Article

Micro-Blog Sentiment Classification Method Based on the Personality and Bagging Algorithm

Wenzhong Yang, Tingting Yuan * and Liejun Wang

College of Information Science and Engineering, Xinjiang University, Urumqi, Xinjiang 830046, China;
564638827@stu.xju.edu.cn (W.Y.); wljxju@xju.edu.cn (L.W.)

* Correspondence: 1243058154@stu.xju.edu.cn; Tel.: +86-175-9002-1663

check for
updates

Abstract: Integrated learning can be used to combine weak classifiers in order to improve the effect of emotional classification. Existing methods of emotional classification on micro-blogs seldom consider utilizing integrated learning. Personality can significantly influence user expressions but is seldom accounted for in emotional classification. In this study, a micro-blog emotion classification method is proposed based on a personality and bagging algorithm (PBAL). Introduce text personality analysis and use rule-based personality classification methods to divide five personality types. The micro-blog text is first classified using five personality basic emotion classifiers and a general emotion classifier. A long short-term memory language model is then used to train an emotion classifier for each set, which are then integrated together. Experimental results show that compared with traditional sentiment classifiers, PBAL has higher accuracy and recall. The F value has increased by 9%.

---

1. Introduction

With the rapid development and maturity of Internet technologies, many online social platforms have gradually become the primary medium for people to obtain information and communicate with each other. Emerging social platforms, such as Weibo and WeChat, allow users to interact with information quickly and easily.

Weibo is not only a medium for people to communicate with each other, but also a way to express personal emotions. However, while expressing opinions, spreading thoughts, and expressing personal emotions, users also generate a large amount of information with personal subjective emotional characteristics. This information contains emotional characteristics of different tendencies. These emotional characteristics reflect the user's hobbies and interests. At the same time, it may also have a huge impact on the spread of Internet public opinion. Therefore, the sentiment analysis of Weibo text can understand the use's preferences and users' views on some hot events in the real society and make trend predictions.

Personality is a unique characteristic of an individual and profoundly affects the user's psychological state and social behavior. Personality research mainly focuses on the correlation between various personalities and the relationships between personalities and performance, creativity, and others, most of which are analyzed by self-reporting and regression algorithms. With developments in psychological research, people with the same personality have been found to exhibit similarities in writing and expressions. This feature is the basis for introducing personality into sentiment analysis. At present, personality-based sentiment analysis is still in its exploratory stages. Sentiment analysis does not differentiate the various ways of expressing emotions based on user individuality, nor does it consider the combination of sentiment analysis and personality analysis. In order to address this

problem, this paper proposes a personality-based Weibo sentiment analysis model, which introduces personality judgment rules to study the influence of personality on sentiment analysis.

2. Related Work

A number of studies have been conducted to improve traditional sentiment analysis methods. Bermudezgonzalez et al. [1] proposed building a comprehensive Spanish sentiment repository for subjective analysis of emotions. Cai et al. [2] solved the polysemy of emotional words by constructing a sentiment dictionary based on a specific domain. It is experimentally confirmed that the accuracy of using two superimposed classifiers, Support Vector Machine (SVM) and Gradient Boosting Decision Tree (GBDT) is better than that of a single model. Xu et al. [3] effectively constructed the sentiment classification of text by assembling an extended sentiment dictionary containing essential sentiment words, scene sentiment words, and polysemous sentiment words. Yang and Zhou [4] compare the processing speed and accuracy of Bayesian classifiers and support vector machine classification algorithms that implement sentiment mining for microblogs. Pang et al. [5] used emotional polarity determination for film reviews through three different supervised machine learning methods, namely support vector machine, naive Bayes, and maximum entropy. In the experiment, Pang et al. used unigram to construct vector features and then carried out chapter-level emotional polarity discrimination. The experimental results show that both the SVM and naive Bayes could achieve better emotional scores. Kamal and Abulaish [6] proposed an emotion analysis system based on a combination of rules and machine learning methods to identify feature-opinion pairs and their emotional polarity, in order to achieve user evaluation in different electronic products and attain user's emotional polarity. Song et al. [7] developed a new emotional word embedding technique. The primary framework differences are the joint code of morphemes and the part-of-speech tags. Under the proposed method, only important morphemes in the embedding space are trained to address the problem. This overcomes the traditional limitations of contextual word embedding methods and significantly improves the performance of sentiment classification. Sharma and Dey [8] proposes a hybrid sentiment classification model based on enhanced support vector machines. This model makes full use of the classification performance of boosting and support vector machines in sentiment-based online review classification. Experimental results show that in terms of sentiment-based classification accuracy, support vector machine integration using bagging or boosting is significantly better than a single support vector machine. Sharma et al. [9] proposes a method of emotion classification based on machine learning. The experimental results show that the combination of multiple emotion classifiers can further improve the accuracy of classification. Rong et al. [10] proposed an auto-encoder-based bagging prediction architecture (AEBPA), which has been shown to have huge potentials by experimental studies on commonly used datasets. Lin et al. [11] proposed a method to improve sentiment classification by adding weights to highlight emotional features for the first step. Bagging is then used to construct multiple classifiers on different feature spaces and are combined into one aggregate classifier. The results showed that the method could significantly improve the performance of sentiment classification. Wang and Han [12] propose a micro-blog sentiment analysis method that integrates an explicit semantic analysis algorithm. Wikipedia is regarded as an external semantic knowledge base, which improves the previous text representation method of micro-blog emotion analysis and improves the effectiveness of emotion classification. Waila et al. [13] used the SO-PMI-IR algorithm, based on unsupervised semantic orientation, to evaluate the classification method based on machine learning (Naive Bayes and SVM) in order to realize the emotion analysis in movie reviews. Mladenovic et al. [14] established a framework (SAFOS) using emotional dictionaries with emotion polarity scores and thesaurus of Serbian WordNet (SWN) in the feature selection process in order to execute emotion analysis in Serbian.

