KNDTE: A System for Textual Entailment and Fact Validation Tasks at the NTCIR-11 RITE-VAL

Tao-Hsing Chang
Dept. of Computer Science and Information Engineering, National Kaohsiung University of Applied Sciences
changth@kuas.edu.tw

Shih-Kuei Kao
Dept. of Computer Science and Information Engineering, National Kaohsiung University of Applied Sciences
1102308114@kuas.edu.tw

Yen-Wen Lin
Dept. of Computer Science and Information Engineering, National Kaohsiung University of Applied Sciences
1100108132@kuas.edu.tw

Li-Jia Wei
Dept. of Computer Science and Information Engineering, National Kaohsiung University of Applied Sciences
1100108152@kuas.edu.tw

Chung-Yin Tsai
Center of Learning Technology for Chinese, National Taiwan Normal University
cytsai@ntnu.edu.tw

ABSTRACT
Using a decision tree as the predictive model, two algorithms are proposed in this paper: one for detecting textual entailment relationships, and the other for confirming textual fact. Features proposed by previous studies are improved upon, and new features are introduced. This enhances the effectiveness of the proposed method, which was tested using the RITE-VAL task of NTCIR 11. The proposed method predicts entailment relationships between Chinese textual pairs in the data set of the system validation (SV) subtask. The predictions are then verified for accuracy using sentences from the data set of the fact validation (FV) subtask. When applied to the binary-class (BC) and multi-class (MC) tasks of the SV subtasks, the average macro-F1 rates of the proposed method are 54.1% and 39.08%, respectively. For the FV subtask, the average macro-F1 rate of the proposed method is 33.97%.

Team Name
KUAS

SubTasks
RITE-VAL(CN): Fact Validation, System Validation

Keywords
Textual entailment, fact validation, decision tree, linguistic feature, RITE-VAL

1. INTRODUCTION
The basic task of many studies and applications involves the use of textual entailment techniques. Using the issue of textual fact as illustration, given that Internet users can receive a significant amount of information from such source daily, it is important for users to be able to determine and distinguish between factual and contradictory information. Textual entailment techniques can be applied to manage this issue. These techniques are also currently applied to various fields, including question answering systems [2], information extraction [3], and machine translation [4]. Many methods for the detection of entailment relationships have already been proposed in previous studies [5-11,13,14]. However, the existence of linguistic differences means that the detection method for one language does not necessarily apply to another. Furthermore, the effectiveness of detection also varies between texts of different languages.

Detecting textual entailment in the Chinese language is relatively more difficult compared to such detection in the English language for two main reasons: (a) Chinese characters in sentences are not separated by spaces, and (b) there is a relative lack of linguistic tools for analyzing the Chinese language; moreover, those tools that do exist have low effectiveness. Because of the first reason, performing word segmentation and part-of-speech (POS) tagging is required prior to the determination of textual entailment relationships. However, any errors made during pre-steps can affect detection effectiveness. On the other hand, the limitations that arise from the second reason inhibit the accurate calculation of the linguistic features used to detect entailment relationships. Hence, finding the means to improve effectiveness when detecting Chinese textual entailment relationships has become an important research topic.

Most detection methods for Chinese textual entailment [10,11,13,14] use predictive models based on linguistic features, followed by model training. The trained model is then used to detect Chinese textual entailment relationships. There is considerable diversity in the selection of features used for these studies, whereas a support vector machine (SVM) and decision tree are used mainly for a predictive model.

Although many studies [10,11,13,14] used SVM as the predictive model, such a model faces three main issues. First, SVM explores overall model optimization, so there is no way of showing the predictive ability of individual features or the applicable data types. Second, previous research found that SVM is quite susceptible to the influence of the training data set. Consequently, the predictive ability of SVM for partial entailment relationships is significantly weaker when test data are used. Third, SVM optimization requires the trial of different parameters. However, it is extremely difficult to analyze interaction between parameters and features. Hence, when new features is added to train a new model, it is difficult to predict the response and performance of the new model.

Unlike SVM, analyzing the response of features when using a decision tree is much easier. The impact of newly introduced features can also be observed more effectively with a decision tree than with SVM. Findings in previous study [1] indicate that the difference between SVM effectiveness and decision tree post-optimization is limited. Hence, the study by Chang et al. [1] is used as the basis for our study. A decision tree is similarly adopted as the predictive model, although further refinements are made to the computational method for features. New features are also introduced to enhance the effectiveness of predicting Chinese textual entailment relationships.

