



**C&LING  
2020**

TextGraphs 2020 Shared Task

# LIT : LSTM-Interleaved Transformer for Multi-Hop Explanation Ranking

**Yew Ken Chia**  
ken@RedDragon.ai

**Sam Witteveen**  
sam@RedDragon.ai

**Martin Andrews**  
martin@RedDragon.ai

## Summary

### Shared Task :

- ▲ Rank explanation sentences for elementary school science questions

### Data Used :

- ▲ WorldTree V2 Corpus
- ▲ 'Common sense' embedded in BERT

### Ideas :

- ▲ Improve BM25 ranking incrementally
- ▲ Use interaction between explanations
- ▲ LSTM chains for rank-aware interaction

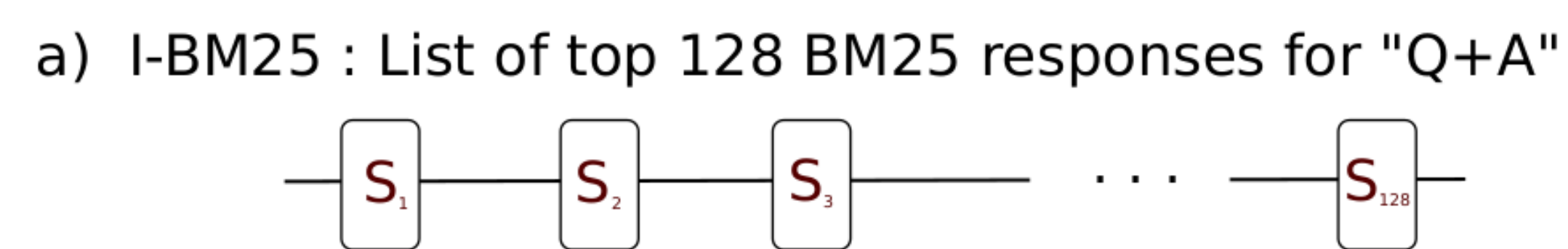
### Results :

