Affective Concept-Based Encoding of Patient Narratives via Sentic Computing and Neural Networks

Hanane Grissette[1] · El Habib Nfaoui[1]

Abstract

The automatic generation of features without human intervention is the most critical task for biomedical sentiment analysis. Regarding the high dynamicity of shared patient narrative data, the lack of formal medical language sentiment dictionaries prevents retrieval of the appropriate sentiment, which is unapproachable and can be prone to annotator bias. We propose a novel affective biomedical concept-based encoding via sentic computing and neural networks. The main contributions include four aspects. First, a biomedical embedding, in which a medical entity is defined, normalized, and synthesized from a text, is built using online patient narratives after being combined with label propagation from a widely used comprehensive biomedical vocabulary. Second, considering the dependence on biomedical definitions, drug reaction sample selection based on general matching is suggested. These feature settings are then used to build and recognize affective semantics and sentics based on an extreme learning machine. Finally, a semisupervised LSTM-BiLSTM model for biomedical sentiment analysis is constructed. There was a massive influx of patient self-reports related to the COVID-19 pandemic. A study was conducted in this direction, and we tested the validity, medical language familiarity, and transferability of our approach by analyzing millions of COVID-19 tweets. Comparisons to affective lexicons also indicate that integrating extreme learning machine cognitive capabilities has advantages over biomedical sentiment analysis. By considering sentics vectors on top of the formed embeddings, our semisupervised LSTM-BiLSTM achieved an accuracy of 87.5%. The evaluations of unsupervised learning approximated the results of the previous model when dealing with a serious loss of biomedical data. In this paper, we demonstrate the effectiveness of integrating deep-learning-based cognitive capabilities for both enhancing distributed biomedical definitions and inferring sentiment compositions from many patient self-reports on social networks. The relevant encoding of affective information conveyed regarding medication subjects clearly reveals defined roles and expectations that can have a positive impact on public health.

Keywords Sentic computing · Pandemic COVID-19 · Biomedical sentiment analysis · Distributed biomedical vocabularies · Affective computing · Social networks

Introduction

Patient narratives posted on social networks and health-related platforms reflect a patient's health status in terms of observations and contain much objective information, such as descriptions of examinations and interventions. Sentiment analysis (SA) models are mostly used to evaluate the sentiments conveyed by patient narratives gathered from multiple information sources, determine positive or negative clinical outcomes and judge the impact of drugs and medical conditions. Their effectiveness depends mainly on how well they recognize entities, concepts, events, and their aspects.

Among the most successful SA models are those based on deep learning techniques. They perfectly perform within the setting in which they were trained, but their performance decreases sharply when they are used to analyze online patient stories. They often lose ground in the face of more flexible knowledge-based challenges [1]. This failure is due to many problems, particularly the lack of sentiment dictionaries in the bio-medical domain, which are

✉ Hanane Grissette
  hanane.grissette@usmba.ac.ma

  El Habib Nfaoui
  elhabib.nfaoui@usmba.ac.ma

1 LISAC Laboratory, Faculty of Sciences Dhar EL Mahraz,
  Sidi Mohamed Ben Abdellah University, Fez, Morocco

unapproachable and can be prone to annotator bias, and the lack of discriminatory techniques for medical data objects from online patient content.

The corresponding growth of this field has resulted in the emergence of various subareas, each addressing a different level of analysis regarding various aspects and targets (entities, events or persons) [2]. Recently, the domain of aspect-based sentiment analysis (AbSA), in which sentiments are analyzed, aggregated over time, and synthesized for given aspects, has received much attention, with increasing feedback on public health on social media [3, 4]. Issues and complex scenarios for AbSA have appeared regarding the very large number of medical concept citation medication-related reports that are affectively and semantically associated in different contexts. Notably, few AbSA approaches consider medication-related concepts and their aspects and hardly take into account affective levels regarding medical concepts vs. conditions [5, 6]. They mostly neglect unclear words and detect emotional valence only by scoring the text at the sentence- level or word- level by describing how words and phrases match the sentiment vocabulary.

Word-embedding techniques are widely used to capture richer contextual semantics because of their ability to learn deep patterns and relationships through distributed representations. Existing embeddings focus on the structure and meaning of words to encode the semantic and syntactic information of words that would be used as the input of neural networks in sentiment analysis tasks [7]. The best accuracy is achieved by exploring bidirectional contextual information from a large knowledge-based dataset since the accuracy of sentiment classification mainly relies on the quality of the annotation, word score, and entity tagging techniques [8–10]. Despite their success in modeling varied linguistic structures, these methods face many difficulties in capturing, extracting and quantifying bio-medical entities from unstructured text. Patient narratives can represent a combination of meaning, structure, and events. They might exhibit additional natural language challenges, such as excessive use of abbreviations, acronyms, alternatives, and misspellings of medical entity citations [11]. For example, in the dataset used in this study, we have found 1489 COVID-19-related spellings, e.g., "postcovid", "covidiots", "non-covid", and "covidale". In addition, the annotated corpora proposed in this area are very few. Moreover, current annotation systems primarily focus only on common concept types such as genes and diseases [12]. Consequently, the limited amount of training data available increases biomedical data loss. To overcome this limit, controlled medical languages and their data object components, which contain rich semantic information, are incorporated to build and learn medical word embeddings [12–14] from generated narratives.

This study addresses the following research questions: (1) How can the exactness of medical fact assessments be enhanced by studying varied medical concepts from many self-reported experiences? (2) What embedding strategy should be considered for deploying entity interactions, dependencies, contextual semantics, and their sentics for an improved biomedical sentiment analysis by considering the aspects in which they are extracted? (3) How can salient online information be identified to focus on a given person, target or entity?

In this paper, we propose a novel method that defines and normalizes biomedical concepts and events and synthesizes contextual semantics and their assembled sentics. Contextualized embedding is a critical step in biomedical sentiment analysis. To support the credibility and correctness of our biomedical representations, we use label propagation at the first model layer by integrating a large number of well-coordinated samples of highly controlled vocabulary from PubMed. Additionally, we enhance our connotation by integrating extreme learning machine cognitive capabilities into our neural network model. The obtained embeddings are then used to build and recognize affective semantics and sentics based on extreme learning machines. Finally, a semi-supervised LSTM-BiLSTM model for biomedical sentiment analysis is constructed.

The proposed model performance meets several requirements, such as the use of large and balanced labeled data for training models, which further yields good sentiment inference results. Efficient biomedical sentiment inference involves drug-related component distinction within text. In this regard, the proposed method has the advantage of considering drug reaction identification, which consists of natural biomedical concept comprehension and drug reaction extraction.

The remainder of this paper is structured as follows: "Background and Related Work" provides a summary of the literature review and the background of this study by illustrating the importance of including biomedical aspects in analyzing related medication text. "Proposed Approach" presents the proposed architecture with more problem definition details. "Validation and Results" describes large social network experimentation output and comparisons with state-of-the-art sentiment lexicons and sentence-level biomedical distributed systems. "Linguistic Gap Discussion" discusses some linguistic gaps. Finally, "Conclusion" concludes this paper.

Background and Related Work

This section presents relevant works related to sentiment analysis in bio-medical domain. First, we present the importance of extracting aspects and bio-medical entities for enhancing SA. Indeed, we review the relevant literature on the models applied upon the social media. Second, we cover the necessary distributed biomedical background for understanding the remainder of the paper, which are exploited to

build the embeddings. Finally, we introduce transfer learning theory adaptation for transferable characteristics of semantic models applied to biomedical sentiment annotation.

Affective Computing and Sentiment Analysis

The basic tasks of affective computing and sentiment analysis are emotion recognition and polarity detection [15]. These two tasks are highly interrelated. Indeed, [16] stated that "both the terms of affective computing and sentiment analysis relate to the computational interpretation and generation of human emotion or affect, whereas the former mainly relates to instantaneous emotional expressions and is more commonly associated with speech or image/video processing, the later mainly relates to longer-term opinions or attitudes and is more commonly associated with natural language processing."

Aspect-Based Sentiment Analysis

Sentiment analysis concerns the study of capturing, quantifying, and measuring dynamic public sentiments through different methods, tools and techniques. SA can be performed at three levels: at document level, at sentence level, and at aspect level.

An important part of research effort focuses on aspect-based SA. It provides more detailed information than general SA [17], which allows to formulate interactions, contextual dependencies, semantics, and relationships of different data objects, e.g., aspect, event, person, target, etc. Refers to [2], three processing steps can be distinguished when performing AbSA: identification, classification, and aggregation.

In bio-medical domain, supervised AbSA methods are the most accurate approaches [15], especially lexicon-based approaches, but many unseen entities still appeared regarding the social networks dynamicity. Although their effectiveness in selecting relevant and original features, they still remain not well-adapted to unlabeled examples [10]. Owing to the limited availability of lexicons that specifically offer medical contextual definitions and sentimental correlations between varied medical concepts, they fail to address affective aspects regarding biomedical contexts. Unsupervised AbSA methods are also used to discover medical patterns from real-time examples. The problem is that these methods might result in redundant and irrelevant features leading to a serious loss of biomedical data [1]. Aspect extraction (AE) is the most challenging task for biomedical analysis. Numerous contributions only consider explicit aspects such as drugs or disease where implicit aspects have been neglected. For example, [18] used two aspects: drugs and doctors over two different corpora extracted from the social web. Authors in [19] focused on detecting the mood of patients with cancer by analyzing their messages in online communities.

Despite multiple attempts, existing unsupervised approaches generally suffer from suboptimal entity recognition methods due to the frequent use of informal medical language, non-standard format, poor spelling and biomedical abbreviation forms, as well as typos in social media. Two main research questions have been involved: (i) how to perform the process of AE effectively regarding these unique characteristics of social networks toward specific biomedical targets ? and (ii) how to map relationships between different data objects for improved biomedical aspect-based sentiment analysis ?

The extraction of explicit/implicit aspects, medical entities, aspect categories of multi-word medical concepts may require involvement of external controlled medical knowledge and vocabularies. Refers to many studies [11, 20–22], this knowledge-based integration ensures and enhances the credibility and transparency of analysis conveyed toward these medical concepts patterns. Since it relies on sufficiently large biomedical training corpora to define the appropriate medical definitions, a deeper understanding of transfer learning and semisupervised learning becomes apparent.

Semisupervised learning is a fundamental solution that allows learning from heterogeneous unlabeled examples spread via social network in combination with typically sets of labeled data. In this sense, several studies used semisupervised techniques to primarily develop automated corpora [23] by utilizing word representation capabilities in two ways: (1) generating context-depend embeddings, (2) transferring knowledge from pretrained embeddings or fusing weighted distributed features [24–26]. The ultimate problem is that these corpora result in low medical entities recognition recall due to the failure of recognizing drug-related components and their aspects such as drug reactions and indications

Extracting drug-related explicit aspects may be failed due to the existence of adverse drug reaction (ADR) multi-word expressions, leading to a higher accumulation of false positives and negatives. Few studies have been carried out to consider side effects in raw text, i.e., annotation of ADR lexicons and corpora. Table 1 shows a list of baselines models used for drug reactions lexicons with relevant details. The problems with these lexicons are as follows: (1) They are annotated for local context and target. For example, [27] considered an adverse drug events benchmark reference dataset consisting of 5,645,336 English-language articles, but it regarded to specific targets that mentioned at least one of six medicinal products of interest such as insulin glargine, levetiracetam, methylphenidate, sorafenib, terbinafine, and zolpidem. (2) The major limits are concerned with the data. Twitter data are unbalanced regarding time, where there is no relationship with other drug-related concepts, such as indications from formal medical language. Word embedding features created from a large non-domain-specific

Table 1 An overview of baselines models used for drug reactions lexicon construction regarding varied benchmarks algorithms

ADR Lexicon	corpus	Algorithm	Accuracy (%)
[76]	Twitter Corpus	Multinomial logistic regression	69 %
		Bidirectional GRU + GoogleNews embedding	55 %
		hierarchical-based-Character LSTM (Char-LSTM)	47 %
		hierarchical-based-Character CNN (Char-CNN)	77 %
[77]	AskPatient corpus	CNN	78 %
		DNorm	73 %
		RNN	68 %
		Multi-class logistic regression	57 %

Twitter dataset have been proposed as a solution that may disambiguate medical-sense directions. (3) Another significant problem is how efficiently retrieve formal formats of drug reaction (combinations of words) from imbalanced and unstructured text. This explains why sentiment analysis systems are able to perform reasonably well towards given entities and events but far poorly on clarifying sentiment towards biomedical subjects. To overcome these limitations, a medical normalization step [12, 28, 29] is carried out by our system for efficient meaning disambiguation as a whole, then extract and define medical entities and aspects.

The main contributions of this article include the following three goals: (1) using semisupervised learning to unlock new paths towards conquering annotated sentiment lexicons deficiency for biomedical sentiment analysis, (2) building contextual semantics for aspect-based sentiment of a given concept or entity, and (3) investigating which way of tackling unstructured medication-related data for enhancing biomedical sentiment analysis training robustness, consistency, and reliability.

Biomedical Sentiment Analysis on Twitter

Patient-reported outcome measures are useful to evaluate patients health status and their experiences by documenting the impact of a disease on functional performance as well as the effectiveness of medical and rehabilitation interventions [30–34]. Nowadays, online social networks have emerged as a new platform that provides an arena for people to share their views and perspectives on different issues and subjects with their friends, family, relatives, etc [35]. Because of the emergence of deep learning techniques, the use of Twitter data has become valuable and popular in analyzing the sentimental attitudes and measuring patient-reported outcomes in recent years.

Traditional approaches, including dictionary-based and machine learning tools, are less efficient in terms of learning for biomedical sentiment analysis on Twitter. Supervised algorithms are commonly used including naive Bayes, K-nearest neighbors (KNN), logistic regression, support vector machine (SVM), and artificial neural networks (ANNs).

For example, [36] used Bayesian network classifier to perform Twitter Sentiment Analysis (TSA) on critical events such as natural disasters or social movements. The latter used the Bayes factor measure that improved the classification performance better than SVM and random forests, given a sufficient number of training examples. The most significant challenge for the machine learning approaches is feature selection, which is often domain dependent. They often result in a highly sparse input matrix and produce low accuracy measures in biomedical sub-tasks. To reduce sparsity, authors in [37] provided a novel supervised learning technique that identifies a lexicon set unique to the TSA. This identification process generates a collection of vectorized tweets for input to machine learning tools, which is further reduced into a small set of seven meta-features, but they lose word orders and long-term dependencies.

Deep learning methods have increased the accuracy of several sentiment analysis tasks, which offer a rich representation that allows identifying the relations among words, offering interesting qualitative information to historically, semantically, and contextually consider the main features. While these methods have been considered as state-of-the-art solutions, there still remain several challenges and limitations confronting their use on medication-related text analysis. Various contributions rely on combining techniques and fusing datasets to conquer most of these limitations and boost models performance [27, 37–39].

