electronics

MDPI

Article

Sentiment-Target Word Pair Extraction Model Using Statistical Analysis of Sentence Structures

Jaechoon Jo [1], Gyeongmin Kim [2] and Kinam Park [2,*]

[1] Division of Computer Engineering, Hanshin University, Osan 18101, Korea; jaechoon@hs.ac.kr
[2] Department of Computer Science and Engineering, Korea University, Seoul 02841, Korea; totoro4007@korea.ac.kr
* Correspondence: spknn@korea.ac.kr

Abstract: Product information has been propagated online via forums and social media. Lots of merchandise are recommended via an expert system method and is considered for purchase by online comments or product reviews. For predicting people's opinions on products, studying people's thoughts via extracting information in documents is referred to as sentiment analysis. Finding sentiment-target word pairs is an important sentiment mining research issue. With the Korean language, as the predicate appears at the very end, it is not easy to find the exact word pairs without first identifying the syntactic structure of the sentence. In this study, we propose a model that parses sentence structures and extracts sentiment-target word pairs from the parse tree. The proposed model extracts the sentiment-target word pairs that appear in the sentence by using parsing and statistical methods. For extracting sentiment-target word pairs, this model uses a sentiment word extractor and a target word extractor. After testing data from 4000 movie reviews, the applicable model showed high performance in both accuracy 93.25 (+14.45) and F1-score 82.29 (+3.31) compared with others. However, improvements in the recall rate (−0.35) are needed and computational costs must be reduced.

Keywords: opinion mining; parser; word extraction; data mining; statistical model

Citation: Jo, J.; Kim, G.; Park, K. Sentiment-Target Word Pair Extraction Model Using Statistical Analysis of Sentence Structures. *Electronics* **2021**, *10*, 3187. https://doi.org/10.3390/electronics10243187

Academic Editor: Taeshik Shon

Received: 11 November 2021
Accepted: 16 December 2021
Published: 20 December 2021

1. Introduction

Traditionally, some information has been spread orally but recently it has also been propagated online via forums such as blogs, Twitter, Facebook, etc. Recent research on the consumption patterns of more than 2000 adults produced the following results: 81% of Internet users conduct their consumption activities on the Internet; 20% conduct consumption activities on the Internet daily; reviews written by the opinion readers exert influence on the consumer activities of others at 73%; 80% of consumers prefer goods that are rated five stars to goods rated four stars; and 32% of the overall grades of merchandises are decided online via an expert system method and about 30% or so are determined by the online comments or product reviews, etc. [1,2]

Studying other people's thoughts by extracting online documents is referred to as opinion mining, an operation that deciphers and extracts subjective information or opinions from the source material. Research on opinion mining is concerned with two important tasks: Making dictionaries with words tagged with opinions and finding target words represented by opinions. For example, when analyzing "This cellphone's LCD is bright", through opinion mining, it can be analyzed as having the sentiment word "bright" and the target word "LCD". Building a dictionary of words tagged with sentiments permits the determination of whether the sentiment word "bright" was a word used in a positive or negative sense. And distinguishing that the word pointed to by the sentiment word "bright" is in fact "LCD" is what solves the issue of finding the target word. In opinion mining for target words, the predicate and object that represent attributes and sentiments have distinctly significant meanings. Since the predicate contains different meanings, depending

on the attribute part of the sentence, it needs to be handled together with the attribute part. For instance, looking at the sentence "This cellphone's size is large", and the sentence "This car has a trunk that's large", the verb "is large" can be thought of as negative in the former sentence however, it can be thought of as positive in the latter sentence. Here, the verb "is large" has a dependency on the words "size" and "trunk". As such, in order to determine whether the verb "is large" is positive or negative, the part that has a dependency should be considered together.

Since the word order in the Korean language typically is of a structure in which the predicate appears in the last part of the sentence, a particular approach is needed in order to accurately find the target word that the predicate points. In order to accurately find the sentiment-target pairs, we propose in this article a model that can reflect the characteristics of the syntactical structure of the Korean language. The proposed model has found, in structurally analyzed sentences, the words with the possibility of being sentiment words and the words with a possibility of being target words using statistical data.

