课程阅读报告要求

1. 使用中文撰写，报告篇幅应在 5000 字以上，内容建议涵盖：

1) 背景介绍（论文所针对的 NLP 任务）；

2) 现有的方法及其局限性，论文方法的优越性；

3) 论文提出的方法（模型架构、训练策略等）；

4) 实验结果与分析（数据集、对比模型、数据分析等）；

5) 该方法的局限性，可能的拓展方向。

2. 有开源代码的论文，如果在报告中体现复现结果和进行改进优化，会有适当加分。

3. 3 人一组自行组队完成。要求：最终报告需要指明每个组员的负责内容，最终成绩将结合小组整体表现，以及具体每个组员的工作量综合评定。

4. 4 月 11 日前私信联系助教告知分组组员名单并选择调研论文（报论文编号），每组选择一篇，至多 2 组同学选择同一篇论文，先到先得。

课程阅读备选论文

- 1. 词向量

1.1 word2vec

Efficient Estimation of Word Representations in Vector Space

http://arxiv.org/pdf/1301.3781v3.pdf

备注：开源代码 https://code.google.com/p/word2vec/

1.2 doc2vec

Distributed Representations of Words and Phrases and their Compositionality

http://papers.nips.cc/paper/5021-distributed-representations-of-words-and-phrasesand-their-compositionality.pdf

备注：开源代码 https://code.google.com/p/word2vec

1.3 GloVe

GloVe: Global Vectors for Word Representation

https://aclanthology.org/D14-1162.pdf

备注：开源代码 https://nlp.stanford.edu/projects/glove/

1.4 DSG

Directional Skip-Gram: Explicitly Distinguishing Left and Right Context for Word Embeddings

https://aclanthology.org/N18-2028.pdf

https://ai.tencent.com/ailab/nlp/en/embedding.html

- 2. 基本网络模型

2.1 CNN

Convolutional Neural Networks for Sentence Classification

http://arxiv.org/abs/1408.5882

备注：开源代码 https://github.com/yoonkim/CNN_sentence

2.2 LSTM

Attention-based LSTM for Aspect-level Sentiment Classification

https://aclanthology.org/D16-1058.pdf

备注：开源代码 https://github.com/songyouwei/ABSA-PyTorch#atae-lstm-atae_lstmpy

2.3 LSTM-RNN

A Critical Review of Recurrent Neural Networks for Sequence Learning

https://arxiv.org/abs/1506.00019

- 3. 语言模型

3.1 RNN-LM

Recurrent neural network based language model

https://www.cs.cmu.edu/~hiroakih/pdf/RNNLM_hiroakih.pdf（Slides）

备注：开源代码 https://github.com/mspandit/rnnlm

3.2 LSTM-Char-CNN

Character-Aware Neural Language Models

https://www.aaai.org/ocs/index.php/AAAI/AAAI16/paper/view/12489/12017

备注：开源代码 https://github.com/yoonkim/lstm-char-cnn

- 4. 自然语言处理框架

4.1 Natural Language Processing (Almost) from Scratch

http://www.jmlr.org/papers/volume12/collobert11a/collobert11a.pdf

- 5. 命名实体识别

5.1 Neural Architectures for Named Entity Recognition

https://arxiv.org/pdf/1603.01360.pdf

备注：开源代码 https://github.com/glample/tagger&https://github.com/clab/stack-lstm-ner

5.2 HSCRF

Hybrid semi-Markov CRF for Neural Sequence Labeling

https://aclanthology.org/P18-2038.pdf

备注：开源代码 https://github.com/ZhixiuYe/HSCRF-pytorch

5.3 Survey

A Survey on Deep Learning for Named Entity Recognition

https://arxiv.org/pdf/1812.09449v3.pdf

5.4 HSCRF + softdict

Towards Improving Neural Named Entity Recognition with Gazetteers

https://aclanthology.org/P19-1524/

备注：开源代码 https://github.com/lyutyuh/acl19_subtagger

5.5 cross-lingual

Single-/Multi-Source Cross-Lingual NER via Teacher-Student Learning on Unlabeled Data in Target Language

https://aclanthology.org/2020.acl-main.581.pdf

备注:开源代码 https://github.com/microsoft/vertpapers/tree/master/papers/SingleMulti-TS

