16 May CS224n: NLP with deep learning Week9 Yukun J @nyu sh

Lecture 17 - multi task learning

- Machine learning with feature engineering
- Deep learning for feature learning
- Deep learning architecture for single tasks

The limit of single-task learning

- need continuous learning in a single model instead
- start randomly or pre-trained
- we can hill-climb to local optima

There is blocking task! i.e. classification in C.V.

why not many weight & model sharing happen in NLP?

- NLP needs different types of reasoning.
- divide into separate tasks.
- Require short and long term memory.

Language requires supervision in nature.

Multi-task learning is a blocker for general NLP systems.

3 NLP tasks framework

- Sequence tagging
- Text classification
- Seq 2 seq

3 supertasks for NLP

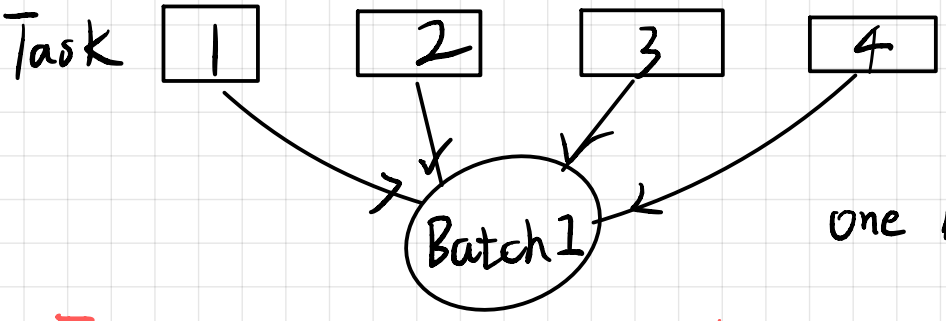
- Language Modeling
- Question answering
- Dialogue

MQAN (multitask Question answering Network)

- Toansformer lager yield benefits in single-task and multi-task setting.

Training strategy

- Join training Fully



one by one feed in tasks

- Tasks should be fed in decreasing order of difficulty.

difficult first

Lecture 18 - constituency Parsing, Tree RNNs

The spectrum of CS

- Bag of words

Semantic interpretation of language — Not just word vectors

Snowboarder ⟷ person on a snowboard

Constituency Sentence Parsing: What we want

Recursive Neural Networks for structure Prediction

- semantic representation
- score of meaning representation

△ Scene Parsing

similar principle of compositionality

Setiment analysis: is the tone of a piece of text positive, neutral or negative?