To consider the issue of confirming textual fact, this paper further propose an algorithm based on a predictive method for textual entailment relationships. Textual fact is defined as follows: for any given a sentence *s* and a large corpus (such as Wikipedia), *s* is tagged as entailment (E) if it can be deduced from the information in the database, or as contradiction (C) if it conflicts with the available information. If the authenticity of *s* cannot be determined because of insufficient information, it is tagged as unknown (U).

The rest of this paper is organized as follows: previous related studies and methods are explored in the next section; an elaboration of both improved and new features is done in Section 3, with an introduction to the structure of the decision tree used as the predictive model; Section 4 introduces the proposed algorithms to address the issue of confirming facts; the results obtained from the application of our proposed method to the RITE-VAL data sets of NTCIR 11 are presented in Section 5; finally, a discussion is made of the proposed method and experimental findings in the last section.

2. RELATED WORKS

There is substantial research on the detection of English textual entailment relationships. Androutsopolos and Malakasiotis [5] believed that lexical similarity between sentences can be used to determine textual entailment relationships, although they could not consider the issue of synonyms. To improve such limitation, Bos and Market [6] proposed the deep semantic analysis (DSA) technique, which uses WordNet to determine whether words and phrases contained within a set of information have similar or opposite meanings. To illustrate, there is relevance between the words "賣家" (sellers), "顧客" (customers), and "消費" (consume), where the third word can be deduced from the second, whereas the semantics of the first two words are opposite.

Another common method used for the detection of English textual entailment is the analysis of parsing tree [7–9]. Carbrio et al. [7] proposed a distance-based algorithm to calculate the difference between the parsing trees of sentences in a textual pair, and then used that difference to determine the entailment relationship of the textual pair. After analyzing two sentences, study [8] and [9] took the parsing tree of one sentence and transformed it to that of the other sentence through insertion, deletion, replacement, and other steps. In turn, the number of steps required to complete the transformation process determined the degree of difference between the two sentences, and following from that, their textual entailment relationship.

Previous studies [10,11,13–17] also proposed various deduction methods for Chinese textual entailment, some of which are similar to those for English textual entailment mentioned above. For example, Huang et al. [11] used the Stanford Parser [12] to analyze the syntax trees of sentences in textual pairs, and tagged the verbs and nouns. Four main features were obtained from the analysis of these verbs and nouns, which were then used to determine textual entailment relationships. There are greater difficulties involved in the grammatical analysis of Chinese sentences and in the extraction of Chinese synonyms, compared to English. Consequently, experimental results from the use of the aforementioned methods indicated that their degree of effectiveness was lower when applied to Chinese text, compared to their analysis of English text.

Currently, most deduction methods for Chinese textual entailment work by extracting linguistic features from the text. These features are then used as input for classification models to predict entailment relationships. For example, Han and Ku [10] used features such as degree of similarity and length to determine the textual entailment relationship between sentences a textual pair. Shih et al. [13] chose other features found in the text, including named entities, Chinese tokens, dependency word, and sentence length. Lin and Tu [14] used 20 features for training before using SVM for determination. Our proposed method similarly employs feature framework and a predictive model to predict textual entailment relationships.

Textual fact is one of the important applications to deduce textual entailment relationships. Hsu [18] used textual entailment relationships to examine whether contradictions exist among the sentences of different subjects within Wikipedia. This paper also applies features commonly used to deduce textual entailment to determine whether the authenticity of sentences could be confirmed in Wikipedia.

3. DETER MINING TEXTUAL ENTAILMENT RELATIONSHIPS

The entailment relationship of a textual pair can be one of the following: forward (F), bidirection (B), contradiction (C), or independence (I) [1]. The structure of the decision tree that we used to predict textual entailment relationships is shown in Figure 1. Six features of textural pairs as proposed by Chang et al. [1] are adopted and three of six features are improved by this paper. These features are used as the basis for the proposed decision tree. We further introduced another four features. These two categories of features are elaborated in Sections 3.1 and 3.2, respectively.