Numerous attempts have also been made to improve sentiment analysis techniques using deep learning. Yin et al. [15] propose a semantic enhanced convolutional neural network (SCNN) for sentiment analysis. Based on sentiwordnet, a widely used emotional vocabulary resource, two methods

of word embedding and emotion embedding are input into a convolutional neural network classifier, and good experimental results are obtained. Dan and Jiang [16] proposed a long short-term memory language model (LSTM) for sentiment analysis. Lu et al. [17] propose a p-lstm model based on long-term memory recurrent neural network (LSTM). The experimental results show that p-lstm has good performance in emotion classification task. In order to cope with the limitations of existing pre-trained word vectors which are used as inputs for CNN, Rezaeinia et al. [18] propose a novel method, Improved Word Vectors (IWV). The IWV improves the accuracy of CNNs which are used for text classification tasks. Jabreel and Moreno [19] combines two different methods for sentiment analysis. The first is N-Stream ConvNets, which is a deep learning method, and the second is XGboost regression based on a set of embedded and dictionary-based features. Abdi et al. [20] propose a method based on deep learning to classify the emotions expressed by users in comments (called RNSA). This method uses a unified feature set to analyze emotions, which represents word embedding, emotional knowledge, emotional transfer rules, statistics and speech knowledge. The experimental results show that the unified feature set learning method can achieve more significant performance than the feature set learning method. Liu and Chen [21] further studies deep learning and microblog sentiment analysis, extracts data from microblog by crawler, preprocesses it by corpus, takes it as the input sample of the convolutional neural network, establishes a classifier based on SVM / RNN, and finally judges the sentiment orientation of each sentence in a given test set. The experimental results show that the scheme can effectively improve the accuracy of emotional orientation, and the verification results are ideal. Hyun et al. [22] proposed a target-dependent convolutional neural network (TCNN) method of TLSA (target-level sentiment analysis) tasks. This method uses distance information on target words and neighboring words to understand their importance and achieve the classification task of extracting emotions from text targets. This approach is able to achieve better performance on single-target and multi-target datasets. Chen et al. [23] used BiLSTM and CNN neural network methods to improve the effect of sentiment analysis. In this approach, the BiLSTM-CRF sequence model is used to classify sentences into three types (no target, one target, multiple targets) based on the number of targets appearing in the sentence. Each set of sentences is then sent to a one-dimensional convolutional neural network of emotional classification. The experimental results show that the proposed method is able to improve the performance of sentence-level sentiment analysis and achieve the latest results from several benchmark datasets. Rezaeinia et al. [24] proposes an improved word vector (IWV) method for sentiment analysis. This method is based on part of speech tagging technology, word-based method, word location algorithm and word2vec/glove method. The experimental results show that the improved word vector (IWV) is very effective for emotion analysis. Sun et al. [25] utilized a deep neural network based on convolutional expansion features to perform sentiment analysis on Chinese micro-blogs. The posts and comments on Chinese micro-blogs are integrated to form a micro-blog session. Then, the automatic convolutional encoder is used for training to obtain the integrated features, and a deep belief network is used for the final sentiment classification. The experimental results show that under the proper structure and parameters, the performance of the deep belief network is better than that of SVM or NB. In order to solve the problem of mismatches between reviews and ratings on Amazon, Shrestha and Nasoz [26] used paragraph vectors to transform product reviews and used vectors to train a circular neural network of gated recursive units. This model combines the semantic relationship between review text and product information in implementing emotion analysis. Bijari et al. [27] developed a sentence-level graphical representation, which includes stop words that consider semantic and term relationships. The representation learning method of the sentence combination graph is employed to extract the underlying and continuous features of the document. Then, the learning characteristics of the document were entered into the deep neural network used for the emotion classification task. Hassan and Mahmood [28] proposed a neural network structure using convolutional neural networks (CNN) and long-short-term memory (LSTM) on pre-trained word vectors. In this approach, the ConvLSTM makes use of the LSTM to replace the

pool layer on CNN in order to reduce the loss of local detailed information and capture long-term dependencies on sentence sequences.

At present, most sentiment analysis is mainly based on text. However, with the rise of the picture sharing mode in social platforms, multi-modal sentiment analysis research on pictures, texts, and emoji has emerged. In a multimodal sentiment analysis method, Poria et al. [29] propose a new method to extract features from visual and text patterns by using deep convolution neural network. By inputting these features into the multi-core learning classifier, the performance of the emotion analysis task is better. You et al. [30] argue that pictures and texts should be jointly analyzed in a structured way. They developed a semantic tree mechanism, where the word and image areas in the text are mapped in implementing sentiment classification of image fusion. Jianzhong et al. [31] characterized Weibo messages using manual features (such as emotional word frequency, use of negative words, and emoji) and employed SVM for classification. Han and Ren [32] carried out sentiment classification by improving the Fisher discriminator of the kernel function. The use of latent semantic information with probabilistic characteristics as classification features is able to improve the classification effect of support vector machines. Cai and Xia [33] pre-trained text CNN and image CNN to obtain text and image representations, and then used CNN to connect two feature vectors. Yu et al. [34] used pre-trained CNNs to represent text and images and performed sentiment classification using logistic regression. Huang et al. [35] proposed the deep multimodal attention fusion (DMAF) method as a new image and text sentiment analysis model, which utilizes a hybrid fusion framework to mine distinguishing features and intrinsic relationships of visual and semantic contents. Xu et al. [36] developed a new bi-directional multi-level attention (BDMLA) model, using the complementary and comprehensive information between image modality and text modality to realize the joint classification of visual-text modality. Poria et al. [37] used multimodal cues that blended speech, video, and text for sentiment analysis. In this approach, the video is first collected from the website and is processed to obtain the features of the video, voice, and text. The three modes are then merged to obtain the final emotional polarity.

In terms of personality prediction, a number of psychological and computational scientific studies have been conducted exploring the relationship between people's language use and personality traits in the Big Five model [38]. Golbeck et al. [39] analyzed Twitter using structural and linguistic features and applied two regression algorithms to predict user personality traits. Bai et al. [40] suggested using multi-task regression and incremental regression to predict user personality in the online behavior among Sina micro-blog (weibo.com) users. They found that the Mean Absolute Error (MAE) on this particular microblog platform is between 0.1 to 0.2. In addition, Nowson et al. [41] applied a machine translation model to solve multi-language problems with text-based personality prediction. Their study achieved a root mean square error (RMSE) between 0.08 and 0.25.

Several studies have adopted integrated learning methods in emotional classification work, but the classification and personality prediction are in different research fields. Sentiment classification does not take into account the different emotional expressions of different personalities, nor does it couple in sentiment and personality analyses. Psychological research has shown that personality affects people's writing and speaking styles, and people having similar personalities tend to exhibit similar emotional expressions. Considering the potential relationship between emotion and personality, this paper proposes a microblog emotion analysis method based on a personality and bagging algorithm (PBAL).