Moreover, Twitter data exhibit depend-platform properties or features, and many additional features have been considered in the literature. Prior studies propose to develop hybrid features regarding Twitter data particularities, which proves its subjectivity and effectiveness for polarity detection with regard to linguistic features such as stylistic part-of-speech [40]. [41] presented a pattern-matched Twitter analysis of US cancer-patient sentiments that mainly characterizes the content of tweets authored by patients with cancer in the USA. Indeed, they utilize patient tweets to compute the average happiness of patients with cancer for each cancer diagnosis. A large sample of English tweets was used and the tweets were filtered by the cancer diagnosis. The patients' Twitter identification numbers were used to

gather all the tweets for each patient where happiness values for the patients' tweets were calculated using a quantitative hedonometric analysis. Others proposed other salients such as [42] that used the citation count as the dependent variable (number of unique Twitter users), with 6,482,260 dependent tweets linked to 1,083,535 publications, which was considered an indicator of the regression model. It improved the adjusted R-squared value of regression analysis in several disciplines.

Biomedical Sentiment Annotation and Sentic Computing

There are currently numerous English-language sentiment knowledge bases and state-of-practice benchmarks such as SenticNet [43] and SentiWordNet [44], ANEW [45], and VADER [38]. Few medical sentiment-annotated lexicons are trained on medication-related examples with regard to cross-domain and cross-lingual measures, such as unprecedented shifts in patient attitudes [46]. In particular, biomedical sentiment analysis requires particular attention to be paid to the sentiment annotation scale regarding biomedical natural entities and aspects [47]. Indeed, limited labeled training data fails to cover the entire input space.

Automatic sentiment lexicon annotation techniques involve other substantial requirements such as aspect extraction and medical entity normalization [48], especially, to bridge the gap between patients' language and formal medical language. The multidimensional way of mining their feelings and their sufferings concerning treatments and symptoms is of critical importance. This task is extremely difficult, as it involves a deep understanding of most of the explicit and implicit meanings. It mainly relies on parts of the text and context in which sentiment aspects are explicitly expressed, such as polar terms, emotion words, and co-occurrence frequencies. Indeed, patient opinions are not just reviews of a product/service but more like small donations of experience [34, 49]. In addition, the implicit meanings are important. Referring to many studies [8], opinions and sentiments are often conveyed implicitly through latent semantics, which makes purely syntactical approaches ineffective, especially when aiming to find new exciting outcomes hidden in the unstructured narratives on social networks.

Sentic computing is one of the most crucial techniques used for this purpose [15, 50] and has been exploiting in many patient-centered applications [51–56], which exploits semantics and their sentics based on the relation dependency, and linguistics properties to annotate automatically concept-level sentiment of an input phrase. Where semantics are concepts that are most semantically-related to the input concept, and sentics are emotion categorization values expressed in terms of varied affective dimensions such as pleasantness, attention, sensitivity, and aptitude.

Since polarity is strongly connected to attitudes and feelings and involves quantifying the affective reaction of a human to a document, a media item or an event, we proposed to use sentic computing to encode contextual semantics and sentics to synthesize the full range of emotional medication experiences in terms of pleasantness, attention, sensitivity, and aptitude. Then after, these generated former sentics is efficiently deployed using our extreme learning machine model, and therefore a better understanding of the contextual role of each concept within a sentence is achieved in this study. Indeed, the proposed model leverages semantic and contextual relatedness to deeply understand the world and social aspects, extract medical components, and define cultural awareness and common-sense knowledge.

Biomedical Distributed Representations

Embeddings have received substantial attention because of their capabilities to capture the semantics of biomedical concepts. To date, methods have rapidly adopted word embeddings and combined different types of medical dictionaries from controlled biomedical ontologies, vocabularies and medical resources, such as PubMed[1] and MedLine[2] databases. Many studies have contributed to revolutionizing the use of word embedding techniques to represent biomedical text as a reduced dimensional vector. Primary methods trained the embeddings based on (1) the average surrounding context words, such as the continuous bag-of-words (CBOW) model in Word2Vec [57], (2) weighted context words, such as the Skip-Gram model in Word2Vec, (3) global cooccurrence statistics, such as GloVe [58], and (4) word n-grams, such as fastText [59]. For example, authors in [13] applied Word2Vec to a corpus of 10,876,004 English abstracts of biomedical articles from PubMed to estimate the relatedness of two words. Also, authors in [60] provided biomedical word embeddings that combine subword information from unlabeled biomedical text with a widely used biomedical controlled vocabulary called medical subject headings (MeSH). However, word embeddings are traditionally computed at the word level from a large corpus of unlabeled text, ignoring the information present in the internal structure of words or any information available in domain-specific structured resources such as ontologies. Here these knowledges hold the potential for significantly improving the quality of word representations, as suggested in some recent studies in the public domain.

Recently, the use of embedding with advanced computational linguistics techniques has shown promising performance in biomedical concept extraction and tagging

---

[1] https://PubMed.ncbi.nlm.nih.gov/

[2] https://www.nlm.nih.gov/bsd/pmresources.html

[12], which employed other tools such as PubTator [29] (a state-of-the-art named entity recognition system), N-gram counts, word sequence occurrences, word sequence probabilities [61], and medical patterns of word semantics. Many pre-trained embeddings require no human-labeled data but they make use of resources for weak supervision (e.g., UMLS semantic types). They have been shown to speed up manual data curation by providing human annotators with automated pre-annotations generated by rules or machine learning models. Authors in [12] deployed PubTator via four different word embedding techniques on 30 million Pub-Med abstracts. It is the largest among the publicly available biomedical concept embeddings to date. It covers over 400,000 biomedical concepts mentioned in the literature and journal life pharma sciences. Additionally, many researchers have built continuous vectors space by using a training set of sentence embeddings with over 30 million documents from both scholarly articles in PubMed and clinical notes from MIMIC-III clinical Database [14]. Their purpose is to prove the effectiveness of discriminating expression features by employing pair similarity tasks in different biomedical text genres. For example, in [62], resources were derived from the combination of all publication abstracts from Pub-Med and all full-text documents from the PubMed Central[3] (PMC) open access subset and others. Our method is implemented on the top of training samples as a combination from PubMed and MIMIC-III clinical Notes.

Patient narratives on social network might contain different spellings of varied bio-medical entities such as genes, drug-interactions, and diseases. The capture of the semantic relatedness and contextual visibility of these bio-medical entities plays a vital role in retrieving hidden information and clarifying meanings and their attached sentiments. Various contributions have been done in this context: (1) bag of meta-words [63], (2) improved embeddings [64], (3) and one-hot character vectors for aspect-based sentiment analysis [65]. Deep learning models have slightly improved these representations performance. For example, authors in [65] joined multiple convolutions to generate suitable word representation for typical settings and tasks. Other distributions have combined unbiased representations from CNN with semantics from multiple BiLSTMs for generating contextual vectors [66]. Even if most of them got good accuracy, they are not robust because (i) medication-related texts require learning from real-world data, (ii) they present additional language challenges that may require processing by a word-expert annotation system for efficient meaning disambiguation, especially for reducing biomedical data loss. To conquer these limitations, we attempt to define suitable transfer learning setting over the biomedical resource combination

assessed through the experiments, and therefore this setting fit our extreme learning machines to define biomedical patterns on social networks.

Transfer Learning

One of the most powerful ideas in deep learning is that sometimes you can take the knowledge that the neural network has learned from one task and apply it to a separate task. This process is called transfer learning. Transfer learning makes sense when you have a large amount of data for the problem you are transferring from and usually relatively less data for the problem you are transferring to. Transfer learning is widespread in deep learning, given the enormous resources required to train deep learning models or the vast and challenging datasets on which deep learning models are trained. Transfer learning can be categorized under one of three subcategories: inductive transfer learning, transductive transfer learning, and unsupervised transfer learning. Each subcategory is based on different situations between the source and target domains and tasks [67]. Many contributions have been made in this scope. Instance transfer approaches aim to reweight some labeled data in the source domain for use in the target domain [68–70]. Feature representation transfer approaches have been used to find a good and reliable feature representation that reduces the difference between the source and the target domains and the error of classification and regression models [71, 72]. Relational knowledge transfer builds a map of the relational knowledge between the source domain and the target domains [73, 74], with both areas being relational domains and the independent and identically distributed (i.i.d.) assumption being relaxed in each field.

In the past few years, deep learning has seen incredible progress and has largely removed the requirement of strong domain knowledge. However, existing approaches in natural language processing still require task-specific modifications and training from scratch [75]. Knowledge transfer with deep neural networks offers vast improvements to many challenging tasks, e.g., language understanding, word embedding, and sentiment classification. Sentiment classification has seen full applications because of its transfer learning capability to transfer knowledge or lexicons from one domain to other domains. It has played a vital role in improving various meta-word dictionaries by expanding pretrained word embedding vectors such as Word2Vec and GloVe. In our study, we apply a semisupervised feature representation transfer approach to the inductive transfer learning mechanism, whereby we aim to find distributional feature representations under medical settings that minimize domain divergence and classification model error. Technically, the source and target domains are different, but they are dedicated to the same task. Indeed,
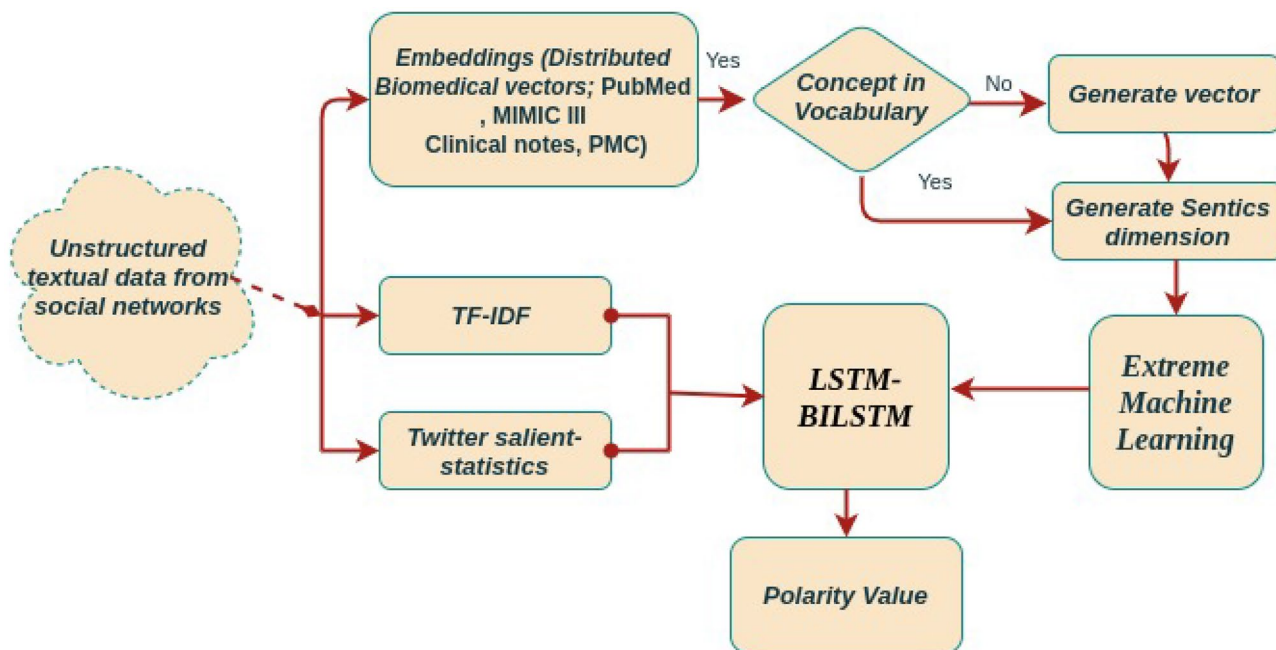
---

[3] https://www.ncbi.nlm.nih.gov/pmc/

Fig. 1 The overall architecture of the proposed system

no labeled sentiment dictionaries in the medical context data are available; thus, we adopted a massive amount of labeled data in a general context. To this end, the source domain is represented under the following settings:

$\mathcal{D} = \{\mathcal{F}_S, P(\mathcal{F}_S)\}$; where $\mathcal{F}_S$: *Feature Space* and $P(\mathcal{F}_S)$: distribution over the feature space.

Transductive transfer learning must satisfy the following criteria:

Rule (*i*) **:** the feature spaces between the source and target domains are different, $\mathcal{F}_S \neq \mathcal{F}_T$
Rule (*ii*) **:** the marginal probability distributions of the input data are different, $P(\mathcal{X}_S) \neq P(\mathcal{X}_T)$
Rule (*iii*) **:** the label prediction tasks between two tasks are different, $\mathcal{Y}_S \neq \mathcal{Y}_T$
Rule (*iv*) **:** Data imbalance exists with regard to the positive and negative classes for the source and target domains, $P(\mathcal{Y}_S|\mathcal{X}_S) \neq P(\mathcal{Y}_T|\mathcal{X}_T)$

$\mathcal{Y}$ is the label space, and $P(\mathcal{Y}|\mathcal{X})$ is the conditional probability distribution over the feature space.

This setting concerns domain adaptation for knowledge transfer in medication-related processing and sample selection bias or covariate shift, which we want to imitate positive/negative polar facts in given social assumptions. Many low-level features learned from enlarged datasets might help our learning to better understand metawords and make sense of medication-related semantics and affective levels from well-understood patterns and to

recalculate new multi-word expression semantics and contextual patterns.

Proposed Approach

As illustrated in the introduction section, our main goal is to consider, extract, and exploit biomedical data for efficient biomedical sentiment analysis. Figure 1 shows the overall architecture of the proposed system. It allows the contextual semantics of a given entity, its data-object properties and its relationships to be exploited in order to compute the attached sentiment based on the formed sentics. The subsections below give more details about each component.

