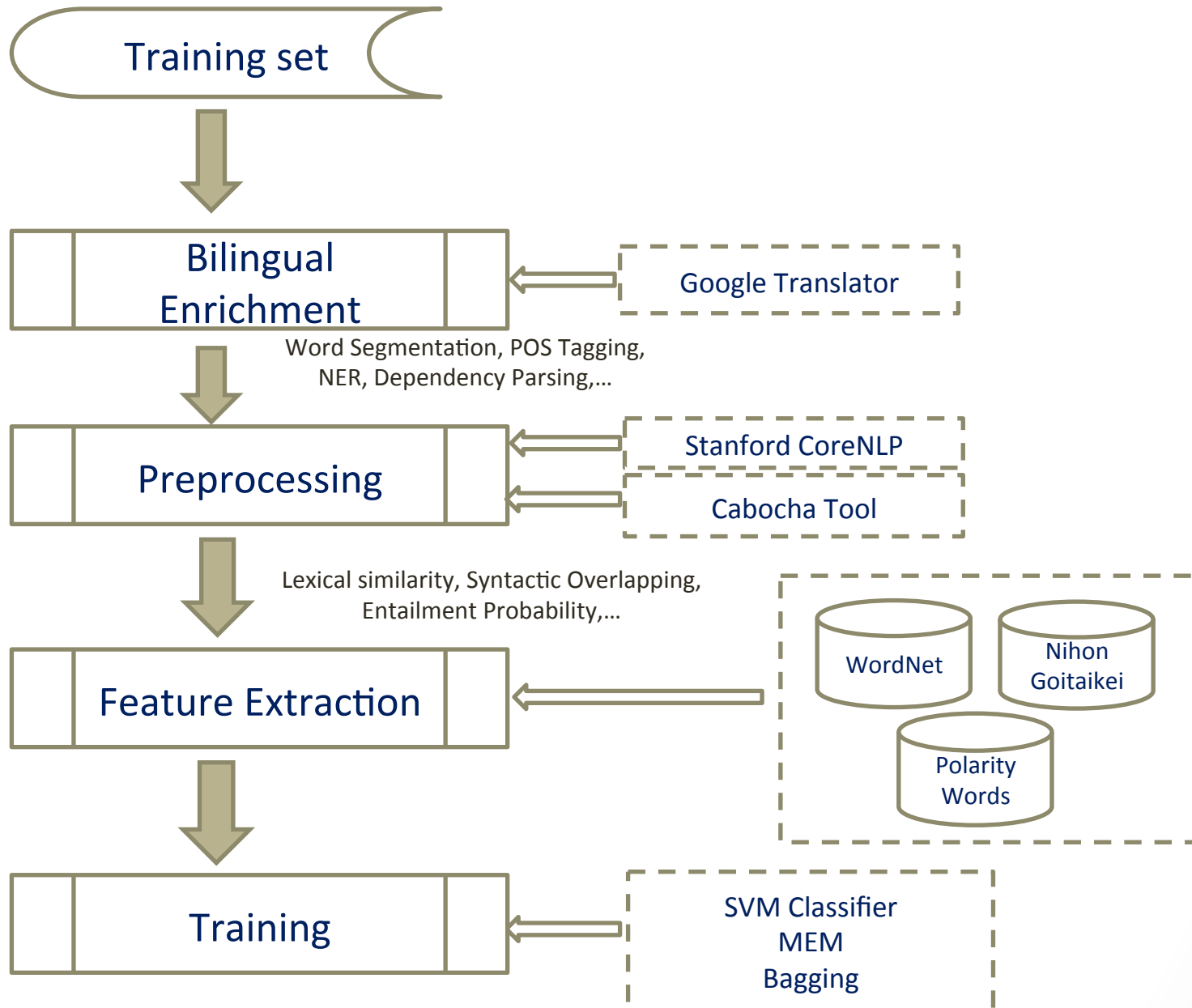


JAIIST Participation at NTCIR-10 RITE-2

Minh Pham, Minh Nguyen, Akira Shimazu

Japan Advanced Institute of Science and Technology

System



JAIST System at NTCIR-10

- Based on the system used at NTCIR-9
 - Machine learning-based system
 - Exploit bilingual features extracted from translation pairs
- Differences
 - Employ ensemble learning method
 - Bagging technique with Ripper rule learners as the based learners.
 - Extract similarity features from character-based representations of the text T and the hypothesis H
 - String-edit distances
 - Longest common subsequence string
 - BLEU measures

Feature Design

Feature	Japanese	English
Word Overlap	x	x
Levenshtein distance	x	x
1-gram, 2-gram, 3-gram based BLEU measures (baseline)	x	x
1-gram, 2-gram, 3-gram based BLEU measures (modified)	x	x
Longest common subsequence string	x	x
Jaccard coefficient	x	x
Dice coefficient	x	x
Manhattan distance	x	x
Jaro-Winkler distance	x	x
Cosine similarity	x	x
Entailment Probability	x	x
Dependency Relation Overlap	x	x
Levenshtein distance based on characters	x	
Longest common subsequence string based on characters	x	
BLEU measures (modifier) based on characters	x	