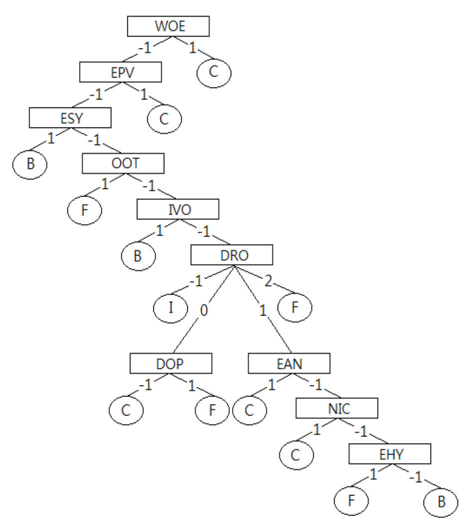

Figure 1 Proposed decision tree for predicting textual entailment relationships

3.1 Improved features

3.1.1 Existence of privatives(EPV)

This feature, as defined by [1], is as follows: if a privative appears in one sentence of the textual pair but not in the other, the value for this feature is 1. It is -1 for the opposite situation. The characteristic of this feature is the determination of the existence of a privative, which causes a conflict between the two sentences in terms of their description of an event. In turn, this leads to a contradictory relationship between the textual pair.

There are two limitations to the original feature. First, for it to work effectively, the premise is that there must be similarity in information between the two sentences. If the difference in information is significant, using this feature to determine entailment relationships directly is not logical. Next, there is a shortage of privatives contained within the vocabulary used originally, with some privatives not even included. To increase the effectiveness of determining such sentences, this paper expanded the number of privatives in the vocabulary.

In addition, this paper further broadened the definition of this feature as follows: if the value for feature “consistency in nouns” of a textual pair is 1, and a privative appears in one sentence but not the other, the value for this feature is 1; otherwise, it is -1. Please refer to [1] for the definition of the feature “consistency in nouns.”

3.1.2 Difference in repetition rates of words (DRO)

Chang et al. [1] used this feature to indicate that the amount of information between two sentences is similar. A lack of similarity indicates a significant discrepancy in the semantics being expressed, although a case of independence could be likely. If two sentences have similar information, but the amount of information provided by one is greater than the other, a forward relationship between the textual pair could exist.

During our study, we observed that textual pairs, of which this feature could not deduce the entailment relationship, could be subdivided into two categories. For the first category, there is great similarity in terms of information, but only minute difference. The probability that the entailment relationships of these pairs are either bidirectional or contradictory is extremely high, with only a few cases being forward. Thus, we expanded and revised the definition of this feature as follows:

$$DRO = \begin{cases} 2, if\,(RWF \geq TI \text{ or } RWB \geq TI) \text{ and } |RWB - RWF| \geq TQ \\ 1, if\,(RWF \geq TE \text{ or } RWB \geq TE) \text{ and } (|RWF - RWB| \leq TD) \\ 0, if\,(RWF \geq TI \text{ or } RWB \geq TI) \text{ and } (DRO \neq 2 \text{ and } DRO \neq 1) \\ -1, otherwise \end{cases}$$

where the definition of RWF and RWB is the same as [1], whereas TI, TE, TQ, and TD are threshold with the value of TE ≥ TI.

3.1.3 Difference in repetition rates of POS (DOP)

The principle of this feature is extremely similar to the above-mentioned feature DRO. Originally, the issue of calculation order for RPF and RPB, two parameters used by this feature, was not considered. We reviewed this and amended the definition as follows:

$$DOP = \begin{cases} 1, if (RPF \ \geq TP \text{ or } RPB \ \geq TP) \text{ and } |RPF - RPB| \ \geq TK \\ -1, otherwise \end{cases}$$

where the definition of *RWF* and *RWB* is the same as [1], whereas *TP* and *TK* are threshold values.

3.2 New features introduced

3.2.1 Inconsistency in voice (IVO)

By observing textual pairs in the training data, we found that for some bidirectional textual pairs, the difference between two sentences is that of voice. For example, the majority of the vocabulary that appears in both sentences of the following textual pair (S1, S2) is similar, while S1 and S2 are in the active and passive voices, respectively.

S1: 歷史上沒有吉力馬札羅山火山噴發的記錄。

(There is no historical record of any Mount Kilimanjaro eruptions.)

S2: 歷史上吉力馬札羅山火山的噴發沒有被記錄。

(Historically, Mount Kilimanjaro eruptions were not recorded.)

Because voice can affect the accuracy of algorithms for deducing entailment relationships, this paper introduced “inconsistency in voice” (IVO) as a new feature, which is defined as follows: if the value for feature “consistency in nouns” of a textual pair is 1, with one sentence containing the character “被” (by, a passive indicator) but not the other, the value for this feature is 1. For the opposite situation, the value is -1.