For this paper, in Section 2, research related to the sentiment-target pairs model that can reflect the characteristics of the syntactical structure of the Korean language is outlined, and previously developed technologies are explained. In Section 3, the methodology of the proposed model is explained. In Section 4, experimental methods and results are described. In Section 5, the conclusion of this study is presented.

2. Related Works

Sentiment analysis in the Korean language and sentiment analysis in the English language are clearly different. The first difference is in the word order. The existence of different word orders can compound the difficulty of an opinion mining study in which the pacing of the subject, predicate, and object in a random sentence needs to be performed. For general plain English texts, the typical structure is of a structure similar to "Subject + Predicate + Object" where the predicate is in the center, the subject in the front, and the object in the rear. Representing the "Subject + Predicate + Object" structure in phrase units, it can be expressed as "NP + VP + NP" where NP is the noun phrase and VP is the verb phrase. The noun phrase is comprised of two or more words and it refers to the unit acting as the noun. The verb phrase consists of two or more words, and it refers to the unit that serves as the verb. In such a sentence structure, on the basis of the verb phrase, the subject part and the object part can effectively be separated. This is so because each part can be separated effectively with only a morpheme analysis. However, in the case of the Korean language, the structure does not allow for distinguishing the subject part and object part based on the verb phrase. The typical structure of the word order is "Subject phrase + Object phrase + Verb phrase". Representing the corresponding sentence in phrase units, it can be expressed as "NP + NP + VP". In a structure that has the noun phrase appearing nested, it is difficult to distinguish the subject phrase and the object phrase with only a morpheme analysis. This implies that in doing a morpheme analysis in the Korean language, there is a possibility that sequential nouns may appear, as shown in Equation (1). N is the noun, V is the verb, and X is the selection problem of the noun to be used as an attribute:

$$\underbrace{N + N + N}_{x} + N + \cdots + V. \quad (1)$$

As shown in Equation (1), finding the attribute-sentiment pair in a sentence can cause a problem in selecting "up to which part" from the sequential nouns. Regarding the typical prior sentiment mining studies [3,4], there has been a number in which the target word is found by the PMI method or by applying the rule after tagging parts of a speech [5], when the target word is not determined or when it is determined [6,7]. In order to accurately find the parts of a sentence that can be the target word and sentiment word, a statistical model that analyzes the sentence structure and effectively extracts the target-sentiment word pair from the analyzed structure is proposed. In order to find the target words, B. Liu used a pattern in which various commas, periods, semi-colons, hyphens, &, and, but,

etc. appeared in review sentences summarized by users [8,9]. An example of the review sentences is shown in Table 1, as well as the pros and cons of the item in the example. Then in Table 2, we show how the review sentences were analyzed.

Table 1. An example review.

My SLR is on the shelf
by camerafun4. Aug 09'04
Pros: Great photos, easy to use, very small
Cons: Battery usage; included memory is stingy.
I had never used a digital camera prior to purchasing this Canon A70. I have ...
Read the full review

Table 2. The pros in Table 1 can be separated into three segments.

great photos -> ⟨photo⟩
easy to use -> ⟨use⟩
very small -> ⟨small⟩ ⇒ ⟨size⟩

By using the Web-PMI method, Popescu and Etzioni attempted to find the target words. The typical PMI method is the same as in Equation (2). As with the $P(w)$ calculation method, it is used to count the number of documents containing the word (w) in Equation (3). When Equation (4) is substituted into Equation (2) it then becomes Equation (4), which is called the Web-PMI [10].

$$PLM(w_1, w_2) = log \frac{p(w_1, w_2)}{p(w_1)p(w_2)} \tag{2}$$

$$p(w) = \frac{1}{N} hits(w) \tag{3}$$

$$Web - PLM(w_1, w_2) = log \frac{\frac{1}{N} hits(w_1 \wedge w_2)}{\frac{1}{N} hits(w_1) \frac{1}{N} hits(w_2)} \tag{4}$$

w_1 and w_2 are words; w_1 is used as a candidate element for identification and w_2 is used as an identifier. By confirming the co-occurrence information between w_1 and w_2, an attempt was made to determine whether or not w_1 was a target word. The elements of the sentence used as identifiers are a pattern between structured morphemes and elements in WordNet [11,12].