- ▲ Submitted score : 0.4793
- ▲ Better methods submitted soon after

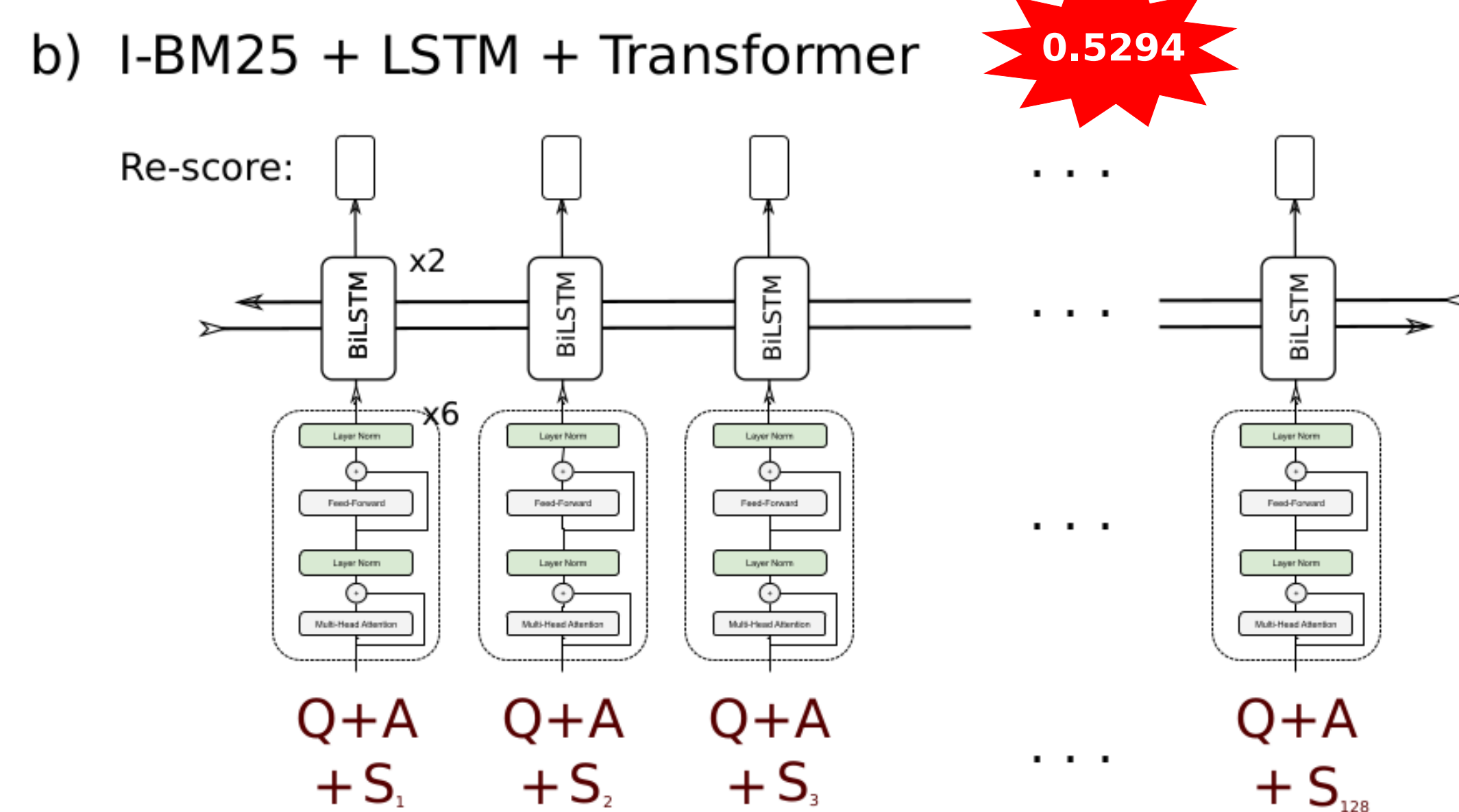
## Key References

- ▲ "TextGraphs 2020 Shared Task on Multi-Hop Inference for Explanation Regeneration" - Jansen and Ustalov (2020)
- ▲ "Colbert: Efficient and effective passage search via contextualized late interaction over BERT" - Khattab and Zaharia (2020)
- ▲ "Modeling document interactions for learning to rank with regularized self-attention" - Sun and Duh (2020)
- ▲ "Parameter-efficient transfer learning for NLP" - Houshy et al. (2019)

