IEEE Access
Multidisciplinary : Rapid Review : Open Access Journal

Received 12 January 2024, accepted 1 February 2024, date of publication 26 February 2024, date of current version 1 March 2024.

Digital Object Identifier 10.1109/ACCESS.2024.3369033

RESEARCH ARTICLE

Investigate How Market Behaves: Toward an Explanatory Multitasking Based Analytical Model for Financial Investments

SARMISTHA DAS [1], UJJWAL CHOWDHURY [2], N. S. LIJIN [3], ATULYA DEEP [4], SRIPARNA SAHA [1], (Senior Member, IEEE), AND ALKA MAURYA [5]

[1]CSE Department, IIT Patna, Patna, Bihta 801106, India
[2]Ramakrishna Mission Vivekananda Educational and Research Institute, Belur, West Bengal 711202, India
[3]IISER Bhopal, Bhopal, Madhya Pradesh 462066, India
[4]SRM Institute of Science and Technology, Chennai 603203, India
[5]Crisil Pvt., Ltd., Mumbai 400076, India

Corresponding author: Sarmistha Das (sarmistha_2221cs21@iitp.ac.in)

This work was supported in part by Crisil Data Science Private Ltd.

ABSTRACT In recent years, the intersection between financial market behavior and social media has emerged as a sought-after source of information, meeting the requirements of investors, institutions, regulators, researchers, and policymakers. Assessing sentiment and emotions aids in evaluating public psychology on particular stocks, assets, or the overall market, with shifts often aligning with market movements. Previously, machine learning, both traditional and deep learning methods, targeted discerning stock market sentiment and emotion without conducting studies to offer comprehensive explanations for these behavioral factors. In this study, we introduce a multitasking sequence-to-sequence model that integrates financial investment analysis with sentiment and emotion analysis from tweets upheld by an explanation mechanism. We also present the *FinEMA* dataset, featuring sentiment, emotion, and cause labels on financial stock market changes. Our study highlights how joint learning improves performance in discerning sentiment and emotion by utilizing interrelated features, enhancing task effectiveness. Our proposed model, the Emotion-Sentiment Attention Network (ESAN), achieved 89% accuracy in sentiment identification and 79% accuracy in emotion recognition, outperforming conventional machine learning methods. Furthermore, our findings indicate a positive outlook for the stock market in the latter half of 2023, which has intensified investor optimism, though some individuals still harbor uncertainties. Conclusively, our results suggest that regenerating existing computational tools can open up new research opportunities to address relevant novel tasks. The primary aim of this study is to elucidate the diverse dimensions of financial market behaviour and offer explanatory insights for the research community. The authors maintain impartiality towards specific stocks. It's essential to note that stock market investments inherently carry market risks and potential losses. The market information within the research findings remains independent of the authors' viewpoints.

INDEX TERMS Financial market, emotion and sentiment analysis, explainability, multi-tasking, sequence-to-sequence, social sentiment on the stock market.

I. INTRODUCTION

The recent amalgamation of artificial intelligence and finance has instigated a paradigm shift in the financial domain, notably fueling the surge in demand for online investment avenues rooted in financial market dynamics [1]. Utilizing fintech tools, investors can access up-to-the-minute financial data, enabling them to scrutinize market patterns and make well-informed investment choices [2]. Digital currencies like Bitcoin, Ethereum, Litecoin, Ripple, and Tether are emerging currencies that have started to gain significant popularity along with financial stocks [3]. In September 2023, the investment participation of individuals aged $18-20$ abruptly escalated to 16.1 million, marking a notable increase

The associate editor coordinating the review of this manuscript and approving it for publication was Arianna DUlizia.

from the 3.4 million recorded in 2021.[1] Earlier, in 2011, Chong and Sheng first demonstrated the power of financial blogs [4] to determine the sentiment of posts related to a particular stock and used this information to predict its future price movements. However, due to the temporal nature of financial data, keeping track of each event becomes difficult. In 2015, Preethi et al. illustrated through their work [5] the significance of causal relationships between events based on sentiment and time to understand cause and effect. Gradually, Inamdar et al. [6] pioneered how investors utilized cryptocurrency sentiment analysis with Explainable AI to inform their investment decisions. Subsequently, Fior et al. [7] illustrated explainable AI algorithms can leverage cryptocurrency predictions and aid crypto investors with practical guidance. In the domain of digital currency, optimistic tweets hold significant importance in forecasting market trends, as evidenced in research by Wolk [8]. In recent times, several extensive language models like BloombergGPT [9] and FinMA [10] fulfil diverse functions such as named-entity identification, question response, sentiment assessment, among others.

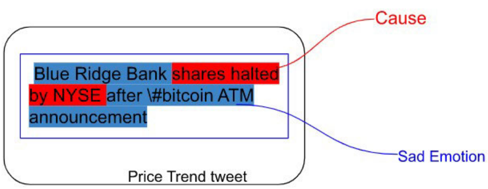

FIGURE 1. Example of explainable financial market behaviour.

Moreover, Narman and Uulu [11] conducted a study that revealed the effects of social media comments, such as those on Twitter, Facebook, and Instagram, on cryptocurrency prices. Our research intends to incorporate the fundamental development of recognizing the causal span of expressed emotions in automatic reasoning about human emotions [12] to understand the strategy behind financial investment reviews shared on social media. Figure 1 illustrates that shares halted by the NYSE of Blue Ridge Bank are indeed sad news; hence, the emotion is recognized as sad, and the cause of such an instance is NYSE halted the shares. Therefore, by employing the suggested framework, we can examine the emotions associated with currency trends and understand the underlying reasons that contribute to their downturn nature. However, the following challenges need to be considered:

- Assessing an investment's social or environmental impact is challenging due to insufficient market data, hindering comparisons between investments and the evaluation of associated risks.
- The intrinsic temporal fluctuations in financial data, including seasonal patterns, liquidity during trading hours, policy and regulatory shifts, and cyclical trends, amplify the unpredictability of financial market behaviour.
- Accurate forecasting in finance requires ample up-to-date training data for the model. Yet, the restricted access to public data, stemming from financial institutions' privacy concerns, presents a major obstacle in obtaining an adequate volume of data for training.

Research Objectives: Following are the research objectives of the current study:

- Our main objective is to study financial market behaviour by understanding public psychology and correlating these insights to pinpoint the underlying causes behind this behaviour. Through causal explanations, we intend to improve entrepreneurs' and investors' understanding of finance, aiding in more informed decision-making in the future.
- Following that, we intend to explore diverse task combinations that encompass sentiment categorization (SC), emotion recognition (ER), and cause extraction (CE) to evaluate their implications on predictive accuracy within the financial domain.

Contributions: Major contributions of the current work are as follows:

- We constructed our customized *FinEMA* dataset by merging publicly available datasets with supplementary scraped online tweets. These tweets underwent a rigorous examination process, during which we manually annotated the causal spans for sentences and assigned sentiment and emotion labels. The *FinEMA* dataset comprehensively represents the intricate relationship between sentiment, emotion, and causal factors in the financial domain.
- Additionally, we develop an Emotion-Sentiment Attention Network (ESAN) model specifically designed to analyze financial market behaviour. Utilising tweet reviews as its base, this model is crafted to elucidate the underlying factors influencing public sentiment systematically.
- Furthermore, we present a hierarchical attention-based approach that tackles two main sub-problems. The first one involves handling emotion recognition (ER) and sentiment categorization (SC) together, while the second deals with cause extraction (CE). We explore various setups for each problem and assess accuracy performance in both single-task and multi-task scenarios.

II. RELATED WORKS

Present-day Generation Z and Millennials exhibit heightened caution regarding financial affairs, particularly in the context of the stock market. They actively employ diverse AI tools to analyze market behaviours [13]. Vanstone et al. [14] demonstrated the sentiment relationship between news articles and stock price. Similarly, in 2020, the work [15] briefly investigated that positive tweet sentiment leads to stock market price increment. Social media like Twitter, Reddit, and Instagram provide an accessible platform for discussing

[1] https://young investors

TABLE 1. The highlighted portions represent the causal span duration marking in the *FinEMA* data samples. Further classification labels pertain to tasks involving Sentiment and emotion.

