

# Linsen Li

Email: lli23@tulane.edu Tel: +5513238898  
Contact: 4025 W Napoleon Ave, Metairie, LA 70001

## EDUCATION BACKGROUND

### Tulane University

*Doctor of Philosophy in Computer Science*

### Stevens Institute of Technology

*Master of Science in Data science*

### Southern University of Science and Technology

*Bachelor of Science in Statistics*

### The University of California, Berkeley

*2018 Summer Session*

**Sept. 2021-May. 2026(Expected)**

*New Orleans, Louisiana, U.S.*

**Sept. 2019-Jun. 2021**

*Hoboken, New Jersey, U.S.*

**Sept. 2015-Jun. 2019**

*Shenzhen, Guangdong, China*

**Jun. 2018-Aug. 2018**

*Berkeley, CA, USA*

## RESEARCH EXPERIENCE

### Tulane University

*TAPI (Text Analytics for Public Interest) Lab in Computer Science Department*

Advisor: Aron Culotta, Nicholas Mattei

**Sep. 2021 - Present**

*New Orleans, Louisiana, U.S.*

## RESEARCH PROJECT

**Using Text-Based Causal Inference to Disentangle Factors Influencing Online Review Ratings** **June. 2023 - Aug. 2024**

*Advisor: Aron Culotta, Nicholas Mattei*

*Submitted to AAAI 2024*

- Enhanced CausalBERT to isolate individual review aspects' effects on overall perception
- Validated the approach on over 600K U.S. K-12 school reviews. Identified key drivers of overall school ratings, such as perceptions of administration and test performance.

**Online Reviews Are Leading Indicators of Changes in K-12 School Attributes(Paper link)**

**Jan. 2022 - Oct.2022**

*Advisor: Aron Culotta, Douglas M. Harris, Nicholas Mattei*

*Accepted by WWW 2023*

- Applied language processing models over 300K reviews of 70K U.S. schools from a popular ratings website to predict whether schools will significantly increase or decrease in an attribute of interest over a future time horizon. Attributes included test scores and the socio-demographic makeup of a school.
- Found that using the text improves predictive performance significantly over a baseline model that does not include text but only the historical time series of the indicators themselves, suggesting that the review text carries predictive power.

**Tweet Sentiment Extraction on Kaggle competition (NLP) (Report link)**

**Jan. 2020 - June 2020**

- Investigating text classification and text summary problems on Kaggle, with Multi-channel CNN, Stacked LSTM, and attention using Keras. We designed a new deep neural network model based on the idea of an autoencoder for text summary problems.
- Experiment on our deep learning models reached a significantly higher BLEU score and model accuracy.

**Gaming Industry - STEAM and NLP Applications (Report link)**

**Sep. 2019 - Dec. 2019**

- Led a team to scrape game reviews from Steam using Python and BeautifulSoup.
- Implement topic analysis models for information extraction based on the Latent Dirichlet allocation (LDA) model and the Biterm topic model (BTM). Developed a system classifying reviews by topic for industry stakeholders.

**Undergraduate Thesis: Research in tracking diseases (Report Link)**

**Aug. 2018 - May. 2019**

- Worked on a real word classification problem of Disease(Low birth weight) prediction.
- Applied the linear regression, logistic regression, SVM, and Naïve Bayes on this problem. Evaluated by the confusion matrix and ROC curve to compare different models.
- Enhanced the performance prediction of disease by using bagging and boosting for test sets using the Scikit-learn library

## PUBLICATIONS

- Linsen Li, Aron Culotta, Douglas M. Harris and Nicholas Mattei.** "Online Reviews Are Leading Indicators of Changes in K-12 School Attributes". In Proceedings of The Web Conference(TheWebConf) 2023. (Acceptance Rate : 19.2%)

## EXPERIENCE

**Teaching Assistant, CMPS 3140/6140 Intro to Artificial Intelligence**

*Instructor: Prof. Aron Culotta, Tulane University*

**New Orleans, Louisiana, U.S.**

*Jan. 2022 - May. 2022*

**Teaching Assistant, CMPS 3240/6240 Introduction to Machine Learning**

*Instructor: Prof. Jihun Hamm, Tulane University*

**New Orleans, Louisiana, U.S.**

*Aug. 2021 - DEC. 2021*

**Teaching Assistant, CMPS 3160/6160 Intro. to Data Science**

*Instructor: Prof. Nicholas Mattei, Tulane University*

**New Orleans, Louisiana, U.S.**

*Aug. 2021 - Dec. 2021*

## SKILLS

- Python, Pytorch, Hugging Face, Matlab, R, L<sup>A</sup>T<sub>E</sub>X, HTML, Hadoop, SQL Server, Oracle, PostgreSQL, Spark