Article

Short Text Aspect-Based Sentiment Analysis Based on CNN + BiGRU

Ziwen Gao ¹, Zhiyi Li ^{1,2,*} , Jiaying Luo ¹ and Xiaolin Li ¹

¹ School of Economics and Management, South China Normal University, Guangzhou 510006, China; ziwengao666@gmail.com (Z.G.); 2021020324@m.scnu.edu.cn (J.L.); cervusnippon_lea@foxmail.com (X.L.)
² Faculty of Information Technology, Macau University of Science and Technology, Macau 999078, China
* Correspondence: leeds@scnu.edu.cn

Abstract: This paper describes the construction a short-text aspect-based sentiment analysis method based on Convolutional Neural Network (CNN) and Bidirectional Gating Recurrent Unit (BiGRU). The hybrid model can fully extract text features, solve the problem of long-distance dependence on the sequence, and improve the reliability of training. This article reports empirical research conducted on the basis of literature research. The first step was to obtain the dataset and perform preprocessing, after which scikit-learn was used to perform TF-IDF calculations to obtain the feature word vector weight, obtain the aspect-level feature ontology words of the evaluated text, and manually mark the ontology of the reviewed text and the corresponding sentiment analysis polarity. In the sentiment analysis section, a hybrid model based on CNN and BiGRU (CNN + BiGRU) was constructed, which uses corpus sentences and feature words as the vector input and predicts the emotional polarity. The experimental results prove that the classification accuracy of the improved CNN + BiGRU model was improved by 12.12%, 8.37%, and 4.46% compared with the Convolutional Neural Network model (CNN), Long-Short Term Memory model (LSTM), and Convolutional Neural Network (C-LSTM) model.

Keywords: short text; aspect-level; sentiment analysis; convolutional neural network (CNN); bidirectional gating recurrent unit (BiGRU)

Citation: Gao, Z.; Li, Z.; Luo, J.; Li, X. Short Text Aspect-Based Sentiment Analysis Based on CNN + BiGRU. *Appl. Sci.* **2022**, *12*, 2707. <https://doi.org/10.3390/app12052707>

Academic Editor: Fabrizio Granelli

Received: 4 January 2022

Accepted: 2 March 2022

Published: 5 March 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Copyright: © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

With the development of e-commerce, more and more people are becoming willing to post their opinions and comments on certain products after consumption, forming a large number of comment texts. These short texts generally have a strong subjective nature, and sometimes contain different emotional tendencies within a sentence. Additionally, short text comments are highly colloquial. This makes the topic of the text vague and difficult to find, semantically incoherent, and, more disturbingly, difficult for researchers to use directly. At the same time, most current research on the sentiment analysis of short comment texts involves analysis from a coarse-grained perspective—that is, only one sentiment tendency is obtained from a paragraph. However, current user reviews generally contain a variety of complex emotions. For example, a product comment in Chinese read as follows: “I love this nice new look of the phone and its glass case, which makes it look fashionable. But it may be lagging after two week”. In a coarse-grained sentiment analysis, this comment would be marked as simply positive. In fact, the user only gave favorable comments on the appearance of the product, and negative comments on the product quality. It can be seen that coarse-grained sentiment analysis cannot accurately reflect the specific aspects that users really care about. Aspect-based sentiment analysis can determine the sentimental tendencies of different aspects in certain product review ontologies. Based on this advantage, aspect-based sentiment analysis of short texts of user reviews can better help consumers to make judgments and decisions, and can also help businesses to make targeted product improvements and increase user satisfaction.

In response to the above problems, this aim of this study was to improve the accuracy of aspect-based sentiment analysis through deep learning research methods. Convolutional neural network (CNN) can extract local feature information in a Chinese comment corpus very well, but can easily miss the long-distance features of the text in the extraction process. Recurrent neural network (RNN) has good memory ability and it is often used to extract the long-distance dependent information of a review corpus, as it can make up for the shortcomings of the convolutional neural network. At the same time, the bidirectional gating recurrent unit (BiGRU) is an “upgraded” version of the RNN, which can better solve the problems of gradient explosion and gradient disappearance. Based on the complementary advantages of the above two neural networks, this paper describes the construction of a hybrid model based on CNN and BiGRU, referred to simply as the CNN + BiGRU model. This model can extract the required local feature information and capture the aspect level of comments. Additionally, it can avoid the problems of gradient explosion and gradient disappearance, improve the accuracy of the model, and reduce the computational overhead. At the same time, in practical application, this model can also be used in the emotional analysis of short texts such as microblog comments and social public opinion posts, which play a guiding role in relevant fields of all walks of life.

2. Literature Review

At present, researchers in text sentiment analysis mostly use three types of methods: (1) the method of using the sentiment polarity dictionary for statistics; (2) the method of building a machine learning framework; (3) the method of deep learning based on a hierarchical model. However, the sentiment dictionary approach requires researchers to define the judgment rules, and the model cannot break away from these fixed sentiment word restrictions. Moreover, the training of machine learning sentiment analysis is a long task, which is too dependent on the categories marked by the text in the training set. In general, both the sentiment dictionary and machine learning approaches have existing problems, so the deep learning methods came into being. Among them, common deep learning models include CNN, RNN, and Long-Short Term Memory (LSTM), which are often used in the field of sentiment analysis and have made certain contributions.

In related research about deep learning methods, Hinton et al. [1] put forward the concept of the deep network model; this model uses a layer-by-layer greedy algorithm to overcome the problems of the deep network, so the performance of deep learning in all aspects is effectively improved. Zhang et al. [2] proved that convolutional neural networks can extract local n-gram features from text, have strong learning power for local features, perform well in feature extraction and text classification, and are a method that can relatively reduce the amount of calculation. However, a traditional CNN cannot deeply learn pooled features, and has the following shortcomings in feature extraction:

- The use of sigmoid causes the problem of gradient disappearance and slow convergence [3].
- The deeper the learning layer, the more serious the overfitting problem may be [4].
- The adoption of a gradient descent strategy may lead to an increase in cumulative error [5,6].

Therefore, this study used a hybrid model based on CNN and BiGRU (CNN + BiGRU) for feature learning.

