Address : UTA 5.410, 1616 Guadalupe St
University of Texas at Austin, TX 78701, U.S.
Email : ajayjaiswal@utexas.edu
Homepage : https://ajay1994.github.io/
Google Scholar : https://shorturl.at/duzBM
Contact No : +1-(512)-781-4993

Ajay Kumar Jaiswal

PhD Student, Information Science, UT Austin

Education

University of Texas at Austin, TX, United States
Ph.D., Information Science (Machine Learning in Healthcare) *January' 21 (Started)*
GPA: 4.00/4.00 (Overall)

Indian Institute of Technology Kharagpur, West Bengal, India
Master of Technology, Computer Science *April' 15 - May' 17*
GPA: 9.02/10

Interests

Medical Imaging, Multi-Modal Systems, Model Interpretability and Optimization

Publications

Ajay Jaiswal, Tianhao Li, Cyprian Zander, Yan Han, Justin Rousseau, Yifan Peng, Ying Ding
SCALP - Supervised Contrastive Learning for Cardiopulmonary Disease Classification and Localization in Chest X-rays using Patient Metadata.
Submitted to *21st IEEE International Conference on Data Mining (ICDM 2021)*

Ajay Jaiswal, Liyan Tang, Meheli Ghosh, Justin F Rousseau, Yifan Peng, and Ying Ding
Factually-Aware Contrastive Learning For Radiology Report Classification.
Submitted to *21st IEEE International Conference on Data Mining (ICDM 2021)*

Yan Han, Chongyan Chen, Liyan Tang, Mingquan Lin, **Ajay Jaiswal**, Ying Ding, Yifan Peng
Using Radiomics as Prior Knowledge for Thorax Disease Classification and Localization in Chest X-rays. To appear in *American Medical Informatics Association Annual Symposium 2021*

Ajay Jaiswal, Meijun Liu, and Ying Ding
Understanding Parachuting Collaboration. Published at *iConference 2021*

Mayank Raj*, **Ajay Jaiswal***, Ankita Gupta, Sudeep Kumar Sahoo, Rohit R.R, Vertika Srivastava, Yeon Hyang Kim. Be Reasonable : Exploiting large-scale models for commonsense reasoning. *In Proceedings of International Workshop on Semantic Evaluation, Barcelona, Spain (SemEval-2020)*

Vertika Srivastava, Sudeep Kumar Sahoo, Yeon Hyang Kim, Rohit R.R, Mayank Raj, **Ajay Jaiswal**. Ensemble Architecture for Fine-Tuned Propaganda Detection in News Articles. *In Proceedings of International Workshop on Semantic Evaluation, Barcelona, Spain (SemEval-2020)*

Mayank Singh, **Ajay Jaiswal**, Priya Shree, Arindam Pal, Animesh Mukherjee, and Pawan Goyal. Understanding the Impact of Early Citers on Long-Term Scientific Impact. *In Proceedings of ACM/IEEE Joint Conference on Digital Libraries, Toronto, Ontario, Canada (JCDL 2017)*

US Patents

Novel Method for Long Range, Non-contact Body Temperature Estimation Using Near-Infrared and RGBD Camera. *Patent Application SRIB-20190211-005, filed on March, 2019. [Patent Issued] owned by Samsung Electronics-Bangalore.*

Awards & Achievements

Winner of the AI Health Data Challenge at UT Austin. (First Prize - $1000)
Recipient of Kilgarlin Fellowship ($2,500/semester - 2021 to 2024)
Recipient of William and Margaret Kilgarlin Endowed Scholarship ($54,750 - 2021 to 2022)
Promoted to Lead Engineer (ML Research) for exceptional performance at Samsung Electronics.
Secured 99.99 percentile GATE Computer Science, 2015 and received "M.Tech Fellowship" by Government of India

Work Expperience

[Lead Engineer] Advanced Technology Lab, Samsung Electronics, Bangalore, India
Supervisor : Dr. Vijay Narayan Tiwari and Dr. Aloknath De *Aug '17 - Dec '20*

- **Cognitive Vigilance Team:** Responsible for formulating and scoping the solution *News Content Quality and Authenticity Analyser* for Samsung Browser under proof of concept. Developed modules and prototypes using state-of-the-art DL/ML algorithms that evaluate content quality such as article bias, sensationalism, logical fallacy, etc. I also worked with testing partners for rolling out large scale user-trails for estimating system accuracy.
- **Input Intelligence Team:** Responsible for building on-device deep learning models for Samsung Keyboard – swipe-based input and suggestive text. These keyboard solutions get pushed to millions of flagship devices (Samsung Galaxy 9/10 and Note 9/10) that contribute to the revenue of Samsung. I manage Global Ownership of swipe-based continuous input for 13 languages with specialized decoding techniques that can run on Samsung smartphones.
- **Connected Health And Fitness Team:** Developed a data-driven, personalized Actionable Health Insights Engine which provides cognitive actionable health insights by analyzing food, sleep and exercise records of 65 million Samsung Health users. This engine provides meaningful actionable insights and uses real-time data to make dynamic suggestions that help people maintain their daily and weekly progress.