3.2.2 Existence of antonyms (EAN)

Textual pairs in our training data indicated that those with a contradictory relationship often contain antonyms. For example, by observing the textual pair (S3, S4), we can determine that the character “高” (high) and “低” (low) in the two respective sentences are antonyms. Antonyms cause the semantics of the two sentences to be opposite, thus making the textual pair contradictory.

S3: 水蘊草原產地是在南美洲氣溫較高的區域，包括巴西的東南部、阿根廷、烏拉圭等地。”

(The egeria densa originated from high-temperature areas in South America, including southeastern Brazil, Argentina, and Uruguay.)

S4: 水蘊草原產地是在南美洲氣溫較低的區域，包括巴西的東南部、阿根廷、烏拉圭等地。”

(The egeria densa originated from low-temperature areas in South America, including southeastern Brazil, Argentina, and Uruguay.)

This feature is defined as follows: if antonyms exist in the two sentences of a textual pair, the value for this feature is 1; otherwise, it is -1. The vocabulary of antonyms used in our study comprises a total of 445 groups of words collected from multiple sources.

3.2.3 Numerical inconsistencies (NIC)

Observing our training data reveals that contradictory textual pairs often contain numerical inconsistencies when describing facts. It is obvious from the textual pair (S5, S6) that, although both sentences describe the same event, the latitudes stated are different (“3°” versus “300°”). Based on this phenomenon, it can be deduced that this textual pair is contradictory.

S5: 吉力馬札羅山位於赤道與南緯 3 度之間，在東非大裂谷以東 160 千米處。”

(Mount Kilimanjaro is located between the equator and latitude 3° south, and 160 kilometers east of the Great Rift Valley.)

S6: 吉力馬札羅山位於赤道與南緯 300 度之間，在東非大裂谷以東 160 千米處。”

(Mount Kilimanjaro is located between the equator and latitude 300° south, and 160 kilometers east of the Great Rift Valley.)

In this study, this feature is defined as follows: if the numerals that appear in the same position for both sentences are inconsistent, the value is 1; otherwise, it is -1.

3.2.4 Existence of hyponyms (EHY)

We found that some textual pairs with a forward relationship have a phenomenon similar to that of numerical inconsistencies. Basically, the majority of the vocabulary for both sentences is similar, but different words are used in a minor number of positions. These words are neither synonyms nor antonyms, but one is semantically a hyponym of the other. The textual pair (S7, S8) contains similar words, except for “廣東” (Guangdong) and “中國” (China). In terms of the hierarchy of knowledge, the former is a hyponym of the latter. Thus, it can be concluded that semantically, the textual pair (S7, S8) has a forward relationship.

S7: 乾炒牛河是廣東菜色的一種。

(Stir-fried beef noodles is a Cantonese cuisine.)

S8: 乾炒牛河是中國菜色的一種。

(Stir-fried beef noodles is a Chinese cuisine.)

This feature is defined as follows: for different words in two sentences of a textual pair, if one word is a hyponym of the other, the value for this feature is 1. If that is not the case, the value is -1. For two different words located in the same position in the sentences, their knowledge categories are separately checked using the Wikipedia query function to determine whether one is a hyponym of the other. If their categories both appear under the same upper knowledge category within a limited number of layers, a hyponym exists.

4. CONFIRMATION OF TEXTUAL FACT

For a text *T* and a knowledge database, the issue of confirming textual fact is to determine whether *T* is an entailment (E), contradiction (C), or an unknown (U). Each page of a knowledge database usually represents a particular knowledge. Hence, we assume that if the text is an entailment, there must be a page that provides sufficient information for such determination.

Because the main constituents of textual fact are nouns and verbs, the text could be an entailment if that single page contains sufficient numbers of nouns and verbs appear in the text. If the majority are nouns, but the verbs differ from the text or are antonyms, the text could be a contradiction. If none of the pages contain sufficient nouns and verbs, it is not possible to known whether the text is factual (i.e., it is an unknown).

The following algorithm is proposed based on the various assumptions stated above:

Step 1: Use WECAn [11] to segment the text into words and tag the POS of the words, while retaining all nouns.

Step 2: Collate from the knowledge base all pages that contain those nouns.

Step 3: Tag the text according to the following rules:

Step 3.1: E if any of the pages contains more than two-thirds of its nouns and half of its verbs.

Step 3.2: C if any of the pages contains more than two-thirds of its nouns, but less than half of its verbs.

Step 3.3: U for all remaining situations.

Nouns refer to the POS Na, Nb, and Nc as defined in [20], whereas verbs refer to VA, VB, and VC. Text S9 below is used in this study to illustrate the operation of the algorithm.