3. Related Concepts

3.1. Ensemble Learning

Ensemble Learning is widely used for classification and regression tasks. Its principal concept is simple: different methods are used to change the distribution of the original training samples to build multiple different classifiers, which are then combined linearly to get a more robust classifier. The two main types of integrated learning are boosting and bagging.

Bagging is a popular integration approach that uses a bootstrapping algorithm to get multiple copies of the training set, which are then used to train different models. Voting schemes are then employed to incorporate projection and forecasting. Because the training sets are slightly different from each other, each model trained for these training sets would have different weights and focus, thus obtaining different generalization errors. By combining the models, the overall generalization error is expected to decrease to a certain extent. Figure 1 shows the flowchart of the bagging algorithm.
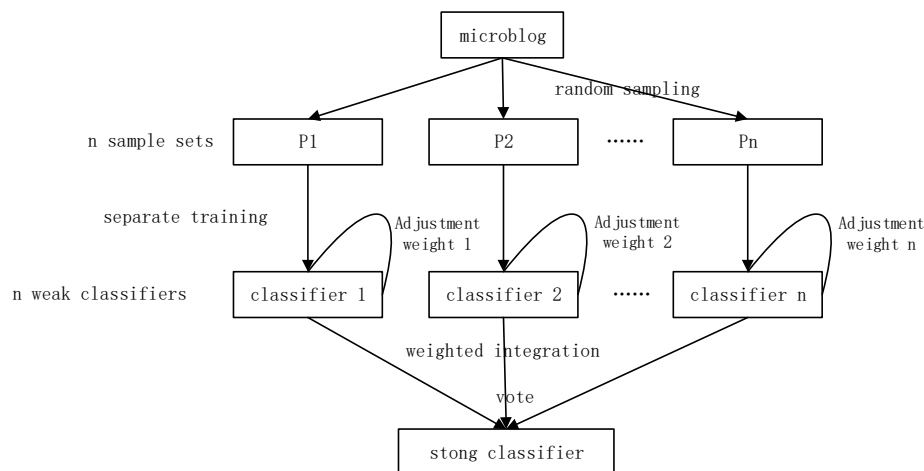

Figure 1. Flow chart of bagging algorithm.

3.2. Personality Model

In the field of psychology, some personality models have been proposed, such as the Big Five model and the MBTI (Myers–Briggs Type Indicator) model [42]. The Big Five model is a more authoritative personality model [38], which has been widely used in psychology and artificial intelligence [43]. The Big Five model characterizes personality based on five aspects: openness, conscientiousness, extroversion, agreeableness, and neuroticism. People with high openness are imaginative, creative, and curious. Conscientiousness reflects the degree of self-discipline and adequate preparedness for opportunities. Highly conscientious people are keen on work and eager for achievement. Extraverted people like to associate with others, while introverts prefer to be alone. Agreeable people are generous and trustworthy and are more inclined to help others. Nervousness reflects the degree of instability in human emotions.

3.3. LSTM

In-text sentiment analysis, the order relationship of words is crucial. Mikolov et al. [44] proposed a language model called the Recurrent Neural Network (RNN), which has been recognized as particularly suitable for processing text sequence data. However, when the interval between the related information of the text and the current position to be predicted increases, the number of backward propagation layers of the time-based direction propagation algorithm also grows. This would result in loss of historical information and gradient attenuation or explosion during training [45].

LSTM can be viewed as an improvement to the traditional RNN language model, which calculates the model error using a textual statement as an input sequence. The smaller the error, the higher the confidence in the text expression. However, as the text sequence information gets longer, the LSTM model becomes more effective in overcoming the attenuation problem of the sequence information.

4. Micro-Blog Sentiment Analysis Method Based on Personality and Bagging Algorithm

Since microblog texts published by users having the same personality type often contain similar emotional features, microblog texts are first classified into different character sets according to the personality. For each personality dataset, an emotion classifier is trained using the LSTM model, and

five basic personality emotion classifiers and one general emotion classifier are obtained. The voting method is then used to integrate learning on several basic classifiers. The sentiment classification process based on the personality and bagging algorithm is shown in Figure 2, where C, A, and E refer to extroversion, agreeableness, and conscientiousness, respectively. H is used to indicate high, and L is used for low. For example, HA means high agreeableness, while LA indicates low agreeableness.
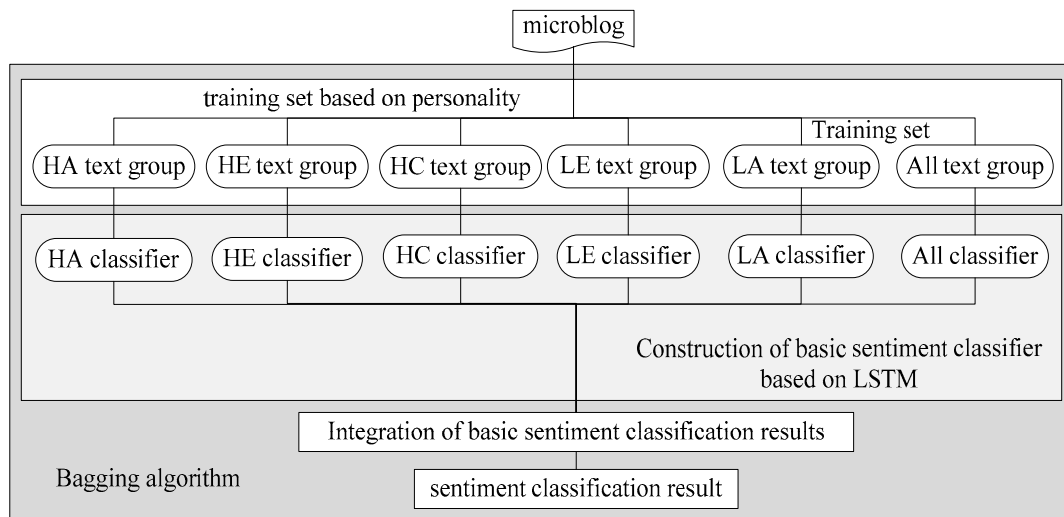

Figure 2. The sentiment classification process based on personality and bagging algorithm.

