Expert Systems With Applications 218 (2023) 119611

Contents lists available at ScienceDirect

Expert Systems With Applications

journal homepage: www.elsevier.com/locate/eswa

Targeted aspect-based emotion analysis to detect opportunities and precaution in financial Twitter messages

Silvia García-Méndez *, Francisco de Arriba-Pérez, Ana Barros-Vila, Francisco J. González-Castaño

Information Technologies Group, atlanTTic, University of Vigo, EI Telecomunicación, Campus Lagoas-Marcosende, Vigo, 36310, Spain

ARTICLE INFO

Keywords:
Aspect-based emotion analysis
Machine learning
Natural language processing
Opinion mining
Personal finance management
Portfolio optimisation

ABSTRACT

Microblogging platforms, of which Twitter is a representative example, are valuable information sources for market screening and financial models. In them, users voluntarily provide relevant information, including educated knowledge on investments, reacting to the state of the stock markets in real-time and, often, influencing this state. We are interested in the user forecasts in financial, social media messages expressing opportunities and precautions about assets. We propose a novel Targeted Aspect-Based Emotion Analysis (TABEA) system that can individually discern the financial emotions (positive and negative forecasts) on the different stock market assets in the same tweet (instead of making an overall guess about that whole tweet). It is based on Natural Language Processing (NLP) techniques and Machine Learning streaming algorithms. The system comprises a constituency parsing module for parsing the tweets and splitting them into simpler declarative clauses; an offline data processing module to engineer textual, numerical and categorical features and analyse and select them based on their relevance; and a stream classification module to continuously process tweets on-the-fly. Experimental results on a labelled data set endorse our solution. It achieves over 90% precision for the target emotions, financial opportunity, and precaution on Twitter. To the best of our knowledge, no prior work in the literature has addressed this problem despite its practical interest in decision-making, and we are not aware of any previous NLP nor online Machine Learning approaches to TABEA.

1. Introduction

In this section, the context of the work, as well as the research problem, will be discussed, paying special attention to sentiment and emotion analysis fields towards targeted aspect-based emotion analysis, in which the proposed solution is framed. Then, the contributions of the research will be described along with the paper organisation.

1.1. Application context

Crowdsourcing in Web 2.0 and beyond and, particularly, the rapid development of social media platforms, have produced massive digital content in many sectors, including finance (Sohangir et al., 2018). Microblogging and social trading networks allow users to track the stock markets' behaviour from the valuable comments by expert users and other screening data. These sources are widely used by financial screening solutions owing to the real-time data provided. They are useful to generate indicators on asset pricing (Houlihan & Creamer, 2021), loan and insurance underwriting (Bee et al., 2021), and financial risk prevention (Gao, 2021; Li et al., 2021) for novice traders and stockholders (Ge et al., 2020).

It has been reported that user contributions to social media platforms such as Twitter[1] clearly influence the behaviour of stock markets (Li, van Dalen, & van Rees, 2018; Mai et al., 2018; Ronaghi et al., 2022). This type of content has been analysed to successfully predict sales (Kim et al., 2021; Pai & Liu, 2018; Yuan, Xu, et al., 2018). Some social trading platforms such as eToro[2] and XTB[3] also keep users' reviews and comments along with past pricing data on the assets they trade, which can be processed. For example, the Thomson

* Corresponding author.
E-mail addresses: sgarcia@gti.uvigo.es (S. García-Méndez), farriba@gti.uvigo.es (F. de Arriba-Pérez), abarros@gti.uvigo.es (A. Barros-Vila), javier@det.uvigo.es (F.J. González-Castaño).

[1] Available at https://twitter.com, January 2023.
[2] Available at https://www.etoro.com, January 2023.
[3] Available at https://www.xtb.com, January 2023.
[4] Available at https://eikon.thomsonreuters.com, January 2023.

https://doi.org/10.1016/j.eswa.2023.119611
Received 26 March 2022; Received in revised form 17 January 2023; Accepted 23 January 2023
Available online 31 January 2023

Reuters Eikon platform[4] claims to analyse financial news and market sentiments.

1.2. Research problem

The user opinions on the web pages and social platforms can only be analysed with automatic methods (Alamoudi & Alghamdi, 2021; Da'u et al., 2020; Hemmatian & Sohrabi, 2019; Kang et al., 2018; Nilashi et al., 2021; Wang et al., 2022) based on Artificial Intelligence (AI) techniques such as Natural Language Processing (NLP) and Machine Learning, due to the noise and complexity of the messages.

From an application perspective, our work pursues tools used to detect market speculations in social media posts, whether expressing opportunities or precautions, about the several financial assets or stock markets that may be present in the same social media post, to support decision-making in finance (Li & Tang, 2020).

From an analytical perspective, the problem in this work is modelled as a Targeted Aspect-Based Emotion Analysis (TABEA), which derives from the homologous problem in sentiment analysis and is related to Opinion Target Extraction (OTE). There exist two approaches within Aspect-Based Sentiment Analysis (ABSA): Aspect-Category Sentiment Analysis (ACSA) (Liao et al., 2021) and Aspect-Term Sentiment Analysis (ATSA) (Peng et al., 2022). ACSA seeks to extract the polarity of predefined aspect categories even though most of the time these aspects are not explicit in the text (*e.g.*, "they offer high-quality steaks", referring to aspect category "food" of implicit target "hotel"). ATSA extracts aspects related to certain target aspects/entities (for example, "Hilton"). When ABSA considers several different entities, it is referred to as Targeted Aspect-Based Sentiment Analysis (TABSA). The only difference with TABEA is that our work focuses on emotions (Chen et al., 2018; Polignano et al., 2021; Rout et al., 2018; Shahin et al., 2022), rather than on polarities, although it is also applied at aspect/entity levels. This is necessary because financial news or social media posts often contain different aspects and emotions, and we seek an accurate classification system conducting fine-grained emotion analysis of stock market assets.

1.3. From basic sentiment and emotion analysis to online TABEA

Sentiment and emotion analysis in their more general forms, whether applied at a coarse or fine-grained level to finance, are relevant to our research problem. They have inspired us to take certain practical considerations as the Machine Learning approach selected.

Sentiment analysis of user opinions has been a prolific research field (Madhu, 2018; Nurifan et al., 2019; Yue et al., 2019). Nowadays, sentiment analysis can automatically process large-scale data from social media platforms (Yue et al., 2019), blogs (Madhu, 2018), online communities and fora such as wikis (Nurifan et al., 2019) and other collaborative media. Sentiment analysis often relies on NLP techniques (Dogra et al., 2021), mostly on English texts. It usually classifies texts into three polarity levels (negative, neutral, and positive) and can be performed at document (Rhanoui et al., 2019), sentence (Arulmurugan et al., 2019) and aspect/entity (Mowlaei et al., 2020) levels. Both coarse and fine-grained sentiment analysis approaches have traditionally relied on Machine Learning algorithms whose accuracy depends on the availability of large labelled data sets with enough context information for the target domain (Elnagar et al., 2018). Document-oriented sentiment analysis seeks an overall document polarity score, regardless of the fact that documents are composed of many sentences where distinct targets and polarities coexist. The same applies to the sentence level. Different topics (asset *tickers* in this research) in a sentence may have different respective polarities. Depending on the depth of the level, sentiment analysis may be more complex. Coarse-grained document and even sentence-level sentiment analysis cannot handle the different aspects of users' comments and their associated polarities. This is the goal of entity/aspect-based sentiment analysis (Ozyurt & Akcayol, 2021), which focuses on the polarities of users' opinions about specific features (aspects) of the products or services (topics) under evaluation.

Previous work on financial emotion analysis to characterise the behaviour of stock markets, on which our research focuses, is scarce (De Arriba-Pérez et al., 2020; Duxbury et al., 2020). In fact, most previous works have exclusively focused on human emotions such as anger, fear, joy, love, sadness, and surprise (Nandwani & Verma, 2021). Nowadays, the interest in emotion analysis is widely recognised in academic and industrial research. Thus, we contribute to the solution of this problem with the particular approach of TABEA, as previously stated.

At a processing framework level, online or streaming Machine Learning (Benllarch et al., 2021; Mohawesh et al., 2021; Seth et al., 2021) is useful to handle real-time data in business applications (Baier et al., 2020), compared to Machine Learning batch approaches that handle changing input datasets with overall periodical retraining, even if this involves outdated data (Pugliese et al., 2021). In finance, unlike batch classification, online classification allows discovering relevant events as soon as they occur and taking early decisions (Ko & Comuzzi, 2021, 2022).

1.4. Contributions

We present a Machine Learning system for TABEA supported by NLP techniques specifically designed to detect two new financial emotions on textual content of the widely used Twitter platform, "opportunity" (De Arriba-Pérez et al., 2020) and "precaution". In the terminology by Plutchik (2004), "opportunity" would roughly correspond to positive emotions such as "expectation" or "anticipation". "Precaution" would be related to "disapproval", a negative opinion and, in a financial environment, it could correspond to a pessimistic feeling about a company or asset. These two new emotions allow identifying tweets that speculate about asset rises and falls.

To the best of our knowledge, no prior work has applied NLP to TABEA in streaming mode, either in finance or any other field. Thus, our problem and its analytical solution are the original contributions of this work. Besides, with the exception of our preliminary research on opportunity detection (De Arriba-Pérez et al., 2020), the two new emotions we consider are also novel as here considered.

1.5. Paper organisation

The rest of this work is organised as follows. Section 2 discusses related work on Opinion Mining, sentiment analysis and emotion analysis from different perspectives (the use of NLP techniques, Machine Learning algorithms, etc.). Section 3 presents our classification problem and the architecture of our solution. Section 4 provides details about the financial data set and its features, and validates our approach with experimental results. Finally, Section 5 concludes the paper.

2. Related work

As previously mentioned, sentiment and emotion analysis related works will be discussed in this section, with special consideration to their application in the financial domain. Finally, explainability and data analysis in streaming will be discussed as relevant related fields.

2.1. Sentiment analysis

Researchers use techniques such as feature-engineering to characterise textual content like user comments (Carrillo-de Albornoz et al., 2018; Nawangsari et al., 2019). In this regard, Weichselbraun et al. (2017) used an aspect lexicon that was built from training corpora, including ConceptNet,[5] Wikipedia[6] and Wordnet[7] to identify aspects.

[5] Available at https://conceptnet.io, January 2023.
[6] Available at https://es.wikipedia.org, January 2023.
[7] Available at https://wordnet.princeton.edu/, January 2023.