Biomedical Embedding Method

The embedding strategy followed at this point in the process is designed to create a contextual embedding for biomedical concept dependencies on very large shared data. We use semisupervised learning, which uses label propagation to automatically label unlabeled data. As described in Fig. 2, an embedding strategy is defined to represent the entire sentential context representation and is primarily constructed on two levels as follows:

1. **Enrich the training use case:** This is a straightforward technique for mutating predefined knowledge to draw upon pretrained word embeddings (Table 2). An
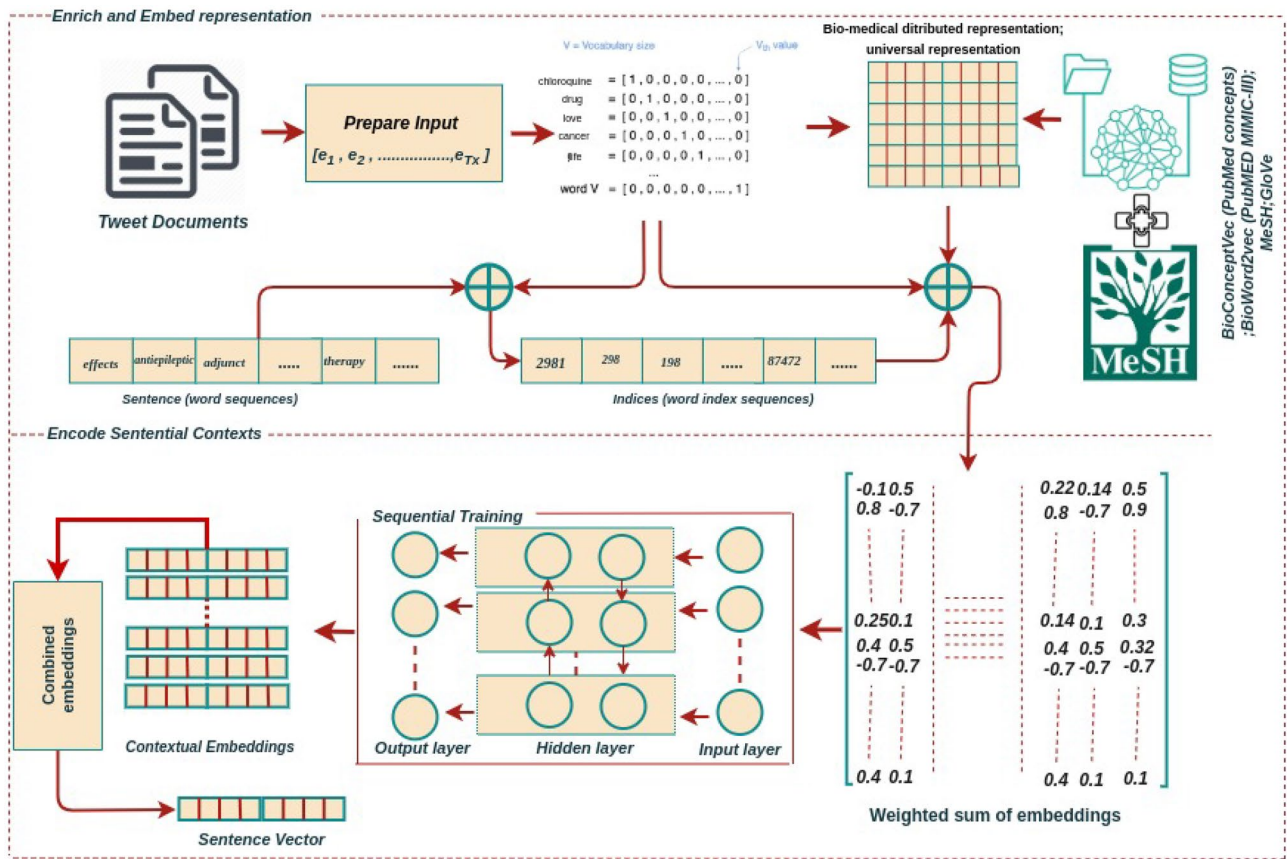
Fig. 2 The process designed to build embeddings

embedding strategy with metaword features is used to support the large vocabularies of social networks. Three main biomedical components were considered: (i) the MeSH relationship network, (ii) distributed representations, and (iii) biomedical concepts and their data-object attributes. Behind the biomedical embedding strategy, a universal metaword dictionary is used by employing GloVe vectors that are trained on very large general-scope data. Glove vectors are incorporated to adjust the general context of some biased medical concepts and biased words, e.g., oils, love, and cancer.

2. **Transfer learning use case:** We followed the setting described in "Transfer Learning", which concerns the domain adaptation of knowledge transfer for medication-related processing and sample selection. The main objectives are to (i) extract medical entities from social network content, (ii) define their correspondence on formal vocabularies, (iii) adjust the medical training contexts to reduce serious biomedical data loss, and (iv) consider the long-term dependencies of n-gram medical components from highly defined abstracts, such as entries referring to the side effects of a given drug.

This method assigns a weight to each entry based on Eq. 1. The context embeddings are computed using the weighted sum of the embeddings of the neighbor words of the target word w_i and thereby sequentially generate the augmented embedding for representing the entire sentential context.

$$e(i) = \sum_{j=i-t;j\neq i}^{j=i+t} e(j)\frac{t - |i - j|}{t} \qquad (1)$$

The performance of transfer knowledge depends mainly on how the sufficient and adequate training corpus is used. Table 2 summarizes the statistics of the distributed vectors we use in this study, which are used for constructing both models and evaluations. They represent an open set of biomedical word vectors (embeddings) that combines subword information from PubMed and MIMIC-III clinical notes with a widely used controlled biomedical vocabulary. Table 3 provides statistics on the different medical language systems and corpora used for training our system. This combination provides well-learned vectors for unseen entries. For each entry that appears in the corpus for the first time, we define dependency vectors that capture

Table 2 A summary of pre-trained embeddings for biomedical words and phrases using medical databases such as PubMed (abstracts, concepts, etc.) and MIMIC-III clinical notes

Biomedical Word Embedding	Embedding	Parameter settings	Text Copora
BioWordVec[13]	FastText	200-dimensional word embeddings, where BioWordVec vector 13GB in Word2Vec bin format and BioWordVec model 26GB.	PubMed and clinical note from MIMIC-III clinical Database
BioSentVec[14]	Sent2Vec	700-dimensional sentence embeddings. We used the bigram model and set window size to be 20 and negative examples 10.	PubMed
BioASQ3[60]	Word2Vec	1,701,632 distinct word vectors and biomedical types	10,876,004 English abstracts of biomedical articles from PubMed.
BioConceptVec[12]	CBOW skip-gram glove fastText	It covers genes, mutations, chemicals, diseases and cellines. The trained embeddings contain over 400,000 concepts.	The entire PubMed biomedical concepts. PubTator to annotate biomedical concepts in the PubMed.

https://github.com/ncbi-nlp/BioWordVec

https://github.com/ncbi-nlp/BioSentVec

http://bioasq.org/news/bioasq-releases-continuous-space-word-vectors-obtained-applying-Word2Vec-PubMed-abstracts

https://github.com/ncbi-nlp/BioConceptVec

semantic and biomedical concept connections and contexts, as described in Fig. 2 ; this is not necessarily training with additional formal data but retraining in a new model with a corpus that has the same vocabulary properties as those previously employed. By considering the MeSH relationship network, the embedding gain greatly improves the performance of the downstream computational prediction models [78]. Indeed, this mechanism enhances the accuracy of our model in comparison with other biomedical embedding baselines in the literature, as we will discuss in "Experiment Design".

For a better generalization performance, we choose to let the training of the new model propagate back to these word vectors to gain understanding of the intrinsic semantic and conceptual connections. Consequently, we can achieve a valid and substantial embedding generation. Then, the underlying transfer is deployed sequentially to generate augmented embeddings that leverage adequate contextual biomedical representations for biomedical sentiment analysis.

As shown in Table 4, a given patient narrative may represent a side effect citation, which can be detected as a drug reaction subword that may be neglected or not considered. In the next section, we aim to define a drug reaction normalization component to embed drug-related entries from given shared contents.

Table 3 Statistics of biomedical corpora and medical ontologies used for pre-trained biomedical embeddings training

Sources	Documents	Sentences	Tokens
PubMed	28,714,373	181,634,210	4,354,171,148
MIMIC-III Clinical Notes	2,083,180	41,674,775	539,006,967
PubMed Central	2,083,180	41,674,775	539,006,967

Drug Reaction Normalization Component

The strong motivations behind adding the drug reaction normalization component are (i) obtaining a distinct set of words that refer to drug reactions and (ii) clearly negotiating the negativity of these elements to reduce the number of false positives and false negatives obtained by our classifier, which are caused mostly by mistaken categorization of ADRs (e.g., headache) and beneficial effects (e.g., hair loss reversal) (which are also known as medical events). We consider the aspect-based sentiments of these medical events by exploring embedding characteristics through groups of drug reactions, where the subgroups consist of

Table 4 Samples of annotated tweets and their drug reaction detected

Twitter document	Drug reaction detected
@carolewhelan i find this drug worse, it has left me in a tired/sleepy fog all day - the codeine i can (mostly) deal with.	Sleepy
@shipmom cipro is awful makes me sick as a dog too	Sick
@jameetmiller i didn't have side effects when i started taking cymbalta. after about 6 mths had to stop because it raised my blood pressure.	raised my blood pressure
brisdelle dangerous. paxil is supposed to treat depression and suicidal thoughts. i took paxil and got depressed. went off & ok in 2 wk.	depression
09.26 day 12 rivaroxaban diary: headache, right shoulder and neck ache, lower back pain, weak knees, limping when walking.taken paracetamol	headache

the defined elements of a group for which the group events with regard to a given event are commutative (common drug interactions).

Drug reaction normalization is the task of retrieving and mapping word combinations that represent drug reactions and finding their corresponding formats from external drug knowledge such as the side effect resource (SIDER). This task frequently requires analysis at the feature extraction level, which can be difficult without sufficient background knowledge of the formal medical context. For this purpose, we have deployed external multicorpus drug reaction knowledge from a state-of-the-art design for adverse drug reaction normalization [76] and the (SMM4H) 2017 shared task [79], data and systems for medication-related text classification and concept normalization from Twitter. The automatic construction of the annotated drug reaction dataset explores many learning algorithms through groups for drug reaction detection (25678 annotated tweets and 9150 ADR phrases and identifiers). Ensembles of system combinations obtained the best score of 88.7%, outperforming individual systems for the normalization of ADR expressions on social networks. The problems are that (1) the systems frequently misclassify rarely occurring ADRs due to suboptimal entity recognition methods in drug reaction detection, and (2) the lack of context in length-limited posts prevents explicit drug reaction formats from being defined. With the dependence on biomedical definitions described in "Biomedical Embedding Method", drug reaction sample selection based on general matching and neural networks is developed. We have paired the drugs and side effects as samples, thereby modeling the problem by using a neural network component that serves to extract related drug reaction contextual semantics, as described in Fig. 5. Indeed, this component consists of the shared layers with the previous construction embedding model. We aim to add related side effect semantics to our weight matrix (embedded biomedical definitions) by using an annotated dataset primarily to seek the possible side effect semantics of unknown words such as those described in Table 4. This component also allows the long-term dependencies of drug reactions on large patient reports and their related attributes to be defined, such as indications, diseases, and drugs.

Each tweet is a candidate for reporting a drug reaction or medical event reaction that can be translated into formal medical language using our neural-network-based detector. The drug reaction detector takes as input a set of n-grams and the corresponding medical format, and it is trained principally with our proposed model to obtain automatic connections with positive and negative medical facts. The model learns the weights via the portable multicorpus training defined earlier. Indeed, a simple recurrent neural network is used on top of the augmented embedding from the previous

component and achieves stable accuracy in detecting drug reaction references in a given raw text.

Inductive Sentiment Method

This section presents a novel method to deduce sentiment from unlabeled data obtained from social networks. We describe the core steps of the proposed analogical affective reasoning method as follows: (1) We first describe in detail how we adapt the sentic computing approach for the task of encoding the affective information conveyed towards medication-related targets. Then, (2) we propose leveraging TF-IDF statistical features in our proposed feature framework. Finally, (3) we create an LSTM-BiLSTM sentiment computing component based on the proposed feature framework.

By adding an extreme learning machine algorithm into our neural network model, the proposed method obtains a good understanding of medication-related affective aspects and generates suitable sentics. Furthermore, we address the question raised in "Introduction" regarding how the deficiency of recognition tasks for word combinations (large medical concepts) in the text may fail to provide relevant insight. By convolving the [1-4] gram training setting, it is estimated that accurate recognition capabilities can be achieved for efficient affective reasoning tasks.

Encoding Sentics and Semantics via a Neural Network

Owing to the need to analyze large volume of patient narratives and extract medication-related concepts, we consider two levels of abstraction: a single entry and a set of entries under various learning problems as the context and target formalization. In particular, we explain the sentic computing approach to clarify common-sense cognitive aspects and affect-related information, especially for medication-related natural language concepts. Human emotions and their modeling are increasingly understood to be crucial tasks, especially under particular circumstances, e.g., patients with particular health conditions.

According to the hourglass of emotions model that is used to form the full spectrum of human emotional experiences, the mind is to a large extent emotionally controlled to obtain independent resources. Moreover, shared information can be relatively quantified to generate affective vectors by specifying emotion compounds through four sentic dimensions that measure the degree to which (i) the user is amused by interaction modalities (pleasantness); (ii) the user is interested in the interaction content (attention), (iii) the user is comfortable with the interaction dynamics (sensitivity), and (iv) the user is confident in the interaction benefits (aptitude). In particular, sentic computing-based encoding is used to synthesize the full range of emotions

conveyed towards medication-related experiences and to make sense of medical facts. The model's output is in terms of the polarity, which is a floating number between -1 and $+1$ (where -1 is extreme negativity and $+1$ is extreme positivity) [15], and is calculated according to the following formula:

$$p = \sum_{i=1}^{10} \frac{Pl(c_i)| + |At(c_i)| - |Se(c_i)| + |Ap(c_i)|}{3N} \quad (2)$$

where c_i is an input concept, N is the total number of concepts, and 3 is the normalization factor.

Each concept may appear in a negative context (negative intensity) or in a positive context (positive intensity). Therefore, we consider the absolute value; for example, in terms of the attention dimension, "surprise" is negative in the sense of lack of attention but positive in terms of the point of view.

The existing sentiment generation approaches discussed in "Affective Computing and Sentiment Analysis" are mainly focused on the identification of parts of the text that express sentiment as polar terms, expressions, and statements that express emotions. In our proposal, we attempt to build analogy-based sentics and semantics within a neural network learning strategy that strives to provide an understanding of a clear connection between different affective aspects. Indeed, a semantic configuration is first used to deconstruct medical natural language text into concepts. Second, linguistic patterns are used in conjunction with sentic dimensions to infer polarity from sentences. If no match is found in the vocabulary or in the linguistic patterns, the neural network model is used.

The large number of patient narratives discussed on social networks spread large amounts of information citing a large number of medication-related concepts. Indeed, they are contextually and semantically associated in different contexts, which further require a higher entity recognition performance to better capture conceptual and sentimental patterns in a sparse dynamic affective space. Moreover, a large social network solution requires particular computing characteristics that allow shared relevant information insights to be captured. First, the process should have low computational complexity, and fast affective learner aspects are required regarding varied targets. In particular, the sentic vector encoding is computed every time for various entries, and multiword expressions are also considered. Otherwise, multiple neural network learning with numerous parameters in each time step will lead to issues in backpropagation network training, while ordinary backpropagation-based neural networks have to deal with the local minima problem, including retraining strategies, convergence rates, etc.

All these features are examples of good extreme learning machine (ELM) training points. Extreme learning machine algorithm training shows appreciable results

with very low computational cost, which is especially relevant when dealing with many patterns defined in a high-dimensional space, where it always converges to the global optimal solution. In this study, we proceed by feedforward neural networks with multiple hidden layers that powerfully identify affective patterns as well as those associated with medication-related conceptual sentics as a function approximation. As depicted in Fig. 3, for a coactivated unsupervised learning problem, each concept (target) is first learned or explained by general concept matching, the semantic and sentic dimensions of which are well explained. Otherwise, concept associations regarding specific biomedical contexts are recalculated every time a new multiword expression is inserted into the vocabulary in Fig. 1, which is further designed to be the input to the LSTM-BiLSTM sentiment inductive model, received as a training set X of N labeled pairs, where (x_i, y_i) is the input.

– $x_i \in \mathbb{R}^{T_x}$: Input embedding vector. Here, T_x is the maximum length of the inputs.
– $y_i \in \mathbb{R}$: The expected target polarity value p is calculated according to four sentics dimensions (pleasantness, attention, sensitivity, and aptitude).

