



# Deep associative learning approach for bio-medical sentiment analysis utilizing unsupervised representation from large-scale patients' narratives

Hanane Grissette<sup>1</sup> · El Habib Nfaoui<sup>1</sup>

Received: 1 July 2020 / Accepted: 29 June 2021 / Published online: 11 August 2021  
© The Author(s), under exclusive licence to Springer-Verlag London Ltd., part of Springer Nature 2021

## Abstract

Owing to the quick spread of minute health-related experiences, the distillation of knowledge from such unstructured narratives is an extremely challenging task. In spite of the success of neural networks methods in improving learning structural reliability, they result in inadequate accuracy of the bio-medical sentiment classification when employing less useful features sets. Therefore, they lack discriminatory potential. In this study, we propose to add a deep associative memory into neural networks for an effective sentiment decomposition, which emphasizes correctly on bio-medical entities related to the extraction of different data-object properties, and contextual-semantics dependencies for a given aspect. The underlying trust of these measures is behind the ability to compute the completion of unseen medical patterns, where comprehensive bio-medical distributed representations are used for representing the formal medical connections from PubMed databases. Experiments on a biomedical sentiment analysis task show that the model provides comprehensive embeddings with meaningful medical patterns. It achieved an average performance of 87% on varied large online datasets. It also outperforms baselines in discovering and identifying medical natural concepts. We provide meaningful support to bio-medical sentiment analysis applications in social networks. Indeed, the facets of this study might be used in many health concerns such as analyzing change in health status or unexpected situations.

**Keywords** Deep associative learning · Neural networks · Biomedical distributed representation · Sentiment analysis · Aspect-based sentiment

## 1 Introduction

Daily patient self-reports may reveal true facts based on life experiences, which alert us to various medical information such as drug abuse or drug reactions. Sentiment analysis (SA) is widely used to uncover negative and/or positive evaluations of patients towards a given entity or condition from online discussions [1–4]. Such a large amount of citations of medical concepts, events, addictions, and semantic relationships in drug-related thesis are affectively associated in different contexts. Massive examples of deep

learning sentiment services have been provided which explore rich embedding benefits, dimensionality reduction, and long-term dependencies on large dataset conversion (training) examples. Each working perfectly in the setting in which they were trained, but their performance decreased sharply when they became accustomed to analyzing patient stories online. They often lose ground in the face of more flexible knowledge-based challenges. This failure can be strongly associated with many problems, in particular the lack of formal medical language sentiment dictionaries, which is unapproachable and can be prone to annotator bias.

Since the last decade, several learning methods based on neural networks have been proposed, ranging from supervised methods to fully unsupervised methods. In particular, unsupervised methods associated with expert-annotated knowledge/lexicons or rule-based methods are considered to be more accurate and credible for assessing positive/negative patient reports. In particular, neural networks (NNs) are used effectively to solve many text problems related to the linguistics of drugs. The problem

---

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

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

<sup>1</sup> LISAC Laboratory, Faculty of Sciences Dhar El Mahraz,  
Sidi Mohamed Ben Abdellah University, Fez, Morocco

**Table 1** Epilepsy samples from collected dataset

Samples	Medical-related concepts
Ambiguate incidence and high case fatality rate of #COVID19 in patients with active #epilepsy	Epilepsy, COVID-19
“Dr Pot” stashes away a #cannabis fortune with some NHS help: Pharma boss earns £5.8m as drug for #epilepsy is approved for use in UK	D064420 drug for epilepsy, NHS, cannabis, Disease_MESH.D064420
amepilepsysoc terrific leadership team great paners advancing epilepsy awareness treatment thank sharing chi amp...	Epilepsy, amphetamine (AMP)
@aspiringheadr frequency takotsubo cardiomyopathy #epilepsy-related hospitalizations among adults impact in-hospit...	Takotsubo, cardiomyopathy
@adrienneolivier @tiaank89 year ago today able add vagus nerve implant peters treatment condition been...	Vagus nerve
aeds antiepileptic drugs besides #epilepsy also treat mood disorders #chronicpain &...	AEDs, antiepileptic, mood disorders

is that these methods fail because they do not properly focus on the medical components related to extracting different properties, concepts, medical events, and people from data objects, which they should be used to formulate interactions, semantics, dependencies, and inter-relations. Text accumulation errors and linguistic subproblems such as noisy (insignificant) text often occur due to the inability to resolve the ambiguity of these medical elements and define drug-related entities, resulting in a lower recall than desired. Most of the existing tools require highly skilled natural concept recognition capabilities, or huge labeled medical data, which are unaffordable. Extraction techniques based on medical concepts should be involved so as to collectively detect different concepts of related drugs regarding the dynamism of various medical situations. An intuitive solution is to take into account the properties and attributes of medical data objects, dependent domain semantics, and additional online attributes such as natural language concepts. Recent studies improve the accuracy of recognition of medical entities by linking social media data to their correspondence in health-related databases and pharmaceutical clusters, where it is computational and software costly to centralize all this knowledge from resources around the world.

Furthermore, referring to a previous qualitative analysis in this sense provided by [1], most SA approaches generally classify subjective medical events or sensitive medical entities as a negative sign of every polar story, which can seriously fail to recover the corresponding feeling. Thus, no study considers these entities as an analysis target, which aims to detect the relevant sentiment towards these entities for a given aspect. As an example in [5], the feelings conveyed towards the concept of “Epilepsy”<sup>1</sup> are all considered negative where this disease can also appear in positive quotes. As shown in the Table 1, it has appeared

in both several positive patient reports, as well as negative patient reports, due to the frequent associations with SUDEP concepts as a life-threatening complication that generally seemed to arise when a person with epilepsy dies so unexpected and was previously in its usual state<sup>2</sup> of health. An estimated study reported that SUDEP takes more lives each year in the USA than other epilepsy-related deaths such as sudden infant death syndrome (SIDS), and their concepts have appeared frequently in accounts of epilepsy in different ways; Table 3 highlighted these related concepts and their specific meaning in the context of epilepsy. Additionally, many related natural remedies are flagged up in online conversations as treatment issues or usage descriptions, as shown in the Table 2; unfortunately, related concepts have been ignored by controlled biomedical vocabulary such as manganese, taurine, seizure reduction, and CBD oil, which is variously cited in terms of side effects and problems treatment without formal reports by medical experts.