The more sophisticated solution by Liu et al. (2021), a Global Semantic Memory Network for ABSA, enriches the representation of the targeted aspects with context information. Polarity lexica such as WordNet (Jiménez et al., 2019) and SentiWordNet (Madani et al., 2020) allow extracting the most likely sentiment that is associated with a certain word (Mumtaz & Ahuja, 2018). Then, sentiment propagation methods can be applied to provide a final score (Li, Guo, et al., 2018). For example, Fernández-Gavilanes et al. (2018) presented an unsupervised Machine Learning algorithm to automatically generate sentiment lexica from sentiment scores that were propagated across syntactic trees. Nevertheless, unsupervised approaches have been exclusively applied to coarse-grained document-level sentiment analysis so far. A promising strategy both for supervised and unsupervised approaches is stacking, which consists in combining different Machine Learning techniques to improve performance (Mehmood et al., 2018; Wang et al., 2019) and allows handling neutrality and sentiment ambivalence (De Arriba-Pérez et al., 2020; Macdonald & Birdi, 2019; Valdivia et al., 2018) by filtering neutral opinions to improve the performance of binary polarity classification. Moreover, Wang et al. (2020) presented a multi-level fine-scaled sentiment analysis system with ambivalence handling at the sentence level using keywords like *extremely* and *minor*, and then transferring polarity to paragraph/article levels with aggregation methods (*i.e.*, sum function) and manual rules.

Sentiment analysis research has recently begun to apply graph neural models such as Convolutional Neural Networks (Liang et al., 2022; Zhao et al., 2022). Ongoing research also exploits syntactic and semantic structures into graph neural networks (An et al., 2022; Phan et al., 2022; Xiao et al., 2022). For example, Xiao et al. (2022) employed information of syntactic dependency trees from part-of-speech (POS) graphs. Then, contextual semantic dependencies are extracted using a syntactic distance attention mechanism over a densely connected graph convolutional network. In Liang et al. (2022), a graph convolutional network considers sentence-aspect affective dependencies extracted from SenticNet.[8] Moreover, the heterogeneous graph neural network by An et al. (2022) is composed of the word, aspect, and sentence nodes. The resulting structure and semantic data allow updating feature embeddings. Zhao et al. (2022) presented an aggregated graph convolutional network with an attention mechanism to model sentiment dependencies. Phan et al. (2022) exploited syntax-, semantic-, and context-based graphs as part of a convolutional neural network with an attention mechanism for aspect-level sentiment analysis.

2.2. Emotion analysis

Matsumoto et al. (2022) have performed an emotional breakdown analysis for chatbot solutions. They detect emotions in users' utterances with deep neural networks with sentence representation vectors. Guo (2022) has also developed a deep learning-based solution by applying semantic text analysis for human emotion detection. Thus, no domain-focused emotions have been defined for finance, with the exception of our previous research on opportunity detection (De Arriba-Pérez et al., 2020). For example, Razi et al. (2017) presented a multi-layer perceptron to analyse the influence of the aforementioned "general" emotions on investing behaviour. This was also the case of Zhou et al. (2018), who studied the correlation between the trends of the Chinese stock market and five relevant features derived from "general" emotions. Other related problems can be found in Ahn and Kim (2021), Chung and Zeng (2020), Taffler (2018).

Yuan, Lau, et al. (2018) proposed a novel Emotion Latent Dirichlet Allocation (ELDA) model for credit rating prediction from Twitter data. Akhtar et al. (2020) presented a stacked ensemble Machine Learning solution to predict sentiments and emotions, but they only considered sentiment analysis nor emotion analysis in the financial domain. ASPENDYS (Theodorou et al., 2021) is an AI investment platform that mines financial text from Twitter and Stocktwits[9] searches to assist investors with personalised recommendations. However, it is only based on sentiment analysis. Chun et al. (2021) employed a simple Emotion Term Frequency–Inverse Emotion Document Frequency (ETF- IEDF) technique derived from the classical Term Frequency–Inverse Document Frequency (TF- IDF) (Qaiser & Ali, 2018) instead of sentiment analysis to predict stock market prices using Machine Learning models, with good results. Finally, Valle-Cruz et al. (2022) analysed the influence of stock markets on investors emotions.

2.3. Financial domain

The knowledge extraction system by Wang (2017) analysed sentiment time series from microblogs for stock market forecasting. Dridi et al. (2019) proposed a supervised Machine Learning sentiment analysis model for finance with competitive performance based on semantic features. Khan et al. (2020) proved that public sentiment and the political context affect stock market trends. They studied data from Yahoo! Finance,[10] Twitter messages and Wikipedia political information to differentiate between stock markets that are hardly predictable and those that are easily influenced by financial news and social media. They combined spam filtering and feature selection with ensemble classifiers. The more sophisticated solution by Dogra et al. (2021) applies deep contextual language representation into the DistilBERT supervised sentiment analysis model (Hew et al., 2020).

In our work, based on the fact that a given financial text from the news or social media posts often contains different aspects and polarities, we seek an accurate classification system conducting fine-grained emotion analysis of stock market assets. As previously said, no previous work has considered financial opportunity and precaution from a TABEA perspective.

2.4. Interpretability and explainability

Few recent state-of-the-art AI solutions have explored the interpretability and explainability of the models to explain their outcome to the end users to foster trustworthiness (Gaur et al., 2021). They apply symbolic techniques (Cambria et al., 2022) (*e.g.*, auto-regressive language models, counterfactual explanations, kernel methods, post-hoc interpretability, rule-based explanations and statistical methods).

2.5. Data streaming analysis

To avoid the disadvantages of batch retraining (Ortiz-Martínez, 2016; Turchi et al., 2017), there exist real-time data streaming analysis alternatives, such as Massive Online Analysis (MOA),[11] Scikit-multiflow[12] and StreamDM.[13] These methods are appropriate for the fast, continuous and imbalanced real-time data from dynamic environments like social media platforms. They also allow for analysing the temporal evolution of input data. However, they have been largely unexplored in financial use cases, with some exceptions like the work by Nagarajan and Gandhi (2019). They describe an online sentiment analysis system for Twitter messages that outperforms traditional batch Machine Learning classifiers. However, they do not consider emotions, either financial or any other.

[8] Available at https://sentic.net, January 2023.

[9] Available at https://stocktwits.com, January 2023.

[10] Available at https://finance.yahoo.com, January 2023.

[11] Available at https://moa.cms.waikato.ac.nz, January 2023.

[12] Available at https://scikit-multiflow.readthedocs.io/, January 2023.

[13] Available at http://huawei-noah.github.io/streamDM, January 2023.

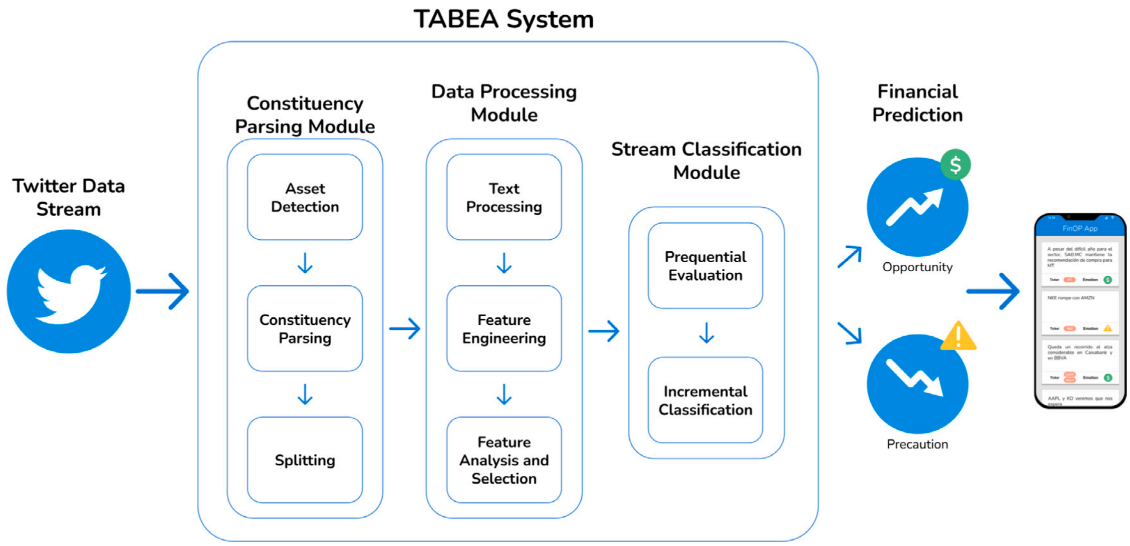

Fig. 1. System scheme.

2.6. Summary

Summing up, no previous authors have considered TABEA of financial, social media to detect investment opportunities or precautions about financial assets. In fact, as previously said, no prior work has applied NLP nor online Machine Learning to TABEA, either in finance or any other field.

3. System architecture

In the following sections, we describe in detail our novel TABEA system for fine-grained detection of financial opportunities and precautions on Twitter messages about stock markets and assets. Fig. 1 shows the system scheme. Unlike our previous work (De Arriba-Pérez et al., 2020) as well as other prior work in the literature, it is a streaming scheme handling continuous flows of information and thus facilitating retraining of Machine Learning models. It is composed of (*i*) a constituency parsing module for tweet parsing and splitting; (*ii*) an offline data processing module to engineer features and select the best ones based on their relevance; and (*iii*) a stream classification module that processes tweets on-the-fly.

3.1. Constituency parsing module

This module supports TABEA through constituency parsing (tweet structure tree) and boundary splitting based on asset detection and linguistic features (morphology and syntax). Particularly, we detect Simple Declarative Clauses or SDC (*i.e.*, those not introduced by a subordinating conjunction or a wh-word, and without subject-verb inversion) within the textual content of the tweets. The outcome of the module is a hierarchical dependency graph of the words that compose the tweet, as shown in Fig. 2.

Firstly, all assets in the tweet are identified using our financial lexica,[14] and the tweets are split into simple declarative clauses. Then, the resulting segments are grouped if the following rules hold and forward propagated otherwise:

1. The next segment does not contain any asset nor the additive conjunction *y* ‘and’ nor a comma (,) nor a hyphen (–).

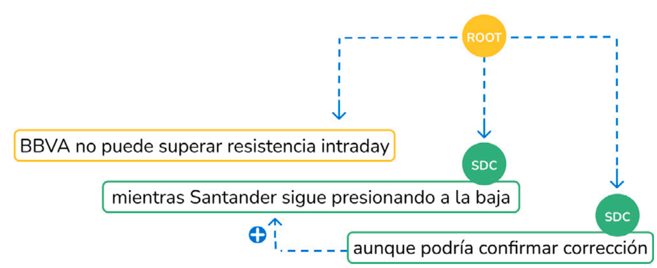

Fig. 2. Tweet segments after applying forward propagation.