Text	Sentiment	Emotion	Span
Blue Ridge Bank shares halted by NYSE after #bitcoin ATM announcement	negative	sad	shares halted
"This network is secured by 9 508 nodes as of today. Soon, the biggest bears will recognise: #BTC"	negative	disgust	biggest bears
"#Bitcoin to Hit $60-64K Soon, massive price dips ahead. #BTC #Binance #BNB"	negative	disgust	massive price dips
dip down to 39kish tomorrow? should stir enough panic $btc #btc #Bitcoin	negative	confused	dip down
Annnd #btc #Bitcoin is headed even higher now ...	positive	relief	headed even higher now
Bitcoin institutional demand accelerates in 2021 says @Grayscale CEO	positive	relief	demand accelerates
#Bitcoin #BTC #ADA #DOT Mastercard Will Let Merchants Accept Payments in Crypto This Year	positive	happy	Accept Payments in Crypto

and sharing cryptocurrency information [16]. Recently, Huynh [17] demonstrated how cryptocurrency market behaviour significantly relies on social media. Various machine-learning and deep-learning methods have already been implemented for cryptocurrency sentiment prediction [18], [19]. The general public's psychology regarding cryptocurrency is known as "crypto sentiment". Traders and investors use crypto sentiment analysis as a key indicator of future price fluctuations, where cryptocurrency-related tweets and news headlines are utilized. Good positive posts regarding the crypto market are typically linked to increased demand for cryptocurrencies and vice versa. For instance, Tang et al. [20] suggested that the value of cryptocurrencies can be estimated using machine learning methods. The Chinese social networking site, Sina-Weibo, has mined for cryptocurrency-related posts for examination. According to the study [21], there is a correlation between the bitcoin fluctuation price and the tone of tweets. In 2022, Valle-Cruz et al. [22] proposed that tweeter has a significant effect on the stock market. Moreover, emotion analysis can be useful in aiding investors to decide on the appropriate time to buy or sell their cryptocurrency investments [23]. In [24], Haryadi and Kusuma showed variants of emotions in tweet data.

Although sentiment analysis and emotion detection are widely used in financial market behaviour forecasting, there has been limited research focused on integrating both aspects together. To the best of our knowledge, only one recent publication by Aslam et al. [25] explored both sentiment analysis and emotion detection using a multitask learning framework. Due to the temporal nature of financial data, Explainable AI (XAI) can be utilized to determine the most significant features or inputs that contribute to predicting market trends [26] and making investment decisions [27] thereby mitigating risks [28] associated with financial investments.

Until now, no research has specifically focused on identifying and extracting causative factors from tweet texts, especially in scenarios where public sentiment strongly influences investment choices. This exploration could offer valuable guidance for entrepreneurs and investors. In addressing this intricate challenge, we present an explainable cause analysis method developed upon a dynamic attention mechanism. Our approach is configured with a multitasking strategy, encompassing both word-level and sentence-level analyses. Additionally, we introduce a specialized dataset focused on financial market behaviour, offering comprehensive and explicable data. This dataset, detailed in Table 1 encapsulates the specifics of the FinEMA dataset, showcasing nuanced information concerning market behaviours and trends.

III. CORPUS GENERATION

Lately, social media has become increasingly important as a crucial information repository utilized by various financial institutions. It serves as an easily accessible platform where individuals openly share their sentiments, opinions, reactions, and experiences regarding a wide range of financial assets, such as stocks, cryptocurrencies, and commodities. Despite the abundant availability of data sources, there is a limited number of publicly accessible explainable financial market corpora specifically designed for predicting the explainable feelings of financial investments. This scarcity is evident in Table 2. Given the uniqueness and novelty of our work, we have developed a market-generalized dataset specifically designed to train and evaluate our model. This dataset will facilitate an examination of the influence of social media on digital currency prices, financial stock markets, and investments. It will offer valuable insights into forecasting market behaviour influenced by public emotional and sentiment-driven factors.

In this research, the dataset formulation was conducted in three phases. The initial phase involved collecting existing datasets containing review texts with sentiment class labels. Subsequently, in the second phase, we gathered tweet reviews from Twitter and performed manual annotation for sentiment, emotion, and cause extraction tasks in the last phase.

TABLE 2. **Comparative study of existing Cryptocurrencies and *FinEMA* dataset.**

Dataset	Labels		
	Emotion	Sentiment	Cause
Cryptocurrency-review [29]	✓	×	×
Kaggle Cryptocurrency dataset [30]	✓	✓	×
TweetEval benchmark [31]	✓	✓	×
FinEMA(proposed dataset)	✓	✓	✓

A. PHASE-1: EXISTING DATASETS FOR SENTIMENT ANALYSIS

Initially, we targeted three primary domains to procure textual data: news articles, tweets related to the stock market, and content from the financial banking sector. Eventually, we acquired three human-annotated datasets comprising texts and associated sentiments, which were utilized as secondary data sources. These three datasets include:

1) SentFiN 1.0: ENTITY-AWARE SENTIMENT ANALYSIS FOR FINANCIAL NEWS: [32]

This human-annotated dataset contains 10,753 news headlines with entity-sentiment annotations, of which 2,847 headlines contain multiple entities, often with conflicting sentiments. For our work purpose, we have dropped the conflicting sentiments.

2) STOCK MARKET TWEETS DATA -SENTIMENT ANALYSIS: [33]

This dataset contains 1,300 stock market-related tweets, which were manually annotated and reviewed.

3) FinancialPhraseBank-v1.0: [34]

This dataset is the Polar sentiment dataset of sentences from financial news. The dataset consists of 4 text files with around 5000 human-annotated sentences categorized according to sentiment. The dataset annotation task was mutually distributed among 5-8 annotators.

However, the aforementioned datasets did not include current financial market reviews. Recognizing that the financial market exhibits temporal dependencies, we collected additional recent tweets through web scraping. By incorporating these recent tweets into our dataset, we aim to capture the dynamic nature of the financial market and enhance the timeliness of our analysis.

B. PHASE-2: SCRAPPED DATA FROM TWITTER

Despite broad data access, social media comes with its own challenges where noise, sarcasm, and context dependency in social media posts make the analysis process troublesome. Sanitizing the raw data and processing them qualitywise to create a generalised corpus is required for efficient model training and prediction performance. In this phase, we used the Tweepy library and a Twitter developer account to collect more than 26,000 tweets related to bonds, cryptocurrency, the stock market, and mutual funds over three months and created the manual dataset. Standard preprocessing techniques were used to omit unnecessary information such as conflicting statements, code mixed reviews and unformed sentences. After precise filtering, we considered 6500 valid tweets from our manual dataset.

Ultimately, we merged the 5072 tweets obtained from the existing datasets in phase-1 with the scraped tweets from phase-2, constituting our *FinEMA* dataset in English. Each instance was meticulously annotated with appropriate domain knowledge under the supervision of financial experts.[2]

C. ADDITIONAL CONSIDERATIONS

The text classification labelling using pre-trained models depends upon the sequence of words in a particular text. After analyzing our corpus, we found that some commonly repetitive words are present in different classes of sentences. For example, words like ‘stock’, ‘future’, ‘demand’, ‘buy’, and ‘may’ are common in all three (sentiment) classes of texts. Words like ‘profit’, ‘quality’, ‘jump’, ‘well’ etc. commonly appeared in both ‘neutral’ and ‘positive’ sentiment classes of texts.

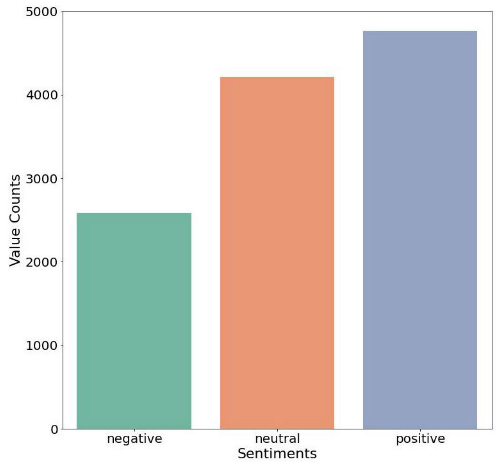

FIGURE 2. **Sentiment class label distributions.**

Words like ‘contract’, ‘sell’, etc., have commonly appeared in both ‘neutral’ and ‘negative’’ sentiment classes of texts. However, ‘performance’, ‘expect’, ‘quality’ etc. are commonly appear in the ‘relief’, ‘happy’, and ‘confused’ emotion classes of texts. An overlapping between emotion and sentiment classes formed due to words like ‘contact’, ‘may’, ‘expect’, etc., commonly appeared in the ‘confused’, ‘happy’, and ‘disgust’ emotion class label texts.

