

Integration of Fuzzy and Deep Learning in Three-Way Decisions

L.D.C.S. Subhashini^{1,2,3,4}, Yuefeng Li*^{1,5}, Jinglan Zhang^{1,6}, Ajantha S. Atukorale^{2,7},

¹ Queensland University of Technology, Australia

² University of Colombo, Sri Lanka

³ University of Sri Jayewardenepura, Sri Lanka

⁴ subhashini@sjp.ac.lk, shashikala.charles@hdr.qut.edu.au

⁵ y2.li@qut.edu.au

⁶ jinglan.zhang@qut.edu.au

⁷ aja@ucsc.cmb.ac.lk

Abstract—The problem of uncertainty is a challenging issue to solve in opinion mining models. Existing models that use machine learning algorithms are unable to identify uncertainty within online customer reviews because of broad uncertain boundaries. Many researchers have developed fuzzy models to solve this problem. However, the problem of large uncertain boundaries remains with fuzzy models. The common challenging issue is that there is a big uncertain boundary between positive and negative classes as user reviews (or opinions) include many uncertainties. Dealing with these uncertainties is problematic due in many frequently used words may be non-relevant. This paper proposes a three-way based framework which integrates fuzzy concepts and deep learning together to solve the problem of uncertainty. Many experiments were conducted using movie review and ebook review datasets. The experimental results show that the proposed three-way framework is useful for dealing with uncertainties in opinions and we were able to show that significant F-measure for two benchmark dataset.

Index Terms—Opinion Mining, Fuzzy Logic, Three-way Decision, Classification, Deep Learning

I. INTRODUCTION

The decision-making process for analysing customer reviews using machine learning algorithms has become a challenging issue [1]. Today, people write large numbers of opinions and make them available via the Internet. Therefore, online customers find it challenging to make decisions based on such reviews. However, the growth of online customers has led to a reliance on reading previous customer opinions before new customers make decisions. At the same time, organizations generate useful information from customer reviews by using text mining techniques. This concept is known as sentiment analysis or opinion mining [2]. Many models have been proposed in recent years for opinion mining [3], [4]. However, because of uncertainties and vagueness of opinions within the wide range of reviews usually posted, these models can create unexpected results in some cases [5]. Therefore, the ability to handle uncertainty is important in opinion classification models.

*Corresponding Author

The concept of uncertainty refers to unclear and inconsistent opinions. Therefore, the ability to handle uncertainty is important in opinion classification models [5]. When it comes to opinion classification models, the mechanism for managing these uncertainties is a challenging task [6] as a lot of user generated opinions contain uncertainties. The main challenge is to automatically deal with these uncertainties to enhance the performance of machine learning algorithms. For example, a customer expresses review for a camera as "The camera's resolution is ok but the sound quality looks not fine". This review is uncertain as the frequent features can appear in both the negative or positive categories. However, machine learning algorithms do not know how to wisely use them; therefore, it is very difficult to decide the relevant category. In this example, it is hard to decide the relevant category with the uncertainty of using two frequent features, "ok" and "fine", particularly the review is "not fine" rather than "fine". Also, uncertainties are a hard issue in opinion mining since many frequently used words may be non-relevant to a class. Therefore, we argue that using patterns or concepts to link several features together to make the meaning clear is a possible solution for dealing with uncertainties in opinions. Another alternative way is to put uncertain reviews into an uncertain boundary provisionally, and then design a more specific method to make a right decision.

Fuzzy logic has been used by researchers to solve the problem of uncertainties in opinion mining [5], [7], [8]. It represents uncertainties of features using fuzzy values between zero and one. Fig. 1 and Fig. 2 show the process of how to reduce the impacts of noisy features by using fuzzy logic. In the figures, opinion words "fine" and "ok" were frequently used in both positive and negative categories so they are noisy features. Although they have large weights, they are not useful for understanding the difference between the positive and negative categories. By using the fuzzy composition, we can reduce the large weights of noisy features to "1" if they are great than "1". The advantage of using fuzzy composition is to reduce impacts of noisy features, e.g., *OK* and *fine* in Fig. 1 and Fig. 2. Fuzzy composition can reduce their impacts

by adjusting their values (e.g., frequency) to one.

After analyzing many experimental results, we found that there is a critical problem for using fuzzy concepts to make binary decisions when the fuzzy similarity values of two categories are very close. To solve this hard research issue, in this paper, we propose a three-way decision framework to identify the uncertain boundary firstly. This framework classifies reviews into three regions in the first stage: the positive region, negative region, and uncertain boundary. It then presents a new feature selection model for using deep learning to further split the uncertain boundary into positive and negative regions. The three-way decision framework provides an elegant way to integrate fuzzy logic and deep learning in a single umbrella. There are many approaches that come up with different feature selection methods in opinion mining [5], [9], [10]. To find more specific features, the popular way is using three feature selection indexes. For using deep learning, people usually transfer textual features into numerical vectors (e.g, using word2vec, word embedding technique, to generate a vector for each word). This process requires a huge collection of documents that are related to the given topic; however, normally the very large collection is not designed specially for this opinion classification task. Therefore, the input vectors to a deep learning algorithm describe more general knowledge for the opinion classification. Also, deep learning does not have the capacity for dealing with uncertainties in features; therefore, it is very hard to produce satisfactory results for using deep learning directly. To make the proposed three-way decision framework more effective, we also propose a new feature selection model. It firstly discover closed patterns using the selected features, and then select fuzzy formal concepts (higher-level concepts) from the discovered closed patterns. The selected concepts are also used for the three-way decision framework to classify reviews into three regions in the first stage. After that, in the second stage, we integrate word-embedding vectors with the selected concepts and identify the difference between vectors and their mean value.

