Wilson Jusuf

Email: wilsonjusuf1998@gmail.com **Github:** github.com/willyspinner **LinkedIn:** linkedin.com/in/wilsonjusuf

EDUCATION

University of California, Los Angeles **(B.S. Computer Science)**

CGPA: 3.80/4.00 Expected graduation: Jun. 2021

Courses: Data Structures & Algorithms, Computer Organization/Architecture, Software Engineering Lab, OS (POSIX), Databases, Distributed Systems, Computer Networks, GIS Systems, IoT, ML & Deep Learning, Computer Vision, Linear Algebra, Stochastic Processes

WORK EXPERIENCE

Paypal Inc. - Software Engineering Intern **Los Angeles, CA** *Jun. 2020 - Sept. 2020*

Worked at the Product Catalog team of Honey, responsible for developing a **scalable, fault-tolerant** pipeline using NodeJS, Docker & Kube, Scala, and Google Dataflow that keeps track of millions of product info affecting all of **20M+** shoppers. Major projects I undertook:

- Created a Dataflow Job that batch-updates product category data at **scale** (more than **40M+**), used for downstream ML initiatives to ensure compact product categorization. This **improves** data quality for **millions of products** in the catalog.
- Led the **GraphQL Federation** migration initiative for our user-facing API, promoting **autonomy/decoupling** from other services.

Mathpix Inc. - Software Engineering Intern - Deep Learning **Santa Monica, CA** *Oct. 2019 - Jun. 2020*

Mathpix develops market-leading Math OCR-as-a-Service that digitizes STEM images into LaTeX. My responsibilities include:

- **Foreign-Languages OCR** - led the OCR of **high-demand foreign** languages such as Korean, Hindi, Chinese, Japanese, Thai, and Viet to expand to **key Asian markets**, ensuring **optimal vocabulary compactness** - **reducing 10k+ tokens** to **400 in the model**.
- **Bootstrapped scalable ML training** for any foreign language by automating generation of realistic language training data, **eliminating** need for **manual labor labeling** by developing scripts that **generate 1M+** training samples.
- **Led Dataset Analytics & Continuous Training** initiative to obtain more **visibility** into model's **weaknesses** using ES & Kibana.

Textpert Inc. - Machine Learning Intern **Los Angeles, CA** *Sept. - Dec. 2018*

- Led the design and fine-tuning of Audio Neural Networks for AIME (Artificial Intelligence Mental Evaluation - a 10-minute Mental Health Test) to analyze depression risk from patients' vocal expression.
- Achieved a **90% testing accuracy** just from raw audio data in Proof of Concept Audio Network, using speech feature extraction.
- Integrated Networks with text and video to ultimately form AIME's engine used in **LA and SD clinics** to measure anxiety risk.

Tokopedia.com - Fullstack Engineering Intern **Jakarta, Indonesia** *Jun. - Sept. 2018*

- Led the **React ES6** Migration Project for Tokopedia's chat service. Tokopedia is Indonesia's largest online marketplace.
- Designed entire project infrastructure and strategized build & deployment with DevOps team.
- Significantly decreased load by achieving a **20% (75 KB) size reduction** for production assets.
- Enabled **80M+** end-users to chat with sellers easily in a new homepage chat feature created with migrated stack.
- Eased the path of frontend development for the chat team of **8 people** with an optimized workflow and stack.

Gojek Indonesia - Machine Learning Intern **Jakarta, Indonesia** *Jul. - Sept. 2017*

- Architected a deep learning model for Gojek (one of the fastest-growing Indonesian on-demand/logistics Decacorns) to conduct sentiment analysis of Indonesian feedback comments to improve customer service.
- Achieved an **unprecedented testing accuracy score of 78%** for sentiment analysis in the Indonesian language by utilizing LSTM cells, Neural Networks, and word2vec embedding.

PROJECTS

Missing Child

- Re-uniting parents and children separated by disaster by using state-of-the-art kinship verification and transfer learning. IBM Call for Code 2019. Achieved a **90.2% kinship verification testing accuracy** on the FIW dataset. (Available on Github)

tepuknyamuk.com

- An online, real-time multiplayer card game I fully architected front-end to back-end. Built with **React, Node, Redis, PostgreSQL, WebSockets, Nginx,** and hosted on **DigitalOcean**. Used **Datadog** for game monitoring. (Available on Github)

sharded Paxos KV store

- A distributed, replicated and sharded fault-tolerant Key-Value Store using the Paxos Consensus Algorithm.

LEADERSHIP & VOLUNTEERING

Cyber Shanty

- Co-founded Cyber-shanty, a digital literacy initiative to teach **a cohort of 15 kids** living in impoverished areas of Jakarta to use computers to study, contact each other, have fun, and empower their lives.

SKILLS

Programming Languages: Go, Python C, C++, JS, Bash, Java, Scala **Frontend**: React ES6, Webpack, Redux, HTML & CSS

Backend & Cloud: Node, PostgreSQL, BigQuery & BigTable, Docker, Redis, Nginx, Datadog, ElasticSearch & Kibana, Linux, AWS, DigitalOcean, GCP

Machine Learning: Pytorch, Keras, Regression & Classification, NNs, RNNs, Computer Vision, CNNs, NLP, Probabilistic Methods

Languages: English, Indonesian, Japanese, Chinese, German, Korean **Interests**: Cooking, Music & Jazz Improvisation, Learning Languages!