Modeling of Online Learners' Sentiments About Multigranularity Knowledge

Anping Zhao and Yu Yu

Abstract—**To provide insight into online learners' interests in various knowledge from course discussion texts, modeling learners' sentiments and interests at different granularities are of great importance. In this article, the proposed framework combines a deep convolutional neural network and a hierarchical topic model to discover the hidden structure of online learners' sentiments about knowledge topics. The approach is to capture multigranularity knowledge of topics of interest to learners with the hierarchical topic model and to identify information about learners' different sentiments with the convolutional neural network. This approach not only models knowledge of hierarchical interest from general to specific but also identifies learners and their sentiment orientations to better correspond to the different granularities of knowledge. The experimental results and analysis of real-world datasets show that the proposed approach is effective and feasible.**

Index Terms—**Multiple granularity knowledge, sentiment analysis, topic model.**

I. Introduction

ONLINE learning has become an integral part of the educational landscape. It can be described as learning via the Internet, which offers students access to educational resources, content, and support, allowing meaningful interaction and facilitating the creation of knowledge. Online learning platforms produce a large amount of unstructured textual data through the application of interactive tools embedded platforms, such as course reviews and discussion forums. Online learners usually express themselves by means of these interactive tools when they engage in learning activities.

Sentiment analysis is a widely conducted natural language processing task, wherein the sentiment orientation of a text unit is judged. It involves the identification and extraction of subjective expressions from text sources and the computational analysis of people's opinions about target entities. Researchers use sentiment analysis to understand learners' engagement in a course and discover their personalized needs based on a large amount of learner data generated from online learning platforms with the aim of predicting future learning modes and responding to learning behavior [1], [2]. These data produced by learners provide ample opportunities to obtain an in-depth understanding of their implicit psychological features and learning requirements.

However, in real applications, it is not easy to identify learners' needs for different granularities of knowledge based on their interests because a knowledge representation with a multigranularity structure must be learned. Different learners need different knowledge and hence are focused on different granularities of knowledge. For example, some learners care about general knowledge, such as *Practice*, while others may be more concerned with more specific knowledge, such as *Coding*. Modeling knowledge granularity and incorporating sentiments make the identification of learners' needs even more challenging, as both the multigranularity structure and sentiment orientations depending on specific knowledge granularities must be learned. Furthermore, learners who share the same sentiment orientation and the same granularity of knowledge should be grouped hierarchically. A high-quality model of learner sentiments at different levels of knowledge granularity has many valuable applications in the areas of learning resource recommendation and personalized learning. To delve into learners' feedback to identify multigranularity knowledge and to determine learners' sentiment orientation regarding the textual content, the granularity of knowledge and a corresponding sentiment expression need to be extracted and judged in an efficient manner.

To address these challenges, motivated by the recent success of deep learning, an artificial intelligence method based on the deep learning model is an effective solution to the problem. Deep learning models, such as convolutional neural network (CNN), recurrent neural network (RNN), and bidirectional encoder representations from transformers (BERT), are very effective tools in the fields of text sentiment analysis. BERT has a stronger semantic feature extraction ability than CNN and RNN, and CNN has better parallel computing ability than RNN and BERT. For our multigranular learner sentiment modeling problem, a deep model is used for the sentence-level sentiment classification task. The sentences in the learners' review are usually short and concise with clear meaning. CNN is enough to achieve satisfactory classification performance and the calculation speed is faster than BERT and RNN. Therefore, a joint deep CNN [3] and hierarchical latent Dirichlet allocation (HLDA) [4] method is explored in this article to discover the multigranular structure of knowledge and sentiments from unstructured textual data produced by learners. Specifically, we explore a sentiment polarity classification method based on the CNN architecture

Manuscript received July 4, 2020; revised March 26, 2021 and February 24, 2022; accepted March 12, 2022. Date of publication March 16, 2022; date of current version June 20, 2022. This work was supported by the Basic Scientific Research Projects of Wenzhou under Grant G2020024. *(Corresponding author: Anping Zhao.)*

Anping Zhao is with the College of Computer and Artificial Intelligence, Wenzhou University, Zhejiang 325035, China (e-mail: apzhao@wzu.edu.cn).

Yu Yu is with the College of Education, Wenzhou University, Zhejiang 325035, China (e-mail: yuyu@wzu.edu.cn).

Digital Object Identifier 10.1109/TLT.2022.3159627

with distributed word embedding that can be applied to multigranularity knowledge-based sentiment analysis tasks without any feature engineering effort. The unsupervised topic modeling approach may then be especially beneficial for the performance of multigranularity knowledge detection with sentiment in the same model without the use of labeled training data.

The major contributions of this article are as follows.

1) Our method is the first attempt to model both knowledge of hierarchical interest from general to specific and learners with their sentiment orientations to better correspond to the different granularities of knowledge.
2) The proposed model is capable of processing key tasks in sentiment analysis for fine-grained knowledge simultaneously by combining hierarchical topic model with deep learning.
3) We use representations of multigranular sentiment to capture learners' interest in fine-grained knowledge to enable modeling of learners' requirements.

The rest of this article is organized as follows. In Section II, we review the related work. Section III introduces the basic idea and defines the problem addressed in our work. The joint model of hierarchical knowledge and sentiment identification is described in Section IV. We show the effectiveness of the presented approach through evaluations and the presentation of the experimental results in Section V. Finally, we summarize our conclusions and future research in Section VI.

II. Related Work

Sentiment modeling is widely used in various interactions between online learners to record knowledge that is of common interest to learners, to detect learners' feedback, and to provide targeted support to learners. This work focus on a method for modeling online learners' sentiments about multigranularity knowledge from an online course discussion forum.

