IMTKU Textual Entailment System for Recognizing Inference in Text at NTCIR-9 RITE

Min-Yuh Day [1, *], Re-Yuan Lee [1, 2], Cheng-Tai Liu [2], Chun Tu[1] , Chin-Sheng Tseng [1], Loong Tern Yap [1], Allen-Green C.L. Huang [1], Yu-Hsuan Chiu [1], Wei-Ze Hong [1], Chien Lun Hsu [1], Hou Cheng Vong [1], Ruo Hua Wu [1], Joon Kah Man Ho [1]

* myday@mail.tku.edu.tw

[1] Department of Information Management, Tamkang University, New Taipei City, Taiwan
[2] Department of Computer Science, National Tsing Hua University, Hsinchu, Taiwan

In this paper, we describe the IMTKU (**I**nformation **M**anagement at **T**am**K**ang **U**niversity) textual entailment system for recognizing inference in text at NTCIR-9 RITE (Recognizing Inference in Text). We proposed a textual entailment system using a hybrid approach that integrate knowledge based and machine learning techniques for recognizing inference in text at NTCIR-9 RITE task. We submitted 3 official runs for both BC and MC subtask. In NTCIR-9 RITE task, IMTKU team achieved 0.522 in the CT-MC subtask and 0.556 in the CT-BC subtask. For RITE4QA subtask, the best MRR of IMTKU team is 0.3998 in CS-RITE4QA and 0.3992 in CT-RITE4QA subtask.
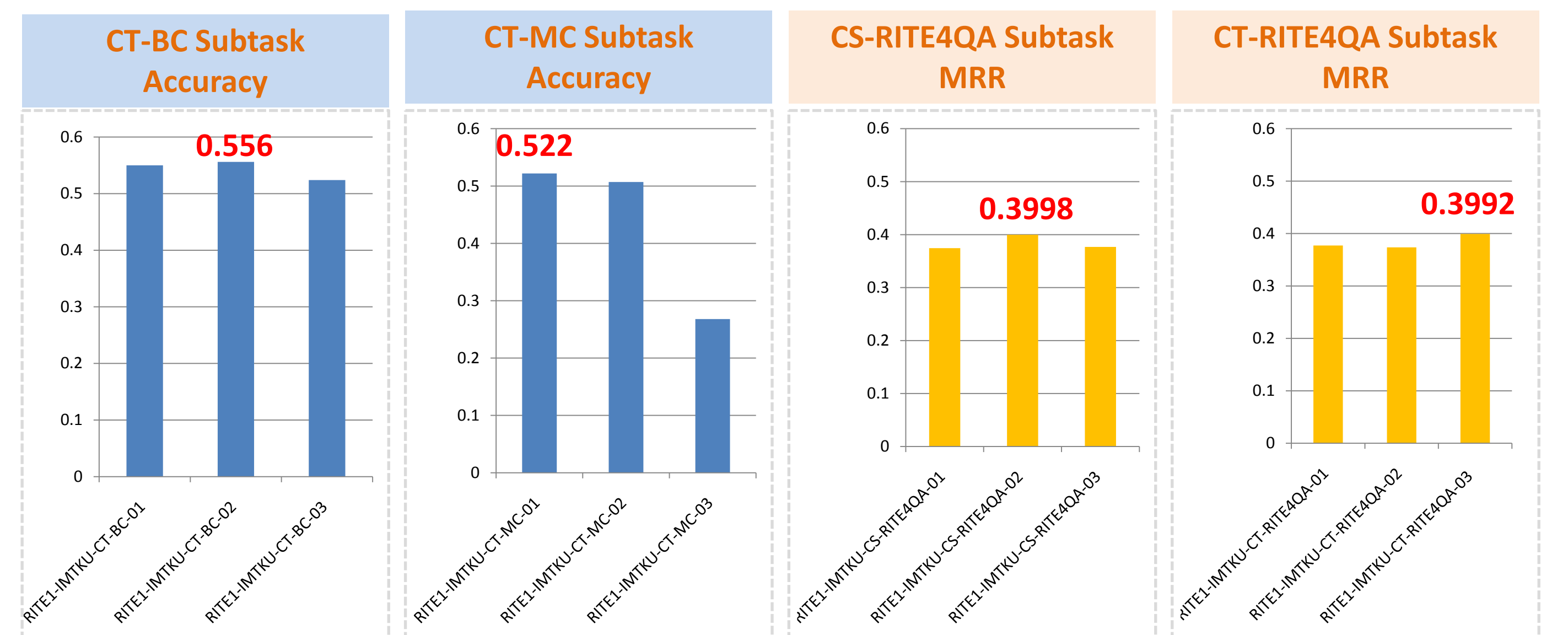
System Architecture
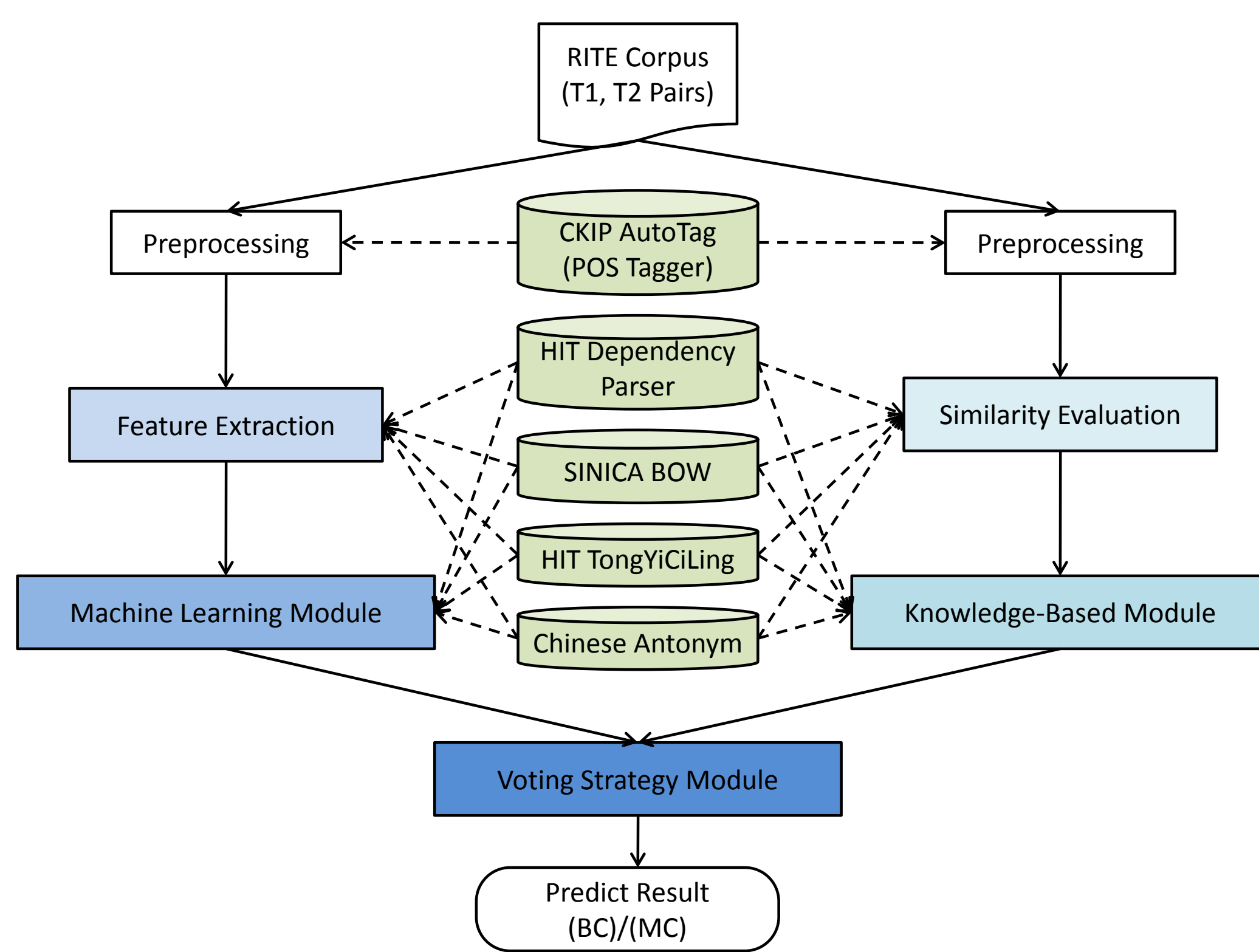
Performance

IMTKU **BC** Subtask Official Runs	Accuracy
RITE1-IMTKU-CT-BC-01	0.550
RITE1-IMTKU-CT-BC-02	0.556
RITE1-IMTKU-CT-BC-03	0.524

IMTKU **MC** Subtask Official Runs	Accuracy
RITE1-IMTKU-CT-MC-01	0.522
RITE1-IMTKU-CT-MC-02	0.507
RITE1-IMTKU-CT-MC-03	0.268