The input layer has T_x time steps and is connected to N_h neurons of the hidden layer through a set of weights $\{\hat{W}_j \in \mathbb{R}^{T_x} j = 1, \ldots\ldots, N_h\}$, which is initialized arbitrarily. The jth hidden neuron embeds a bias term b_j that is activated using a nonlinear activation function g(.).

The coactivation of different affective levels allows different compound emotions to be captured in varied medication-related discussions. Effectively, Fig. 4 highlights the ELM-based network that generates the sentic vectors, where the use of a scalar output implies that the network has one output
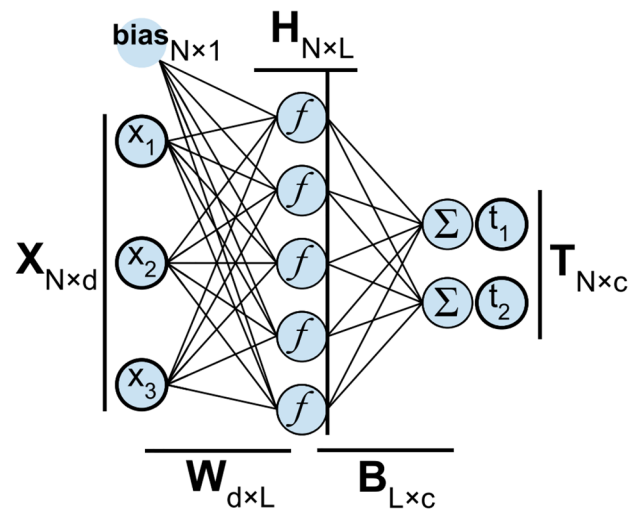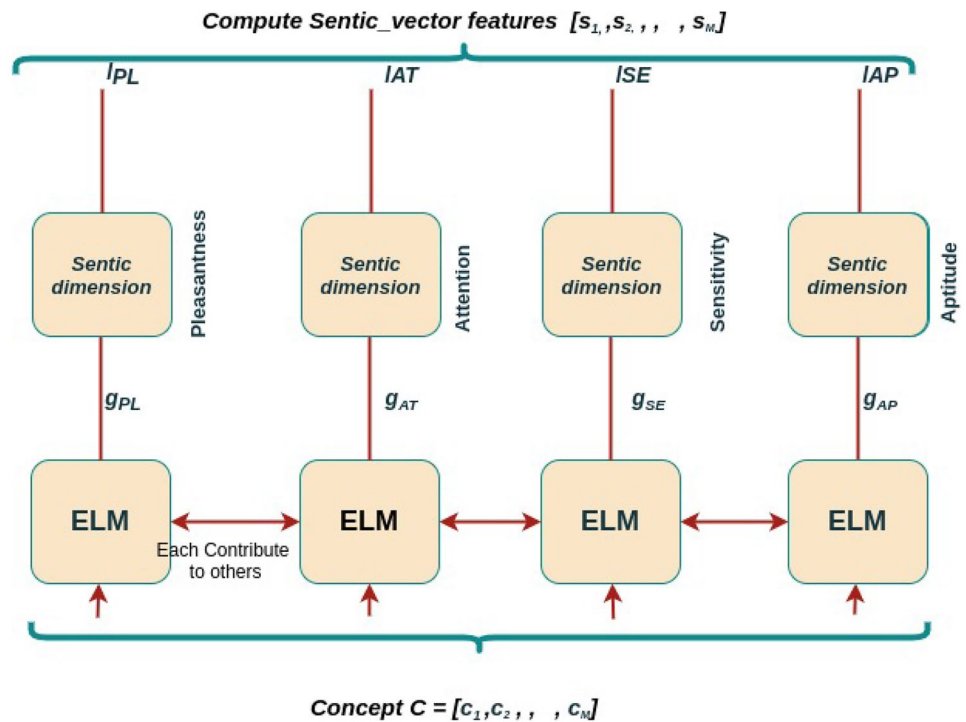

Fig. 3 Multiple hidden layer extreme learning machine network

Fig. 4 ELM-based network to generate sentic vectors by specifying compound emotion into four sentic dimensions

unit, without loss of generality. The sentic vectors are recalculated independently into four predefined levels, allowing us to deductively propagate the long-term affective dependencies of each concept in each dimension and then measure how each sentic level contributes to the others. Many performance evaluations [80] have been carried out on various inference problems, and ELM has demonstrated learning abilities leading to crucial results. The random weights in the hidden layer endow a network with a notable representation ability that embeds the compound emotions of non sentimental expressions such as medical concepts.

The overall output function, f(x), of the network is:

$$f(x) = \sum_{j=1}^{N_h} \bar{W}_j g(\hat{W}_j.x + \hat{b}_j) \tag{3}$$

where $\bar{W}_j$ is a vector of weighted links, $\bar{W}_j \in \mathbb{R}^{N_h}$.

TF-IDF Component

After expanding the sentic vectors with the aim of leveraging the seminal affective meaning of medical concepts, related bag-of-sentimental-concepts features are formed that are similar to bag-of-concepts features. Each dimension in the feature vector represents a concept, and each concept is assigned a value by multiplying the TF-IDF and the polarity value of the concept. The theoretical foundations are considered to be less than firm by

information theory experts. The inverse document frequency for any given term is defined as:

$$tf(term, tweet) = \frac{f_{tweet}(term)}{\max_{word \in tweet} f_{tweet}(word)} \tag{4}$$

$f_d(t)$: frequency of term t in tweet d

$$idf(term) = \log(\frac{n_{tweet}}{n_{allTweets}}) \tag{5}$$

The $n_{allTweets}$ are tweets that contain term. The Twitter TF-IDF matrix is defined as:

$$W_{term_j, tweet_i} = tf(term_i, tweet_j) * Idf(term_i) \tag{6}$$

Each coefficient $W_{i,j}$ associated with term i in tweet j describes and quantifies how important various terms are in a tweet that is part of a Twitter corpus in a slice of time.

Sentiment Computing Component

In this step, we leverage our feature framework from different core steps and bidirectionally propagate them to efficiently compute sentiment-related information toward a particular target t_i in a given context c_i. The overall sentiment inductive method is shown in Fig. 6. The proposed architecture is designed to receive as an input a large number of tweets represented according to an m-dimensional contextual embedding space and to predict the corresponding overall sentiment with regard to the various affective dimensions

Fig. 5 Neural network-based drug reaction normalizer process
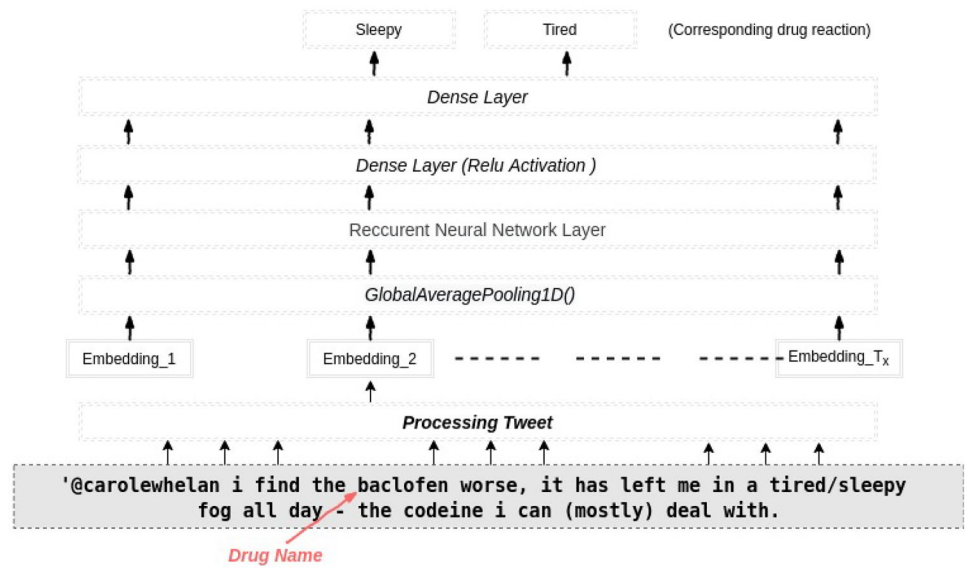
involved, i.e., pleasantness, attention, sensitivity, and aptitude, as discussed in the last section. For the outputs, in principle, the inductive polarity covers three main salient features derived from minute Twitter data and is characterized by an analog value in the range $[-1, 1]$, which represents the intensity of the expressed or received emotion under a medical policy, which is synthesized every time through learning procedures.

However, two important aspects should be taken into consideration. (1) The architecture relies on an embedding strategy that aims to transfer embeddings from large medical ontologies already adopted for social network data ("Biomedical Embedding Method") and to learn connections between new multiword expressions that may refer to medical concepts such as drugs, symptoms, and ADRs. The design of a reliable sentiment predictor regarding biomedical contexts is

Fig. 6 ELM-based network for computing sentimental outputs
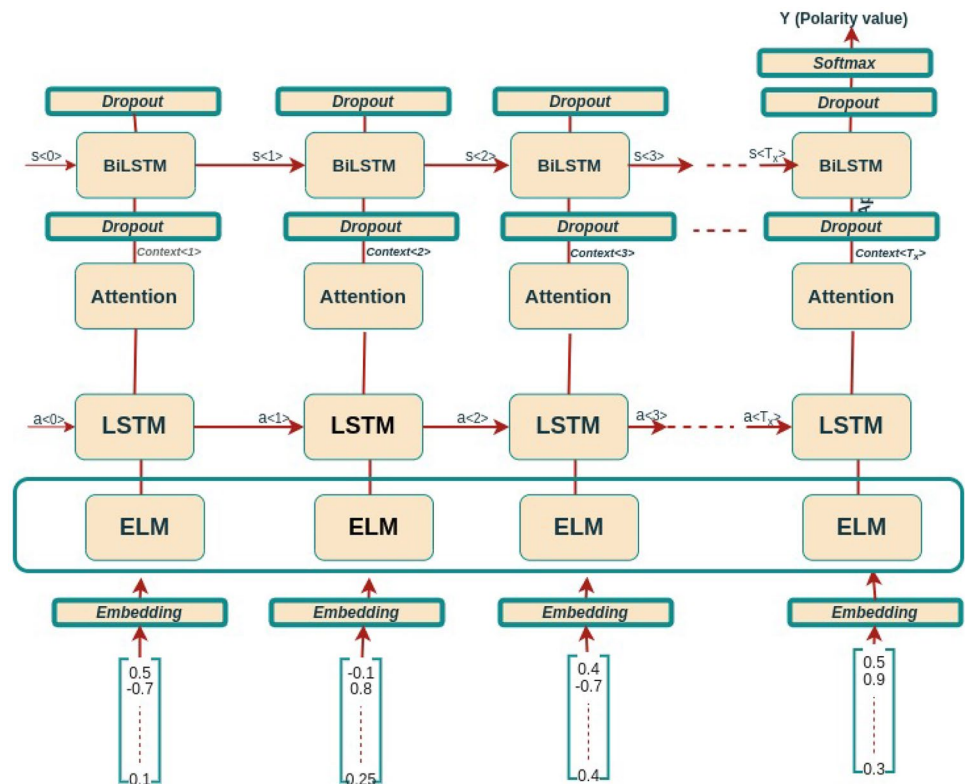
important, especially when considering that few alternative schemata have been provided in the literature without clearly adopted medication-related concept features or a target of analysis. (2) The affective reasoning model discussed previously is considered an input to the sentiment sequential model. Indeed, the values from each sentic dimension are ultimately remapped to obtain the expected concept polarity. Neural computing spans each affective dimension separately, where various dimensions contribute to each other. More complex emotions can be synthesized by using three or even four other sentic levels, e.g., *joy + trust + anger = jealousy*, as discussed by the authors of [43]. Conceivably, each affective dimension may be handled by a dedicated ELM, which addresses learning and explains semantics for a given sentic dimension independently, and then, by using formula 2, the semantic associations of attitudes and feelings are defined as polarity values. These polarity values are joined to another feature $\hat{y}$. Then, these values are fed into a BiLSTM sentiment classifier that is designed primarily to learn efficient connections and long-term dependencies using the policies from previous steps and to loop them into a network, as shown in Fig. 6, which provides the overall scheme of the proposed approach. It yields as output the analog polarity value that will eventually lead to the corresponding sentiment insight.

The problem definition should also state what kind of performance is desired. In this study, automatic sentiment induction of biomedical concepts based on sentic computing and neural networks is carried out to principally clarify truly affective information regarding given medication-related aspects to better fit the medical conceptual aspect and affective associations to detect the "fakeness" or "genuineness" of a given medical fact. Indeed, the ultimate motivation behind the choice of BiLSTM was to learn long-term dependencies, especially those for medication-related terms. Another reason is that multiple BiLSTM layers allow the exploitation of sentic semantics under medical vector dependency at the same time steps and the collection of affective common-sense knowledge through sequential computing, designing the computation of the four affective sentic levels. Moreover, we employ the attention mechanism to help memorize long normalized term dependencies and identify how much attention should be given to particular medication-related concepts through the proposed neural networks. Attention is arguably one of the most powerful concepts in deep learning. The attention layer is a component of memory networks, which focuses their attention on external memory storage rather than a sequence of hidden states.

Validation and Results

In this section, we first present the relevant COVID-19 problem for this study. Afterwards, we conduct experiments in the following directions: (1) we respond to the questions raised

and formulated in this study; (2) we perform a sentence-level evaluation with the clinical semantic textual similarity benchmarking resource to prove model transferability; and (3) we discuss the model performance over the embedding levels, where it is compared with BioSentVec (a sentence-level biomedical distributed representation [14]) in combination with general-scope embeddings. We then evaluate the models using a set of deep learning algorithms based on our proposed feature framework in comparison with other sentiment dictionary baselines. Furthermore, as discussed in "Introduction", the biomedical knowledge combination advantages in improving the system accuracy are assessed through various experiments. In addition, we demonstrate the use of additional features, e.g., the frequency weighting feature for providing significant gains in accuracy and recall over the corpus-based classifier. Finally, (4) we present and discuss different medication-related short text perspectives and the evaluation results.

## Experiment Design

COVID-19 Case Study

The current pandemic is widely known owing to the large number of patient narratives on social networks, which are characterized by vigorous activity and dynamism in terms of global pandemic combat. Many platforms, such as Twitter, are becoming inundated with information and misinformation and are used to share serious side effects of vaccinations and drugs, which might disturb public health.

To validate the proposed case study, the social network Twitter was used. Twitter is a microblogging site that generates a constant stream of communication, some of which concerns events of general topics. In the design of the experiments, a large number of narratives on the COVID-19 crisis are extracted using a set of relevant keywords dedicated to the filtering of information (the most-used drug data groups are used in "Results"). Hence, by posting a tweet, it is possible to analyze the information behavior and then synthesize the affective components (sentics) in the sense of judgments or intentions according to a specific target.

Twitter-Specific Corpus

This contribution also addresses medical domain transferability challenges for biomedical sentiment analysis on Twitter. Due to the limited annotated data, we built our Twitter-based corpus to efficiently define adequate features to generate sentiment semantics and sentics. COVID-19 tweets were collected from January 2020 to May 2020 regarding three properties: (i) the time axis, (ii) drug-related keywords, and (iii) medication-related evidence. We collected tweets and replies in COVID-19-related discussions. The methods of gathering and preprocessing data are described in Fig. 7.