Sentiment analysis is a process of analyzing, processing, concluding, and inferring subjective texts with sentiments [5]; it aims to discover structured opinions from unstructured text and identify their sentiment polarities [6]. Researchers have recently examined the interplay between deep learning methods and sentiment analysis. Deep learning has already been applied to sentiment analysis owing to its ability to identify high-level features. Deep learning models have been used to achieve better performance as they can automatically learn syntactic and inherent semantic information from the data. The convolutional multihead self-attention memory network was proposed for aspect-based sentiment classification tasks; it is an improved model based on memory networks [7]. CNNs have been applied to sentiment analysis, which explicitly models coherent relations within certain syntactic structures [8]. Araque *et al.* integrated deep learning with traditional surface approaches to improve the performance of sentiment analysis [9]. Sentiment analysis was utilized in a knowledge-based recommendation system to discover users with potential psychological disturbances [10]. An aspect-level neural network for the sentiment was presented that extracts a higher level phrase representation sequence from the embedding layer by using a CNN [11]. A deep neural network-based sentiment analysis methodology was proposed for opinion mining with big social data; this methodology is an adaptable sentimental analysis mechanism [12]. The authors combined the strength of linguistic resources with a gating mechanism to propose an effective CNN-based model for aspect-based sentiment analysis [13]. The latest trends in the ensemble application of symbolic and subsymbolic AI for sentiment analysis have been presented [14], integrating logical reasoning into deep learning architecture to build a new common-sense knowledge base for sentiment analysis. Recent efforts for a stacked ensemble method have been proposed to predict sentiment intensity by combining the outputs obtained from several deep learning and classical feature-based models using a multilayer perceptron network [15].

In recent years, efforts have been made to combine machine learning with sentiment analysis for online learning application fields. S. Liu *et al.* presented the behavior-sentiment topic mixture model to automatically reveal potential information from 50 online courses to detect the topics that learners discuss most, as well as how learners interact with these topics [16]. The study adopted sentiment analysis and hierarchical linear modeling to analyze the course features of MOOCs by a supervised machine learning algorithm [17]. The study investigated learners' explicit behaviors and implicit discourse content derived from reviews by using a mixed approach of text mining and statistical analysis [18]. Detecting and interpreting emotional information will be applied to sentiment analysis in the next generation of sentiment computing [19]. Emotional intelligence prediction models were presented based on the sentiment analysis of social networking data [20]. A conceptual framework of indistinguishability was proposed as the key component of the evaluation of computerized decision support systems [21].

Few models effectively discover topic hierarchies and perform sentiment analysis. The hierarchical user sentiment topic model (HUSTM) was designed to discover the hidden structure of topics and users while performing sentiment analysis in a unified way [22]. The multigranularity model of process knowledge was established in the form of a tuple, which helps to clarify the hierarchical structure and internal relations [23]. A joint aspect-sentiment model was presented to jointly extract multigrained aspects and opinions by modeling aspects, opinions, sentiment polarities, and granularities simultaneously [24]. Peng *et al.* proposed explicitly modeling the aspect target and conducting sentiment classification directly at the aspect target level via three degrees of granularity [25]. The hierarchical aspect-sentiment model (HASM) [26] and the structured sentiment model [27] are latent structures of aspects and sentiments that can naturally be organized into a hierarchy, where each of the nodes is made up of an aspect and the sentiment polarities associated with it.

The above-mentioned methods attempt to either explore unsupervised machine learning models to determine both aspects and their sentiment simultaneously within the same model or to employ deep learning models to present the syntactic and semantic information to capture both semantic and sentiment information encoded in the textual content for sentiment analysis. However, most of the existing methods lack sentiment information considering the specific granularity of knowledge and learners' interests in knowledge with multigranularity structures. Compared with previous research, our work focuses on the interplay between hierarchical topic models and the deep learning method for learners' multigranularity knowledge sentiment analysis. The novelty of the proposed method is that learners' discussion content is not simply classified based on the sentiment orientation but instead is used to generate knowledge-dependent sentiment granularity for each selected knowledge topic.

III. Problem Definitions

A. Basic Idea

In real sentiment analysis-based online learning intelligent applications, identifying online learner interests and sentiments at different levels of granularity is crucial. It can provide insights into learner interests with respect to a variety of topics of knowledge and help analyze learners' behaviors, allowing the identification of influential learners at any level of granularity based on their sentiment information. The hierarchical structure is more in line with the human cognitive process and provides a better solution for capturing learners' interests and requirements. The use of a hierarchical structure can be considered a global methodological approach to visualize and navigate learners' sentiments about their knowledge of topics of interest. The HASM and HUSTM hierarchical models were used to analyze different features of products and organize them into hierarchies while identifying the sentiment polarity for those features. However, there is a key difference between our model and the HUSTM. The HUSTM performs both topic sentiment term identification and sentiment polarity analysis, while only topic sentiment terms are identified and organized into a hierarchical structure by our joint model. In our problem, we believe that the HASM can be used to discover learners' focused knowledge, as well as different granularities of knowledge-oriented sentimental tendencies, which will construct a hierarchical structure space with the ability to naturally distinguish the learners' multigranularity knowledge-dependent sentiments.

To address this problem of the multigranularity structure of learners' knowledge of interest and sentiments, we extend the HUSTM by adding the deep CNN architecture that captures learners' sentiment information from discussion forum text. The proposed framework is shown in Fig. 1.We impose a deep CNN architecture on the hierarchical topic model to identify sentence orientation. Sentiment information can then be fed into the hierarchical semantic space to align the knowledge of the corresponding hierarchy. Thus, we feed learners' discussion sentences with the sentiment orientation information from the CNN into the basic HUSTM framework to capture knowledge sentiments at different granularities. This simple joint framework has the ability to identify both multigranularity knowledge-dependent sentiments and learners who share identical sentiments about the same knowledge based on the

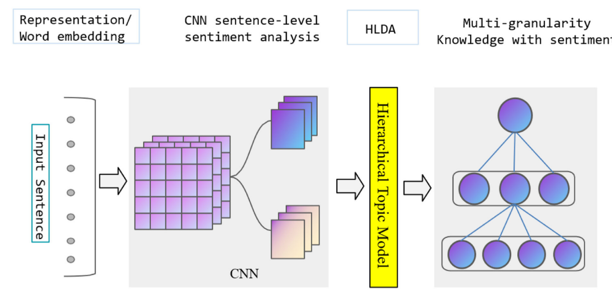

Fig. 1. Basic framework of the proposed model.

hierarchical structure. We use the hierarchical structure to represent the different granularity of knowledge and sentiments based on the assumption that a corpus of textual data from a learner discussion forum contains a latent structure of knowledge and sentiments that can naturally be organized into a hierarchy. To capture the idea that each different granularity of knowledge is associated with a set of sentiments, a path in the hierarchical structure selects an infinite collection of knowledge and sentiments of different learners. Both multigranularity knowledge features and sentiment are taken into account in the exploitation of latent hierarchical semantic topical space of textual data from the discussion forum.

