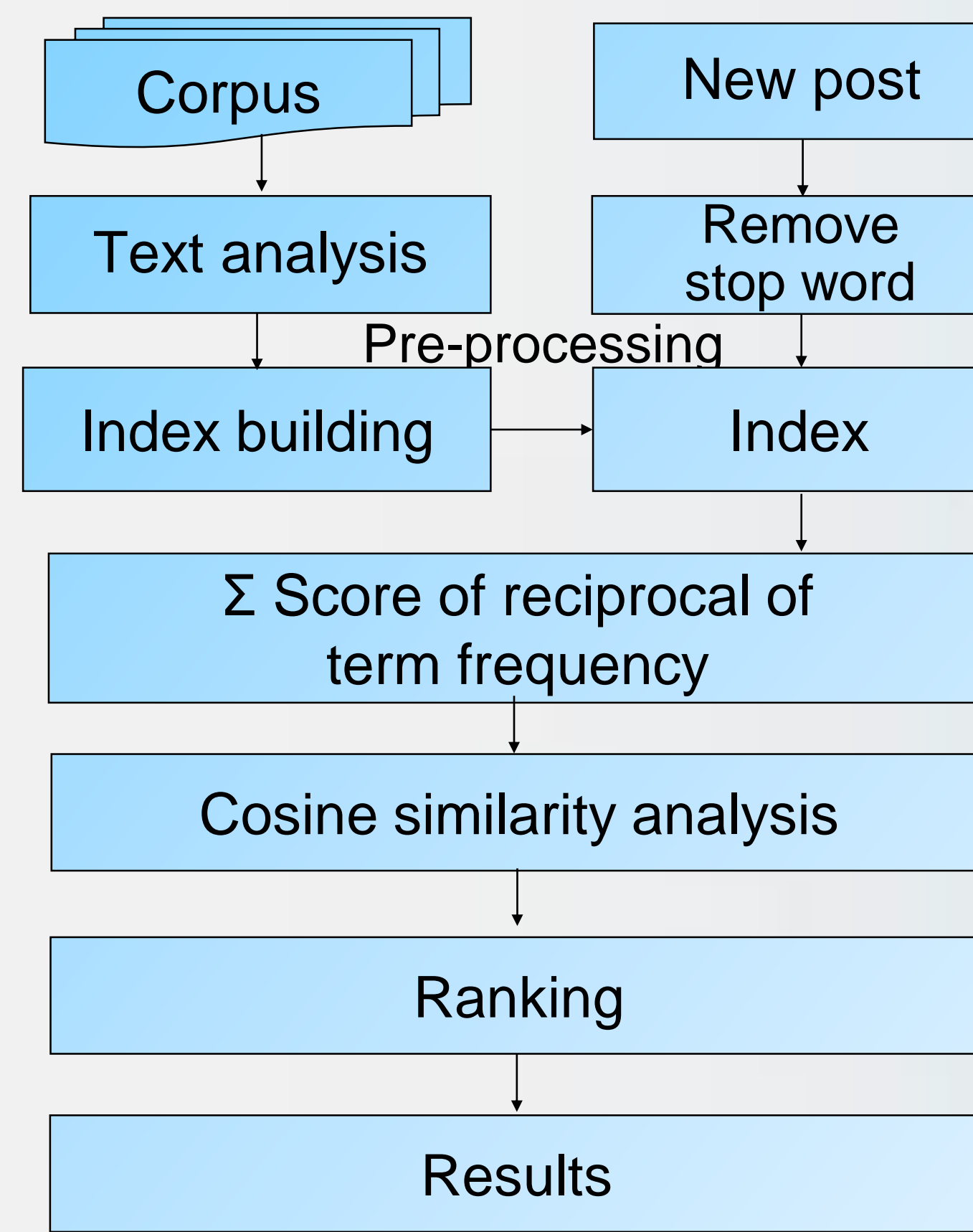


Introduction

Retrieval-based and generative-based approaches were used for emotional short text conversation in the study. In the retrieval-based approach, we adopted Apache Solr and used cosine similarity to rank the retrieval results. In the generative-based model, we used deep learning models, namely, MLP, GRU, LSTM, BiGRU, and BiLSTM to develop the emotion classification model. Beside, we used Emotional General Purpose Response (EGPR) to improve the generative-based response performance.