Wen and Wan tried to extract target words by using label sequential rules [13]. Label sequential rules describe the combination that has the possibility of the elements of the sentence being seen when analyzed morphologically. Pertinent rules are applied to find the target words. Kang Han-hoon established a review pattern database for product reviews and used it to extract product attribute-specific positives/negatives. By comparing the positive/negative information extracted from the data (i.e., monitor, laptop, digital camera, and MP3 player), the rate of accuracy was then calculated. The approach was tested using a method that applied rules in morphologically analyzed sentences [6]. Yang Jung-Yeon obtained the product feature words (e.g., battery) and product sentiment words (e.g., short) from product reviews by using the data containing both the product reviews and product scores, via PMI, and determined whether the sentiment words appearing in the product features are positive or negative by using the review scores [14]. Long Jiang undertook classifying Twitter data according to sentiment words nouns extracted through the PMI method and words showing more than the threshold were deemed as one chunk of data. In addition, in order to overcome the difficult parts of the analysis that were caused by frequent occurrences of short sentences, the Tweets deemed as having been posted by one person were considered altogether in the form of a graph. The words processed in this way were classified into positive, negative, or neutral sentiments [7]. All of the applicable models are subject to issues in either the PMI method or in a method that applies rules after morpheme analysis: For the PMI method, words that are either high or low in

frequency; for methods that apply rules, sentences that fall outside the categorized rule statements [15,16]. The issues stem from relying on simple probability information while the syntactic structures of sentences are not identified as of yet, or from trying to express everything with rules that solve the syntactic structures of sentences manually. In this paper, in order to solve problems such as these, a method is used that finds target words by identifying the sentence structure [17].

3. Materials and Methods

The model extracts the sentiment-target word pairs that appear in the sentence by using parsing and statistical methods. The model is comprised of two parts: One that parses sentences in the inputted documents and one that extracts word pairs. The part that parses the sentence structure consists of a morphological analyzer, a speech tagger component, and a syntactic structure analyzer; the other part that extracts word pairs consists of a sentiment word extractor and a target word extractor.

3.1. Sentence Structure Analysis

The sentence structure analysis part of the proposed model is comprised of a parts-of-speech tagger and a parser. First, the parts-of-speech tagging model uses a general probability model similar to Equation (5) where T is the parts-of-speech tagging function of W, M represents the morpheme candidate, T is the parts-of-speech candidate, and W represents words of the sentence:

$$T(W) \stackrel{\text{def}}{=} \underset{M,T}{argmax} P(M, T \vee W). \tag{5}$$

By using the applicable model, the parts-of-speech of neutral words are attached properly [18,19]. What this is referring to is the fact that in the sentence "The sailor dogs the barmaid", the word "dogs" is not used as the frequently used noun form, but a verb form is determined and attaches the appropriate parts-of-speech [20–22]. Similar to the parts-of-speech tagging, the parser model also commonly uses a Probabilistic Context-Free Grammar model [23–25]. The Probabilistic Context-Free Grammar model can be expressed as shown in Equation (6). T_{best} is a function that selects the syntax structure with the highest generation probability from the syntax structure trees, T represents words that comprise the parse tree, G is the grammar rules, t is the sentence, $rule_i$ is the ith grammar rule in the parse tree, and h_i is the history of the appearance of the ith grammar rule:

$$T_{best}(G, T_{1n}) = \underset{T}{argmax} P(T|G, t_{1,n}) = \underset{T}{argmax} \prod_i P(rule_i \vee G, t_{1,n} h_i). \tag{6}$$

3.2. Extraction of Word Pairs

For input sentences that are of the parse tree [26], the extraction of the sentiment word and target word is done in two kinds of processes. The extraction of the sentiment word is to find the verb or adjective that applies to the top-level node verb phrase in the tree that is being analyzed at the sentence structure analysis level; the extraction of the target word is finding the noun word of the noun phrase that is the most dependent with the found sentiment word. This can be represented with an equation as shown in Equation (7):

$$WordPair(W) = \underset{S,A}{argmax} P(S, A|T). \tag{7}$$

In Equation (7), S represents the sentiment words and A is the target words, and T refers to the parse tree. Eventually, the extraction of the pair words refers to finding S and A, which have the highest probability values for the sentiment words S and target words A that are seen in the phrase-analyzed parse tree T. Equation (7) can be expressed