As stated above, we focus on modeling learners' knowledge interest and sentiment at different granularities, enabling us to create a hierarchical structure to discover learners' sentiments and interests of different sentiment-polar knowledge topics. We explore the idea that the deep learning approach can be applied to multigranularity knowledge-based sentiment analysis tasks without any feature engineering effort, which might be beneficial for intelligent online learning applications to recommend knowledge of interest to learners.

B. Problem Definition

In our problem, the joint model consists of two main components. The first is to identify the sentiment orientation of sentences based on CNN with different word embedding representations. The second component is to discover the hierarchical knowledge sentiments from a large volume of text based on the HLDA framework. Sentences with sentiment polarity classification from the CNN are fed into the HLDA model to discover knowledge sentiments with different granularities from general to specific. We incorporate the tree structure of the HUSTM into our joint model to represent learners' sentiments and knowledge of interest at different granularities in the hierarchy. The tree structure arranges the topics of knowledge into a hierarchy with the desideratum that more general knowledge concepts should appear near the root and more specialized knowledge concepts should appear near the leaves. Meanwhile, each group of learners and sentiments can be associated with the knowledge of interest at different granularities. In the topic model, knowledge terms that convey homogeneous semantic themes are grouped together, as learners who share the same opinions and knowledge should be hierarchically grouped together. As a result, a corpus of course discussion texts contains a latent structure of knowledge and sentiments that can naturally be organized into a multigranularity structure of nodes. Each of the nodes comprises a knowledge feature and the sentiment polarities associated with it. Each node in the multigranularity structure of the hierarchical learners' knowledge sentiment model can be represented as two subnodes: the knowledge sentiment node and the learner sentiment node.

Definition 1: Multigranularity knowledge-sentiment elements can be represented by nodes in a hierarchical tree structure T, which is constructed by reorganizing multiple granularities of knowledge sentiment information extracted from a corpus $N_d = \{d_1, d_2, \ldots, d_N\}$ and a set of learners $L_d = \{l_1, l_2, \ldots, l_L\}$. All nodes in T are organized from general knowledge concepts to specific concepts. Each node consists of a knowledge-sentiment node Φ_k and a learner-sentiment node Θ_l to capture the idea that each learner and topic of knowledge is associated with a set of sentiments, which reflect the essential features of knowledge hierarchies and the main concern of learners, where $\Phi_{k,s}$ models the word distribution over knowledge k and sentiment polarity s, which is represented by a multinomial distribution of the words in the vocabulary of course discussion texts to describe popular knowledge that learners are focused on. Θ_l models a group of learners sharing a common sentiment about the same knowledge topic k and sentiment polarity s, which captures learners' attitudes and the respective sentiment information and is represented by a multinomial distribution over all learners.

Therefore, these are also necessary attributes for constructing a multigranularity representation of knowledge sentiment.

Definition 2: Given a corpus N_d and a set of learners L_d, the problem of the discovery of learners' multi-granularity knowledge sentiments is used to group the learners into a hierarchical structure based on their opinions and the coherent knowledge topics in those groups. The problem can be represented as a distribution over pairs of knowledge sentiment and learner sentiment for each sentence in the course discussion textual data, which transforms the problem into hierarchical topic discovery to investigate learners' sentiments about topics of knowledge at different granularities. The mixed learner, sentiment, and knowledge components are organized into a hierarchy at different granularities, and each group of data points can be associated with any node in the hierarchical structure T. Then, the constructed multigranularity structure of knowledge sentiment models learners' requirements at different granularities of knowledge.

IV. Multigranularity Knowledge Sentiment Modeling

A. CNN-Based Sentiment Identification

This section introduces a deep learning approach to effectively identify the sentiment orientation of online discussion content. To achieve this aim, we adopt word embedding to understand the sentiment contributions and their association with the discussion content, which combines a word and its context to represent words as real valued, dense, and low-dimensional vectors and greatly alleviates the data sparsity problem. Words are projected onto a lower dimensional vector space that potentially encodes syntactic or semantic features of words in their dimensions.

We use pretrained word embedding in the CNN framework for the sentiment classification task. Pretrained word embedding can capture meaningful syntactic and semantic regularities. We feed the word embedding representations of sentences into the deep learning framework CNN to predict sentiment polarity. In sentiment analysis, polarity refers to the scale of orientation prediction, i.e., either a binary (positive or negative) or multivariate scale. We use the readily available word embedding from external sources in the CNN framework as the only feature to avoid manual feature engineering efforts for sentence sentiment orientation classification. The resulting sentences with sentiments from the CNN are fed into the HLDA model for learners' multigranularity knowledge sentiment identification.

A CNN utilizes layers with convolution filters that are applied to local features to compute the distributed vector representations. The resulting high-level distributed representations are used as the only features to classify the sentiment of each sentence. CNN used in this work is a simple CNN architecture proposed by Kim [3]. An input sentence can be represented as

$$s_{i:i+j} = x_i \oplus x_{i+1} \oplus \cdots \oplus x_{i+j} \tag{1}$$

where x_i is the d-dimensional word embedding and $x_i \oplus x_{i+1} \oplus \cdots \oplus x_{i+j}$ is the compositional context matrix by concatenating each word embedding in the sentence $s_{i:i+j}$. A convolutional layer applies the filter w to each possible window of h words in the matrix of sentences. The resulting new feature map $c = [c_i] \in \mathbb{R}^{n-h+1} (i = 1, 2, \ldots, n - h + 1)$, where $c_i = f(w \cdot s_{i:i+h-1} + b)$ is a feature generated from a window of sequence words $s_{i:i+h-1}$, $b \in \mathbb{R}$ is a bias term, and f is a nonlinear function. This process aims to capture semantic context dependencies between words. The learned weights in the filter w correspond to a semantic feature that learns to recognize a specific sentiment class of n-grams.