Sentiment class distribution is illustrated in Figure 2, while Figure 3 presents the distribution of emotion classes.

D. PHASE-3: ANNOTATIONS

In the final phase, three annotators were involved, including an individual pursuing a doctoral degree and two undergraduates specializing in computer science with robust linguistic

[2]Code&Dataset: https://github.com/sarmistha-D/Financial-Market-with-ESAN

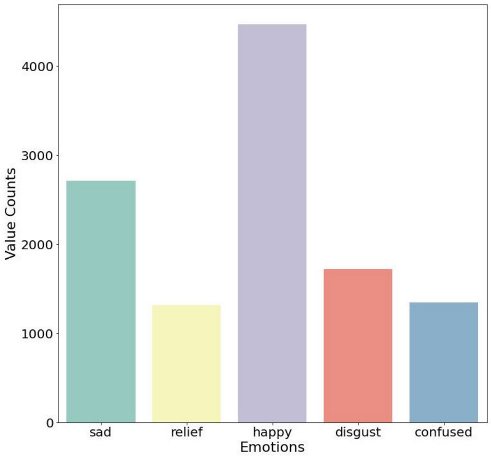

FIGURE 3. Emotion class label distributions for sad, relief, happy, disgust, and confused features.

foundations for this task. Drawing upon their significant domain expertise, we constructed the corpora and meticulously annotated causal spans for each instance in the dataset.

Initially, we provided some financial investment-related tweet instances with precise labels and descriptions for better understanding. To avoid annotation bias, each annotator was instructed to annotate data samples in isolation. At first, they identified the causal span,$X(i)$, that further represented the ground of market behaviour for each instance (i) in the *FinEMA* dataset. After manual annotation, as per the previous studies on span extraction, we employ the macro-F1 [35] metric to assess inter-rater agreement. Eventually, we obtained 0.81 F1 score for decent-quality annotation.[3] The explainable price trend dataset used in this work consists of the tweet text, sentiment and emotion label, and the corresponding annotated causal span for each record. Table 1 shows a few example instances of causal span annotation.

IV. METHODOLOGY

This section describes preprocessing steps like feature extraction, embedding techniques used in our dataset, and the architecture and training process of the multitasking model. The Workflow of the proposed model is demonstrated in Figure4, and the overall framework is shown in Figure 5.

A. PROBLEM DEFINITION

We approach the task of financial market behaviour analysis using a multitasking architecture where sentiment and emotion classifications are jointly learned, followed by cause extraction task

We define the cause extraction task for any instance as follows: Given a text instance $X_t = \{x_0, x_1, \ldots, x_n\}$ with a sentiment label s_t and emotion label em_t, and determine the causal span ce_t of instance X_t. Here $X_t \in n$, and n is the number of text instances. $s_t \in S$ (sentiment labels), $em_t \in E$ (emotion labels) and $ce_t \in CE$ (cause extraction).

[3]The authors of the work mentioned the agreement ratings of emotion and sentiment tasks as 0.68 and 0.82, respectively, suggesting good annotations.

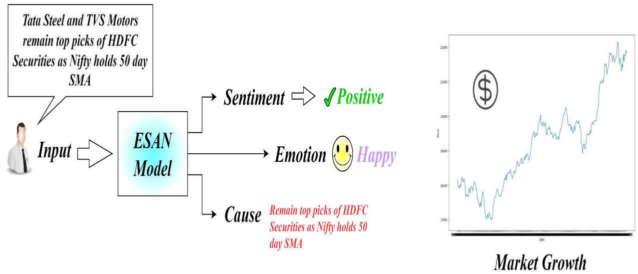

FIGURE 4. Workflow of the proposed model.

Our multitask model jointly learns to maximize prediction that links each instance of X_t with appropriate sentiment label s_t, emotion label em_t, and cause extraction label ce_t.

$$argmax_{\Theta}\left(\prod_{t=1}^{n} P(s_t, em_t, ce_t | X_t, \theta)\right) \tag{1}$$

Here the model wants to optimize with parameter θ.

B. PREPROCESSING

Preprocessing is the sanitization of raw data that enhances the efficiency of machine learning and natural language processing algorithms. The techniques used here are described as follows:

1) NUMBER REMOVAL

Numbers in the text don't offer valuable data for training models, and removing them shrinks the feature space. By removing the numbers, the complexity of the model is reduced while efficiency is increased. As a result, machine learning algorithms perform better.

2) SPECIAL SYMBOL AND PUNCTUATION REMOVAL

Special symbols and punctuation are sets of special characters often used in phrases and remarks for increasing readability and clarity. However, it hinders machine learning algorithms' capacity for learning and must be eliminated to improve the learning process. Since Python's string library only contains a few special characters, a custom dictionary of all the symbols and emojis have been created. By implementing Pyhton's ReGex package, we removed these unnecessary characters from the texts.

3) STOP WORDS REMOVAL

For machine learning models, stop words do not carry further significance in a phrase. Stop words can therefore be eliminated without changing the meaning of a statement. Eliminating stop-words reduces the input's complexity while improving the model's performance. We removed them using the pre-defined NLTK English stopwords corpus.

4) SPELL CORRECTION

Spelling correction is the process of changing incorrectly spelt words. The spelling checker is used in this stage to look for misspelt words and swap them out for the appropriate

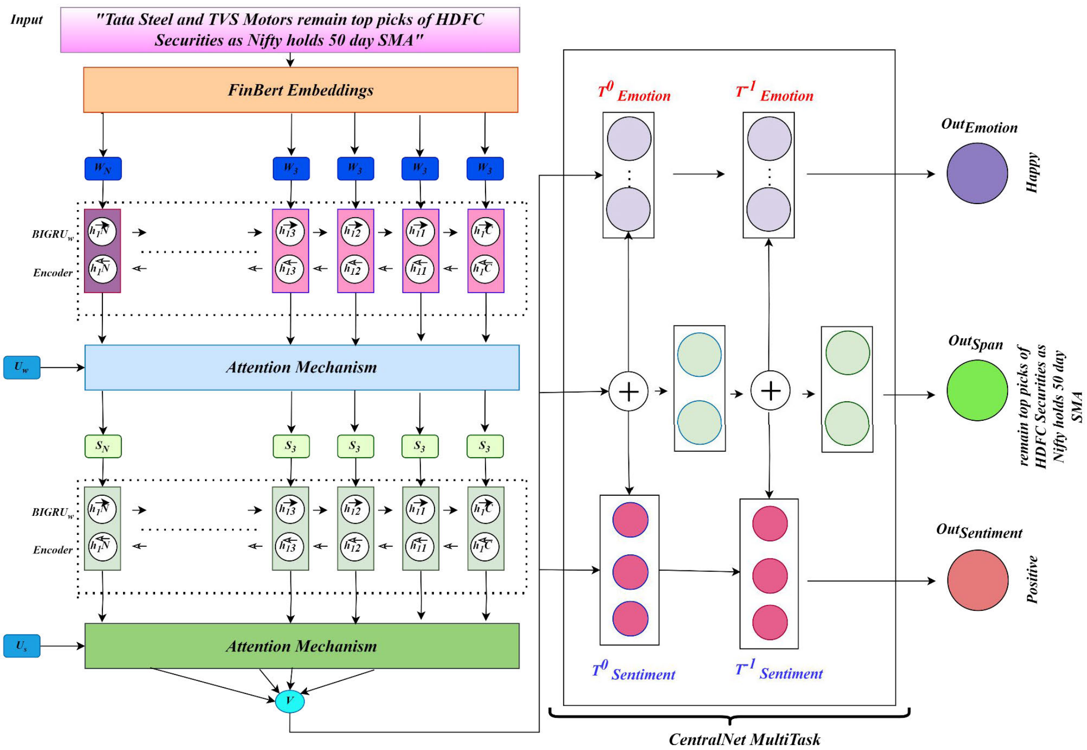

FIGURE 5. The architectural representation on W: word-level and S: sentence-level of proposed hierarchical attention multitask framework for simultaneous detection of sentiment, emotion and span.

ones. We have utilized the Python library TextBlob, which provides the crucial capability for identifying misspelt words.

5) LEMMATIZATION

The base or dictionary form of a word is known as the lemma. The algorithmic process of lemmatization in NLTK determines a word's lemma based on its meaning and context [36]. The NLTK lemmatization method is based on WorldNet's built-in "WordNetLemmatizer" function [37]. Lemmatizer reduces ambiguity in texts. For example, the root word is "like", and it's variations are "likes", "liked", "likely", "liking," and so on. The "WordNetLemmatizer" function [37] converts back the variations to their root words.

