Bert-Based Reinforcement Training on Part-Of-Sentence Task And Knowledge Transferring

Data Science Institute
COLUMBIA UNIVERSITY

Zhirui Yang, Qiran Li, Zheng Wu, Chenqi Wang, Zonghan Yue, Xianmeng Wang
Industry Mentor: Akshat Gupta

Data Science Capstone Project with JPMC

Part-Of-Sentence Task and Basic Generalization of Online Training

Part-of-sentence (POS tagging or POST) is a widely used NLP task. The target is to mark up a word in a corpus (text) as corresponding to a particular part of the sentence based both on its definition (meaning) and context (semantic information). POS tags describe the characteristic structure of lexical terms within a corpus; it can be used for making assumptions about semantics. A basic method is online training.

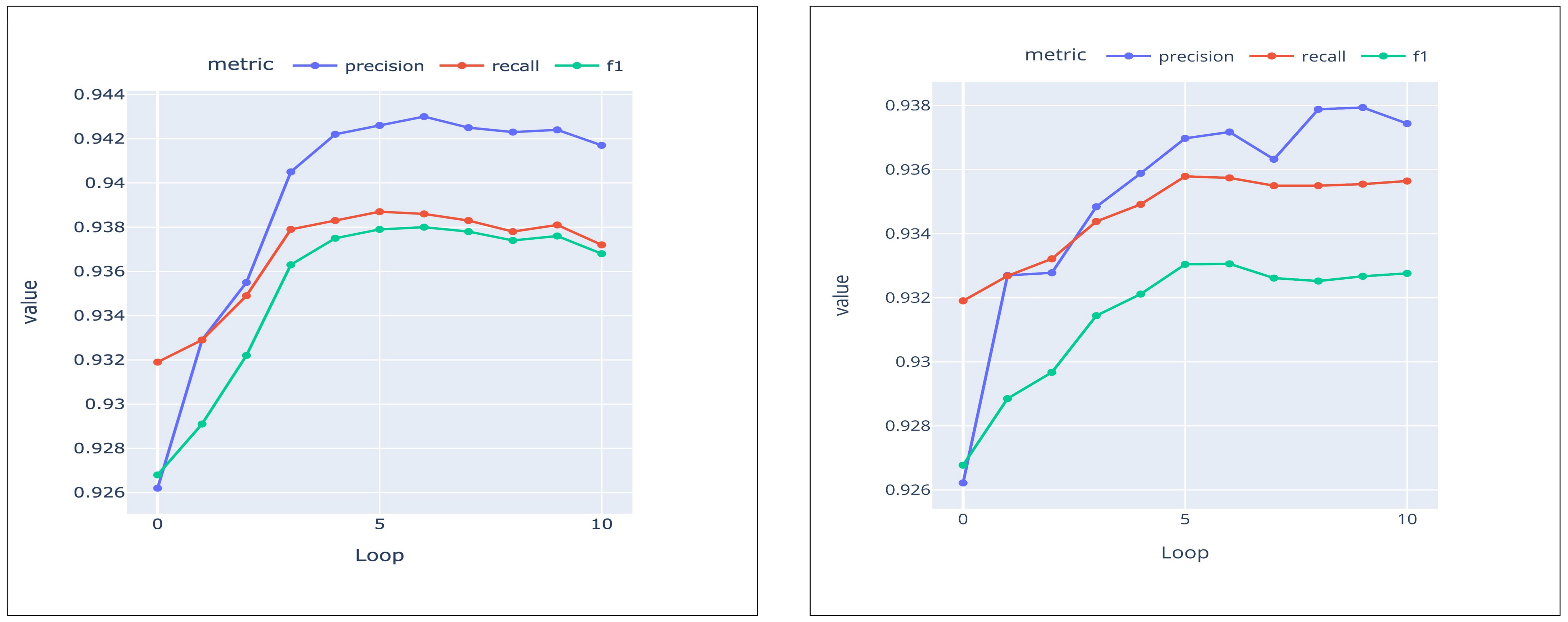

Figure 1&2. Model Evaluation of Online Fixed & Non-Fixed Domain Transferring (top 10% sentences)

Reinforcement Model Trained from Scratch: Pseudo Code & Result

The second section shows the results of reinforcement models trained from scratch.