This study addressed the following research question: Considering the various medical components in the text, associative learning mechanisms can be proposed offering better results in terms of capacities of conceptualization, generalization, recognition of medical familiarity, and categorization based on aspects. In this work, we discuss the need to capture, quantify, and memorize a dependency measure considering online dynamicity by formulating bidirectional associative properties, interactions, and data objects. Associative learning is mainly about expressing ideas and experiences reinforcing each other, which are furthermore interrelated, making it a powerful learning strategy [6]. It increases the levels of abstraction with communication between compound patterns, as well as patterns of learning feeling and inference. In particular, this paper investigated an associative learning approach through multiple convolutions to build a dynamic configuration

<sup>1</sup><https://www.mayoclinic.org/diseases-conditions/epilepsy/symptoms-causes>

<sup>2</sup><https://www.epilepsy.com/learn/early-death-and-sudep/sudep/>

**Table 2** Some related concepts have been cited in many stated natural remedies used for some ailments

Concept	Description
CBD Oil	CBD is one of over 100 chemical compounds known as cannabinoids found in the cannabis or marijuana plant, <i>Cannabis sativa</i> .
Cannabidiol	Cannabidiol is a popular natural remedy used for many common ailments.
THC	Tetrahydrocannabinol (THC) is the main psychoactive compound in marijuana that gives the high sensation.

space from unsupervised medical concepts representations. We approach a deep associative neural network based on biomedical distributed representation which imitates such a learning process to learn the network parameters. Furthermore, the model promised to allow modelling medical patterns by identifying every time and inserted into space with corresponding sentimental value with little support evidence. The model also has the ability to calculate the completion of various medical patterns that have been missed or observed informally by convolving multiple convolutional neural network channels on the basis of old associative characteristics based on concepts. Those former patterns will be deployed later by our bio-medical sentiment decomposition aspect-based algorithm. The contributions of this work are summarized as follows:

- 1. Collect, standardize, and correct content shared online by patients. A representation of biomedical inclusion (embedding) for various medical components is involved.
- 2. Define a deep associative learning method to learn the properties of data objects from unstructured data, in conjunction with with medical comprehensive medical vocabularies. An associative neural network computation in which future values mimic such a learning process from well-understood biomedical patterns.
- 3. Deploy the former patterns for sentiment decomposition regarding the aspects that are extracted. A statistical approach is also used to improve performance accuracy.

In addition, we are conducting a comparative study between state-of-the-art deep learning algorithms utilizing our embeddings-based model for bio-medical sentiment analysis task. Indeed, we examined four cases of efficient vector derivation based on biomedical inclusion, whereby we discuss the quality of the study by performing an explanatory

analysis using the positive and negative messages of the patients on Twitter. The experiments have proven the effectiveness of the abstraction of medical concepts and demonstrated how good our proposed model is.

The rest of this article is structured as follows. Section 2 presents a summary of the literature review in the context of sentiment analysis for healthcare. Section 3 presents the proposed architecture with more details on the problem definition. Section 4 offers extensive results of experimentation on social networks and comparisons with the medical vocabularies gain through convolving convolutions on biomedical language inclusion. Section 5 concludes this article and provides some perspective.

## 2 Background and literature review

This section describes and reviews previous research work done in the area of bio-medical sentiment analysis approaches. Besides, we will discuss the theory of deep associative learning for unsupervised representations recognition.

### 2.1 Bio-medical sentiment analysis

Sentiment analysis (SA) has witnessed huge contributions in recent years. Healthcare organizations and pharmaceutical companies have already used this potential of SA techniques applied to narratives and stories generated on social media in many tasks: capturing adverse drug reactions, detecting product problems, categorizing patients’ status, and therefore building interpretable sentiment analysis models for personalization of various patient statements, etc. Technically, SA is the task that aims to automatically quantifying the sentiment expressed towards a giving entity, usually by given predefined sentiment assumptions. Exploring the

**Table 3** Frequent medical concept used for epilepsy treatment

Concept	Description
SUDEP	SUDEP sudden unexpected death in epilepsy
AEDS	Anti-epileptic drugs (AEDs), and therapeutic drugs monitoring for phenytoin, carbamazepine, phenobarbital, etc.
SCN2A	Movement disorders seem to be common in patients with epileptic encephalopathy. SCN2A encodes an alpha subunit in a voltage-gated sodium channel

methods used for sentiment analysis in healthcare research may help us better understand the options available for future research in this growing field.

Social networks are increasingly popular as a platform for sharing personal health information [7]. This information can be used for public health surveillance tasks, in particular for extracting and tracking drug-related data, such as drug efficacy or drug abuse incidents. Patients' narratives on social media can reveal critical public health issues [3]. In particular, sentiment is a commonly used metric to investigate the positive or negative reports contained in these messages. For example, authors in [8] analyzed a tweet of risky behavior related to drug abuse as a positive tweet, or a tweet of safe behavior as a negative tweet. A rich variety of techniques have been proposed in this context ranging from dictionary-based methods using annotated knowledge, supervised machine learning methods to fully unsupervised methods, in conjunction with statistical methods such as affinity analysis and cluster occurrences of words, thereby inducing word sentiment intensities. Among these, supervised learning approaches have been the most successful algorithms to date.

In the bio-medical contexts, the lack of meaning of words, the inability to properly match the medical context and the clarification of medical entities are the most crucial challenges in bio-medical SA [1]. Deep learning techniques for sentiment analysis have become very popular; in particular, they provide automatic feature extraction and both richer representation capabilities and better performance than traditional feature-based techniques [9]. In particular, they are widely used to build distributed vocabulary and to overcome the limitations of manually labeled lexicons. In [1], this is an N-gram-based convolution vocabulary scheme, which is primarily dedicated to featuring text in a medical setting and clarifying associated feelings at the same level. Nevertheless, most existing deep learning-based SA approaches largely focus on grammar directions such as semantic direction or focus only on extracting sentiment words [10]. They have widely adopted pre-trained incorporations to exploit semantics as much as possible and contextual information, especially pre-trained contextualized word vector representations, have been widely used to create contextual and semantic vocabularies [11], and therefore used to induce sentiment-related information by deploying these representations and their attributes. Two main approaches involved in this case, (i) supervised methods applying an induction update based on lexicons. For example, count positive and negative term scores to determine sentiment direction, and improvement is presented by creating a dataset of relevant features using predefined sources such as SentiWordNet [12]. Unsupervised methods, in particular neural network-based approaches, exploit unstructured data to generate and retrieve hidden sentiment

information by identifying the constraints of conjunctions on the positive or negative semantic orientations [13, 14]. The latter have also been used to expand the size of sentiment lexicons and learn from new raw data, which further succeeded in the accuracy of many tasks of disambiguating the meaning of words, analyzing sentiment, and find out new knowledge.

## 2.2 Distributed medical language

Referring to [10], the heart of the effective quantification of aspect-based sentiments relating to drugs/treatments hardly involves a degree of medical language comprehension. Most of the existing research, in this case, assumes the existence of hand-annotated sentiment dictionaries, which is time consuming and may be subject to annotator/expert bias.

Recently, many medical organizations have built a format of continuous biomedical representation format of their ontologies such as PubMed<sup>3</sup>. It is a free search engine that primarily accesses the MEDLINE<sup>4</sup> database of references and abstracts on life sciences and biomedical topics that comprises more than 29 million citations for biomedical literature, journal of life sciences, and books online. To date, the methods have rapidly adopted and combined different types of medical dictionaries. The authors in [15] applied embedding approaches on a corpus of 10,876,004 English abstracts of biomedical articles from PubMed to estimate the relatedness of two words, and combine subword information from unlabeled biomedical text with a widely used biomedical controlled vocabulary called Medical Subject Headings (MeSH), where [16] provided a training set of sentence embeddings with over 30 million documents from both scholarly articles in PubMed and clinical notes in the MIMIC III Clinical database to prove the effectiveness of features of discriminating expressions by employing pair similarity tasks on different genres of biomedical texts.