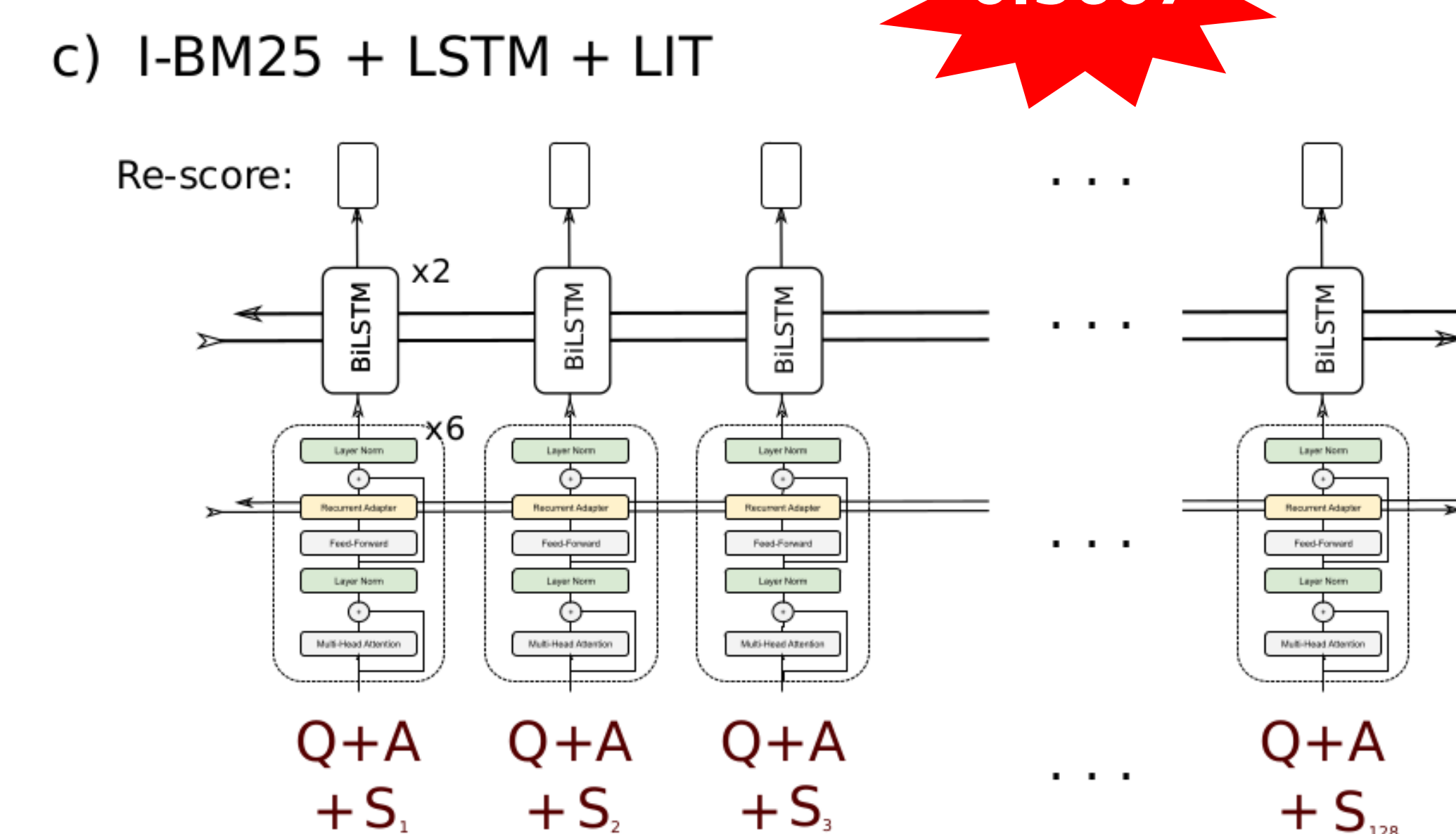
## Three Methods with Increasing Test Scores