A pooling operation captures the most important feature for each feature map. The maximum value $c = \max\{c\}$ is taken as the feature corresponding to the particular filter w. Multiple features can be obtained by using multiple filters in the convolutional architecture. For m filters w, the final feature vector $z = [\hat{c}_1, \ldots, \hat{c}_m]$ with the concatenated word embedding is passed to the output layer for classification. The output of the softmax layer is the probability distribution over labels of sentiment orientation. The likelihood of softmax classification within the output is calculated as

$$p(y_t = k|s, \theta) = \frac{\exp(W^{(S)} z + b)}{\sum_{k=1}^{K} \exp(W^{(S)} z + b))}. \tag{2}$$

Based on this CNN architecture, we can capture the sentiment polarity information of the sentences from the text of the discussion forum.

B. Discovery of Learners' Multigranularity Knowledge Sentiment

Different from existing hierarchical topic models for aspect-based sentiment analysis, which discover topic and sentiment polarity in a unified topic modeling framework, we feed the sentences with sentiment orientation information from the CNN into the extended HLDA model to capture different granularities of learners' knowledge sentiments to extract a more meaningful hierarchical structure. We can naturally distinguish knowledge-dependent sentiments based on the hierarchical structure. This simple joint framework can identify both the different sentiment polarities of the knowledge topics and the learners' attitudes at different granularities.

Based on the definition of the HLDA model, topics and the relationships between them are identified by using probabilistic inference. For our problem of the multigranularity structure of learners' knowledge topics and sentiments, we use the extended HLDA model based on the recursive Chinese restaurant process (rCRP) [28]. The rCRP is an extension of the CRP, in which a mixture of components are organized in an infinite tree, and each group of data points can be associated with any node in the tree. In the topic model, words that convey homogeneous semantic themes are grouped together, as customers with similar tastes sit at the same table. The sentiment polarity information comes directly from the sentiment of the sentences identified by the CNN before the performance of the hierarchical topic-sentiment term identification.

More formally, the infinite hierarchical tree defined by the rCRP can be considered to automatically extract both the structure and parameters of the hierarchical tree. A document corresponds to a discussion text. The output of the model could be a probability distribution over pairs of learners' knowledge sentiments, which integrate the learners' sentiment information to enhance its hierarchical structure. We use the generative method as follows.

1) Draw the global infinite learners' topic tree of knowledge and sentiment $T \sim \mathrm{rCRP}(\gamma)$.
2) For each term node Φ_k of knowledge topic k and sentiment terms node in the topic tree T, draw a word distribution $\boldsymbol{\phi}_k \sim \mathrm{Dirichlet}(\beta)$.
3) For each sentence i in discussion document d:
    a) draw a knowledge-sentiment terms node $c \sim T$;
    b) draw a subjectivity probability distribution $\theta \sim \mathrm{Beta}(\alpha)$;
    c) for each word j :
        i) draw a learner $l_{di} \sim \mathrm{Uniform}(l_d)$;
        ii) probabilistically draw a word distribution of subjectivity $p \sim \mathrm{Binomial}(1, \theta)$;
        iii) draw the word $w \sim \mathrm{Multinomial}(\boldsymbol{\phi}_{k,p})$.

Each word with a sentiment label is associated with a latent learner and topic variables by a multinomial distribution over knowledge topics $\Phi_{k,s,w}$. Each learner is associated with a multinomial distribution over knowledge topics and sentiments $\Theta_{k,s,l}$ by sampling at random for each word token in the discussion text. A knowledge topic is chosen from the hierarchical structure for each word associated with the learner of that word. The subjectivity of every word indicates whether the word has a sentiment tendency, and nonsubjective words represent knowledge topics with different granularities. This generative method defines a likelihood distribution across possible corpora. Each knowledge sentiment topic could be a multinomial probability distribution over the total vocabulary. The multigranularity structure provides information about the knowledge topics that a group of learners cares about, which not only captures learners' interests but also shows the words that learners use to describe their opinions.

We use sentence-level sentiment classification by CNN to determine the sentiment polarity of topics. The total probability of the hierarchical structure from HLDA is

$$\mathcal{L} = P(T|\gamma) \prod_{k=0}^{K} \prod_{i=1}^{N_d} \prod_{l=1}^{L_d} p(w_{di}|p, \Phi_k, \beta) P(\Phi_{kdi}|l_{di}, T) P(l_{di}|l_d). \tag{3}$$

The probability of producing sentence i in discussion document d from word subjectivity p and knowledge sentiment node Φ_k is

$$p(w_{di}|p, \Phi_k, \beta) \propto \prod_{s=0}^{1} \left(\frac{\Gamma(n_{k,s,-i}^{w,(.)} + \hat{\beta}_{s_i})}{\prod_w \Gamma(n_{k,s,-i}^{w,(w)} + \beta_{s_i,w})} \times \frac{\prod_w \Gamma(n_{k,s}^{w,(w)} + \beta_{s_i,w})}{\Gamma(n_{k,s}^{w,(.)} + \hat{\beta}_{s_i})} \right) \tag{4}$$

where $n_{k,s,-i}^{w,(v)}$ is the size of words having v_{th} vocabulary allocated to topic k and subjectivity s. n_{-i} represents counter variable does not include index i. We use the Gibbs sampling process for subjectivity that the update is identical to that in HASM.