For instance, the use of embeddings has shown promising performance in the extraction and tagging of biomedical concepts. For example, BioConceptVec [17] employed PubTator [18] (a state-of-the-art Named Entity Recognition system) via four different word embeddings techniques on 30 million PubMed abstracts. It covers more than 400,000 biomedical concepts mentioned in the literature and is to date one of the most important among the biomedical concepts embeddings available to the public. However, word embeddings are traditionally computed at the word level from a large body of unlabeled text, ignoring information present in the internal structure of words or any information available in structured domain-specific resources

<sup>3</sup><https://pubmed.ncbi.nlm.nih.gov/>

<sup>4</sup><https://www.primarynlm.nih.gov/bsd/pmresources.html>

such as ontologies. Such knowledge has the potential to dramatically improve the quality of word representation, as suggested in some recent research [10, 19]. Nevertheless, these vocabularies are obviously not able to distinguish an unseen natural medical concept such as a multi-word phrase indicating undefined drug abuse. In this article, we investigate an associative embedding mechanism that mainly consists of processing distributed medical concepts in natural language to better match the formal semantics of a widely used controlled biomedical vocabulary and contribute to define unrelated items.

### 2.3 Deep associative memory for unsupervised learning

As mentioned earlier, the success of these learning algorithms generally depends on data representation [11]. Due to the complexity in understanding shared patient experiences on social media, the frequent use of informal medical language, non-standard format, poor spelling, and abbreviation forms, as well as typos in social media posts, should be taken into consideration. However, adopting an existing lexicon for sentiment analysis in health-related datasets provides inadequate classification precision, which uses less useful feature sets and therefore lacks discriminatory potential [7]. There has been limited progress in addressing these challenges, and so far advanced machine learning-based entity recognition and NLP techniques have been underutilized, especially for the huge textual challenges of medication-related texts. Unsupervised methods have received less attention and there is less work on automatic labeling of medical components and sentiment for given aspects. It is widely believed that end-to-end training with the backpropagation algorithm helps solve most of these challenges, and plays the main role in retrieving informative feature detectors, with reference to [20]. These detectors are useful for the task performed by the upper layers of that neural network. Additional unsupervised parameters are achieved by calculating hidden units or by convolving complementary neural components by following particular learning rules. To conclude, the generalized version of the well-known Hopfield model and its extensions are considered good examples that inhibit the early feature detector in a completely unsupervised manner.

The deep associative memory is one of the most successful approaches which greatly increase the modelling and abstraction capabilities of networks, although training a deep network is much more difficult [21]. In this paper, we aim at investigating deep associative neural networks based on learning unsupervised representations, which is primarily devoted to learning features from large-scale unstructured data in conjunction with pre-trained biomedical embeddings. This family corresponds to the modeling

associative memory for pattern recognition by taking advantage of the greatest computational properties of deep learning. Referring to [22], we could categorize them into two main subset algorithms: (1) probability-based associative memory. (2) neural associative memory. Probability-based associative memory consists of a multilayer array of processing elements (PEs) [23], where neural associative memory stores patterns in-memory storage matrix [24].

We argue that, while associative learning models often assume an arbitrary connectionist architecture to define relationships, deep networks learn representations of local patterns that may be less efficient at defining a pattern of unobserved entities. Recent distributions have followed this line and have implemented acceptable solutions by improving learning ability using hierarchical representations of learning [22], convolutional and ultimately more precise predictive learning models [20, 21], although there is an on-going debate on the explanatory power of associative learning theory. For this, we were inspired by the process of learning [22] and learning from errors [23], and we proved the role of deep associative learning in defining concept-based biomedical variants. The main point of the controversy, however, does not call into question the experimental evidence, of which there is abundance, but whether this evidence is proven by recent models within the framework of the traditional theory of associative learning.

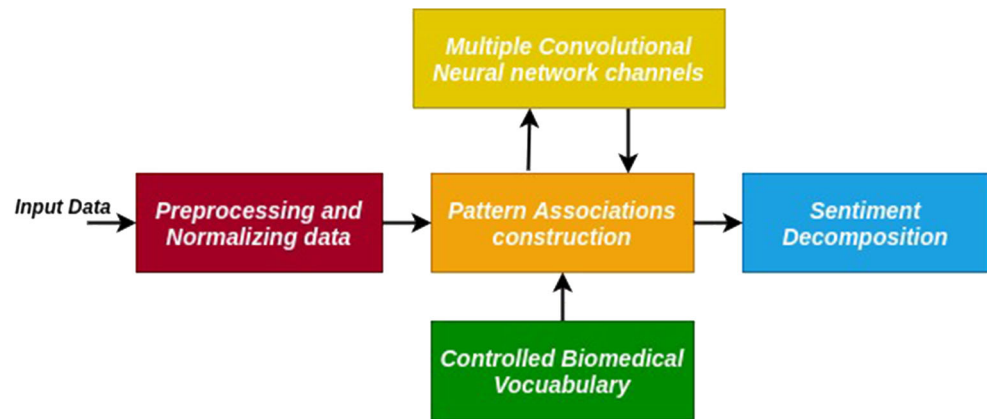
### 3 Proposed approach

A key problem in sentiment analysis is its sensitivity to the domain from which the training data originates or on which a sentiment lexicon is built. Existing pre-formed world knowledge provides insufficient context to reliably determine the sentiment polarity and builds upon personal knowledge and understanding.

The design principle of our approach is to add approximate associative reasoning into a neural network model for medication-related sentiment decomposition task (Fig. 1). Let  $D$  be a collection of patient narratives made up of  $N$  documents  $d_k$ ,  $k \in \{1 : N\}$ , heuristic sentiment intensity  $v_k$  which would be projected into the target space of three class labels, entitled “**POSITIVE**,” “**NEGATIVE**,” or “**NEUTRAL**.” A given patient report  $d_k$ ,  $d_k \in D$  is formed by a set of  $m$  sentences  $p_1, p_2, \dots, p_m$ . Each sentence  $p_j$  is a sequence of words  $p_j = w_{j1}w_{j2}\dots w_{jl}$ , where  $l$  is the size of  $p_j$ , the goal is to (i) extract medical patterns, (ii) project medical sentences into the distributional medical target space, and (iii) calculate their attached sentiments.

Meaningful learning involves in-depth knowledge of the context of the facts learned. This is particularly important for understanding the patients’ experiences in the many contexts sampled in text. In particular, we ensure the credibility

**Fig. 1** The overall process of the proposed approach



of our model comprehension by incorporating knowledge from formal biomedical ontologies and journals of life science. In different word-level abstractions, as depicted in Fig. 3, we explore a basic knowledge transfer mechanism that studies both biomedical distributed representation and universal word representation. The system has been initialized in addition to the following predefined basic knowledge used to represent the connections of basic medical patterns connections.