4.1. Text Personality Classification

In order to accurately assign micro-blog text to different collections, the personality characteristics of the text have to be predicted accurately. At present, personality prediction mainly involves three personality aspects in the Big Five model: extroversion, agreeableness, and conscientiousness. The two other dimensions, openness and neuroticism, are difficult to predict, as shown in previous studies [39,43], and therefore would not be included in this study. In scoring the personality, each dimension of personality can either be high or low. However, due to the low number of low conscientiousness texts in micro-blog, this study considers only five aspects: high extroversion, low extroversion, high agreeableness, low agreeableness, and high conscientiousness.

A rule-based personality classification method is used to evaluate the aspects of extraversion, agreeableness, and conscientiousness. The textual characteristics of each personality group reflect the commonalities of the corresponding emotional expressions. For example, conscientiousness expressions are often perceptions of achievement (e.g., results, persistence). Expressions of agreeableness are often related to love and praise (e.g., "love", "beautiful") and sympathy (e.g., "poor"). In contrast, the expressions of low-level agreeableness usually involve accusations or insults (e.g., "fool", "stupid"). Extraversion expressions convey positive (e.g., "happy") or negative (e.g., "lonely") emotions directly.

For example, if the number of words (Cword) or emoticons (Cemoction) expressed by a high degree of conscientiousness (HC) in the text is high, the value of text conscientiousness would be considered as high. Table 1 presents the main text features of personality prediction.

Table 1. Text characteristics of a personality prediction.

Feature Symbol	Symbol Meaning
HC_Cword	number of highly conscientiousness words in the text
HC_Cemoction	number of highly conscientiousness emoji in the text
HA_Cword	number of highly agreeableness words in the text
HA_Cemoction	number of highly agreeableness emoji in the text
LA_Cword	number of low agreeableness words in the text
LA_Cemoction	number of low agreeableness emoji in the text
HE_Cword	number of highly extraversion words in the text
HE_Cemoction	number of highly extraversion emoji in the text
LE_Cword	number of low extraversion words in the text
LE_Cemoction	number of low extraversion emoji in the text

Table 2 summarizes the set of rules for personality determination (where p1, p2... p10 are the thresholds used by the rule, and the size of the threshold is determined experimentally). Since the texts are evaluated using several aspects of personality, each text can belong to multiple personality sets.

Table 2. Rule of personality determination.

Rule Name	Rule	Rule Meaning
High conscientiousness judgment rule	IF $HC\_Cword \geq p1 \lor$ $HC\_Cemoction \geq p2$ THEN $C = HC$	When the number of words in the text containing the high conscientiousness dictionary exceeds p1, or the number of highly conscientiousness emoji in the text exceeds p2, the text is determined to be highly conscientiousness.
High agreeableness judgment rule	IF $LA\_Cword \geq p3 \lor$ $LA\_Cemoction \geq p4$ THEN $A = LA$	When the number of words in the text containing the high agreeableness dictionary exceeds p3, or the number of highly agreeableness emoji in the text exceeds p4, the text is determined to be highly agreeableness.
Low agreeableness judgment rule	IF $LA\_Cword \geq p5 \lor$ $LA\_Cemoction \geq p6$ THEN $A = LA$	When the number of words in the text containing the low agreeableness dictionary exceeds p5, or the number of low agreeableness emoji in the text exceeds p6, the text is determined to be low agreeableness.
High extraversion judgment rule	IF $HE\_Cword \geq p7 \lor$ $HE\_Cemoction \geq p8$ THEN $E = HE$	When the number of words in the text containing the high extraversion dictionary exceeds p7, or the number of highly extraversion emoji in the text exceeds p8, the text is determined to be highly extraversion.
Low extraversion judgment rule	IF $LE\_Cword \geq p9 \lor$ $LE\_Cemoction \geq p10$ THEN $E = LE$	When the number of words in the text containing the low extraversion dictionary exceeds p9, or the number of low extraversion emoji in the text exceeds p10, the text is determined to be low extraversion.

4.2. Ensemble Learning of Basic Emotion Classifier

A base sentiment classifier for each personality set is constructed using a tagged dataset. Micro-blog texts are sequenced data composed of words, which can have significant long-range dependencies, especially those reflecting sentiment and personality. LSTM is a sequence model, which can be used to construct a sentiment classifier and explore long-range dependencies.

In this study, the text is divided into different categories according to personality to obtain multiple training sets. One text may belong to multiple personality sets, and the data can be regarded as a choice with a return. A sentiment classifier is trained for each personality dataset using the LSTM. Each sentiment classifier is used to predict the sentiment tendency of the microblog text, and the results of all six basic classifiers are integrated. The integration process of the sentiment classifier is shown in Figure 3.
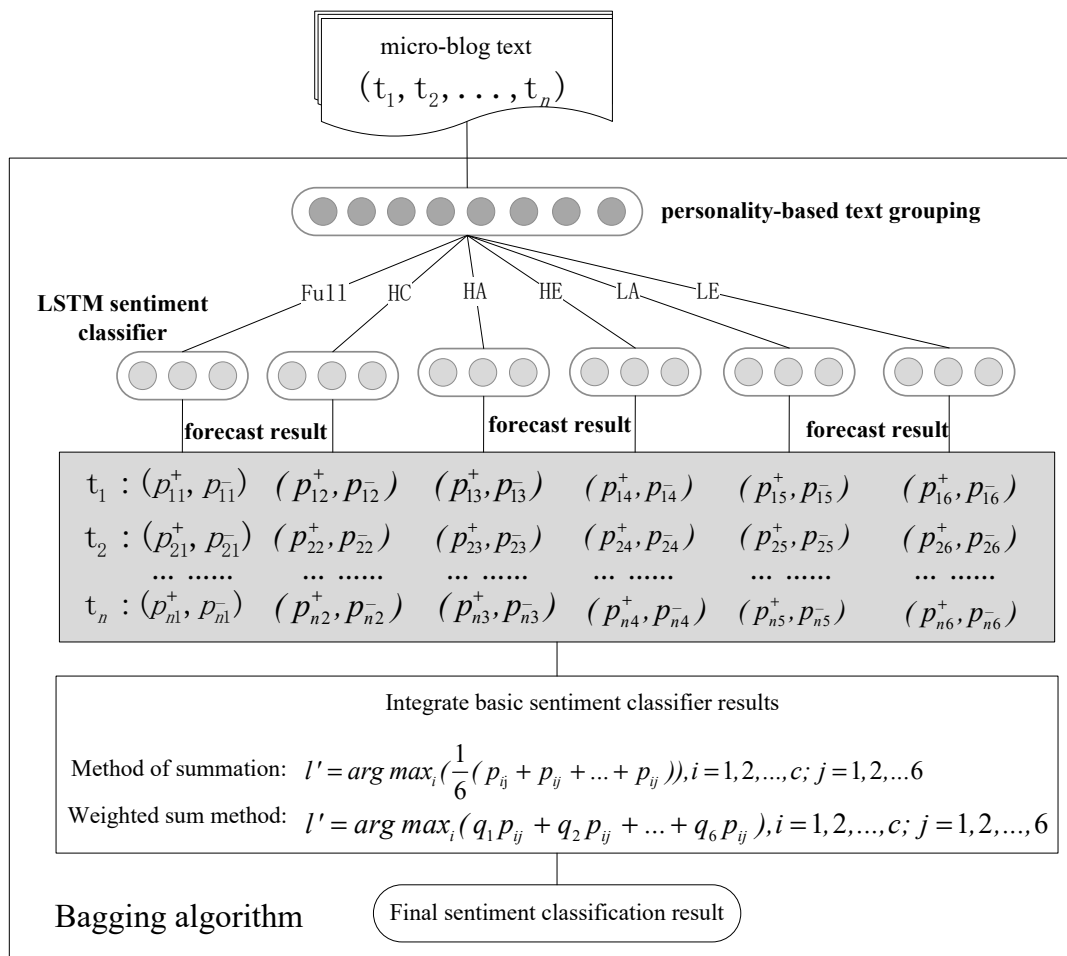
Figure 3. The process of integration of sentiment classifiers.