as Equation (8), in which each of the needed elements is calculated by using Equation (9) and Equation (10):

$$\underset{S,A}{argmax} P(S, A \vee T) = \underset{S,A}{argmax} P(S|T) P(A \vee S, T) \tag{8}$$

$$P(S|T) = \underset{i}{argmax} P(d_i \vee node_{1,n}) \tag{9}$$

$$P(A \vee s_i, T) = \underset{i}{argmax} P(ds_i \vee s_i, w_{d_{1,n}}, w_{co_{1,n}}, p_{co_{1,n}}). \tag{10}$$

The w_d is the distance information apart from the selected sentiment word; w_{co} is the probability information for words that can appear together with the sentiment word; and p_{co} is the probability information for the parts-of-speech that can appear together with the sentiment word. ds are the dependency strength that is calculated into w_d, w_{co}, p_{co} [27,28]. The detailed calculation process of Equation (7) is as follows.

The rule of information from the parse tree.

- Step 1: Select the node with the highest dependence in the parse tree (generally, root node).
- Step 2: Register as target word candidates the two nodes close to the root node.
- Step 3: Calculate the distance information with the selected node word from the two candidates, co-concurrence information of words, and co-concurrence information of parts-of-speech.
- Step 4: Select the candidate with higher calculated dependency strength from the two candidates.
- Step 5: Extract the selected two nodes into sentiment word and target word.
- Step 6: If the parse tree can be turned into a sub-tree, then proceed to turn it into a sub-tree and repeat the above steps.

The distance between words is measured by how close the root node word and candidate words are in the input sentence. The distances between the selected root node and the candidate words w_1 and w_2 are calculated, and by interchanging the calculated values, how close each word is to the root node is then quantified. For example, if the first candidate word is the one-word distance from the root node and the second candidate word is three-word distances away from the root node word, then the first candidate word has a value of 3 and the second candidate word has a value of 1. The corpus dependent pattern is comprised of the word co-occurrence frequency and POS co-occurrence frequency [29]. Co-occurrence frequency means the value of the measure of frequency in which the root node word and a candidate word appear together, whereas the POS co-occurrence frequency is the value of the measure of frequency in which the POS of the root node word and the POS of a candidate word appear together. Finally, the measurement is made via "Dependency Strength = (word position) × (word co-occurrence) × (POS co-occurrence)", and as a measured value, the target word is selected from the candidates. Table 3 shows the extraction of the sentiment word and target word.

Table 3. Extraction of sentiment-target word.

Korean POS	English POS
Mina-ga	Mina: personal pronoun, ga: a subjective case
Ye-peun	pretty: adjective
In-hyeong-ui	In-hyeong: doll, ui: a noun modifier
Jip-eul	Jip: house, eul: an objective case
Sat-da	Sat: buy, da: a finishing final ending

Figure 1 shows the parsing results of the input sentence and shows a parse tree that expresses the example sentence in the nodes of verb phrase (VP) and noun phrase (NP).

The target-sentiment word pairs that are possible to be extracted from Figure 1 are shown in Table 4.

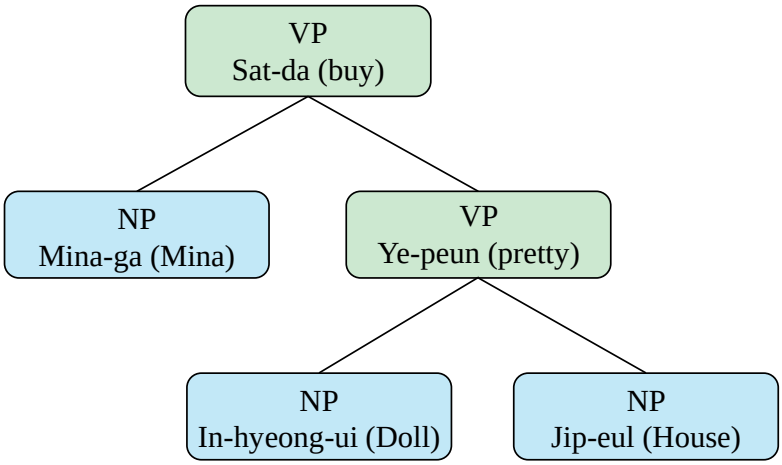

Figure 1. Diagram of the analysis sentence, "Mina buy a pretty doll house".

