence (Noordewier and Breugelmans, 2013).

In terms of negative emotions, for the "Sadness" emotion, we can observe words like "fuck," "pierce," and "doomad," indicating that investors are feeling down. This emotion is more pronounced in the "Fear" category. It's noteworthy that the "Fear" category features a significant number of words like "hopefully," "opportunities," and "confidence," reflecting investors' inner uncertainty. When investors feel fear, they tend to be hesitant, lose confidence in their investments, and are more susceptible to herd behavior (Bekiros et al., 2017). Finally, under the "Anger" emotion, we can clearly sense the strong emotions of the investors. These word cloud visuals vividly illustrate the clear distinctions between the emotion categories we classified. There are evident boundaries between different emotions.

[Fig. 6 about here.]

To gain a deeper understanding of these emotions, we utilize the SHAP explanation method proposed by Lundberg and Lee (2017). This method assists us in comprehending the impact of each vocabulary in the text on the final emotion classification. The core concept of SHAP is derived from Shapley values in game theory (SHAPLEY, 1953), and has made significant progress in explaining the predictions of machine learning models, particularly when dealing with black-box models. Python's SHAP provides tools for text explanation. We select a text instance from each of the six different emotions obtained through the BERT pre-trained classifier. Fig. 7 presents the explanation results for the obtained text.

[Fig. 7 about here.]

In the figure, a higher SHAP value indicates a more significant contribution to that category. Red segments represent words contributing positively to the SHAP value, while blue segments represent words that suppress the SHAP value. This interpreter helps us gain a deeper understanding of how large language models (LLMs) comprehend semantic meanings. Taking the "Joy" label instance as an example, words like "excited," "up," and "good" all make significant contributions to this text leaning towards the "Joy" category. In the "Sadness" instance, "desperate" has the most substantial contribution, indicating that the investor was in a state of sadness. For the "Fear" emotion, we can infer that the investor was feeling fear through the words "worrying" and "dangerous." These explanations help us gain a deeper insight into the relevant emotions and tones for each emotion category, allowing us to better understand the emotions and behavior patterns of these investors.

5.3 Emotion-Based Investment

To explore whether different emotional atmospheres affect investments, we conduct a bitcoin investment simulation. The specific investment strategy is as follows: we select a specific emotion and use its median value as the threshold for that emotion throughout the entire period. If the emotion for a particular day is higher than the threshold, we choose to buy one bitcoin and hold it for a day; otherwise, we choose not to buy. This ensured that during the entire experimental period, we buy bitcoin on half of the days. We conduct investment simulations for six different emotions. Additionally, to establish a baseline, we also execute a strategy of buying bitcoin throughout the entire period (and holding it for a day).

Table 1 presents the return results for different bitcoin investment strategies. The baseline return, obtained by adopting a daily purchase strategy, is 35.79%. There are significant differences in returns under different emotional strategies. Surprisingly, regardless of which emotion classifier is used, investing in bitcoin when investor sentiment is generally high (Joy or Love) resulted in negative returns. However, choosing to invest when investor sentiment is relatively low (Sadness or Fear) yielded returns higher than the baseline. In particular, when using the "Fear" classification results obtained from the BERT pre-trained model, the return rate even reached an astonishing 187.79%. To provide a more detailed depiction of return rate variations under different strategies during the selected period, Fig. 8 illustrates the changes in return rates throughout the entire period based on the emotion classification results from the BERT pre-trained model.

[Table 1 about here.]

[Fig. 8 about here.]

The results presented in the figure reaffirm the findings from the table. It's evident that we obtain the highest investment returns when investor sentiment leans towards sadness, while we suffer the most significant losses in the "Joy" emotional atmosphere. This result is counterintuitive because, in general, when investors are in a positive emotional feature, it's commonly expected that market returns will increase (Shu, 2010). However, this counterintuitive outcome corroborates Warren Buffett's famous saying: "Be fearful when others are greedy, and be greedy when others are fearful" (Buffett, 2008). When investors are in high spirits, they tend to overestimate market prices and, as a result, are inclined to invest. This can lead to some investors blindly following, causing a herd effect. In such cases, it may not be the best investment opportunity because the market cannot be profitable for everyone (Edelman et al., 2016).

However, when investors are in a low emotional feature, especially when they experience fear, they will underestimate market prices and feel aversion to risks, hence choosing not to invest, which can also trigger a herd effect. But as mentioned earlier, when everyone is feeling fearful, the market's low point may not have arrived yet, and we can still achieve higher returns that others are unwilling to try. Rational investors who are not easily influenced by others' emotions can purchase bitcoin at lower prices and profit when market sentiments recover. This point has also been mentioned in the research by Day and Ni (2023).

In the above discussion, we don't mention two other emotions: "Surprise" and "Anger." For the "Surprise" emotion, we observe different outcomes under the two classifiers. Under the BERT classifier, it resulted in a loss (-2.32%) in the investment strategy, while under the RoBERTa classifier, it leads to significant returns (65.93%). This difference might be related to the instability of the "Surprise" emotion. As previously mentioned, Surprise is not a black-and-white emotion (Noordewier and Breugelmans, 2013). We believe this uncertainty contributes to the different results obtained under the two models. Regarding the "Anger" emotion strategy, we obtain returns lower than the benchmark under both models. This could be because the "Anger" emotion is too extreme. When people feel angry, their irrational tendencies are often amplified by emotions (Oh et al., 2021). In this emotional feature, they may exhibit unpredictable behavior (Jansz and Timmers, 2002), making it challenging to predict their investment strategies. Generally, such irrational behavior may lead to increased losses (Zizzo, 2008).

5.4 Binary Emotion Investment

To further validate the reliability of our previous conclusions, we conduct emotion classification again, simplifying the classification into two emotional categories: positive and negative emotions. In this section, we retrain our pre-trained model, utilizing the *tweet_sentiment_extraction* dataset provided by *HuggingFace*, which consists of nearly 30,000 tweets already labeled as positive or negative emotions.