Given a set of microblog texts $(t_1, t_2, \cdots, t_n)$, six LSTM basic sentiment classifiers are used to generate output p_{ij}^{+}, p_{ij}^{-} for each of the microblog texts, where p_{ij}^{+} and p_{ij}^{-} represent the positive and negative probabilities of the microblog text computed by the *j*-th classifier, respectively. Based on the output of each basic sentiment classifier, the results are integrated to obtain the final emotional polarity. The base classifier results are then integrated using direct and weighted summation, which are expressed as follows:

$$l' = argmax_i\left(\frac{1}{6}\left(p_{ij} + p_{ij} + \cdots p_{ij}\right)\right), i = 1, 2, \ldots c; j = 1, 2 \ldots 6 \tag{1}$$

$$l' = argmax_i\left(q_1 p_{ij} + q_2 p_{ij} + \cdots q_6 p_{ij}\right), i = 1, 2, \ldots, c; j = 1, 2, \ldots, 6 \tag{2}$$

where l' represents the final emotional polarity; *i* is the number of emotional categories; p_{ij} is the output probability score of the six classifiers; and, $q_1, q_2, \cdots, q_6$ are the weights of each basic classifier. For this paper, *i* is equal to 2, and the emotional categories are positive and negative.

5. Micro-Blog Sentiment Classification Experiment

5.1. Experimental Data

The training data in this paper is from the literature [46], including the text published by Sina micro-blog users from 21 October 2009, to 15 December 2014. The dataset contains 10,474 texts, and each text is labeled with emotional polarity. Among them, there are 7562 positive and 2912 negative.

Test dataset 1 used the Chinese Weibo sentiment analysis evaluation data provided by the 2012 CCF Natural Language Processing and Chinese Computing Conference, which took 1100 texts, 600 positive texts, and 500 negative texts. The test dataset 2 consisted of a text crawled on Weibo using a crawler, with a total of 1500 pieces, 750 pieces each being positive and negative.

5.2. Basic Sentiment Classifier Comparison

This experiment was mainly used to verify the classification precision of each basic sentiment classifier, including five personality basic emotion classifiers and a general basic emotion classifier. The dataset contains 10,474 texts, comprising 3151 high conscientiousness datasets (HC), 3188 high agreeableness datasets (HA), 3204 low agreeableness datasets (LA), 5585 high extroversion datasets (HE), and 3154 low extroversion dataset (LE). For comparison purposes, 3600 data points were randomly extracted from the universal set to form the dataset (ALLpart) that has an equal number of samples for each category. The experimental results are shown in Table 3.

Table 3. Comparison of basic sentiment classifiers.

Personality	ALL	HA	HC	HE	LA	LE
F1score	88.31	89.19	89.59	87.39	72.28	83.25

As shown in Table 3, the F1 values of the HA classifier and HC classifier in the five personality basic emotion classifiers are higher than the ALL classifier. This indicates that the targeted sentiment analysis is effective on the different personality sets.

HC has the highest accuracy among the five sentiment classifiers. However, the overall accuracy is not particularly high, which could be the result of having a small dataset. Another possible explanation could be that, since the personality set is divided according to user personality, some errors may exist in classifying the user's personality, which could decrease the accuracy of the trained sentiment classifier.

5.3. Comparison of Bagging Algorithm Integration Methods

Experiments were used to verify the feasibility of applying the bagging algorithm. In the direct summation method, the weight of each classifier is kept equal (i.e., $q_j = \frac{1}{6}$). In the weighted sum method, the weight of each personality set is determined using cross-validation. The main criterion is to make precision as high as possible. In order to limit the computational workload, we restricted the number of experiments to 30 and picked the set with the highest precision as the approximate optimum. The optimal weights obtained for each personality set are: $q_1 = 0.16$, $q_2 = 0.23$, $q_3 = 0.16$, $q_4 = 0.16$, $q_5 = 0.14$, $q_6 = 0.15$. q_1, q_2,..., q_6 are the weight parameters of ALL, HC, HA, HE, LA, and LE. For comparison purposes, Table 4 lists some of the other weight settings.

Table 4. Partial weight setting comparison.

