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

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

*Big Data Tools****:*** Spark, Hadoop, Pig, Hive, Drill

# 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

**Breast Cancer Malignant or Benign Diagnosis**  [*hjmok.github.io/josephmok\_portfolio/#/BC*](https://hjmok.github.io/josephmok_portfolio/#/BC)

* Applied Logistic Regression and K-Nearest Neighbor analysis using Scikit-Learn and PySpark to a Breast cancer classification dataset with 32 columns of patient data
* Logistic Regression model achieved F1-score of 98% and K-Nearest Neighbour achieved F1-score of 97%

**Stock Prices Prediction**  [*hjmok.github.io/josephmok\_portfolio/#/StockPriceRNN*](https://hjmok.github.io/josephmok_portfolio/#/StockPriceRNN)

* Created an RNN with LSTM model on Tensorflow to predict AMD and Google Stock prices by training on daily stock price data from May 2009 to August 2018.
* Resulting models were able to follow trend and scale of stock prices for quarterly and annual sequence sizes

**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 a F1-score of 98% on the SMS dataset and 86% on the Amazon Reviews Dataset

**User Movie Rating Prediction with AutoEncoders**  [*hjmok.github.io/josephmok\_portfolio/#/AE*](https://hjmok.github.io/josephmok_portfolio/#/AE)

* Created a Stacked AutoEncoder model to predict what rating a user will give to a film.
* Trained on a Grouplens dataset with 1 million rows of move ratings from 6040 users across 3952 rows
* Achieved a train loss of 0.808 and test loss of 0.896, meaning every prediction is off by ±1 star

# Education

University of Waterloo,

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