- **Clinical Notes MIMIC III and PubMed database [15]**: This knowledge consists of 10,876,004 English abstracts of biomedical articles. Authors in [15] provide distributed representations in many formats such as Word2Vec; they also combine subword information from unlabeled biomedical texts with a widely used controlled biomedical vocabulary called Medical Subject Headings (MeSH).
- **Central PubMed resources [17]**: It covers over 400,000 biomedical concepts that involve a state-of-the-art named entity recognition system on 30 million PubMed abstracts, including concepts, drug-gene interactions, gene-gene interactions, protein-protein interaction, and drug-drug interaction.

The proposed medical semantic network allows bidirectional associative memory to complement cognitive structures with sufficient meaning and drug-related semantics, as well as the sentential context in which it occurs. Thus, the combined word representation captures both distributional and compositional semantics to calculate attached sentiment. To formalize a medical new pattern, an input vector  $x_k$  associated with pattern vector  $y_k$  is used to calculate the corresponding associative memory by summing the total related pattern experiences of learning  $W_k$  (depend weights matrix). Second, a convolutional component is used to deploy associate features and leverage properties of data objects effectively complete patterns. The entire architecture is illustrated in Fig. 2. To sum up, our approach attempts to create a related medical configuration space that focuses

on harnessing both sentimental and biomedical associations through unsupervised neural network associative memories.

### 3.1 Deep associative neural network model

Deep learning typically uses non-dynamic feedforward neural networks<sup>5</sup>. Truly distributed representation methods can capture various arbitrary aspects of the word depending on certain contexts as well as the semantics on the words. But, they are limited in relating and defining domain-dependent models and neglect incomplete (ambiguous) definitions and concepts such as the definition of drug reaction. Thus, standard embedding representations can lead to huge and deeper memory, as well as simultaneous and repeated recall of similar memory contents.

This article examines these questions by deploying neural associative memories and additional theory of collective memory, including generalization ability, familiarity recognition, categorization, error correction, and temporal sequence retention [25]. Our motivation behind using associative memory is that it allows us to store a much larger number of pattern and memory associations. In this work, we aim to add other salient points to (i) address the relationship between different properties of data objects and medical concepts and (ii) define the two models that link one model to another (hetero-associative memory) through bidirectional memory and also associates a pattern with itself (auto-associative memory) or by convolution with the corresponding medical pattern. This philosophy provides a model with the ability to recover missing medical models. When it comes to the huge shared patient experiences, a simple neural associative memory model fails to meet the text requirements of drugs, and dealing with complex data at scale on social media has been crucial because most associative models are not reported. Since the convolutional neural network (CNN) is a feedforward neural network with convolutions and parameterizable targets, it has a

<sup>5</sup><http://meetings.aps.org/link/BAPS.2017.MAR.P5.9>

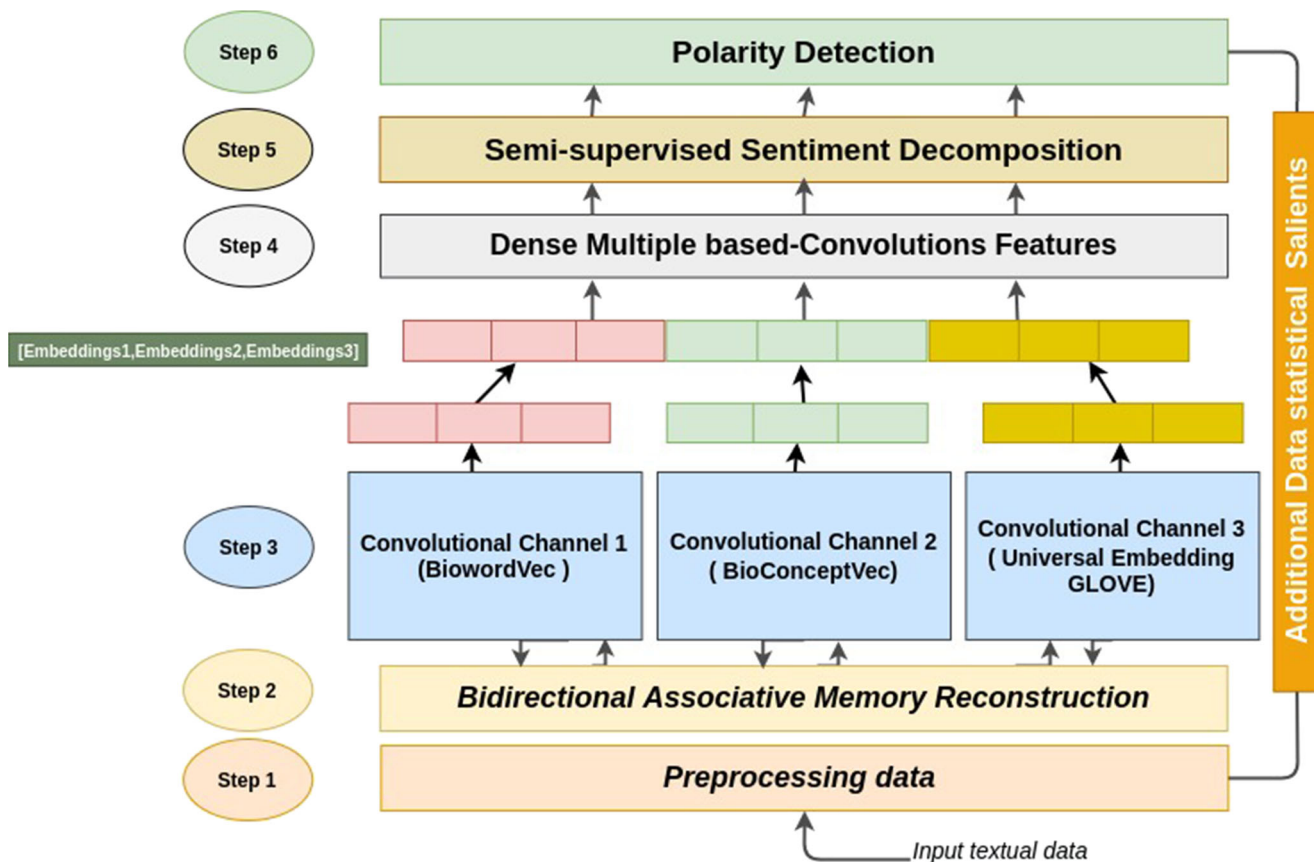


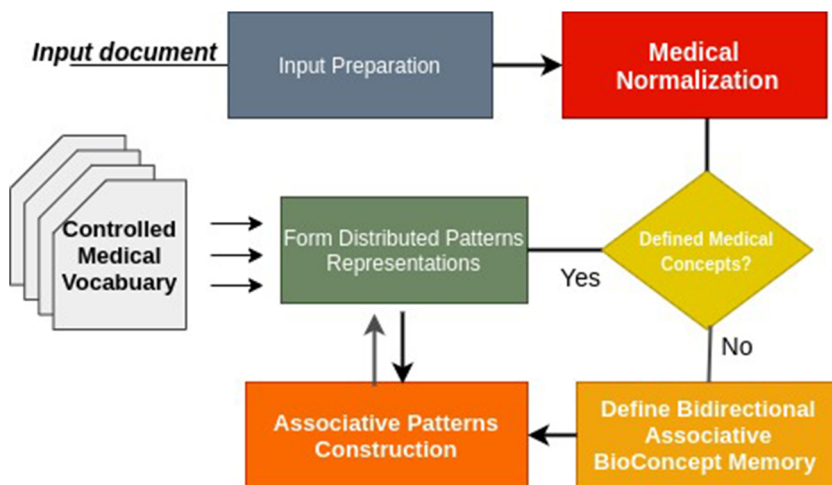
Fig. 2 The overall architecture

special spatial-local correlation by enforcing a model of local connectivity of pattern between neurons in adjacent layers, where associative memory algorithm retrieves patterns more than the number of neurons [26].