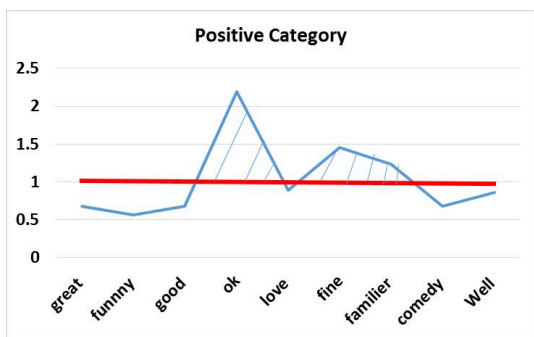


Fig. 1. Feature weights in positive reviews

The updated feature vectors are then used for a deep learning algorithm to split the boundary region clearly to negative and positive regions. The proposed framework provides valuable insights for dealing with uncertainties in user opinions.

This research contributes to opinion classification with a new feature selection model and a three-way decision framework. It presents a two-stage feature selection model to integrate an uncertainty processing model (e.g., fuzzy composition) and deep learning technique using a three-way decision framework. This research contributes new feature selection model and a three-way decision framework for opinion classification. To conduct a comprehensive evaluation for the proposed framework, we selected ve state-of-the-art models for performance comparison. We have conducted many experiments on the two datasets and the proposed model achieved impressive performance on F-measure (the F-measure value is 94.67% and 97.44% in movie review and ebook review datasets, respectively).

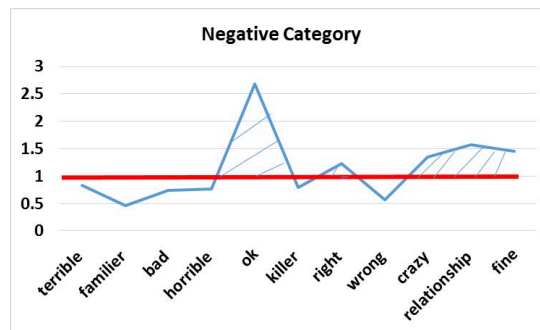


Fig. 2. Feature weights in negative reviews

II. RELATED WORK

A. Opinion Mining

Opinion mining (also known as sentiment analysis) refers to the use of natural language processing and text analysis to systematically identify, extract, quantify, and study affective states and subjective information. Opinion is widely applied to customer reviews and social media. Opinion mining contains the goal (entity) of the opinion, the attributes of the goal that the opinion is aimed at, and the sentiment (polarity) of a positive, negative or neutral opinion [11], [12].

B. Feature Selection

In feature selection there are term based methods such as TF-IDF, bm25, Uniformity (Uni) and Inverted Conformity Frequency (ICF). The bm25 is a ranking function used to rank matching documents according to their relevance to a given query. Li and Tsai used TF-IDF for feature selection in their model with Uniformity (Uni) and Inverted Conformity Frequency (ICF) [5].

Eq.1 and Eq.2 are the equations for Uni and ICF, respectively [5].

$$Uni(t_i) = \max_j \left[\frac{d_{ij}}{t_{ij}} \times \frac{d_{ij}}{\sum_{j=1}^k d_{ij}} \right], k = 2 \quad (1)$$

Where t_i is a term i , d_{ij} is the number of documents in which term i appears in category j . t_{ij} indicates the number of times

where term i appears in category j . When the Uni value in the Eq.1 is larger, it means the that term is more distinctive in a specific category. Li and Tsai used $\text{Uni} > 0.2$ as a threshold value for feature selection [5].

$$ICF(t_i) = \sum_{j=1}^k P_{ij} \log_2 P_{ij}, P_{ij} = \frac{d_{ij}}{|j|} \quad (2)$$

Where t_i indicates a term and j indicates category. d_{ij} is the number of documents which contain term t_i in category j . $|j|$ is the total number of reviews in category j . The ICF in Eq.2 indicates a term should appear frequently in a specific category instead of others. The smaller the ICF value indicates term frequently appears in a specific category. Li and Tsai used $ICF < \log(2)$ [5].

C. Knowledge Discovery

Discovered knowledge from opinions can be represented in different forms, such as terms, patterns, phrases, concepts, rules, relations, or ontologies. Some of them are combined with others to increase the accuracy.

For knowledge discovery pattern based methods are used as a higher level method [13]. In opinion mining, pattern mining can discover sequencing terms that frequently co-occur in a customer review, and such set of terms can represent the knowledge in reviews effectively. Frequent patterns and closed patterns are frequently employed to represent knowledge and trends in a dataset [13], [14]. Selecting reliable pattern is very vital, which enhances the efficiency of generating the frequent itemsets without losing any item [15]. In order to enhance the efficiency of pattern identification, researchers proposed several techniques: maximum frequent pattern mining [16], closed frequent pattern mining [17], top-k closed pattern mining [18]. Closed patterns [17] were proposed for handling large number of frequent patterns. A closed pattern is also a frequent pattern but it is not included in another sequential pattern that has the exactly same support. Therefore, the computational time of finding closed patterns may be reduced and it can also largely reduce the number of frequent patterns.

Fuzzy Logic can be used to extract knowledge from text data where a lot of uncertainties need to be coped with. Fuzzy Formal Concept Analysis (FFCA) [5], [7] is a theory which combines fuzzy logic and Formal Concept Analysis (FCA) to represent the uncertainties in data. FCA is a technique based on lattice theory [19], [20]. Formal concept is used to define the relationships between objects and attributes in a domain.

