



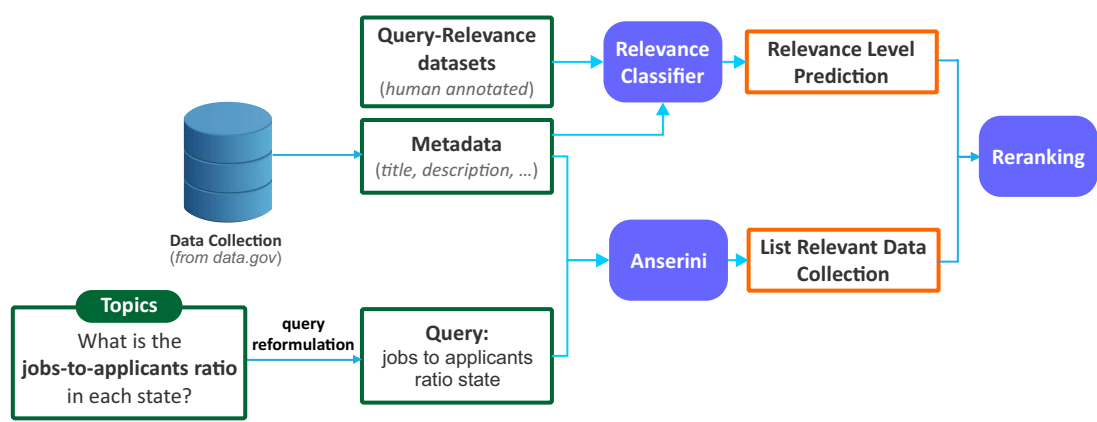
STIS at the NTCIR-DataSearch Task: Document Retrieval Reranking

Lya Hulliyyatus Suadaa (lya@stis.ac.id), Lutfi Rahmatuti M (lutfirm@stis.ac.id),
Isfan Nur Fauzi (isfan@stis.ac.id), and Siti Mariyah (sitimariyah@stis.ac.id)