**0.4745**

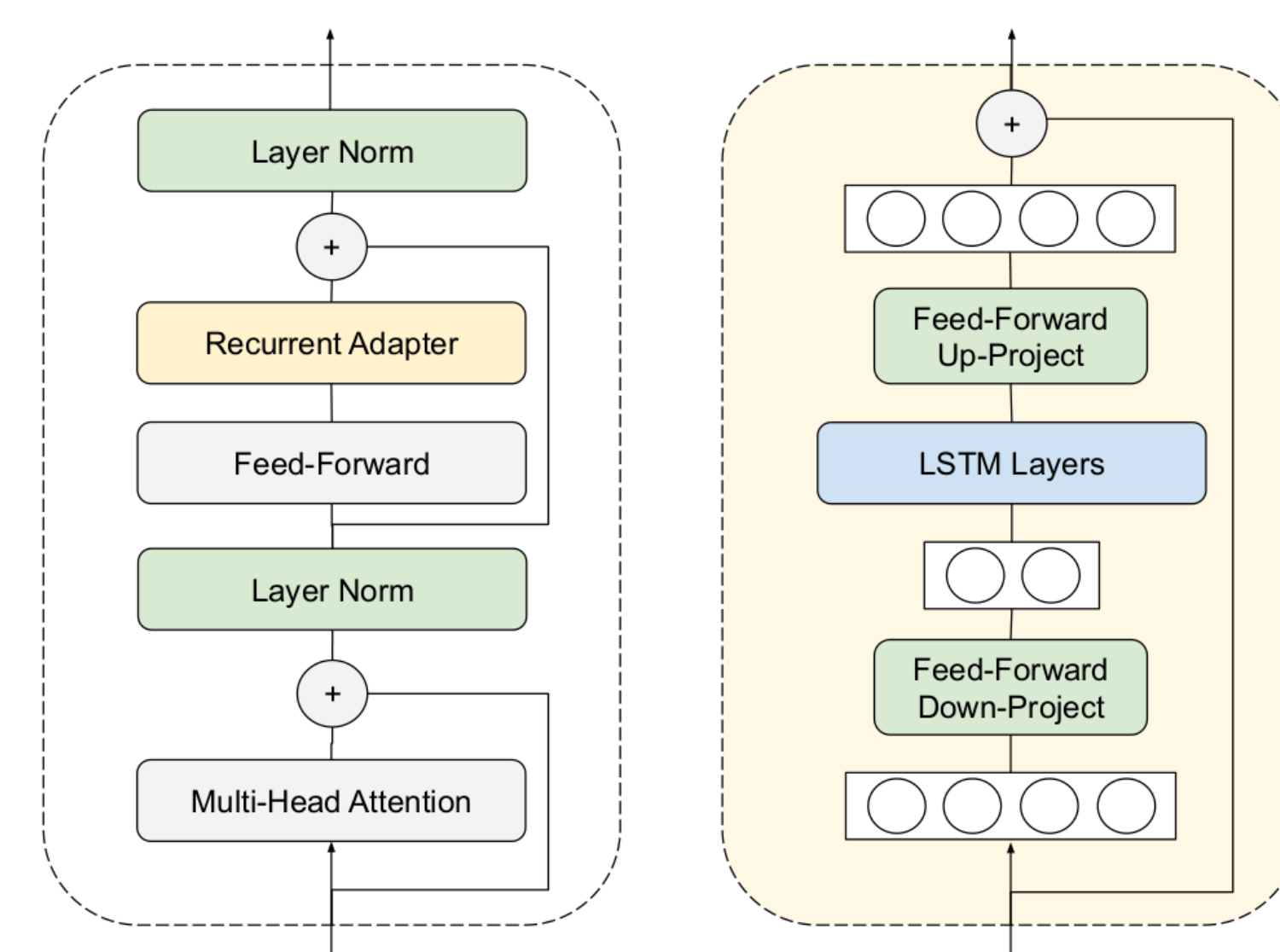


**0.5294**



**0.5607**

## LIT Detail



### Enhanced Information Distribution:

- ▲ Use adapter layer in Transformer modules
- ▲ Link document representations using LSTM
- ▲ Earlier information flow improves results

## Results

Model	Dev MAP	Test MAP
BM25	0.4615	
Iterative BM25 (Chia et al., 2019)	0.4704	
I-BM25	0.4861	0.4745
I-BM25 + LSTM + Transformer	0.5470	0.5294
I-BM25 + LIT	0.5680	0.5607

Table 1: Main score comparison on WorldTree V2 dataset

- ▲ Investigation of different Loss functions:

Loss Function	Dev MAP
LambdaLoss	0.4970
APLoss	0.5187
Binary Crossentropy	0.5680

Table 2: Loss function comparison on WorldTree V2 dataset

## Discussion

### Updated Dataset :

- ▲ Larger set of Q&A and facts
- ▲ Larger training set / more 'distractors'
- ▲ Still not totally clean

### Preprocessing :

- ▲ Use spaCy for lemmatisation
- ▲ I-BM25 is enhanced from 2019 version
- ▲ "Combo statements" still W.I.P.

### Focus on Transformer Reranking :

- ▲ DistilBERT used for 'common sense'
- ▲ Novel LIT architecture
- ▲ Tried GNN methods, but observed same problems as other participants

### Future directions :

- ▲ Still don't have solid grounding for Graph-based methods
- ▲ LIT architecture shows promise as a drop-in replacement for other Q&A tasks

## Code & Contact

### Source code is on GitHub, see:

- ▲ <http://RedDragon.ai/research>

### Contact:

martin@RedDragon.ai  
+65 8585 1750  
<http://RedDragon.ai>