D. Classification

Classification can be supervised, unsupervised or semi-supervised learning [21]. Semi-supervised learning falls between unsupervised learning (with no labeled training data) and supervised learning (with only labeled training data). Therefore, supervised learning algorithm is very costly process, especially when dealing with large volumes of data. Semi-supervised was introduced [21]–[26]. Word polarity disambiguation is one important part of recent efforts on

semi-supervised learning for social data analysis. At the pre-processing stage, a vector of context features is built for each word w based on all its occurrences in the positive polarity corpus and another vector on its contexts in the negative polarity corpus [27], [28]. An ensemble framework for intensity prediction of sentiment and emotion is developed using three deep learning models based on LSTM, CNN and GRU and one feature-driven classical supervised model based on SVR [29]. Convolutional Neural Networks (CNNs) is the most popular neural network model used in recent years for opinion mining with significant performance [30]–[35]. Long Short Term Memory LSTM employs a bidirectional variant to capture sentiment for sentiment classification [36]. Recursive Tensor Neural Network is used to model correlations between different dimensions of child node vectors [37]. Hierarchical Convolutional Attention Networks (HCAN) was developed for text classification [38]. A novel attention model was recently implemented to incorporate user and product information for sentiment classification [39]. The newest technique applied in attention models has been reinforcement-Learning which implemented the learning based on how to best-react to situations through trial and errors [40]. New type of model called transformers are developed for sentiment analysis. Transformer-based method for sentiment analysis that encodes representation from a transformer and applies deep intelligent contextual embedding to enhance the quality of data by removing noise while taking word sentiments, polysemy, syntax, and semantic knowledge into account [41], [42].

Fuzzy composition was used to retrieve the relationships between concepts and categories in opinion classification [5]. They have shown high accuracy than other existing models [5], [43] in opinion mining. Rule-based fuzzy logic systems were built by several researchers [44], [45].

III. PROPOSED MODEL

As there are a lot of uncertainties in user reviews, there is an uncertain boundary, a margin that includes both positive and negative samples [6], between positive and negative reviews.

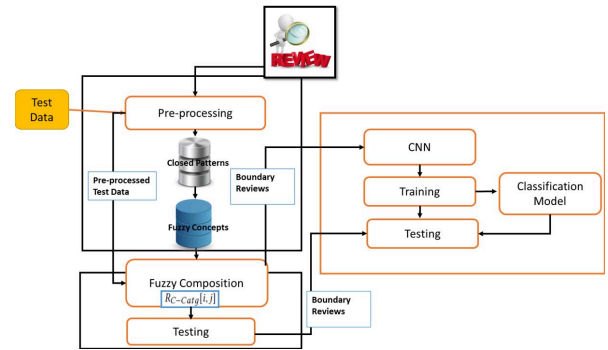


Fig. 3. Three-way Decision Framework

Therefore, it is very hard for a binary classifier to deal with uncertain boundary by using the traditional methods. A three-way based classification framework was proposed to address

the problem of uncertain boundary for binary classification [6] as in Fig. 3.

A. Fuzzy-based concept discovery and three-way decisions

It includes two transformations: ‘two-way to three-way’ (obtaining three regions: the positive region, negative region and uncertain boundary), and then ‘three-way to two-way’ (classifying the uncertain boundary into positive and negative regions). It also used a vector space model to find the uncertain boundary and then classify the uncertain boundary. The problem with using this existing framework for opinion mining is that the uncertain boundary is very big as the vector space model is not suitable for representing short text documents, such as user reviews.

In this paper, we extend the three-way decision framework for opinion mining. We firstly proposed a fuzzy-based model to infer the positive region, negative and uncertain boundary. We also integrate the fuzzy-model with a deep learning process to accurately classify the big uncertain boundary into positive and negative classes.

Term-based feature selection approaches suffer from poly-semy and synonymy [46] when used to deal with uncertainties within user reviews. Currently, a popular way for solving this problem in text mining is to extend low-level term spaces to higher level patterns or concepts [13]. In the following sections, we firstly define formal concepts within user reviews. We then provide a method to weight associations between concepts and selected term features. At last, we discuss how to use fuzzy composition to make the classification.

1) *Formal concepts*: Formally, let R be a set of reviews (a training set), T be a set of attributes (or terms that are used to describe reviews) and $I \subseteq R \times T$ be a binary relation between R and T (e.g., rIt means $t \in r$). A formal context is a triple $K := (R, T, I)$.

For a review sub-set $X \subseteq R$ and a term sub-set $Y \subseteq T$, derivation operators are defined as follows [47]:

$$X^I := \{t \in T | rIt \text{ for all } r \in X\}$$

$$Y^I := \{r \in R | rIt \text{ for all } t \in Y\}$$

The two derivation operators have the following properties:

$$Z_1 \subseteq Z_2 \implies Z_1^I \supseteq Z_2^I,$$

$$Z \subseteq Z^{II}, \text{ and } Z^{III} = Z^I.$$

A pair $\langle A, B \rangle$ is called a formal concept of K if $A \subseteq T$, $B \subseteq R$, $A = B^I$ and $B = A^I$. A is called its intent and B is called its extent.

The sub-concept (or super-concept) relation between two pairs of formal concepts can be formally described as follows:

$$\langle A_1, B_1 \rangle \preceq \langle A_2, B_2 \rangle \Leftrightarrow A_1 \subseteq A_2 \Leftrightarrow B_1 \supseteq B_2$$

Let $B(K)$ be the set of all formal concept of K , then $B(K)$ is a complete lattice (the concept lattice of K) for the above sub-concept relation.

Let A be a pattern (a sub-set of terms). Its *coverset* is a sub-set of reviews that include pattern A , and its closure

is $[\text{coverset}(A)]^I$. We call A be a closed pattern if $A = \text{closure}(A)$ [48].

Based on the above definition, a pair $\langle A, \text{coverset}(A) \rangle$ is a formal concept if pattern A is a closed pattern. We can prove $A^I = \text{coverset}(A)$ as $r \in \text{coverset}(A)$ iff for all $t \in A$ we have $t \in r$. Also, we have $[\text{coverset}(A)]^I = \text{closure}(A) = A$.