S9: 柏拉圖是一位古希臘哲學家。

(Plato was a philosopher from ancient Greece.)

First, three nouns (“Plato,” “ancient Greece,” and “philosopher”) and a verb (“was”) are derived using word segmentation. All pages in the database that contain those three nouns are then collated. Among these, one page contains all the nouns and the verb; hence, the text is tagged as E.

5. EXPERIMENTAL RESULTS

The test data used for this study are from the formal run dataset of the RITE-VAL task recorded during the NTCIR 11 Conference. The RITE-VAL task comprises two subtasks: system validation (SV) and fact validation (FV). The latter is used to verify the effectiveness of our proposed method for confirming textual fact. In turn, SV is subdivided into two other tasks: binary-class (SV-BC) and multi-class (SV-MC). SV-BC and SV-MC are used to verify the effectiveness of our proposed method for determining whether a textual pair has a deductive relationship and textual entailment relationships, respectively. For the FV, SV-BC, and SV-MC subtasks, NTCIR provided the FV-CT, SV-CT-BC, and SV-CT-MC data sets, respectively. These were used to train the model and test the effectiveness of the proposed method. Hence, for each of the three data sets, there are data sets for the two stages. Those for the training and testing stages are known as the dry run and formal run stage data sets, respectively.

When the proposed method is applied to the formal run stage data set of the SV subtask, the results of the predictions are as listed in Table 1. The parameters TI, TE, TQ, TD, TP, and TK mentioned in Section 3 were set at 0.6, 0.85, 0.15, 0.1, 0.7, and 0.15, respectively. For the SV-CT-BC and SV-CT-MC data sets, the average marco-F1 rates for our proposed method are 54.1% and 39.08%, respectively.

Table 1 Data set from formal-run stage of SV subtask: prediction results using proposed method

Task / Indicator	SV-CT-BC		SV-CT-MC			
	Y	**N**	**B**	**C**	**F**	**I**
F1	62.59	45.60	50.50	43.00	49.09	13.74
Precision	54.14	58.99	43.95	42.04	42.21	39.06
Recall	74.17	37.17	59.33	44.00	58.67	8.33

Table 1 demonstrates that the performance of the proposed method to determine relationship I is relatively poor. A possible cause is that the decision tree has only one leaf node under I. Thus, the number of text pairs categorized as I is naturally rather limited. Although C has the most number of leaf nodes in the decision tree, the performance to identify it is even lower than the identification of B and F. This is because in terms of semantics, the textual pairs in C contain many contradictions, making identification difficult.

Table 2 Data set from formal-run stage of FV subtask: prediction results using proposed method

Indicator \ Task	FV-CT		
	E	**C**	**U**
F1	48.99	25.59	27.33
Precision	41.43	26.92	37.27
Recall	59.91	24.38	21.58

When the proposed method is applied to the formal run stage data set of the FV subtask, the results of the predictions are as listed in Table 2. For this data set, the average marco-F1 rate for the proposed method is 33.97%. Similar to the results of the SV subtask, the performance of the proposed method is better when the equivalent is the entailment of B and F in SV. However, its performance is poorer when the equivalents are the contradiction of C in SV and the unknown of I in SV.

6. DISCUSSION AND FURTHER WORKS

The experimental results indicate that, regardless of the SV or FV subtasks, the most difficult issue continues to be the identification of contradiction and independence. To improve the identification, it is necessary to use the concept of the DSA technique proposed by Bos and Market [6] to develop the features of deep-seated meanings in the Chinese language. Only then can the effectiveness of the proposed method in detecting textual entailment relationships be further improved.

It should also be noted that the proposed algorithm for confirming textual fact is still at the preliminary stage of development, and thus, it is limited in its effectiveness. Originally, the deduction of entailment relationships, proposed by this paper, involved the use of two sentences as the unit of identification. As such, the proposed decision tree method cannot be applied directly. However, there is preliminary effectiveness when a small number of features are used. This indicates that the issue of confirming textual fact is worthy of further study. One possible direction is to integrate the algorithm for confirming textual fact with the method for determining entailment relationships.

In this study, the task of optimizing the decision tree was not performed because we would like to observe and test the effectiveness of individual features. Moving forward, we will consider the use of the ID3 algorithm to propose an optimal decision tree. Classification models (including SVM) that are currently commonly used will be adopted as predictive models, so that their effectiveness at identification can be compared. This will facilitate the derivation of mathematical methods that are more suitable for the determination of textual entailment relationships.