By fine-tuning the new classifier, we obtain daily binary emotion distribution on

Bitcointalk. Similar to our previous approach, we formulate investment strategies under positive and negative emotions, and the return results are displayed in Table 2. The results align with our expectations, where we obtain higher-than-benchmark returns when choosing to invest in bitcoin under negative investor emotions. However, when we select to invest in bitcoin during positive investor sentiments, our returns are below the expected benchmark, even leading to losses. Under the classification results from the BERT classifier, our return rate is -1.41%. This result once again supports our hypothesis that making investment decisions when investors are generally in a negative emotional feature does not necessarily imply foolish choices.

[Table 2 about here.]

6 Bitcoin Forecasting

Using sentiment information provided by the BERT pre-trained classifier, we established a model to forecast bitcoin returns during a specific time frame. By incorporating various emotional information, we explored the impact of emotional factors on predictions.

6.1 Data and Evaluation

Table 3 displays the daily returns of bitcoin along with the descriptive statistics for the six emotion scores. It can be observed that the "Joy" emotion component is generally high in most of the textual data from the bitcoin forum. This contributes to the positive skewness in other emotions, with skewness values greater than 1, indicating a right-skewed distribution. The Jarque-Bera statistic (Jarque and Bera, 1980) rejects the null hypothesis at the 1% significance level for all variables, emphasizing that these variables follow non-Gaussian distributions. Additionally, the augmented Dickey-Fuller (ADF) test (Dickey and Fuller, 1979) suggests stationarity in these metrics.

[Table 3 about here.]

In order to facilitate a more comprehensive performance comparison of the predictive models, we employ two fundamental evaluation metrics: mean absolute error (MAE) and root mean squared error (RMSE), which are defined as follows:

$$\mathrm{MAE} = \frac{1}{n}\sum_{i=1}^{n}|(y_i - \hat{y}_i)| \tag{3}$$

$$\mathrm{RMSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2} \tag{4}$$

Where n is the number of observed data points during the testing period, y_i denotes the actual value of bitcoin return on day i, and $\hat{y}_i$ is the return forecast obtained using the forecasting model. During the training phase, we used a rolling window approach to assess predictions. The rolling window was set to 2500 days.

6.2 Forecasting Performance

Given the significantly different returns we obtain under various emotional conditions compared to the benchmark investment strategy, we have good reason to believe that emotions contain some information that can explain the bitcoin market. We introduce various daily emotional scores obtained earlier into the predictive models to forecast bitcoin's future returns. Considering the presence of lag effects in financial markets, we also include the past five periods of bitcoin returns in our consideration.

Given the high non-linearity in financial markets, according to Li et al. (2019), we select machine learning models, specifically support vector regression (SVR) and random forest (RF) regression, as our forecasting models for this study. Table 4 presents the predictive performance of the models under different input indicators. Following the research by Maciel (2021), we set the forecasting horizons to $h = 1$ and $h = 5$ to observe the model's performance in the short term and long term. Short-term forecasting focuses on recent changes, while long-term forecasting pays more attention to future trends. For comparative analysis, we also establish a baseline model that does not consider emotional information and only focuses on the lag effect of bitcoin price returns. This helps evaluate the degree of improvement in model performance after introducing emotional information.

[Table 4 about here.]

Based on the results analysis, emotional information can lead to a certain degree of improvement in predictive models. Particularly, the "Joy" emotional information has the most significant impact on results for both prediction models. In the case of a forecasting horizon of $h = 5$, the MAE loss under "Joy" emotional information is 81.98% of the baseline SVR model, showing the greatest improvement. In comparison, the impact of "Sadness" and "Fear" emotional indicators is somewhat lower, with their contributions to predictive model enhancement being fairly similar. These three emotional features show consistent performance with the previous investment results. In portfolio strategies, they have all led to higher (or lower) returns, and they also exhibit stronger explanatory power in their predictions compared to several other features.

Overall, emotional indicators have a more substantial impact on the SVR model compared to the RF model. Additionally, it's worth noting that especially in long-term forecasting, emotional information has a more pronounced effect on model performance.

This result further underscores the potential value of emotional information in practical applications (Whiting et al., 2011).In line with the ideas of Edmans et al. (2022), emotions exhibit a certain lag effect in the financial markets. Our research findings demonstrate the predictive power of emotional information in forecasting Bitcoin market changes.

6.3 Combined Factor

Finally, we consider combining the "Joy" and "Sadness" emotional indicators and introduced them together into the predictive model. The reason for choosing these two indicators is that, after comparison, we find that these two types of emotions best represent positive and negative emotions. Similarly, to emphasize the contributions of these indicators, based on the previous comparative results, we select SVR as the sole prediction model. In Table 5, we showcase the performance improvement of our combined factor over the SVR baseline model under different forecasting horizons. Comparing the results, we find that positive emotion ("Joy") and negative emotion ("Sadness") have complementary effects. When they are combined, the impact on the performance of the predictive model is significantly greater than when they are used separately. Under the influence of the combined factors, the loss of our predictive model decreased by nearly 20%. Especially in the long-term forecasting horizon ($H = 5$), the effect of this combined factor on predictions is even more significant. This result highlights the inconsistency in the economic information contained in different emotions. In line with the findings of Dias et al. (2022), when we consider multiple aspects of investor emotions comprehensively, we can more accurately predict future return trends.

Based on the above results, we have confirmed the significant role that emotional factors play in predicting the bitcoin market. Additionally, the diverse performance of various emotions is evident. Integrating them might offer a more comprehensive understanding of the market situation.

[Table 5 about here.]

7 Conclusion

"When others are fearful, be greedy; when others are greedy, be fearful" (Buffett, 2008). Perhaps Warren Buffett's famous quote indeed holds the magic that led to his success. In this study, we utilize a Large Language Model (LLM) for sentiment analysis of bitcoin forum data, categorizing it into six different emotional features. Through in-depth analysis of the textual content under these six emotional features, we reveal that the bitcoin market is influenced by various irrational factors, and investors, especially retail ones, often exhibit pronounced herd behavior. When investors fall into positive emotional features (especially joy or love), they tend to mistakenly believe that the bitcoin market has a bright future, leading to collective investment decisions. However, our simulation results indicate that these joyful investors may ultimately incur losses. On the contrary, when investors' emotions hit a low point, especially when they are generally feeling fearful, they tend to avoid risk and are reluctant to invest in bitcoin. Our results suggest that this hesitation may lead to missed profit opportunities. This intriguing finding reaffirms a fundamental principle: investment decisions should not be blindly following the crowd, but rather require investors to maintain a clear and rational mindset (McMullen, 2015).