In order to reduce the impact of the vanishing gradient problem in CNN, scholars have designed Gated Recurrent Unit (GRU) to alleviate it. GRU introduces an update gate and a reset gate. The update gate is used to control the extent to which the state information from the previous moment is brought into the current state, and the reset gate controls the amount of information from the previous state that is written to the current candidate set. Both of these gates alleviate the problems of vanishing gradient and long dependence [7]. However, there are some problems in standard GRU, such as incomplete learning of the eigenmatrix and inconsistent influence of the beginning and end of a sentence sequence on the state of the one-way hidden layer [8]. For this reason, the BiGRU network was

proposed by scholars. BiGRU is a neural network model jointly determined by the states of two unidirectional and opposite GRUs. At each moment, the input provides two GRUs in opposite directions simultaneously, and the output is determined jointly by the two unidirectional GRUs. For Chinese text, Wang et al. [9] proposed the use of the CNN and BiGRU joint network model to learn text features and extract the feature representation of sentences, thereby improving the accuracy of text sentiment analysis and the calculation speed of the model. Wang et al. [10] constructed a neural network model based on BiGRU which used BiGRU to extract features from deep information of text, and proved through experiments that the model had better accuracy and lower loss rate compared with classical models. Geng et al. [11] proposed a model based on BiGRU and an attention mechanism for the prediction of the novel coronavirus epidemic, and conducted an experiment to prove that BiGRU could reduce the computational cost and make full use of two-way data. In terms of fine-grained text sentiment analysis, Feng et al. [12] established a fine-grained feature extraction model based on BiGRU and attention, and proved through experiments the role of the BiGRU model in improving the accuracy of sentiment analysis.

Chinese text sentiment analysis is a process of analyzing sentences and judging the subjective feelings, opinions, and attitudes of the sentence users through a series of methods. Deep learning has been widely used in the field of text sentiment analysis. Traditional deep networks such as RNN and LSTM have been applied by many scholars in aspect sentiment analysis [13]. The function of sentiment analysis technology is to judge the emotional tendencies of Chinese sentences, judge whether the reviewer is positive or negative, and divide the text into several categories according to the reviewer's attitude. According to the delicacy and emphasis of the evaluated content, it can be divided into three levels: discourse level, sentence level, and aspect level. Zhu et al. [14] combined a multi-hop inference network to transform a sentiment analysis task into a reading comprehension task, and proposed a text sentiment analysis model based on multi-hop inference. Furthermore, scholars have refined the content of the research objects to study the sentiment analysis of sentence-level text. Wang et al. [15] proposed an algorithm based on the contribution of emotional polarity which predicts the sentence polarity of a corpus based on the position of words in the sentence, and proved the effectiveness of the algorithm through experiments. At present, the discourse-level and sentence-level sentiment analysis methods and technologies have been developed to a level of relative maturity, but both of these groups focus on overall sentiment analysis. This can lead to omission of details and miscalculations in application. Aspect-level sentiment analysis technology can be used to discover different objects in an aspect and to identify the emotional information expressed in a text for each aspect, effectively solving the appeal problem.

Aspect sentiment analysis was only proposed in 2010, and there have been few studies on this topic so far. It consists of two subtasks: aspect item extraction and aspect sentiment classification [16]. Specifically, aspect item extraction aims to extract the attributes of goods or services from a comment text; aspect-level sentiment classification should judge the emotional tendency corresponding to each aspect.

In aspect extraction, Paltoglou et al. [17] solved aspect extraction as a sequence labeling problem, and then used a linear chained conditional random field to deal with this problem. Traditional methods (such as constructing sentiment dictionaries) completely separate text representation from feature extraction and model training, and focus on text representation and feature extraction. Due to the randomness, high ambiguity, irregularity, and other characteristics of short text, this can easily lead to the problems of feature dispersion and context independence in the process of text representation and feature extraction. All of these factors may lead to lower accuracy of feature extraction and disconnection of contextual semantic relations when using traditional sentiment analysis methods [18]. To improve the accuracy of sentiment analysis based on sentiment lexicographs, Bravo-Marquez et al. [19] proposed a time-varying sentiment lexicographs based on incremental word vectors that train an incremental word sentiment classifier from a dynamic word vector to automatically update the sentiment lexicographs.

In aspect-based sentiment classification, there are multiple emotions co-existing in the short-text aspect-based sentiment analysis. Although this problem can be solved by refining the types of sentiment labels, this approach may also lead to a problem in that the model is too complex and the essence of the problem is not solved [20]. If the number of network layers in RNN is too great, the problem of gradient explosion or gradient disappearance will occur [21]. At the same time, existing heuristic methods cannot extract the semantic features of polysemous words efficiently, resulting in a poor classification effect and poor generalization of existing deep learning classification models. Therefore, how to effectively solve the above problems and improve the accuracy and generalization of aspect sentiment analysis is attracting extensive attention. In order to solve the above technical problems, Zhang [22] proposed a short-text sentiment analysis algorithm based on Bi-LSTM, which aimed at solving the problem that statistics-based feature selection methods ignore semantic information and deep learning methods do not contain the statistical and sentiment information of the feature. Tran et al. [23] proposed a model which uses BiGRU and did an experiment by training GloVe on the SemEval 2014 to prove the effectiveness of this model. Han [24] proposed an sentiment classification model based on BiGRU and knowledge transfer which uses BiGRU to classify sentiments more accurately according to the semantics of aspect words, and obtains domain knowledge by combination with the knowledge transfer method. Song et al. [25] used a network model based on a bidirectional gating cyclic neural network which used BiGRU to ensure that the model would have fewer network parameters and work faster than CNN.

In recent years, aspect-level sentiment analysis has also been applied in other fields. Alamoodi et al. [26] applied aspect-level sentiment analysis to research on the public's acceptance of vaccines during the COVID-19 pandemic. Alam et al. [27] applied aspect-based sentiment analysis based on parallel dilated convolutional neural networks to smart city applications.

In summary, aspect-based sentiment analysis has become a hot topic in the field of sentiment analysis in the past two years, and has attracted the attention of many sentiment analysis scholars. In actual application scenarios, aspect-based sentiment analysis has also been improved. The advantage of accuracy is resulting in the gradual replacement of sentence-level and text-level sentiment analysis. Chinese short text reviews usually contain both explicit aspect-level information and implicit aspect-level information. Therefore, aspect-level sentiment analysis technology requires not only explicit structural analysis, but also that attention be paid to implicit expression. Therefore, this paper combines CNN and BiGRU for aspect-level semantic sentiment analysis of short texts.