IMTKU **CS RITE4QA** Subtask Official Runs	Accuracy	Top1	MRR
RITE1-IMTKU-CS-RITE4QA-01	0.3319	0.2450	0.3744
RITE1-IMTKU-CS-RITE4QA-02	0.4090	0.2953	0.3998
RITE1-IMTKU-CS-RITE4QA-03	0.4716	0.2550	0.3768

IMTKU **CT RITE4QA** Subtask Official Runs	Accuracy	Top1	MRR
RITE1-IMTKU-CT-RITE4QA-01	0.3246	0.2517	0.3772
RITE1-IMTKU-CT-RITE4QA-02	0.3392	0.2517	0.3736
RITE1-IMTKU-CT-RITE4QA-03	0.4003	0.2953	0.3992

CT-BC Subtask Accuracy — 0.556

CT-MC Subtask Accuracy — 0.522

CS-RITE4QA Subtask MRR — 0.3998

CT-RITE4QA Subtask MRR — 0.3992

Methods for Official Runs

• RITE-IMTKU-CT-BC Subtask

RITE-IMTKU-CT-BC-01
Tools: CKIP AutoTag, HIT Dependency Parser, LibSVM
Resources: Bilingual Wordnet (SINICA BOW), HIT TongYiCiLing (HIT-TYCL)
Method: Hybrid approach (Integrated Knowledge Base and Machine Learning Approach)
Feature Extraction from normalized t1 and t2. Measure similarity match between t1 and t2.
Voting strategy from multiple prediction result.

RITE-IMTKU-CT-BC-02
Tools: CKIP AutoTag, HIT Dependency Parser, LibSVM
Resources: Bilingual Wordnet (SINICA BOW), HIT TongYiCiLing (HIT-TYCL)
Method: Hybrid approach (Integrated Knowledge Base and Machine Learning Approach)
Feature Extraction from normalized t1 and t2. Measure similarity match between t1 and t2.
Multiple Features used (Word Based Similarity, Token Based Similarity, Lexical overlap, Text Pair Length, Token Length) in SVM.

RITE-IMTKU-CT-BC-03
Tools: CKIP AutoTag, LibSVM
Resources: NONE
Method: Machine Learning Approach for NTCIR-9 RITE.
Feature Extraction from normalized t1 and t2.
4 Features used (Word Based Edit Distance, Token Based Edit Distance, Text Length, Identical POS Token Rate) in SVM.

• RITE-IMTKU-CT-MC Subtask

RITE-IMTKU-CT-MC-01
Tools: CKIP AutoTag, HIT Dependency Parser, LibSVM
Resources: Bilingual Wordnet (SINICA BOW), HIT TongYiCiLing (HIT-TYCL)
Method: Hybrid approach (Integrated Knowledge Base and Machine Learning Approach).
Feature Extraction from normalized t1 and t2. Measure similarity match between t1 and t2.
Multiple Features used (Word Based Similarity, Token Based Similarity, Lexical overlap, Text Pair Length, Token Length) in SVM.

RITE-IMTKU-CT-MC-02
Tools: CKIP AutoTag, HIT Dependency Parser, LibSVM
Resources: Bilingual Wordnet (SINICA BOW), HIT TongYiCiLing (HIT-TYCL)
Method: Hybrid approach (Integrated Knowledge Base and Machine Learning Approach).
Feature Extraction from normalized t1 and t2. Measure similarity match between t1 and t2.
4 Features used (Word Based Similarity, Token Based Similarity, Lexical overlap, Text Pair Length) in SVM.

RITE-IMTKU-CT-MC-03
Tools: CKIP AutoTag, LibSVM
Resources: NONE
Method: Machine Learning Approach.
Feature Extraction from normalized t1 and t2.
4 Features used (Word Based Edit Distance, Token Based Edit Distance, Text Length, Identical POS Token Rate) in SVM.