After our brief bitcoin return predictions, we are pleasantly surprised to discover that the results underscore the crucial role of emotional factors in more accurately predicting bitcoin market performance. Different types of emotions carry distinct economic information, emphasizing the need to consider the complementary nature of these emotions. This finding further underscores the importance of sentiment analysis in the bitcoin market, providing valuable insights for market analysis and forecasting. Moreover, the application potential of sentiment analysis is not limited to the bitcoin market but can also be extended to other financial markets and various asset classes.

As a non-traditional financial market, bitcoin is full of elements that challenge conventional financial theories. From the emergence of blockchain technology to the introduction of the concept of virtual currency by Nakamoto (2008), the close connection between bitcoin and investors is an unprecedented phenomenon in traditional markets. However, our research only reveals a part of the puzzle. In the field of sentiment analysis, there are still many aspects to be improved and explored further. Just as Demszky et al. (2020) proposed 27 different ways to categorize emotions in more detail, we can consider introducing more emotional features to provide a deeper explanation of differences in investor behavior. Furthermore, with the ongoing development of large-scale pre-trained models, audio and video data may become new avenues to gain a deeper understanding of emotional associations. But whatever the future holds, similar to the perspective presented by Duxbury et al. (2020), emotions will always remain an indispensable topic in finance.

References

Airani, R. and Karande, K. (2022). How social media effects shape sentiments along the twitter journey? a bayesian network approach. *Journal of Business Research*, 142:988–997.

Alain, G. and Bengio, Y. (2016). Understanding intermediate layers using linear classifier probes. *arXiv preprint arXiv:1610.01644*.

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.

Bekiros, S., Jlassi, M., Lucey, B., Naoui, K., and Uddin, G. S. (2017). Herding behavior, market sentiment and volatility: will the bubble resume? *The North American journal of economics and finance*, 42:107–131.

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

Buffett, W. E. (2008). Buy american. i am. *The New York Times*, 16.

Bukhari, A. H., Raja, M. A. Z., Sulaiman, M., Islam, S., Shoaib, M., and Kumam, P. (2020). Fractional neuro-sequential arfima-lstm for financial market forecasting. *Ieee Access*, 8:71326–71338.

Cao, Z., Shen, J., Wei, X., and Zhang, Q. (2023). Anger in predicting the index futures returns. *Journal of Futures Markets*, 43(4):437–454.

Chen, M., Qin, C., and Zhang, X. (2022). Cryptocurrency price discrepancies under uncertainty: evidence from covid-19 and lockdown nexus. *Journal of International Money and Finance*, 124:102633.

Christensen, K., Siggaard, M., and Veliyev, B. (2022). A machine learning approach to volatility forecasting. *Journal of Financial Econometrics*, page nbac020.

Consoli, S., Pezzoli, L. T., and Tosetti, E. (2021). Emotions in macroeconomic news and their impact on the european bond market. *Journal of International Money and Finance*, 118:102472.

Corbet, S., Hou, Y. G., Hu, Y., Larkin, C., and Oxley, L. (2020). Any port in a storm: Cryptocurrency safe-havens during the covid-19 pandemic. *Economics letters*, 194:109377.

Dang, H. V. and Lin, M. (2016). Herd mentality in the stock market: On the role of idiosyncratic participants with heterogeneous information. *International Review of Financial Analysis*, 48:247–260.

Day, M.-Y. and Ni, Y. (2023). Be greedy when others are fearful: Evidence from a two-decade assessment of the ndx 100 and s&p 500 indexes. *International Review of Financial Analysis*, 90:102856.

Demszky, D., Movshovitz-Attias, D., Ko, J., Cowen, A., Nemade, G., and Ravi, S. (2020). Goemotions: A dataset of fine-grained emotions. *arXiv preprint arXiv:2005.00547*.

Devlin, J., Chang, M.-W., Lee, K., and Toutanova, K. (2018). Bert: Pre-training of deep bidirectional transformers for language understanding. *arXiv preprint arXiv:1810.04805*.

Dias, I. K., Fernando, J. R., and Fernando, P. N. D. (2022). Does investor sentiment predict bitcoin return and volatility? a quantile regression approach. *International Review of Financial Analysis*, 84:102383.

Dickey, D. A. and Fuller, W. A. (1979). Distribution of the estimators for autoregressive time series with a unit root. *Journal of the American statistical association*, 74(366a):427–431.

Duxbury, D., Gärling, T., Gamble, A., and Klass, V. (2020). How emotions influence behavior in financial markets: a conceptual analysis and emotion-based account of buy-sell preferences. *The European Journal of Finance*, 26(14):1417–1438.

Edelman, B., Jaffe, S., and Kominers, S. D. (2016). To groupon or not to groupon: The profitability of deep discounts. *Marketing Letters*, 27(1):39–53.

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.

Garcia, D. (2013). Sentiment during recessions. *The journal of finance*, 68(3):1267–1300.

Gu, S., Kelly, B., and Xiu, D. (2020). Empirical asset pricing via machine learning. *Review of Financial Studies*, 33(5).

Guégan, D. and Renault, T. (2021). Does investor sentiment on social media provide robust information for bitcoin returns predictability? *Finance Research Letters*, 38:101494.

Güler, D. (2023). The impact of investor sentiment on bitcoin returns and conditional volatilities during the era of covid-19. *Journal of Behavioral Finance*, 24(3):276–289.

Jansz, J. and Timmers, M. (2002). Emotional dissonance: When the experience of an emotion jeopardizes an individual's identity. *Theory and Psychology*, 12(1):79–95.

Jarque, C. M. and Bera, A. K. (1980). Efficient tests for normality, homoscedasticity and serial independence of regression residuals. *Economics letters*, 6(3):255–259.

Jegadeesh, N. and Wu, D. (2013). Word power: A new approach for content analysis. *Journal of financial economics*, 110(3):712–729.

Jha, M., Liu, H., and Manela, A. (2021). Natural disaster effects on popular sentiment toward finance. *Journal of Financial and Quantitative Analysis*, 56(7):2584–2604.