2. The next segment is a relative clause, *i.e.*, it starts with the relative conjunction *que* ‘that’ followed by a syntagm that contains any asset.

In the example in Fig. 2, the two green clauses are grouped due to condition 1. The final segments after applying forward propagation are: *BBVA no puede superar resistencia intraday* ‘BBVA cannot overcome intraday resistance’ and *mientras Santander sigue presionando a la baja aunque podría confirmar corrección* ‘while Santander continues to press down although it could confirm a correction’.

After the grouping is completed, a final segmentation based on regular expressions is performed to address the specific case of tweets containing lists of several assets (*e.g.*, ALUA.BA *−2,57%* EDN *+8,08%* CRES.BA *−4,86%*...). In this particular and rather typical scenario, segments are split again right where the next asset is detected.

In the end, only segments containing at least one asset remain, and the other segments are discarded by the constituency parsing module. Algorithm 1 summarises the actions performed by the module.

3.2. Data processing module

This module processes the text, generates and engineers features and selects the most relevant ones for classification.

3.2.1. Text processing

Text processing improves the efficiency of classification. It comprises the detection of asset and numeric values, filtering, cleaning and removing unnecessary information; hashtag splitting; and text lemmatisation tasks.

[14] Available at https://www.gti.uvigo.es/index.php/en/resources/14-resources-for-finance-knowledge-extraction, January 2023.

Algorithm 1: Constituency parsing and forward propagation

```
Input: Financial lexica and textual content from the tweet
Data: financial_lexica, tweet_text, numbers (regular expression to detect numerical characters)
Result: List of segments with a single asset or few strongly related assets
begin
    tweet_segmented = declarative_clause_segmenter(tweet_text);    /* List of segments from the tweet text obtained by
     constituency parsing. Assets are identified by the ticker tag.*/
    segment_grouped = []
    segment_aux=""
    foreach segment in tweet_segmented do
        if TICKER OR "and" OR "," OR "-" in segment then
            segment_grouped.add(segment_aux)
            segment_aux=segment
        else
            segment_aux+=segment                                   /* Group segments based on consideration 1.*/
    segment_grouped.add(segment_aux)
    segment_grouped.remove_empty_segments()
    list_segments= segment_grouped.copy()
    segment_grouped = []
    foreach r, r+1 in range(len(list_segments)-1) do
        if TICKER in list_segments[r] and TICKER in list_segments[r+1] and list_segments[r+1] starts_with "that" then
            list_segments[r+1]=(list_segments[r] + list_segments[r+1])   /* Group segments based on consideration 2.*/
        else
            segment_grouped.add(list_segments[r])
        segment_grouped.add(list_segments[len(list_segments) - 1])
    result = []
    foreach segment in segment_grouped do
        if segment.regex("numbers")>1                    /* Regular expression to detect numerical characters.*/
         then
            result.addAll(regex.split_when_found(segment,TICKER))   /* To address the specific case of those tweets reporting
             the state of several assets.*/
        else
            result.add(segment)
// Returns the tweet segments
result.keep_segments_with_ticker()
return result
```

- **Asset identification**: we tag all financial assets with the TICKER tag (note that this is consistent with the fact that the rules in Algorithm 1 refer to any asset). For this purpose, we use the same dictionaries mentioned in Section 3.1. Moreover, numbers and percentages are replaced by NEGATIVE or POSITIVE tags, depending on the mathematical signs that precede them, and, otherwise, by the NUMBER tag.
- **Filtering, cleaning and removal**: we detect and remove symbols $, @ and #. Meaningless words such as connectors[15] are also removed from the text along with URLS and retweet (RT) tags. We keep words *no* 'not', *sí* 'yes', *muy* 'very' and *poco* 'few' due to their semantic load since they provide nuances of assets' trends.
- **Hashtag splitting**: in order to maximise relevant textual data in the tweets and to fix spelling mistakes of the users, we use our lexica (García-Méndez et al., 2018, 2019) and the Spanish CREA frequency reference corpus by *Real Academia Española de la Lengua*[16] to decompose hashtags and compounds such as *mayorcaída* 'biggestfall' into *mayor caída* 'biggest fall'.
- **Text lemmatisation**: the textual content is first tokenised into words and checked using a Spanish dictionary to keep only correct words. Finally, the tokens are lemmatised, and spelling mistakes are corrected with our text distance algorithm by replacing them with the most likely candidate (using again the CREA corpus).

[15] Available at https://www.ranks.nl/stopwords/spanish, January 2023.
[16] Available at http://corpus.rae.es/lfrecuencias.html, January 2023.

Table 1 illustrates the operation of the text processing stage.

3.2.2. Feature engineering

This module generates and engineers features from the text or from external quantitative data sources related to stock markets and assets. Table 2 summarises the features we consider:

- **Textual features**: char-grams, word-grams and word tokens, plus a bag of words (BOW) composed of the most frequent words and bigrams that are unique for each emotion category.
- **Numerical features**: these features characterise each entry by tweet length; amount of numerical values and percentages (negative, positive and total); amount of financial abbreviations, exclamation and interrogative symbols; amount of adverbs (total, negative, positive, expressing doubt and intensifiers); and amount of words with polarity.[17] (negative, neutral and positive) and describing general emotions[18] To create these features, we perform morphological and syntactic parsing and exploit our lexica.
- **Boolean features**: we compute the variations of stock market assets as the differences between their previous and posterior prices by considering the temporary window between the working day before the tweet was posted and the end of the following working day. Then, we indicate if the trend is upwards or downwards.

[17] Available at https://www.gti.uvigo.es/index.php/en/resources/8-lexicon-of-polarity-and-list-of-emojis-by-polarity-and-emotion-for-application-in-the-financial-field, January 2023.
[18] Available at https://www.cic.ipn.mx/~sidorov/SEL.txt, January 2023.

Table 1
Example of tweet after text processing.

	Tweet
Before	*$Bankia sigue el crack bursátil. −1,925 euros, del IBEX35 #mayorcaída* *https://t.co/S73BxUSKiR* '$Bankia follows the stock market crash. −1,925 euros, the biggest fall in IBEX35 #biggestdrop https://t.co/S73BxUSKiR'
After	TICKER*#1 seguir bursátil* NEGATIVE *euros* TICKER*#2 mayor caída* 'TICKER#1 follow stock market NEGATIVE euros TICKER#2 biggest fall'

Table 2
Features and target of the Machine Learning models.

Type	Num.	Feature name	Description
Textual	1	CHAR_GRAMS	CharacterL *n*-grams
	2	WORD_GRAMS	WordL *n*-grams
	3	WORD_TOKENS	Character *n*-grams only from text inside word boundaries
	4	PRE_BOW	textscbow of most frequent words and bigrams that are exclusive of precaution emotion
	5	NEU_BOW	textscbow of most frequent words and bigrams that are exclusive of neutral emotion
	6	OPP_BOW	textscbow of most frequent words and bigrams that are exclusive of opportunity emotion
Numerical	7	LEN_TWEET	TweetL length
	8	NEG_NUM	AmountL of negative numerical values
	9	POS_NUM	AmountL of positive numerical values
	10	TOTAL_NUM	AmountL of numerical values
	11	NEG_PERC	AmountL of negative percentages
	12	POS_PERC	AmountL of positive percentages
	13	TOTAL_PERC	AmountL of percentages
	14	FIN_ABBR	AmountL of financial abbreviations
	15	EXCLAMATION	AmountL of exclamation marks
	16	INTERROGATION	AmountL of interrogation marks
	17	ADVERBS	AmountL of adverbs
	18	ADVERBS_NEG	AmountL of negative adverbs
	19	ADVERBS_POS	AmountL of positive adverbs
	20	ADVERBS_DOUBT	AmountL of adverbs expressing doubt
	21	ADVERBS_INT	AmountL of intensifying adverbs
	22	NEG_POLARITY	AmountL of words with negative polarity
	23	NEU_POLARITY	AmountL of words with neutral polarity
	24	POS_POLARITY	AmountL of words with positive polarity
	25	NEG_EMOTION	AmountL of words that express sadness, anger and other negative emotions
	26	POS_EMOTION	AmountL of words that express happiness, surprise and other positive emotions
Boolean	27	TREND	IndicatesL if the asset exhibits an upward or downward trend
Target	28	EMOTION	FinancialL emotion tag

Table 3 provides some examples of tweets and their corresponding feature values.

3.2.3. Feature analysis and selection

We compute the Pearson Correlation Coefficient (Eq. (1)) to measure the correlations between the features in Table 2. For any two features x, y:

$$r_{xy} = \frac{\sum (x_i - \overline{x})(y_i - \overline{y})}{\sqrt{\sum (x_i - \overline{x})^2}\sqrt{\sum (y_i - \overline{y})^2}} \tag{1}$$

where $r_{xy} \in [-1, 1]$. Then, we apply a selection algorithm to identify the features with higher correlation to reduce the original dimensionality of the data space and, thus, the computational requirements of the models.

3.3. Stream classification module

We designed a multi-class stacking (Malmasi & Dras, 2018) classifier model. It was implemented in streaming (online) mode (Montiel et al., 2018), which seemed adequate given the nature of social data.

In the stacking ensemble, the first classifier stage differentiates among precaution, opportunity, and neutral (that is, any other) emotions. The second classifier re-evaluates this prediction to improve its accuracy by differentiating again between the corresponding emotion (opportunity or precaution) and the neutral emotion. Fig. 3 shows the stacking scheme.

We evaluated several Machine Learning algorithms for the stacking ensemble based on their promising performance in similar classification problems (Barrón Estrada et al., 2020; Kaur, 2020; Muneer & Fati, 2020):

- Naive Bayes (NB)[19] (Berrar, 2019)
- Decision Tree (DT)[20] (Trabelsi et al., 2019)
- Random Forest (RF)[21] (Parmar et al., 2019)
- Stochastic Gradient Descent (SGD)[22] (Mehlig, 2021)

4. Evaluation

All the experiments were executed on a computer with the following specifications:

- **Operating System**: Ubuntu 18.04.2 LTS 64 bits
- **Processor**: Intel@Core i9-9900K 3.60 GHZ
- **RAM**: 32 GB DDR4
- **Disk**: 500 Gb (7200 rpm SATA) + 256 GB SSD

[19] Available at https://scikit-multiflow.readthedocs.io/en/latest/api/generated/skmultiflow.bayes.NaiveBayes.html, January 2023.