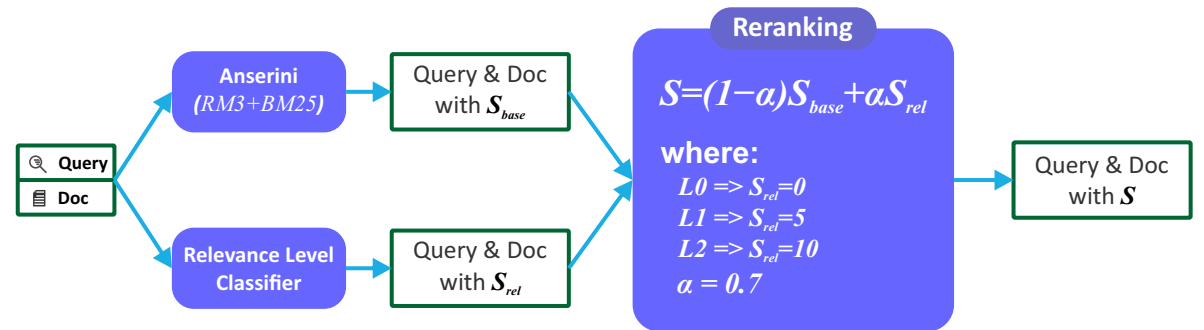
Politeknik Statistika STIS

Indonesia

Task and Motivation

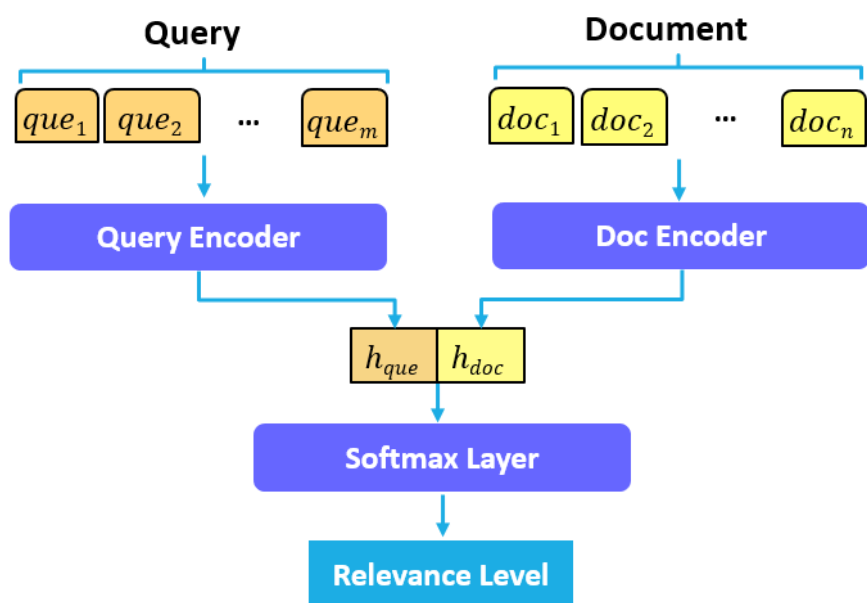


Reranker Model

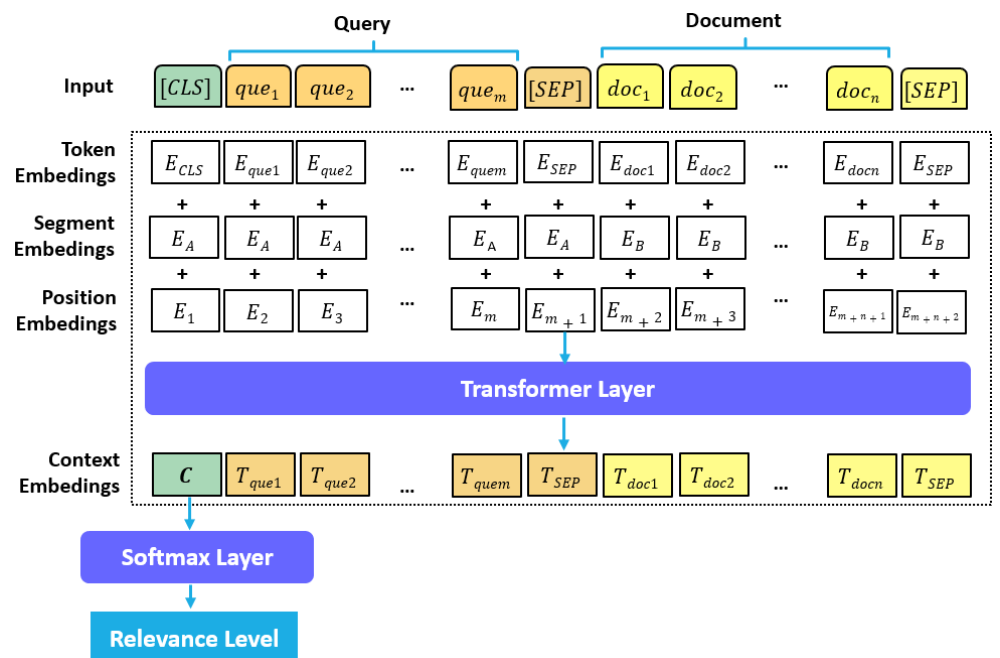


Relevance Classification Models

BiLSTM Classifier



BERT-based Classifier



Experiment Settings

Used metadata (**title, description, tags**) as document features

Retrieve **100 hundred** documents using **RM3+BM25** of Anserini

Split original train set into **90% train** set and **10% val** set

BiLSTM Classifier : two layer BiLSTM; 128 hidden sizes; initialized by Glove; dropout 0.5; Adam LR 0.003

BERT-based Classifier : dropout 0.1; BertAdam LR $3 \cdot 10^{-5}$

Implement our models using the **AllenNLP** library; batch size: 5; patience: 3

Experiment Results

Relevance Classifier	Doc Feature	P_{L0}	R_{L0}	F_{L0}	P_{L1}	R_{L1}	F_{L1}	P_{L2}	R_{L2}	F_{L2}	P	R	F
BiLSTM	title + desc	0.94	0.65	0.77	0.12	0.54	0.19	0	0	0	0.8664	0.6341	0.7148
BiLSTM	title + tag	0.96	0.46	0.62	0.11	0.76	0.19	0	0	0	0.886	0.4822	0.5847
Finetuned BERT-base	title + desc	0.97	0.70	0.82	0.18	0.75	0.30	0	0	0	0.9051	0.7039	0.771
Finetuned BERT-base	title + tag	0.98	0.73	0.84	0.21	0.85	0.34	0	0	0	0.9133	0.7376	0.794
Finetuned BERT-large	title + desc	0.99	0.64	0.78	0.18	0.91	0.30	0	0	0	0.916	0.6595	0.735
Finetuned BERT-large	title + tag	0.99	0.57	0.73	0.16	0.93	0.27	0	0	0	0.9167	0.5982	0.6847
Finetuned roBERTa-base	title + desc	0.98	0.42	0.59	0.12	0.89	0.21	0	0	0	0.9085	0.4572	0.5569
Finetuned roBERTa-base	title + tag	0.98	0.61	0.75	0.15	0.83	0.26	0	0	0	0.9038	0.6191	0.7031

Table 1: Performance of relevance level classifier.

Anserini	Relevance Classifier	Doc Feature	nDCG@3	nDCG@5	nDCG@10	nERR@3	nERR@5	nERR@10	Q-mea
RM3+BM25	-	title + desc	0.195	0.202	0.213	0.230	0.201	0.215	0.228
RM3+BM25	BiLSTM	title + desc	0.165	0.175	0.197	0.187	0.182	0.194	0.211
RM3+BM25	BiLSTM	title + tag	0.167	0.163	0.172	0.164	0.186	0.192	0.201
RM3+BM25	Finetuned BERT-base	title + desc	0.201	0.201	0.221	0.199	0.227	0.234	0.249
RM3+BM25	Finetuned BERT-base	title + tag	0.230	0.228	0.237	0.217	0.248	0.255	0.264
RM3+BM25	Finetuned BERT-large	title + desc	0.189	0.195	0.211	0.202	0.202	0.214	0.226
RM3+BM25	Finetuned BERT-large	title + tag	0.172	0.171	0.192	0.185	0.190	0.199	0.212
RM3+BM25	Finetuned roBERTa-base	title + desc	0.155	0.151	0.177	0.171	0.175	0.181	0.198
RM3+BM25	Finetuned roBERTa-base	title + tag	0.165	0.175	0.197	0.187	0.182	0.194	0.211

Table 2: Results from NTCIR-15 Data Search. Values in bold outperform the baseline. There are no significant differences between run pairs.