V. Experimental Results and Evaluations

A. Dataset

The evaluation and experimental results shown in this section demonstrate the utility of combining CNN with a hierarchical topic model for the analysis of online learners' sentiments about multigranularity knowledge. We analyze the experimental results and evaluate the performance of our method against that of other similar models from prior related work. The two datasets investigated in this work were collected from MOOC offerings at the Chinese University MOOC platform. We organized posts into two groups by course type: Java programming and Python programming. Hereafter, we refer to these two datasets as Java and Python. We focus on tracking the knowledge hierarchy of discussions related to the course content. This resulted in 23 956 posts by 4725 different learners for Java programming and 30 784 posts by 5234 learners for Python programming. We performed a series of preprocessing steps, including separating sentences and removing stop words. The final datasets comprise 37 824 sentences with 4982 unique terms for Java and 41 178 sentences with 5173 unique terms for Python.

B. Experimental Settings

1) Convolutional Neural Network: The simple CNN architecture has been shown to be effective and to perform well in multiple-sentence classifications for sentiment analysis.

We use the same hyperparameters as the simple CNN: we set filter windows (h) of 3–5 with 100 filters for each filter region and a stochastic dropout rate p of 0.5 on the penultimate layer. Optimization is performed on minibatches of size 50.

2) Pretrained Word Embedding: The Chinese Word Vectors [29] project provides more than 100 Chinese word embeddings trained with different representations (dense and sparse), context features (words, n-grams, characters, and more), and corpora. One can easily obtain pretrained vectors with different properties and use them for downstream tasks. We use the pretrained 300-dimensional dense word embedding trained with skip-gram with negative sampling (SGNS) and sparse word embedding trained with positive pointwise mutual information (PPMI), respectively.

3) Hierarchical Topic Model: The hierarchical topic model based on the rCRP is established to find sentiment topics of different granularity with positive and negative sentiments for review sentences with a subjectivity value of 1. The hierarchy of aspect topics can be found for words in sentences with a subjectivity of 0. We set the Dirichlet hyperparameter α as 25.0 and set the rCRP prior γ as 0.01.

C. Evaluation Metrics

We aim to assess whether deep learning-based sentiment analysis can provide support for learners' hierarchical knowledge topic discovery by offering a comparison with the other hierarchical topic sentiment models. There are two parts to our evaluation. First, the sentiment analysis result is evaluated by using standard measurements. Second, we evaluate learners' knowledge topic hierarchy according to two import criteria for discovering the optimal topic hierarchy from the text.

1) AUC Accuracy: We define an accuracy expression to evaluate the quality of the sentiment classification results, which is a useful quantity to record during training and testing. A common analysis index for binary classification issues is the area under the curve (AUC). We use the AUC to indicate the changes in accuracy from the CNN with totally different sizes of filter windows and to check the result with different sentiment classifications.

We utilize two metrics to determine the quality of the hierarchical learners' knowledge topic sentiment generated by the HLDA model.

2) Relatedness: This metric assesses how well the extracted hierarchical knowledge topics represent real knowledge granularity, which shows the related relationship between different levels of knowledge topics. We use the same method to compute the relatedness score as shown in the HUSTM.

3) Knowledge-Dependent Sentiment Consistency: This metric evaluates how well the sentiment polarities expressed for a specific knowledge topic are extracted from the multigranularity knowledge sentiment structure. We evaluate knowledge-dependent sentiment using precision, recall, and the F-measure. A knowledge sentiment topic is defined as a true positive or true negative if it is correctly extracted by the model, and it is compared with the result of manually identified from a course discussion text. A knowledge-sentiment topic is defined as a false positive or false negative if it is incorrectly identified by the model, and it is compared with the manually extracted result.

D. Experimental Results

The different sentiment analysis methods for text permit us to compare the results of comparable approaches to investigate the effectiveness of each approach. Recent in-depth studies have indicated that CNN-based sentiment classification techniques are better than previously proposed models. However, we principally focus on comparing the HASMs, which are applicable to online learning for learners' multigranularity knowledge topic discovery from course discussion text. The HASM and the HUSTM naturally organize topics and sentiments into a hierarchy and identify the sentiment polarities associated with them, which incorporate both hierarchical topic modeling and sentiment analysis of the same model. To show the effectiveness of sentiment classification by CNN compared to that of related hierarchical methods, we compare the HUSTM and the HASM. Our evaluation compares the results of the two models for sentiment classification of the datasets and uses the HASM as a baseline.

In our framework, hierarchical knowledge topic sentiment classification is performed by CNN. This sentiment classification information can be used to discover knowledge with positive and negative learners' sentiments. By comparing the learner knowledge sentiment classification extracted by our approach with that extracted using the HASM and HUSTM, we aim to show the effectiveness of incorporating other linguistic features into the context of the deep CNN method to better leverage semantic data to benefit the discovery of knowledge sentiment polarity. For this purpose, we evaluate the sentiment classification accuracy of the model by taking the different effects of the CNN with different pretrained word embeddings into consideration. We assess the different effects of CNN with a multiple region size of (3,4,5) by using the different word embeddings. The results are shown in Fig. 2.

It can be seen from the results that the accuracy obtained using CNN on the datasets is better than that of both the HASM and AUSTM. This may be because the classifier employed in this technique has distinct characteristics. The HASM and HUSTM neither consider the sentence structure nor construct any internal sentence representations, while the CNN constructs internal representations with word embedding and operates on sentences with word sequences before predicting the sentiment class distribution. Thus, CNN architecture can consider intrasentence relations and provide valuable clues for the sentiment prediction task, while the HASM and HUSTM take each feature or sentiment polarity as a distribution of words.