[20] Available at https://scikit-multiflow.readthedocs.io/en/stable/api/generated/skmultiflow.trees.ExtremelyFastDecisionTreeClassifier.html, January 2023.

[21] Available at https://scikit-multiflow.readthedocs.io/en/stable/api/generated/skmultiflow.meta.AdaptiveRandomForestClassifier.html, January 2023.

[22] Available at https://scikit-learn.org/stable/modules/sgd.html, January 2023.

Table 3
Examples of tweet textual content and corresponding feature values. Character "–" represents blank spaces.

		Tweet sample 1	Tweet sample 2
		30-07-2019 #Ibex35 −2,48% sigen llegando resultados llega agosto mucho cuidado con los piratas de guante blancoveremos si es movido o no... https://t.co/3IX2zsNpEC" 07-30-2019 #Ibex35 −2.48% The results keep coming, August arrives, be very careful with the white glove pirates, we will see if it moves or not... https://t.co/3IX2zsNpEC	*#IBEX35 La superación, otra vez, del 9375 del índice, aupado posiblemente por una recuperación de la banca, que ha estado muy castigada, podría darle fortaleza para ir a cotas más altas.* #IBEX35 Again, the overcoming of the 9375 reference, possibly boosted by a recovery of the banking system, which has been hit hard, could give the index strength to reach higher levels.
Features	1	[number,stock, ...]	[stock,superación, ...]
	2	[numb,umbe,mber,ber_, ...]	[stoc,tock,ock_,ckv_s, ...]
	3	[_num,numb,umbe,mber,ber_, ...]	[_sto,stoc,tock,ock_, ...]
	4	[mucho cuidar]	–
	5	–	–
	6	–	[vez number]
	7	135	139
	8	0	0
	9	0	1
	10	0	1
	11	1	0
	12	0	0
	13	1	0
	14	0	0
	15	0	0
	16	0	0
	17	2	2
	18	1	0
	19	0	0
	20	0	1
	21	1	1
	22	1	0
	23	0	0
	24	2	2
	25	0	0
	26	0	0
	27	downward	upward
	28	precaution	opportunity

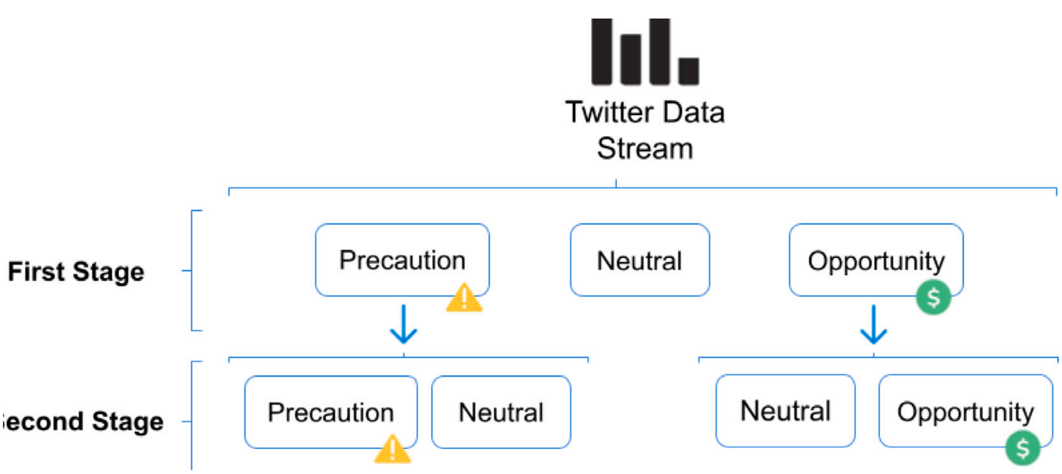

Fig. 3. Streaming classification on stacking.

4.1. Experimental data set

The experimental data set[23] is composed of 5,000 tweets manually tagged as explained in Section 4.3. It was collected using the Twitter API[24] from January 14th to June 21st 2020. It is similar in size to the data sets in other studies (Al-Smadi et al., 2019; Plaza-del Arco et al., 2021; Chatzis et al., 2018; Simester et al., 2019; Tuke et al., 2020). Table 4 shows some examples of the texts of the entries.

4.2. Constituency parsing module

For constituency parsing, we used the CoreNLP library.[25] along with the techniques and considerations explained in Section 3.1

4.3. Annotation of segments

The constituency parsing module may produce segments with one or several different assets. In the second case, each segment in the training

[23] Available from the corresponding author on reasonable request.
[24] Available at https://developer.twitter.com/en/docs, January 2023.
[25] Available at https://stanfordnlp.github.io/CoreNLP, January 2023.

Table 4
Examples of the texts of the entries of the data set.

A pesar del difícil año para el sector, SAB.MC mantiene la recomendación de compra para MT
'Despite the difficult year for the sector, SAB.MC maintains a buy recommendation for MT'
Queda un recorrido al alza considerable en Caixabank y en BBVA
'There is a considerable upward path in Caixabank and BBVA'
Bankia sigue el crack bursátil. 1,925 euros, mayor caída del IBEX35
'Bankia follows the stock market crash. 1,925 euros, the biggest fall in IBEX35'
NKE rompe con AMZN
'NKE breaks with AMZN'
Asíva el IBEX35 en lo que llevamos del mes de agosto
'This is how IBEX35 has gone so far in August'
AAPL y KO veremos que nos espera
'AAPL and KO we will see what awaits us'

Table 5
Distribution of annotated segments in the data set by financial emotion.

Emotion	Number of segments
Precaution (P^-)	1644
Neutral (N)	4172
Opportunity (O^+)	2392
Total	8208

Table 6
Coincidence matrix of the annotation process.

	P^-	N	O^+
P^-	2827	341	102
N	341	7505	662
O^+	102	662	3466

Table 7
Inter-agreement Alpha-reliability of emotion tag by pairs of annotators.

	Ann. 1	Ann. 2	Ann. 3
Annotator 1	–	0.792	0.703
Annotator 2	0.792	–	0.819
Annotator 3	0.703	0.819	–

Table 8
Inter-agreement accuracy of emotion tag by pairs of annotators.

	Ann. 1	Ann. 2	Ann. 3
Annotator 1	–	0.874	0.822
Annotator 2	0.874	–	0.890
Annotator 3	0.822	0.890	–

set is replicated as many times as the different assets it contains to ensure a correct annotation. For each replica, only one asset is tagged as TICKER while the rest are tagged as OTHER_TICKER. Each replica is annotated as opportunity, precaution or neutral by focusing on the role of TICKER.

The segmented entries of the data set were manually labelled by three experts in both NLP and finance. Table 5 shows the resulting distribution of the segments in the experimental data set by emotion category tag.

To assess the consistency of the annotation processes of financial emotions precaution, opportunity and neutral, we calculated Alpha-reliability and accuracy (Krippendorff, 2018). Table 6 shows the coincidence matrix of the annotators. Note that the elements in the diagonal of the matrix count the tweets on which the three annotators fully agreed. Tables 7 and 8 show the inter-agreement by pairs of annotators. For reference purposes, the literature considers that values above 0.667 are acceptable and that those near 0.8 are optimal, as in our case (Kilicoglu et al., 2021; Rash et al., 2019; Salminen et al., 2019; Seité et al., 2019).

4.4. Data processing module

Next, we describe the implementations of the text processing, feature engineering, and feature analysis and selection modules.

4.4.1. Text processing

Text processing is based on our financial lexica, and the techniques explained in Section 3.2.1. Particularly, for the lemmatisation process, we used the Freeling library.[26]

4.4.2. Feature engineering

Regarding feature engineering, we applied the methodology explained in Section 3.2.2. Table 2 shows the list of features. Specifically, to produce the textual features (1 to 3 in Table 2) we used `CountVectorizer`.[27] and `GridSearchCV`,[28] both from the Scikit-Learn Python library.[29] with the parameter ranges in Listing 1 The optimal parameter settings were `max_df` = 0.5, `min_df` = 0.001 and `ngram_range` = (1,4) Also, for BOW features (4 to 6 in Table 2) we used `CountVectorizer`, by applying frequency sorting and keeping the 500 most frequent words and bigrams that were exclusive to each emotion category. Finally, asset trends (feature 27 in Table 2) were obtained from Yahoo! Finance.

Listing 1: Parameter ranges for the generation of n-grams.

```
max_df: (0.3,0.35,0.4,0.5,0.7,0.8,1)
min_df: (0,0.001,0.005,0.008,0.01)
ngram_range: ((1,1),(1,2),(1,3),(1,4),(1,4),
(1,5),(1,6),(1,7))
```

4.4.3. Feature analysis and selection

First, we estimated the contributions of the features to the target using the Pearson correlation coefficient, as explained in Section 3.2.3. Table 9 shows the most correlated features. They correspond to percentage values, polarity lexica and assets' trends.

Consequently, both financial historical data and knowledge extracted from social media are relevant to predict negative (precautions) and positive (opportunities) emotions. The moderate correlation levels indicate that the classification problem can be addressed with Machine Learning algorithms.

We then selected the features using the transformer wrapper `SelectPercentile`[30] method from Scikit-Learn with χ^2 score function

[26] Available at http://nlp.lsi.upc.edu/freeling/node/1, January 2023.
[27] Available at https://scikit-learn.org/stable/modules/generated/sklearn.feature_extraction.text.CountVectorizer.html, January 2023.
[28] Available at https://scikit-learn.org/stable/modules/generated/sklearn.model_selection.GridSearchCV.html, January 2023.
[29] Available at https://scikit-learn.org, January 2023.
[30] Available at https://scikit-learn.org/stable/modules/feature_selection.html, January 2023.

Table 9
Most correlated features with the target (feature identifiers from Table 2).

Feature num.	Correlation
24	0.10
12	0.14
27	0.21

Table 10
Performance of diverse single-stage Machine Learning classification models for financial emotions, initial tests.

Classifier	Accuracy	Precision			Recall		
		P^-	N	O^+	P^-	N	O^+
NB	46.82	79.14	**80.65**	36.17	59.55	14.89	**93.77**
DT	60.23	74.06	64.43	49.64	25.54	76.16	57.60
RF	71.82	**80.08**	71.05	69.75	48.42	**88.28**	59.20
SGD	**75.06**	67.83	79.67	**71.83**	**67.45**	80.51	70.78

and 15th percentile threshold, *i.e.*, the features above that threshold were considered relevant. In the end, the features we selected were 5, 8–10, 19, 21, 22 and 24 in Table 2, plus 1,734 out of 11,555 n-gram features.

