aip.github.io Jesse Cai

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

EXPERIENCE

UCLA NLP June 2019 - Present

Undergraduate ML/NLP Researcher

- Used BERT and contrastive learning to train better sentence embeddings.
- Created new BERT based topic models for short-text topic mining.
- Supervised by Professor Kai-Wei Chang.

Blend

September 2017 - September 2018

Software Engineer - Machine Learning

- Deployed a LSTM-RNN model to identify user complaints about app glitches, improved P/R and accuracy from 70% to 90%.
- Applied deep learning NLP methods to categorize and cluster user reviews.
- Created and maintained lambda architecture event/log ingestion pipeline with PySpark.

Blend

June 2017 - September 2017

Software Engineering Intern

- Trained an RNN to <u>predict loan application submission</u> with 90% accuracy.
- Maintained 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 Intern

- Used DeepDive and MITIE to create RDF triples for a natural language QA 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 <u>here</u>.

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.

PDB Evaluator

- Wrote query parser and first-order lifted inference for probabilistic database queries.
- Used SQLite to achieve up to 4x speedup on inference on decomposable queries.

EDUCATION

UCLA

Expected March 2020

B.S. Computer Science, Upper Division GPA: 3.57

- Coursework: ECE 239 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.