John (Jonghyun) Chung

Location: Ithaca, NY
Email: jc2739@cornell.edu
Phone: (607) 379-1488

SUMMARY

Software Engineer with hands-on experience leading the design, development, and deployment of innovative software solutions from scratch. Motivated self-starter focused on working with diverse cross-functional teams to build automated, data-driven services and products that optimize the user experience and maximize performance, quality, stability, and reliability at scale.

SKILLS

- **Data Science & Engineering:** Roadmap Development, Architecture Design & Coding, Quantitative Analysis, ML/DL, Automation, Algorithm Design, Web Development, Data Mining/Cleansing/Modeling, Data Visualization & Dashboard Design, ETL Development & Optimization, Business Intelligence & Analytics, Feature Engineering, Cloud Services
- **Leadership:** Strategic Planning & Execution, Project Management, Prioritization, Process Improvement, Collaboration
- **Software:** Python, Tensorflow, PyTorch, Flask, Django, Java, OCaml, Javascript, React, Next, D3, Firebase, C, Airflow, Docker, Kubernetes, MongoDB, MySQL, Google Analytics, Big Query, FB Ads, MS Office, Notion, Jira, Slack, Git
- **Languages:** English, Korean

EDUCATION & CERTIFICATIONS

- Master of Engineering in Computer Science, Cornell University — Expected December 2022
- Bachelor of Arts in Computer Science, Cornell University — Expected May 2022

RELEVANT EXPERIENCE

Research/Software Engineering Intern, Samsung Research — **June 2020 - August 2020**

- Tested curriculum learning with difficulty metrics, including sentence length and word rarity, during training of neural machine translation models using TensorFlow's seq2seq module to generate a 1.3 increase in BLEU score.
- Applied morphology, subword, and BPE-based tokenizers to Korean and English raw text files and trained transformer-based seq2seq binary language models to test the performance of tokenizers on Ko-En Neural Machine Translation model.

Software Engineering/MLops Intern, Riiid — **May 2021 - August 2021**

- Achieved a 45% decrease in model size and 2/3 decrease in model inference time by applying PyTorch and Onnx quantization to user score prediction models within a language learning app for English exams with 3MM+ global users.
- Managed CI/CD of the AI inference server, A/B test dashboard, and OneFlow repositories and implemented integration test automation with lint and test coverage using Github Action to decrease overall time for model deployment.
- Developed an in-company dashboard for Riiid that displays company product info, market statistics (KPIs), and HR information using Typescript, React.js, Next.js, Vercel, Airflow, and Notion/Asana/Google Sheets API.
- Built an in-company platform that enabled 150 users at Riiid to express opinions on specific parts of online/offline meetings with a web app and helped optimize content for presentations using Typescript, React.js, Next.js, Vercel, and Firebase.

Founder/Data Engineer/Frontend Developer, poner:se — **August 2020 - January 2021**

- Founded an online female clothing store by building, enhancing, and maintaining a web/mobile eCommerce site that generated $1,500 in monthly profits.
- Analyzed user data using Google Analytics/Big Query and conducted multiple A/B tests to optimize product photos, information, and social media content to maximize time spent on landing pages and drive new conversions.
- Developed a product naming algorithm for SEO by combining keywords collected from keyword rankings.
- Built and managed a 5-person team consisting of a developer, designer, marketer, merchandiser, and photographer.

Undergraduate Researcher, Cornell University — **January 2020 - June 2020; August 2021 - Present**

- Conducted research under the supervision of Professor Cardie with a focus on building few-shot named entity recognition models for low resource languages using pre-trained models, including RoBERTa and LUKE, and open-source text datasets.
- Summarized and presented research papers on Attention mechanisms, Transformers, pre-trained models, and other NLP-related topics during weekly meetings and discussions to encourage conversation on relevant research topics.

SELECT PROJECTS

- **Cryptocurrency Price Prediction:** Building a bidirectional LSTM-based Bitcoin price prediction model with 5 years of time-series BTC price data crawled from the web using Python, Tensorflow, and Selenium. Creating a web application to show stock price predictions over the next days, minutes, and hours with real-time updates using Azure ML, Apache Airflow for automatic data collection, and Azure CosmosDB for data storage.
- **Daily Self-Development Helper:** Developing the backend and recommendation system for an application that recommends articles and videos related to user professions based on similar user preferences and user history to drive professional development using Django, PyTorch, Selenium, bs4, Airflow, and AWS.