Table 4. Extractable sentiment-target word pair.

Sentiment Word	Target Word
Buy	House
Buy	Doll
Buy	Mina
Pretty	House
Pretty	Doll
Pretty	Mina

The node with the highest dependency in the parse tree in Figure 1 is the node after the word "buy". The words "Mina" and "pretty" located in the lower part of the node selected after the applicable node are selected as candidates for the target words. The dependency strength of the selected candidate words is calculated as follows:

- Dependency strength ("Mina") = 3 × 2000 × 100,000;
- Dependency strength ("pretty") = 4 × 4000 × 5000.

For the distance calculation, "Mina" gets a value of 4 and "pretty" gets a value of 3, and by interchanging the corresponding distance values, the degree of closeness of each of the words and the selected root node is measured. The calculation for the word co-occurrence frequency uses the statistics appearing in the corpus, the pronoun "Mina" is relatively lacking in the occurrence frequency as compared to the word "pretty" being used as an adjective. Calculation of POS co-occurrence frequency, as expected, uses the statistics appearing in the corpus, and the noun-verb combinations appear much more than the verb-verb combinations. Through the calculations of the applicable numerical values, the word "bought" is extracted as the sentiment word and the word "Mina" is extracted as the target word. By calculating with the same method, the sub-tree in which "pretty" is the root node in such ways, the word pair of "pretty" and "house" is obtained.

4. Experiment And Results

4.1. Evaluation Metric

To evaluate the effectiveness of our proposed method, we use the F1-score, which is to evaluate the effectiveness for evaluation. The F1-score is a harmonic value of precision (P) and recall (R) as a standard indicator to compensate for the shortcomings caused by using only accuracy for evaluation. It is calculated by Equation (11), where E_p represents the set of

predicted correct answers, E_r denotes the ground-truth answer collection, and $C = E_p \bigcap E_r$ are the correct answers:

$$P = \frac{|C|}{|E_p|}, R = \frac{|C|}{|E_r|}, F1 = (\frac{R^{-1} + P^{-1}}{2})^{-1} = 2 \cdot \frac{P \cdot R}{P + R}. \quad (11)$$

4.2. Experimental Results

The data used in the experiment consisted of 4000 movie reviews written in Korean, and for comparison with other studies tested in English. Functional words that appear in the Korean language were removed and comparative experiments were conducted. For the experiments, the performances were compared using a method that measures accuracy and recall rates.

In order to conduct a comparative experiment, using the same data, the method proposed in Long Jiang's model was implemented [30]. Table 5 shows the accuracy and recall rates of Long Jiang's model and the model proposed in this study. For the experiment, a comparative experiment was conducted with the proposed model, of which, the measured results are shown in the table below.

Table 5. Results of performance comparison of the two models. For more precise results, we employ the metric of F1-score with precision and recall.

Models	Accuracy	Recall	Precision	F1-Score (%)
Long Jiang's Model	78.80	76.27	81.89	78.98
Proposed Model	**93.25 (+14.45)**	75.92 (−0.35)	89.84 (+7.95)	**82.29 (+3.31)**

Results of significant improvements were seen in the proposed model as compared to the existing model. The recall rates were about the same. The reason why the accuracy increased greatly from the existing model seems to be attributable to the fact that the analysis took into consideration the structure of the sentence. However due to a large number of calculations, the execution speed falls slightly compared to the existing models. An example of a specific miss-analysis of data actually analyzed is shown in Table 6. In the analysis results, the part in front of the symbol "-" is the property and the rear part is sentiment. Although the first miss-analysis result is a case finding, only the most representative pair in a sentence, in this case, the system extracts other properties and sentiments in addition to finding the most representative pair. The case of the second miss-analysis result is a case in which the results are not extracted due to the fact that the noun in the verb part does not show a dependency relationship in the syntax structure analysis results, where the verb phrase is in the front and the noun is in the back.

Table 6. Incorrect analysis example.