The algorithm 1 illustrates the typical computational goals discussed above. Associative memory is distributed in nature; this assumption is at the origin of the execution of storage mechanisms for a large number of models. A

normal associative memory task is to test participants on their recall of pairs of unrelated items. Given an input vector  $x_k$  associated with pattern vector  $y_k$ , we build memory space formalism by summing the total pattern experiences of learning  $W_k$ ( weights matrix). Each pattern structure is convolving  $m$ -neurons that are associated with the synoptic weights matrix to define the exact retrieval pattern, where  $m$  is the number of defined patterns fed to

Fig. 3 Illustration of the construction of distributed patterns representations



the network. Indeed, the exact retrieval aspect from the pool of memories, as shown in the algorithm 2, is calculated in the recursive form of the associative memory, and the associated weights concerning the pattern itself are strongly estimated. This means that we involve data entry and exit to retrieve the associated weight matrix, which each time passes to recalculate the output of unrelated medical phrases or correct associative links.

---

**Algorithm 1** Auto-associative neural network algorithm.
 

---

**Result:** Set of associated patterns  $x_i : c_i$   
 Initialize the weights  $i = 1$  to  $n$  and  $j = 1$  to  $n$ ;  
 $W_{i,j} = 0$ ;  
**for** Each vector  $x_i$  (to be stored) **do**  
 Set activation for each input Unit and Output unit respectively:  
**for**  $i = 1$  and to  $n$  **do**  
 |  $x_i = s_i$   
**end**  
**for**  $j = 1$  and to  $n$  **do**  
 |  $y_i = s_i$   
**end**  
**for**  $i = 1$  and to  $n$  and  $j = 1$  and to  $n$  **do**  
 |  $W = \sum_{p=1}^n S_{(p)}^T S_{(p)}$ ;  
 | activation vectors following (1);  
 |  $W_{new} = W_{old} + x_{i,j}$  (Adjust the weights)  
**end**  
 ;  
**end**

---

$$x_i = \begin{cases} 1 & \text{if } \sum_{p=1}^n w_i x_i > 0 \\ x_i & \text{if } \sum_{p=1}^n w_i x_i = 0 \\ -1 & \text{if } \sum_{p=1}^n w_i x_i < 0 \end{cases} \quad (1)$$

As shown in Fig. 2, we propose to design a related concept abstraction through multiple convolution channels that mainly deploy a distributed biomedical representation. Each channel generates a feature vector in which the powerful properties are summarized, and also stores and completes ambiguous patterns. This joint model used the potential of CNN in pattern extraction to powerfully create local and deeper dense associative representation maps, and help process/memorize data at scale where the components of convolutions manage scarcity and reduce dimensionality through the downsampling step.

---

**Algorithm 2** Testing and recorrecting associative neural network algorithm.
 

---

**Result:** Set of associated patterns  $x_i : x_i$   
**for** input  $x_i = s_i$  **do**  
 |  $W = S.S^T$  (Find the weight Matrix);  
**end**  
**while** well activated **do**  
 - Test the I/P vector with one mistake in the input vector.  
 - Test the network with one missing in the input vector.  
 - Test the network with two mistake in the input vector.  
 - activation vectors following (1)  
**end**

---

As deep learning increases learning ability with massive layer parameters, training difficulty and error rate increase respectively. Many contributions have been proposed to reduce errors and effectively build deeper networks with regard to the associative property of patterns. To this end, [24] has provided reliable storage and definition of many more models through dense associative memory learning techniques. We explore this method which consists of feedforward neural networks with one hidden layer and various activation functions to analyze the computational properties of neural networks.

### 3.2 Modelling sentiment decomposition

Neural network methods have seen huge distributions, where they have mainly deployed embeddings features to track sentiment information. But, by taking these approaches to getting sentiment insight from patient self-reports on social media, get as slow as desired accuracy as the model has no knowledge of medical entities. We propose to deploy both the approximate modeling obtained in the previous steps and the approximate reasoning for the composition of feelings; this improved our chances of clarifying the intensity of sentiments from large complex data. We have developed a semi-supervised sentiment scale which consists both in exploiting associations of embeddings and in incorporating lexical and semantic relationships between SentiWordNet<sup>6</sup> and SenticNet<sup>7</sup>. Table 9 summarizes the statistics and the level of automation of these.

In addition, SenticNet has also been adopted to improve the accuracy of sentiment classification on Twitter by

<sup>6</sup><http://SentiWordNet.isti.cnr.it/>

<sup>7</sup><https://sentic.net/>



extracting sentiment indicators from natural language texts at the syntactic and semantic levels, which is called sentic computing, which aims to bridge the gap between statistical processing of natural language and many other disciplines necessary to understand human languages, such as linguistics, common sense reasoning, and affective computing. Where Sentiwordnet has been used for the sake of disambiguation of meaning, which explores the semantic knowledge of WordNet [27], WordNet can therefore be seen as a combination of a dictionary and a thesaurus.

As Algorithm 3 highlights, the generation of an adequate sentiment lexicon consists in defining the sentimental aspects integrated in three stages:

**Algorithm 3** Semi-supervised sentiment lexicon generation algorithm.

---

**Data:** dictionary of words and their indices  
**Result:** Dense Embedd Sentiment value of combined vectors([Cluster])

**Input:**  
 $n_x$ : Size of vocabulary  
 $l$ : embedding dimension  
 $W \in \mathbb{R}^{l*n_x}$ : embedding Matrix  
 $n$ : refers to positive and negative classes.  
iter:max iterations of Kmeans.  
init:Method for kmeans initialization.  
polarity: +1 positive semantics, - 1 negative semantics.  
distances: a set of value between 0 and 1 resumes distances from each class centroids.  
 $coef_s$ : sentiment coefficient from semantics induction.  
Compound Matrix: Matrix  $C \in \mathbb{R}^{m*n_x}$   
 $m$ : max length of predefined lexicons  
model: KMEANS model ( $n, iter, init$ ) that fit defined embeddings.

```

for Each vector  $x_i \in W$  do
  if  $x_i$  in cluster0 then
    polarity = +1;
    distances =  $\frac{1}{\min_{words, cluster_0} distances(x_i)}$ ;
     $coef_s$ =polarity*distances;
  else
    polarity =-1;
    distances =  $\frac{1}{\min_{words, cluster_1} distances(x_i)}$ ;
     $coef_s$ =polarity*distances;
  end
end
for  $i=0$  to  $n_x$  do
   $O_i$ = OneHotEncoding(Data.indices[i]);
  compound =  $O_i$ *C;
end
return [ $coef_c, compound$ ]

```

---

- **Step (i):** Reassemble the embeddings in two organized groups using K-means; typically, we group the vectors into two separate contexts and semantics according to the definition of embeddings.
  - Group (1): Positive or positively connected semantics.
  - Group (2): Negative semantics.
- **Step (ii):** I inspect the sentiment coefficient by approaching the minimum proximity distance through positive or negative semantics. At each sentiment dimension, each sentiment word will take a value between -1 and +1 (-1 is extreme negativity and +1 is extreme positivity) by element-by-element multiplication as follows:  $sentiment_{coef} * Cluster_{polarity}$ . Where the cluster of positive semantics has a polarity value equal to +1, otherwise, negatives -1.
- **Step (iii):** The words of sentiment composed note the average of the lexicons summarized in Table 9 to the formalism’s initial sentiment rate matrix. We used the internal product of the rate matrix and the elementary vectors from step (ii) to define the sentiment score at the word level.
- **Step (iv):** Define a sentiment scale between the words for each semantic group and matrix composed by simple element-by-element multiplication.