- 6. 记忆&注意力网络

6.1 End-To-End Memory Networks

https://arxiv.org/abs/1503.08895

备注：开源代码 https://github.com/facebook/MemNN

6.2 Hierarchical Attention Networks for Document Classification

https://www.aclweb.org/anthology/N16-1174

备注：开源代码 https://github.com/richliao/textClassifier

6.3 A Structured Self-Attentive Sentence Embedding

https://arxiv.org/abs/1703.03130v1

备注：开源代码 https://github.com/ExplorerFreda/Structured-Self-Attentive Sentence-Embedding

- 7. 机器翻译

7.1 RNN-based

Neural Machine Translation by jointly learning to align and translate

http://arxiv.org/pdf/1409.0473v6.pdf

备注：开源代码 https://github.com/ihsgnef/Groundhog

7.2 CNN-based

Convolutional Sequence to Sequence Learning

https://arxiv.org/abs/1705.03122

备注：开源代码 https://github.com/facebookresearch/fairseq

7.3 Transformer-based

Attention Is All You Need

https://arxiv.org/abs/1706.03762

备注：开源代码 https://github.com/tensorflow/tensor2tensor

7.4 Pre-training-based

BART: Denoising Sequence-to-Sequence Pre-training for Natural Language Generation, Translation, and Comprehension

https://aclanthology.org/2020.acl-main.703.pdf

备注：开源代码 https://github.com/pytorch/fairseq/tree/master/examples/bart

- 8. 文本摘要

8.1 A Neural Attention Model for Abstractive Sentence Summarization.

http://arxiv.org/abs/1509.00685

备注：开源代码 https://github.com/facebookarchive/NAMAS

8.2 Get To The Point: Summarization with Pointer-Generator Networks

https://aclanthology.org/P17-1099.pdf

备注：开源代码 https://github.com/abisee/pointer-generator

8.3 Text Summarization with Pretrained Encoders

https://aclanthology.org/D19-1387.pdf

备注：开源代码 https://github.com/nlpyang/PreSumm

8.4 Learning to Summarize from Human Feedback

https://papers.nips.cc/paper/2020/file/1f89885d556929e98d3ef9b86448f951-Paper.pdf

备注：开源代码 https://github.com/openai/summarize-from-feedback

8.5 GSum: A General Framework for Guided Neural Abstractive Summarization

https://aclanthology.org/2021.naacl-main.384.pdf

备注：开源代码 https://github.com/neulab/guided_summarization

- 9. 对话生成

9.1 encoder-decoder-based

Neural responding machine for short-text conversation

https://arxiv.org/abs/1503.02364

9.2 latent-variable-based

A Hierarchical Latent Variable Encoder-Decoder Model for Generating Dialogues

https://www.aaai.org/ocs/index.php/AAAI/AAAI17/paper/viewPaper/14567

备注：开源代码 https://github.com/Ravi-Jay/VHRED-implementation-in-Tensorflow

9.3 Language-model-based

DIALOGPT : Large-Scale Generative Pre-training for Conversational Response Generation

https://aclanthology.org/2020.acl-demos.30.pdf

备注：开源代码 https://github.com/microsoft/DialoGPT

- 10. 自然语言推理

10.1 ESIM

Enhanced LSTM for Natural Language Inference

https://aclanthology.org/P17-1152.pdf

备注：开源代码 https://github.com/lukecq1231/nli

10.2 KIM

Neural Natural Language Inference Models Enhanced with External Knowledge

https://aclanthology.org/P18-1224.pdf

备注：开源代码 https://github.com/lukecq1231/kim

10.3 SemBERT

Semantics-aware BERT for Language Understanding

https://arxiv.org/pdf/1909.02209.pdf

备注：开源代码 https://github.com/cooelf/SemBERT

10.4 CA-MTL

Conditionally Adaptive Multi-Task Learning: Improving Transfer Learning in NLP Using Fewer Parameters & Less Data

https://arxiv.org/pdf/2009.09139.pdf

备注：开源代码 https://github.com/CAMTL/CA-MTL

- 11. 预训练语言模型

11.1 ELMo

Deep contextualized word representations

https://aclanthology.org/N18-1202.pdf

备注：开源代码 https://allennlp.org/elmo

11.2 GPT

Improving Language Understanding by Generative Pre-Training

https://www.cs.ubc.ca/~amuham01/LING530/papers/radford2018improving.pdf

备注：开源代码 https://github.com/openai/finetune-transformer-lm

11.3 BERT

BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding

https://aclanthology.org/N19-1423.pdf

备注：开源代码 https://github.com/google-research/bert

11.4 GPT-2

Language Models are Unsupervised Multitask Learners

https://d4mucfpksywv.cloudfront.net/better-language-models/languagemodels.pdf

备注：开源代码 https://github.com/openai/gpt-2

11.5 XLNet

XLNet: Generalized Autoregressive Pretraining for Language Understanding

http://papers.nips.cc/paper/8812-xlnet-generalizedautoregressive-pretraining-forlanguage-understanding.pdf