	Example 1	Example 2
Original Sentence	Of the recent movies seen, this is most fun	Movie that doesn't quite satisfy
Accurate Analysis	Movie-is fun	Movie-not satisfying
System Analysis	Movie-is fun	-

5. Conclusions

Since the word order in the Korean language typically is of a structure in which the predicate appears in the last part of the sentence, a particular approach is needed in order to accurately find the target word that the predicate points to. In order to accurately find the sentiment-target pairs, we proposed in this article a model that can reflect the characteristics of the syntactical structure of the Korean language. The proposed model found,

in structurally analyzed sentences, the words with the possibility of being sentiment words and the words with a possibility of being target words using statistical data. Experimental results show a 93.25 (+14.45) accuracy and 82.29 (+3.31) F1-score, as compared to the test set.

However, due to a large number of calculations, the parts of the model that tax computational resources would need to be improved. In addition, the recall rate was not improved over the rate that is achieved by other models. These two shortcomings suggest a need for further studies. We also chose corpora with very different structure styles (such as the Korean language) for the experimental setting, and more extensive evaluations will be required to confirm that the presented results are applicable across domains. Therefore, we are conducting research to show the strength of generalized processing by adding linguistic rules that consider Korean characteristics and use deep learning technology.

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

Funding: This work was supported by Institute of Information & Communications Technology Planning & Evaluation (IITP), grant funded by the Korean government (MSIT) (no. 2020-0-00368, A Neural-Symbolic Model for Knowledge Acquisition and Inference Techniques).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

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

References

1. Lipsman, A. *Online Consumer-Generated Reviews Have Significant Impact on Offline Purchase Behavior*; Comscore, Inc. Industry Analysis: Reston, VA, USA, 2007; pp. 2–28.
2. Horrigan, J.B. *Internet Users Like the Convenience but Worry About the Security of Their Financial Information*; PEW Internet & American Life Project: Washington, DC, USA, 2008; pp. 1–32.
3. Bakshi, R.K.; Kaur, N.; Kaur, R.; Kaur, G. Opinion mining and sentiment analysis. In Proceedings of the 2016 3rd International Conference on Computing for Sustainable Global Development (INDIACom), New Delhi, India, 16–18 March 2016; pp. 452–455.
4. Devika, M.; Sunitha, C.; Ganesh, A. Sentiment analysis: A comparative study on different approaches. *Procedia Comput. Sci.* **2016**, *87*, 44–49. [CrossRef]
5. Sadredini, E.; Guo, D.; Bo, C.; Rahimi, R.; Skadron, K.; Wang, H. A scalable solution for rule-based part-of-speech tagging on novel hardware accelerators. In Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining, London, UK, 19–23 August 2018; pp. 665–674.
6. Kang, H.; Yoo, S.; Han, D. Design and Implementation of System for Classifying Review of Product Attribute to Positive/Negative. In Proceedings of the 36th KIISE Fall Conference, Boston, MA, USA, 2–7 August 2009; Volume 36, pp. 1–6.
7. Jiang, L.; Yu, M.; Zhou, M.; Liu, X.; Zhao, T. Target-dependent twitter sentiment classification. In Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics: Human Language Technologies, Portland, OR, USA, 19–24 June 2011; pp. 151–160.
8. Liu, B. *Web Data Mining: Exploring Hyperlinks, Contents, and Usage Data*; Springer: Berlin/Heidelberg, Germany, 2011; Volume 1.
9. Liu, B.; Hu, M.; Cheng, J. Opinion observer: Analyzing and comparing opinions on the web. In Proceedings of the 14th International Conference on World Wide Web, Chiba, Japan, 10–14 May 2005; pp. 342–351.
10. Popescu, A.M.; Etzioni, O. Extracting product features and opinions from reviews. In *Natural Language Processing and Text Mining*; Springer: Berlin/Heidelberg, Germany, 2007; pp. 9–28.
11. Smith, G.G.; Haworth, R.; Žitnik, S. Computer science meets education: Natural language processing for automatic grading of open-ended questions in ebooks. *J. Educ. Comput. Res.* **2020**, *58*, 1227–1255. [CrossRef]
12. Emadi, M.; Rahgozar, M. Twitter sentiment analysis using fuzzy integral classifier fusion. *J. Inf. Sci.* **2020**, *46*, 226–242. [CrossRef]
13. Wen, S.; Wan, X. Emotion classification in microblog texts using class sequential rules. In Proceedings of the Twenty-Eighth AAAI Conference on Artificial Intelligence, Québec City, QC Canada, 27–31 July 2014.
14. Yang, J.y.; Myung, J.S.; Lee, S.G. A sentiment classification method using context information in product review summarization. *J. KIISE Databases* **2009**, *36*, 254–262.