4.5. Streaming classification

In this section, we evaluate the final performance of our system to detect financial opportunities and precautions. The results were computed using two different streaming approaches. The first is a single-stage scheme as a baseline, using the classifiers in Section 3.3. The second is the multi-class stacking ensemble described in the same section. Since both were implemented in streaming mode, they were progressively tested and trained by sequentially using each sample from the experimental data set to test the model (*i.e.*, to predict) and then to train the model (*i.e.*, for a partial fit). Performance metrics are obtained as their incremental averages. In particular, we employed the `EvaluatePrequential`[31] library.

As shown in Table 10, RF and SGD exhibited promising performance by attaining accuracies close to 70% in preliminary tests. However, the recall of RF in the detection of precaution and opportunity emotions, and the precision of SGD for precaution, were rather poor.

These results motivated us to build the more complex second approach, a multi-class stacking classification scheme. Tables 11 and 12 show the confusion matrix of the single-stage RF and SGD classifiers. We can observe (in bold) that most errors are concentrated around neutral predictions. Thus, the stacking approach was designed to maximise precision when discerning non-neutral from neutral financial emotions.

In each test, we independently optimised the hyperparameters of the classifiers with `GridSearchCV`. Listings 2 and 3 respectively show the parameter ranges for the RF and SGD models.

Table 13 shows the results for the two best classifiers, RF and SGD, using the single-stage and stacking approaches. Particularly, row 1 of Table 13 contains the results of the single-stage classifiers with the selection of features in Section 4.4.3. Hyperparameter optimisation further improved the best classifiers (with a ~10% precision increase for the RF classifier compared to Table 10[32]). By applying the stacking approach (see row 2), accuracy was similar for better precision in the detection of opportunities and precautions, especially with the SGD classifier, at the cost of some degradation of the precision for neutral emotions. Consequently, the recall of neutral emotions improved.

[31] Available at https://scikit-multiflow.readthedocs.io/en/stable/api/generated/skmultiflow.evaluation.EvaluatePrequential.html, January 2023.

[32] Results obtained without hyperparameter optimisation.

Table 11
Single-stage RF confusion matrix, without BOW features.

	P^-	N	O^+
P^-	796	**620**	228
N	**103**	3683	**386**
O^+	95	**881**	1416

Table 12
Single-stage SGD confusion matrix, without BOW features.

	P^-	N	O^+
P^-	1109	**342**	193
N	**342**	3359	**471**
O^+	184	**515**	1693

We further improved the system by incorporating features 4–6 in Table 2. Rows 3 and 4 of Table 13 show the performance of the single-stage and stacking approaches of the experiments in rows 1 and 2, by adding those extra BOW features. This modification enhanced the system significantly in general (*i.e.*, 10% improvement to 80% in the recall of opportunities with the single-stage SGD classifier, 6% improvement to 91% in the precision of precaution with the stacked RF classifier, and a dramatic 18% improvement in the recall of precaution with the stacked RF classifier, from 56% to 74%). Overall, the best method was the stacked RF classifier with additional BOW features, except for the precision of neutral emotions, for which the single-stage SGD classifier was superior (although its performance was comparable to that of the stacked SGD classifier).

Listing 2: RF hyperparameter configuration.

```
estimators: (10,35,50,100)
max_features: (auto,35,50,100)
lambda: (6,35,50,100)
```

Listing 3: SGD hyperparameter configuration.

```
penalty: (l1,l2,elasticnet)
l1range: (0.05,0.15,0.9)
alpha: (0.001,0.0001,0.00001)
max_iter: (100,1000,10000)
tol: (1e−1,1e−3,1e−5)
```

Summing up, the ~70% accuracy baseline by the RF and SGD models was overcome by the use of BOW features (features 4–6 in Table 2). Note the 6% and 18% improvements with the stacked RF classifier to reach 91% and 74% in precaution precision and recall, respectively. Consequently, we consider the RF classifier the best choice in our scenario, attaining around 85% accuracy, over 90% precision and around 75% recall both for financial precautions and opportunities. These promising results endorse the use of the system to detect financial emotions on Twitter about specific stock market assets as well as its practical interest in decision-making.

The single-stage and stacking results with BOW features (rows 3 and 4 in Table 13) reflect the importance of the latter. The performance increase can be explained by the fact that under-represented textual features (word n-grams and bigrams) are discarded because their document frequency is strictly lower than the given threshold (see `min_df` hyperparameters in Listing 1), prior to model training. BOW features, thanks to the potent corpus they provide, allow generating direct and effective prediction rules for otherwise undetectable patterns. They include word n-grams with a high semantic load. A manual inspection of the BOW corpus reveals meaningful bigram examples such as *ser bajista* 'be bearish', *TICKER quiebra* 'TICKER bankruptcy' and *abajo sin*

Table 13
Performance of RF and SGD single-stage and stacking models in streaming mode.

Configuration	Model	Accuracy	Precision			Recall		
			P^-	N	O^+	P^-	N	O^+
Single-stage (1)	RF	77.68	84.24	75.36	79.93	57.24	91.99	66.76
	SGD	76.05	70.77	79.58	73.02	67.15	82.55	70.82
Stacking (2)	RF	77.27	85.73	73.09	84.97	55.54	94.49	62.17
	SGD	76.82	76.27	76.68	77.49	62.96	87.78	67.22
Single-stage plus BOW features (3)	RF	84.49	90.39	81.73	86.97	74.39	93.39	75.92
	SGD	83.47	79.82	**86.07**	81.38	**80.11**	86.77	**80.02**
Stacking plus BOW features (4)	RF	**84.54**	**91.43**	80.45	**90.24**	73.97	**95.28**	73.08
	SGD	84.33	84.34	84.08	84.85	78.29	90.48	77.76

Table 14
Stacking plus BOW features RF confusion matrix.

	P^-	N	O^+
P^-	1216	**367**	61
N	**69**	3975	**128**
O^+	45	**599**	1748

Table 15
Stacking plus BOW features SGD confusion matrix.

	P^-	N	O^+
P^-	1287	**268**	89
N	**154**	3775	**243**
O^+	85	**447**	1860

Table 16
Relative prediction improvement (percentage) with stacking plus BOW features compared to single-stage, RF model.

	P^-	N	O^+
P^-	–	−40.81%	−73.25%
N	−33.01%	–	−66.84%
O^+	−52.63%	−32.01%	–

Table 17
Relative prediction improvement (percentage) with stacking plus BOW features compared to single-stage, SGD model.

	P^-	N	O^+
P^-	–	−21.64%	−53.89%
N	−54.97%	–	−48.41%
O^+	−53.80%	−13.20%	–

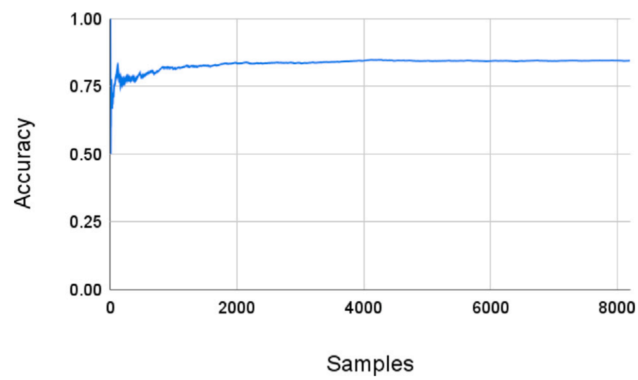

Fig. 4. Evolution of the accuracy of the stacked RF streaming classifier.

'down without' with respective frequencies of appearance 0.00098, 0.00098 and 0.00085, for financial precaution entries; and *experimentar espectacular* 'experience spectacular', *mayor ganancia* 'highest profit' and *alcista TICKER* 'bullish TICKER', with the same respective frequencies of appearance for financial opportunity entries. If BOW features are not forced, they would be excluded from training because their frequency of appearance is less than 0.001.[33]

Another relevant aspect to consider to explain the better performance of our approach, including stacking and BOW features (see row 4 in Table 13) is the handling of neutral tweets. The bold entries in the confusion matrices in Tables 14 and 15, in comparison with the respective values in Tables 11 and 12, show the reduction of categorisation errors due to misclassified neutral tweets as precautions and opportunities, and the other way round (see Tables 16 and 17). These are the hardest to avoid in principle since mutually misclassified precautions and opportunities are much less frequent. In particular, with these modifications, there is 67% and 33% prediction improvement for neutral entries that were previously incorrectly predicted as opportunities and precautions, respectively (see Table 16).

Besides, the combined introduction of stacking and BOW features, compared to the single-stage approach (row 1 in Table 13), further reduces the misclassification of precautions and opportunities up to 73% for the case of precautions incorrectly classified as opportunities with the RF model, 53% in the opposite case. The rationale behind the stacking approach is that its structure insists on the differentiation between neutral and non-neutral emotions. The latter entries are typically expressed more assertively, so it makes sense to expect that there will be low mutual overlapping between precautions and opportunities. Logically, the entries that are separated in the first stage as neutral may include misclassified precaution and opportunity entries, but this is not relevant if the precision of precautions and opportunities is high at the output of the second stage. Consequently, the stacking approach also seems crucial to improve precision, which is the most relevant performance metric from an application perspective in financial investment decision-making (as wrong suggestions are much more harmful than missing suggestions (Akhtar & Das, 2019), *i.e.*, offering fewer recommendations can be acceptable if they are truthful). This effect is well illustrated by the increased precautions precision for the SGD algorithm with stacking by comparing rows 3 and 4 of Table 13 (from 79.82% to 84.34% for precautions).

Finally, Fig. 4 shows the evolution over time of the accuracy of the stacked RF streaming classifier, whose cold start ends by the thousandth sample. The performance level is satisfactory to produce usable indicators even for inexpert users. Fig. 5 represents a possible view of how these indicators would be used in an investment application. In this example, each entry consists of a text segment, its tickers, and an icon symbolising predicted investment emotion, for opportunity and for precaution.

5. Conclusions

Microblogging platforms such as Twitter provide valuable information for market screening and financial models. They often carry

[33] Note that `min_df` = 0.001 in Listing 1 after hyperparameter optimisation.

Fig. 5. Possible integration of the system in a mobile application.

tractable knowledge on investment options that either influence the market or react to it in real-time. Quite often, they include positive and negative educated forecasts, which we term opportunities and precautions in our emotion analysis.