The results are also assessed for the effect of the distributed representations of words as the input for the CNN on the dataset. We replace SGNS word embedding with PPMI word embedding representations. However, this change does not have much effect on the performance of the dataset. The accuracy is affected only very slightly, and the difference may be because pretrained word embedding may not always be available for specific words in the datasets. With the use of word

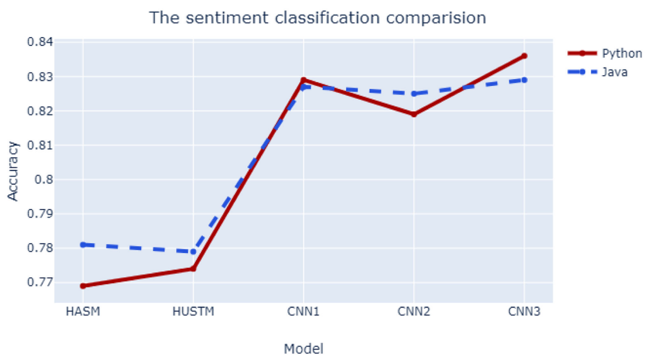

Fig. 2. Sentiment classification comparison for the different models. CNN1-CNN with SGNS embedding. CNN2-CNN with PPMI embedding. CNN3-CNN with concatenating SGNS and PPMI embedding.

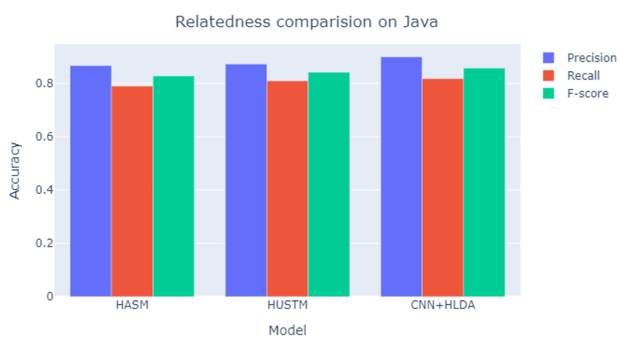

Fig. 3. Relatedness comparison among HASM, HUSTM, and CNN+HLDA on Java.

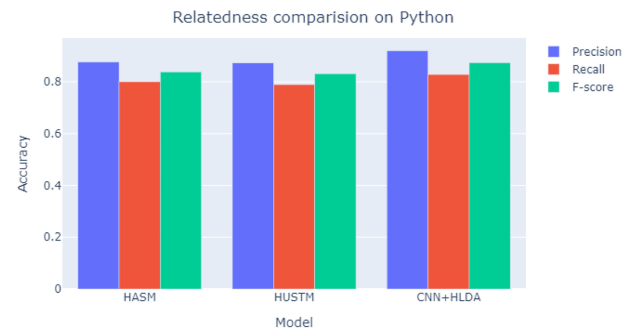

Fig. 4. Relatedness comparison among HASM, HUSTM, and CNN+HLDA on Python.

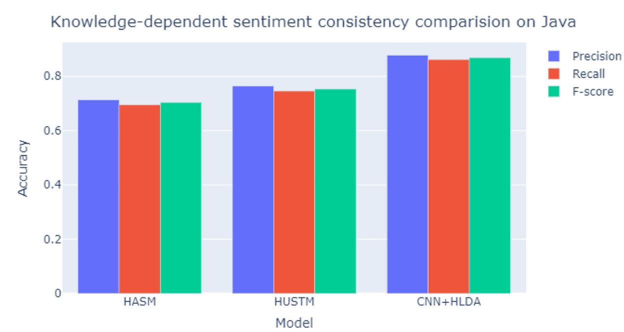

Fig. 5. Knowledge-dependent sentiment consistency comparison among HASM, HUSTM, and CNN+HLDA on Java.

embedding, some composition-based methods are presented to capture the semantic representation of sentences and texts. Furthermore, combining different pretrained word embeddings with deep neural networks has offered new inspiration for various sentiment analysis tasks. We also experiment with the simple concatenating of different representations as the input to CNN to incorporate other linguistic features into the context of word input representation to benefit sentence sentiment polarity classification. SGNS embedding uses a shallow neural network to learn the low-dimensional dense embedding, while the PPMI model is a method for representing a sparse bag of features, which uses positive point-by-point mutual information to weight features. The concatenating representation is helpful for performance on the Python dataset, but it also does not have much effect on the Java dataset. As a result, it is more likely to depend on the dataset.

We use hierarchical topic modeling to discover learners' knowledge sentiment granularity. To measure the relatedness of knowledge topics between the different granularities, we manually check ten of the most popular extraction knowledge items for each of the 50 knowledge topics generated by HLDA. We use content analysis techniques to create a ground truth set of knowledge topics referred to in the course discussion text and their associated sentiments. We compare the results of our approach with those of the manual analysis of hierarchical knowledge topic relatedness and knowledge-dependent sentiment. The goal of this evaluation is to analyze the semantic relatedness between the knowledge topics of different hierarchies. It is based on the assumption that parent topics should have more similarities with their direct child topics than with the child topics of other parent topics.

Figs. 3 and 4 show the results of the relatedness comparison of the three models: 1) ASUM; 2) HUSTM; and 3) CNN+HLDA.

The results show that the three models have good performance in terms of relatedness to multigranularity topics of knowledge. The extracted knowledge topics are usually words describing actual knowledge concepts. This is because both our model and the HUSTM identify the aspect-sentiment hierarchy granularity using similar models.A further metric is used to evaluate the quality of the knowledge sentiment with different granularities generated by the hierarchical topic model. Knowledge-dependent sentiment consistency assesses how coherent and consistent the sentiments and the knowledge topic within a knowledge sentiment node are. To qualitatively measure the knowledge-dependent sentiment consistency of the multigranularity knowledge sentiments generated by our approach, we evaluate the consistency of each knowledge sentiment node by analyzing whether the identified knowledge sentiment hierarchy can show different sentiment orientations depending on the specific knowledge at different granularities. The result is shown in Figs. 5 and 6.

