

# Haoyuan Li

[haoyuanl@ad.unc.edu](mailto:haoyuanl@ad.unc.edu) | <https://scholar.google.com/citations?hl=en&user=BdfQ44MAAAAJ>

## EDUCATION

<b>Renmin University of China, Beijing, CN</b> <i>Major: Statistics, Bachelor of Science, School of Statistics</i> • Cumulative GPA: 3.74/4.0	Sep.2016–July.2020
<b>Harvard University, Cambridge, US</b> <i>Major: Health Data Science, Master of Science, School of Public Health</i> • Cumulative GPA: 3.79/4.0	Sep.2020-Jan.2022
<b>University of North Carolina at Chapel Hill, Chapel Hill, US</b> <i>Major: Computer Science, Doctor of Philosophy, Department of Computer Science</i> • Advisor: Snigdha Chaturvedi	Sep.2022-Present

## INTERNSHIP

<b>Stanford University</b> • Position: Research Assistant	Jun.2019-Sep.2019 Advisor: Ram Rajagopal
<b>Yale University</b> • Position: Research Assistant	Jun.2020-Sep.2020 Advisor: Dragomir R. Radev
<b>Microsoft Research Asia</b> • Position: Research Assistant	Jan.2021-July.2021 Advisor: He Di
<b>Dana-Farber Cancer Institute</b> • Position: Data Science Trainee	Sep.2021-Dec.2021 Advisor: Jacob Rosenthal

## PUBLICATIONS

Wang, Z.\*, **Haoyuan Li\***, Rajagopal, R., (2019) Urban2Vec: Incorporating Street View Imagery and POIs for Multi-Modal Urban Neighborhood Embedding, AAAI 2020, (\*contribute equally)

Fabbri, Alexander R., Simeng Han, **Haoyuan Li**, Haoran Li, Marjan Ghazvininejad, Shafiq Joty, Dragomir Radev, and Yashar Mehdad. "Improving Zero and Few-Shot Abstractive Summarization with Intermediate Fine-tuning and Data Augmentation.", NAACL 2021

Bi, Qiwei, **Haoyuan Li**, and Hanfang Yang. 2021. "Boosting Few-Shot Abstractive Summarization with Auxiliary Tasks.", CIKM 2021

Qiwei Bi, **Haoyuan Li**, Kun Lu, Hanfang Yang, Augmented Abstractive Summarization with Document-Level Semantic Graph, J. data sci. 19(2021)

**Haoyuan Li**, Somnath Basu Roy Chowdhury, Snigdha Chaturvedi, 'Aspect-aware Unsupervised Extractive Opinion Summarization', ACL 2023 Findings

## RESEARCH & PROJECTS

### Urban2Vec: Incorporating Street View Imagery and POIs for Multi-Modal Urban Neighborhood Embedding

Advisor: Ram Rajagopal, Stanford University Jun.2019-Sep.2019

Summary: We develop unsupervised methods to generate the embedding of neighborhoods based on the street view image and POIs. In the project, I explore the way to combine image and text and tune parameters and evaluate its performance.

### Augmented Abstractive Summarization with Document-Level Semantic Graph

Advisor: Hanfang Yang, Asso Prof. School of Statistics, Renmin University of China

May.2020-Jun.2020

Summary: We generate a graph of entities in each document using distant supervision. Then, we apply GNN to generate the entity embedding and insert an attention sub-layer in each decoder layer to incorporate entity embedding into generation. In the project, I tune the hyper-parameter and explore the variant of generating the graph.

### Improving Zero and Few-Shot Abstractive Summarization with Intermediate Fine-tuning and Data Augmentation

Advisor: Dragomir R. Radev, Yale University

Jun.2020-Sep.2020

Summary: We generate pseudo training pair from the wiki dataset based on the characteristics of the target dataset. We also introduce back-translation and consistency loss to further improve the performance. In the project, I explore the possibility of using consistency loss between original documents and augmented documents given pseudo-target summaries.

### Boosting few-shot abstractive summarization with auxiliary tasks

Advisor: Hanfang Yang, Renmin University of China

Nov.2020-Feb.2021

Summary: We develop 3 auxiliary tasks, sentence extraction, object prediction and triple entailment and trained the abstractive summarization model with these tasks in a multi-task learning framework. In this project, I help build BART model for abstractive summarization, introduce adapter for multi-task learning, run experiment on xsum and write the paper.

### Aspect-aware Unsupervised Extractive Opinion Summarization

Advisor: Snigdha Chaturvedi, University of North Carolina at Chapel Hill

Summary: We introduce a method for unsupervised extractive opinion summarization that automatically identifies the aspects described in the review sentences and then extracts sentences based on their aspects. It identifies the underlying aspects of the review sentences using the roots of noun phrases and adjectives appearing in them.

## HONORS

Scholarship for Academic Achievement, 2017,2019(**Top 20%**), Prize of National English Contest of University Students,2017(**Top 1%**), First Prize of China Undergraduate Mathematical Contest in Modeling, 2018(**Top 1%**)