Motivated by these facts and the interest in helping investors in their decision processes, we propose a novel TABEA system that combines NLP techniques and Machine Learning algorithms to predict financial emotions. More in detail, we rely on sophisticated linguistic features, such as sentiment and emotion lexica, along with quantitative financial data and specialised bags of words, and apply them in streaming Machine Learning stacked classification models that process a continuous flow of tweets on-the-fly. This approach is a novel contribution to this work.

The final system achieves over 90% precision when discerning financial opportunities and precautions on Twitter. The TABEA approach considers relevant text segments by focusing on specific assets and provides insightful indicators, as illustrated in Fig. 5.

In future work, we plan to extend our system to a multilingual framework, including English, taking into account that financial jargon is rather technical and very similar in Spanish and other languages, so language-dependent aspects should be easy to handle. We will also further analyse neutrality and ambivalence handling, as well as the potential of our architecture for automatic explainability since the RF algorithms we found so successful are intrinsically adequate to support natural language descriptions of the inner mechanisms of their predictions.

CRediT authorship contribution statement

Silvia García-Méndez: Conceptualization, Methodology, Software, Validation, Formal analysis, Investigation, Data curation, Writing – original draft. **Francisco de Arriba-Pérez:** Conceptualization, Methodology, Software, Validation, Formal analysis, Investigation, Data curation, Writing – original draft. **Ana Barros-Vila:** Software, Validation, Investigation, Resources, Data curation, Writing – review & editing. **Francisco J. González-Castaño:** Conceptualization, Methodology, Data curation, Writing – review & editing, Supervision.

Declaration of competing interest

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

Data availability

Data will be made available on request.

Acknowledgements

This work was partially supported by Xunta de Galicia, Spain grants ED481B-2021-118 and ED481B-2022-093, Spain; and University of Vigo/CISUG for open access charges.

References

Ahn, Y., & Kim, D. (2021). Emotional trading in the cryptocurrency market. *Finance Research Letters*, *42*, Article 101912. http://dx.doi.org/10.1016/j.frl.2020.101912.

Akhtar, F., & Das, N. (2019). Predictors of investment intention in Indian stock markets. *International Journal of Bank Marketing*, *37*(1), 97–119. http://dx.doi.org/10.1108/IJBM-08-2017-0167.

Akhtar, M. S., Ekbal, A., & Cambria, E. (2020). How intense are you? Predicting intensities of emotions and sentiments using stacked ensemble. *IEEE Computational Intelligence Magazine*, *15*(1), 64–75. http://dx.doi.org/10.1109/MCI.2019.2954667.

Al-Smadi, M., Al-Ayyoub, M., Jararweh, Y., & Qawasmeh, O. (2019). Enhancing aspect-based sentiment analysis of Arabic hotels' reviews using morphological, syntactic and semantic features. *Information Processing & Management*, *56*(2), 308–319. http://dx.doi.org/10.1016/j.ipm.2018.01.006.

Alamoudi, E. S., & Alghamdi, N. S. (2021). Sentiment classification and aspect-based sentiment analysis on yelp reviews using deep learning and word embeddings. *Journal of Decision Systems*, *30*(2–3), 259–281. http://dx.doi.org/10.1080/12460125.2020.1864106.

Carrillo-de Albornoz, J., Rodríguez Vidal, J., & Plaza, L. (2018). Feature engineering for sentiment analysis in e-health forums. *PLoS One*, *13*(11), Article e0207996. http://dx.doi.org/10.1371/journal.pone.0207996.

An, W., Tian, F., Chen, P., & Zheng, Q. (2022). Aspect-based sentiment analysis with heterogeneous graph neural network. *IEEE Transactions on Computational Social Systems*, 1–10. http://dx.doi.org/10.1109/TCSS.2022.3148866.

Plaza-del Arco, F. M., Molina-González, M. D., Ureña-López, L. A., & Martín-Valdivia, M. T. (2021). Comparing pre-trained language models for spanish hate speech detection. *Expert Systems with Applications*, *166*, Article 114120. http://dx.doi.org/10.1016/j.eswa.2020.114120.

Arulmurugan, R., Sabarmathi, K. R., & Anandakumar, H. (2019). Classification of sentence level sentiment analysis using cloud machine learning techniques. *Cluster Computing*, *22*(S1), 1199–1209. http://dx.doi.org/10.1007/s10586-017-1200-1.

Baier, L., Reimold, J., & Kuhl, N. (2020). Handling concept drift for predictions in business process mining. In *Procceedings of the IEEE conference on business informatics, vol. 1* (pp. 76–83). IEEE, http://dx.doi.org/10.1109/CBI49978.2020.00016.

Barrón Estrada, M. L., Zatarain Cabada, R., Oramas Bustillos, R., & Graff, M. (2020). Opinion mining and emotion recognition applied to learning environments. *Expert Systems with Applications*, *150*, Article 113265. http://dx.doi.org/10.1016/j.eswa.2020.113265.

Bee, M., Hambuckers, J., & Trapin, L. (2021). Estimating large losses in insurance analytics and operational risk using the g-and-h distribution. *Quantitative Finance*, *21*(7), 1207–1221. http://dx.doi.org/10.1080/14697688.2020.1849778.

Benllarch, M., Benhaddi, M., & El Hadaj, S. (2021). Enhanced Hoeffding anytime tree: A real-time algorithm for early prediction of heart disease. *International Journal on Artificial Intelligence Tools*, *30*(03), Article 2150010. http://dx.doi.org/10.1142/S021821302150010X.

Berrar, D. (2019). Bayes' theorem and Naive Bayes classifier. In *Encyclopedia of bioinformatics and computational biology, vol. 1–3* (pp. 403–412). Elsevier, http://dx.doi.org/10.1016/B978-0-12-809633-8.20473-1.

Cambria, E., Liu, Q., Decherchi, S., Xing, F., & Kwok, K. (2022). Senticnet 7: a commonsense-based neurosymbolic AI framework for explainable sentiment analysis. In *Proceedings of the international conference on language resources and evaluation*. European Language Resources Association.

Chatzis, S. P., Siakoulis, V., Petropoulos, A., Stavroulakis, E., & Vlachogiannakis, N. (2018). Forecasting stock market crisis events using deep and statistical machine learning techniques. *Expert Systems with Applications*, *112*, 353–371. http://dx.doi.org/10.1016/j.eswa.2018.06.032.

Chen, C. H., Lee, W. P., & Huang, J. Y. (2018). Tracking and recognizing emotions in short text messages from online chatting services. *Information Processing & Management*, *54*(6), 1325–1344. http://dx.doi.org/10.1016/j.ipm.2018.05.008.

Chun, J., Ahn, J., Kim, Y., & Lee, S. (2021). Using deep learning to develop a stock price prediction model based on individual investor emotions. *Journal of Behavioral Finance*, *22*(4), 480–489. http://dx.doi.org/10.1080/15427560.2020.1821686.

Chung, W., & Zeng, D. (2020). Dissecting emotion and user influence in social media communities: An interaction modeling approach. *Information & Management*, *57*(1), Article 103108. http://dx.doi.org/10.1016/j.im.2018.09.008.

Da'u, A., Salim, N., Rabiu, I., & Osman, A. (2020). Recommendation system exploiting aspect-based opinion mining with deep learning method. *Information Sciences*, *512*, 1279–1292. http://dx.doi.org/10.1016/j.ins.2019.10.038.

De Arriba-Pérez, F., García-Méndez, S., Regueiro-Janeiro, J. Á., & González-Castaño, F. J. (2020). Detection of financial opportunities in micro-blogging data with a stacked classification system. *IEEE Access*, *8*, 215679–215690. http://dx.doi.org/10.1109/ACCESS.2020.3041084.

Dogra, V., Singh, A., Verma, S., Kavita, Jhanjhi, N. Z., & Talib, M. N. (2021). *Lecture notes in networks and systems*: *vol. 248*, *Analyzing DistilBERT for sentiment classification of banking financial news* (pp. 501–510). Springer, http://dx.doi.org/10.1007/978-981-16-3153-5_53.

Dridi, A., Atzeni, M., & Reforgiato Recupero, D. (2019). FineNews: fine-grained semantic sentiment analysis on financial microblogs and news. *International Journal of Machine Learning and Cybernetics*, *10*(8), 2199–2207. http://dx.doi.org/10.1007/s13042-018-0805-x.

Duxbury, D., Gärling, T., Gamble, A., & 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. http://dx.doi.org/10.1080/1351847X.2020.1742758.

Elnagar, A., Khalifa, Y. S., & Einea, A. (2018). *Studies in computational intelligence*, *Hotel Arabic-reviews dataset construction for sentiment analysis applications* (pp. 35–52). Springer, http://dx.doi.org/10.1007/978-3-319-67056-0_3.

Fernández-Gavilanes, M., Juncal-Martínez, J., García-Méndez, S., Costa-Montenegro, E., & González-Castaño, F. J. (2018). Creating emoji lexica from unsupervised sentiment analysis of their descriptions. *Expert Systems with Applications*, *103*, 74–91. http://dx.doi.org/10.1016/j.eswa.2018.02.043.

Gao, B. (2021). The use of machine learning combined with data mining technology in financial risk prevention. *Computational Economics*, *1*, 1–21. http://dx.doi.org/10.1007/s10614-021-10101-0.

García-Méndez, S., Fernández-Gavilanes, M., Costa-Montenegro, E., Juncal-Martínez, J., & González-Castaño, F. J. (2018). Automatic natural language generation applied to alternative and augmentative communication for online video content services using SimpleNLG for spanish. In *Proceedings of the internet of accessible things* (pp. 1–4). ACM, http://dx.doi.org/10.1145/3192714.3192837.

García-Méndez, S., Fernández-Gavilanes, M., Costa-Montenegro, E., Juncal-Martínez, J., & Javier González-Castaño, F. (2019). A library for automatic Natural Language Generation of Spanish texts. *Expert Systems with Applications*, *120*, 372–386. http://dx.doi.org/10.1016/j.eswa.2018.11.036.

Gaur, M., Faldu, K., & Sheth, A. (2021). Semantics of the black-box: Can knowledge graphs help make deep learning systems more interpretable and explainable? *IEEE Internet Computing*, *25*(1), 51–59. http://dx.doi.org/10.1109/MIC.2020.3031769.

Ge, Y., Qiu, J., Liu, Z., Gu, W., & Xu, L. (2020). Beyond negative and positive: Exploring the effects of emotions in social media during the stock market crash. *Information Processing & Management*, *57*(4), Article 102218. http://dx.doi.org/10.1016/j.ipm.2020.102218.

