

# Joseph Mok

hojinmok@gmail.com • (647) 975-5126

Portfolio Website: [https://hjmok.github.io/josephmok\\_portfolio](https://hjmok.github.io/josephmok_portfolio)

Linked In: <https://www.linkedin.com/in/hojinjosephmok/>

GitHub: <https://github.com/hjmok>

---

## Summary of Qualifications

---

- Proficiency in utilizing TensorFlow/Keras, PyTorch, Scikit-Learn, OpenCV, NumPy, and Pandas libraries in Python
  - Programmed various machine learning models including linear/logistic regression, CNN, RNN, NLP models, etc.
  - Well versed in data preprocessing and feature engineering in preparation of machine learning model training
  - Experienced with SCADA and PLC software packages, including Allen Bradley and Ignition Automation Software
  - Experienced integrating MySQL, Python scripting, and Ignition for capturing SCADA trend data and data analysis
  - Strong cross-functional project management experience from leading various multidisciplinary projects
- 

## Professional Experience

---

Yaya Foods Corp.

Toronto, ON

**Senior Automation Engineer**, July 2019 – Present

- Lead the automation design for several SCADA systems for beverage manufacturing clients by integrating Rockwell Automation PLCs/HMIs, Ignition Automation Software, MySQL, and ethernet IP networks
- Developed data mining process by storing transmitter data from PLCs to MySQL database using Ignition SCADA packages. Developed client applications within Ignition for data analysis (heavy Python scripting)
- Gained strong troubleshooting skills for hardware and software to determine root cause of production halts
- Sourced transmitters, motors, VFDs, valves, and necessary hardware for new automation systems

Apple Inc.

Cupertino, CA

**Recycling R&D Engineer Intern**, September 2018 – April 2019

- Managed development of recycling systems for Apple products. Involved heavy coordination with vendors and internal Product Design teams to organize documentation and deadlines
  - Developed optimal cycle time processes for de-manufacturing of various Apple products. Resulted in a 90% increase in the Unit-per-Hour output compared to the current manual methods
  - Designed semi-autonomous prototypes to showcase cycle time and disassembly improvements. Designs continuously integrated feedback to improve operator ergonomics and meet California waste regulations
- 

## Projects

---

The following listed are NLP projects. To see all my projects, please visit: [https://hjmok.github.io/josephmok\\_portfolio](https://hjmok.github.io/josephmok_portfolio)

### Topic Modeling for Question and Article Categories

[https://hjmok.github.io/josephmok\\_portfolio/#/TM](https://hjmok.github.io/josephmok_portfolio/#/TM)

- Used Latent Dirichlet Allocation (LDA) and Non-Negative Matrix Factorization (NMF) methods to form pre-determined number of clusters that acted as assigned topics to a Quora questions and NPR articles dataset
- Created a document term matrix, then fit onto Scikit-Learn's NMF and Latent Dirichlet Allocation imports
- Resulting model was able to assign each article/question to one of 12 topics, which the end user interprets

### TFIDF Text Classification

[https://hjmok.github.io/josephmok\\_portfolio/#/TFE](https://hjmok.github.io/josephmok_portfolio/#/TFE)

- Created a supervised learning model to classify Positive/Negative reviews in an Amazon Reviews dataset and Ham/Spam text messages in an SMS dataset. Data-preprocessing involved removing null rows
- Utilized Scikit-Learn's TfidfVectorizer to Count Vectorize each unique word in the training set, then applied Term Frequency-Inverse Document Frequency feature extraction to said words. Then used Scikit-Learn's LinearSVC (Support Vector Classifier) to return the best fit hyperplane to categorize the data.
- Achieved a 98% accuracy on the SMS dataset and 86% accuracy on the Amazon Reviews Dataset

### Novel Text Generation Model

[https://hjmok.github.io/josephmok\\_portfolio/#/TG](https://hjmok.github.io/josephmok_portfolio/#/TG)

- Used PyTorch to create a deep learning model that uses novels such as Shakespeare and Tom Sawyer as inputs, then outputs texts that match the tone/vocabulary similar to the input novel.
- Encoded every unique word in the novels and prepared them in batches with experimental sequence lengths
- Model utilized LSTM layers and dropout layers, which input texts in batches to help the model understand the grammatical structure of the novels

### Facebook Babi Dataset Chatbot

[https://hjmok.github.io/josephmok\\_portfolio/#/CB](https://hjmok.github.io/josephmok_portfolio/#/CB)

- Created a chatbot by implementing End-to-End Memory Networks and LSTM layers with Keras
- Trained on the Facebook Babi Dataset, which consists of a Story, Question about the story, and Answer. As such, the chatbot took takes a Story and Question as inputs, then outputs the Answer.
- Resulting model achieved close to 95% accuracy on the training data and up to 90% on the test data

---

### Education

---

University of Waterloo,  
**Bachelor of Applied Science, Honours Mechanical Engineering,** Graduated June 2019