3. Construction of CNN + BiGRU

Based on CNN and BiGRU, this study modeled short texts from an e-commerce platform and realized aspect-based sentiment analysis. The CNN + BiGRU model constructed in this paper mainly involves the following three technologies:

- One-hot word embedding technology: This technology vectorizes the preprocessed word segmentation.
- An improved CNN structure: CNN constructs a multilayer neural network, performs multilayer calculations, and realizes multilevel feature extraction, which greatly improves the accuracy of the original neural network. The convolution structure of CNN greatly reduces the amount of data operations.
- BiGRU model: BiGRU can correlate the output at the current moment with the state at the previous moment and the state at the next moment, which is more conducive to the extraction of deep features of the text, so as to obtain a more complete text feature vector.

3.1. One-Hot Word Embedding Technology

One-hot technology uses the N-bit status register to encode N states, and converts the words in corpus sentences into a vectorized representation that can be operated and

understood by the computer, which it takes as the input of the model. In this study, one-hot was used to transform the segmented words into word vector form. One-hot technology first builds a dictionary G with size N . When encoding any word G , the position of the word in G is 1, and all other words are 0. For example, if the word embedding technique is performed on two short sentences, with one being “I like you” and the other being “I like your phone”, the process would be as follows. Firstly, a dictionary G is built of all Chinese words that appear in these two sentences, and this dictionary contains four words [I, like, you, mobile phone]. Secondly, according to one-hot encoding, the word “I” would be encoded as [1,0,0,0] and “like” would be encoded as [0,1,0,0], and so on. Finally, the one-hot vectors of each word in the sample are directly added, and the results of sentence vectorization are as follows: “I like you”: [1,1,1,0], “I like your mobile phone”: [1,1,1,1]. In this way, text is converted into a one-hot vector [28].

3.2. Convolutional Neural Network

Convolutional neural network is a kind of deep neural network which is often used in natural language processing and can also achieve good results in sentence classification and sentiment analysis. CNN technology generally includes a convolution layer and a pooling layer. The convolution layer continuously carries out sliding convolution of input data and outputs a feature mapping matrix. The pooling layer carries out pooling operations. In order to meet our research needs, this study adopted the pooling method of maximum pooling, and obtained the optimal solution of local value by taking the point with the maximum value in the locally accepted domain.

In the context of this article, we represent the text input as a matrix $S \in R_{n \times d}$ (n is the number of words in the sentence, d is the word vector dimension), and the convolution kernel as $W \in R_{h \times d}$ (h is the convolution kernel width, d is the word vector dimension). The feature map vector $O \in R_{n-k+1}$ is obtained through the convolution operation [24].

3.3. Bidirectional Gating Recurrent Unit

The Bidirectional Gating Recurrent Unit Network model has a simpler structure compared with LSTM. Thus, the training of it is more easy. Firstly, the BiGRU model uses gates to suppress the loss of information. The structure of the BiGRU model integrates the output gate and the forget gate in LSTM to form an update gate. Therefore, its structure is streamlined, which can save disk space to a large extent [9]. Secondly, the BiGRU does not control and retain internal memory, which reduces the consumption of memory space. The BiGRU highlights the key information of a text through output in two directions, and assigns corresponding weights to the extracted deep-level features to obtain a better feature extraction effect.

The GRU model can automatically learn which resources are useful and which resources can be discarded through training. Therefore, in extended texts, it may perform better. However, in a one-way GRU network, the state is always output from the front to the back and only one-way time series can be processed, which makes it easy to miss information when performing text sentiment analysis. There are also related studies that have proved that using a Bi-GRU bidirectional model performs better than GRU. Ay-oobi et al. [29] predicted the time series of new COVID-19 cases and new deaths through LSTM, convolutional LSTM, and GRU, and the results show that the error of the two-way model was lower than that of the one-way structure model. Hou et al. [30] proposed an attention mechanism recognition model based on BiGRU, which was applied to ship fault recognition, and proved through experiments that the time consumption of the bidirectional GRU network was smaller than that of other models, and the accuracy and recall rate were better. Therefore, this study used the two-way GRU model to obtain text vectors in two flow directions. In this model, another layer is added on the basis of a layer of GRU, and the directions of the two layers are different, so that not only the above information can be processed, but the following information can also be processed at the same time and plays an intermediate transition role. The positive and negative features are merged to

obtain more complete feature vector information of the text, and the input of the model is optimized. The BiGRU model is composed of two GRUs that are unidirectional and opposite in direction, combined to form a neural network model. All resources flowing through this network model are simultaneously used by the two layers of GRU.

BiGRU consists of two layers of GRU: forward output GRU and reverse output GRU.

The GRU model includes a reset gate and an update gate. The role of the reset gate is equivalent to that of the forgetting gate and input gate in LSTM, which determines how much the information at the previous moment is related to that of the current one. That is, there are some resources that do not need to be memorized. These meaningless resources are discarded by resetting the gate. The calculation formula is as follows:

$$R_t = \sigma(W_r \cdot [h_{t-1}, X_t]) \quad (1)$$

where R_t represents the reset gate and W_r is the weight matrix. X_t represents the input at time t and h_{t-1} is the output at the previous moment. In this part, the Hadamard operation is calculated to determine which information needs to be discarded and which information is retained. The operation can give a result within the interval $\{0, 1\}$. If the value of an element is 0, then it is completely useless; if the score of a value is closer to 1, then it indicates that it is more important.

The update gate determines when to update the state of the cell. For example, in the long sentence "the ipone which . . . , is red", in order to correctly generate the word "is", it needs to be in the input "ipone" to keep its information, so that the model will know that the input is a singular number. Thus, a sigmoid function is needed to map the value to $[0, 1]$, as shown in the following formula:

$$Z_t = \sigma(W_z \cdot [h_{t-1}, X_t]) \quad (2)$$

where Z_t represents the update gate and W_z is the weight matrix. X_t represents the input at time t and h_{t-1} is the output at the previous moment. After X_t is transformed by the sigmoid function and then operates with W_z , the output result at time t can be obtained. This result of whether the current state is the updated state or the previous state can be determined by the update gate. When the update gate is 1, the state changes, and when the update gate is 0, the current state is retained and the transmission continues.