Fixed Learning from Scratch:

Repeat:

1. *Make predictions **on the remaining unlabeled data***
2. *Use mean probability to select Top N sentences*
3. *Generate new training data*
4. *Use **[new training data + WSJ data]** to train the base_model from scratch*
5. *Remove Top N sentences from **unlabeled data***

Stop: No remaining data

Non-Fixed Learning from Scratch:

Repeat k times:

1. *Make predictions **on the whole unlabeled data***
2. *Use mean probability to select Top M tokens*
3. *Ignore other token in the sentences*
4. *Use **[whole unlabeled data]** to train the base_model*

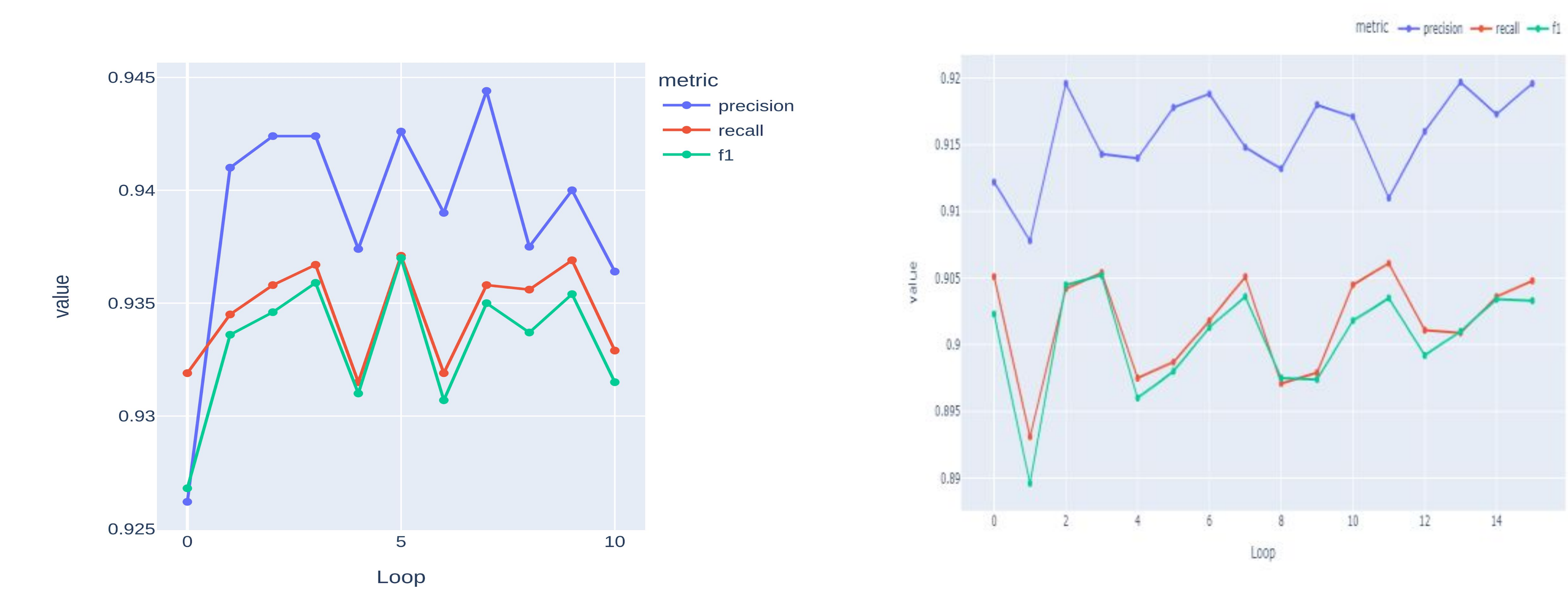

Figure 3&4. Model Evaluation of Scratch Fixed & Non-Fixed Domain Transferring (top 10% sentences)

Online Fixed/Non-fixed Model (Sentence-level): Pseudo Code

The algorithm of online fixed/non-fixed models (figure 1&2) are shown as below.

Online Fixed Learning:

Repeat:

1. *Make predictions **on the remaining unlabeled data***
2. *Use mean probability to select Top N sentences*
3. *Generate new training data*
4. *Use **[new training data]** to train the base_model*
5. *Remove Top N sentences from **unlabeled data***

Stop: No remaining data

Online Non-Fixed Learning:

Repeat:

1. *Make predictions **on the whole unlabeled data***
2. *Use mean probability to select Top N sentences*
3. *Generate new training data*
4. *Use **[new training data]** to train the base_model*

Stop: The number of changes in Top N sentences is less than 20% of N

Exploratory Reinforcement Model (Token-level): Result

We also derived a token-level model which might be a further research direction.