Retrieval-based System



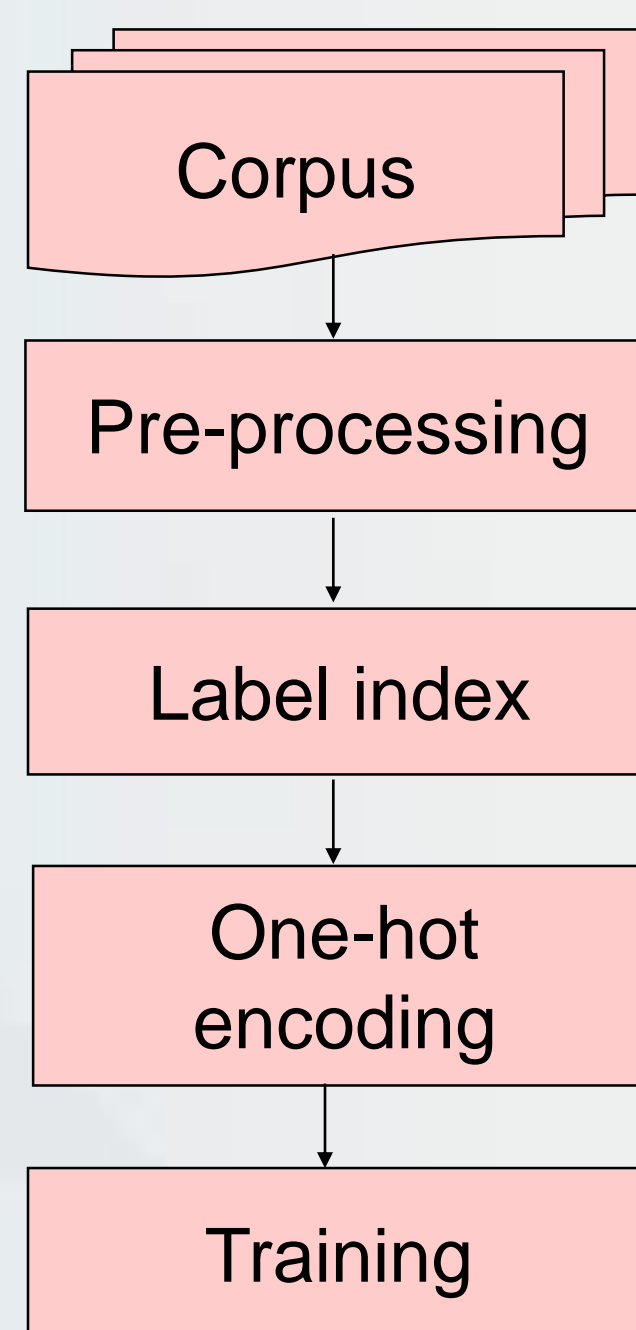
Resource: our research

In the **Retrieval-based method**, we used the term of provided new post to search the indexed posts in Apache Solr. We used the fetched search results as potential candidate comments. Then accumulated the inverse term frequency for each candidate comments and computed the cosine similarity between the new post and candidate comments. And we multiplied accumulated inverse term frequency by cosine similarity as the relevance score. Finally, the candidate comment that match the assigned emotion and is with highest relevance score is treated as the generated comment.