The GRU neural network forward propagation formula is as follows [31]:

$$\tilde{h}_t = \tanh(W^h X_t + U^h (h_{t-1} * r_t)) \quad (3)$$

$$h_t = (1 - z_t) * \tilde{h}_t + z_t * h_{t-1} \quad (4)$$

where $\tilde{h}_t$ represents the candidate hidden layer, W^h and U^h are the weight matrix of GRU, $*$ is the multiplication element, and h_t represents the hidden layer.

The BiGRU model combines two unidirectional GRUs. At each moment, there is input to two GRUs in opposite directions at the same time, and the output is jointly determined to make the result more accurate.

The network structure of BiGRU is shown in Figure 1 below.

The output of the BiGRU can be described as

$$H = \vec{h}_t \oplus \overleftarrow{h}_t \quad (5)$$

where H represents the output of BiGRU, $\vec{h}_t$ and $\overleftarrow{h}_t$ represent two unidirectional GRUs, $\oplus$ and is the addition element.

Figure 1. BiGRU Network Structure.

3.4. CNN + BiGRU Experimental Model Construction

The model designed in this study is based on a CNN and BiGRU network and is divided into six levels: input layer, convolutional layer, pooling layer, BiGRU network layer, fully connected layer, and output layer. The structure of the BiGRU network layer is shown in Figure 2 below.

Figure 2. CNN + BiGRU Network Structure.

Input layer: this layer is responsible for data preprocessing and word vector training. (1) Data preprocessing. This layer performs deduplication, data cleaning, and word segmentation on the input comment sentences, and combines continuous character sequences into word sequences according to certain specifications through a series of processing steps.

(2) Word vector training. The input layer uses one-hot technology to convert the result of word segmentation into word vector form. Assuming the dimension of the word vector is N , the input vector of a comment text of length n after one-hot encoding can be expressed as follows:

$$X_{1:n} = X_1 \oplus X_2 \oplus \dots \oplus X_n \quad (6)$$

where X_n is the word vector of the n th word, $\oplus$ is the concatenation operator, and $X_{1:n}$ represents the feature vector composed of word vectors of the words $X_1 \sim X_n$.

The output word vector output after one-hot encoding can be expressed as follows:

$$W_{1:n} = X_{1:n} \cdot P \quad (7)$$

where P represents the weight of the feature value of the input vector and $W_{1:n}$ is the output word vector (that is, the result of multiplying the input vector and the weight matrix).

(3) Data vector transfer. By using a three-layer neural network to transform language into a spatial vector form, natural language is transformed into a machine-recognizable spatial vector.

Convolution processing: This layer passes the word vector set to the original sentence convolution layer for training. Feature extraction of the input text is completed through the set filter. We use a convolution kernel with dimensions $h \times k$ to perform convolution operations, where h is the height of the convolution kernel and n is the dimension of the word vector. In order to capture as much context information as possible, multiple sets of convolution kernels with different heights are generally set for operation, but as the convolution kernels increase, the training efficiency will decrease accordingly. Therefore, in the model training part, the sizes of the convolution windows were set to 2, 3, 4, and 5 respectively; the number of convolution kernels was set to 100. The calculation formula is as follows:

$$C_{hi} = f(W_h X_{i:i+h-1} + b), h = 2, 3, 4, 5 \quad (8)$$

where C_{hi} represents the generated feature sequence, W_h represents the weight matrix of different convolution kernels, b represents the bias vector, and $X_{i:i+h-1}$ represents the word vector matrix from i to $i+h-1$ in the sentence matrix. f represents the activation function, and the feature set is obtained after calculation:

$$C_1 = C_1 \oplus C_2 \oplus \dots \oplus C_{i+h-1} \quad (9)$$

Pooling layer: this layer samples the features obtained from the convolution layer and obtains the optimal solution of local values so as to realize the secondary screening of features and output the feature matrix with a fixed size to reduce the dimension of the results. The calculation of the convolution layer uses convolution kernels of different sizes, which will lead to inconsistency of the vector dimensions, and the maximum value of each vector can generally be considered the most important feature of the vector, so the method of maximum pooling was adopted to carry out the calculation of the pooling layer. The calculation formula for the feature vector graph after maximum pooling of different convolution kernels is as follows:

$$C_{pool} = \text{Max}(C_1, C_2, \dots, C_{n-h+1}), h = 2, 3, 4, 5 \quad (10)$$

BiGRU network layer: BiGRU layer is mainly composed of input, forward GRU, reverse GRU, and forward GRU output [8]. The input layer transmits data to both the forward GRU and the reverse GRU, and the output sequence is jointly determined by the two GRUs. The structure of the BiGRU network layer is shown in Figure 3 below.

Figure 3. BiGRU Network Layer.

The BiGRU layer expression is as follows:

$$\vec{h}_t = GRU(x_t, \vec{h}_{t-1}) \quad (11)$$

$$\overleftarrow{h}_t = GRU(x_t, \overleftarrow{h}_{t-1}) \quad (12)$$

$$h_t = f(W\vec{h}_t + W\overleftarrow{h}_t + b_t) \quad (13)$$

where $\vec{h}_t$ and $\overleftarrow{h}_t$ are the states of the positive and negative hidden layers at time t , respectively; $W\vec{h}_t$ and $W\overleftarrow{h}_t$ are the weights of the positive and negative hidden layers, respectively; and b_t is the bias of the hidden layer at time t .

Full connection layer: this layer connects the pooled result sequence structure to form an eigenmatrix which conforms to BiGRU's input specification. The full connection layer uses the tanh function as its activation function, and its output can be expressed as follows:

$$O = \tanh(W_k C_{pool} + b_k) \quad (14)$$

Output layer: This layer first connects the feature vector matrix of the BiGRU layer to the output layer, and uses the Softmax function to complete the short-text classification process. The specific process is expressed as follows [28]:

$$p(y_j) = \exp(u_j \cdot f_{sg} + b_j) / \sum_{j=1}^n \exp(u_j \cdot f_{sg} + b_j) \quad (15)$$

where $p(y_j)$ is the output of the short text in the i th category, and u_j and b_j are the weight matrix and bias corresponding to $p(y_j)$ in the i th category.

4. Experimental Study, Performance Evaluation, and Comparison

This experimental section mainly discusses the dataset used in the experiment, pre-processing, evaluation index, and method of comparison. This article used 5G mobile phone evaluation information under JD Mall. The main indicators used were accuracy rate, recall rate, and F1 value, which were used as standards to compare three classic models to evaluate the effect of the model.