C. EMBEDDING GENERATION WITH FinBert

BERT is a multi-layer bidirectional transformer encoder [29] based on the primary work described in [38]. It is trained on masked language modelling and leverages masking input word tokens arbitrarily to predict certain contextual information. In 2019, FinBert was proposed for tackling NLP tasks in the finance domain, especially for social media analysis tasks [39]. FinBert is an extension of BERT [40], which enhances model performances and is trained on enormous financial data with hyperparameters, demonstrating financial text mining. In this work, we use the *FinBert-base-uncased* model that generates 768 embedding for the words within each financial tweet instance.

D. EMOTION-SENTIMENT ATTENTION NETWORK (ESAN)

We enhanced our model configuration to a hierarchical form to represent context-rich data samples. We used a sentence-level encoder to compute sentence representation and consider a tweet texts collection as our input, $T_x = [ST_1, ST_2, ST_3, \ldots, ST_n]$ from the source data. For each given sentence, ST_i is made up of word tokens $[w_{l-k+1}, w_{l-k+2}, w_{l-k+3}, \ldots, w_l]$. Each word token of the entire preprocessed tweet text is considered in the sequential input for Bi-GRU, the word-level encoder. Let's assume for each time stamp, H_t^{wl} is the output of input word token wl_t (here sl and wl are superscripted to indicate sentence-level and world-level) respectively. Since the semantic contributions of all words are not equal to the texts, word-level attention (WAtt) has been introduced. To form a sentence vector, we aggregate the word-level attention representation as follows:

$$H_{sT_i}^{wl} = \sum_{t=1}^{k} a_{l-k+t}^{wl} H_{l-k+t}^{wl} \tag{2}$$

Thus, H_{Tx}^{wl} became the word-level sentence embedding of the instance and obtained as $H_{Tx}^{wl} = [H_{sT_1}^{wl}, H_{sT_2}^{wl}, H_{sT_3}^{wl}, \ldots, H_{sT_n}^{wl}]$. Later this word-level embedding is fed into Bi-GRU as a sentence-level encoder. Afterwards, the sentence-level embedding is obtained as, $H_{Tx}^{sl} = [H_{sT_1}^{sl}, H_{sT_2}^{sl}, H_{sT_3}^{sl}, \ldots, H_{sT_n}^{sl}]$ and aggregated with attention. This computed sentence-level tweet representation is defined as follows:

$$H_{sT}^{Tx} = \sum_{t=1}^{n} a_{sT_t}^{sl} H_{sT_t}^{sl} \quad (3)$$

To investigate the cause extraction performance further, we fed the final aggregated tweet representation into the multitask module, where sentiment categorization and emotion recognition are the auxiliary tasks to enhance the performance of the main task (cause extraction).

E. MULTITASK FRAMEWORK

As per the proposed framework's prerequisite, we modified the CentralNet [41] framework into a multitasking model. CentralNet is an advanced centralised version of a neural network with one central core network where each task-specific network is divided into n separate networks described in Figure 6. For performing multi-tasking tasks like ER(emotion recognition), SC(sentiment categorization) and CE(cause extraction), CentralNet is the standard layer, where emotion and sentiment are task-specific layers. CentralNet performs the multilayer tasks by analysing the weighted summation of each task-specific network along with their preceding layers and aggregates the features obtained from different studies. The multitask layers can be defined as follows:

$$ML_{i+1} = \alpha n ML_i + \sum_{k=1}^{n} \alpha s_i^k ST_i^k \quad (4)$$

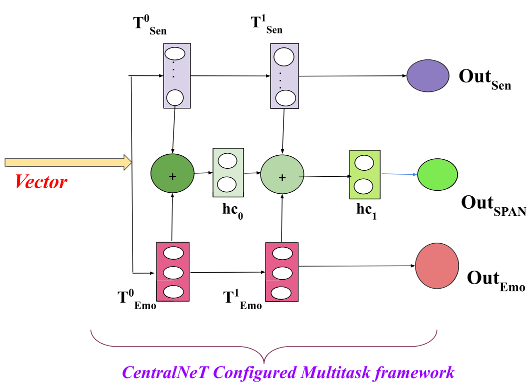

FIGURE 6. Brief overview on Layerwise Multi-tasking framework.

Here n denotes the number of task-specific networks and α is scalar trainable weights for n task-specific networks. ST_i^k is the hidden representation of k^{th} task-specific network at i^{th} layer, and the hidden central representation of the main task is denoted by ML_i and the final layer ML_{i+1} is fed to an operating layer (dense layer accompanied by an activation function).

F. ESAN + CentralNet

We obtained features from FinBert and passed them to the ESAN module. Then ESAN model generated rich word-level and sentence-level attention encodings of the tweet instances. Based on their contextual aspect, encoded tweets were further passed to the emotion and sentiment-aided CentralNet multitasking module for cause extraction.

G. POST-PROCESSING FOR CAUSE EXTRACTION TASK

Based on the binary class label(0,1), post-processing of CE task was conducted. The proposed model would find the corresponding word token where $0 \implies no-cause$ and $1 \implies cause$. Similarly, $label \in \{0, 1\}$ will be mapped according to each tweet instance. From a given text tweet sample, token length is utilized to decode the beginning and end of the text. For instance:

Tata Steel/[0] and/[0] TVS/[0] Motors/[0] remain/[1] top/[1] picks/[1] of/[1] HDFC/[1] Securities/[1] as/[1] Nifty/[1] holds/[1] 50/[1] day/[0] SMA/[0]

H. LOSS FUNCTION

- For loss calculation of the emotion recognition(ER) and sentiment categorization (SC) tasks, we considered cross entropy(l_{CE}) as the loss function. The definition of the integrated loss function $J(\theta)$ is given below:

$$J(\theta) = l^{ER}{}_{CE}(\theta) + l^{SC}{}_{CE}(\theta) \quad (5)$$

  where θ is denoted as optimized model parameters.
- In the cause extraction evaluation task, where each instance may belong to multiple classes, the Sparse Categorical Cross-Entropy loss (l_{SCE}) is used. This loss function allows the model to determine whether a given instance solely belongs to a specific class or not. Therefore, the loss of the cause extraction task is $J_{SCE}(\theta) = l_{SCE}$.
- The merged loss function (L) of the ESAN model is defined as the sum of the primary loss function $J(\theta)$ and the cause extraction loss J_{SCE}:

$$L = J(\theta) + J_{SCE}(\theta) \quad (6)$$

By minimizing the merged loss L during training, the model aims to optimize both the primary classification objectives and the cause extraction task simultaneously.

V. EXPERIMENTAL RESULTS & ANALYSIS

This section illustrates the proposed approach's performance, efficiency, and results on *FinEMA* dataset for sentiment categorization, emotion recognition, and cause extraction.

A. METRICS

To evaluate our model's performance, we have utilized several metrics such as accuracy, precision, recall, F1 score, training loss, validation loss, and total model loss.

B. INTRODUCTION OF THE BASELINES AND STATE OF THE ART (SOTA)

We have set different baselines for fair comparisons with our proposed model. The following baselines were executed on our dataset.

TABLE 3. Comparison studies of the ER, SC and CE tasks in terms of macro-F1 score (F1) and Accuracy (A) values.JS: Jaccard Similarity, HD: Hamming Distance, and ROS: Ratcliff-Obershelp Similarity predictions by the Baselines and the proposed ESAN model. All of the metrics values are in %, and we highlighted the maximum achieved scores in bold.

