# **Joowon Kim**

5615 Upper Creek Ct, Suwanee, GA 30024 • (678)-938-4939 joowonk2@illinois.edu • github.com/jookimmy • www.joowonkim.me

#### **EDUCATION**

#### University of Illinois at Urbana-Champaign

*May 2021* 

Bachelor of Science, Computer Science and Economics

GPA: 3.4/4.0

#### SKILLS

**Skills:** Python, C++, Javascript, Statistics, Numerical Analysis, Machine Learning, Data Science, Full Stack Development **Technologies:** PyTorch, Pandas, NumPy, Django, PostgreSQL, Linux, AWS, Git, Jupyter Notebook

**Coursework:** Algorithms and Models of Computation, Linear Algebra, Regression Analysis, Statistics, Computer Architecture, Data Science, Machine Learning Artificial Intelligence, Econometrics

#### **EXPERIENCE**

## **Software Engineering Intern,** *Pure Storage*

*May* 2020 – *August* 2020

- ⇒ Constructed data engineering module in Python to convert low-density data to high-density in FlashArrays
- ⇒ Designed a new row compression algorithm that squashes real-time data stream from Pure Storage hardware
- ⇒ Achieves 100% lossless compression and effectively reduces up to 97% on 10,000,000+ rows of information
- ⇒ Reported as primary author for the algorithm, which will outline invention disclosure and patent application

### **Undergraduate Researcher**, ECE Department

September 2020 - Present

- ⇒ Working as an undergraduate researcher in Sanmi Koyejo's lab on machine learning models for cognitive science
- ⇒ Researching relations between cognitive intelligence and brain measures using hierarchical mixture-of-experts
- ⇒ Improving expectation-maximization for MoE-layer's optimal expert to create more robust predictions of data samples

#### Data Science Intern, State Farm

August 2019 – December 2019

- ⇒ Congregated different claims processes based on agglomerative clustering, vector distance, shingling
- ⇒ Solo-built entire LDA-simulation pipeline in Python for team to utilize, improving future research efficiency
- ⇒ Integrated various clustering methods to optimize runtime for >18GB files and ~400 different vector combinations

#### Software Engineering Intern, State Farm

May 2019 – August 2019

- ⇒ Constructed corporate data analytics portal using Javascript, Python, Django, PostgreSQL, JSON, and pandas
- ⇒ Developed analysis tool in software, increasing the efficiency of data search functionality by 100%
- $\Rightarrow$  Redesigned information storage in PostgreSQL database to improve analysis for  $\sim \! 1000$  data-layered maps

#### Director of Project Management, CS 196 at UIUC

August 2018 – December 2019

- ⇒ Directing all 35 project managers for the 210 computer science students in the CS 196 Honors course
- ⇒ Installed new and rigid grading system and hiring/interview methods, while keeping an effective 1:7 student-PM ratio

#### **PROJECTS**

# **NCAA March Madness Prediction**

Present

- ⇒ Utilizing PyTorch, pandas, and NumPy to predict the best performers for the 2020 NCAA Tournament
- ⇒ Implementing multi-layered classification neural network using Kaggle data to identify the result of a matchup

AWS Email Bot July 2019

- ⇒ Built a Python script using libraries like smtplib to develop a subscription-type email bot on AWS EC2 instance
- ⇒ Implemented structural methods supporting constant runtime, making user data processing easier and more efficient

#### **MISCELLANEOUS**

Placed 5<sup>th</sup> Internationally for HOSA Creative Problem-Solving event at ILC Placed 1<sup>st</sup> Nationally for Academic Games League of America in On-Sets event Love playing basketball or golf, losing in fantasy sports, going to social events, and gaming