2) *Feature selection for generating formal concepts*: Formal concepts have elegant properties as described above; however, closed patterns only discuss terms’ binary appearances and long-closed patterns have very low frequency [49]. To solve these issues, in this paper, we select some useful term features firstly to reduce the time complexity for finding concepts; then using fuzzy composition to find the associations between concepts and categories.

In this research, reviews are pre-processed and applied *bm25*, *Uni* and *ICF* (three feature selection techniques) to find the useful terms. This is a very important step as we need to conduct many experiments to decide the combination strategy of using multiple feature selections. Preprocessing is very much vital for improving the performance of the classification process of customer reviews. Natural Language Processing (NLP) techniques of stemming, stop words removal and tokenization for pre-processing were applied as they have been identified by literature as the best techniques presently available. The output will be a set of terms as shown in Table I. Feature selection indexes of *bm25*, *Uni* and *ICF* are applied for the term selection. We have applied feature selection separately for both categories with the same method.

TABLE I
REVIEWS AFTER FEATURE SELECTION

Review	funny	good	pretty	great	comedy	awful
r1	x		x			
r2	x	x		x		
r3	x		x			
r4		x		x		x
r5	x		x		x	
r6	x	x		x		
r7	x	x				x

In this model multiple indexes of $bm25 > 30$, $Uni > 0.2$ and $ICF < \log(2)$ are used. The condition of $bm25 > 30$ was suggested based on the experiment results and the other two were based on the past literature [5]. The three conditions were checked simultaneously and if one condition did not satisfy, the term is eliminated. Table I shows an example of feature selection, where we assume that term set $T = \{\text{funny}, \text{good}, \text{pretty}, \text{great}, \text{comedy}, \text{awful}\}$ and review set $R = \{r_1, r_2, \dots, r_7\}$.

The frequency of terms might vary according to user reviews’ size. The normalization values of *bm25* weights are calculated in Eq. 3 [50] for a given category j (for binary classification, j is either the positive class or negative class).

$$nbm25(t_i) = bm25(t_i) \log_2 \left(1 + c \frac{avgR}{df_{ij}} \right) \quad (3)$$

where $avgR$ is the average review length in the whole collection; t_i is the original term frequency; df_{ij} is the number of

reviews in term i in category j and c is the free parameter of the normalization method which is 0.75.

The next task is to define a threshold to reduce the number of noisy terms. The significance of the feature selection method is to identify relevant features for mining closed patterns, and also reduce the size of reviews for generating closed patterns efficiently.

Closed patterns were generated using frequent itemset generation in FTree algorithm by defining a minimum support [51].

Identified patterns are larger than the terms. Hence, how to effectively deal with large amount of discovered patterns was the next challenge. Closed pattern mining algorithm with a minimum support value which is greater than 20 was introduced in our experiments. Thereafter, the change of the minimum support is used to control the size of discovered patterns. The discovered closed patterns are further processed to determine the formal concepts using a suitable minimum support as the threshold. Table II shows the selected formal concepts from closed patterns $CP = \{p_1, p_2, p_3, p_4, p_5\}$.

TABLE II
SELECTED FORMAL CONCEPTS

Concept	Intent	Extent	Original Pattern
c1	funny, pretty	r1, r3, r5	p1
c2	funny, good	r2, r6, r7	p5

3) *Fuzzy-based three-way decisions*: According to the above analysis, we can find a set of concepts C for a given set of reviews R , where R is either a set of positive reviews or a set of negative reviews, and we have

$$C = \{c = (Terms(p), coverset(p)) | p \in CP, supp_r(p) \geq \theta\}$$

where parameter θ is a minimum support; $Terms(p)$, the set of terms used in pattern p , is the intent of c and $coverset(p)$ is the extent of c .

In this sub-section, we discuss the relation between concepts and categories (e.g., the positive class and negative class for binary classification) using fuzzy composition [5] in order to classify reviews into three regions: the positive, negative and uncertain boundary.

The relation between concepts and categories is denoted as (I_{C-Catg}) which can be evaluated by using the fuzzy composition to integrate relation (I_{C-T}) between concepts and terms and relation (I_{T-Catg}) between terms and categories as described in Eq. 4.

$$I_{C-Catg} = I_{C-T} \circ I_{T-Catg} \quad (4)$$

where $I_{C-Catg}(i, j) = \max_{k \in T} \min[I_{C-T}(i, k), I_{T-Catg}(k, j)]$

The I_{C-T} relation is used to describe the uncertain factor of terms in a concept. In this paper, we use the normalized bm25 value $nbm25(t_k)$ as the uncertain factor for term k in concept i (see Eq. 5).

$$I_{C-T}[i, k] = nbm25(t_k) \quad (5)$$

The relation of Term-Category (I_{T-Catg}) can be obtained using the fuzzy composition over the Term-Review relation (I_{T-R}) and Review-Category relation (I_{R-Catg}) as shown in Eq. 6. In this paper, we use $nbm25(t_k)$ values to represent the relevant strength between terms and reviews and reviews and categories, respectively. Therefore (see the Eq. 7), $I_{T-Catg}(k, j)$ can be calculated using $nbm25$ weights for a given category j .

$$I_{T-Catg} = I_{T-R} \circ I_{R-Catg} \quad (6)$$

$$I_{T-Catg}(k, j) = \begin{cases} nbm25(t_k), & \text{if } nbm25(t_k) < 1 \\ 1 & \text{otherwise} \end{cases} \quad (7)$$

Algorithm 1 illustrates the idea for classifying user reviews into three regions: the positive region (POS), negative region (NEG) and the uncertain boundary (BND).