备注：开源代码 https://github.com/zihangdai/xlnet

11.6 GPT-3

Language Models are Few-Shot Learners

https://papers.nips.cc/paper/2020/file/1457c0d6bfcb4967418bfb8ac142f64aPaper.pdf

备注：开源代码 https://github.com/openai/gpt-3

11.7 Tinybert

TinyBERT: Distilling BERT for Natural Language Understanding

https://arxiv.org/pdf/1909.10351.pdf

代码 https://github.com/huawei-noah/Pretrained-Language-Model/tree/master/TinyBERT

11.8 InstructGPT

Training language models to follow instructions with human feedback

https://arxiv.org/pdf/2203.02155.pdf

11.9 ERNIE

ERNIE: Enhanced Language Representation with Informative Entities

https://arxiv.org/pdf/1905.07129.pdf

11.10 LLaMA 2

Llama 2: Open Foundation and Fine-Tuned Chat Models

https://arxiv.org/pdf/2307.09288.pdf

备注： 开源代码 https://github.com/meta-llama/llama

11.11 ChatGLM

GLM: General Language Model Pretraining with Autoregressive Blank Infifilling

https://arxiv.org/abs/2103.10360

备注：开源代码 https://github.com/THUDM/ChatGLM3

- 12. 生成对抗网络

12.1 SeqGAN

SeqGAN: Sequence Generative Adversarial Nets with Policy Gradient

https://arxiv.org/abs/1609.05473

备注：开源代码 https://github.com/LantaoYu/SeqGAN

12.2 IRGAN

IRGAN: A Minimax Game for Unifying Generative and Discriminative Information Retrieval Models

https://arxiv.org/pdf/1705.10513.pdf

备注：开源代码 https://github.com/geek-ai/irgan

- 13. 强化学习

13.1 RL for Dialogue Generation

Deep Reinforcement Learning for Dialogue Generation

https://aclweb.org/anthology/D16-1127

备注：开源代码 https://github.com/liuyuemaicha/Deep-Reinforcement-Learning-forDialogue-Generation-in-tensorflow

13.2 RL for Information Extraction

Improving Information Extraction by Acquiring External Evidence with Reinforcement Learning

https://aclanthology.org/D16-1261.pdf

备注：开源代码 https://github.com/karthikncode/DeepRL-InformationExtraction

13.3 RL for Extractive Summarization

Ranking Sentences for Extractive Summarization with Reinforcement Learning

https://aclanthology.org/N18-1158.pdf

备注：开源代码 https://github.com/EdinburghNLP/Refresh

13.4 RL for Extractive Summarization

Multi-document Summarization with Maximal Marginal Relevance-guided Reinforcement Learning

https://aclanthology.org/2020.emnlp-main.136.pdf

备注：开源代码 https://github.com/morningmoni/RL-MMR

13.5 RL for QA

Few-Shot Complex Knowledge Base Question Answering via Meta Reinforcement Learning

https://aclanthology.org/2020.emnlp-main.469.pdf

13.6 RLHF

Training language models to follow instructions with human feedback

https://proceedings.neurips.cc/paper_files/paper/2022/file/b1efde53be364a73914f58805a001731-Paper-Conference.pdf

- 14. 问答

14.1 knowledge graph for QA

QA-GNN: Reasoning with Language Models and Knowledge Graphs for Question Answering

https://arxiv.org/pdf/2104.06378.pdf

开源代码: https://github.com/michiyasunaga/qagnn

14.2 Multi-hop QA

Hierarchical Graph Network for Multi-hop Question Answering

https://arxiv.org/pdf/1911.03631.pdf

代码： https://github.com/yuwfan/HGN

14.3 RAG

Retrieval-Augmented Generation for Knowledge-Intensive NLP Tasks

https://proceedings.neurips.cc/paper_files/paper/2020/file/6b493230205f780e1bc26945df7481e5-Paper.pdf

代码： https://huggingface.co/docs/transformers/model_doc/rag

14.4 Self-RAG

Self-rag: Learning to retrieve, generate, and critique through self-reflection

https://arxiv.org/pdf/2310.11511.pdf

代码：https://github.com/AkariAsai/self-rag

14.5 Chain of Thoughts

Chain-of-Thought Prompting Elicits Reasoning in Large Language Models

https://arxiv.org/pdf/2201.11903.pdf