4.1. Dataset and Preprocessing

In the data collection stage, this study used the Scrapy code and the 6573-comment text obtained by Octopus software from JD Mall. The original format of the data was an XLS file. First, the original data were converted into a TXT file, and then the text was converted into comments line by line. The length of the obtained text was at the sentence level. We then deleted the comment texts with fewer than five words and deleted the repetitive and various irrelevant texts, leaving 6233 comment texts as the final research object. The data were then preprocessed.

In the data preprocessing stage, the data were first cleaned to filter out all special symbols, punctuation, English, numbers, etc. This study used the `re.sub` function to remove special characters, and the `re.sub` ("[a-zA-Z0-9]", text) statement to remove English and numbers. Next, data segmentation is performed, and the Jieba segmentation in python was used for the operation. The Jieba word segmentation first recognizes which strings are involved in the Chinese text, and then uses a number of expressions to filter out the characters recognized in the previous step. Next, we performed the operation of removing stop words to remove useless words that have actual meanings but are extremely common in Chinese (such as: ah, ah, ah, ah). This study referred for this step to the stop word list of the Chinese Academy of Sciences. There were two main steps in the process of removing stop words: (1) Read the Chinese stop word list; (2) traverse the previously processed sentence, match the words in it with the stop word list of the Chinese Academy

of Sciences, and delete any word if it appears the same. Finally, part-of-speech screening was performed.

This study used the Jieba word segmentation component to mark the parts of speech. After marking according to the needs of the present research, other types of word segmentation were deleted, leaving only the four types of speech part words, namely idioms, nouns, other nouns, and noun verbs, namely $pos = ['n', 'nz', 'vn', 'l']$. Data preprocessing was completed through four steps: data cleaning, data segmentation, stop word removal, and part-of-speech filtering. The results are shown in Table 1.

Table 1. Preprocessing Results.

Glass, Color, Beautiful, A Kind of, Cocktail, Feeling
screen, quality, very, good, color, reduction degree, high, refresh rate, impressions, good, fluency,
take photographs, clear, boy, enough, BeautyCam, result, reducibility, high
Samsung, Android, camp, flagship, processor, benchmarking, ten thousand, daily, game, nice
stand-by-time, heavy, use, one day, long time, game, once, quick charge, one, hour, about, full,
take a bus, payment, convenient, mobile phone, overall, nice, cost performance, high, type,
favorite, configuration, nice
sound, nice, sound effect, good, double trumpet, not, common, nice, enough
...

4.2. Aspect-Level Feature Extraction Based on TF-IDF Vectorization

4.2.1. TF-IDF Algorithm

The TF-IDF algorithm is a feature extraction method recognized by academia. Compared with other algorithms, TF-IDF extraction is more accurate. Therefore, this study used the TF-IDF feature extraction algorithm to implement vectorization processing for 5G mobile phone review text. Through TF-IDF calculation, we were able to find whether a certain word was critical in this text sample. The specific calculation formula is as follows:

$$f(\omega) = TF(\omega) * IDF(\omega) = TF(\omega) * \frac{\log N}{n(\omega) + 1} \quad (16)$$

The values of $TF(\omega)$ and $IDF(\omega)$ are calculated separately using Formula 3, and then the total TF-IDF weight value is obtained and sorted in descending order. Choosing the first few keywords with the highest scores played a role in dimensionality reduction. The workflow is shown in Figure 4.

4.2.2. TF-IDF Keyword Table

The Scikit-learn machine learning library in Python contains a variety of functions for numerical operations, and also provides the `TfidfTransformer` function required by the TF-IDF algorithm in this article. The weight was calculated through the above process, so as to filter out suitable keywords. In this paper, 90 keywords from 5G mobile phone reviews were finally screened out, and their respective weights were obtained.

Figure 4. TF-IDF Workflow Chart.

4.2.3. Feature Induction

After summarizing and sorting out the keywords extracted by TF-IDF, the consumer review characteristics of 5G mobile phones were summarized into the following six categories: battery, appearance, function, performance, price, and service, which represent the most important considerations of consumers when buying 5G mobile phones. The six major elements of, the specific extracted feature words are shown in Table 2 below.

Table 2. Evaluation Feature Vocabulary.

Battery	Endurance, Battery Capacity, Charging Speed, Standby Time
appearance	screen, appearance, feel, screen, color, color, refresh rate, weight
function	camera, pixel, picture quality, lens, game, reaction, sound, sound effect, sound quality, effect
performance	run, system, quality, fingerprint, unlock, memory, CPU, cpu, configuration, signal
price	cost performance, price-insurance, price protection
serve	customer service, gifts, discounts, invoices, express delivery, delivery

First of all, if there is a problem with the battery of a mobile phone, the product must be repaired or discarded. Therefore, the performance of the battery is a major factor that consumers pay attention to. Secondly, as the user group of electronic products becomes younger, appearance has also become a major factor influencing purchase decisions. Thirdly, the comprehensiveness of the functions of mobile phones and the superiority of performance are the internal driving forces that determine the vast majority of users' purchase decisions. Moreover, the price will also impact consumer choices. These key elements are considered when using mobile phones. Consumers will have different evaluations of mobile phones with different prices and different grades. Finally, service is also a major factor considered by consumers, one which best reflects the sense of responsibility and service attitude of this merchant. It also affects consumers' evaluations of mobile phones.

4.2.4. Annotating Emotional Polarity

Using a manual labeling method, we referred to the summarized evaluation features and labeled the feature subject and emotional polarity of each 5G mobile phone review datum. The emotional polarity was replaced by the numbers “0” and “1”. This study adopted the two-polarity classification method, where “0” represents negative emotions, and “1” represents positive emotions. Finally, 7003 pieces of data were marked in this study, of which 6002 were training data and 1001 were test data. A total of 2777 negative emotions and 3215 positive emotions were marked in the training set. The positive and negative emotions were balanced and suitable as input data for model training. Examples of comment annotation are shown in Table 3.

Table 3. Evaluation data labeling table.

Comment	Feature	Polarity
The sound effect of watching movies is better than that of a single speaker. It will be easy to use after a period of using. But it is recommended to buy a large capacity one, this one is too small.	performance	0
	function	1
The mobile phone was bought at 8 o'clock yesterday and arrived at noon today. The logistics speed is really comfortable. It can be called Chinese speed. The mobile phone is beautiful and feels good. It runs better than the one I used before. It is still 5G.	service	1
	performance	1
	appearance	1
...	...	...