Weight (q_1,q_2,q_3,q_4,q_5,q_6)	Precision
0.16,0.23,0.16,0.16,014,0.15	96.91
0.15,0.15,0.23,0.16,0.15,0.16	96.81
0.14,0.16,0.25,0.15,0.15,0.15	96.81
0.15,0.15,0.15,0.24,0.15,0.16	96.63
0.15,0.27,0.15,0.14,0.15,0.14	96.55
0.14,0.15,0.15,0.15,0.16,0.25	96.45
0.16,0.16,0.16,0.21,0.15,0.16	96.43
0.2,0.16,0.16,0.16,0.16,0.16	96.42
0.16,0.16,0.15,0.16,0.21,0.16	96.41
0.14,0.16,0.15,0.16,0.24,0.15	96.36
0.14,0.15,0.14,0.15,0.14,0.28	96.27
0.25,0.15,0.15,0.15,0.15,0.15	96.27

The weight is usually proportional to the precision of the classifier on the training set; the higher the precision, the greater the weight. As shown in Table 3, the precision of the HC personality set is the highest, which consequently resulted in its weight being then largest. This is consistent with the results obtained in Table 4.

The experimental results of the two integration methods are shown in Table 5. The weighted sum method has a higher F score. The summation method assumes that all classifiers have equal weights. However, in practical applications, different classifiers should have varying weights given their different degrees of importance. Therefore, the weighted sum method is recommended as the final integration method.

Table 5. Comparison of integration methods.

Method	Precision	Recall	F Score
Direct sum	96.58	97.00	96.78
Weighted sum	96.91	97.00	96.95

5.4. Comparative Experiment

The precision, recall, and F1 values are used as evaluation indicators in comparing the proposed PBAL approach with other benchmark techniques:

(1)　SVM: The basic model of a Support Vector Machine (SVM) is to find the best-separated hyperplane in the feature space to maximize the positive and negative sample spacing on the training set. SVM is a supervised learning algorithm used to solve the two-class problem. It can also be used to solve nonlinear problems after the introduction of the kernel method.

(2)　LSTM: Proposed by Hochreiter and Schmidhuber (1997), LSTM is a special type of RNN that helps in learning long-term dependency information.

(3)　CNN-rand: The CNN-rand was proposed by Kim in 2014 [47] where all the words are randomly initialized and then modified during training.

(4)　CNN-static: presented also by Kim in 2014, the CNN static includes a pre-training vector model from word2vec. All words, including randomly initialized unknown words, are kept static, and only the other parameters are learned in the model.

(5)　CNN-non-static: Similar to CNN-static, the CNN–non-static includes a fine-tuned vector pre-trained for each task.

(6)　PBAL: See Table 6 for experimental parameter settings.

Table 6. Experimental parameters.

Parameter Name	Parameter Value
wordvector dimension	128
batch_size	100
learning_rate	0.01
dropout_ prob	0.5
num_epochs	750
hide layer nodes	75

The experimental results in test dataset 1 are summarized in Table 7. The F1 value of the proposed method was found to be higher compared with other benchmark methods. This suggests the effectiveness of personality and integrated learning in improving the emotional classification effect. Users with the same personality type are likely to be consistent with their expressions, and the sentiment classifiers trained towards a specific personality set are more precise than the general sentiment classifiers. The use of integrated learning can also make less-used personality-related

features more effective while improving the classification with the use of multiple classifiers. At the same time, we applied PBAL to test dataset 2. As shown in Table 8, the results are consistent with the experimental conclusions on test dataset 7.

Table 7. Experimental results on test dataset 1.

Method	Precision	Recall	F-Score
SVM	58.91	87.6	70.45
CNN-rand	73.03	85.37	78.72
CNN-static	73.52	88.49	80.31
CNN-non-static	73.55	88.36	81.79
LSTM	88.82	87.80	88.31
PBAL (proposed)	96.91	97.00	96.95

Table 8. Experimental results on test dataset 2.

Method	Precision	Recall	F-Score
SVM	61.97	85.26	79.69
CNN-rand	67.48	82.13	74.09
CNN-static	67.35	83.04	74.38
CNN-non-static	69.83	86.62	77.32
LSTM	78.41	84.67	81.42
PBAL (proposed)	84.54	91.87	88.05

In order to further study the influence of each personality on the final result of sentiment classification, different personality characteristics were combined and compared with the results of PBAL method. The experimental results are shown in Table 9. PBAL-HC indicates that the HC personality is not considered, PBAL-HA indicates that HA personality is not considered, PBAL-LA indicates that the LA personality is not considered, PBAL-HE indicates that HE personality is not considered, and PBAL-LE indicates that LE personality is not considered.

Table 9. Experimental results of the personality and bagging algorithm (PBAL) method lacking a set of personality in dataset 1.

	PBAL	PBAL-HA	PBAL-HC	PBAL-HE	PBAL-LA	PBAL-LE
precision	96.91%	94.45%	95.18%	95.18%	94.31%	96.18%
recall	97.00%	95.00%	94.40%	94.80%	94.26%	96.00%
F-Score	96.95%	94.72%	94.79%	94.99%	94.28%	96.09%

Comparing the results shown in Tables 9 and 10, the F1 value of the PBAL method is the highest. This suggests that in the ensemble learning process, the result of the basic emotion classifier without any character affects the final classification quality. Each personality contributes to the final classification effect, which means that the training of personality-based sentiment classifiers should make full use of the text collection of the various personality. Considering personality characteristics as comprehensively as possible would improve the performance of sentiment analysis.

Table 10. Experimental results of the PBAL method lacking a set of personality in dataset 2.

	PBAL	PBAL-HA	PBAL-HC	PBAL-HE	PBAL-LA	PBAL-LE
precision	84.54%	76.95%	80.00%	80.58%	82.70%	82.34%
recall	91.87%	90.80%	90.13%	92.40%	92.40%	92.00%
F-Score	88.05%	83.30%	84.76%	86.09%	87.28%	86.90%

6. Conclusion and Future Work

Significant similarities in user expressions have been shown for users with the same personality type. Thus, microblog texts can be divided into groups based on user personality, and a basic sentiment classifier can be trained for each personality type and then combined using integrated learning to achieve a stronger classifier. Based on the analysis of user personality, this paper constructs five types of personality classifiers, introduces the concept of personality based on traditional sentiment analysis, and proposes a microblog sentiment classification method based on personality and bagging algorithms. The results show that personality characteristics significantly improve the sentiment analysis of the text and provide new ideas for social network sentiment analysis.

While the use of sentiment analysis was found to be better than other typical machine learning methods, this study has several shortcomings. The rule-based personality classification method proposed in this paper may affect the accuracy of sentiment analysis because it uses fewer features. Future studies should further develop personality determination rules, extract more fine-grained personality characteristics, and improve classification accuracy. In our proposed approach, only the LSTM model is used to train the basic emotion classifier. Future research ought to consider using other neural network methods to determine if they can yield better results. The multimodality-based sentiment analysis method is one of the future development trends in this field. This means that future work should consider mining user personality characteristics in multimodal data and further study the relationship between emotion and personality.