15. Joshi, A.; Prabhu, A.; Shrivastava, M.; Varma, V. Towards sub-word level compositions for sentiment analysis of hindi-english code mixed text. In Proceedings of the Coling 2016, the 26th International Conference on Computational Linguistics: Technical Papers, Osaka, Japan, 11–16 December 2016; pp. 2482–2491.
16. Van De Mieroop, D.; Miglbauer, M.; Chatterjee, A. Mobilizing master narratives through categorical narratives and categorical statements when default identities are at stake. *Discourse Commun.* **2017**, *11*, 179–198. [CrossRef]
17. Lee, S.; Potter, R.F. The impact of emotional words on listeners' emotional and cognitive responses in the context of advertisements. *Commun. Res.* **2020**, *47*, 1155–1180. [CrossRef]
18. Jindal, N.; Liu, B. Mining comparative sentences and relations. *AAAI* **2006**, *22*, 9.
19. Zhao, N.; Wu, M.; Chen, J. Android-based mobile educational platform for speech signal processing. *Int. J. Electr. Eng. Educ.* **2017**, *54*, 3–16. [CrossRef]
20. Farooq, U.; Mansoor, H.; Nongaillard, A.; Ouzrout, Y.; Qadir, M.A. Negation Handling in Sentiment Analysis at Sentence Level. *J. Comput.* **2017**, *12*, 470–478. [CrossRef]
21. Kim, J.D.; Lim, H.S.; Rim, H.C. Twoply hidden Markov model: A Korean POS tagging model based on morpheme-unit with Eojeol-unit context. In Proceedings of the 1997 International Conference on Computer Processing of Oriental Languages, Ulm, Germany, 18–20 August 1997; pp. 144–148.
22. DeRose, S.J. *Stochastic Methods for Resolution of Grammatical Category Ambiguity in Inflected and Uninflected Languages*; Brown University: Providence, RI, USA, 1989.
23. Kim, Y.; Dyer, C.; Rush, A.M. Compound probabilistic context-free grammars for grammar induction. *arXiv* **2019**, arXiv:1906.10225.
24. Raghavan, S.; Kovashka, A.; Mooney, R. Authorship attribution using probabilistic context-free grammars. In Proceedings of the ACL 2010 Conference Short Papers, Uppsala, Sweden, 11–16 July 2010; pp. 38–42.
25. Charniak, E. Immediate-head parsing for language models. In Proceedings of the 39th Annual Meeting of the Association for Computational Linguistics, Toulouse, France, 6–11 July 2001; pp. 124–131.
26. Massung, S.; Zhai, C.; Hockenmaier, J. Structural parse tree features for text representation. In Proceedings of the 2013 IEEE Seventh International Conference on Semantic Computing, Irvine, CA, USA, 16–18 September 2013; pp. 9–16.
27. Liu, Z.; Chen, G. Remote sensing image landmark segmentation algorithm based on improved GSA and PCNN combination. *Int. J. Electr. Eng. Educ.* **2020**,
28. Yu, S.E.; Kim, D. Landmark vectors with quantized distance information for homing navigation. *Adapt. Behav.* **2011**, *19*, 121–141. [CrossRef]
29. Brunellière, A.; Perre, L.; Tran, T.; Bonnotte, I. Co-occurrence frequency evaluated with large language corpora boosts semantic priming effects. *Q. J. Exp. Psychol.* **2017**, *70*, 1922–1934. [CrossRef] [PubMed]
30. Li, Y.; Jin, Q.; Zuo, M.; Li, H.; Yang, X.; Zhang, Q.; Liu, X. Multi-neural network-based sentiment analysis of food reviews based on character and word embeddings. *Int. J. Electr. Eng. Educ.* **2020**.