The evaluation text feature–polarity distribution diagram in Figure 5 below shows the ranked aspects of consumer focus on 5G mobile phone elements. It can be seen that the most concerning 5G mobile phone element for consumers is appearance, followed by performance and function. There were relatively few comments on the price. This may be due to the fact that consumers have a detailed understanding of the price before purchasing, so there is no excessive evaluation.

Figure 5. Length Distribution of Comment Text Data.

4.3. Experimental Settings and Evaluation Criteria

In this section, we first set up the experimental environment and built the experimental platform. The experimental environment parameters are shown in Table 4.

Table 4. Experimental Environment and Configuration.

CPU	Intel Core i5-8250U
RAM	32G
video card	NVIDIA 1070
operation platform	Window 10
Python	Python3.8
Deep learning framework	TensorFlow
	jieba = 0.42.1
	numpy = 1.18.5
	pandas = 1.0.1
Python tools library	matplotlib = 3.1.3
	scikit-learn = 0.23.2
	gensim = 3.8.3
	tensorflow = 2.3.0

This study used Python as the model implementation language and Tensorflow as the experimental framework. For the experiment, we first set up an experimental environment and built an experiment platform. The Chinese comments were vectorized using the one-hot model, and then the parameters of the model were set. The dimension of the word vector and the size of the hidden layer were 300. The hyperparameters in the model that needed to be adjusted were determined using the grid search method. After much iteration, the hyper parameter settings required for the experiment are shown in Table 5.

Table 5. Experimental Model Parameters.

Training parameters	epochs: int = 2	number of network training rounds
	batch_size: int = 64	network training batch size
	train_size: float = 0.9	training set splits the size
	test_size: float = 0.1	test set split size
	data_nums_min: int = 0	sample dataset size
Mesh parameter	maxlen: int = 50	the maximum length of the entered sentence
	vocab_size: int = 10,000	size of the word set
	embedding_dim: int = 100	embedded layer dimension, general size: $n > 8.33 * \log(N)$
	embedding_matrix: int = None	embedded layer pre-training
	dropout_rate: float = 0.3	dropout rate
	class_num: int = 2	classification number of category
	title_maxlen: int = 57	the onehot dimension entered by the title

In order to verify the effect of the CNN + BiGRU model proposed in this section, this study also used accuracy, precision, and F1 value (F1 measure) as experimental evaluation indicators. The specific formulas of each evaluation index are shown in Table 6.

Table 6. Calculation Formula of Experimental Indices.

Evaluation Index	Positive Emotion	Negative Emotion
accuracy	$\frac{P1}{P2} \times 100\%$	$\frac{N1}{N2} \times 100\%$
recall	$\frac{P1}{P3} \times 100\%$	$\frac{N1}{N3} \times 100\%$
F1 value	$\frac{2 \times \frac{P1}{P2} \times \frac{P1}{P3}}{\frac{P1}{P2} + \frac{P1}{P3}} \times 100\%$	$\frac{2 \times \frac{N1}{N2} \times \frac{N1}{N3}}{\frac{N1}{N2} + \frac{N1}{N3}} \times 100\%$

4.4. Comparison and Analysis of Experimental Results

After many iterations, the CNN + BiGRU model constructed in this article reached its optimal state and the accuracy and loss curve of the model was obtained, as shown in Figure 6 below. It can be seen from observation that as the number of batches increased,

the accuracy of the model continued to rise, the loss continued to drop, the learning rate was suitable, and there was no overfitting situation where loss and acc decreased simultaneously, indicating that the application effect of the model was relatively good. Another important point is that the experiment found that the CNN + BiGRU model has a relatively simple structure and the network training batch size was 64, which will not cause memory explosion and may reduce the training difficulty. Therefore, the CNN + BiGRU model reduced the calculation time and improved the operation efficiency.

Figure 6. Acc-Loss chart.

At the same time, in order to prove the superior reliability of the CNN + BiGRU model described in this paper, it was necessary to show that this model is superior to other methods. Therefore, this study included a comparative experiment in which the CNN + BiGRU model was compared with the traditional CNN model, LSTM model, and C-LSTM model. All experimental models can be simply divided into three parts, namely input, processing, and output. The experimental environment, experimental parameters, and number of iterations were all the same. This ensured that the internal structure of the model was unique and made the comparison results more convincing. The experimental results are shown in Table 7 below.

Table 7. Experimental Index Score.

Model	Accuracy	Recall	F1 Value
CNN	0.8358	0.8428	0.8375
LSTM	0.8733	0.8665	0.8742
C-LSTM	0.9124	0.8917	0.9245
CNN + BiGRU	0.9570	0.9638	0.9549

Among all the models, the CNN model had the worst performance in accuracy, recall, and F1 value. The reason may be that, on one hand, the number of 5G mobile phone reviews was relatively small, but they were all more than 30 words. CNN is limited by its inability to solve the problem of long-distance dependence on the context of the sentence, resulting in the loss of some information and making the performance of the CNN model very poor. For the LSTM model, for a comment sentence to enter the current processing structure, all previous units need to be traversed, which not only increases the workload of sentiment analysis tasks, but also easily causes the problem of gradient disappearance. The C-LSTM model uses CNN as an auxiliary, which improves the model's ability to extract local features, so the model's information extraction was also more perfect and the three evaluation index scores were improved. However, there were still some problems with the model, namely that when such models do not need to be trained, they still require a lot of resources, which increases the difficulty of training and causes a waste of resources. Finally, the experiment found that the accuracy of the CNN + BiGRU model was improved by 12.12%, 8.37%, and 4.46% compared with the other three models. It verified that the

model could not only extract the local features of the evaluated text, but also solve the contextual text dependence problem. The structure is relatively simple while ensuring the effectiveness of the model; taking into account the calculation speed and reducing the computational consumption, it can be seen that the CNN + BiGRU model constructed in this study is far better than other traditional models.