• RITE-IMTKU-CT-RITE4QA Subtask

RITE1-IMTKU-CT-RITE4QA-01
Tools: CKIP AutoTag, HIT Dependency Parser, LibSVM
Resources: Bilingual Wordnet (SINICA BOW), HIT TongYiCiLing (HIT-TYCL)
Method: Hybrid approach (Integrated Knowledge Base and Machine Learning Approach)
Feature Extraction from normalized t1 and t2. Measure similarity match between t1 and t2.
Multiple Features used (Word Based Siilarity, Token Based Similarity, Lexical overlap, Text Pair Length, Token Length) in SVM.
Machine Learning Training Dataset: Dev Dataset (421)

RITE1-IMTKU-CT-RITE4QA-02
Tools: CKIP AutoTag, HIT Dependency Parser, LibSVM
Resources: Bilingual Wordnet (SINICA BOW), HIT TongYiCiLing (HIT-TYCL)
Method: Hybrid approach (Integrated Knowledge Base and Machine Learning Approach)
Feature Extraction from normalized t1 and t2. Measure similarity match between t1 and t2.
Multiple Features used (Word Based Siilarity, Token Based Similarity, Lexical overlap, Text Pair Length, Token Length) in SVM. Machine Learning Training Dataset: Dev Dataset + Test Dataset (1321)

RITE1-IMTKU-CT-RITE4QA-03
Tools: CKIP AutoTag, LibSVM
Resources: NONE
Method: Machine Learning Approach for NTCIR-9 RITE.
Feature Extraction from normalized t1 and t2.
13 Features used (Longest Common Substring, Word Based Edit Distance, T1 Token Based Edit Distance,T2 Token Based Edit Distance,T1 Noun Number-T2 Noun Number, T1 Verb Number-T2 Noun Number,T1 Text Length,T2 Text Length, Text Length Rate,T1 Text Length-T2 Text Length, Text Length Ratio, T1 Token Based Edit Distance-T2 Token Based Edit Distance, Identical POS Token Rate) in SVM.

• RITE-IMTKU-CS-RITE4QA Subtask

RITE1-IMTKU-CS-RITE4QA-01
Tools: CKIP AutoTag, HIT Dependency Parser, LibSVM
Resources: Bilingual Wordnet (SINICA BOW), HIT TongYiCiLing (HIT-TYCL)
Method: Hybrid approach (Integrated Knowledge Base and Machine Learning Approach) for NTCIR-9 RITE.
Feature Extraction from normalized t1 and t2.
Measure similarity match between t1 and t2.
Multiple Features used (Word Based Siilarity, Token Based Similarity, Lexical overlap, Text Pair Length, Token Length) in SVM.

RITE1-IMTKU-CS-RITE4QA-02
Tools: CKIP AutoTag, LibSVM
Resources: NONE
Method: Machine Learning Approach for NTCIR-9 RITE.
Feature Extraction from normalized t1 and t2.
13 Features used (Longest Common Substring, Word Based Edit Distance, T1 Token Based Edit Distance,T2 Token Based Edit Distance,T1 Noun Number-T2 Noun Number, T1 Verb Number-T2 Noun Number,T1 Text Length,T2 Text Length, Text Length Rate,T1 Text Length-T2 Text Length, Text Length Ratio, T1 Token Based Edit Distance-T2 Token Based Edit Distance, Identical POS Token Rate) in SVM.

RITE1-IMTKU-CS-RITE4QA-03
Tools: CKIP AutoTag, LibSVM
Resources: NONE
Method: Machine Learning Approach for NTCIR-9 RITE.
Feature Extraction from normalized t1 and t2.
9 Features used (Word Based Edit Distance, T1 Token Based Edit Distance,T2 Token Based Edit Distance,T1 Noun Number-T2 Noun Number, T1 Verb Number-T2 Noun Number,T1 Text Length, Text Length Ratio, T1 Token Based Edit Distance-T2 Token Based Edit Distance, Identical POS Token Rate) in SVM.) in SVM.

Discussion

Issues of Definition and Data Conversion in RITE MC and BC datasets:
• Definition of MC subtask :"A 5-way labeling subtask to detect (forward / reverse / bidirection) entailment or no entailment (contradiction / independence) in a text pair."
(Incorrect conversion: "F/R/B"→Y, "C/I"→N)
• Definition of BC subtask : "Given a text pair (t1,t2) identify whether t1 entails (infers) a hypothesis t2 or not."
(Correct conversion: "F/B" → Y, "R/C/I" →N).

Cross Validation of Development and Test datasets of RITE CT-BC Subtask

Datasets	10 Fold CV Accuracy
RITE1_CT_dev_bc_g.txt (gold standard) (BC Development Dataset: 421 pairs)	76.48%
RITE1_CT_test_bc_g.txt (BC Test Dataset: 900 pairs)	66.33%
RITE1_CT_dev_test_bc_g.txt (BC Dev+Test Dataset: 421+900 =1321 pairs)	67.67%

DEMO: http://rite.im.tku.edu.tw

Information Management, Tamkang University

Tamkang University
淡江大學