Algorithm 1 Three-way Classification

Require: $C, I_{C-T}, I_{T-Catg}, U$, two categories: $j = 1$ and $j = 0$ and an experimental coefficient δ

Ensure: I_{C-Catg} and three regions: POS, NEG, BND

```

1: for each  $j$  do
2:   for  $c_i \in C$  do
3:      $I_{C-Catg}(i, j) =$ 
4:      $\max_{k \in T} \min[I_{C-T}(i, k), I_{T-Catg}(k, j)]$ 
5:   end for
6: end for
7: for  $r \in U$  do
8:    $C_r = \{c \in C | intent(c) \subseteq r\}$ 
9:   for  $j$  do
10:     $f_j(r) = \max_{c_i \in C_r} [I_{C-Catg}(i, j)]$ 
11:   end for
12: end for
13: for  $r \in U$  do
14:    $POS = \{r \in R, f_{j=1}(r) - f_{j=0}(r) > \delta\}$ 
15:    $NEG = \{r \in R, f_{j=0}(r) - f_{j=1}(r) > \delta\}$ 
16:    $BND = \{r \in R, |f_{j=0}(r) - f_{j=1}(r)| \leq \delta\}$ 
17: end for

```

It firstly describes the training process (the first for loop) to calculating the relation I_{C-Catg} (a concept-category matrix) by using the fuzzy composition. It then calculates a fuzzy value for each new review r to a category j (see the second for-loop). At last, it determines the three regions (POS, NEG and BND) based on these fuzzy values for a given unlabeled review set U . The time complexity for the first for loop is $O(|C| * |T|)$ as we only two categorizes for binary opinion classification. For the second for loop, the time complexity is $O(|U| * |C|)$ as $|C_r| \leq |C|$; and the last for loop needs $O(|U|)$. Therefore, the time complexity of Algorithm 1 is $O(|C| * (|U| + |T|))$.

B. Boundary Classification

Algorithm 2 uses a parameter δ (a minimal value) to classify reviews into three regions as it is very difficult to decide the class of a review when its fuzzy values to the two classes are very closed.

Algorithm 2 Deciding parameter δ

Require: R, I_{C-T} two categories: $j = 1$ and $j = 0$ **Ensure:** δ

```
1: for  $r \in R$  do
2:   for each  $j$  do
3:      $f_j(r) = \sum_{t_k \in r, intent(c_i) \subseteq r} I_{C-T}(i, k)$ 
4:   end for
5:    $f(r) = \frac{f_0(r) + f_1(r)}{2}$ 
6: end for
7: for  $r \in R$  do do
8:   let  $\mu = \frac{1}{|R|} \sum_{r \in R} f(r)$ 
9: end for
10: let  $\delta = \sqrt{\frac{\sum_{r \in R} (f(r) - \mu)^2}{|R| - 1}}$ 
```

It clearly shows that the boundary region is a big problem for the classification. In this sub-section, we firstly propose a method to decide parameter δ . Different from other methods that use a probability-based loss function to decide δ [6], in this paper we use a standard derivation to decide parameter δ as it is a very small value. Algorithm 2 describes the process for calculating the value for parameter δ based on a training set R and the concept-term relations. The time complexity of Algorithm 2 is $O(|T| * |C| * |R|)$. As the numbers of terms and concepts are not very big, the algorithm for deciding δ is efficient.

It is obvious that the uncertain boundary (BND) includes a lot of uncertain reviews. Table III and Table IV show the size and the percentage of the uncertain boundary in both movie review and ebook review datasets, respectively.

In this sub-section, we also develop a new method to further classify the uncertain boundary (BND) as the data in BND are not linearly separable. In this paper, we assume that it may be possible to map the data into a higher-dimensional space resulting in a linearly separable set. There are many methods to do the mapping. For example, we may map a set of terms (a n -dimensional vector) $[t_1, t_2, \dots, t_n]$ to a $2n$ -dimensional vector $[t_1, t_2, \dots, t_n, t_1^2, t_2^2, \dots, t_n^2]$. However, it is likely to find the dependency between features. In this paper we use word embedding to do the mapping. It can find a higher dimensional space easily. Word embedding models map each word from the vocabulary to a vector of real numbers. We use word2vec model [52] in our experiments and each word was encoded by 8 dimension vector, that is,

$$\vec{w} = (x_1, x_2, \dots, x_8)$$

for all words $w \in \Omega$, where x_i are real numbers.

To solve the hard issue, in this paper, we integrate word embedding vectors and the terms (with assigned fuzzy values) that we selected from the intents of the formal concepts; and then use a CNN classifier to classify the uncertain boundary.

CNN consists with input layer, two convolution layers, two max-pooling layers, and a fully connected layer with softmax as the activation function. This is a more powerful CNN architecture when compared to standard CNN architecture [31]. This architecture has achieved high accuracy than CNN

architecture of Kim. The outputs of new word embedding are fed to the CNN for the input layer. The first convolution layer consists of 100 feature maps with filter size 4. The second convolution layer had 50 feature maps with filter size 2. The stride in each convolution layer is 1 as we wanted to tag each word. A max-pooling layer followed each convolution layer. The pool size we used in the max-pool layers was 3. We used regularization with drop-out on the penultimate layer with a constraint on norms of the weight vectors, with 50 epochs. The output of each convolution layer was computed using a non-linear function; in our case we used the hyperbolic sine.

TABLE III
BOUNDARY REVIEWS-MOVIE REVIEW DATASET

Category	Boundary Reviews	Percentage
Positive	115	11.5%
Negative	102	10.2%
Total	217	21.7%

TABLE IV
BOUNDARY REVIEWS-EBOOK REVIEW DATASET

Category	Boundary Reviews	Percentage
Positive	104	10.4%
Negative	101	10.1%
Total	205	20.5%

IV. EXPERIMENT RESULTS

Many experiments were conducted to evaluate the effectiveness of the proposed model. The purpose of these experiments is to evaluate our model and compare it with other state-of-the-art classification models. The performance of the model is measured by the F-measure. In addition, pair wise t-test is performed to verify the significance.