5. Conclusions and Future Outlook

In this paper, we propose an aspect-level text sentiment analysis method based on a convolutional neural network model and a bidirectional threshold recurrent neural network. By cleaning 5G crawler text, Chinese word separation, deleting meaningless discontinued words, and filtering lexicality, the original dirty data were turned into a corpus that could be directly input into the model; in the aspect ontology extraction session, the TF-IDF algorithm was used to obtain feature word vector weights; in the model building session, the one-hot word embedding technique was used to vectorize the subtext, CNN extracted local feature information and GRU extracted the long-distance dependency information, and both below and above information were processed through a bidirectional structure. Based on this, this study used CNN + BiGRU to obtain the contextual information of the text, managing to fully extract the local features of the review text, improve the accuracy of sentiment classification, avoid the gradient explosion and gradient disappearance problems, and reduce the loss value of the model by constantly adjusting the experimental parameter settings. Finally, by analyzing the accuracy, recall, and F1 score metrics, the CNN + BiGRU model constructed in this paper was determined to have significantly improved in each metric compared with the CNN model, LSTM model, and C-LSTM model.

In the comparison with the most basic CNN network model, the CNN + BiGRU model in this paper improved the accuracy by 12%; compared with the LSTM model and C-LSTM model, the CNN + BiGRU model in this paper improved the accuracy by 8% and 4%, respectively, and the model in this paper not only improved the classification effect and achieves the optimum result in all criteria, but also reduced the operation time significantly. It can be seen that the model in this paper has some applicable advantages in aspect-based sentiment analysis research of short comment texts, which can provide reference for the future direction of sentiment analysis.

Aspect-level sentiment analysis can extract all the sentiment tendencies expressed by consumers in all aspects of a product, and merchants can formulate targeted policies on this basis. Applying aspect-level sentiment analysis to recommendation systems can improve their practicability. Therefore, aspect-level sentiment analysis has high research value. This paper constructed a CNN + BiGRU model for aspect-level sentiment analysis, which improved relevant indicators to a certain extent. Regarding aspect-level sentiment analysis, there are some other different and valuable research works; Zhang [32] used graph neural networks to capture the implicit features between nodes in sentence relation graphs for sentiment analysis. Considering the problem of processing graph-structured data, Wu et al. [33] proposed a sentiment analysis model based on distance and graph-structured convolutional neural networks.

The authors' academic ability, research time, and academic energy are limited. Future studies could conduct in-depth research in the following areas:

- In the sentiment classification section, this paper only divided sentiments into positive and negative ones. Further research could carry out specific grading calculations for certain emotions, and further refine the classification of words of praise and derogation.
- In specific short-text contexts, different emotional entities are usually given different degrees of importance. Therefore, further research could introduce an attention mechanism to help exclude useless information and find key information from big data quickly.
- The CNN + BiGRU model could be used for feature fusion and applied to other fields, such as the technology for text backdoor attacks [34], mode recognition [35], etc.

- Similar models have been applied in sentiment analysis in other languages. Ayoobi et al. [36] proved the effect of GRU structure on improving the accuracy of sentiment analysis of Arabic text through experiments, and introduced the Multilingual Universal Sentence Encoder machine to improve the accuracy of sentiment analysis. However, the adaptability in other languages needs to be further studied.

Author Contributions: Conceptualization, Z.G. and Z.L.; methodology, Z.L.; software, Z.G.; validation, Z.G., Z.L.; formal analysis, Z.G.; investigation, Z.G.; resources, Z.L.; data curation, Z.G.; writing—original draft preparation, Z.G.; writing—review and editing, Z.L., J.L. and X.L.; visualization, Z.G.; supervision, Z.L.; project administration, Z.L.; funding acquisition, Z.L. All authors have read and agreed to the published version of the manuscript.

Funding: This paper is supported by the National Office for Philosophy and Social Sciences Project “Research on Cross-Modal Retrieval Model and Feature Extraction Based on Representation Learning” (No. 17BTQ062).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Hinton, G.E.; Salakhutdinov, R.R. Reducing the Dimensionality of Data with Neural Networks. *Science* **2006**, *313*, 504–507. [[CrossRef](#)] [[PubMed](#)]
2. Zhang, X.; Zhao, J.; LeCun, Y. Character-Level Convolutional Networks for Text Classification. *Adv. Neural Inf. Processing Syst.* **2015**, *28*, 649–657.
3. Almási, A.D.; Woźniak, S.; Cristea, V.; Leblebici, Y.; Engbersen, T. Review of Advances in Neural Networks: Neural Design Technology Stack. *Neurocomputing* **2016**, *174*, 31–41. [[CrossRef](#)]
4. Xu, G. Face Recognition Research Based on Convolutional Neural Network. Master’s Thesis, Harbin University of Science and Technology, Harbin, China, 2021.
5. Zhang, C. The Research of Bone Age Assessment Method Based on Convolutional Neural Networks and Multi-scale Feature Fusion. Master’s Thesis, Hefei University of Technology, Hefei, China, 2020.
6. Li, Z.; Xu, X.; Zhang, D.; Zhang, P. Cross-Modal Hashing Retrieval Based on Deep Residual Network. *Comput. Syst. Sci. Eng.* **2021**, *36*, 383–405. [[CrossRef](#)]
7. Irie, K.; Tüske, Z.; Alkhoul, T.; Schlüter, R.; Ney, H. Lstm, Gru, Highway and a Bit of Attention: An Empirical Overview for Language Modeling in Speech Recognition. In Proceedings of the 17th Annual Conference of the International-Speech-Communication-Association (INTERSPEECH 2016), San Francisco, CA, USA, 8–12 September 2016; pp. 3519–3523.
8. Cui, Y. Research on Sentiment Analysis of Product Reviews Based on Dual-channel Mixed Neural Network. Master’s Thesis, Xinyang Normal University, Xinyang, China, 2020.
9. Wang, L.; Liu, C.; Cai, D.; Lu, T. Chinese text sentiment analysis based on CNN-BiGRU network with attention mechanism. *J. Comput. Appl.* **2019**, *39*, 2841–2846.
10. Wang, W.; Sun, Y.; Qi, Q.; Meng, X. Text sentiment classification model based on BiGRU-attention neural network. *Appl. Res. Comput.* **2019**, *36*, 3558–3564.
11. Geng, H.; Sun, J.; Li, Y.; Wei, Y. Prediction of COVID-19 epidemic based on BiGRU-Attention network. *J. Wuhan Univ. Sci. Technol.* **2022**, *45*, 75–80.
12. Feng, X.; Liu, X. Sentiment Classification of Reviews Based on BiGRU Neural Network and Fine-Grained Attention. In Proceedings of the 3rd International Conference on Machine Vision and Information Technology (CMVIT), Guangzhou, China, 22–24 February 2019.
13. Jiang, H.; Jiao, R.; Wang, Z.; Zhang, T.; Wu, L. Construction and Analysis of Emotion Computing Model Based on LSTM. *Complexity* **2021**, *2021*, 8897105. [[CrossRef](#)]
14. Zhu, M.; Ban, H.; Zhao, L. A Textual Sentiment Analysis Model Based on Multi-Hop Reasoning. *Electron. Devices* **2021**, *44*, 628–632.
15. Wang, H.; Zheng, L.; Yi, P.; He, S. Sentence-Level Sentiment-Based Sentiment Polarity Classification of Chinese Online Reviews. *J. Manag. Sci.* **2013**, *16*, 64–74.
16. Yan, W. Research on Text Sentiment Analysis Based on Deep Learning. Master’s Thesis, Harbin University of Science and Technology, Harbin, China, 2021.
17. Paltoglou, G.; Thelwall, M. A Study of Information Retrieval Weighting Schemes for Sentiment Analysis. In Proceedings of the 48th Annual Meeting of the Association-for-Computational-Linguistics (ACL), Uppsala, Sweden, 11–16 July 2010; pp. 1386–1395.