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

Kaplanski, G. and Levy, H. (2010). Sentiment and stock prices: The case of aviation disasters. *Journal of financial economics*, 95(2):174–201.

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.

Leippold, M. (2023). Thus spoke gpt-3: Interviewing a large-language model on climate finance. *Finance Research Letters*, 53:103617.

Li, J., Li, G., Liu, M., Zhu, X., and Wei, L. (2022). A novel text-based framework for forecasting agricultural futures using massive online news headlines. *International Journal of Forecasting*, 38(1):35–50.

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.

Li, Y., Bu, H., Li, J., and Wu, J. (2020). The role of text-extracted investor sentiment in chinese stock price prediction with the enhancement of deep learning. *International Journal of Forecasting*, 36(4):1541–1562.

Liu, Y., Ott, M., Goyal, N., Du, J., Joshi, M., Chen, D., Levy, O., Lewis, M., Zettlemoyer, L., and Stoyanov, V. (2019). Roberta: A robustly optimized bert pretraining approach. *arXiv preprint arXiv:1907.11692*.

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.

Lundberg, S. M. and Lee, S.-I. (2017). A unified approach to interpreting model predictions. *Advances in neural information processing systems*, 30.

Maciel, L. (2021). Cryptocurrencies value-at-risk and expected shortfall: Do regime-switching volatility models improve forecasting? *International Journal of Finance & Economics*, 26(3):4840–4855.

McMullen, J. S. (2015). Entrepreneurial judgment as empathic accuracy: A sequential decision-making approach to entrepreneurial action. *Journal of Institutional Economics*, 11(3):651–681.

Nakamoto, S. (2008). Bitcoin: A peer-to-peer electronic cash system. *Available at SSRN 3440802*.

Nassirtoussi, A. K., Aghabozorgi, S., Teh, Y. W., and Ngo, D. C. L. (2014). Text mining for market prediction: A systematic review. *Expert Systems with Applications*, 41(16):7653–7670.

Noordewier, M. K. and Breugelmans, S. M. (2013). On the valence of surprise. *Cognition and emotion*, 27(7):1326–1334.

Oh, S.-H., Lee, S. Y., and Han, C. (2021). The effects of social media use on preventive behaviors during infectious disease outbreaks: The mediating role of self-relevant emotions and public risk perception. *Health communication*, 36(8):972–981.

Park, J. and Seo, Y.-S. (2023). Twitter sentiment analysis-based adjustment of cryptocurrency action recommendation model for profit maximization. *IEEE Access*, 11:44828–44841.

Poongodi, M., Nguyen, T. N., Hamdi, M., and Cengiz, K. (2021). Global cryptocurrency trend prediction using social media. *Information Processing and Management*, 58(6):102708.

Ren, R. and Wu, D. (2018). An innovative sentiment analysis to measure herd behavior. *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, 50(10):3841–3851.

Saggu, A. and Ante, L. (2023). The influence of chatgpt on artificial intelligence related crypto assets: Evidence from a synthetic control analysis. *Finance Research Letters*, page 103993.

Saravia, E., Liu, H.-C. T., Huang, Y.-H., Wu, J., and Chen, Y.-S. (2018). Carer: Contextualized affect representations for emotion recognition. In *Proceedings of the 2018 conference on empirical methods in natural language processing*, pages 3687–3697.

SHAPLEY, L. (1953). A value for n-person games. *Annals of Mathematics Studies*, 28:307–318.

Shleifer, A. and Summers, L. H. (1990). The noise trader approach to finance. *Journal of Economic perspectives*, 4(2):19–33.

Shu, H.-C. (2010). Investor mood and financial markets. *Journal of Economic Behavior and Organization*, 76(2):267–282.

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

Stevenson, M., Mues, C., and Bravo, C. (2021). The value of text for small business default prediction: A deep learning approach. *European Journal of Operational Research*, 295(2):758–771.

Sun, X., Liu, M., and Sima, Z. (2020). A novel cryptocurrency price trend forecasting model based on lightgbm. *Finance Research Letters*, 32:101084.

Tian, X., Pavur, R., Han, H., and Zhang, L. (2022). A machine learning-based human resources recruitment system for business process management: using lsa, bert and svm. *Business Process Management Journal*, 29(1):202–222.

Toyoda, K., Mathiopoulos, P. T., and Ohtsuki, T. (2019). A novel methodology for hyip operators' bitcoin addresses identification. *IEEE Access*, 7:74835–74848.

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.

Wang, A., Singh, A., Michael, J., Hill, F., Levy, O., and Bowman, S. R. (2018). Glue: A multi-task benchmark and analysis platform for natural language understanding. *arXiv preprint arXiv:1804.07461*.

Whiting, A., Donthu, N., and Baker, A. M. (2011). Investigating the immediate and long-term effects of job stressors on frontline service employees. *International Journal of Research in Marketing*, 28(4):319–331.

Wu, D. D., Zheng, L., and Olson, D. L. (2014). A decision support approach for online stock forum sentiment analysis. *IEEE transactions on systems, man, and cybernetics: systems*, 44(8):1077–1087.

Xiao, D., Su, J., et al. (2022). Research on stock price time series prediction based on deep learning and autoregressive integrated moving average. *Scientific Programming*, 2022.

Xiong, J., Liu, Q., and Zhao, L. (2020). A new method to verify bitcoin bubbles: Based on the production cost. *The North American Journal of Economics and Finance*, 51:101095.

Xu, Y. and Zhao, J. (2022). Can sentiments on macroeconomic news explain stock returns? evidence form social network data. *International Journal of Finance and Economics*, 27(2):2073–2088.

Zhou, G. (2018). Measuring investor sentiment. *Annual Review of Financial Economics*, 10:239–259.

Zhou, Y., Myrick, J. G., Farrell, E. L., and Cohen, O. (2023). Perceived risk, emotions, and stress in response to covid-19: The interplay of media use and partisanship. *Risk Analysis*, 43(8):1572–1586.

Zizzo, D. J. (2008). Anger and economic rationality. *Journal of Economic Methodology*, 15(2):147–167.