Fig. 7 The process of gathering Tweets from '2020-03-16' to '2020-05-02' and preparation process
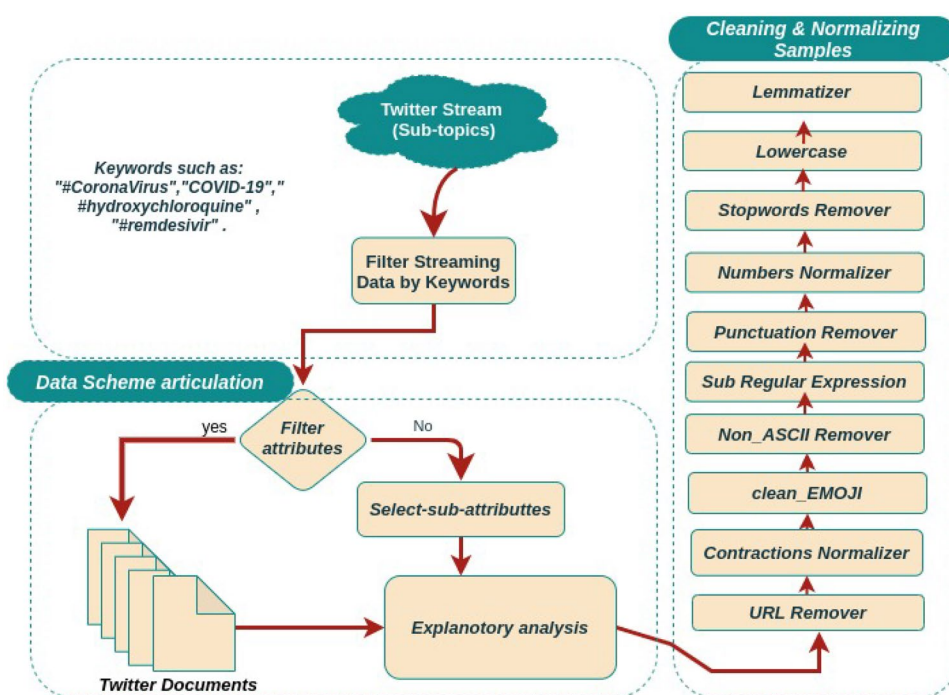
First, we streamed data related to a set of keywords, and we built a data scheme with regard to two types of attributes:

– **Attributes from Twitter:** ["id", "created_at", "source", "original_text", "lang", "favorite_count", "retweet_count", "original_author", "hashtags", "user_mentions", "place", "place_coord_boundaries"].
– **Encoded attributes :** ["cleaned_text", "Sentics", "polarity", "medical concepts", "Bag_of_ADRs"].

The "Sentics" attribute represents affective dimensions including sentiment intensity, semantics, and subjectivity; the "polarity" attribute stands for a polarity value, which is a floating number between -1 and +1 (where -1 is extreme negativity and +1 is extreme positivity); "medical concepts" indicates the set of defined medical entities within a text; and "Bag_of_ADRs" concerns drug effects or indications stated in the Twitter document.

Second, we have stored tweets in a set of files (dataset groups) regarding drugs vs. disease components and a time axis, allowing more conceptual personalization in the same semantic configuration subspace. We applied an explanatory analysis to filter the relevant data into drug-related groups to make sense of medication-related semantics and affective levels from well-understood contextual concept associations. Finally, these tweet samples were passed through a set of normalization and processing steps. Ultimately, we kept 9 M COVID-19-related tweets. Based on the previous settings of the prepared and encoded attributes, we computed the attached polarities and affective dimensions of each tweet and its replies. Then, they

were annotated automatically using the former sentics, and we used them for training and testing purposes.

Additional Twitter Salience

Real-time Twitter-based documents present typical particularities, including limited text size, noise, informal language, and hashtags, which must be taken into consideration. Owing to the character limit, the authors of tweets are usually straightforward. The shortness and lack of biomedical context of length-limited posts cause problems in defining complicated medical components such as drug reactions or events.

Due to the high level of communication rates regarding the urgent COVID-19 pandemic, people spend much time sharing the same ideas about a given drug or committed drug reaction. Observing the three tweets below, each was repeated more or less than the average of 596,730 times (in 9 million tweets). Certain narratives have received much attention and have been frequently retweeted.

Tweet 1: Take #hydroxychroloquine followed by a 500ml glass of Dettol. (As per Presidential Instructions) Lie down in a comfortable position and cover..
Tweet 2: #Hydroxychroloquine in #COVID_19 . The drug is untested the benefits unknown, and the risks not negligible, especially at the mass prophylaxis level of use.
Tweet 3: Im at a complete loss for words. Fox joined by Geraldo Rivera are still PUSHING hydroxychroloquine as a magic bullet for Covid_19 The FDA issued a WARNING

I ve taken the drug for Lupus for 6 years. It has side-effects. ENOUGH!

To determine the appropriate boundary of the affinity of each term, we define the average affinity value as an additional feature to adjust the polarity with regard to how many representatives are in the corpora by the formula:

$$Affinity(P) = \frac{f(P)}{min_{\forall w_i \in P}(f(W_i))} \tag{7}$$

The $f(P)$ is the frequency of phrase P and $min(f(w_i))$ is the minimum frequency across the words in phrase P. Referring to [37], the application of the affinity method allows us to select a set of n-grams that have "higher collocation frequencies relative to the individual occurrence frequencies of the constituent unigrams."

Evaluation Metrics

The evaluation was performed on two levels: (1) automatic biomedical annotation and (2) sentiment classifiers.

Classification Metrics

Precision - The number of definitional sentences correctly labeled by our model divided by the number of sentences marked by our model as definitional.
Recall - The number of definitional sentences correctly labeled by our model divided by the number of definitional sentences.
F1-measure - $2PR/P + R$, where (P) is the precision, and (R) the recall.

Medical Annotation Evaluation

We conduct a sentence-level medical annotation evaluation for clinical sentence pair similarity tasks. MedSTS[4] and BIOSSES[5] are two benchmark datasets and resources for determining clinical semantic textual similarity and are widely used for medical annotation evaluation tasks. Table 5 highlights the statistics of these datasets. Alternatively, a method that is widely used for evaluating multiword concept normalization, such as drug reactions, is to compare the number of correctly normalized drug reactions against the ground-truth value. We adopt this method to evaluate our drug reaction normalizer, defined as follows:

$$Accuracy = \frac{N_{correct}}{N_{T_g}} \tag{8}$$

where $N_{correct}$ is the number of correctly normalized drug reactions and N_{T_g} is the length of drug annotations that are true samples.

Parameter setting

This precomputation strategy allows a nice compromise between GPU (NVIDIA (R) CUDA compiler driver) and implementation simplicity. For the ELM, the wide lower layer and each of the 4 bidirectional-LSTM layers, we used the Adam optimizer with 128 units for both the LSTM and BiLSTM folds. In addition, we trained the model from 1-30 epochs, computed on the GPU[6], and then moved the precomputed hidden weights to the CPU for the prediction task. This makes many things much easier. To boost our model performance, we train a sufficiently large neural network in terms of parameters and connections to take advantage of the very large number of shared experiences on social networks in a matter of seconds.

Results

The experiments conducted assess the performance of our model. We mainly compare our model with other methods in three aspects: (1) We compare the performance of our proposed feature framework with biomedical embedding baselines. (2) We explore the performance of our approach in a biomedical sentiment analysis case study. (3) We compare different ways to encode sentiment regarding medical aspects and related components. The most crucial challenge of this study is the lack of a labeled training dataset including medication-related concepts since the drug reaction detector model requires much label training data. However, with word embeddings, the model can build good classifiers even with only modestly sized label training sets. We cover large-scale biomedical semantic indexing and map medical concepts to natural language expressions cited in patient narratives on social networks, especially drug reaction expressions and PubMed concepts.

In fact, we have a potent biomedical continuous space of 10, 876, 004 English abstracts of biomedical articles from PubMed, 400,000 concepts, and 15, 677 resulting vectors of 1, 701, 632 distinct words that form drug reaction expressions, PubMed abstracts, symptoms, drugs, etc. All words were converted to lowercase, the embedding was converted to the same dimension, and the vector dimension was 300. Our sequential model considers tweets as a sequence of tokens, and the outputs are floating numbers in $[-1, 1]$ with clear medical concept tagging. The settings of the features used in the model are as follows:

Table 5 Two banchmarking datasets used for medical annotation evaluation tasks

Dataset	Description	settings	Accuracy
MedSTS	Resource for clinical Semantic Textual Similarity	MedSTS (MedSTS_ann) containing 1,068 sentence pairs was annotated by two medical experts with semantic similarity scores of 0-5 (low to high similarity).	0.81
BIOSSES	Semantic Sentence Similarity Estimation System for the bio-medical domain	The BIOSSES data set comprises total 100 sentence pairs all of which were selected from the TAC2 biomedical Summarization Track Training Data Set .	0.79

https://arxiv.org/ftp/arxiv/papers/1808/1808.09397.pdf

https://tabilab.cmpe.boun.edu.tr/BIOSSES/DataSet.html

1. The biomedical distributed representation, which shuffles pad vectors into the same dimension.
2. GloVe vectors, which define universal words and adjust the general context of some biased medical concepts and biased words, e.g., oils, love, and cancer.
3. Sentic vectors, which define sentimental semantics and recontribute to the affective dimensions every time, namely, the sentics.
4. The LSTM, which learns embeddings for new multiword expressions and represents long-term dependencies in the very large vocabulary of medication-related concepts, e.g., drug names, symptoms, pandemics, and diseases.
5. The BiLSTM, which deploys precomputed features and learns long-term sentiment dependencies regarding medication aspects and contexts.
6. The attention mechanism, which defines attention parameters that propagate the amount of attention that should be given to those features of the input sentence.

For each token, we consider an embedding vector hybridizing both a universal format and biomedical information into a 300-dimensional vector. Our baseline model can actually work fairly well using only biomedical vectors but leaves out universal and concept-related contexts such as drugs or symptoms. Therefore, we developed dynamic medical corpora, which are further dedicated to learning deeper representations under target medical settings, and we collected 8,593,940 examples. The first version of our model used 9 million COVID-19-related tweets learned by the BiLSTM and our sentic computing-based sentiment architecture. This gain is sometimes not significant regarding the generalization aspect. Moreover, these deep architectures work significantly better than the classical approaches. We have obtained degrees of error from approximately 8% to 5% when we used n-grams with TF-IDF. This number of Twitter examples can be addressed efficiently by learning from scratch based on the literature, which works best for large datasets, where our method beats classical approaches such as the bag of words and part-of-speech feature selection techniques. To convert our input sequence into a vector of fixed size, we apply some

dense layers. We apply a multilayer positron on top of the 300 features and train them using a long-term memory algorithm that uses GloVe vectors to ensure that universal words are defined in connection with the biomedical expressions.

We apply neural networks at different architecture levels, and we prove how we can embed our medication-related concepts efficiently, especially unlabeled examples. We propose analyzing two-grams and three-grams using the ELM neural network to propagate the sentiment aspects for new multiword medication expressions that are inserted into the vocabulary, which are all fast operations. Therefore, it works even faster than classical sentiment lexicon approaches. Even drug reactions that are not in the vocabulary may be selected as possible drug reactions. They are used to define embeddings and may still perform this correctly and generalize much better. The activation step of the predicted sentics has approximately four levels but concomitant dimensions. Theoretically, the model synthesizes the full range of emotional experiences in terms of pleasantness, attention, sensitivity, and aptitude, as the different combined values of the four affective dimensions can also model the affective states. In this way, we trained our model to achieve good performance and obtain a deep sense of concepts.

Additionally, popular non-domain-specific sentiment lexicons as well as state-of-the-art machine-learning and deep-learning models are chosen as benchmarks, and the experimental results show that our sentiment feature framework outperforms the benchmarks with statistically significant differences, thus proving the effectiveness of the proposed approach. An input sequence was first processed for embedding generation to train the specific features that this neural network needs to classify the conveyed sentiment. The LSTM and Bi-LSTM had different operational learning procedures, which were learned, memorized, and hybridized outputs that served as the inputs of the next level (Fig. 6). Indeed, the weight matrix of each level was stored for evaluation. Deep learning models have slightly different results regarding the embedding strategy offered in this study vs sentiment dictionaries; we also evaluated the models in our model groups based on comparisons with deep learning (DL) algorithms

(LSTM, BiLSTM, multiple RNNs, and CNN) and SVM on COVID-19 Twitter datasets, as described in Table 6.

The BiLSTM algorithm on one- and two-grams, in conjunction with the SenticNet prior sentiment knowledge, achieves an accuracy range of [83-85]%. Particularly, when we merge [1-4]-grams through the proposed architecture with BiLSTM on top of those that depend on features, we obtain an accuracy of 8.2 bumps, which gives us an almost 87.5% accuracy with the potential for clarifying the corresponding medical concepts as targets. Then, we ensure that false positives and negatives are avoided. We train the model for 1–30 epochs, and then we choose a set of optimal parameters from the trained model under this setting. Subsequently, we choose the optimal settings that yield the minimum loss. Therefore, we can capture the information about three-, four-, and five-grams, and for each n-gram, new multiword expressions are inserted into the vocabulary.

