

(Vincent) **Zihan Guo**
5629 Hempstead Road, Pittsburgh, PA, 15217
zihang@andrew.cmu.edu | (412)-265-7671
github.com/gzhami/projects

OBJECTIVE

To obtain Full Time Software Engineer / Data Engineer Positions

EDUCATION

Carnegie Mellon University
Bachelor of Science in Machine Learning
Double Major: Mathematics
Triple Major: Logic and Computation
Dean's List

Pittsburgh, PA
Expected May 2018

COURSEWORK & SKILLS

RELATED COURSES

10-701 Advanced Machine Learning, 11-411 Natural Language Processing, 15-150 Functional Programming, 15-122 Imperative Computation, 36-315 Statistical Visualization, more on LinkedIn: www.linkedin.com/in/zihan-guo

SKILLS

Programming: Python, R, Scala, SQL, Hive, Java, Javascript
Multi-thread Environment: Apache Spark, Hadoop, Scalding Scala

EXPERIENCE

Tumblr, Yahoo (Now Oath)

Software Engineer Intern - Data Engineer @ Content Intelligence Team

New York, NY
2017 Summer

- Used Scalding (Scala) to construct Hadoop Map-reduce jobs for big data extraction.
- Performed model selection and visualization with R ggplot2.
- Built a Random Forest Follow Spam Classifier using Spark (Scala)
- Optimized algorithm and reduced Type 1 error from 30.6% to 0.9%.

Microsoft

Research Assistant - Lean Interactive Theorem Prover | <https://leanprover.github.io/>

Pittsburgh, PA
2016 Summer and Present

- Built probability library for Lean Interactive Theorem Prover
- Organized and documented the combinatorics documentation

CMU Summit on US - CN Innovation and Entrepreneurship

Co-founder and Vice President of Marketing

Pittsburgh, PA
Since 2016

- Invited Dr. Kai-fu Lee and Dr. Jaime Carbonell for A.I. Panel Discussion
- Built marketing team of 35+ members

Carnegie Mellon University

Teaching Assistants

Pittsburgh, PA
Since 2016

- 15-110 Introduction to Computer Science (Python)
- 36-225 Probability Theory (R)
- 36-315 Statistical Visualization (R, ggplot2, Shiny App)

PROJECTS

Capital One - Data Science Competition Final Round

- Reduced the overall error rate to 2.1 percent and mis-classified bad load rate to 82.4 percent. (github.com/gzhami/projects)

Khan Academy - Version Infection Algorithm Simulation

- Implemented a Python simulation for web version infection algorithm using Breadth First Search. (github.com/gzhami/projects)

Hidden Markov Model on PoS(Part of Speech) Tagging

- Built a Hidden Markov Model using Forward, Backward and Viterbi Algorithms to perform part of speech tagging with nearly 0.0% error rate. (github.com/gzhami/projects)

Neural Network Using Back-propagation Algorithm

- Implemented a ANN with one hidden layer to predict academic grade with 99.84% accuracy. (on my Github)