Model	Sentiment Analysis				Emotion Detection				Span Extraction		
	P	R	F1	A	P	R	F1	A	ROS	JS	HD
BERT$_{Baseline_1}$	0.82	0.80	0.80	0.81	0.68	0.60	0.62	0.66	-	-	-
CNN+BiLSTM$_{Baseline_2}$	0.76	0.74	0.76	0.76	0.59	0.52	0.56	0.57	-	-	-
Ensamble LSTM-GRU$_{Baseline_3}$	0.78	0.74	0.75	0.77	0.59	0.55	0.58	0.56	-	-	-
SOTA	0.79	0.78	0.77	0.78	0.65	0.60	0.62	0.62	0.66	0.59	0.53
ESAN$_{\text{Sent}}$	0.82	0.84	0.83	0.84	-	-	-	-	-	-	-
ESAN$_{\text{Emo}}$	-	-	-	-	0.71	0.73	0.68	0.72	-	-	-
ESAN$_{\text{Emo+Span}}$	-	-	-	-	0.72	0.72	0.74	0.73	0.64	0.67	0.62
ESAN$_{\text{Sent+Span}}$	0.85	0.85	0.84	0.87	-	-	-	-	0.79	0.74	0.69
ESAN$_{\text{All}}$	**0.91**	**0.89**	**0.86**	**0.89**	**0.80**	**0.79**	**0.81**	**0.79**	**0.82**	**0.73**	**0.70**
SpanBERT$_{\text{Cause-Baseline}}$	-	-	-	-	-	-	-	-	0.83	0.74	0.71

- *Baseline*$_1$**-BERT:** We considered a transformer model BERT [29] as our first baseline model that stands for Bidirectional Encoder Representations from Transformers and is most suitable for texts. It is a deep learning-based framework that extracts contextual semantic meanings from the text by considering surrounding texts. We reimplemented *Baseline*$_1$ on our dataset and applied a multiclass classification head on the top of the BERT model to perform emotion recognition and sentiment categorization tasks jointly.
- *Baseline*$_2$**-CNN+BiLSTM:** Inspired by multi-tasking approaches, we implemented CNN-BiLSTM [42] Attention model for the multitasking text classification task since the essential semantic information is extracted from convolution and pooling operation on the text. Subsequently, the contextual information is extracted by BiLSTM framework. After that, attention layers assign more weights to those words, which are important for the classification.
- *Baseline*$_3$**-Ensembled LSTM-GRU:** Ensemble techniques enhance the prediction model's optimality with multiple model considerations. As our proposed framework is based on multi-tasking on sentiment categorization(SC) and emotion recognition(ER), we consider [25] as our third baseline for SC and ER. LSTM and GRU are both suitable choices for text data, though GRU is an advanced version of recurrent neural network(RNN).
- **Baseline for Cause Extraction Task:** In finance domain, cause extraction task is novel and popular, especially in financial complaint use cases. Cause extraction helps to strain insightful information among ambiguous data. We took our motivation from Poria et al. [12] emotion recognition work which was trained on SpanBERT base pre-trained model. SpanBERT was fine-tuned on SQuAD 2.0 [35] dataset and used explicitly for the CE task.
- **SOTA [43]:** To substantiate our research claim concerning sentiment and emotion analysis via a generative task, we utilized an alternative sequence-to-sequence generative method, incorporating a financial complaint detection model named ConCS2I, which is informed by common sense. For this investigation, we adapted the model to our research requirements, emphasizing the identification of sentiment and emotion, subsequently proceeding to cause generation.
- **Ablation Model:** We implemented our proposed ESAN model as a hierarchical multi-tasking model. Additionally, we develop multiple ablation models to understand the impact of jointly learnt emotion recognition and sentiment categorization tasks with cause extraction. Initially, we kept emotion recognition (ER) and sentiment categorization (SC) as individual tasks and evaluated model performance accordingly. Later, we add a cause extraction task. Finally, we develop multiple ablation models by considering combinations of each (SC and ER) task, where each combination must include the CE task. The various combinational models have been explicitly explained in Table 3, and they share similarities with the *ESAN* framework in other aspects.

C. EXPERIMENTAL SETUP

In the proposed model, we considered a few parameters such as test data size, dropout, optimizer, learning rate, epochs, and GPU for training that are requisites of our experimental setup, and their corresponding values are mentioned in the given Table 4.

We implemented and compared our model against baseline models using Python 3 programming language with PyTorch 4, Tensorflow 5, and Scikit-learn 6 libraries. The experiments were performed on an RTX2080Ti GPU with 200GB storage. Consistent training configurations were applied across all experiments. The training loop involved iterating through data batches from the data loader, computing a combined loss (emotion and sentiment), backpropagating gradients, and updating model parameters using an optimizer. We calculated accuracy metrics for both tasks during training. Model evaluation occurred using a separate validation dataset without updating parameters. We adopted a strategy of minimizing validation loss over 20 training epochs to select the optimal model. The input sequence was fixed at a maximum length of 25. The training utilized a batch size

TABLE 4. **Experimental setup of emotion sentiment attention-based network model.**

Parameters	Values
Test Data Size	2600
Dropout	0.2
Optimizer	Adam
Learning Rate	0.005
Batch size	32
Activation	ReLU
Max input Sequence	20
Epochs	25
GPU for Training	RTX2080Ti

of 32, employing the Adam optimization algorithm with a learning rate (lr) set to 5e-5. To mitigate overfitting, ReLU activation function was used, and dropout regularization was applied at a rate of 40%. The dropout rate was selected from {20%, 30%, 50%} after dense layers using grid search. For emotion recognition and sentiment analysis tasks, output layers employed a softmax activation function with 5 and 3 neurons, respectively. In the cause extraction task, the output layer featured 20 neurons with a Sigmoid activation function. Our baseline model, FinBERT, was sourced from the Hugging Face Transformer open-source repository.

D. ABLATION STUDIES

This study aims to enhance the outcomes of emotion recognition (ER) and sentiment categorization (SC) tasks by incorporating joint learning with the cause extraction (CE) task. Consequently, the results and analysis are presented with a focus on CE as a primary task in all task combinations. This work introduces unique versions of the proposed architecture, *ESAN* multitasks model. The primary motivation behind this is that not all portions of a text are equally significant for addressing a problem. Identifying which portions are relevant requires modelling the associations of the words instead of their existence in isolation. Table 3 depicts the classification results from the various experiments for the ER, SC, and CE tasks.

As can be observed, the proposed model ESAN_{All} outperforms all the other baselines for ER and SC tasks. The results signify that the proposed model is able to capture contextual information. It's noteworthy that when focusing solely on the emotion classification task, the F1 score tends to be lower compared to when performing sentiment classification alone. This discrepancy can be attributed to the fact that emotion class labels are heavily reliant on context. Despite attempts to enhance performance by including causal spans, the improvements remain marginal, with an increase of merely one percentage point in accuracy. Moreover, $\text{ESAN}_{\text{SC+CE}}$ outperforms $\text{ESAN}_{\text{ER+CE}}$ model. It can be entailed by the fact that emotion recognition itself, a fine-grained task, is highly context-dependent, and compared to sentiment, it contributes less to the multitask model alone. For the quantitative assessment of the CE task, we used several evaluation metrics such as Jaccard Similarity (JS),

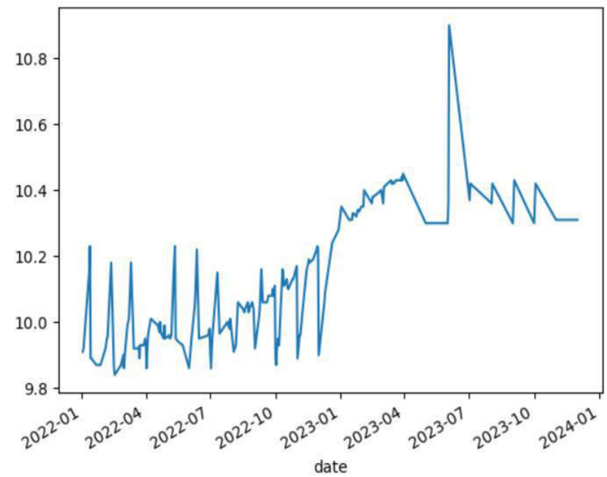

FIGURE 7. **Resultant stock price predictions.**

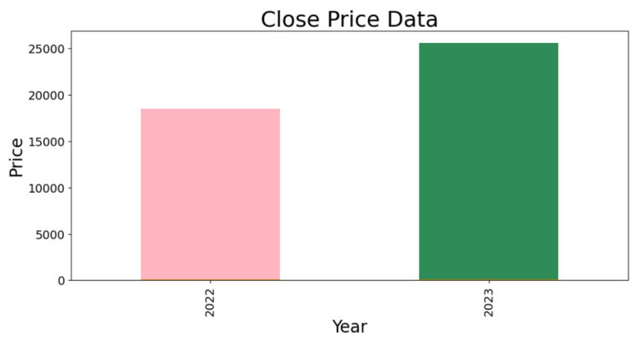

FIGURE 8. **Analysis of daily bitcoin data using LSTM model: A validation of positive increment in cryptocurrency closing values of 2022 and 2023.**

Hamming Distance (HD), and Ratcliff-Obershelp Similarity (ROS) metrics as mentioned in [44]. All are based on token distance, edit distance, and sequence distance metrics. The ROS metric, which relies on the most extended sequence matching approach, is appropriate for the training objective of the CE task and provides a more accurate indication of the framework's effectiveness. It can be observed from Table 3, with an accuracy of 89%, the ESAN_{All} model exceeds all the other models for the cause extraction task.