References

1. Bermudezgonzalez, D.; Mirandajiménez, S.; Garcíamoreno, R.; Calderónnepamuceno, D. Generating a Spanish Affective Dictionary with Supervised Learning Techniques. In *New Perspectives on Teaching and Working with Languages in the Digital Era*; Research-publishing.net: Dublin, Ireland, 2016.

2. Cai, Y.; Yang, K.; Huang, D.; Zhou, Z.; Lei, X.; Xie, H.; Wong, T.L. A hybrid model for opinion mining based on domain sentiment dictionary. *Int. J. Mach. Learn. Cybern.* **2019**. [CrossRef]

3. Xu, G.; Yu, Z.; Yao, H.; Li, F.; Meng, Y.; Wu, X. Chinese Text Sentiment Analysis Based on Extended Sentiment Dictionary. *IEEE Access* **2019**, *7*, 43749–43762. [CrossRef]

4. Yang, Y.; Zhou, F. Microblog Sentiment Analysis Algorithm Research and Implementation Based on Classification. In Proceedings of the 2015 14th International Symposium on Distributed Computing and Applications for Business Engineering and Science (DCABES), Guiyang, China, 18–24 August 2015.

5. Pang, B.; Lee, L.; Vaithyanathan, S. Thumbs up? In Sentiment Classification using Machine Learning Techniques. In Proceedings of the Empirical Methods in Natural Language Processing, Philadelphia, PA, USA, 6–7 July 2002; pp. 79–86.

6. Kamal, A.; Abulaish, M. Statistical Features Identification for Sentiment Analysis Using Machine Learning Techniques. In Proceedings of the 2013 International Symposium on Computational and Business Intelligence, New Delhi, India, 24–26 August 2013; IEEE: Piscataway, NJ, USA, 2014.

7. Song, M.; Park, H.; Shin, K. Attention-based long short-term memory network using sentiment lexicon embedding for aspect-level sentiment analysis in Korean. *Inf. Process. Manag.* **2019**, *56*, 637–653. [CrossRef]

8. Sharma, A.; Dey, S. A boosted SVM based sentiment analysis approach for online opinionated text. In Proceedings of the Research in Adaptive and Convergent Systems, Montreal, QC, Canada, 1–4 October 2013; pp. 28–34.

9. Sharma, S.; Srivastava, S.; Kumar, A.; Dangi, A. Multi-Class Sentiment Analysis Comparison Using Support Vector Machine (SVM) and BAGGING Technique-An Ensemble Method. In Proceedings of the 2018 International Conference on Smart Computing and Electronic Enterprise (ICSCEE), Shah Alam, Malaysia, 11–12 July 2018; IEEE: Piscataway, NJ, USA, 2018; pp. 1–6.

10. Rong, W.; Nie, Y.; Ouyang, Y.; Peng, B.; Zhang, X. Auto-encoder based bagging architecture for sentiment analysis. *J. Vis. Lang. Comput.* **2014**, *25*, 840–849. [CrossRef]

11. Lin, D.; Chen, H.; Li, X. Improving Sentiment Classification Using Feature Highlighting and Feature Bagging. In Proceedings of the 2011 IEEE 11th International Conference on Data Mining Workshops (ICDMW), Vancouver, BC, Canada, 11 December 2011.

12. Wang, C.H.; Han, D. Sentiment Analysis of Micro-blog Integrated on Explicit Semantic Analysis Method. *Wirel. Pers. Commun.* **2018**, *102*, 1–11. [CrossRef]

13. Waila, P.; Singh, V.K.; Singh, M.K. Evaluating Machine Learning and Unsupervised Semantic Orientation approaches for sentiment analysis of textual reviews. In Proceedings of the 2012 IEEE International Conference on Computational Intelligence & Computing Research (ICCIC), Coimbatore, India, 18–20 December 2012.

14. Mladenovic, M.; Mitrovic, J.; Krstev, C.; Vitas, D. Hybrid sentiment analysis framework for a morphologically rich language. *Intell. Inf. Syst.* **2016**, *46*, 599–620. [CrossRef]

15. Yin, R.; Li, P.; Wang, B. Sentiment Lexical-Augmented Convolutional Neural Networks for Sentiment Analysis. In Proceedings of the IEEE International Conference on Data Science in Cyberspace, Shenzhen, China, 26–29 June 2017; pp. 630–635.

16. Dan, L.; Jiang, Q. Text sentiment analysis based on long short-term memory. In Proceedings of the 2016 First IEEE International Conference on Computer Communication and the Internet (ICCCI), Wuhan, China, 13–15 October 2016.

17. Lu, C.; Huang, H.; Jian, P.; Wang, D.; Guo, Y. A P-LSTM Neural Network for Sentiment Classification. In Proceedings of the Pacific-Asia Conference on Knowledge Discovery and Data Mining, Jeju, Korea, 23–26 May 2017; pp. 524–533.

18. Rezaeinia, S.M.; Ghodsi, A.; Rahmani, R. Text Classification based on Multiple Block Convolutional Highways. *arXiv* **2018**, arXiv:1807.09602.

19. Jabreel, M.; Moreno, A. EiTAKA at SemEval-2018 Task 1: An ensemble of n-channels ConvNet and XGboost regressors for emotion analysis of tweets. *arXiv* **2018**, arXiv:1802.09233.

20. Abdi, A.; Shamsuddin, S.M.; Hasan, S.; Piran, J. Deep learning-based sentiment classification of evaluative text based on Multi-feature fusion. *Inf. Process. Manag.* **2019**, *56*, 1245–1259. [CrossRef]

21. Liu, Y.; Chen, Y. Research on Chinese Micro-Blog Sentiment Analysis Based on Deep Learning. In Proceedings of the 2015 8th International Symposium on Computational Intelligence and Design (ISCID), Hangzhou, China, 12–13 December 2015.