BC Subtask

Run	Training data	Machine Learning	Features
JAIST-JA-BC-01	Dev set (RITE2_BC)	SVM	mono + bilingual
JAIST-JA-BC-02	Dev set (RITE2_BC)	Bagging	mono +bilingual
JAIST-JA-BC-03	Dev set + test set (RITE1_BC) + Dev set (RITE2_BC)	Bagging	mono

Table 2: Results on BC subtask (JA)

Run	MacroF1	Acc.	Y-F1	Y-Prec.	Y-Rec.	N-F1	N-Prec.	N-Rec.
JAIST-JA-BC-01	75.56	76.23	71.51	71.94	71.09	79.61	79.27	79.94
JAIST-JA-BC-02	76.47	76.89	73.35	71.06	75.78	79.59	81.60	77.68
JAIST-JA-BC-03	73.08	73.77	68.75	68.75	68.75	77.40	77.40	77.40
Baseline	62.53	63.93	55.28	57.63	53.13	69.78	67.91	71.75
1 st -rank system	80.49	81.64	75.76	84.95	68.36	85.22	79.95	91.24

MC Subtask

Run	Training data	Machine Learning	Features
JAIST-JA-MC-01	Dev set (RITE2_BC)	SVM	mono
JAIST-JA-MC-02	Dev set (RITE2_MC) + Dev set (RITE1_MC) + Test set (RITE1_MC)	SVM	mono
JAIST-JA-MC-03	Dev set (RITE2_MC) + Dev set (RITE1_MC) + Test set (RITE1_MC)	SVM	mono + bilingual

Table 3: Results on MC subtask (JA)

Run	MacroF1	Acc.	B-F1	B-Prec.	B-Rec.	F-F1	F-Prec.	F-Rec.	C-F1	C-Prec.	C-Rec.	I-F1	I-Prec.	I-Rec.
MC-01	52.60	66.97	66.67	64.86	68.57	74.55	69.79	80.00	0.00	0.00	0.00	69.20	65.68	73.11
MC-02	52.27	65.33	63.51	60.26	67.14	71.84	59.68	90.24	5.97	33.33	3.28	67.76	80.52	58.49
MC-03	51.48	65.33	67.92	60.67	77.14	71.13	58.49	90.73	0.00	0.00	0.00	66.86	83.69	55.66
Baseline	26.61	45.44	0.00	0.00	0.00	56.18	43.01	80.98	5.41	15.38	3.28	44.88	54.36	38.21
1 st -rank	59.96	69.53	67.18	72.13	62.86	76.47	76.85	76.10	21.15	25.58	18.03	75.06	70.54	80.19

ExamBC Subtask

Run	Training data	Machine Learning	Features
JAIST-JA-ExamBC-01	Dev set (RITE2_ExamBC)	SVM	mono + bilingual
JAIST-JA-ExamBC-02	Dev set + test set (RITE1_ExamBC) + Dev set (RITE2_ExamBC)	SVM	mono +bilingual
JAIST-JA-ExamBC-03	Dev set (RITE2_ExamBC)	SVM	mono

Table 4: Results on ExamBC subtask (JA)

Run	MacroF1	Acc.	Corr. Answer Ratio	Y-F1	Y-Prec.	Y-Rec.	N-F1	N-Prec.	N-Rec.
JAIST-JA-ExamBC-01	57.55	63.17	40.74	42.11	53.57	34.68	73.00	66.37	81.09
JAIST-JA-ExamBC-02	59.04	63.39	41.67	45.70	53.49	39.88	72.39	67.40	78.18
JAIST-JA-ExamBC-03	58.65	64.96	42.59	42.49	58.00	33.53	74.80	66.95	84.73
Baseline	54.77	56.47	32.41	45.98	44.15	47.98	68.55	65.38	61.82
1 st -rank system	67.15	70.31	55.56	56.96	64.71	50.87	77.34	72.76	82.55

UnitTest Subtask

Run	Training data	Machine Learning	Features
JAIST-JA-BC-01	Dev set (RITE2_BC)	SVM	mono + bilingual
JAIST-JA-BC-02	Dev set (RITE2_BC)	Bagging	mono +bilingual
JAIST-JA-BC-03	Dev set + test set (RITE1_BC) + Dev set (RITE2_BC)	Bagging	mono

Table 5: Results on UnitTest subtask (JA)

Run	MacroF1	Acc.	Y-F1	Y-Prec.	Y-Rec.	N-F1	N-Prec.	N-Rec.
JAIST-JA-UnitTest-01	67.36	79.67	87.40	96.05	80.19	47.31	34.38	75.86
JAIST-JA-UnitTest-02	74.52	89.21	93.87	93.87	93.87	55.17	55.17	55.17
JAIST-JA-UnitTest-03	29.46	30.71	38.83	86.89	25.00	20.10	11.67	72.41
Baseline	51.70	86.31	92.58	88.41	97.17	10.81	25.00	6.90
1 st -rank system	77.77	90.87	94.84	94.39	95.28	60.71	62.96	58.62

Summary & Conclusions

- Based on the system used at NTCIR-9 RITE-1
 - Machine learning-based system
 - Utilize bilingual features
- Employ ensemble learning method
 - Bagging method
- Obtain competitive results at four tasks: BC, MC, ExamBC, UnitTest.