E. MULTILABEL CLASSIFICATION VIEW

To support our predictions, we evaluate our model's performance across various class-specific predictions, as outlined in Table 5. It is evident that the model demonstrates robust performance in the positive class, achieving an F1 score of 0.89, and for the negative class, it achieves an F1 score of 0.84. However, due to the inherent challenges in comprehending sarcasm and neutral sentiments, the model's performance confidence decreases in these cases. Figure 9 and Figure 10 demonstrate the significant performance of sentiment and emotion class labels. Table 6 represents the performance of the proposed multitasking model ESAN for individual class levels of the emotion recognition task. Our model achieved

TABLE 5. **Quantative evaluation of the proposed model on different classes for the sentiment categorization task.**

Sentiment				
Class Labels	**Accuracy**	**Precision**	**Recall**	**F1-score**
Negative	0.85	0.82	0.86	0.84
Neutral	0.77	0.76	0.78	0.77
Positive	**0.89**	**0.88**	**0.90**	**0.89**

TABLE 6. **Performance evaluation of the proposed multi-tasking model for individual class-levels of the emotion recognition task.**

Emotions				
Class Labels	**Accuracy**	**Precision**	**Recall**	**F1-score**
Sad	0.76	0.74	0.77	0.75
Relief	0.82	0.80	0.83	0.81
Happy	**0.89**	**0.87**	**0.90**	**0.88**
Disgust	0.80	0.78	0.81	0.79
Confused	0.75	0.73	0.76	0.74

a high F1 score of 0.89 for predicting happy class labels. Additionally, the performances of the model for the relief and disgust classes are relatively consistent. However, the contextual dependency of the confused and sad class labels results in lower accuracy compared to the other emotional class labels.

F. REAL-TIME DATA SALIENCE

To demonstrate the relevance of our model's performance in real-time financial market scenarios, we conducted an analysis of stock prices using historical data. From Table 6, it was evident that the prediction performance of the happy class label has 89% accuracy value, which entails people having positive sentiment towards financial investments. To validate our research claim, we obtained the NASDAQ daily stock prices [45] dataset from Kaggle, which covers the period from January 3rd 2022, to March 30th 2023. We applied the ARIMA Model [46] to predict stock prices for the year 2023. The plot in Figure 7 showed the relationship between data and the highest stock value. The graph obtained from the ARIMA model predicts a substantial increase in stock prices for the year 2023, leading to a positive outlook for investors and indicating a favourable trend in the stock and cryptocurrency market, hence validating the confidence of our model's prediction. In this study, we utilized daily Bitcoin data sourced from[4] covering the period from 01/09/2022 to 07/07/2023 and applied an LSTM model as our chosen machine learning approach. The thorough examination of the data presented in Figure8 consistently shows a positive increment in cryptocurrency closing values, effectively validating the clarity of our model's predictions. *All of the results are statistically significant*[5] Moreover, Figure 11 shows correlations between different cryptocurrency variables over time.

[4]https://in.investing.com/crypto/bitcoin/historical-data

[5]For the significance test, we used the Student's t-test (p-value < 0.04).

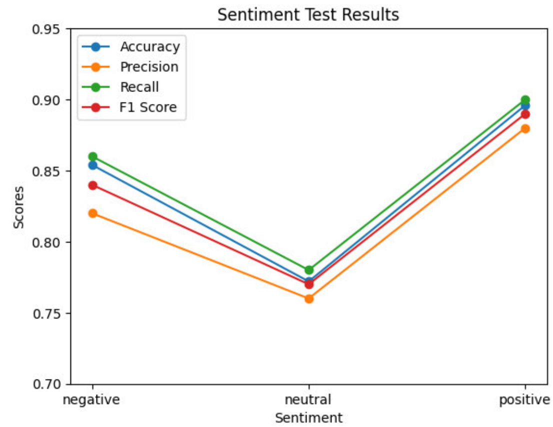

FIGURE 9. **Analysis of sentiment performance: Effectiveness evaluation of the proposed ESAN model for sentiment categorization on test data (Highlighted colors indicate performance variations in evaluation metrics).**

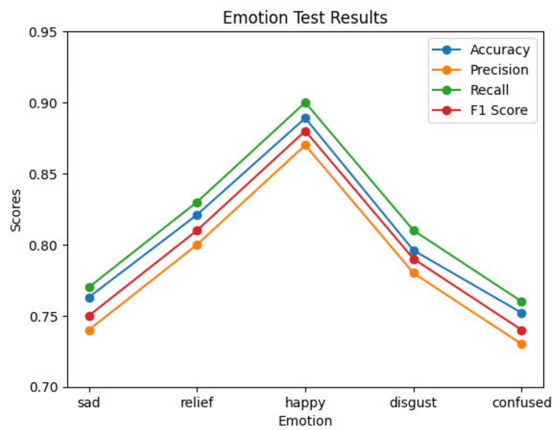

FIGURE 10. **Performance evaluation of the proposed multitasking ESAN model for emotion recognition task (Highlighted colors indicate performance variations in evaluation metrics).**

G. QUALITATIVE ANALYSIS

Tweets including firm price trend indications, such as expressions of currency, or emotions-related words, are less misclassified. Table 7 explains the qualitative study of the cryptocurrency price trend predictions on a few tweet samples from our dataset. Prediction results are obtained by the $Baseline_3$ and ESAN multitask framework. Previously in Table 3, the sentiment categorization (SC) task and emotion recognition (ER) task were jointly learnt, and with each other's shared features, the combined ER and SC tasks improved the prediction performance and outperformed $Baselin_3$. The probable reason for such improvement was to filtrate the unnecessary data by cause-extraction module, which enhanced emotion label predictions. Table 7 also shows both $Baseline_3$ and ESAN multitask framework correctly predicted text sample instances as positive sentences; however, $Baseline_3$ framework failed to predict emotion class labels as per their respective actual label.

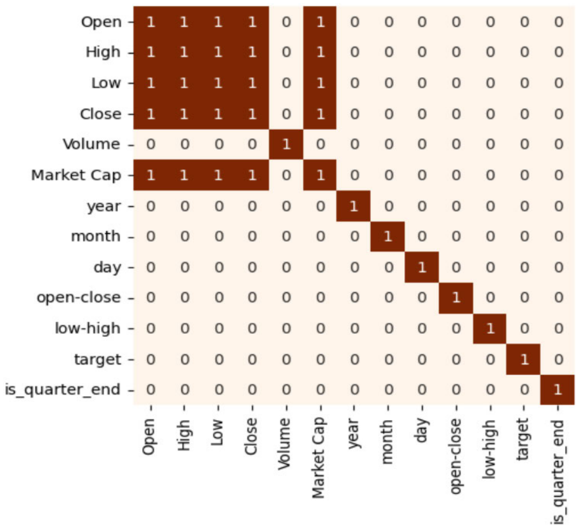

FIGURE 11. Generated heatmap on daily cryptocurrency variables.

TABLE 7. Example of qualitative study with the Baseline3 and the proposed ESAN multitasking model on emotion recognition(ER) and sentiment categorization(SC). Here 'Actual Label' indicates the true labels for ER and SC tasks; the causal span annotation is highlighted in yellow in the text of the sentence. (Best viewed in colour.)

Tweet text	$Baseline_3$	Proposed	Actual
Cryptocurrencies are gaining popularity all over the world and Australia is no exception.	Positive relief	Positive happy	Positive happy
The difference can be explained by the fact that two shipping companies have stopped operating	negative disgust	negative sad	negative sad

TABLE 8. Comparison of cause instance predicted by human annotators and the ESAN model. Predicted spans are highlighted in red for both the human annotator and the model prediction use case. The yellow highlight indicates relevant tokens to the model. (Best viewed in colour.)

Model	Text
Human Annotator	<USER>Ethereum mining alternative raven coin jumps ahead of the merge.
$ESAN_{All}$	<USER>Ethereum mining alternative raven coin jumps ahead of the merge.

In Table 8, we performed a qualitative analysis for the CE task where the relevant causal span that the human annotator picked is represented in the first row. The spans obtained by the proposed model and the human annotator are shown in the following two rows in Table 8.