A. DataSets

The movie review dataset and ebook review dataset are used for the evaluation since two sets are standard datasets in opinion mining research [53]. The movie review dataset and ebook review dataset consist of 2000 movie reviews, with 1000 reviews for each category (negative and positive) for both testing and training purpose.

B. Baseline Models

In this research, we provide a comprehensive evaluation of the proposed model, we have selected eight baseline models. The models, including the state-of-the-art and closely related work, are used as baselines to evaluate the performance of the proposed model.

- Li and Tsai's [5] fuzzy model was developed and evaluated on movie review and ebook review datasets.
- SVM is used since it is an outstanding model in opinion mining in the literature [5].
- CNN model of Kim [31] is a deep learning model for opinion mining. It only used a movie review dataset for training and testing.
- Graph Convolutional Networks (GCN) [55] is a recent model developed using GCN and has shown significant

results. It has shown the accuracy of 76.74% for the movie review dataset.

- Attention Network [56] used Multi-sentiment-resource Enhanced Attention Network (MEAN) to alleviate the problem by integrating three kinds of sentiment linguistic knowledge(e.g., sentiment lexicon, negation words, intensity words) into the deep neural network via attention mechanisms.

C. Results

In order to evaluate the effectiveness of the proposed model, we compare the results with existing classification models as in Table V and Table VI. It is evident that the proposed model has the best performance in the two datasets comparing with models: FFCM [5], CNN [31], SVM [5], SVM [54], GCN [55] and Attention Network [56]. In Table VI we have not included CNN [31], GCN [55] and Attention Network [56] because those models were not evaluated with ebook review dataset. We also conducted the statistical significance testing (a two-tailed t-test). The results for the above models are shown in Table VII. The t-test results show that the proposed model is better than the four baseline models. It has shown from the Table VII our model is more significant than CNN and Attention Network which indicated that increment is > 0.5 .

TABLE V
F MEASURE OF STATE-OF-THE-ART MODELS: MOVIE REVIEW

Model	Precision	Recall	F measure
Proposed Model	0.9404	0.9509	0.9467
FFCM	0.8870	0.8840	0.8800
CNN	Not Given	Not Given	0.8150
SVM	0.8770	0.8690	0.8730
GCN	Not mention	Not mention	0.7674
Attention	Not mention	Not mention	0.8450

TABLE VI
F MEASURE OF STATE-OF-THE-ART MODELS: EBOOK REVIEW

Model	Precision	Recall	F measure
Proposed Model	0.9710	0.9777	0.9744
FFCM	0.9509	0.9508	0.9509
SVM	0.9382	0.9381	0.9382

TABLE VII
P-VALUES-MOVIE REVIEW-TWO TAILED

Model	P values -Movie	P values-ebook
Proposed Model	0.0023	0.0013
FFCM	0.0376	0.0317
CNN	0.0213	-
SVM	0.0199	0.0123
GCN	0.0087	-
Attention	0.0075	-

V. CONCLUSION

This study proposes a three-way based new framework for binary opinion classification by integrating fuzzy concepts and deep learning. This framework addresses the problem of uncertainties in opinion classification. This framework provides a promising way to construct an opinion classifier by identifying boundary region. This model uses statistical features for opinion classification. The experiment results show that the model

can significantly improve the performance of binary opinion classification in terms of F-measure. The results indicate that the proposed three-way based new framework performs binary classifications of user reviews with high accuracy.

The contributions made by the proposed framework are as follows:

- An innovative statistical feature selection method (combination of three feature selection indexes bm25, ICF and Uni) is proposed to retrieve fuzzy concepts.
- An effective classifier is implemented using fuzzy concepts and deep learning to classify three regions negative, positive, and boundary.

One potential future direction is to explore the effects of semantic structure of text data on the same architecture. Semantic structure might help to exploit the additional representation of fuzzy concepts and to investigate the impact of semantic patterns in fuzzy concepts is vital in opinion mining.

ACKNOWLEDGMENT

This research was supported by the Queensland University of Technology (QUT), Australia, University Grant Commission (UGC), Sri Lanka and University of Colombo School of Computing (UCSC), Sri Lanka.

