

Pranav Agrawal

✉ cs.pranav.a@gmail.com | 🏠 pranav-a.github.io | 📁 pranav-ust | 🗣️ pranav_nlp

Experience

Dayta AI

Hong Kong

MACHINE LEARNING LEAD

Sep 2018 - Jun 2024

- Responsible for leading Computer Vision and NLP tasks and designing experiments
- Developed question answering systems for Traditional Chinese based text corpora.
- Developed script conversion methods for conversions between Traditional Chinese and Simplified Chinese.
- Developed paraphrasing systems for Traditional Chinese based news texts

Education

Hong Kong University of Science and Technology

Hong Kong

MASTERS IN BIG DATA TECHNOLOGY

Sep. 2017 - Sep. 2018

- Awarded entrance scholarship for top performance.
- Top 15% of the class. *Detailed list of exams and projects are available here (Link)*

VIT University

India

BACHELORS IN COMPUTER SCIENCE AND ENGINEERING

Sep. 2013 - Sep. 2017

- Consistently awarded high performing awards in academics
- THESIS: Exploration of Heuristics in IR in Applications of Approximate String Similarity (Link)

Conference Publications

Comparing Static and Contextual Distributional Semantic Models on Intrinsic Tasks: An Evaluation on Mandarin Chinese Datasets

LREC-COLING 2024

- First authored paper with Yan Cong, Emmanuele Chersoni, Yu-Yin Hsu and Alessandro Lenci.
- Analyzes the performance of static vs. contextualized models on distributional semantic models on Chinese datasets.

Queer In AI: A Case Study in Community-Led Participatory AI

FACCT 2023, PARTICIPATORY WORK WITH QUEER IN AI

- Won the best paper award at FACCT 2023
- Led the work on conference inclusivity and trans-inclusive publishing

2kenize: Tying Subword Sequences for Chinese Script Conversion

ACL 2020, ARXIV LINK

- First authored paper published in ACL 2020
- Supervised by Dr. Isabelle Augenstein
- Work done on increasing accessibility to Traditional Chinese and Cantonese.
- Developed script conversion methods for converting between Chinese scripts.
- Developed state of the art classifiers for topic classification in Chinese NLP.

Workshop publications, preprints and projects

Check out my github repository *pranav-ust* for more projects and resources.

Making Conferences Queer Friendly

WIDENING NLP 2021, [LINK](#), WORK PUBLISHED WITH QUEER IN AI

- Widening NLP 2021
- Outlining best practices for making conferences queer-friendly, which include inclusive designs of registration forms and virtual spaces, strengthening queer diversity and developing trans-inclusive publication workflow.

Learn Languages First and Then Convert: Towards Effective Simplified to Traditional Chinese Conversion

NAACL STUDENT RESEARCH WORKSHOP 2019, NON-ARCHIVAL TRACK

- Version of ACL 2020 paper, where I developed script conversion methods for converting between Chinese scripts.

Alignment Analysis of Sequential Segmentation of Lexicons

ACL STUDENT RESEARCH WORKSHOP 2018, [PUBLICATION LINK](#)

- Work based off undergrad thesis.
- Retrieving similar cognates in cross lingual applications using subword tokenization and graph similarity
- Published in ACL-SRW 2018

Bit Partitioning Schemes for Multicell Zero-Forcing Coordinated Beamforming

ICIIS 2017, [ARXIV LINK](#)

- Awarded best undergrad paper
- Optimization methods developed for bit partitioning in network cells.

Inflo: News Categorization and Keyphrase Extraction in an Aggregation System

[ARXIV LINK](#)

- Automatic keyphrase extraction and topic classification for news articles. Developed novel algorithms for implementation.

Exploration of Proximity Heuristics in Length Normalization

[ARXIV LINK](#)

- Exploring heuristics such as length similarity for retrieval systems.

BREAK, MAKE and TAKE: An Information Retrieval Approach

[ARXIV LINK](#)

- Work based off undergrad thesis.
- Work on fast string retrieval methods using subword tokenization and graphical similarity.

Keyword Extraction and Classification in Scientific Documents

[REPOSITORY LINK](#)

- Usage of SciBERT to identify keyword boundaries and classify them onto scientific terminologies.

Security Analysis of Bitcoin Transactions through Statistical Graph Annotation

[REPOSITORY LINK](#)

- Work done for Blockchain course in HKUST
- Developed transaction parsing algorithms. This is used for fraud detection using graphical flow methods.

Volunteer Experience and D&I advocacy _____

NAACL 2022

D&I ADVISOR

- Leading and advising on diversity and inclusion matters at NAACL 2022

NAACL 2021

SENIOR D&I CHAIR

- Highly prestigious position in a top-tier NLP conference
- Implemented several D&I measures including financial access, D&I awareness programs, established forums for research and affinity group socials, worked on increasing accessibility for conferences.
- Coordinated and helped in organizing various affinity groups including Black in AI, Queer in AI, Masakhane and North Africans in NLP.

Queer in AI

CORE ORGANIZER

- Working on D&I and safety initiatives
- Helped in organizing many socials in top-tier conferences for facilitating discussions and cohort forming

Skills

Proficient in Python, \LaTeX , Pytorch, Bash Scripting, Git, Cloud Based Systems (AWS and GCP)

Familiar with R, Tensorflow, MySQL, C++, PySpark, Data Visualization platforms (Tableau, d3.js)

Research Interests Natural Language Processing, Ethics in AI, Computer Vision

Awards

- 2023 **Best Paper Award**, FAccT 2023 for Queer in AI
- 2020 **Financial Award**, ACL 2020 scholarship for research presentation
- 2019 **Financial Award**, NAACL 2019 scholarship for research presentation
- 2018 **Financial Award and Travel Grants**, ACL 2018 scholarship for research presentation
- 2017 **Entrance Scholarship**, HKUST, for outstanding academic performance
- 2017 **Best Undergrad Paper**, ICIIS for outstanding undergrad presentation
- 2017 **Best Thesis Award**, VIT University, for outstanding thesis
- 2013-17 **High GPA awards**, VIT University, for outstanding academic performance

Program Committees and Service

- 2023 **Reviewer**, Pacific Asia Conference on Language, Information and Computation 2023
- 2024 **Reviewer**, LREC-COLING 2024
- 2023 **Organizer**, Pacific Asia Conference on Language, Information and Computation 2023
- 2022 **Organizer**, Queer in AI Workshop at NAACL 2022
- 2022 **Chair**, D&I chair at NAACL 2022
- 2021 **Organizer**, Queer in AI Workshop at NAACL 2021
- 2022 **Chair**, D&I chair at NAACL 2021
- 2021 **Journal Reviewer**, Computer Speech and Language
- 2020 **Session Chair**, AACL Student Research Workshop 2020
- 2020 **Event Organizer at COLING**, D&I in NLP
- 2020 **Organizer**, SIGTYP 2020
- 2020 **Mentor**, Widening NLP 2020
- 2020 **Organizer**, Queer in AI at EMNLP 2020
- 2020 **Organizer**, Queer in AI at ACL 2020
- 2019 **Reviewer**, NAACL Student Workshop 2019
- 2020 **Reviewer**, AACL Student Workshop 2020
- 2019 **Reviewer**, ACL Student Workshop 2019