H. ERROR ANALYSIS

We investigate the possible reasons for the proposed model's errors:

- **Disperse Causes**: The proposed framework cannot identify multiple causes spread across a tweet instance. For example, *DB realty may find bank loans drying up stock falls.* The causal span predicted: *loans drying up*, where the actual span of cause is *stock fails*. Often, causes are scattered across a complete tweet, and the casual annotations of this present work are based on the first encounter with a strong expression of impactful reason in the tweet. Our model is incapable of identifying dispersed causes in a tweet.
- **Misclassified Emotional Labels**: Often, neighbouring emotion class label instances got misclassified due to their semantic nature. For example, *Asian shares pressured by Wallstreet geopolitical tensions.* The actual emotional level for the given instance is sad, but the model predicts it as disgust.

VI. CONCLUSION

This study delves into the influence of sentiment and emotions on the financial market. To enhance understanding, we integrated an explanatory element reflecting factors guiding investors' decisions. We created the FinEMA dataset, featuring annotated causes, to facilitate research on comprehensible financial market behaviour identification. Our ESAN model, designed with hierarchical attention-based multitasking, concurrently handles sentiment, emotion identification, and causal extraction. Through varied experiments, $ESAN_{All}$ demonstrated 89% accuracy, surpassing single-task and baseline models. Results suggest potential stock market improvements in late 2023, albeit influenced by emotions like sadness or disgust. Moreover, discerning financial market behaviour transcends a mere classification task; the algorithm's role and prediction confidence demands careful consideration. Our aim is to uncover investment behaviour through public psychology, aiming to mitigate investment-related disappointment.

VII. LIMITATIONS

The multitasking model proposed in this study exhibits certain limitations when it comes to capturing comprehensive textual understanding.

- Challenges in Navigating Complexity:** The model faces hurdles in processing complex financial jargon and nuanced contextual situations. For instance, terms like "Arbitrage," referring to profiting from price differences across markets or assets, and "Leverage," involving using borrowed funds to amplify returns, pose challenges due to their intricate implications and risks. These complexities could potentially limit the model's accuracy in capturing the entire spectrum of sentiments and emotions.
- Causal Span Identification Limitations:** The current method relies on pinpointing causal factors primarily from explicit statements in the initial segments of financial reviews. This approach might overlook crucial causes scattered throughout the entirety of a tweet, reducing the model's capacity to comprehend the financial context fully.

Addressing Limitations Through Future Research: Acknowledging these constraints, there's a need for further exploration and refinement to augment the model's capabilities. Enhancements should focus on accommodating the diverse and intricate language patterns inherent in the analysis of financial market behaviour.

ACKNOWLEDGMENT

The authors express their gratitude for the invaluable support provided by CRISIL and the multiple dataset annotators for their tireless dedication throughout the project. Furthermore, special acknowledgment is extended to Mohit Kumar Wankhade for his invaluable assistance in collecting time series data.

REFERENCES

[1] S. Zakaria, S. M. A. Manaf, M. T. Amron, and M. T. M. Suffian, "Has the world of finance changed? A review of the influence of artificial intelligence on financial management studies," *Inf. Manage. Bus. Rev.*, vol. 15, no. 4, pp. 420–432, Nov. 2023.

[2] J. Almeida and T. C. Gonçalves, "A systematic literature review of investor behavior in the cryptocurrency markets," *J. Behav. Exp. Finance*, vol. 37, Mar. 2023, Art. no. 100785.

[3] D. Karthikeyan, B. A. Sivamani, P. K. Tummala, and C. Arumugam, "Time series for forecasting stock market prices based on sentiment analysis of social media," in *Proc. 21st Int. Conf. Comput. Sci. Appl. (ICCSA)*, Cagliari, Italy. Cham, Switzerland: Springer, Sep. 2021, pp. 353–367.

[4] C. Oh and O. Sheng, "Investigating predictive power of stock micro blog sentiment in forecasting future stock price directional movement," in *Proc. ICIS*, 2011, p. 17.

[5] P. G. Preethi, V. Uma, and A. Kumar, "Temporal sentiment analysis and causal rules extraction from tweets for event prediction," *Proc. Comput. Sci.*, vol. 48, pp. 84–89, Jan. 2015.

[6] A. Inamdar, A. Bhagtani, S. Bhatt, and P. M. Shetty, "Predicting cryptocurrency value using sentiment analysis," in *Proc. Int. Conf. Intell. Comput. Control Syst. (ICCS)*, May 2019, pp. 932–934.

[7] J. Fior, L. Cagliero, and P. Garza, "Leveraging explainable AI to support cryptocurrency investors," *Future Internet*, vol. 14, no. 9, p. 251, Aug. 2022.

[8] K. Wołk, "Advanced social media sentiment analysis for short-term cryptocurrency price prediction," *Expert Syst.*, vol. 37, no. 2, Apr. 2020, Art. no. e12493.

[9] S. Wu, O. Irsoy, S. Lu, V. Dabravolski, M. Dredze, S. Gehrmann, P. Kambadur, D. Rosenberg, and G. Mann, "BloombergGPT: A large language model for finance," 2023, *arXiv:2303.17564*.

[10] Q. Xie, W. Han, X. Zhang, Y. Lai, M. Peng, A. Lopez-Lira, and J. Huang, "PIXIU: A large language model, instruction data and evaluation benchmark for finance," 2023, *arXiv:2306.05443*.

[11] H. S. Narman and A. D. Uulu, "Impacts of positive and negative comments of social media users to cryptocurrency," in *Proc. Int. Conf. Comput., Netw. Commun. (ICNC)*, Feb. 2020, pp. 187–192.

[12] S. Poria, N. Majumder, D. Hazarika, D. Ghosal, R. Bhardwaj, S. Y. B. Jian, P. Hong, R. Ghosh, A. Roy, N. Chhaya, A. Gelbukh, and R. Mihalcea, "Recognizing emotion cause in conversations," *Cogn. Comput.*, vol. 13, no. 5, pp. 1317–1332, Sep. 2021.

[13] A. Rao and M. R. Hossain, "The future of finance: Artificial intelligence's influence on behavioral investment decisions," in *Leveraging AI and Emotional Intelligence in Contemporary Bus. Organizations*. Hershey, PA, USA: IGI Global, 2024, pp. 166–186.

[14] B. J. Vanstone, A. Gepp, and G. Harris, "Do news and sentiment play a role in stock price prediction?" *Appl. Intell.*, vol. 49, pp. 3815–3820, Apr. 2019.

[15] X. Li, P. Wu, and W. Wang, "Incorporating stock prices and news sentiments for stock market prediction: A case of Hong Kong," *Inf. Process. Manage.*, vol. 57, no. 5, Sep. 2020, Art. no. 102212.

[16] G. Ranibaran, M.-S. Moin, S. H. Alizadeh, and A. Koochari, "Analyzing effect of news polarity on stock market prediction: A machine learning approach," in *Proc. 12th Int. Conf. Inf. Knowl. Technol. (IKT)*, Dec. 2021, pp. 102–106.

[17] T. L. D. Huynh, "When Elon Musk changes his tone, does Bitcoin adjust its tune?" *Comput. Econ.*, vol. 62, no. 2, pp. 639–661, Aug. 2023.

[18] S. Rahman, J. N. Hemel, S. J. A. Anta, H. A. Muhee, and J. Uddin, "Sentiment analysis using R: An approach to correlate cryptocurrency price fluctuations with change in user sentiment using machine learning," in *Proc. Joint 7th Int. Conf. Informat., Electron. Vis. (ICIEV), 2nd Int. Conf. Imag., Vis. Pattern Recognit. (icIVPR)*, Jun. 2018, pp. 492–497.

[19] Y. Wang and R. Chen, "Cryptocurrency price prediction based on multiple market sentiment," in *Proc. Annu. Hawaii Int. Conf. Syst. Sci.*, 2020, pp. 1–9.

[20] X. Tang, M. Zhang, J. Surbiryala, V. Iosifidis, Z. Liu, X. Huang, W. Zhang, and J. Zhang, "LSTM based sentiment analysis for cryptocurrency prediction," Ph.D. thesis, 2021.

[21] S. M. Raju and A. M. Tarif, "Real-time prediction of BITCOIN price using machine learning techniques and public sentiment analysis," Ph.D. thesis, 2020.