## 4 Experiments and output result

### 4.1 Datasets

In our exploration system, raw tweets are collected through the Twitter APIs. For the collection of targeted Twitter data, we use a list of keywords, e.g., epilepsy, Parkinson’s disease, and HeartAttack. Table 4 summarizes the statistics of tweets based on the context of a medical concept. However, the data is very noisy, as social media posts relieve various natural language concepts such as the spelling of many related medical concepts, everyday expressions, excessive use of abbreviations, acronyms and alternatives, and incorrect medical spellings. The main most crucial

**Table 4** Statistics of Twitter datasets regarding main aspects used for the development and evaluation

Main aspects	Number of tweets
Epilepsy	20,192
Parkinson’s Disease	208,419
COVID-19 pandemic	489,302
Heart attack disease	390,291

challenges are as follows: (1) There was no guidance on how to distinguish between various related drugs meanings such as drug misuse or abuse and legitimate use in collected tweets, (2) patient reports may refer to an unsuspected drug reaction, which deserves great awareness to detect, and (3) the intentions of users when they report problems.

Textual data may require a credible annotated system to clarify medical knowledge and cover the truth about the patient’s status. The standardization of the medical concept required massive efforts to cover important aspects. It should be noted that there are many recent distributions offering annotated benchmark data sets for analyzing medical sentiment, but it had been built into the local window of the predefined set of drugs or diseases, for example, eDisease [29] Annotated a set of over 3,500 posts on online health forums about breast cancer, Crohn’s and different allergies, respectively. Each sentence in a message is manually labeled as “experience,” “fact,” or “opinion,” and as “positive,” “negative,” and “neutral.” In this study, we proceed through word processing steps which consist of tokenizing, cleaning, and normalizing the data in relation to the identification of the biomedical concept to enter into our neural network architecture, as illustrated in Fig. 4.

## 4.2 Implementation and output results

In this section, we go through the steps and parameters for implementing the architecture. The proposed joint model of multiple CNN channels joins both concrete biomedical definitions and a deep dive into the downstream task of

associative memory processes to define new associations from real-time medical patterns. Therefore, these features will be input into the sentiment generation algorithm to derive an aspect-based sentiment decomposition with additional statistical salients to accurately approach sentiment. We aim to prove how we embed our own from a large-scale unsupervised representation for learning affective associations related with these embeddings on top to analyzing two grams or three grams using convolutions. It works even faster than traditional word-bag and part-speech representation techniques.

Our use case scenarios focus on how a large amount of unstructured data from online health forums is characterized by the details of patient treatment. This study has been validated on many crucial bed cases: epilepsy, Parkinson’s disease, heart attack, etc. It consists of an automatic online method that generates efficient sentiment-related features, identifies, and stores natural patterns of drugs linked to various patient spaces. So far, our sentiment analysis model uses the encapsulated sentiment, context, and semantics information differently to parameterize new drug concepts by downstream deep associative memory methodology.

The results of our experiment show that the improvement of joint biomedical distributed vectorization helps identify formal associations of medical structures from real-time data. These abstractions range from lower level nodes representing the building blocks of discrete word associations to complex data such as concepts of multi-word medical associations. Pattern associations and representations would be learned through multiple convolutional connections between hidden layers, backpropagation, and a set of

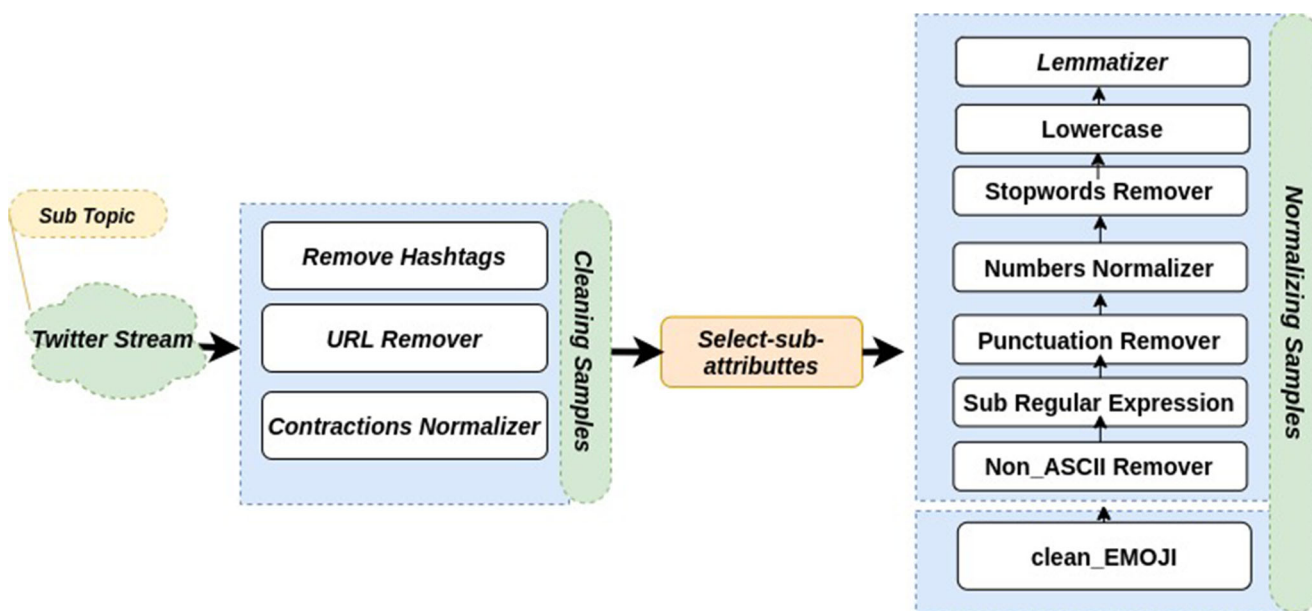


Fig. 4 Preprocessing steps

**Table 5** Embedding evaluation using two bio-medical pretrained embeddings

Dataset	Embedding strategy	Accuracy (%)
Small dataset	Proposed strategy	84%
	Hybrid convolutions of biomedical distributed from [15] and Bioconcept[17]	73%
	Joint convolutions of biomedical distributed from [15] and Bioconcept[17]	81%
Big dataset	<b>Proposed strategy</b>	<b>87%</b>
	Hybrid biomedical distributed from [15] and Bioconcept[17]	77%
	Joint biomedical distributed from [15] and Bioconcept[17]	79%

The bold entries highlight the achieved results of the proposed approach

(1) Hybrid convolutions of biomedical distributed from [15] and Bioconcept[17]

(2) Joint convolutions of biomedical distributed from [15] and Bioconcept[17]

concatenated feature layers running on learning rules and other processes used in classical conditioning.