Observing sentiment information regarding three pertinent drug names reveals their high rate of appearance in our experimentation. A statistical analysis was performed on the inference results for three drug components to visually

Table 6 Comparison of proposed feature framework with different ensemble classifiers using other sentiment benchmark dictionaries

| Corpus	Sentiment dictionary	Algorithm	Accuracy
Twitter datasets	Our proposed feature	LSTM-BiLSTM	**0.87**
	framework	LSTM	0.72
		Multiple RNN	0.72
		CNN	0.68
		SVM	0.58
	SenticNet [43]	BiLSTM	0.79
		LSTM	0.73
		Multiple RNN	0.66
		CNN	0.72
		SVM	0.60
	AFINN [45]	BiLSTM	0.72
		LSTM	0.71
		Multiple RNN	0.62
		CNN	0.68
		SVM	0.45
	VADER [38]	BiLSTM	0.67
		LSTM	0.68
		Multiple RNN	0.65
		CNN	0.63
		SVM	0.52
	TextBlob [81]	BiLSTM	0.67
		LSTM	0.62
		Multiple RNN	0.70
		CNN	0.66
		SVM	0.45
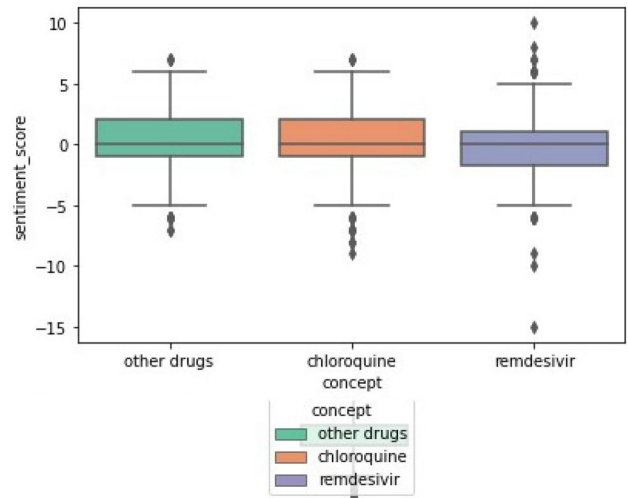
Fig. 8 Box plot of the performances provided on each medical component

show the distribution of numerical data and skewness by displaying the data quartiles (or percentiles) and averages. However, for a different time, Fig. 10 shows a box plot of the performances provided for each component. The number of positive examples was the highest due to the highly recommended use of these drugs to avoid COVID-19 contamination. In particular, we obtained many tweets from US President Trump discussing chloroquine treatment. Thus, we tracked many chloroquine treatment use cases and/or drug effects since we captured drug effects in the text, turning the target of the analysis towards no emotion. Before regenerating sentics-related vectors, the neutrality in the results was explained. Fig. 11 shows the dispersion of sentimental polarities over the three medical components, which helps in
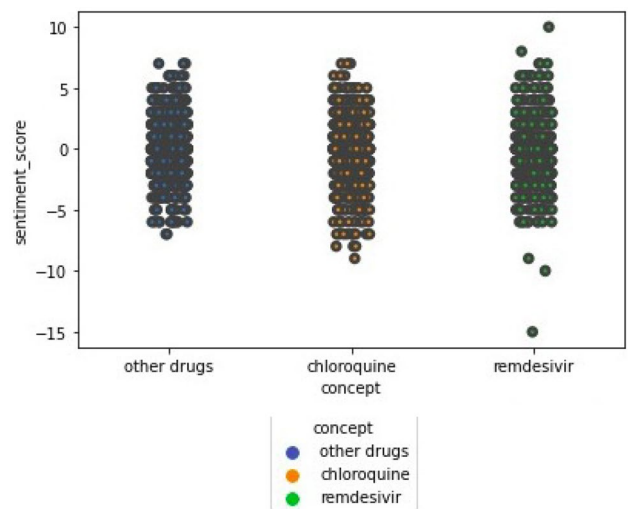

Fig. 9 The dispersity of sentimental polarities vs. training data over three medical components
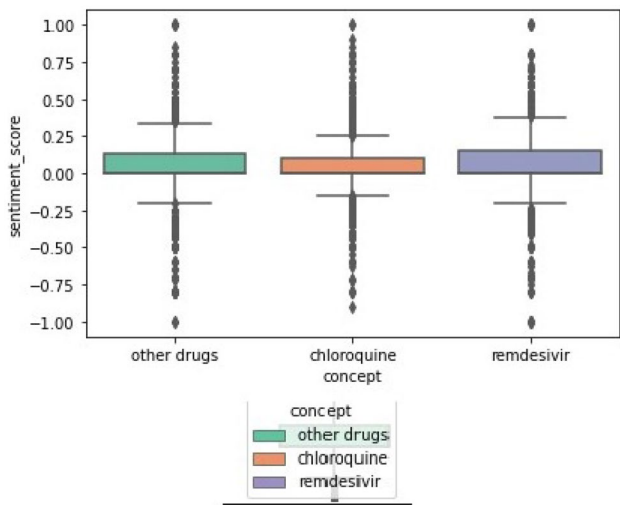
Fig. 10 Box plot of the performances provided on each medical component

Fig. 12 Positives vs. negatives of three most visible medication-related components

seeking the extreme positive values and extreme negatives, which can be used to manually verify whether false positives or negatives occur. The histogram in Fig. 12 shows the positivity and negativity of each component.

Experiments with different distributed biomedical methods and sentiment prediction DL-based algorithms were carried out. Our method shows the effectiveness of considering medication concepts in various situations. An exploratory data analysis is also offered for comparison lexicons. Here, we plot the statistical results for the AFINN model. The box plot in Fig. 8 shows higher inference positivity, while Fig. 9 presents a histogram showing the sentiment discrepancy over three drug-related discussions.
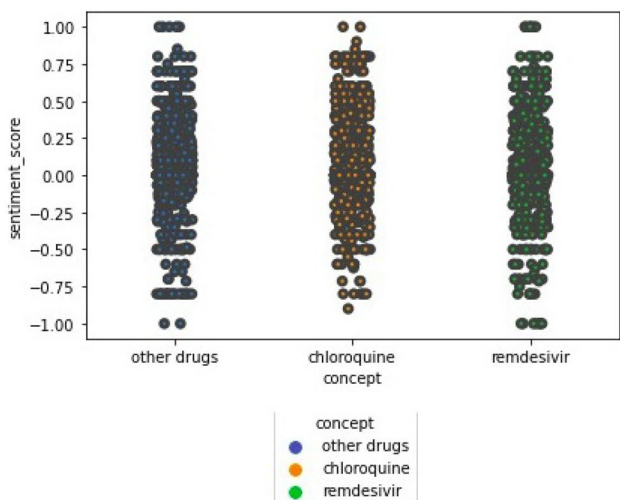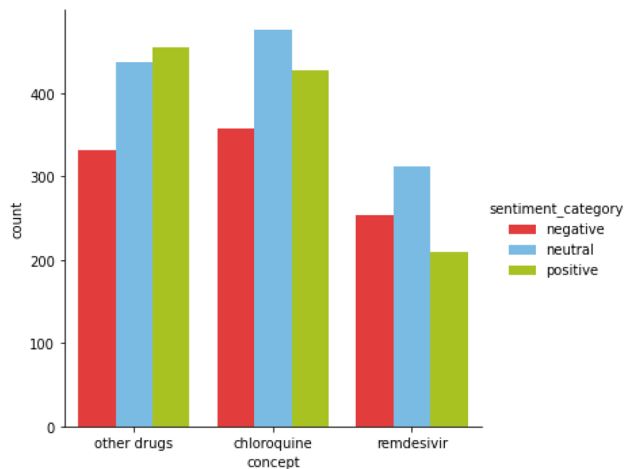
In addition, we visualize the positive and negative numbers in Fig. 13 and a clear indication of the variance behavior. For comparison, we applied literature baseline sentiment lexicons based on the embedding vectors we created. We have used five benchmark sentiment lexicons for the evaluation: (i) SenticNet: we aimed to evaluate our connotated vocabulary with both SenticNet 3 [82], to evaluate the model ability to learn affective information of out-of-vocabulary terms, and SenticNet 6 [43], to compare the overall polarities with a sentiment lexicon that considers other dimensions; (ii) AFINN: a word list for sentiment analysis for short microblog text; (iii) the TextBlob dictionary, and (iv) the valence aware dictionary and sentiment reasoner (VADER): a second rule-based benchmark sentiment analysis lexicon that is specifically attuned to sentiments expressed via social media and works well on texts from other domains. Table 7
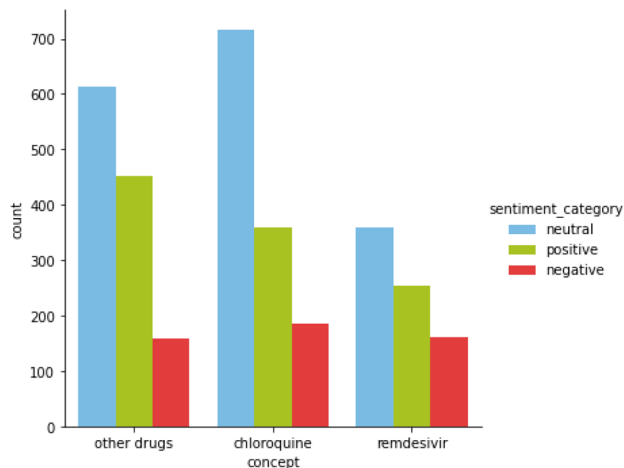

Fig. 11 The dispersity of sentimental polarities vs. training data over three medical components

Fig. 13 Positives vs. negatives of three most visible medication-related components

Table 7 Benchmarking sentiment dictionaries used for evaluation tasks

Lexicon	Description	Level of automation
VADER [38]	A sentiment dictionary, it attuned to sentiment in micro-blog. It combined various lexical features with consideration for five general rules that embody grammatical and syntactical conventions for expressing and emphasizing sentiment intensity. Thus, It has been found to be quite successful when dealing with social media texts e.g. used by Amazon's Mechanical Turk for sentiment ratings.	Semi-automated
AFINN [45]	A sentiment Dictionary (over than 3300)rated for valence with an integer between minus five (negative) and plus five (positive), though not as good as the more elaborate approach found in SentiStrength.	Manual
TextBlob [81]	Sentiment Lexicon goes along finding words and phrases it can assign polarity and subjectivity to, and it averages them all together for longer text. The scores from TextBlob are normalised scale as compare to AFINN.	Automated
SenticNet [43, 50, 82]	The knowledge base which the sentic computing framework leverages on for concept-level sentiment analysis. SenticNet is automatically constructed by clustering the vector space model of affective common-sense knowledge extracted from varied lexical resources.	Automated

summarizes these word-level benchmark datasets in detail. Although they are very effective, these methods have several limitations and need to be improved regarding sentiments for biomedical text. Indeed, the AFINN lexicon gives the largest absolute values, with high positive values. We find similar differences between the methods when looking at other metrics; the proposed sentic computing-based sentiment model is high and has more variance than AFINN and VADER. Table 8 presents all sentiments that appeared various times in finding long stretches of similar text, but all three methods agree roughly on the overall trends in the sentiments through the common shared narratives we obtained through our analysis. An explanatory summary is given in Table 9.

Furthermore, to ensure the effectiveness of our embedding generation, we conducted other experiments using two different sentiment lexicons, where three basic feature embedding strategies were used. As described in Table 10, each embedding feature plays a different role in sentiment determination. GloVe Twitter with a combined biomedical embedding generally performed much better than BioSentVec, which helped resolve noisy Twitter data and further gave meaning to the classification task. However, the proposed biomedical-based embedding clearly enhanced the sentiment inference, outperforming the evaluation techniques. Table 11 presents the polarity value of some tweets from the dataset groups collected.

Linguistic Gap Discussion

In this section, we discuss the importance of distinguishing biomedical components, their aspects, and their relationships in detecting sentiment orientation regarding social network dynamicity and the rapid progress of biomedical language. The poor performance of biomedical sentiment analysis on Twitter may be attributed to many challenging particularities of both tweets and bio-medical entities. Approximately twenty billion tweets are typed every day on varied topics. Distilling the hidden information faces many difficulties and challenges, including language diversity, brevity, shortness, acronyms, named entities, hashtags, emoticons, the frequent use of informal (nontechnical) medical language, nonstandard formats, and abbreviation forms, as well as typos in social network messages. Specifically, the brevity of tweets results in relatively few terms, which contributes to increasing the biomedical disambiguation complexity. The massive bio-medical entity citations, such as drug names, diseases, symptoms, and adverse drug reactions, remain unconsidered by existing tagging and entity recognition methods [6], which hinders understanding of patient assessments. Therefore, this may necessitate a level of biomedical text comprehension. Numerous learning approaches have recently been considered to overcome these limitations (as discussed in "Background and Related Work").

Table 8 Statistics of proposed sentiment inductive model on first Twitter corpus regarding two frequently related drugs components narratives (chloroquine and remdesivir) with other drugs

concept	sentiment score count	mean	std	min	25%	50%	75%	max
Chloroquine	1261.0	0.051449	0.234630	−0.9	0.0	0.0	0.100	1.0
Other drugs	1224.0	0.058930	0.248925	−1.0	0.0	0.0	0.136	1.0
Remdesivir	774.0	0.043039	0.302643	−1.0	0.0	0.0	0.150	1.0

Table 9 Statistics of other lexicons such as AFINN for sentiment inductive model on first Twitter corpus regarding two frequently related drugs components narratives (chloroquine and remdesivir) with other drugs

concept	sentiment score count	mean	std	min	25%	50%	75%	max
Chloroquine	1261.0	0.002379	2.185993	−9.0	−1.00	0.0	2.0	7.0
Other drugs	1224.0	0.261438	2.067610	−7.0	−1.00	0.0	2.0	7.0
Remdesivir	774.0	−0.114987	2.398460	−15.0	−1.75	0.0	1.0	10.0

Most supervised learning approaches are trained on a limited set of word features that result in traditional feature representation and typically diminish the model performance by ignoring other significant insights. Thus, these approaches focus on an individual word's sentiment value, which is independent of the context and insufficient in yielding a better feature subset and hence affects the classification accuracy. The performance of supervised learning depends on the quality data annotation or makes use of resources for weak supervision [11] and results in a further lack of discriminatory tasks. In the bio-medical domain, these models result in inaccurate sentiment prediction performance.

Unsupervised learning approaches result in irrelevant and redundant features. Additional techniques have been considered, such as N-gram techniques, to efficiently leverage language properties. Such techniques include (1) unigram methods used for corpus-based approaches, (2)

n-grams for tracking syntactical aspects, (3) parts of speech, (4) POS tags, and (5) the information gain method, which is mostly used to reduce the number of features. These methods mostly result in a sparse feature representation with ambiguous polarity since some words may have different polarities regarding several contextual aspects. Some researchers have assembled various features from different methods to enhance the accuracy and recall performance [83]. For example, the authors of [84] merged various numbers of features, ranging from 3,066 features (a simple collection of unigrams from the General Inquirer) to 34,718 features (unigrams and bigrams), to boost system performance. However, they made their features completely independent of each other, and features beyond sentences are often corpus dependent, while this may not be the case for online corpora. For this reason, many contributions focus on extracting features from a sentence only by leveraging its definitional structure. Indeed, many previous methods generate features from sentence-based dependency parsing, which can reflect only part of the structure feature. Thus, in comparison with benchmark machine learning algorithms such as SVM, the SVM classifier performed worse than unigrams when using all bigrams as features instead of just the unigrams. The authors of [85] proposed selecting specific bigrams to improve the unigram baseline. The selection mechanism of unigrams, bigrams, and even 3-, 4-, and 5-grams can greatly improve sentiment analysis performance while reducing the modeling complexity, as noted by [37]. At the sentence level, most studies that use unigrams yielded better classification results than those of n-grams for domain-dependent use cases.

N-gram-based techniques have recently achieved remarkable progress in terms of enhancing representation features using neural networks. The most successful type of method is embedding. One possible solution is that we are most likely to focus on ways to improve embeddings of cognitive capabilities, as discussed in "Biomedical Embedding Method". Indeed, we demonstrated the importance of processing and selecting n-grams in enhancing affective reasoning. In this section, we discuss some linguistic gaps confronted through bio-medical entity identification. For this purpose, we perform biomedical embedding comparisons in two cases: (i) a small dataset and (ii) a large dataset.

