Sk Mainul Islam

mainuliitkgp.github.io, skmainul, mainuliitkgp

Email : mainul.islam@iitkgp.ac.in

Mobile : +91-963-5583-571

Research Interests

I am broadly interested in Natural Language Processing and Deep Learning on Graphs. My recent research has focused on aspect based sentiment analysis, interpretability in neural NLP models, controlled text generation, concept mining and debiasing of language models.

Education

- **IIT Kharagpur** — Kharagpur, India
  M.S. by Research Student, Computer Science & Engineering — January 2019 – Present
  Supervisor: Sourangshu Bhattacharya; **Research Areas:** Interpretability, NLP
  Thesis: 'Explainable and Knowledge Enhanced Aspect Level Sentiment Classification'.
- **Maulana Abul Kalam Azad University of Technology** — Kolkata, India
  B.Tech in Computer Science & Engineering; CGPA: 9.13 — July 2014 – May 2018

Publications

- **AR-BERT: Aspect-relation enhanced Aspect-level Sentiment Classification with Multi-modal Explanations**: Sk Mainul Islam and Sourangshu Bhattacharya, TheWebConf (formerly WWW) 2022. [Link]
- **Fair Data Generation using Language Models with Hard Constraints**: Sk Mainul Islam, Abhinav Nagpal, Balaji Ganesan and Pranay Kumar Lohia, CtrlGen Workshop at NeurIPS 2021. [Link]
- **Learnings from Relation Extraction Approaches for Ecommerce**: Sk Mainul Islam and Lucky Dhakad, Convergence-2022. [Flipkart Internal Conference] [Link]

Experience

- **Flipkart** — Bangalore, India
  Summer Intern (Remote), Mentor: Lucky Dhakad — July 2021 - September 2021
  - Clustering based relation surface form extraction between two e-commerce concepts.
  - Adapted a BERT-based self-supervised model to extract relation surface form from internal product review data in lifestyle domain.
  - Mapped relation surface forms to existing relations defined in the Wikidata knowledge graph.
- **IBM Research India** — Bangalore, India
  GRM Intern (Remote), Mentor: Balaji Ganesan — November 2020 - June 2021
  - Built a fair and hard constraint based language model to generate personal data entities in unstructured text. [Demo]
  - Behavioural, fairness and adversarial testing verifies our proposed models' effectiveness in fair augmentation of similar entity in an adversarial setting. Our work got accepted at CtrlGen 2021 (NeurIPS -W).
- **Sponsored Research and Industrial Consultancy, IIT Kharagpur** — Kharagpur, India
  Research Assistant — October 2018 - October 2020
  - Project: "Learning Representations from Network Data", sponsored by Intel Technology India Pvt. Ltd.
  - Principal Investigator: Prof. Niloy Ganguly, IIT Kharagpur.
- **FOSSEE, IIT Bombay** — Mumbai, India
  Software Developer Intern — Summer 2017
  - Developed a web app to port the core functionalities of Xcos (a graphic simulator) to a browser-only version that can be used without installing additional plug-ins or software on the cloud.

Research Projects

- **Explainable and Knowledge Enhanced Aspect Level Sentiment Classification** — IIT Kharagpur
  MS Thesis — June 2019 - Present
  - In this work we introduce a scalable two-level aspect relation information from knowledge graph in the Aspect Level Sentiment Classification (ALSC) task and achieve state-of-the-art performance on the ALSC task.
  - This knowledge incorporation is controlled by an end-to-end incorrect aspect disambiguation detection algorithm.
  - We also introduce a novel post-hoc multi-modal explanation generation to verify the effectiveness of the knowledge incorporation into the ALSC task. The majority portion of this thesis is based on the **WebConf** paper.
- **Dynamic High-level E-commerce Concept Mining** — IIT Kharagpur, Flipkart
  Academic Collaboration Project — February 2021 - Present
  - Recent methods on e-commerce concept extraction fail to capture long-tail and composite concepts to properly understand the user needs.
  - In this work we introduce a constrained soft clustering based concept learning model to extract composite and non-trivial concepts. (work in progress)
- **Debiasing language models with smart prompt design** — Stanford University
  Fatima Fellowship Project, Mentors: Abubakar Abid, Firas Abuzaid — April 2021 - Dec 2021
  - We estimated different attribute specific bias space by controlled swapping of attribute words and contrastive-PCA based estimation to retain user specified desirable bias.
  - We focused on post-hoc language model debiasing by subtracting attribute specific bias space from the sentence representation. This work motivated us to introduce a repository for debiasing auto-regressive language models available in Hugging Face. [Link]
- **Retrieval of Derivational Nouns using Graph Based Learning** — IIT Kharagpur
  Information Retrieval Course Project — Spring 2019
  - In this project we studied the derivational morphology in English and validation of transformations by means of affixation or internal modification to retrieve derivational nouns.
  - We construct a transformation graph where the nodes are source and transformed word pairs and edges are different types of transformations.
  - We built a RGCN based retrieval model to infer whether any source and transformed word pair is valid or not.

Academic Services

- **Teaching Assistant**: Foundations of Algorithm Design & Machine Learning (Spring '22, '21), Computing Lab (Autumn '21), Scalable Data Mining (Autumn '20).
- **Reviewer**: *SEM 2022 (invited), EMNLP 2021, NAACL 2021, EACL 2021, *SEM 2021, EMNLP 2020.

Computer Skills

Languages: Python, C/C++, Java, Spark/Scala, MATLAB
Toolboxes / Frameworks: Pytorch, Tensorflow, scikit-learn, Pandas, numpy, nltk

Coursework

- Deep Learning, Machine Learning, Scalable Data Mining, Information Retrieval, Algorithm Design

OTHER ACTIVITIES

- Selected to attend The Cornell, Maryland, Max Planck Pre-doctoral Research School 2022 with full funding.
- Active member of GEM BENCHMARK, a community driven platform aiming to improve evaluation of natural language generation.
- One of the mentees at ACL Year Round Mentorship program, class of 2022 ACL Mentorship.
- One of the official micro-bloggers at ACL 2021.
- Participated in Amazon Research Day, 2019 at Amazon India, Bangalore.