22. Hyun, D.; Park, C.; Yang, M.; Song, I.; Lee, J.; Yu, H. Target-aware convolutional neural network for target-level sentiment analysis. *Inf. Sci.* **2019**, *491*, 166–178. [CrossRef]

23. Chen, T.; Xu, R.; He, Y.; Wang, X. Improving sentiment analysis via sentence type classification using BiLSTM-CRF and CNN. *Expert Syst. Appl.* **2017**, *72*, 221–230. [CrossRef]

24. Rezaeinia, S.M.; Rahmani, R.; Ghodsi, A.; Veisi, H. Sentiment analysis based on improved pre-trained word embeddings. *Expert Syst. Appl.* **2019**, *117*, 139–147. [CrossRef]

25. Sun, X.; Li, C.; Ren, F. Sentiment analysis for Chinese microblog based on deep neural networks with convolutional extension features. *Neurocomputing* **2016**, *210*, 227–236. [CrossRef]

26. Shrestha, N.; Nasoz, F. Deep Learning Sentiment Analysis of Amazon.Com Reviews and Ratings. *Int. J. Soft Comput. Artif. Intell. Appl.* **2019**, *8*, 1–15. [CrossRef]

27. Bijari, K.; Zare, H.; Kebriaei, E.; Veisi, H. Leveraging deep graph-based text representation for sentiment polarity applications. *Expert Syst. Appl.* **2020**, *144*, 113090. [CrossRef]

28. Hassan, A.; Mahmood, A. Deep Learning approach for sentiment analysis of short texts. In Proceedings of the 2017 3rd International Conference on Control, Automation and Robotics (ICCAR), Nagoya, Japan, 24–26 April 2017.

29. Poria, S.; Chaturvedi, I.; Cambria, E.; Hussain, A. Convolutional MKL Based Multimodal Emotion Recognition and Sentiment Analysis. In Proceedings of the International Conference on Data Mining, Barcelona, Spain, 12–15 December 2016; pp. 439–448.

30. You, Q.; Cao, L.; Jin, H.; Luo, J. Robust Visual-Textual Sentiment Analysis: When Attention meets Tree-structured Recursive Neural Networks. In Proceedings of the 24th ACM International Conference on Multimedia, Amsterdam, The Netherlands, 15–19 October 2016.

31. Xu, J.; Zhu, J.; Zhao, R.; Zhang, L.; He, L.; Li, J. Sentiment analysis of aerospace microblog based on SVM algorithm. *Res. Inf. Secur.* **2017**, *12*, 75–79.

32. Han, K.X.; Ren, W.J. The Application of Support Vector Machine(SVM) on the Sentiment Analysis of Twitter Database Based on an Improved FISHER Kernel Function. *Tech. Autom. Appl.* **2015**, *11*, 7.

33. Cai, G.; Xia, B. Convolutional Neural Networks for Multimedia Sentiment Analysis. In *Natural Language Processing and Chinese Computing*; Springer: Cham, Switzerland, 2015.

34. Yu, Y.; Lin, H.; Meng, J.; Zhao, Z. Visual and textual sentiment analysis of a microblog using deep convolutional neural networks. *Algorithms* **2016**, *9*, 41. [CrossRef]

35. Huang, F.; Zhang, X.; Zhao, Z.; Xu, J.; Li, Z. Image–text sentiment analysis via deep multimodal attentive fusion. *Knowl. Based Syst.* **2019**, *167*, 26–37. [CrossRef]

36. Xu, J.; Huang, F.; Zhang, X.; Wang, S.; Li, C.; Li, Z.; He, Y. Visual-textual sentiment classification with bi-directional multi-level attention networks. *Knowl. Based Syst.* **2019**, *178*, 61–73. [CrossRef]

37. Poria, S.; Cambria, E.; Howard, N.; Huang, G.-B.; Hussain, A. Fusing Audio, Visual and Textual Clues for Sentiment Analysis from Multimodal Content. *Neurocomputing* **2016**, *174*, 50–59. [CrossRef]

38. Hirsh, J.B.; Peterson, J.B. Personality and language use in self-narratives. *J. Res. Personal.* **2009**, *43*, 524–527. [CrossRef]

39. Golbeck, J.; Robles, C.; Edmondson, M.; Turner, K. Predicting Personality from Twitter. In Proceedings of the Privacy Security Risk and Trust, Boston, MA, USA, 9–11 October 2011; pp. 149–156.

40. Bai, S.; Hao, B.; Li, A.; Yuan, S.; Gao, R.; Zhu, T. Predicting Big Five Personality Traits of Microblog Users. In Proceedings of the 2013 IEEE/WIC/ACM International Joint Conferences on Web Intelligence (WI) and Intelligent Agent Technologies (IAT), Atlanta, GA, USA, 17–20 November 2013; pp. 501–508.

41. Nowson, S.; Perez, J.; Brun, C.; Mirkin, S.; Roux, C. XRCE Personal Language Analytics Engine for Multilingual Author Profiling: Notebook for PAN at CLEF 2015. *CLEF (Working Notes)*. Available online: http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.703.87&rep=rep1&type=pdf (accessed on 18 April 2020).

42. IMyers, B.; McCaulley, M.H.; Most, R. *Manual, a Guide to the Development and Use of the Myers-Briggs Type Indicator*; Consulting Psychologists Press: Palo Alto, CA, USA, 1985.

43. Qiu, L.; Lin, H.; Ramsay, J.; Yang, F. You are what you tweet: Personality expression and perception on Twitter. *J. Res. Personal.* **2012**, *46*, 710–718. [CrossRef]

44. Mikolov, T.; Karafiat, M.; Burget, L.; Cernocký, J.; Khudanpur, S. Recurrent neural network based language model. In Proceedings of the Conference of the International Speech Communication Association, Makuhari, Chiba, Japan, 26–30 September 2010; pp. 1045–1048.

45. Bengio, Y.; Simard, P.; Frasconi, P. Learning Long-term Dependencies with Gradient Descent is Difficult. *IEEE Trans. Neural Netw.* **1994**, *5*, 157–166. [CrossRef]

46. Du, Y.; He, Y.; Tian, Y.; Chen, Q.; Lin, L. Microblog bursty topic detection based on user relationship. In Proceedings of the IEEE Joint International Information Technology and Artificial Intelligence Conference, Chongqing, China, 20–22 August 2011; Volume 1, pp. 260–263.

47. Kim, Y. Convolutional Neural Networks for Sentence Classification. In Proceedings of the Empirical Methods in Natural Language Processing, Hong Kong, China, 3–7 November 2014; pp. 1746–1751.