Table 10 Evaluation of proposed sentiment feature framework with existing embedding strategies

Sentiment dicionary	Embeddings	Accuracy
Our proposed sentiment inductive model	Our embeddings framework	**0.87**
	BioWordVec +BioConceptVec	0.79
	BioASQ3 +GloVe	0.71
	BioWordVec +GloVe	0.72
	BioSentVec+ Twitter GloVe	0.72
	BioWordVec +FastText	0.76
SenticNet	Our embeddings framework	**0.79**
	BioWordVec +BioConceptVec	0.75
	BioASQ3 +GloVe	0.76
	BioWordVec +GloVe	0.69
	BioSentVec+ Twitter GloVe	0.73
	BioWordVec +FastText	0.70
AFINN	Our embeddings framework	**0.74**
	Twitter GloVe +BioWordVec	0.70
	BioSentVec + GloVe	0.67
VADER	Our embeddings framework	0.81
	GloVe +BioWordVec	0.77
	BioSentVec + GloVe	0.73
TextBlob	Our embeddings framework	0.71
	GloVe +BioWordVe	0.69
	BioSentVec + GloVe	0.67

Table 11 The overall sentiment (polarity value) of some tweets from dataset groups collected through this study

Twitter document	Medical components	Polarity value
@singleuesguy hydroxychloroquine ninety percent chance help covid19 patient with chronic disease state association	hydroxychloroquine, covid19, patient,chronic disease	0.4
@thekjohnston remember trump amp blowhard fox wouldnat shut hydroxychloroquine miracle covid cure makina	covid,hydroxychloroquine	-0.2
@channelstv bauchi governor mandate use chloroquine treat covid19	chloroquine, covid19,	-0.02
@jamestodaromd fda need randomize controlled trial hydroxy-chloroquine treatment covid19 also fda discourage	trial, hydroxychloroquine treatment covid19	-0.4
@tedlieu anyone read tweet @realdonaldtrump say hydroxy-chloroquine amp azithromycin game changer #covid	hydroxychloroquine, amp, azithromycin	0.1
@ingrahamangle hyperglycemia hydroxychloroquine covida19 pandemic drop brufsky journal medical virology wiley online lia	hyperglycemia, hydroxychloroquine, covida19, medical virology	-0.3

On a small dataset, utilizing BioWordVec as a large distributional biomedical space in the literature, more than 18 drug names in each dataset group were unrecognized. Approximately the same filtering rate was obtained using a Google Word2Vec pretrained embedding. Thus, the Bio-ConceptVec model resulted in poor performance in detecting multiword medical components, and similar results were obtained for all the embedding versions used (CBOW, Skip-Gram, fastText, and GloVe). Otherwise, the GloVe model performed better; in this case, only nine drugs were undefined because GloVe was trained on Wikipedia corpora.

On a large dataset, the evaluation was performed on 441,228 original tweets (with repeated tweets filtered out) from 9 million tweets collected in this study, with 407,578 total words. Table 12 summarizes the results we obtained after generating embeddings using various biomedical models. These models only captured some general medical concepts (such as cancer and tumor) because of the very large general-scope corpora they were trained on, such as Google

News and Wikipedia. After the first filtering attempt, we identified 292,358 words ignored by the GloVe model. We found 79,096 (18%) words defined by the GloVe model and 301,392 (69%) undefined terms. We performed the same process using the Google model to identify general-scope words, and we obtained only 12% of the terms in the vocabulary. After a second filtering with the GloVe model, we identified 282,430 ignored words. A mean of 380,480 undefined medical bigrams was found, and in terms of medical unigrams, only 25.8% of all words were defined, with 74.17% of all words lacking embedding vectors. Even though we enriched the data and retrained the word embeddings, the semantically rich knowledge, contexts, and medical senses were lost.

The nonidentification of the full range of words is made completely clear by all the embeddings that we evaluated. BioConceptVec defined only 77,853 (17.6%) concepts, whereas 302,635 (68.5%) concepts were ignored. Although BioWordVec was trained on PubMed abstracts and the

Table 12 Statistics on over 9 Millions Tweets vs. Distributed representation of general scope data(Word2Vec-GloVe) and biomedical datasets

Emebeddings	Corpus	setting of embedding statistics	words analysis defined words	undefined words
Google Word2Vec model	google news	300dim	48931(12%)	331557(76%)
GloVe model	wikipedia	840B tokens, 2.2M vocab, cased, 300d vectors	79096(18%)	301392(69%)
BioWordVec model	PubMed+MIMIC-III	200dim Word2Vec bin format	87810(19%)	292678(66%)
BioConceptVec with CBOW model	PubMed and Pubtator auto-tagger	contain over 400,000 concepts 100dim	77853(17.6%)	302635(68.5%)
BioConceptVec with Skip-Gram model	PubMed and Pubtator auto-tagger	contain over 400,000 concepts 100dim	77853(17.6%)	302635(68.5%)
BioConceptVec with fastText model	PubMed and Pubtator auto-tagger	contain over 400,000 concepts 100dim	77853(17.6%)	302635(68.5%)
BioConceptVec with GloVe model	PubMed and Pubtator auto-tagger	contain over 400,000 concepts 100dim	77853(17.6%)	302635(68.5%)

Table 13 Example of two different undefined words showing the weakness of BioWordVec to detecting related-drug information

Medical term	Description	Other names	reference
Creatine	Creatineis a chemical that is found in the body. an effective therapy for cancer, which in samples obtained by the NCI and the FDA consisted of creatine monohydrate in mineral oil.	Acetic acid, Cr, Creatin, Creatina, Créatine, Créatine Anhydre, Creatine Anhydrous, Creatine Citrate, Créatine Citrate, Creatine Ethyl Ester, Créatine Ethyl Ester... etc.	FDA(U.S. Food and Drug Administration) and NCI(National Cancer Institute).
Attenta	Attenta belongs to a group of medicines called central nervous system stimulants. Attenta is thought to work by regulating certain chemicals in the brain, which affect behaviour and sleep.	Piperidine acetic acid,α-phenyl-, Ritalina, Rilatine, Attenta, Medikinet, Metadate, Methylin, Penid, Tranquilyn, and Rubifen ...etc.	Medicines

https://pubchem.ncbi.nlm.nih.gov/compound/Creatine#section=Structures

https://www.webmd.com/vitamins/ai/ingredientmono-873/creatine

https://www.fda.gov/home

https://www.cancer.gov/

https://apps.medicines.org.au/files/afcatten.pdf

https://medicines.org.au/files/afpatten.pdf

MIMIC-III clinical Database, it neglected much drug-related information, such as ['wcco', 'partnera', 'warna', 'unimelb', 'mcswan', 'creatina', 'ncdca', 'asu per', 'attenta', and 'tidak'], which were subsequently identified using Bio-ConceptVec vectors. This inability is due to the multiple name types. Another significant challenge in terms of related word names is that a medical concept may be known by its commercial name and reported by its chemical name or vice versa. Each concept may have many names ranging from natural language concepts, commercial names (generic brand names), and chemical names to other spellings depending on the region or speaking culture. For example, methylphenidate drugs are referred to by many names, including Ritalin, attenta, Medikinet, and metadata, as discussed in Table 13.

The semisupervised approach has great adaptability to domain-dependent challenges, offers more portable solutions across various domains and significantly allows domain transferability [37] due to reusable feature-sentiment correlation across domains. Therefore, we deployed a transfer mechanism to obtain generalization benefits from corpora trained on very large datasets to obtain various biomedical components. Indeed, this setting is convolved to leverage knowledge from both the bio-medical domain and general-scope domains to overcome the constraints of limited training data in biomedical contexts.

Conclusion

Information on social media can have a direct influence on patients' decisions to seek a second opinion or choose among potential alternatives such as diagnoses or treatments. Sentiment analysis is widely used for analyzing online patient-generated narratives. In this research, we develop an affective concept-based encoding method via neural networks and sentic computing. It relies mainly on biomedical semantic identification, which is investigated at three levels: (1) a contextual neural network, (2) an embedding matrix, and (3) a biomedical distributed representation obtained from comprehensive and controlled medical databases such as PubMed. Then, the natural medical concept semantic and sentic vectors are calculated. In addition, this paper presents a comprehensive state-of-the-art overview of the studies performed in various areas of biomedical sentiment analysis.

During the worldwide combat of COVID-19, we have suggested validating the proposed approach by analyzing COVID-19 narratives shared on social networks. Experiments have shown the effectiveness of considering biomedical concepts in analyzing medication-related texts. Indeed, the results proved that integrating deep-learning-based cognitive capabilities enhances the biomedical sentiment analysis performance. In particular, our affective encoding approach outperforms benchmarks with statistically significant differences. Furthermore, the current work can provide meaningful support for biomedical sentiment analysis applications in social networks and can truly track patient status. It may also help researchers and professionals in many related discovery tasks, such as (1) inferring unknown side effects of possible drugs and treatments based on the formed embeddings, (2) clarifying the relationships among sets of drugs and conditions and situations to better clarify the "fakeness" or "genuineness" of a given medical fact, and (3) analyzing changes in health status or unexpected situations.

In future work, we plan to investigate the impact of incorporating additional affective dimensions such as frustration to model more complex emotions regarding patient aspects. As a

second future project, we would like to apply our unsupervised algorithm using transformers such as BioBERT and GPT3 to improve the ability to learn and recall factual knowledge.

Declarations

Human and Animal Rights This article does not contain any studies involving human participants and/or animals by any of the authors.

Informed Consent Informed consent was obtained from all individual participants.

Conflicts of Interest The authors declare that they have no conflict of interest.

References

1. Li Z, Yang F, Luo Y. Context Embedding Based on Bi-LSTM in Semi-Supervised Biomedical Word Sense Disambiguation. IEEE Access. 2019. https://doi.org/10.1109/ACCESS.2019.2912584.

2. Schouten K, Frasincar F. Survey on Aspect-Level Sentiment Analysis. IEEE Trans Knowl Data Eng. 2016. https://doi.org/10.1109/TKDE.2015.2485209.

3. Zhou J, Jin S, Huang X. ADeCNN: An improved model for aspect-level sentiment analysis based on deformable CNN and attention. IEEE Access. 2020. https://doi.org/10.1109/ACCESS.2020.3010802.

4. Aydln CR, Gungor T. Combination of recursive and recurrent neural networks for aspect-based sentiment analysis using inter-aspect relations. IEEE Access. 2020. https://doi.org/10.1109/ACCESS.2020.2990306.

5. Yu J, Jiang J, Xia R. Entity-sensitive attention and fusion network for entity-level multimodal sentiment classification. IEEE/ACM Trans Audio Speech Lang Process. 2020;28:429–39. https://doi.org/10.1109/TASLP.2019.2957872.

6. Grissette H, Nfaoui EH. A conditional sentiment analysis model for the embedding patient self-report experiences on social media. In Advances in Intelligent Systems and Computing, 2019. https://doi.org/10.1007/978-3-030-11884-6_6.

7. Yang Q, Xie H, Cheng G, Wang FL, Rao Y. Pronunciation-enhanced chinese word embedding. Cogn Comput. 2021;1:3. https://doi.org/10.1007/s12559-021-09850-9.

8. Zuo E, Zhao H, Chen B, Chen Q. Context-specific heterogeneous graph convolutional network for implicit sentiment analysis. IEEE Access. 2020;8:37967–75. https://doi.org/10.1109/ACCESS.2020.2975244.

9. Ke Z, Sheng J, Li Z, Silamu W, Guo Q. Knowledge-guided sentiment analysis via learning from natural language explanations. IEEE Access. 2021. https://doi.org/10.1109/ACCESS.2020.3048088.

10. Grissette H, Nfaoui EH. Drug reaction discriminator within encoder-decoder neural network model: COVID-19 pandemic case study. 2021. https://doi.org/10.1109/snams52053.2020.9336561.

11. Grissette H, Nfaoui EH. Enhancing convolution-based sentiment extractor via dubbed N-gram embedding-related drug vocabulary. Network Modeling Analysis in Health Informatics and Bioinformatics. 2020;9(1):42. https://doi.org/10.1007/s13721-020-00248-5.

12. Chen Q, Lee K, Yan S, Kim S, Wei CH, Lu Z. Bioconceptvec: Creating and evaluating literature-based biomedical concept embeddings on a large scale. PLoS Comput Biol. 2020. https://doi.org/10.1371/journal.pcbi.1007617.

13. Zhang Y, Chen Q, Yang Z, Lin H, Lu Z. BioWordVec, improving biomedical word embeddings with subword information and MeSH. Sci Data. 2019. https://doi.org/10.1038/s41597-019-0055-0.

14. Chen Q, Peng Y, Lu Z. BioSentVec: creating sentence embeddings for biomedical texts. IEEE International Conference on Healthcare Informatics (ICHI). 2019. https://doi.org/10.1109/ichi.2019.8904728.

15. Satapathy R, Cambria E, Amir H. Sentiment Analysis in the Bio-Medical Domain. Springer International Publishing. 2017.

16. Han J, Zhang Z, Cummins N, Schuller B. Adversarial training in affective computing and sentiment analysis: recent advances and perspectives [Review Article]. IEEE Comput Intell Mag. 2019. https://doi.org/10.1109/MCI.2019.2901088.

17. Nazir A, Rao Y, Wu L, Sun L. Issues and challenges of aspect-based sentiment analysis: a comprehensive survey. IEEE Trans Affect Comput. 2020. https://doi.org/10.1109/TAFFC.2020.2970399.

18. Jiménez-Zafra SM, Martín-Valdivia MT, Molina-González MD, Ureña-López LA. How do we talk about doctors and drugs? Sentiment analysis in forums expressing opinions for medical domain. Artif Intell Med. 2019. https://doi.org/10.1016/j.artmed.2018.03.007.

19. Rodrigues RG, das Dores RM, Camilo-Junior CG, Rosa TC. SentiHealth-Cancer: A sentiment analysis tool to help detecting mood of patients in online social networks. Int J Med Inform. 2014. https://doi.org/10.1016/j.ijmedinf.2015.09.007.

20. Garcia-Pelaez J, Rodriguez D, Medina-Molina R, Garcia-Rivas G, Jerjes-Sánchez C, Trevino V. PubTerm: A web tool for organizing, annotating and curating genes, diseases, molecules and other concepts from PubMed records. Database. 2019. https://doi.org/10.1093/database/bay137.

21. Patel R, Chang T, Greysen SR, Chopra V. Social media use in chronic disease: A systematic review and novel taxonomy. Am J Med. 2015. https://doi.org/10.1016/j.amjmed.2015.06.015.

22. Kwon D, Kim S, Wei CH, Leaman R, Lu Z. EzTag: Tagging biomedical concepts via interactive learning. Nucleic Acids Res. 2018. https://doi.org/10.1093/nar/gky428.

23. Zhang Z, Han J, Deng J, Xu X, Ringeval F, Schuller B. Leveraging unlabeled data for emotion recognition with enhanced collaborative semi-supervised learning. IEEE Access. 2018. https://doi.org/10.1109/ACCESS.2018.2821192.

24. Canales L, Strapparava C, Boldrini E, Martinez-Barco P. Intensional learning to efficiently build up automatically annotated emotion corpora. IEEE Trans Affect Comput. 2020. https://doi.org/10.1109/TAFFC.2017.2764470.

25. Usama M, Xiao W, Ahmad B, Wan J, Hassan MM, Alelaiwi A. Deep learning based weighted feature fusion approach for sentiment analysis. IEEE Access. 2019;7:140252–60. https://doi.org/10.1109/ACCESS.2019.2940051.

26. Gupta R, Sahu S, Espy-Wilson C, Narayanan S. Semi-supervised and transfer learning approaches for low resource sentiment classification. In ICASSP, IEEE International Conference on Acoustics, Speech and Signal Processing - Proceedings. 2018. https://doi.org/10.1109/ICASSP.2018.8461414.

