

Research Projects

Pragmatic competence of pre-trained language models through the lens of discourse connectives *Advised by Dr Allyson Ettinger*

- In this project, we examine aspects of pre-trained LMs competence in pragmatics, with a particular focus on pragmatic reasoning surrounding discourse connectives.
- We use controlled tests, inspired by psycholinguistics, to check if models are able to use high-level pragmatic information rather than relying on superficial lexical cues.
- We have tried several pre-trained models and found that they lack the pragmatic sensitivity to perform well on our controlled tests, and they don't show any strong humanlike preference in terms of temporal dynamics.
- We have submitted a paper(under review) at CoNLL 2021.

Sorting through the noise: Testing robustness of information processing in pre-trained language models *Advised by Dr Allyson Ettinger*

- In this project, we are trying to understand how well pre-trained LM models perform when it comes to processing, retaining and applying information in text.
- We introduce a dataset where each data instance has two sentences. The first sentence contain entity and the profession associated with the entity. The second sentence has cloze-style statement which can be filled correctly if the models retain information about the profession.
- We have tried several pre-trained models and found that these models degrade in performance once we introduce contextual manipulations.
- We have submitted a paper(under review) at EMNLP 2021.

Encoder based Attention Mechanism for Word Sense Disambiguation *Advised by Dr Harish Karnick*

- In this paper we presented the idea of transfer learning using a pre-trained BERT model to fine-tune a model for Word Sense Disambiguation.
- Used a weighted loss function to handle extreme imbalance of senses in datasets.
- Achieved state-of-the-art results on hard, line, serve, and interest datasets.

Word Sense Disambiguation by learning Long Term Dependencies *M.Tech Thesis [Supervised by Dr Harish Karnick] Jan 2016 - Jun 2017*

- We applied bi-directional LSTM to use sequence information to disambiguate senses of a polysemous word.
- Achieved better macro F1 scores than state-of-the-art on several datasets like hard and serve.
- We tried the same approach on selected words from One million sense tagged corpus and achieved encouraging results.

Industry experience

Huawei Technologies India Pvt. Ltd System Architect, *Dec 2018- present*

- Designed the logical replication module which will be used in FileSystem to allow data replication between different FileSystem and Amazon S3.

- Coding of distributed manager for replication of data between Huawei distributed NAS storage clusters particularly focusing on functionality and performance of I/O path.

IQLECT Software Solutions Pvt. Ltd

Data Scientist,

Aug 2018- Nov 2018

Designed and implemented session based recommendation system using Gated Recurrent Unit for e-commerce apps.

NetApp India Pvt. Ltd

Member of Technical Staff II,

July 2017- July 2018

Focused on backup of data from NetApp storage servers to backup devices such as tape using NDMP.

Tata Consultancy Services Pvt. Ltd

System Engineer,

Dec 2012- June 2015

Worked on Java client-server applications dealing with insurance policies.

Teaching Experience

Teaching Assistant

Department of Computer Science, Indian Institute of Technology Kanpur

- Multiagent Systems: Games, Algorithms, Evolution (CS 785) Spring 2017
- Design and Analysis of Algorithms (CS 602A) Fall 2016

Achievements and Awards

- Ranked in top 0.3% (amongst 1,15,425 candidates) in Graduate Aptitude Test in Engineering(GATE), 2015, Computer Science.
- Secured 2nd rank in Computer Science and Engineering Department in Bachelor of Technology, 2012.

Education

Indian Institute of Technology Kanpur

M.Tech. Computer Science and Engineering (CSE)
2015-2017

GPA: 8.33/10.00

Kalasalingam University, Krishnankoil

B.Tech.(Honours) Computer Science and Engineering (CSE)
2008-2012

GPA: 9.27/10.00

Skills

- **Programming Languages:** C, C++, Java, Python
- **ML frameworks and softwares:** PyTorch, numpy, scikit-learn, pandas