Semester Summary

Lucas @ Dec. 5, 2017
Sequence Labeling

• Empower Sequence Labeling with Task-Aware Neural Language Model
  
  • w. Jingbo, Frank, Xiang, etc.
  
  • Co-train neural language model to extract knowledge and prevent over-fitting
  
  • accepted by AAAI, code released to Github:
    
    • 226 stars & 51 folks
Sequence Labeling

• Further improvement:
  • w. Jingbo, Xiaotao, Xiang, etc.
  • decompose Tremendous Language Model
    • into word-level part and character-level part
  • submitted to ICLR; would recycle it to NAACL
    • CoNLL03:
      • AI2 ACL paper: 91.93 +/- 0.19 (pre-trained on 32 GPU for more than half a month)
      • Our results: 92.06 +/- 0.11 (single GPU for several hours)
Named Entity Recognition

• Previous works:
  • Alleviate the dependency of human annotations by incorporating raw text

• Now:
  • Replace human annotations with automatically generated annotations

• Framework:
  • Knowledge Base $\rightarrow$ Phrases in Corpus
  • Phrases $\rightarrow$ Entity Boundary in Sentence
  • Entity Boundary in Sentence $\rightarrow$ Named Entity Recognition

Low-resource Language
Domain-specific Corpus
Downstream Applications
...
Relation Extraction

• Tough Baseline
  • A simple model but achieves the state-of-the-art performance

• Joint Learning with Named Entity Recognition
  • Automatically construct features for Relation Extraction
Collaborations

• GRACE: network clustering with dynamic embedding
  • W. Carl, Mentxiong, etc

• Hierarchical TextCube Construction
  • W. Zoey, Jingbo, Yu Zhang, etc.

• Bridging Knowledge and Text by Learning from Dictionary
  • W. Xiaotao, Jingbo, Xiang, etc.
Thank You

• for all the help, comments, discussions 😊

• hope we can form more collaboration

• reading group?