27. Dietrich J, Gattepaille LM, Grum BA, Jiri L, Lerch M, Sartori D, Wisniewski A. Adverse Events in twitter-development of a benchmark reference dataset: results from IMI WEB-RADR. Drug Safety. 2020. https://doi.org/10.1007/s40264-020-00912-9.

28. Tutubalina E, Miftahutdinov Z, Nikolenko S, Malykh V. Medical concept normalization in social media posts with recurrent neural networks. J Biomed Inform. 2018;84:93–102. https://doi.org/10.1016/j.jbi.2018.06.006.

29. Wei CH, Kao HY, Lu Z. PubTator: a web-based text mining tool for assisting biocuration. Nucleic Acids Res. 2013. https://doi.org/10.1093/nar/gkt441.

30. Bradet-Levesque I, Audet J, Roy J-S, Flamand VH. Measuring functional outcome in dupuytren's disease: A systematic review of patient-reported outcome measures. J Hand Ther. 2021. Available from: https://www.sciencedirect.com/science/article/pii/S0894113021000570. https://doi.org/10.1016/j.jht.2021.04.010.

31. Pérez-Alfonso KE, Sánchez-Martínez V. Electronic patient-reported outcome measures evaluating cancer symptoms: A systematic review. Seminars in Oncology Nursing. 2021;37(2):151145. Available from: https://www.sciencedirect.com/science/article/pii/S0749208121000280. https://doi.org/10.1016/j.soncn.2021.151145.

32. Hansen CF, Obionu KC, Comins JD, Krogsgaard MR. Patient reported outcome measures for ankle instability. an analysis of 17 existing questionnaires. Foot Ankle Surg. 2021. Available from: https://www.sciencedirect.com/science/article/pii/S1268773121000874. https://doi.org/10.1016/j.fas.2021.04.009.

33. Sultan P, Ando K, Sultan E, Hawkins JE, Chitneni A, Sharawi N, Sadana N, Blake LE, Singh PM, Flood P, Carvalho B. A systematic review of patient-reported outcome measures used to assess postpartum pain using consensus based standards for the selection of health measurement instruments (cosmin) guidelines. Br J Anaesth. 2021. Available from: https://www.sciencedirect.com/science/article/pii/S000709122100235X.https://doi.org/10.1016/j.bja.2021.03.035.

34. Cambria E, Benson T, Eckl C, Hussain A. Sentic PROMs: Application of sentic computing to the development of a novel unified framework for measuring health-care quality. Expert Systems with Applications. 2012. https://doi.org/10.1016/j.eswa.2012.02.120.

35. Sailunaz K, Alhajj R. Emotion and sentiment analysis from Twitter text. J Comput Sci. 2019. https://doi.org/10.1016/j.jocs.2019.05.009.

36. Ruz GA, Henríquez PA, Mascareño A. Sentiment analysis of Twitter data during critical events through Bayesian networks classifiers. Futur Gener Comput Syst. 2020. https://doi.org/10.1016/j.future.2020.01.005.

37. Ghiassi M, Lee S. A domain transferable lexicon set for Twitter sentiment analysis using a supervised machine learning approach. Expert Systems with Applications. 2018. https://doi.org/10.1016/j.eswa.2018.04.006.

38. Hutto C, Gilbert E. Vader: A parsimonious rule-based model for sentiment analysis of social media text. Proceedings of the International AAAI Conference on Web and Social Media. 2014;8(1). Available from: https://ojs.aaai.org/index.php/ICWSM/article/view/14550.

39. Pasupa K, Ayutthaya TS. Thai sentiment analysis with deep learning techniques: A comparative study based on word embedding, POS-tag, and sentic features. Sustain Cities Soc. 2019;50:101615. https://doi.org/10.1016/j.scs.2019.101615.

40. Bravo-Marquez F, Mendoza M, Poblete B. Meta-level sentiment models for big social data analysis. Knowl-Based Syst. 2014. https://doi.org/10.1016/j.knosys.2014.05.016.

41. Crannell WC, Clark E, Jones C, James TA, Moore J. A pattern-matched Twitter analysis of US cancer-patient sentiments. J Surg Res. 2016. https://doi.org/10.1016/j.jss.2016.06.050.

42. Hassan SU, Aljohani NR, Idrees N, Sarwar R, Nawaz R, Martínez-Cámara E, Ventura S, Herrera F. Predicting literature's early impact with sentiment analysis in Twitter. Knowl-Based Syst. 2020. https://doi.org/10.1016/j.knosys.2019.105383.

43. Cambria E, Li Y, Xing FZ, Poria S, Kwok K. Sentic-Net 6: Ensemble Application of Symbolic and Subsymbolic AI for Sentiment Analysis. International Conference on Information and Knowledge Management, Proceedings. 2020:105–114. https://doi.org/10.1145/3340531.3412003.

44. Baccianella S, Esuli A, Sebastiani F. SentiWordNet 3.0: An Enhanced Lexical Resource for Sentiment Analysis and Opinion Mining. In: Proceedings of the Seventh International Conference on Language Resources and Evaluation (LREC'10). Valletta, Malta: European Language Resources Association (ELRA). 2010. http://www.lrec-conf.org/proceedings/lrec2010/pdf/769_Paper.pdf.

45. Nielsen FA. A new ANEW: Evaluation of a word list for sentiment analysis in microblogs. CoRR. 2011;abs/1103.2903.

46. Zhong BL, LuoW, Li HM, Zhang QQ, Liu XG, Li WT, et al. Knowledge, attitudes, and practices towards COVID-19 among chinese residents during the rapid rise period of the COVID-19 outbreak: A quick online cross-sectional survey. Int J Biol Sci. 2020. https://doi.org/10.7150/ijbs.45221.

47. Han Y, Liu M, Jing W. Aspect-Level Drug Reviews SentimentAnalysis Based on Double BiGRU and Knowledge Transfer. IEEEAccess. 2020. https://doi.org/10.1109/ACCESS.2020.2969473.

48. Grisstte H, Nfaoui E. Daily life patients sentiment analysis model based on well-encoded embedding vocabulary for related-medication text. Proceedings of the 2019 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining, ASONAM 2019. 2019:921–928. https://doi.org/10.1145/3341161.3343854.

49. Cambria E, Hussain A, Durrani T, Havasi C, Eckl C, Munro J. Sentic computing for patient centered applications. In: International Conference on Signal Processing Proceedings, ICSP. 2010. https://doi.org/10.1109/ICOSP.2010.5657072.

50. Cambria E, Hussain A. Sentic Computing. Springer International Publishing. 2015.

51. Ji S, Pan S, Li X, Cambria E, Long G, Huang Z. Suicidal ideation detection: A review of machine learning methods and applications. IEEE Transactions on Computational Social Systems. 2021;8(1):214–26. https://doi.org/10.1109/TCSS.2020.3021467.

52. Zucco C, Bella S, Paglia C, Tabarini P, Cannataro M. Predicting Abandonment in Telehomecare Programs Using Sentiment Analysis: A System Proposal. In: 2018 IEEE International Conference on Bioinformatics and Biomedicine (BIBM). 2018. p. 1734–1739. https://doi.org/10.1109/BIBM.2018.8621481.

53. Ofek N, Caragea C, Rokach L, Biyani P, Mitra P, Yen J, et al. Improving Sentiment Analysis in an Online Cancer Survivor Community Using Dynamic Sentiment Lexicon. In: 2013 International Conference on Social Intelligence and Technology; 2013. p. 109–113. https://doi.org/10.1109/SOCIETY.2013.20.

54. Jelodar H, Wang Y, Orji R, Huang S. Deep sentiment classification and topic discovery on novel coronavirus or covid-19 online discussions: Nlp using lstm recurrent neural network approach. IEEE J Biomed Health Inform. 2020;24(10):2733–42. https://doi.org/10.1109/JBHI.2020.3001216.

55. Sanglerdsinlapachai N, Plangprasopchok A, Ho TB, Nantajeewarawat E. Improving sentiment analysis on clinical narratives by exploiting umls semantic types. Artif Intell Med. 2021;113:102033 https://www.sciencedirect.com/science/article/pii/S0933365721000269. https://doi.org/10.1016/j.artmed.2021.102033.

56. Singh R, Singh R. Applications of sentiment analysis and machine learning techniques in disease outbreak prediction – A review. Materials Today: Proceedings. 2021. https://www.sciencedirect.com/science/article/pii/S2214785321032764. https://doi.org/10.1016/j.matpr.2021.04.356.

57. Mikolov T, Sutskever I, Chen K, Corrado G, Dean J. Distributed Representations of Words and Phrases and their Compositionality. CoRR. 2013;abs/1310.4546.

58. Pennington J, Socher R, Manning CD. GloVe: Global vectors for word representation. In: EMNLP 2014 - 2014 Conference on Empirical Methods in Natural Language Processing, Proceedings of the Conference; 2014. https://doi.org/10.3115/v1/d14-1162.

59. Mikolov T, Grave E, Bojanowski P, Puhrsch C, Joulin A. Advances in Pre-Training Distributed Word Representations. CoRR. 2017;abs/1712.09405.

60. Kosmopoulos A, Androutsopoulos I, Paliouras G. Biomedical Semantic Indexing using Dense Word Vectors in BioASQ. Supplement On Bio-Medical Information Retrieval: Journal Of Bio-Medical Semantics; 2015.

61. Jing K, Xu J. A Survey on Neural Network Language Models; 2019.

62. Wei CH, Harris BR, Li D, Berardini TZ, Huala E, Kao HY, et al. Accelerating literature curation with text-mining tools: a case study of using PubTator to curate genes in PubMed abstracts. Database : the journal of biological databases and curation. 2012;2012:bas041. http://www.ncbi.nlm.nih.gov/pubmed/23160414, http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC3500520. https://doi.org/10.1093/database/bas041.

63. Fu M, Qu H, Huang L, Lu L. Bag of meta-words: A novel method to represent document for the sentiment classification. Expert Systems with Applications. 2018. https://doi.org/10.1016/j.eswa.2018.06.052.

64. Shuang K, Zhang Z, Guo H, Loo J. A sentiment information Collector–Extractor architecture based neural network for sentiment analysis. Inf Sci. 2018. https://doi.org/10.1016/j.ins.2018.08.026.

65. Pham DH, Le AC. Exploiting multiple word embeddings and one-hot character vectors for aspect-based sentiment analysis. International Journal of Approximate Reasoning. 2018. https://doi.org/10.1016/j.ijar.2018.08.003.

66. Rezaeinia SM, Rahmani R, Ghodsi A, Veisi H. Sentiment analysis based on improved pre-trained word embeddings. Expert Systems with Applications. 2019. https://doi.org/10.1016/j.eswa.2018.08.044.

67. Weiss K, Khoshgoftaar TM, Wang DD. A survey of transfer learning. Journal of Big Data. 2016. https://doi.org/10.1186/s40537-016-0043-6.

68. DaiW, Xue GR, Yang Q, Yu Y. Transferring Naive Bayes Classifiers for Text Classification. In: Proceedings of the 22nd National Conference on Artificial Intelligence - Volume 1. AAAI'07. AAAI Press; 2007. p. 540–545.

69. Adams N. Dataset Shift in Machine Learning. J R Stat Soc A Stat Soc. 2009;173(1):274. https://doi.org/10:1111/j:1467-985x:2009:0062410:x.

70. Jiang J, Zhai CX. Instance weighting for domain adaptation in NLP. In: ACL 2007 - Proceedings of the 45th Annual Meeting of the Association for Computational Linguistics; 2007. p. 264–271.

71. Ando RK, Zhang T. A high-performance semi-supervised learning method for text chunking. In: ACL-05 - 43rd Annual Meeting of the Association for Computational Linguistics, Proceedings of the Conference; 2005. https://doi.org/10.3115/1219840.1219841.

72. Blitzer J, McDonald R, Pereira F. Domain adaptation with structural correspondence learning. In: COLING/ACL 2006 -EMNLP 2006: 2006 Conference on Empirical Methods in Natural Language Processing, Proceedings of the Conference; 2006. https://doi.org/10.3115/1610075.1610094.

73. Davis J, Domingos P. Deep transfer via second-order Markov logic. In: ACM International Conference Proceeding Series; 2009. https://doi.org/10.1145/1553374.1553402.

74. Mihalkova L, Huynh T, Mooney RJ. Mapping and Revising Markov Logic Networks for Transfer Learning. In: Proceedings of the 22nd National Conference on Artificial Intelligence - Volume 1. AAAI'07. AAAI Press; 2007. p. 608–614.

75. Howard J, Ruder S. Universal language model fine-tuning for text classification. In: ACL 2018 - 56th Annual Meeting of the Association for Computational Linguistics, Proceedings of the Conference (Long Papers). 2018. https://doi.org/10.18653/v1/p18-1031.

76. Sarker A, Belousov M, Friedrichs J, Hakala K, Kiritchenko S, Mehryary F, et al. Data and systems for medication-related text classification and concept normalization from Twitter: Insights from the Social Media Mining for Health (SMM4H)-2017 shared task. J Am Med Inform Assoc. 2018. https://doi.org/10.1093/jamia/ocy114.

77. Limsopatham N, Collier N. Normalising medical concepts in social media texts by learning semantic representation. In: 54th Annual Meeting of the Association for Computational Linguistics, ACL 2016 - Long Papers. 2016. https://doi.org/10.18653/v1/p16-1096.

78. Guo ZH, You ZH, Huang DS, Yi HC, Zheng K, Chen ZH, et al. MeSHHeading2vec: a new method for representing MeSH headings as vectors based on graph embedding algorithm. Brief Bioinform. 2020. https://doi.org/10.1093/bib/bbaa037.

79. Nikfarjam A, Sarker A, O'Connor K, Ginn R, Gonzalez G. Pharmacovigilance from social media: Mining adverse drug reaction mentions using sequence labeling with word embedding cluster features. J Am Med Inform Assoc. 2015. https://doi.org/10.1093/jamia/ocu041.

80. Huang GB, Zhu QY, Siew CK. Extreme learning machine: A new learning scheme of feedforward neural networks. In: IEEE International Conference on Neural Networks - Conference Proceedings. 2004. https://doi.org/10.1109/IJCNN.2004.1380068.

81. Knuth D. TextBlob Documentation. 2020. Available from: https://textblob.readthedocs.io/en/dev/.

82. Cambria E, Olsher D, Rajagopal D. SenticNet 3: a common and common-sense knowledge base for cognition-driven sentiment analysis. In: Proceedings of the Twenty-Eighth AAAI Conference on Artificial Intelligence. AAAI'14. AAAI Press; 2014. p. 1515–1521.

83. Gamon M. Sentiment classification on customer feedback data: noisy data, large feature vectors, and the role of linguistic analysis. Proceedings of the 20th international conference on Computational Linguistics. 2004. https://doi.org/10.3115/1220355.1220476.

84. Pang B, Lee L, Vaithyanathan S. Thumbs up? Sentiment classification using machine learning techniques. In Proceedings of the ACL-02 Conference on Empirical Methods in Natural Language Processing, EMNLP '02, Association for Computational Linguistics. 2002;10:79–86. https://doi.org/10.3115/1118693.1118704

85. Elfaik H, et al. Deep bidirectional lstm network learning-based sentiment analysis for arabic text. J Intell Syst. 2020;30(1):395–412.