18. Faruqui, M.; Dodge, J.; Jauhar, S.K.; Dyer, C.; Hovy, E.; Smith, N.A. Retrofitting word vectors to semantic lexicons. *arXiv* **2014**, arXiv:1411.4166.
19. Bravo-Marquez, F.; Khanchandani, A.; Pfahringer, B. Incremental Word Vectors for Time-Evolving Sentiment Lexicon Induction. *Cogn. Comput.* **2021**, *14*, 1–17. [[CrossRef](#)]
20. Yang, J. Research on Fine-grained Sentiment Analysis for Text. Doctoral Thesis, Nanjing University, Nanjing, China, 2019.
21. Thinh, N.K.; Nga, C.H.; Lee, Y.S.; Wu, M.L.; Chang, P.C.; Wang, J.C. Sentiment Analysis Using Residual Learning with Simplified Cnn Extractor. In Proceedings of the 21st IEEE International Symposium on Multimedia (ISM), San Diego, CA, USA, 9–11 December 2019; pp. 335–3353.
22. Zhang, S. Research on Key Issues of Chinese Short Text Sentiment Analysis. Master's Thesis, Northwest University, Xi'an, China, 2021.
23. Tran, T.U.; Hoang, H.T.T.; Huynh, H.X. Aspect Extraction with Bidirectional GRU and CRF. In Proceedings of the IEEE-RIVF International Conference on Computing and Communication Technologies (RIVF), Danang, Vietnam, 20–22 March 2019; pp. 1–5.
24. Han, Y. Research on BiGRU based Aspect-level Sentiment Classification Method for Drug Reviews. Master's Thesis, Northeast Forestry University, Harbin, China, 2021.
25. Song, H.; Zhang, Y. Sentiment classification method based on BiGRU and aspect attention module. *Intell. Comput. Appl.* **2020**, *10*, 83–87.
26. Alamoodi, A.H.; Zaidan, B.B.; Al-Masawa, M.; Taresh, S.M.; Noman, S.; Ahmaro, I.Y.Y.; Garfan, S.; Chen, J.; Ahmed, M.A.; Zaidan, A.A.; et al. Multi-perspectives systematic review on the applications of sentiment analysis for vaccine hesitancy. *Comput. Biol. Med.* **2021**, *139*, 104957. [[CrossRef](#)]
27. Alam, M.; Abid, F.; Guangpei, C.; Yunrong, L.V. Social media sentiment analysis through parallel dilated convolutional neural network for smart city applications. *Comput. Commun.* **2020**, *154*, 129–137. [[CrossRef](#)]
28. Cheng, Y.; Yao, L.; Zhang, G.; Tang, T.; Xiang, G.; Chen, H.; Feng, Y.; Cai, Z. Text sentiment orientation analysis of multi-channels CNN and BiGRU based on attention mechanism. *J. Comput. Res. Dev.* **2020**, *57*, 2583–2595.
29. Ayoobi, N.; Sharifrazi, D.; Alizadehsani, R.; Shoeibi, A.; Gorriz, J.M.; Moosaei, H.; Khosravi, A.; Nahavandi, S.; Chofreh, A.G.; Goni, F.A.; et al. Time series forecasting of new cases and new deaths rate for COVID-19 using deep learning methods. *Results Phys.* **2021**, *27*, 104495. [[CrossRef](#)]
30. Hou, T.; Zhou, L. Chinese Ship Fault Relation Extraction Method Based on Bidirectional GRU Neural Network and Attention Mechanism. *Comput. Sci.* **2021**, *48*, 154–158.
31. Cheng, Y.; Sun, H.; Chen, H.; Li, M.; Cai, Y.; Cai, Z. Text Sentiment Analysis Capsule Model Combining Convolutional Neural Network and Bidirectional GRU. *J. Chin. Inf. Processing* **2021**, *35*, 118–129.
32. Zhang, H.; Gou, G.; Chen, Q. Aspect-based sentiment analysis based on graph neural network. *Appl. Res. Comput.* **2021**, *38*, 3574–3580.
33. Wu, H.; Miao, Y.; Zhang, W.; Zhou, M.; Wen, Y. Aspect level sentiment analysis based on distance and graph convolution network. *Appl. Res. Comput.* **2021**, *38*, 3274–3278.
34. Kwon, H.; Lee, S. Textual Backdoor Attack for the Text Classification System. *Secur. Commun. Netw.* **2021**, *2021*, 2938386. [[CrossRef](#)]
35. Zhang, X.; Zhang, M.; Li, X. Rolling bearing fault mode recognition based on 2D image and CNN-BiGRU. *J. Vib. Shock.* **2021**, *40*, 194–201.
36. Mohammad, A.S.; Hammad, M.M.; Sa'ad, A.; Saja, A.T.; Cambria, E. Gated Recurrent Unit with Multilingual Universal Sentence Encoder for Arabic Aspect-Based Sentiment Analysis. *Knowl. -Based Syst.* **2021**, 107540. [[CrossRef](#)]