jcjessecai@gmail.com • 510-298-2102 • https://jcaip.github.io

Jesse Zexi Cai

EXPERIENCE

UCLA NLP
June 2019 - Present

Undergraduate Researcher
- Investigated learning better sentence representations and contrastive unsupervised representation learning.
- Supervised by Professor Kai-Wei Chang.

Blend
September 2017 - September 2018

Machine Learning Engineer
- Deployed a LSTM-RNN model to identify complaints in user comments, improved P/R and accuracy from 70% to 90%.
- Applied deep learning NLP methods to categorize and cluster user reviews.
- Created and maintained scalable lambda event ingestion pipeline with PySpark.

Blend
June 2017 - September 2017

Software Engineering Intern
- Trained an RNN to [predict loan application submission](#) with 90% accuracy.
- Deployed a multi-user data science toolbox with PySpark support.
- Deployed Airflow with CI/CD to standardize analytics and coordinate ETL jobs.

NASA JPL
June 2016 - January 2017

Software Engineering and Computing Systems Intern
- Used Stanford DeepDive and MITIE to create RDF triples for a natural language question answering system.
- Created several APIs/backends with Flask, Docker, and Elasticsearch.

PROJECTS

QuickThoughts
- Pytorch reimplementation of QuickThoughts paper by Logeswaran and Lee.
- Achieves SOTA performance on several downstream classification tasks.
- Theoretical analysis on the effect of context size on performance [here](#).

DistBelief
- Implemented DownpourSGD (asynchronous distributed training) in PyTorch.
- Reduced training time for AlexNet on CIFAR10 by 50 minutes (45%) vs. single-node SGD
- Wrote small message passing framework on top of PyTorch's native distributed point to point communication.

Technical Blog
- Primarily focused on machine learning, mathematics, and computer science
- Documentation of side/work projects, explanations of interesting concepts.

EDUCATION

UCLA
September 2015 - Expected December 2019

B.S. Computer Science, Upper Division GPA: 3.5
- Coursework: ECE 239AS Deep Learning, CS 269 Natural Language Processing, CS 267A Probabilistic Programming & Relational Learning
- Technical Breadth: Mathematics (Linear Algebra, Real Analysis, Optimization)
- ACM AI President (2017). Received ACM Student Chapter Excellence Award