[22] D. Valle-Cruz, V. Fernandez-Cortez, A. López-Chau, and R. Sandoval-Almazán, "Does Twitter affect stock market decisions? Financial sentiment analysis during pandemics: A comparative study of the H1N1 and the COVID-19 periods," *Cogn. Comput.*, vol. 14, no. 1, pp. 372–387, Jan. 2022.

[23] N. Aslam, F. Rustam, E. Lee, P. B. Washington, and I. Ashraf, "Sentiment analysis and emotion detection on cryptocurrency related tweets using ensemble LSTM-GRU model," *IEEE Access*, vol. 10, pp. 39313–39324, 2022.

[24] D. Haryadi and G. Putra, "Emotion detection in text using nested long short-term memory," *Int. J. Adv. Comput. Sci. Appl.*, vol. 10, no. 6, 2019.

[25] P. B. E. Lee, W. N. Aslam, F. Rustam, and I. Ashraf, "Sentiment analysis and emotion detection on cryptocurrency-related tweets using ensemble LSTM-GRU model," Ph.D. thesis, 2022.

[26] S. Gite, H. Khatavkar, S. Srivastava, P. Maheshwari, and N. Pandey, "Stock prices prediction from financial news articles using LSTM and XAI," in *Proc. 2nd Int. Conf. Comput., Commun., Cyber-Secur. (IC4S)*. Cham, Switzerland: Springer, 2020, pp. 153–161.

[27] J. J. Ohana, S. Ohana, E. Benhamou, D. Saltiel, and B. Guez, "Explainable AI (XAI) models applied to the multi-agent environment of financial markets," in *Proc. 3rd Int. Workshop Explainable Transparent AI Multi-Agent Syst. (EXTRAAMAS)*. Cham, Switzerland: Springer, May 2021, pp. 189–207.

[28] D. V. Kute, B. Pradhan, N. Shukla, and A. Alamri, "Deep learning and explainable artificial intelligence techniques applied for detecting money laundering—A critical review," *IEEE Access*, vol. 9, pp. 82300–82317, 2021.

[29] J. Devlin, M.-W. Chang, K. Lee, and K. Toutanova, "BERT: Pre-training of deep bidirectional transformers for language understanding," 2018, *arXiv:1810.04805*.

[30] *Reddit Cryptocurrency Data for August 2021*, Lexyr, San Jose, CA, USA, Sep. 2021.

[31] F. Barbieri, J. Camacho-Collados, L. Neves, and L. Espinosa-Anke, "TweetEval: Unified benchmark and comparative evaluation for tweet classification," 2020, *arXiv:2010.12421*.

[32] A. Sinha, S. Kedas, R. Kumar, and P. Malo, "SEntFiN 1.0: Entity-aware sentiment analysis for financial news," *J. Assoc. Inf. Sci. Technol.*, vol. 73, no. 9, pp. 1314–1335, Sep. 2022.

[33] R. R. B. Taborda and J. Carlos, "Stock market tweets data-sentiment analysis," Dept. Center Inform. Syst., Univ. Coimbra, Coimbra, Portugal, Tech. Rep.

[34] P. Malo, A. Sinha, P. Korhonen, J. Wallenius, and P. Takala, "Good debt or bad debt: Detecting semantic orientations in economic texts," *J. Assoc. Inf. Sci. Technol.*, vol. 65, no. 4, pp. 782–796, Apr. 2014.

[35] P. Rajpurkar, J. Zhang, K. Lopyrev, and P. Liang, "SQuAD: 100,000+ questions for machine comprehension of text," 2016, *arXiv:1606.05250*.

[36] D. Johnson. (2022). *Stemming and Lemmatization in Python NLTK With Examples*. [Online]. Available: https://www.guru99.com/stemming-lemmatization-python-nltk.html

[37] NLTK Project. (2022). *NLTK Wordnet Documentation*. [Online]. Available: https://www.nltk.org/_modules/nltk/stem/wordnet.html

[38] A. Vaswani, N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A. N. Gomez, Ł. Kaiser, and I. Polosukhin, "Attention is all you need," in *Proc. Adv. Neural Inf. Process. Syst.*, vol. 30, 2017.

[39] D. Araci, "FinBERT: Financial sentiment analysis with pre-trained language models," 2019, *arXiv:1908.10063*.

[40] Z. Liu, D. Huang, K. Huang, Z. Li, and J. Zhao, "FinBERT: A pre-trained financial language representation model for financial text mining," in *Proc. 29th Int. Joint Conf. Artif. Intell.*, Jul. 2020, pp. 4513–4519.

[41] V. Vielzeuf, A. Lechervy, S. Pateux, and F. Jurie, "CentralNet: A multilayer approach for multimodal fusion," in *Proc. Eur. Conf. Comput. Vis. (ECCV) Workshops*, 2018.

[42] G. Liu and J. Guo, "Bidirectional LSTM with attention mechanism and convolutional layer for text classification," *Neurocomputing*, vol. 337, pp. 325–338, Apr. 2019.

[43] S. Das, A. Singh, R. Jain, S. Saha, and A. Maurya, "Let the model make financial senses: A Text2Text generative approach for financial complaint identification," in *Proc. Pacific–Asia Conf. Knowl. Discovery Data Mining*. Cham, Switzerland: Springer, 2023, pp. 58–69.
[44] S. Ghosh, S. Roy, A. Ekbal, and P. Bhattacharyya, "Cares: Cause recognition for emotion in suicide notes," in *Proc. Eur. Conf. Inf. Retr.* Cham, Switzerland: Springer, 2022, pp. 128–136.
[45] SvaningeIgem. (2021). *Nasdaq Daily Stock Prices Dataset*. [Online]. Available: https://www.kaggle.com/datasets/svaningelgem/nasdaq-daily-stock-prices?select=ACBA.csv
[46] A. A. Ariyo, A. O. Adewumi, and C. K. Ayo, "Stock price prediction using the ARIMA model," in *Proc. UKSim-AMSS 16th Int. Conf. Comput. Modelling Simulation*, Mar. 2014, pp. 106–112.

SARMISTHA DAS received the bachelor's degree in information technology from the Narula Institute of Technology, in 2019, and the master's degree in information technology from the University of Calcutta, in 2021. She is currently pursuing the Ph.D. degree with Indian Institute of Technology, Patna. She is a Junior Research Fellow with Indian Institute of Technology, Patna, focusing on machine learning, deep learning, natural language processing, data mining, and computer vision techniques.

UJJWAL CHOWDHURY received the B.Sc. degree in mathematics from Vidyasagar University, in 2020. He is currently pursuing the M.Sc. degree in data science with the Ramakrishna Mission Vivekananda Educational and Research Institute. He is also an AI Research Executive with Vista Intelligence Private Ltd.

N. S. LIJIN is currently pursuing the bachelor's degree in data science and engineering with Indian Institute of Science Education and Research, Bhopal, India. Since December 2022, he has been a Research Intern with the Computer Science and Engineering Department, Indian Institute of Technology, Patna, India. His research interests include artificial intelligence, machine learning, deep learning, and natural language processing. He has previously involved on projects, including early detection of depression over social media, deepfake detection, human activity recognition using smartphones, and facility location planning for solid waste management.

ATULYA DEEP is currently pursuing the degree in enthusiastic artificial intelligence with the SRM Institute of Science and Technology, seeking to deliver state-of-the-art AI solutions for the world. He was involved in algorithms and neural network models for the own personal projects and research papers. His research interests include machine learning, data structures, and computer vision.

SRIPARNA SAHA (Senior Member, IEEE) received the M.Tech. and Ph.D. degrees in computer science from Indian Statistical Institute, Kolkata, Kolkata, India. From September 2009 to June 2010, she was a Postdoctoral Research Fellow with the University of Heidelberg, Germany. From September 2010 to January 2011, she was a Postdoctoral Research Fellow with the Department of Information Engineering and Computer Science, University of Trento, Italy. She is currently the Head of the Department of Computer Science and Engineering, Indian Institute of Technology, Patna, India. She has more than 400 publications with 5000 citations. Her current research interests include text mining pattern recognition, natural language processing, multiobjective optimization, and biomedical information extraction.

ALKA MAURYA was an Executive Program of management with Cornell University, New York, USA. She is currently a seasoned professional with over 14 years of experience in various fields, including operations, analytics, digital analytics, statistics, project management, business development, and strategic planning. She held a senior management positions in various organizations. She holds the position of the Associate Director of Crisil Pvt., Ltd.