7. ACKNOWLEDGEMENTS

This research is partially supported by the "Aim for the Top University Project" and "Center of Learning Technology for Chinese" of National Taiwan Normal University (NTNU), sponsored by the Ministry of Education, Taiwan, R.O.C. and the "International Research-Intensive Center of Excellence Program" of NTNU and Ministry of Science and Technology, Taiwan, R.O.C. under Grant no. NSC 102-2511-S-151-002 and NSC 103-2911-I-003-301.

8. REFERENCES

[1] Chang, T. H., Hsu, Y. C., Chang, C. W., Hsu, Y. C., Chang, J. I. 2013. KC99: A prediction system for Chinese textual entailment relation using decision trees. In Processing of the 10th NTCIR Conference (2013, June), 469-473.

[2] Voorhees, E. M. 1999. The TREC-8 question answering track report. In Proceedings of TREC (Vol. 8, pp. 77-82).

[3] McCallum, A., Freitag, D., & Pereira, F. 2000. Maximum entropy Markov models for information extraction and segmentation. In Proceedings of the 17th International Conference on Machine Learning (Vol. 951, No. 00, pp. 591-598).

[4] Papineni, K., Roukos, S., Ward, T., & Zhu, W. J. 2002. BLEU: a method for automatic evaluation of machine translation. In Proceedings of the 40th Annual Meeting on Association for Computational Linguistics (2002, June), Association for Computational Linguistics, pp. 311-318.

[5] Androutsopoulos, I. and Malakasiotis, P. 2010. A survey of paraphrasing and textual entailment methods. Journal of Artificial Intelligence Research (2010, May), 38:135-187.

[6] Bos, J. and Markert, K. 2005. Recognizing textual entailment with logical inference. In Proceedings of the Conference on Human Language Technology and Empirical Methods in Natural Language Processing, Association for Computational Linguistics (2005, Oct), 628-635.

[7] Cabrio, E., Kouylekov, M., and Magnini, B. 2008. Combining specialized entailment engines for RTE-4. In Proceedings of TAC08, 4th PASCAL Challenges Workshop on Recognizing Textual Entailment (2008).

[8] Kouylekov, M. and Magnini, B. 2005. Recognizing textual entailment with tree edit distance algorithms. In Proceedings of the First Challenge Workshop Recognizing Textual Entailment (2005, Apr), 17-20.

[9] Kouylekov, M. and Magnini, B. 2006. Tree edit distance for recognizing textual entailment: Estimating the cost of insertion. In Proceedings of the PASCAL RTE-2 Challenge (2006, Apr), 68-73.

[10] Han, N. H. and Ku, L. W. 2011. The Yuntech system in NTCIR-9 RITE task. In Proceedings of the NTCIR-9 Workshop (2011, Dec), 345-348.

[11] Huang, H. H., Chang, K. C., Haver II, J. M., and Chen, H. H. 2011. NTU textual entailment system for NTCIR 9 RITE task. In Proceedings of the NTCIR-9 Workshop (2011, Dec), 349-352.

[12] Stanford Parser: A statistical parser. 2002. Available from: http://nlp.stanford.edu/software/lex-parser.shtml.

[13] Shih, C. W., Liu, C., Lee, C. W., Hsu, W. L. 2013. IASL RITE System at NTCIR-10. In Proceedings of the 10th NTCIR Conference (2013, June), 425-429.

[14] Lin, C. J., Tu, Y. C. 2013. The description of the NTOU RITE System in NTCIR-10. In Proceedings of the 10th NTCIR Conference (2013, June), 495-498.

[15] Ku, L. W., Chu, Edward T. H., Han, N. H. 2013. Extracting features for machine learning in NTCIR-10 RITE task. In Proceedings of the 10th NTCIR Conference (2013, June), 457-461.

[16] Day, M. Y., Tu, C., Huang, S. J., Vong, H. C., Wu, S. W. 2013. IMTKU textual entailment system for recognizing inference in text at NTCIR-10 RITE2. In Proceedings of the 10th NTCIR Conference (2013, June), 464-468.

[17] Huang, W. J., Liu, C. L. 2013. NCCU-MIG at NTCIR-10: Using lexical, syntactic, and semantic features for the RITE

tasks. In Proceedings of the 10th NTCIR Conference (2013, June), 430-434.

[18] Hsu, Y. C., Detection of textual contradiction in Wikipedia based on textual entailment, Technical Report NSC 102-2815-C-151-035-E. National Kaohsiung University of Applied Sciences.