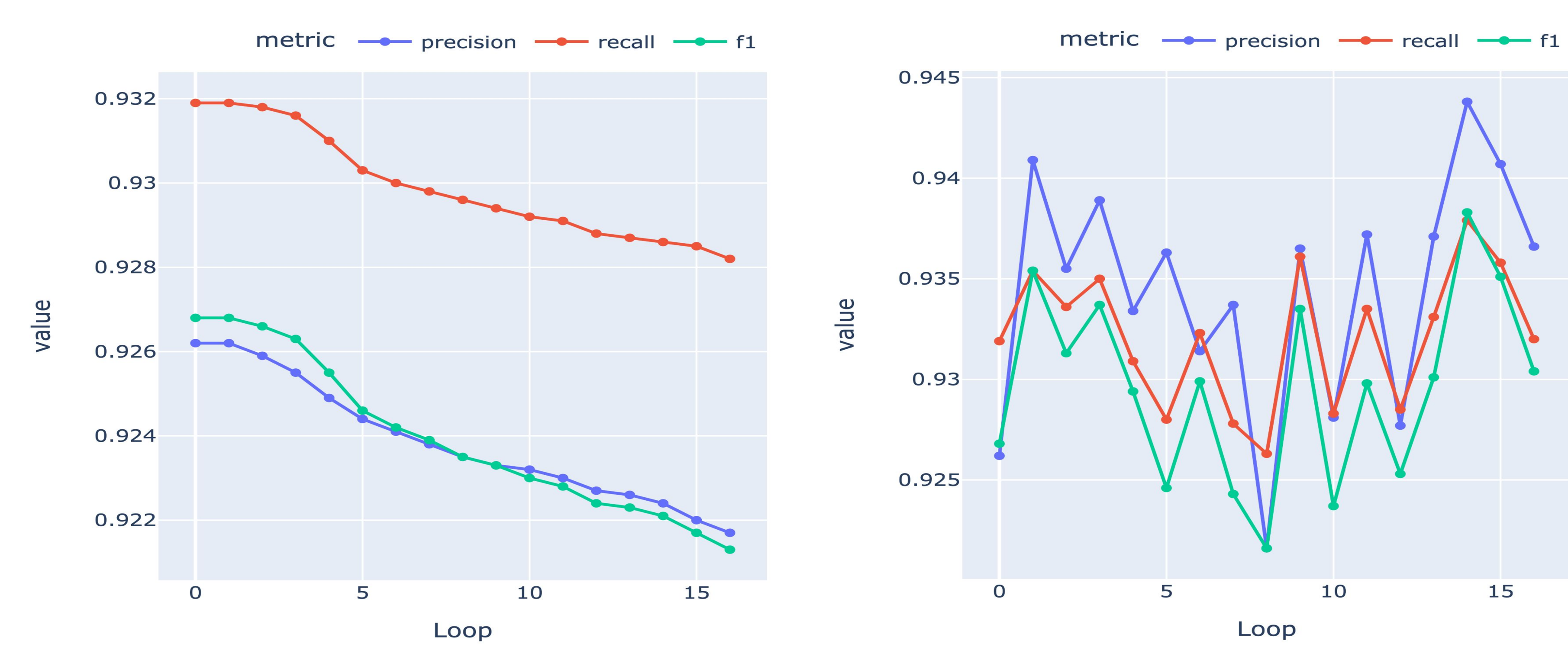

Figure 5&6. Model Evaluation of Online & Scratch Token Self-Learning

Conclusion

Our BERT-based online/reinforcement training models show superiority compared with the traditional Bi-LSTM models in domain-adaptation tasks. We derived 6 different models (3 online, 3 reinforcement). Specifically, the top percentage of sentences, the core parameter, brought approximately a 1.5% increment on F1-score with dramatically less time cost in our fixed/non-fixed online models.

Acknowledgments

We are deeply indebted to Barbara Plank for sharing the complete cross-domain datasets. We are extremely grateful to Columbia DSI for providing us with GCP computational power.

References

Ruder, Sebastian, and Barbara Plank. "Strong Baselines for Neural Semi-Supervised Learning under Domain Shift." ArXiv.org, 25 Apr. 2018, https://arxiv.org/abs/1804.09530.

Petrov, Slav, and Ryan McDonald. "Overview of the 2012 Shared Task on Parsing the Web." https://static.googleusercontent.com/media/research.google.com/zh-CN//pubs/archive/38278.pdf.