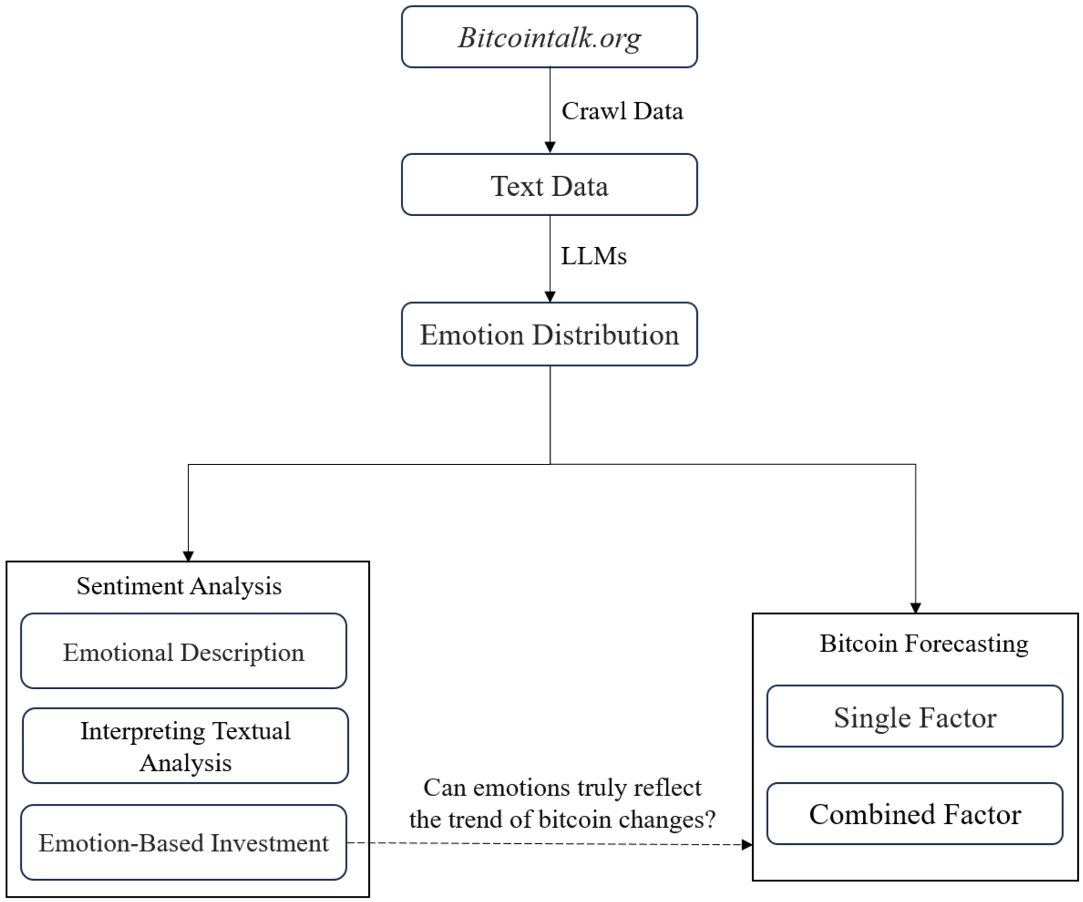


Fig. 1: The Main Technical Research Approach

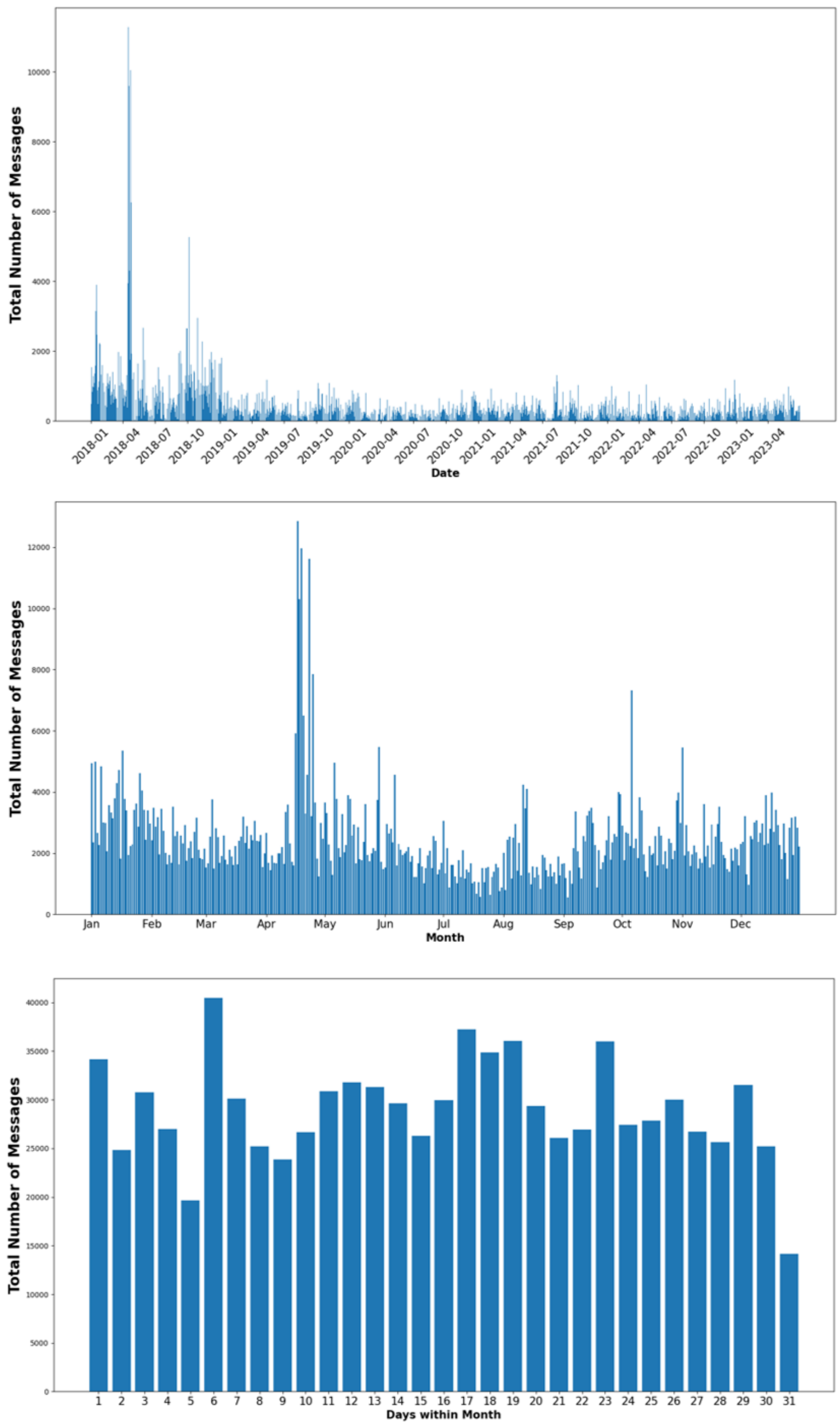


Fig. 2: Bitcointalk Messages Counts