REFERENCES

- [1] E. Cambria, "Affective computing and sentiment analysis," *IEEE intelligent systems*, vol. 31, no. 2, pp. 102–107, 2016.
- [2] P. K. Gupta, S. Gupta, and I. Arora, "Sentiment analysis for design of computing with words based recommender," in *2017 3rd International Conference on Computational Intelligence & Communication Technology (CICIT)*. IEEE, 2017, pp. 1–4.
- [3] E. Cambria and A. Hussain, "Sentic computing," *Cognitive Computation*, vol. 7, no. 2, pp. 183–185, 2015.
- [4] E. Cambria, B. Schuller, B. Liu, H. Wang, and C. Havasi, "Knowledge-based approaches to concept-level sentiment analysis," *IEEE intelligent systems*, vol. 28, no. 2, pp. 12–14, 2013.
- [5] S.-T. Li and F.-C. Tsai, "A fuzzy conceptualization model for text mining with application in opinion polarity classification," *Knowledge-Based Systems*, vol. 39, pp. 23–33, 2013.
- [6] Y. Li, L. Zhang, Y. Xu, Y. Yao, R. Y. K. Lau, and Y. Wu, "Enhancing binary classification by modeling uncertain boundary in three-way decisions," *IEEE Transactions on Knowledge and Data Engineering*, vol. 29, no. 7, pp. 1438–1451, 2017.
- [7] T. Quan, S. Hui, and T. Cao, "A fuzzy fca-based approach for citation-based document retrieval," in *2004 IEEE Conference on Cybernetics and Intelligent System*, vol. 1. IEEE, 2004, pp. 578–583.
- [8] K. Ravi, V. Ravi, and P. S. R. K. Prasad, "Fuzzy formal concept analysis based opinion mining for crm in financial services," *Applied Soft Computing*, vol. 60, pp. 786–807, 2017.
- [9] W. Medhat, A. Hassan, and H. Korashy, "Sentiment analysis algorithms and applications: A survey," *Ain Shams Engineering Journal*, vol. 5, no. 4, pp. 1093–1113, 2014.
- [10] E. Cambria, B. Schuller, Y. Xia, and C. Havasi, "New avenues in opinion mining and sentiment analysis," *IEEE Intelligent Systems*, vol. 28, no. 2, pp. 15–21, 2013.
- [11] M. Hu and B. Liu, "Mining and summarizing customer reviews," in *Proceedings of the tenth ACM SIGKDD international conference on Knowledge discovery and data mining*. ACM, 2004, pp. 168–177.
- [12] B. Liu, "Sentiment analysis and opinion mining," *Synthesis lectures on human language technologies*, vol. 5, no. 1, pp. 1–167, 2012.
- [13] Y. Li, A. Algarni, M. Albathan, Y. Shen, and M. A. Bijaksana, "Relevance feature discovery for text mining," *IEEE Transactions on Knowledge and Data Engineering*, vol. 27, no. 6, pp. 1656–1669, 2015.
- [14] N. Zhong, Y. Li, and S.-T. Wu, "Effective pattern discovery for text mining," *IEEE transactions on knowledge and data engineering*, vol. 24, no. 1, pp. 30–44, 2012.