Guo, J. (2022). Deep learning approach to text analysis for human emotion detection from big data. *Journal of Intelligent Systems*, *31*(1), 113–126. http://dx.doi.org/10.1515/jisys-2022-0001.

Hemmatian, F., & Sohrabi, M. K. (2019). A survey on classification techniques for opinion mining and sentiment analysis. *Artificial Intelligence Review*, *52*(3), 1495–1545. http://dx.doi.org/10.1007/s10462-017-9599-6.

Hew, K. F., Hu, X., Qiao, C., & Tang, Y. (2020). What predicts student satisfaction with MOOCs: A gradient boosting trees supervised machine learning and sentiment analysis approach. *Computers & Education*, *145*, Article 103724. http://dx.doi.org/10.1016/j.compedu.2019.103724.

Houlihan, P., & Creamer, G. G. (2021). Leveraging social media to predict continuation and reversal in asset prices. *Computational Economics*, *57*(2), 433–453. http://dx.doi.org/10.1007/s10614-019-09932-9.

Jiménez, S., González, F. A., Gelbukh, A., & Dueñas, G. (2019). word2set: WordNet-based word representation rivaling neural word embedding for lexical similarity and sentiment analysis. *IEEE Computational Intelligence Magazine*, *14*(2), 41–53. http://dx.doi.org/10.1109/MCI.2019.2901085.

Kang, M., Ahn, J., & Lee, K. (2018). Opinion mining using ensemble text hidden Markov models for text classification. *Expert Systems with Applications*, *94*, 218–227. http://dx.doi.org/10.1016/j.eswa.2017.07.019.

Kaur, C. (2020). Sentiment analysis of tweets on social issues using machine learning approach. *International Journal of Advanced Trends in Computer Science and Engineering*, *9*(4), 6303–6311. http://dx.doi.org/10.30534/ijatcse/2020/310942020.

Khan, W., Ghazanfar, M. A., Azam, M. A., Karami, A., Alyoubi, K. H., & Alfakeeh, A. S. (2020). Stock market prediction using machine learning classifiers and social media news. *Journal of Ambient Intelligence and Humanized Computing*, *1*, 1–24. http://dx.doi.org/10.1007/s12652-020-01839-w.

Kilicoglu, H., Rosemblat, G., Hoang, L., Wadhwa, S., Peng, Z., Malički, M., Schneider, J., & ter Riet, G. (2021). Toward assessing clinical trial publications for reporting transparency. *Journal of Biomedical Informatics*, *116*, 103717–103727. http://dx.doi.org/10.1016/j.jbi.2021.103717.

Kim, H. M., Bock, G. W., & Lee, G. (2021). Predicting Ethereum prices with machine learning based on blockchain information. *Expert Systems with Applications*, *184*, Article 115480. http://dx.doi.org/10.1016/j.eswa.2021.115480.

Ko, J., & Comuzzi, M. (2021). *Lecture notes in business information processing*: *vol. 406*, *Online anomaly detection using statistical leverage for streaming business process events* (pp. 193–205). Springer, http://dx.doi.org/10.1007/978-3-030-72693-5_15.

Ko, J., & Comuzzi, M. (2022). Keeping our rivers clean: Information-theoretic online anomaly detection for streaming business process events. *Information Systems*, *104*, Article 101894. http://dx.doi.org/10.1016/j.is.2021.101894.

Krippendorff, K. (2018). *Content analysis: An introduction to its methodology* (p. 472). SAGE Publications, http://dx.doi.org/10.1177/1094428108324513.

Li, W., Guo, K., Shi, Y., Zhu, L., & Zheng, Y. (2018). DWWP: Domain-specific new words detection and word propagation system for sentiment analysis in the tourism domain. *Knowledge-Based Systems*, *146*, 203–214. http://dx.doi.org/10.1016/j.knosys.2018.02.004.

Li, X., & Tang, P. (2020). Stock index prediction based on wavelet transform and FCD-MLGRU. *Journal of Forecasting*, *39*(8), 1229–1237. http://dx.doi.org/10.1002/for.2682.

Li, T., van Dalen, J., & van Rees, P. J. (2018). More than just noise? Examining the information content of stock microblogs on financial markets. *Journal of Information Technology*, *33*(1), 50–69. http://dx.doi.org/10.1057/s41265-016-0034-2.

Li, Q., Xu, Z., Shen, X., & Zhong, J. (2021). Predicting business risks of commercial banks based on BP-GA optimized model. *Computational Economics*, *1*, 1–19. http://dx.doi.org/10.1007/s10614-020-10088-0.

Liang, B., Su, H., Gui, L., Cambria, E., & Xu, R. (2022). Aspect-based sentiment analysis via affective knowledge enhanced graph convolutional networks. *Knowledge-Based Systems*, *235*, Article 107643. http://dx.doi.org/10.1016/j.knosys.2021.107643.

Liao, W., Zeng, B., Yin, X., & Wei, P. (2021). An improved aspect-category sentiment analysis model for text sentiment analysis based on RoBERTa. *Applied Intelligence*, *51*(6), 3522–3533. http://dx.doi.org/10.1007/s10489-020-01964-1.

Liu, Z., Wang, J., Du, X., Rao, Y., & Quan, X. (2021). GSMNet: Global semantic memory network for aspect-level sentiment classification. *IEEE Intelligent Systems*, *36*(5), 122–130. http://dx.doi.org/10.1109/MIS.2020.3042253.

Macdonald, S., & Birdi, B. (2019). The concept of neutrality: a new approach. *Journal of Documentation*, *76*(1), 333–353. http://dx.doi.org/10.1108/JD-05-2019-0102.

Madani, Y., Erritali, M., Bengourram, J., & Sailhan, F. (2020). A hybrid multilingual fuzzy-based approach to the sentiment analysis problem using SentiWordNet. *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems*, *28*(03), 361–390. http://dx.doi.org/10.1142/S0218488520500154.

Madhu, S. (2018). An approach to analyze suicidal tendency in blogs and tweets using sentiment analysis. *International Journal of Scientific Research in Computer Science and Engineering*, *6*(4), 34–36. http://dx.doi.org/10.26438/ijsrcse/v6i4.3436.

Mai, F., Shan, Z., Bai, Q., Wang, X. S., & Chiang, R. H. (2018). How does social media impact bitcoin value? A test of the silent majority hypothesis. *Journal of Management Information Systems*, *35*(1), 19–52. http://dx.doi.org/10.1080/07421222.2018.1440774.

Malmasi, S., & Dras, M. (2018). Native language identification with classifier stacking and ensembles. *Computational Linguistics*, *44*, 403–446. http://dx.doi.org/10.1162/coli_a_00323.

Matsumoto, K., Sasayama, M., Yoshida, M., Kita, K., & Ren, F. (2022). Emotion analysis and dialogue breakdown detection in dialogue of chat systems based on deep neural networks. *Electronics*, *11*(5), 695. http://dx.doi.org/10.3390/electronics11050695.

Mehlig, B. (2021). Stochastic gradient descent. In *Machine learning with neural networks, vol. 1* (pp. 96–113). Cambridge University Press, http://dx.doi.org/10.1017/9781108860604.006.

Mehmood, A., On, B. W., Lee, I., Ashraf, I., & Sang Choi, G. (2018). Spam comments prediction using stacking with ensemble learning. *Journal of Physics*, *933*, Article 012012. http://dx.doi.org/10.1088/1742-6596/933/1/012012.

Mohawesh, R., Tran, S., Ollington, R., & Xu, S. (2021). Analysis of concept drift in fake reviews detection. *Expert Systems with Applications*, *169*, Article 114318. http://dx.doi.org/10.1016/j.eswa.2020.114318.

Montiel, J., Read, J., Bifet, A., & Abdessalem, T. (2018). Scikit-multiflow: A multi-output streaming framework. *Journal of Machine Learning Research*, *19*, 1–5. http://dx.doi.org/10.5555/3291125.3309634.

Mowlaei, M. E., Saniee Abadeh, M., & Keshavarz, H. (2020). Aspect-based sentiment analysis using adaptive aspect-based lexicons. *Expert Systems with Applications*, *148*, Article 113234. http://dx.doi.org/10.1016/j.eswa.2020.113234.

Mumtaz, D., & Ahuja, B. (2018). *Studies in computational intelligence*, *A lexical and machine learning-based hybrid system for sentiment analysis* (pp. 165–175). Springer, http://dx.doi.org/10.1007/978-981-10-4555-4_11.

Muneer, A., & Fati, S. M. (2020). A comparative analysis of machine learning techniques for cyberbullying detection on Twitter. *Future Internet*, *12*(11), 187. http://dx.doi.org/10.3390/fi12110187.

Nagarajan, S. M., & Gandhi, U. D. (2019). Classifying streaming of Twitter data based on sentiment analysis using hybridization. *Neural Computing and Applications*, *31*(5), 1425–1433. http://dx.doi.org/10.1007/s00521-018-3476-3.

Nandwani, P., & Verma, R. (2021). A review on sentiment analysis and emotion detection from text. *Social Network Analysis and Mining*, *11*(1), 81. http://dx.doi.org/10.1007/s13278-021-00776-6.

Nawangsari, R. P., Kusumaningrum, R., & Wibowo, A. (2019). Word2Vec for Indonesian sentiment analysis towards hotel reviews: An evaluation study. *Procedia Computer Science*, *157*, 360–366. http://dx.doi.org/10.1016/j.procs.2019.08.178.

Nilashi, M., Minaei-Bidgoli, B., Alrizq, M., Alghamdi, A., Alsulami, A. A., Samad, S., & Mohd, S. (2021). An analytical approach for big social data analysis for customer decision-making in eco-friendly hotels. *Expert Systems with Applications*, *186*, Article 115722. http://dx.doi.org/10.1016/j.eswa.2021.115722.

Nurifan, F., Sarno, R., & Sungkono, K. (2019). Aspect based sentiment analysis for restaurant reviews using hybrid ELMoWikipedia and hybrid expanded opinion lexicon-SentiCircle. *International Journal of Intelligent Engineering and Systems*, *12*(6), 47–58. http://dx.doi.org/10.22266/ijies2019.1231.05.

Ortiz-Martínez, D. (2016). Online learning for statistical machine translation. *Computational Linguistics*, *42*, 121–161. http://dx.doi.org/10.1162/COLI_a_00244.

Ozyurt, B., & Akcayol, M. A. (2021). A new topic modeling based approach for aspect extraction in aspect based sentiment analysis: SS-LDA. *Expert Systems with Applications*, *168*, Article 114231. http://dx.doi.org/10.1016/j.eswa.2020.114231.