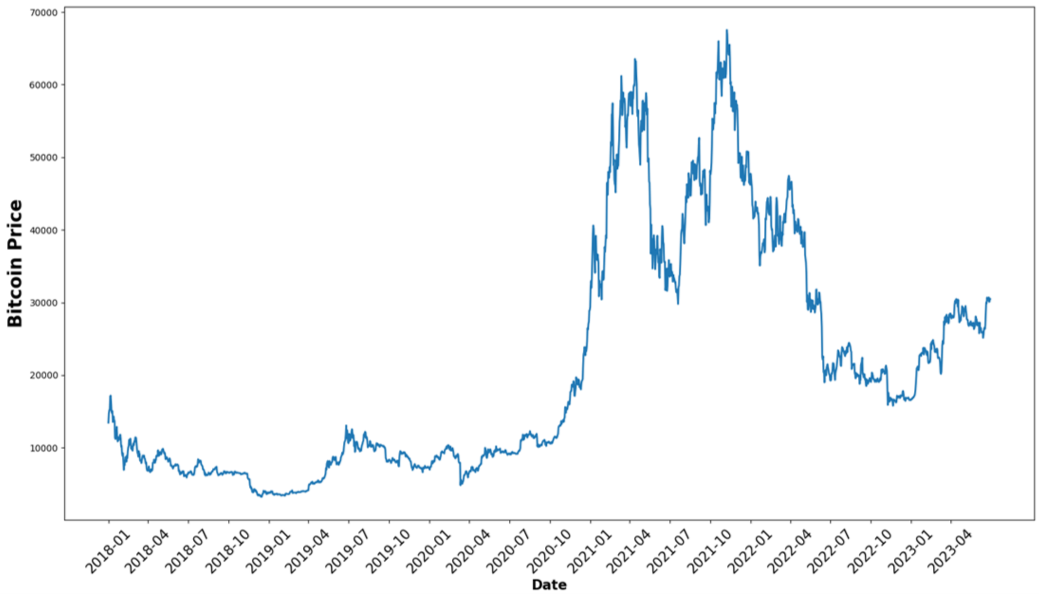


Fig. 3: Bitcointalk Price Trend

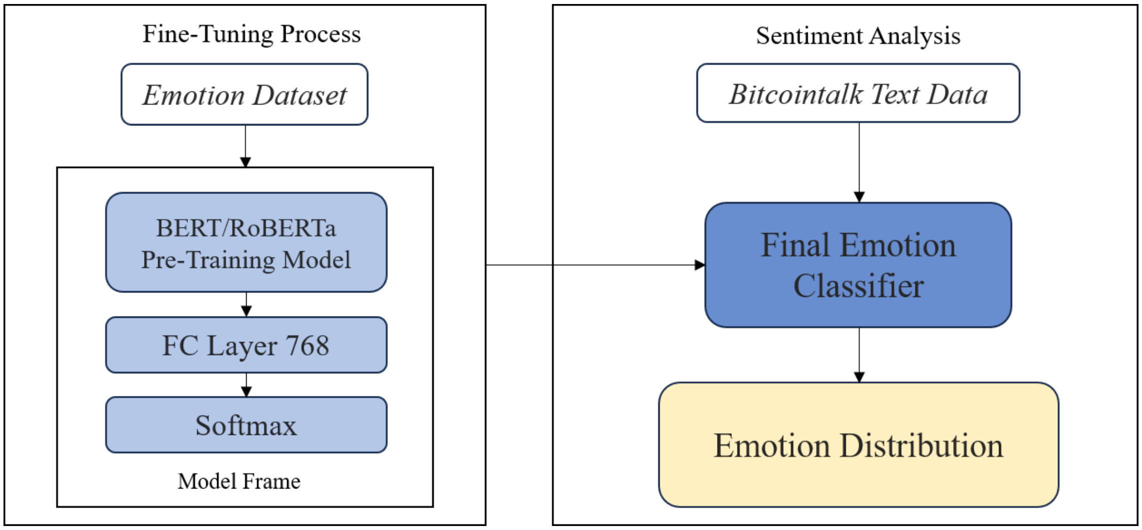
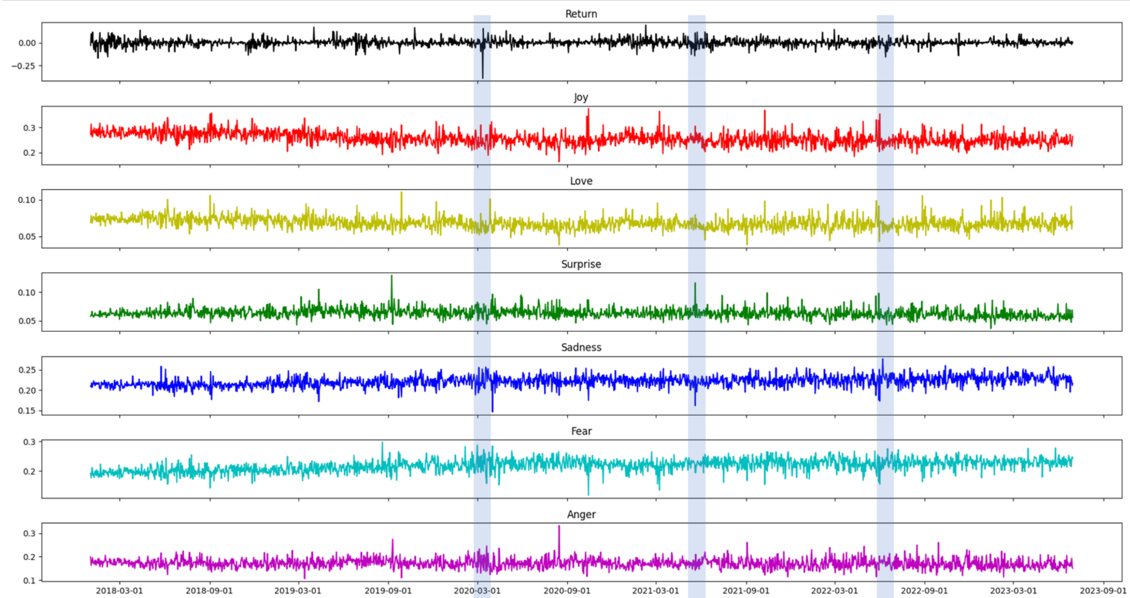


