

# YINHONG LIU

07923592505 ◇ yl535@cam.ac.uk

## EDUCATION EXPERIENCE

---

**Language Technology Lab, University of Cambridge, PhD** 2021.10 - present

- Supervisor: Prof. Nigel Collier
- Research area: Natural language generation, Graph-to-Text generation, Machine reasoning generation.
- Scholarship: CSC Cambridge International Scholarship (Chinese Scholarship Council)

**Computer Lab, University of Cambridge, Mphil** 2020 - 2021

- Subject: **Advanced Computer Science, in Natural Language Processing.**
- Grade: **Distinction.**
- Master project: **Learning Functional Distributional Semantics with Visual Data:** Use Visual Genome data to train a novel probabilistic distributional semantic model which represents words as function rather than vector. The framework provides linguistic interpretability.

**Engineering Department, University of Cambridge; BA, MEng** 2014 - 2018

- Subject: **Information Engineering and Computer Science.**
- Grade in Master of Engineering: **Merit degree** with honours.
- Grade in Bachelor of Arts: **1st Class.**
- Final year project: **Statistical enhancement of multi-channel audio data:** Model multi-channel noisy soundtrack as an additive model with multi-variant Gaussian noises and sound signals. Derive a Maximum Likelihood estimator of the underlying signal and a Bayesian approach to estimate the noise parameters with a Wishart distributed prior.

## WORK, RESEARCH EXPERIENCE

---

**Gambit Research; London.** *Full-time quant-analyst* Oct. 2018 - Jul. 2020

- Working on various statistical football teams & games modelings, market price & volume predictions, strategies developing and back-testing.
- Extensive experience in various programming environments, including Python, SQL and BASH language.

**Jigsaw Unintended Bias in Toxicity Classification, Kaggle competition Jun. 2019**

- Got a 225th place out of 3165 teams (top 8%), a bronze medal, as a solo participant.
- Comments that mention certain minority groups, like black people, women and Asians, have higher chance to be toxic. Therefore, challenge exists in classify those comments' toxicity correctly regardless of their bias in the minority groups.
- Applied the latest language models, including LSTM, BERT and GPT-2. Designed a custom loss function based on the evaluation metrics. Model implemented in PyTorch and trained across multiple cloud GPUs.

## PUBLICATIONS

---

Yinhong Liu and Guy Emerson. **Learning Functional Distributional Semantics with Visual Data**. ACL 2022, main conference.

Yinhong Liu and Simon Godsill. **Multi-channel audio statistical restoration**. IEEE-ICSPCC, 2020.

T. Zhang, Y. Jin, S. Sun, Yinhong Liu and Xiaohua Zhang. **Analysis of Impact Factors of Multiscale Entropy**. IEEE-BIBM, 2018.

## PRIZE & EXTRA-CURRICULUM ACTIVITIES

---

PhD scholarship: CSC Cambridge International Scholarship by Chinese Scholarship Council.

Academic scholarship: Caldwell Scholarship by Corpus Christi College, University of Cambridge

Cambridge Chinese Drama Society: Protagonist in a university play, *Our Assassin*; Producer of two university plays *The Robin Murders* and *The Village*.

## LANGUAGE & SKILLS

---

Languages: Chinese Mandarin (native), English (fluent), Chinese Cantonese (fluent).

Programming Languages:

- Inter-mediate Level: Python, C++, SQL, Matlab, Latex.
- Beginner Level: Javascripts, HTML, Objective-C.