Minnie Nguyen

☎ (438) 764 0121 | ✉ minniexnguyen@gmail.com

🏠 https://nminnie.github.io | ⬛ nminnie | in nminnie

Education

McGill University   *Montreal, QC*

B.Sc. in Mathematics & Computer Science, GPA: 3.73   *Sep 2015 - May 2019*

- Courses: Applied Machine Learning, Database Systems, Probability, Statistics, Stochastic Processes, Algorithms & Data Structures, Software Design
- Awards: Major Entrance Scholarship, Tomlinson Engagement Award for Mentoring (4-time recipient)

Work Experience

Dynamicly Inc. (Startup)   *Montreal, Canada*

Data Scientist Intern   *Jun - Oct 2019*

- Worked on an R&D project in the Data Intelligence team to build a natural language understanding framework that performs intent classification for chatbots.
- Created a data pre-processing pipeline with tokenization using NLTK, lemmatization using spaCy, and spell-checking using pyspellchecker.
- Implemented a flexible, intelligent pattern matching algorithm in regex that allows phrases of variable length and synonym matches using GloVe word embeddings.
- Created a platform for that allows users to easily construct custom-trained word embeddings using Gensim and GloVe.
- Trained a text classification model in Keras with an LSTM and custom-trained word embeddings.

McGill University, Schulich School of Music   *Montreal, Canada*

Student Software Developer   *May - Aug 2018*

- Implemented 8 functionalities for the Interactive Classifier web-based symbol classification software, currently used by musicologists at McGill University and Dalhousie University, using Git, JavaScript, Node.js in a Django framework.
- Analyzed and optimized the runtime performance of zooming and drag-select actions by 90%, using Chrome Developer Tools.

Projects

Book Recommendation System

- kNN model that uses collaborative filtering to recommend the most similar books based on 6 million ratings of 10,000 books on Goodreads.
- Utilized: scikit-learn, SciPy, Pandas, NumPy

Time-series Analysis & Forecasting

- ARIMA time series model that analyzes the trends and seasonality of avocado prices in the last 5 years in the US and predicts future prices.
- Utilized: PyLab, StatsModel, Pandas, NumPy, Matplotlib

Movie Reviews Sentiment

- NB-SVM classification model that uses BoW, bigrams, and NLTK tokenization to classify the sentiment 25,000 IMDb reviews with 91.9% accuracy.
- Utilized: scikit-learn, SciPy, NLTK, Pandas, NumPy, Matplotlib

Handwritten Digit Recognition

- 18-layer ResNet (CNN) for the MNIST handwritten digit dataset with accuracy of 97.5%, implemented with PyTorch and uses OpenCV for noise removal and image transformations for data augmentation.
- Utilized: PyTorch, Keras, OpenCV, scikit-learn, Pandas, NumPy, Matplotlib

Skills

Machine Learning	Classification, Regression, Embeddings, Data pre-processing, Feature engineering, Visualization
Python Modules	scikit-learn, Keras, PyTorch, Gensim, spaCy, NLTK, Pandas, NumPy, Matplotlib
Coding	Python, JavaScript, Java, SQL, R, Matlab
Tools & Frameworks	Git, Node.js, Django, jQuery, HTML5, Android