Fig. 4: Emotion Classification Model Training Framework

(a) BERT



(b) RoBERTa

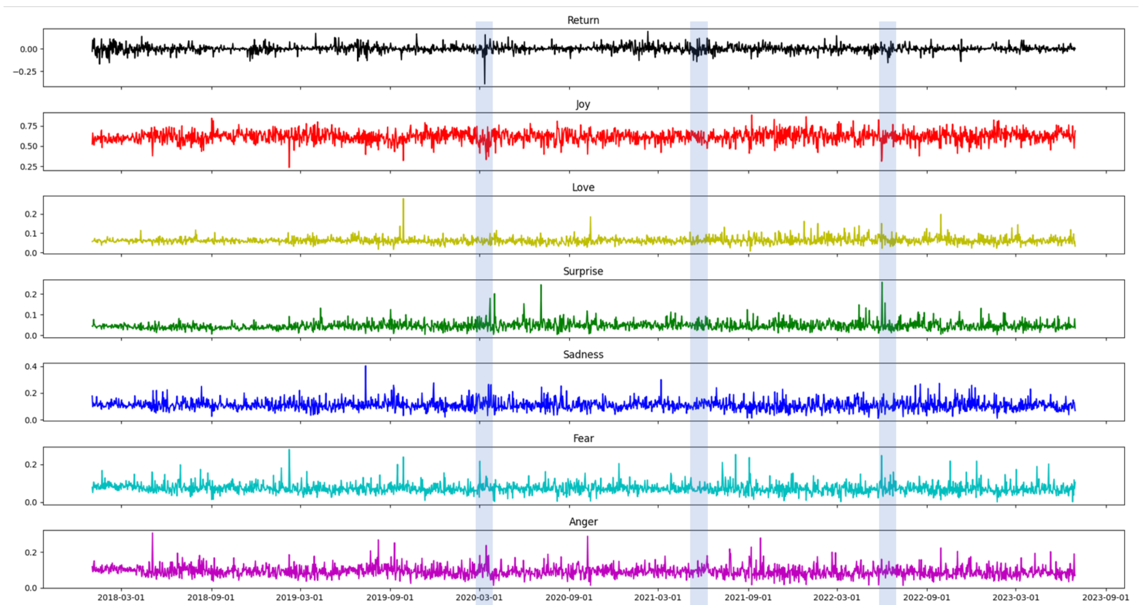
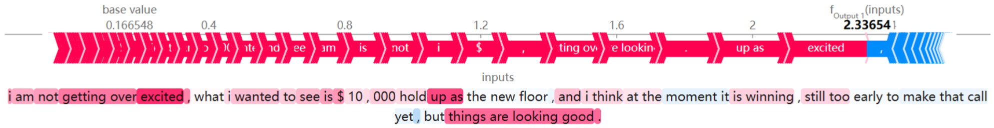


