

Zhisheng Tang

☎ Contact: 626-464-3284
✉ jasontangzs0@gmail.com

Education And Experience

- 2023 – 202* **Ph.D. Student** University of Southern California, Industrial & Systems Engineering
- 2021 – 2023 **M.S.** University of Southern California, Major in Analytics
- 2016 – 2020 **B.S.** University of Washington, Seattle, Major in Mathematics
- 2020 – 2021 **Mobvoi, Intern**, Beijing, China.
My focus was on pre-trained language models, on both experimentation and development.
- Summer 2018 **Irootech, Intern**, Beijing, China.
I worked on software development at their cloud platform, which provides service for their Machine Learning department.

Publications

- 2023 **Zhisheng Tang**, Mayank Kejriwal. Can Language Models be used in Multistep Commonsense Planning Domains? Submitted to AGI-23.
- 2023 **Zhisheng Tang**, Mayank Kejriwal. A Pilot Evaluation of ChatGPT and DALL-E 2 on Decision Making and Spatial Reasoning.
- 2022 **Zhisheng Tang**, Mayank Kejriwal. Can Language Representation Models Think in Bets? Accepted at Royal Society Open Science.

Skills

- Programing Java, Python, R, SQL, NoSQL, Tensorflow, Pytorch, Spark
- Language Chinese (native), and English (fluent)

Projects

- Chinese GPT-2 Training from scratch, a team including myself collected about 500G of Chinese corpus. After data cleaning, we trained a 1.5B parameter GPT-2 using 8 servers, each containing 8 T100 for about 3 weeks. The result shows basic features like the original GPT-2, such as few-shot learning.

Chinese To develop a singing machine, a team including myself adopted the jukebox frame-
Jukebox work by the OpenAI and trained a Jukebox model from scratch. We introduced a novel way of encoding Chinese words into phonic tokens to encode the pronunciation of Chinese better. The model consists of a three-layer VQ-VAE to encode the sound and a small transformer decoder to encode the lyrics and a big transformer decoder to combine the two. The total model parameter is about 1B.