Pai, P. F., & Liu, C. H. (2018). Predicting vehicle sales by sentiment analysis of Twitter data and stock market values. *IEEE Access*, *6*, 57655–57662. http://dx.doi.org/10.1109/ACCESS.2018.2873730.

Parmar, A., Katariya, R., & Patel, V. (2019). A review on random forest: an ensemble classifier. In *Lecture notes on data engineering and communications technologies*: *vol. 26*, (pp. 758–763). Springer, http://dx.doi.org/10.1007/978-3-030-03146-6_86.

Peng, Y., Xiao, T., & Yuan, H. (2022). Cooperative gating network based on a single BERT encoder for aspect term sentiment analysis. *Applied Intelligence*, *52*(5), 5867–5879. http://dx.doi.org/10.1007/s10489-021-02724-5.

Phan, H. T., Nguyen, N. T., & Hwang, D. (2022). Convolutional attention neural network over graph structures for improving the performance of aspect-level sentiment analysis. *Information Sciences*, *589*, 416–439. http://dx.doi.org/10.1016/j.ins.2021.12.127.

Plutchik, R. (2004). The circumplex as a general model of the structure of emotions and personality. In *Circumplex models of personality and emotions* (pp. 17–45). Washington: American Psychological Association, http://dx.doi.org/10.1037/10261-001.

Polignano, M., Narducci, F., de Gemmis, M., & Semeraro, G. (2021). Towards emotion-aware recommender systems: an affective coherence model based on emotion-driven behaviors. *Expert Systems with Applications*, *170*, Article 114382. http://dx.doi.org/10.1016/j.eswa.2020.114382.

Pugliese, V. U., Costa, R. D., & Hirata, C. M. (2021). *Lecture notes in business information processing*: *vol. 417*, *Comparative evaluation of the supervised machine learning classification methods and the concept drift detection methods in the financial business problems* (pp. 268–292). Springer, http://dx.doi.org/10.1007/978-3-030-75418-1_13.

Qaiser, S., & Ali, R. (2018). Text mining: Use of TF-IDF to examine the relevance of words to documents. *International Journal of Computer Applications*, *181*(1), 25–29. http://dx.doi.org/10.5120/ijca2018917395.

Rash, J. A., Prkachin, K. M., Solomon, P. E., & Campbell, T. S. (2019). Assessing the efficacy of a manual-based intervention for improving the detection of facial pain expression. *European Journal of Pain*, *23*(5), 1006–1019. http://dx.doi.org/10.1002/ejp.1369.

Razi, N. I. M., Othman, M., & Yaacob, H. (2017). Investment decisions based on EEG emotion recognition. *Advanced Science Letters*, *23*(11), 11345–11349. http://dx.doi.org/10.1166/asl.2017.10280.

Rhanoui, M., Mikram, M., Yousfi, S., & Barzali, S. (2019). A CNN-BiLSTM model for document-level sentiment analysis. *Machine Learning and Knowledge Extraction*, *1*(3), 832–847. http://dx.doi.org/10.3390/make1030048.

Ronaghi, F., Salimibeni, M., Naderkhani, F., & Mohammadi, A. (2022). COVID-19 adopted hybrid and parallel deep information fusion framework for stock price movement prediction. *Expert Systems with Applications*, *187*, Article 115879. http://dx.doi.org/10.1016/j.eswa.2021.115879.

Rout, J. K., Choo, K. K. R., Dash, A. K., Bakshi, S., Jena, S. K., & Williams, K. L. (2018). A model for sentiment and emotion analysis of unstructured social media text. *Electronic Commerce Research*, *18*(1), 181–199. http://dx.doi.org/10.1007/s10660-017-9257-8.

Salminen, J., Almerekhi, H., Kamel, A. M., Jung, S. G., & Jansen, B. J. (2019). Online hate ratings vary by extremes. In *Proceedings of the conference on human information interaction and retrieval* (pp. 213–217). Association for Computational Linguistics, http://dx.doi.org/10.1145/3295750.3298954.

Seité, S., Khammari, A., Benzaquen, M., Moyal, D., & Dréno, B. (2019). Development and accuracy of an artificial intelligence algorithm for acne grading from smartphone photographs. *Experimental Dermatology*, *28*(11), 1252–1257. http://dx.doi.org/10.1111/exd.14022.

Seth, S., Singh, G., & Chahal, K. K. (2021). Drift-based approach for evolving data stream classification in intrusion detection system. In *Proceedings of the workshop on computer networks & communications, vol. 2889* (pp. 23–30). CEUR.

Shahin, I., Hindawi, N., Nassif, A. B., Alhudhaif, A., & Polat, K. (2022). Novel dual-channel long short-term memory compressed capsule networks for emotion recognition. *Expert Systems with Applications*, *188*, Article 116080. http://dx.doi.org/10.1016/j.eswa.2021.116080.

Simester, D., Timoshenko, A., & Zoumpoulis, S. I. (2019). Targeting prospective customers: Robustness of machine-learning methods to typical data challenges. *Management Science*, *66*, 1–43. http://dx.doi.org/10.1287/mnsc.2019.3308.

Sohangir, S., Wang, D., Pomeranets, A., & Khoshgoftaar, T. M. (2018). Big Data: Deep learning for financial sentiment analysis. *Journal of Big Data*, *5*(1), 3. http://dx.doi.org/10.1186/s40537-017-0111-6.

Taffler, R. (2018). Emotional finance: investment and the unconscious. *The European Journal of Finance*, *24*(7–8), 630–653. http://dx.doi.org/10.1080/1351847X.2017.1369445.

Theodorou, T. I., Zamichos, A., Skoumperdis, M., Kougioumtzidou, A., Tsolaki, K., Papadopoulos, D., Patsios, T., Papanikolaou, G., Konstantinidis, A., Drosou, A., & Tzovaras, D. (2021). An AI-enabled stock prediction platform combining news and social sensing with financial statements. *Future Internet*, *13*(6), 138. http://dx.doi.org/10.3390/fi13060138.

Trabelsi, A., Elouedi, Z., & Lefevre, E. (2019). Decision tree classifiers for evidential attribute values and class labels. *Fuzzy Sets and Systems*, *366*, 46–62. http://dx.doi.org/10.1016/j.fss.2018.11.006.

Tuke, J., Nguyen, A., Nasim, M., Mellor, D., Wickramasinghe, A., Bean, N., & Mitchell, L. (2020). Pachinko prediction: A Bayesian method for event prediction from social media data. *Information Processing & Management*, *57*(2), Article 102147. http://dx.doi.org/10.1016/j.ipm.2019.102147, arXiv:1809.08427.

Turchi, M., Negri, M., Farajian, M. A., & Federico, M. (2017). Continuous learning from human post-edits for neural machine translation. *The Prague Bulletin of Mathematical Linguistics*, *108*, 233–244. http://dx.doi.org/10.1515/pralin-2017-0023.

Valdivia, A., Luzón, M. V., Cambria, E., & Herrera, F. (2018). Consensus vote models for detecting and filtering neutrality in sentiment analysis. *Information Fusion*, *44*, 126–135. http://dx.doi.org/10.1016/j.inffus.2018.03.007.

Valle-Cruz, D., Fernández-Cortez, V., López-Chau, A., & Sandoval-Almazán, R. (2022). Does Twitter affect stock market decisions? Financial sentiment analysis during pandemics: A comparative study of the H1N1 and the COVID-19 periods. *Cognitive Computation*, *14*(1), 372–387. http://dx.doi.org/10.1007/S12559-021-09819-8/TABLES/17.

Wang, Y. (2017). Stock market forecasting with financial micro-blog based on sentiment and time series analysis. *Journal of Shanghai Jiaotong University*, *22*(2), 173–179. http://dx.doi.org/10.1007/s12204-017-1818-4.

Wang, Z., Ho, S. B., & Cambria, E. (2020). Multi-level fine-scaled sentiment sensing with ambivalence handling. *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems*, *28*(04), 683–697. http://dx.doi.org/10.1142/S0218488520500294.

Wang, Y., Liu, S., Li, S., Duan, J., Hou, Z., Yu, J., & Ma, K. (2019). Stacking-based ensemble learning of self-media data for marketing intention detection. *Future Internet*, *11*(7), 155. http://dx.doi.org/10.3390/fi11070155.

Wang, X., Zhou, T., Wang, X., & Fang, Y. (2022). Harshness-aware sentiment mining framework for product review. *Expert Systems with Applications*, *187*, Article 115887. http://dx.doi.org/10.1016/j.eswa.2021.115887.

Weichselbraun, A., Gindl, S., Fischer, F., Vakulenko, S., & Scharl, A. (2017). Aspect-based extraction and analysis of affective knowledge from social media streams. *IEEE Intelligent Systems*, *32*(3), 80–88. http://dx.doi.org/10.1109/MIS.2017.57.

Xiao, L., Xue, Y., Wang, H., Hu, X., Gu, D., & Zhu, Y. (2022). Exploring fine-grained syntactic information for aspect-based sentiment classification with dual graph neural networks. *Neurocomputing*, *471*, 48–59. http://dx.doi.org/10.1016/j.neucom.2021.10.091.

Yuan, H., Lau, R. Y., Wong, M. C., & Li, C. (2018). Mining emotions of the public from social media for enhancing corporate credit rating. In *Proceedings of the IEEE international conference on e-business engineering* (pp. 25–30). IEEE, http://dx.doi.org/10.1109/ICEBE.2018.00015.

Yuan, H., Xu, W., Li, Q., & Lau, R. (2018). Topic sentiment mining for sales performance prediction in e-commerce. *Annals of Operations Research*, *270*(1–2), 553–576. http://dx.doi.org/10.1007/s10479-017-2421-7.

Yue, L., Chen, W., Li, X., Zuo, W., & Yin, M. (2019). A survey of sentiment analysis in social media. *Knowledge and Information Systems*, *60*(2), 617–663. http://dx.doi.org/10.1007/s10115-018-1236-4.

Zhao, M., Yang, J., Zhang, J., & Wang, S. (2022). Aggregated graph convolutional networks for aspect-based sentiment classification. *Information Sciences*, *600*, 73–93. http://dx.doi.org/10.1016/j.ins.2022.03.082.

Zhou, Z., Xu, K., & Zhao, J. (2018). Tales of emotion and stock in China: volatility, causality and prediction. *World Wide Web*, *21*(4), 1093–1116. http://dx.doi.org/10.1007/s11280-017-0495-4.