In terms of embeddings, as described in Section 3, we seek to transfer the formal format of biomedical-controlled vocabulary from pre-trained distributed biomedical ontologies and formal medical resources, such as PubMed, the MIMIC III clinical database, and Central PubMed. It is mainly consulted in the processing of medical concepts distributed in natural language to better correspond to the formal semantics of a widely used controlled biomedical vocabulary. For each input sequence  $[x_1, x_2, \dots, x_m]$ , we propose to imitate embedding vectors for a predefined

configuration. Another slightly better approach is to use 1D convolutions to update and recalculate the embeddings we described in Section 3.1. Each CNN channel formed separately, in particular, the deep neural associative memory model; we further use backpropagation, inspired by [22], and form the corresponding convolutions, to define patterns of related concepts, which the neural network needs to store the matrix of related concepts in the memories matrix for each concept, respectively.

Our proposed model involves three different convolutional components. Each component consists of a flattened layer and a global average pooling 1D. The reason is the

**Table 6** The experiment results and comparison of our proposed sentiment generation approach with the baselines SVM, Simple RNN, Bi-LSTM, and Stacked LSTM

Dataset	Algorithms	Positiveness (%)	Negativeness (%)	Accuracy
Epilepsy	SVM	31 %	26%	0.66
	Simple RNN	45 %	16%	71.3
	BiLSTM	63 %	31%	<b>0.83</b>
	Stacked-LSTM	33 %	27%	0.80
Parkinson’s disease	SVM	31 %	26%	0.61
	Simple RNN	45 %	16%	71.3
	BiLSTM	63 %	31%	<b>0.84</b>
	Stacked-LSTM	33 %	27%	0.75
COVID-19 pandemic	SVM	31 %	46%	0.69
	Simple RNN	36 %	26%	70.9
	BiLSTM	33 %	<b>61%</b>	<b>0.81</b>
	Stacked-LSTM	41 %	37%	0.79
Heart attack disease	SVM	31 %	26%	0.68
	Simple RNN	31 %	16%	72
	BiLSTM	63 %	29%	<b>0.87</b>
	Stacked-LSTM	41 %	37%	<b>0.83</b>

The bold entries highlight the achieved results of the proposed approach

(1) Hybrid convolutions of biomedical distributed from [15] and Bioconcept[17]

(2) Joint convolutions of biomedical distributed from [15] and Bioconcept[17]



**Table 7** Statistics in terms of various SUDEP significant senses, and positivity and/or negativity appearance in the configuration space

Concept	Sense	Description	freq	#pos	#neg
SUDEP	Medical complication	Sudden unexpected death in epilepsy.	550	350	110
	Epilepsy SUDEP awareness	Sets healthcare services and benefits	208	120	66
	SUDEP Institute	Awareness and education services	115	88	5
	SUDEP Action	For raising awareness of epilepsy risks and deaths	298	104	58

that the percentage of positivity and negativity changes from algorithm to algorithm, meaning that context and semantics potentially control the orientation of sentiment in a given plain text where *pandemic related-posts* got huge negativity quantification most times due to the living crisis, where the pandemic period is widely known by tremendous patients’ narratives on social networks, characterized by vigorous activities.

Though testing model, we observe the necessity of multi-properties analysis; it has been proceeded by an explanatory analysis. Those observations hold when controlling for age, income, education, ethnicity, and overall health status through all of which are also associated with doing health-related activities online. Especially, varied linked concepts are fortunately dependent on other conditions and properties such as non-SUDEP, SUDEP, SUDI (sudden unexpected death in infancy (SUDI) is the name for the sudden and unexpected death of a baby), and DUDC (sudden unexplained death in childhood). Figure 5 shows an explanatory analysis that has been done regarding age property, where Fig. 6 illustrates how results change with regard to sexual property.

### 5 Discussion

Referring to Fig. 7, it shows related-medical concepts of SUDEP (sudden unexpected death in epilepsy) concept through the configuration space we proposed earlier. We notice that SUDEP appears three times in different positions or configurations. Each time, this concept has different embedding that clearly resumes such semantics and contextual information, but disambiguates reference and associations with real-world namely entities or medical concepts. Moreover, Table 7 summarizes results regarding

various SUDEP significant senses, which is further needed to be taken into consideration. In future work, we will forese a move from traditional word-based disambiguation approaches, towards semantically disambiguating aspect-based senses regarding various concept variants.

In general context, current accuracy is difficult to state without a host of medical concept abstraction caveats accuracy. Sentiment analysis model consists of many folds; accuracy at supervised sentiment classification for both binary and fine-grained levels is routinely above 90%, especially, with neural networks algorithms on particular benchmarking datasets achieving over 96% (Tables 8 and 9). On finer-grained sense distinctions, top accuracy ranges from 59.1 to 79.0% have been reported in recent evaluations from SemEval and Senseval, where the baseline accuracies of the simplest possible algorithm of always choosing the most frequent sense were 51.4% and 65%, respectively, as demonstrated in Table 5 and 6. Comparison has been done on very large patient narrative datasets, as explained in the dataset’s section, between a model with Naive Bayes on top of one and two grams, and those classical model gave 86.3 accuracies. The proposed neural network associative memory estimated current value based on linked pattern values; we are ambitiously aiming at defining deep associations regarding the combination of features sentiment in future studies. Also, in the future, we aim to conduct experiments towards the following directions, in order to showcase the flexibility of the neural word sense disambiguation, and further improve its performance.

Although very effective, these results have several limits and need to be improved; regarding the daily massive shared patients’ experiences, a dynamism property must be considered respectively to time property. It is substantially very important to consider the time-axis in detecting sentiment information relationships covered towards a set

**Table 8** Baselines models for medical sentiment analysis

Model	Vocabulary	Trainable params	Accuracy
BiLSTM	Universal Embeddings (Free-context)	22,112	[0.6–0.75]
LSTM	Universal Embeddings (Free-context)	12,932	[0.6–0.72]
RNN	Universal Embeddings (Free-context)	11,932	[0.5–0.68]
CNN	Embeddings with GlobalAveragePooling1D	68,573	[0.5–0.7]
SVM	Universal Embeddings		0.638

**Table 9** Summary of two a priori benchmark dictionaries used for sentiment lexicon generation algorithm

Lexicon	Description	Statistics	Automation level
SentiWordNet [12]	SentiWordNet is an opinion lexicon derived from the WordNet database where each term is associated with numerical scores indicating positive and negative sentiment information.	117,000 synsets are linked to other synsets according to their degrees of positivity, negativity, and neutrality.	Automated
SenticNet [28]	Common-Sense-Based knowledge for concept-level Sentiment Analysis, which is automatically constructed by applying graph-mining and multi-dimensional scaling techniques.	It associated with more than 50,000 natural language concepts through a multi-disciplinary approach.	Automated

of drugs or events, e.g., drug reaction, persons, and medical adverse events. Thus, it would be interesting to consider other attributes such as user credibility, harmness rate, and helpfulness rate for accurate patient personalization and high responsiveness.