Generation-based System

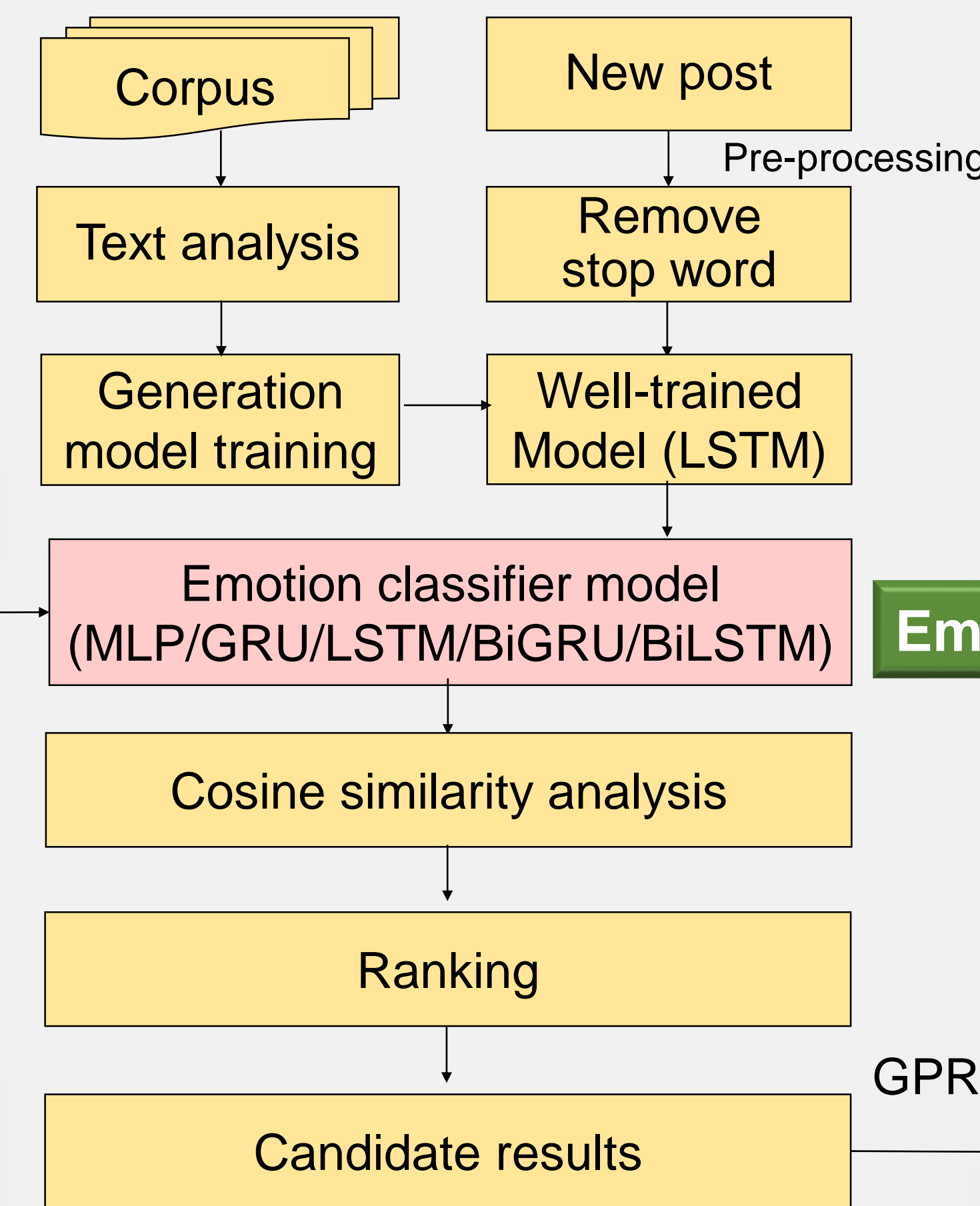
Emotion Classification model

- When a new post provided, we used the Seq2Seq which takes Long Short Term Memory (LSTM) as encoder and decoder of **Generation model** to train the model using the provided corpus to generate candidate comment.
- We also used the emotion classification model to tag the emotion of the generated comments.



In the extension study, we refined the model by 5 different deep learning models including **MLP, GRU, LSTM, BiGRU, and BiLSTM** to improve the performance of **Emotion classification**.