- [15] C.-H. Chee, J. Jaafar, I. A. Aziz, M. H. Hasan, and W. Yeoh, "Algorithms for frequent itemset mining: a literature review," *Artificial Intelligence Review*, vol. 52, no. 4, pp. 2603–2621, 2019.
- [16] R. J. Bayardo Jr, "Efficiently mining long patterns from databases," *ACM Sigmod Record*, vol. 27, no. 2, pp. 85–93, 1998.
- [17] N. Pasquier, Y. Bastide, R. Taouil, and L. Lakhal, "Discovering frequent closed itemsets for association rules," in *International Conference on Database Theory*. Springer, 1999, pp. 398–416.
- [18] J. Han, J. Wang, Y. Lu, and P. Tzvetkov, "Mining top-k frequent closed patterns without minimum support," in *Data Mining, 2002. ICDM 2002. Proceedings. 2002 IEEE International Conference on*. IEEE, 2002, pp. 211–218.
- [19] L. A. Zadeh, "Fuzzy sets," in *Fuzzy Sets, Fuzzy Logic, And Fuzzy Systems: Selected Papers by Lotfi A Zadeh*. World Scientific, 1996, pp. 394–432.
- [20] T. Quan, S. Hui, and T. Cao, "A fuzzy fca-based approach for citation-based document retrieval," in *Cybernetics and Intelligent Systems, 2004 IEEE Conference on*, vol. 1. IEEE, 2004, pp. 578–583.
- [21] Z. Miao, Y. Li, X. Wang, and W.-C. Tan, "Snippext: Semi-supervised opinion mining with augmented data," in *Proceedings of The Web Conference 2020*, 2020, pp. 617–628.
- [22] I. P. Matsuno, R. G. Rossi, R. M. Marcacini, and S. O. Rezende, "Aspect-based sentiment analysis using semi-supervised learning in bipartite heterogeneous networks," *Journal of Information and Data Management*, vol. 7, no. 2, pp. 141–141, 2016.
- [23] B.-M. Nóra, C. Lemmaru, and R. Potolea, "Semi-supervised learning with lexical knowledge for opinion mining," in *Proceedings of the 2010 IEEE 6th International Conference on Intelligent Computer Communication and Processing*. IEEE, 2010, pp. 19–25.
- [24] Y. Lu and C. Zhai, "Opinion integration through semi-supervised topic modeling," in *Proceedings of the 17th international conference on World Wide Web*, 2008, pp. 121–130.
- [25] F. H. Khan, U. Qamar, and S. Bashir, "A semi-supervised approach to sentiment analysis using revised sentiment strength based on sentiwordnet," *Knowledge and information Systems*, vol. 51, no. 3, pp. 851–872, 2017.
- [26] N. Yu and S. Kubler, "Semi-supervised learning for opinion detection," in *2010 IEEE/WIC/ACM International Conference on Web Intelligence and Intelligent Agent Technology*, vol. 3. IEEE, 2010, pp. 249–252.
- [27] H. Zhao, Y. Xia, R. Y. Lau, and Y. Liu, "Word sentiment polarity disambiguation based on opinion level context," in *2012 International Conference on Machine Learning and Cybernetics*, vol. 5. IEEE, 2012, pp. 2007–2012.
- [28] Y. Xia, E. Cambria, A. Hussain, and H. Zhao, "Word polarity disambiguation using bayesian model and opinion-level features," *Cognitive Computation*, vol. 7, no. 3, pp. 369–380, 2015.
- [29] M. S. Akhtar, A. Ekbal, and E. Cambria, "How intense are you? predicting intensities of emotions and sentiments using stacked ensemble," *IEEE Computational Intelligence Magazine*, vol. 15, no. 1, pp. 64–75, 2020.
- [30] M. Hughes, I. Li, S. Kotoulas, and T. Suzumura, "Medical text classification using convolutional neural networks," *Stud Health Technol Inform*, vol. 235, pp. 246–50, 2017.
- [31] Y. Kim, "Convolutional neural networks for sentence classification," *arXiv preprint arXiv:1408.5882*, 2014.
- [32] X. Gu, Y. Gu, and H. Wu, "Cascaded convolutional neural networks for aspect-based opinion summary," *Neural Processing Letters*, vol. 46, no. 2, pp. 581–594, 2017.
- [33] X. Zhang, J. Zhao, and Y. LeCun, "Character-level convolutional networks for text classification," in *Advances in neural information processing systems*, 2015, pp. 649–657.
- [34] G. Chen, D. Ye, Z. Xing, J. Chen, and E. Cambria, "Ensemble application of convolutional and recurrent neural networks for multi-label text categorization," in *2017 International Joint Conference on Neural Networks (IJCNN)*. IEEE, 2017, pp. 2377–2383.
- [35] N. Majumder, S. Poria, A. Gelbukh, and E. Cambria, "Deep learning-based document modeling for personality detection from text," *IEEE Intelligent Systems*, vol. 32, no. 2, pp. 74–79, 2017.
- [36] K. Cho, B. Van Merriënboer, C. Gulcehre, D. Bahdanau, F. Bougares, H. Schwenk, and Y. Bengio, "Learning phrase representations using rnn encoder-decoder for statistical machine translation," *arXiv preprint arXiv:1406.1078*, 2014.
- [37] R. Socher, A. Perelygin, J. Wu, J. Chuang, C. D. Manning, A. Ng, and C. Potts, "Recursive deep models for semantic compositionality over a sentiment treebank," in *Proceedings of the 2013 conference on empirical methods in natural language processing*, 2013, pp. 1631–1642.
- [38] Z. Yang, D. Yang, C. Dyer, X. He, A. Smola, and E. Hovy, "Hierarchical attention networks for document classification," in *Proceedings of the 2016 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, 2016, pp. 1480–1489.
- [39] Z. Wu, X.-Y. Dai, C. Yin, S. Huang, and J. Chen, "Improving review representations with user attention and product attention for sentiment classification," in *Thirty-Second AAAI Conference on Artificial Intelligence*, 2018.
- [40] Y. Tay, L. A. Tuan, and S. C. Hui, "Dyadic memory networks for aspect-based sentiment analysis," in *Proceedings of the 2017 ACM Conference on Information and Knowledge Management*. ACM, 2017, pp. 107–116.
- [41] S. Alaparthi and M. Mishra, "Bidirectional encoder representations from transformers (bert): A sentiment analysis odyssey," *arXiv preprint arXiv:2007.01127*, 2020.
- [42] S. Y. Kim, K. Ganesan, P. Dickens, and S. Panda, "Public sentiment toward solar energy: Opinion mining of twitter using a transformer-based language model," *arXiv preprint arXiv:2007.13306*, 2020.
- [43] L. Subhashini, Y. Li, J. Zhang, and A. Athukorale, "Opinion classification using pattern mining and fuzzy logic," in *2018 18th International Conference on Advances in ICT for Emerging Regions (ICTer)*. IEEE, 2018, pp. 1–7.
- [44] M. Afzaal, M. Usman, A. C. M. Fong, S. Fong, and Y. Zhuang, "Fuzzy aspect based opinion classification system for mining tourist reviews," *Advances in Fuzzy Systems*, vol. 2016, p. 2, 2016.
- [45] L. Bing and K. C. Chan, "A fuzzy logic approach for opinion mining on large scale twitter data," in *2014 IEEE/ACM 7th International Conference on Utility and Cloud Computing (UCC)*. IEEE, 2014, pp. 652–657.
- [46] Y. Li, A. Algarni, and N. Zhong, "Mining positive and negative patterns for relevance feature discovery," in *Proceedings of the 16th ACM SIGKDD international conference on Knowledge discovery and data mining*. ACM, 2010, pp. 753–762.
- [47] J. Poelmans, S. O. Kuznetsov, D. I. Ignatov, and G. Dedene, "Formal concept analysis in knowledge processing: A survey on models and techniques," *Expert systems with applications*, vol. 40, no. 16, pp. 6601–6623, 2013.
- [48] Y. Li and J. Wu, "Interpretation of association rules in multi-tier structures," *International Journal of Approximate Reasoning*, vol. 55, no. 6, pp. 1439–1457, 2014.
- [49] N. Zhong, Y. Li, and S.-T. Wu, "Effective pattern discovery for text mining," *IEEE transactions on knowledge and data engineering*, vol. 24, no. 1, pp. 30–44, 2012.
- [50] B. He and I. Ounis, "Term frequency normalisation tuning for bm25 and dfr models," in *European Conference on Information Retrieval*. Springer, 2005, pp. 200–214.
- [51] M. Mahdi, S. Abdelrahman, R. Bahgat, and I. Ismail, "F-tree: an algorithm for clustering transactional data using frequency tree," *arXiv preprint arXiv:1705.00761*, 2017.
- [52] Y. Goldberg and O. Levy, "word2vec explained: deriving mikolov et al.'s negative-sampling word-embedding method," *arXiv preprint arXiv:1402.3722*, 2014.
- [53] B. Pang and L. Lee, "Seeing stars: Exploiting class relationships for sentiment categorization with respect to rating scales," *arXiv preprint cs/0506075*, 2005.
- [54] R. Moraes, J. F. Valiati, and W. P. G. Neto, "Document-level sentiment classification: An empirical comparison between svm and ann," *Expert Systems with Applications*, vol. 40, no. 2, pp. 621–633, 2013.
- [55] L. Yao, C. Mao, and Y. Luo, "Graph convolutional networks for text classification," *arXiv preprint arXiv:1809.05679*, 2018.
- [56] Z. Lei, Y. Yang, M. Yang, and Y. Liu, "A multi-sentiment-resource enhanced attention network for sentiment classification," *arXiv preprint arXiv:1807.04990*, 2018.