Research Projects

RadBERT-CL: Factually-aware Contrastive Learning Framework for Radiology Report Classification
Supervisor : Prof. Ying Ding *Jan' 21 - Present*

- We propose three novel data augmentation techniques which retain factual and critical medical concepts, identified by a semi-rule based Info-Preservation Module, while generating positive and negative keys for contrastive learning.
- We show that our model RadBERT-CL is able to learn and distinguish fine-grained medical concepts in latent space, which cannot be captured by SOTA pre-trained models like BERT, and BlueBert.
- RadBERT-CL pperformance when a few data labels are available for training and shows that it significantly outperforms baselines by 6-11% improvements in disease classification task.

SCALP - Supervised Contrastive Learning for Cardiopulmonary Disease Classification and Localization in Chest X-rays using Patient Metadata
Supervisor : Prof. Ying Ding *Jan' 21 - Present*

- We propose a novel augmentation technique for contrastive learning utilizes both patient metadata and supervised disease labels to generate clinically accurate positive and negative keys.
- We designed a novel single-staged unified framework to simultaneously improve cardiopulmonary diseases classification and localization by going beyond the conventional two-staged training for CL.
- With SCALP, the average disease classification AUCs improve from 82.8% to 83.9%, while the localization results improve on average by 3.7% over different IoU thresholds in comparable baselines.

Understanding the Impact of Early Citers on Long-Term Scientific Impact
Supervisors : Dr. Animesh Mukherjee and Dr. Pawan Goyal *Mar' 16 - May' 17*

- Proposed an interesting new dimension to the challenging problem of predicting long-term scientific impact (LTSI) usually measured by the number of citations accumulated by a paper in the long-term.
- Demonstrated for the first time, the impact of early citers (EC) on the LSTI of a paper.
- In contrast to popular perception, we find that influential EC negatively affects LSTI possibly owing to attention-stealing. Incorporating EC properties in the state-of-the-art supervised citation prediction models leads to high-performance margins.
- Dataset Used : 2.4 million computer science articles involving 1.2 million authors crawled from Microsoft Academic Research.

Other Academic Projects

Predicting Long Term Citation Profile of Scientific Documents using Stratified Learning
Supervisor : Dr. Animesh Mukherjee and Dr. Pawan Goyal *Dec' 15 - May' 16*

- Implemented a heuristic approach to classify the citation patterns of scientific articles.
- Developed a stratified learning approach (two-staged prediction) for prediction of long term citation count of scientific articles.
- In the first stage, the classification model developed will learn author, venue and content-centric features to classify a query paper in six identified categories. In the second stage, the regression module will predict the future citation count of query paper, utilizing category relevant features.
- Dataset Used : 2.4 million computer science articles involving 1.2 million authors crawled from Microsoft Academic Research.

Abstractive Summarization of Product Reviews Using Discourse Structure
Supervisor : Dr. Pawan Goyal *Aug' 15 - Dec' 15*

- Implemented an abstractive summarization system for product reviews, without using any prior knowledge, utilizing discourse structure of reviews.
- Developed an aspect discourse graph through the aggregation of aspect discourse trees (modified discourse tree such that leaves carry aspect words) and extracted subgraph constituting important aspects and rhetorical relation.
- Implemented a template-based NLG framework (product independent) to generate summary that can convey distribution of opinions.
- Dataset Used: 50 thousand reviews crawled from Flipkart (Indian e-commerce giant) across various product categories.

Extractive Summarization of Documents using Complex System Approaches
Supervisor : Dr. Pawan Goyal *Jan' 16 - Apr' 16*

- Implemented a stochastic graph-based technique for estimation of relative importance of textual units (sentences) of a document, represented as nodes of the graph.
- Implemented a novel method of staged clustering to capture cohesiveness and coherence in the provided query document.
- Relative importance is estimated based on eigenvector centrality in the graph representation of sentences.

Computer Skills

Languages: C, C++, Python, Bash, JAVA
Research Tools: PyTorch, PyTorch-Geometric, TensorFlow(Basics), NetworkX