We observe that the results of the three models are comparable. All three models show that the identified sentiment hierarchy can discover specific knowledge concepts and associated opinions. The joint model outperforms both HUSTM and HASM. This may be mainly because of the sentiment classification

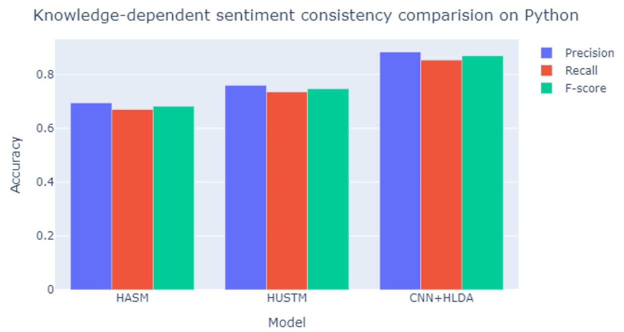

Fig. 6. Knowledge-dependent sentiment consistency comparison among HASM, HUSTM, and CNN+HLDA on Python.

Level0:
Node0：
Node01：
+课程(Curriculum)计算机(Computer)规划(Planning)编程(Programming)体验(Experience)
- 解说(Explaining)课后(After class)效果(Effect)实操(Practice)偏难(Difficulty)
Node02：
+努力_ing HappyMore sdufe_LIXY shjcreat licb122
- 暗星2018 applexujunjie163com 斯滕瓦尔坎 zzy4521yeahnet
Level1:
Node1:
Node10：
+节奏(Rhythm)推荐(Recommendation)干货(Essence)合理(Reasonable)基础(Elementary)
- 测验(Test)参考(Reference)可执行(Executable)类型(Type)系统(System)代码(Code)
Node11：
+KIA_KAHA Lydia0423 jzkcup
- heater 雪花飘荡时 mooc1516296511492 浅笑xr

Fig. 7. Example of multigranularity knowledge-sentiment structure.

results. This result confirms the advantages of the joint model over the other two models for certain features at different granularities. The HLDA model is fed sentences with the sentiment orientation information obtained by CNN, which means that its hierarchical knowledge topics and sentiment analysis are achieved at the sentence level and that all of the words in a sentence are assigned the same sentiment orientation and topic.To provide the experimental results as an example of a knowledge topic with different granularity with sentiment polar identified from the datasets, Fig. 7 shows the learner multigranularity knowledge sentiments discovered by the joint deep CNN and HLDA framework.

The hierarchical nodes show the knowledge topics at various granularities with different knowledge concept-dependent sentiments. Nodes 0 and 1 are organized in a hierarchical structure, in which node 01 and node 02 and node 10 and node 11 are subnodes of node 0 and node 1, respectively. Nodes 01 and 10 correspond to the knowledge sentiment nodes, while nodes 02 and 11 correspond to the learner sentiment nodes. The most general knowledge about some topics with general positive and negative sentiments is at level 0. At level 1, the knowledge concept becomes a specific concept of the topic of knowledge, such as rhythm, type, and code, with positive and negative sentiments, which are closely related to the knowledge topics of level 0. This result also shows that learners are interested in knowledge topics at a specific level. The result confirms that knowledge is organized into a multigranularity structure from general to specific. The knowledge-dependent sentiment polarities are captured mainly because of sentence-level sentiment classification based on the CNN before using the hierarchical topic model. The experimental results show that the knowledge and sentiment automatically identified from the datasets describe the overall hierarchical structure at different granularities.

These experimental results show the effectiveness of the proposed approach for modeling online learners' sentiments about multigranularity knowledge. The qualitative results suggest that learners' multigranularity knowledge sentiment automatically identified from the course discussion text reflected the learners' overall learning interests. We conclude that the mixture of the topic model and deep learning strategies is an effective approach to identifying online learners' interests and sentiments from course discussion text.

VI. Conclusion

Understanding and explicitly modeling the different granularities of knowledge and sentiment are an effective way to discard and meet different learners' interests and requirements. In this article, we proposed the employment of a joint deep CNN and hierarchical topic model to identify sentiment orientation and to discover hierarchical knowledge sentiments for the analysis of learners' potential interests. The main motivation is to facilitate learners' understanding in a manageable way to support intelligent real online learning applications. We presented a method for combining the deep learning CNN with HLDA as a framework for identifying fine-grained knowledge context-dependent sentiment. The methodology revolved around hierarchical topic modeling and sentiment analysis technology to achieve a summary of learners' primary concerns. The work showed the potential of combining deep learning with a topic model to capture learners' multigranularity knowledge interest from a course discussion forum.

Because of the numerous implicit expressions in learners' discussions in a course forum, the terms used to express opinions will differ from person to person in complex semantic contexts. Hierarchical learner knowledge sentiment analysis should be tailored to the implicit linguistic context to extract implicit information and sentiment. This issue is at least worth studying and will be part of our future work.

References

[1] A. S. Sunar *et al.*, "Modelling MOOC learners' social behaviours," *Comput. Hum. Behav.*, vol. 107, 2020, Art. no. 105835.
[2] H. Fasihuddin, G. Skinner, and R. Athauda, "Towards adaptive open learning environments: Evaluating the precision of identifying learning styles by tracking learners' behaviours," *Educ. Inf. Technol.*, vol. 22, no. 3, pp. 807–825, 2017.
[3] Y. Kim, "Convolutional neural networks for sentence classification," in *Proc. Conf. Empirical Methods Natural Lang. Process.*, 2014, pp. 1746–1751.
[4] D. Blei, T. Griffiths, and M. Jordan, "The nested Chinese restaurant process and Bayesian nonparametric inference of topic hierarchies," *J. ACM*, vol. 57, no. 2, pp. 1–30, 2010.
[5] H. Liu, I. Chatterjee, M. Zhou, X. S. Lu, and A. Abusorrah, "Aspect-based sentiment analysis: A survey of deep learning methods," *IEEE Trans. Comput. Social Syst.*, vol. 7, no. 6, pp. 1358–1375, Dec. 2020.
[6] E. Cambria, "Affective computing and sentiment analysis," *IEEE Intell. Syst.*, vol. 31, no. 2, pp. 102–107, Mar./Apr. 2016.
[7] Y. Zhang, B. Xu, and T. Zhao, "Convolutional multi-head self-attention on memory for aspect sentiment classification," *IEEE/CAA J. Automatica Sinica*, vol. 7, no. 4, pp. 1038–1044, Jul. 2020.
[8] L. Zhang, S. Wang, and B. Liu, "Deep learning for sentiment analysis: A survey," *Data Mining Knowl. Discov.*, vol. 8, no. 4, 2018, Art. no. e1253.