## 6 Conclusion

Online conversations are becoming huge, especially patients or their families who are mostly wading through vast amounts of “social noise” looking for treatment second options, disease groups, patient/doctor insights, and more. This distillation of knowledge due to the massive amounts of unstructured information available is of considerable importance for bio-medical sentiment analysis tasks, while existing approaches are limited by the fact that they fail to correctly extract bio-medical entities, events, and contextual senses.

In this research, we investigated a distributed associative learning through multiple convolutions to build a dynamic configuration space from unsupervised medical concepts representations. Thus, these embeddings from this joint model are used for accurate sentiment decomposition. Extensive experiments demonstrate the promise of our work. In particular, the results showed a significant improvement in bio-medical natural concepts recognition and their attached sentiments.

The findings of this study can be useful for many researchers and health professionals to efficiently analyze patients’ narratives in real-world living scenarios. Also, it contributes to a better understanding of patients’ experiences and defining the public health status, outcome, or effectiveness of a treatment.

For future work, we plan to investigate the impact of incorporating additional information (age, sex, psychology status) regarding patient-centric targets. As a second future work, it would be interesting to consider other attributes (user credibility, harmness rate, helpfulness rate, and time) to further improve the performance of deep learning models for bio-medical sentiment analysis.

### Declarations

**Conflict of interest** The authors declare no competing interests.

**Informed consent** Informed consent was obtained from all individual participants.

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

## References

1. Grissette H, Nfaoui EH (2020) Enhancing convolution-based sentiment extractor via dubbed N-gram embedding-related drug vocabulary. *Netw Model Anal Health Inf Bioinform* 9(1):42
2. Rodrigues RG, das Dores RM, Camilo-Junior CG, Rosa TC (2014) SentiHealth-Cancer: a sentiment analysis tool to help detecting mood of patients in online social networks. *Int J Med Inform*
3. Giustini DM, Ali SM, Fraser M, Boulos MNK (2018) Effective uses of social media in public health and medicine: a systematic review of systematic reviews. *Online Journal of Public Health Informatics*
4. Portier K, Greer GE, Rokach L, Ofek N, Wang Y, Biyani P, Yu M, Banerjee S, Zhao K, Mitra P, Yen J (2013) Understanding topics and sentiment in an online cancer survivor community. *Journal of the National Cancer Institute - Monographs*
5. Grissette H, Nfaoui EH (2019) A conditional sentiment analysis model for the embedding patient self-report experiences on social media. *Advances in Intelligent Systems and Computing*
6. Enquist M, Lind J, Ghirlanda S (2016) The power of associative learning and the ontogeny of optimal behaviour. *Royal Society Open Science*
7. Nikfarjam A, Sarker A, O’Connor K, Ginn R, Gonzalez G (2015) Pharmacovigilance from social media: mining adverse drug reaction mentions using sequence labeling with word embedding cluster features. *J Am Med Inform Assoc*
8. Hu H, Phan NH, Geller J, Vo H, Manasi B, Huang X, Di Lorio S, Dinh T, Chun SA (2018) Deep self-taught learning for detecting drug abuse risk behavior in tweets. In: *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*
9. Araque O, Corcuera-Platas I, Sánchez-Rada JF, Iglesias CA (2017) Enhancing deep learning sentiment analysis with ensemble techniques in social applications. *Expert Syst Appl*
10. Grissette H, Nfaoui E (2019) Daily life patients sentiment analysis model based on well-encoded embedding vocabulary for related-medication text. In: *Proceedings of the 2019 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining, ASONAM ’19. Association for Computing Machinery, New York, pp 921–928*

11. Bengio Y, Courville A, Vincent P (2013) Representation learning: a review and new perspectives. *IEEE Transactions on Pattern Analysis and Machine Intelligence*
12. Baccianella S, Esuli A, Sebastiani F (2010) SENTIWORDNET 3.0: an enhanced lexical resource for sentiment analysis and opinion mining. In: *Proceedings of the 7th International Conference on Language Resources and Evaluation, LREC 2010*
13. Mohamad Beigi O, Moattar MH (2021) Automatic construction of domain-specific sentiment lexicon for unsupervised domain adaptation and sentiment classification. *Knowl-Based Syst* 213:106423
14. Wang Y, Huang G, Li J, Li H, Zhou Y, Jiang H (2021) Refined global word embeddings based on sentiment concept for sentiment analysis. *IEEE Access* 9:37075–37085
15. Zhang Y, Chen Q, Yang Z, Lin H, Lu Z (2019) BioWordVec, improving biomedical word embeddings with subword information and MeSH. *Sci Data* 6(1):52
16. Chen Q, Peng Y, Lu Z (2018) BioSentVec: creating sentence embeddings for biomedical texts. arXiv:1810.09302
17. Chen Q, Lee K, Yan S, Kim S, Wei CH, Lu Z (2020) Bioconceptvec: creating and evaluating literature-based biomedical concept embeddings on a large scale. *PLoS Comput Biol*
18. PubTator: a web-based text mining tool for assisting biocuration
19. Grissette H, Nfaoui EH (2020) Drug reaction discriminator within encoder-decoder neural network model: COVID-19 pandemic case study. In: *2020 Seventh International Conference on Social Networks Analysis, Management and Security (SNAMS)*. IEEE
20. Krotov D, Hopfield JJ (2019) Unsupervised learning by competing hidden units. *Proc Natl Acad Sci USA*
21. Demircigil M, Heusel J, Löwe M, Uppgang S, Vermet F (2017) On a model of associative memory with huge storage capacity. *J Stat Phys*
22. Liu J, Gong M, He H (2019) Deep associative neural network for associative memory based on unsupervised representation learning. *Neural Netw* 113:41–53
23. Palm G (2013) Neural associative memories and sparse coding. *Neural Netw*
24. Krotov D, Hopfield JJ (2016) Dense associative memory for pattern recognition. *Advances in Neural Information Processing Systems*
25. Hopfield JJ (1982) Neural networks and physical systems with emergent collective computational abilities. *Proc Natl Acad Sci USA*
26. Zhang L, Wang S, Liu B (2018) Deep learning for sentiment analysis: a survey. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*
27. Fellbaum C (2010) WordNet. In: *Theory and Applications of Ontology: Computer Applications*
28. Cambria E, Olsher D, Rajagopal D (2014) SenticNet 3: a common and common-sense knowledge base for cognition-driven sentiment analysis. In: *Proceedings of the National Conference on Artificial Intelligence*
29. Carrillo-de Albornoz J, Vidal JR, Plaza L (2018) Feature engineering for sentiment analysis in e-health forums. *PLoS ONE*

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.