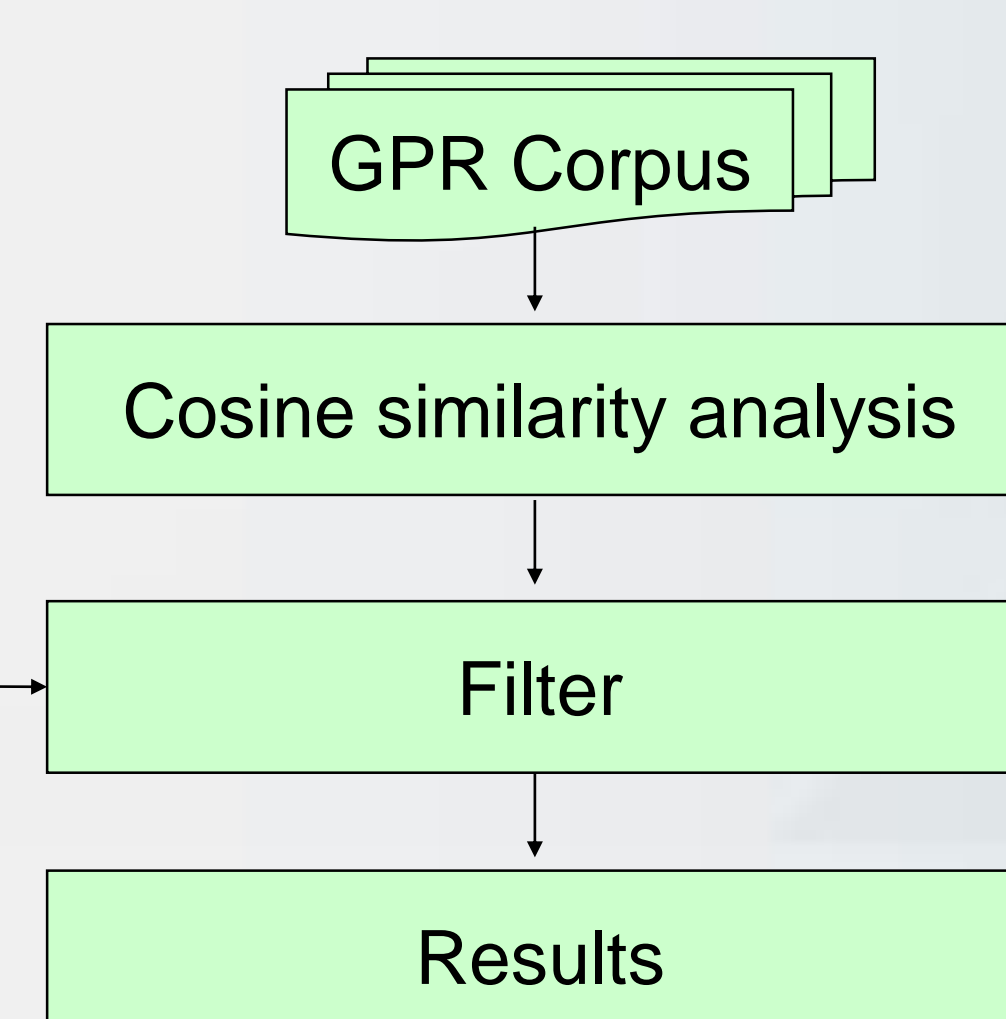
Generative model



Resource: our research

- Besides, we proposed **Emotional General Purpose Response(EGPR)** to improve the generative-based response performance. The proposed EGPR was created and complied with emotion classification model by integrating an open source robot corpus developed by Rediker Software and Chinese language dialogue collected from an educational web blogger.
- The generated comments will be replaced by the GPR at filter stage if the new post and generated comments received a low relevance score computed by cosine similarity.

Emotional General Purpose Response



Post
很想吃多力多滋芝士味儿那个 Want to eat Doritos's Cheese Flavor.
GPR Response example
我也很想吃 I want to eat, too.
我也好想吃 I also want to eat so much.
这味道我想吃都吃不到 I found no one every time I want to eat this flafor.
我每次想吃都吃不到 I found no one every time I want to eat.
我吃过真的很想再吃一次 I had tasted it and I really want to eat once.

Emotion Classification Model

DL model	Batchsize	Dropout	Epochs	Accuracy	Loss
BiGRU	256	0.5	15	0.880	0.333
BiLSTM	256	0.4	10	0.879	0.335
LSTM	256	0.1	20	0.879	0.335
GRU	256	0.4	20	0.872	0.356
MLP	256	0.4	30	0.843	0.451

Evaluation Results

Result	Submission	Chatbot System	Deep Learning	Label 0	Label 1	Label 2	Total	Overall score	Average score
Evaluation result	RUN 1	Retrieval	-	716	200	84	1000	368	0.368
Self-Evaluation	RUN 1	Retrieval	-	560	208	195	1000	598	0.598
Self-Evaluation	RUN 2	Generation	MLP	873	85	42	200	169	0.169
Self-Evaluation	Self-Experiment	Generation(GPR)	BiGRU	727	111	162	1000	435	0.435

Conclusion

AI-NTPU submitted two runs for STC-3(CECG).

- AI-NTPU-1: Retrieval-based model
- AI-NTPU-2: Generation-based model

We also made self-evaluation to check the performance of our extension study.

Future work :

- Bidirectional Encoder Representation from Transformers (BERT) to improve the performance of emotion classification model

Acknowledgements

Our thanks to NTCIR-14 task organizers for their hard work on organizing NTCIR-14 STC-3 CECG subtask