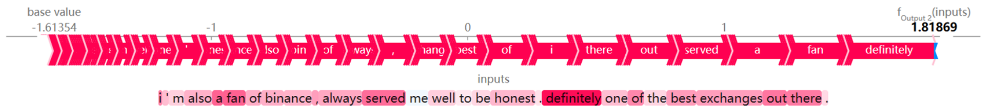
Fig. 5: Daily Trend of Emotion Score

Fig. 6: Word Clouds for Six Emotions

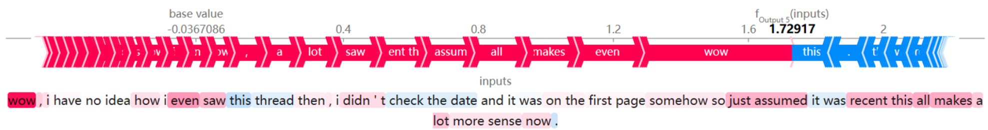
(a) Joy



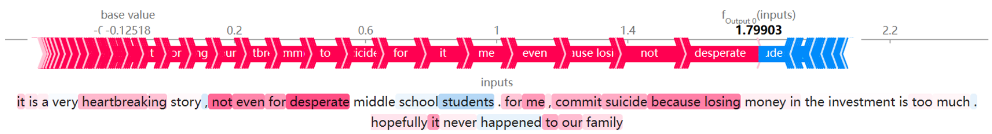
(b) Love



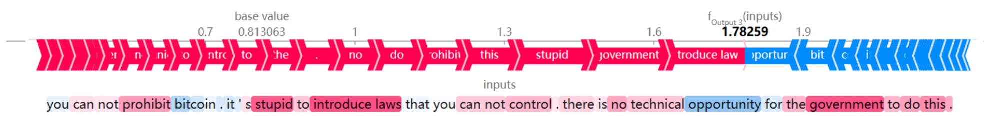
(c) Surprise



(d) Sadness



(e) Anger



(f) Fear

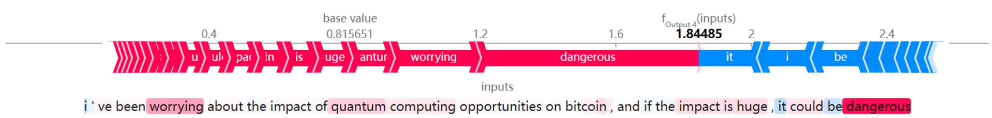


Fig. 7: SHAP Explanations for Examples of Six Emotions

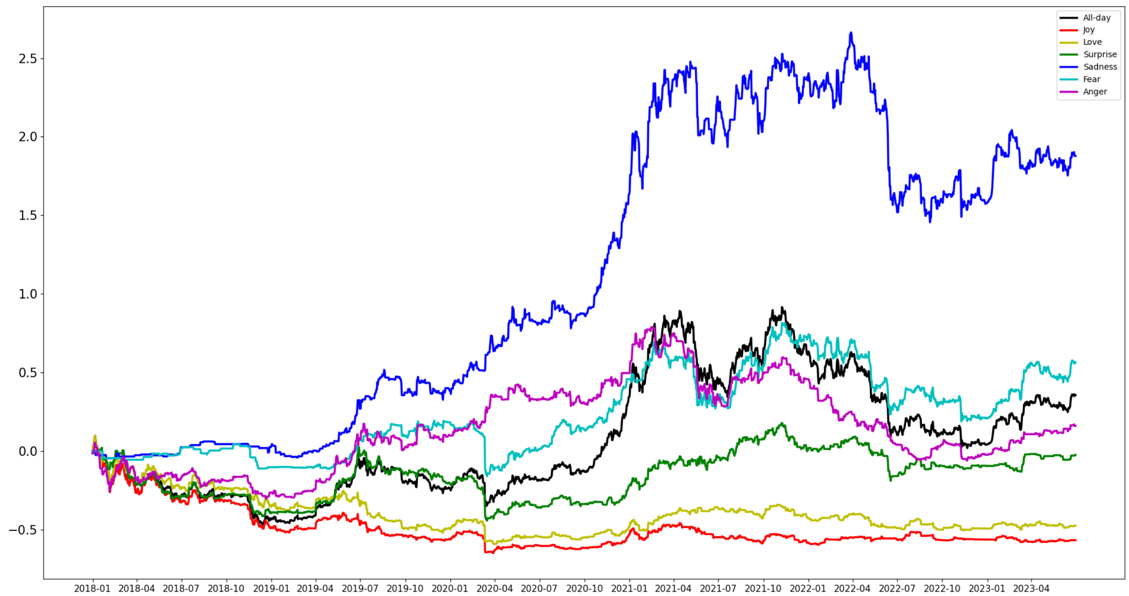


Fig. 8: The Daily Trends of Investment Strategies

Table 1: The Return Results of Different Investment Strategies

	BERT	RoBERTa
Joy	**−56.59**%	**−17.81**%
Love	−47.41%	−15.28%
Surprise	−2.32%	65.93%
Sadness	**187.79**%	45.98%
Anger	16.09%	−16.73%
Fear	56.54%	**126.72**%
Baseline	35.79%	35.79%

Notes: This table displays the final return results under different emotional strategies. The second column represents the emotional classification based on the BERT classifier, and the third column is based on RoBERTa. The last row in the table represents the return under the daily buy-in strategy, serving as the benchmark for this investment.

Table 2: The Return Results of Binary Emotional Strategies

	BERT	RoBERTa
Negative	77.70%	54.12%
Positive	−1.41%	19.72%
Baseline	35.79%	35.79%

Notes: This table displays the final return results under binary emotional strategies. The second column represents the emotional classification based on the BERT classifier, and the third column is based on RoBERTa. The last row in the table represents the return under the daily buy-in strategy, serving as the benchmark for this investment.

Table 3: Descriptive Statistics of Forecasting Variables

	Mean	Median	Std. Dev.	Skewness	Kurtosis	Jarque-Bera	Augmented Dicky-Fuller
Return Rate	0.0010	0.0005	0.0380	-0.4604	8.1288	5525.9141***	−31.1814***
Joy	0.6159	0.6170	0.0688	-0.2897	1.3048	167.8515***	−9.5387***
Love	0.0623	0.0607	0.0167	2.4173	19.8287	34414.8921***	−13.1621***
Surprise	0.0468	0.0436	0.0198	2.4220	16.0113	23119.7542***	−11.4017***
Sadness	00.1093	0.1050	0.0367	1.1133	3.9777	1716.4197***	−30.0642***
Fear	0.0733	0.0702	0.2646	1.7052	7.9071	6126.5446***	−44.0183***
Anger	0.0923	0.0890	0.0296	1.2923	5.1233	2720.1884***	−6.3804***

Notes: This table displays the descriptive statistics for daily bitcoin return and the associated emotional distribution. ***, ** and * denote rejections of the null hypotheses at the 1% , 5% and 10% significance level, respectively.

Table 4: Comparative Analysis of Model Predictive Performance

	H=1			
	SVR		RF	
	RMSE	MAE	RMSE	MAE
Baseline	1.0000	1.0000	1.0000	1.0000
Joy	0.8324	0.8304	0.9643	0.9470
Sadness	0.8483	0.8374	0.9661	0.9503
Fear	0.8439	0.8425	0.9640	0.9494
Love	0.9804	0.9802	0.9624	0.9452
Anger	0.8752	0.8764	0.9624	0.9466
Surprise	0.9767	0.9624	0.9642	0.9489
	H=5			
	SVR		RF	
	RMSE	MAE	RMSE	MAE
Baseline	1.0000	1.0000	1.0000	1.0000
Joy	0.8233	0.8198	0.9697	0.9452
Sadness	0.8369	0.8257	0.9725	0.9452
Fear	0.8342	0.8307	0.9674	0.9411
Love	0.9754	0.9690	0.9704	0.9476
Anger	0.8678	0.8705	0.9693	0.9446
Surprise	0.9707	0.9552	0.9702	0.9478

Notes: The table displays the performance of forecasting models with different emotional indicators for short-term (upper section) and long-term (lower section). The loss for each emotional feature is represented as a ratio compared to the baseline.

Table 5: The Improvement of the Combined Factor on SVR Model

	H=1		H=5	
	RMSE	MAE	RMSE	MAE
Baseline	1.0000	1.0000	1.0000	1.0000
Joy	0.8324	0.8304	0.8233	0.8198
Sadness	0.8483	0.8374	0.8369	0.8257
Joy and Sadness	**0.8189**	**0.8105**	**0.8083**	**0.7986**

Notes: This table displays the performance differences of the model with the combined factor in different forecasting horizons. The loss for each emotional feature is represented as a ratio compared to the baseline.