Joseph Mok

Email: [*hojinmok@gmail.com*](mailto:hojinmok@gmail.com) GitHub: [*github.com/hjmok*](https://github.com/hjmok)

Portfolio: [*hjmok.github.io/josephmok\_portfolio*](https://hjmok.github.io/josephmok_portfolio) LinkedIn: [*linkedin.com/in/hojinjosephmok*](https://www.linkedin.com/in/hojinjosephmok/)

# Skills

*Languages****:*** Python, JavaScript, MATLAB, C++

*ML Libraries****:*** TensorFlow/Keras, PyTorch, Scikit-Learn, NumPy, Pandas, spaCy, NLTK, OpenCV, PySpark, GCP API

*Databases****:*** MySQL, SQL, HBase, Cassandra, MongoDB

*Big Data Tools****:*** Spark, Hadoop, BigQuery, Hive

*Platforms:* GCP, Databricks, Jupyter Notebook

# Professional Experience

Yaya Foods Corp. *Toronto, ON*

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

* Developed data mining process that extracts PLC transmitter data, stores in MySQL database, and presents it to Client application designed in Python and Ignition Automation software
* Lead the automation design for several SCADA systems for beverage manufacturing clients by integrating Rockwell Automation PLCs/HMIs, Ignition, MySQL, and ethernet IP networks
* Determined root cause of production halts through strong troubleshooting skills for hardware and software

Apple Inc. *Cupertino, CA*

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

* Designed semi-autonomous prototypes for optimal disassembly and recycling of Apple products. Designs continuously integrated feedback to improve operator ergonomics and meet California waste regulations
* Prototype resulted in a 90% increase in the Unit-per-Hour output compared to the current manual methods

# Projects

**Topic Modeling for Question and Article Categories**  [*hjmok.github.io/josephmok\_portfolio/#/TM*](https://hjmok.github.io/josephmok_portfolio/#/TM)

* Used Scikit-Learn to implement Latent Dirichlet Allocation and Non-Negative Matrix Factorization methods to form clusters that acted as assigned topics to a Quora questions and NPR articles dataset
* Resulting model was able to assign each article/question to one of 12 topics, which the end user interprets

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

* Created supervised learning models using Scikit-Learn and PySpark to classify Positive/Negative reviews in an Amazon Reviews dataset and Ham/Spam text messages in an SMS dataset.
* Applied Term Frequency-Inverse Document Frequency feature extraction to analyze the vocabulary
* Achieved an F1-score of 98% on the SMS dataset and 86% on the Amazon Reviews Dataset

**Facebook Babi Dataset Chatbot**  [*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
* Resulting model achieved close to 95% accuracy on the training data and up to 90% on the test data

**Novel Text Generation Model**  [*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.
* Model utilized LSTM layers and dropout layers, which fed texts in batches to help the model understand the grammatical structure of the novels

# Education

University of Waterloo,

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