[9] O. Araque *et al.*, "Enhancing deep learning sentiment analysis with ensemble techniques in social applications," *Expert Syst. Appl.*, vol. 77, pp. 236–246, 2017.
[10] R. L. Rosa, G. M. Schwartz, W. V. Ruggiero, and D. Z. Rodríguez, "A knowledge-based recommendation system that includes sentiment analysis and deep learning," *IEEE Trans. Ind. Inform.*, vol. 15, no. 4, pp. 2124–2135, Apr. 2019.
[11] W. Meng, Y. Wei, P. Liu, Z. Zhu, and H. Yin, "Aspect based sentiment analysis with feature enhanced attention CNN-BiLSTM," *IEEE Access*, vol. 7, pp. 167240–167249, 2019.
[12] P. Kaladevi and K. Thyagarajah, "Integrated CNN- and LSTM-DNN-based sentiment analysis over big social data for opinion mining," *Behav. Inf. Technol.*, vol. 40, pp. 11–19, 2019.
[13] D. Zeng *et al.*, "Aspect based sentiment analysis by a linguistically regularized CNN with gated mechanism," *J. Intell. Fuzzy Syst.*, vol. 36, no. 5, pp. 3971–3980, 2019.
[14] E. Cambria *et al.*, "SenticNet 6: Ensemble application of symbolic and subsymbolic AI for sentiment analysis," in *Proc. 29th ACM Int. Conf. Inf. Knowl. Manage.*, 2020, pp. 105–114, doi: 10.1145/3340531.3412003.
[15] M. S. Akhtar, A. Ekbal, and E. Cambria, "How intense are you? predicting intensities of emotions and sentiments using stacked ensemble [application notes]," *IEEE Comput. Intell. Mag.*, vol. 15, no. 1, pp. 64–75, Feb. 2020, doi: 10.1109/MCI.2019.2954667.
[16] S. Liu *et al.*, "Unfolding sentimental and behavioral tendencies of learners' concerned topics from course reviews in a MOOC," *J. Educ. Comput. Res.*, vol. 57, no. 3, pp. 670–696, 2019.
[17] K. F. Hew *et al.*, "What predicts student satisfaction with MOOCs: A gradient boosting trees supervised machine learning and sentiment analysis approach," *Comput. Educ.*, vol. 145, 2020, Art. no. 103724.
[18] X. Peng and Q. Xu, "Investigating learners' behaviors and discourse content in MOOC course reviews," *Comput. Educ.*, vol. 143, 2020, Art. no. 103673.
[19] E. Cambria, S. Poria, A. Hussain, and B. Liu, "Computational intelligence for affective computing and sentiment analysis [guest editorial]," *IEEE Comput. Intell. Mag.*, vol. 14, no. 2, pp. 16–17, May 2019.
[20] X. Wei, G. Xu, H. Wang, Y. He, Z. Han, and W. Wang, "Sensing users' emotional intelligence in social networks," *IEEE Trans. Comput. Social Syst.*, vol. 7, no. 1, pp. 103–112, Feb. 2020.
[21] J. M. Garibaldi, "The need for fuzzy AI," *IEEE/CAA J. Automatica Sinica*, vol. 6, no. 3, pp. 610–622, May 2019.
[22] A. Almars, X. Li, and X. Zhao, "Modelling user attitudes using hierarchical sentiment-topic model," *Data Knowl. Eng.*, vol. 119, pp. 139–149, 2019.
[23] Y. Fan, C. Liu, and J. Wang, "Integrating multi-granularity model and similarity measurement for transforming process data into different granularity knowledge," *Adv. Eng. Inform.*, vol. 37, pp. 88–102, 2018.
[24] F. Tang *et al.*, "Aspect based fine-grained sentiment analysis for online reviews," *Inf. Sci.*, vol. 488, pp. 190–204, 2019.
[25] H. Peng *et al.*, "Learning multi-grained aspect target sequence for Chinese sentiment analysis," *Knowl.-Based Syst.*, vol. 148, pp. 167–176, 2018.
[26] S. Kim *et al.*, "A hierarchical aspect-sentiment model for online reviews," in *Proc. 27th AAAI Conf. Artif. Intell.*, 2013, pp. 526–533.
[27] A. Almars *et al.*, "Structured sentiment analysis," in *Advanced Data Mining and Applications Anonymous*. Cham, Switzerland: Springer, 2017, pp. 695–707.
[28] J. Kim *et al.*, "Modeling topic hierarchies with the recursive Chinese restaurant process," in *Proc. 21st ACM Int. Conf. Inf. Knowl. Manage.*, 2012, pp. 783–792, doi: 10.1145/2396761.2396861.
[29] S. Li *et al.*, "Analogical reasoning on Chinese morphological and semantic relations," in *Proc. 56th Annu. Meeting Assoc. Comput. Linguistics, Vol. 2, Short Papers*, 2018, pp. 138–143.

Anping Zhao received the doctoral degree in artificial intelligence from Southwest University, Chongqing, China, in 2011.

He was a Postdoctoral Researcher with Artificial Intelligence Research Group, University of York, York, U.K., from August 2015 to 2017. He is currently an Associate Professor with the College of Computer and Artificial Intelligence, Wenzhou University, Wenzhou, China. His main research interests include cognitive computation and text mining.

Yu Yu received the doctoral degree in logics from Southwest University, Chongqing, China, in 2010.

She is currently an Associate Professor with the College of Education, Wenzhou University, Wenzhou, China. She was an Academic Visitor to the Department of Computer Science, University of York, U.K., in 2016. Her main research interests include language logic and semantic analysis.