Isaac McKillen-Godfried

207-249-5188 • igodfried@isaac26.com • isaacmg.github.io • github.com/isaacmg

Experiences

Data/ML Engineer Monster Sept-2019 to Present

• Assisted in GCP data lake creation with Terraform, Pub/Sub, and BigQuery to enable advanced analytics
• Set standards for data science reproducibility best practices and trained coworkers in PyTorch.
• Trained and fine-tuned models in PyTorch to automatically add new triplets to the company knowledge graph.

Data Engineer Hudson's Bay Company (Contract) Jan-2019 to Jun-2019

• Developed and optimized machine learning pipelines to forecast in-store retail demand with SparkMLlib
• Researched NLP techniques in PyTorch/Tensorflow to improve product categorization and personalization

Data Analyst Eastern Maine Medical Center Jun-2016 to Jun-2018

• Refactored and implemented RNNs in Keras/Tensorflow to forecast patient length of stay and researched few-shot detection of conditions in Chest X-Rays with RetinaNet and YOLO2.
• Created interactive visualizations for doctors and administrators with Bokeh, Pandas, and Jupyter Notebooks.

Founder PaddleSoft Jun-2015 to Jun-2017

- Created a neural network (NARX) in MATLAB to predict the flow of the Kenduskeag stream.
- Employed Spark to perform analysis and train NLP algorithms like Word2Vec and LDA on textual datasets.

Research Assistant University of Maine Jun-2012 to Aug-2013

Volunteer Work and Open Source Contributions

Lead Machine Learning Researcher CoronaWhy March-2020 to present

- Trained a relation classification models to automatically classify drug treatment pairs and adverse events from COVID-19 research articles. Created a Docker container with unit tests to ship model to production.
- Leading a cross-disciplinary team of data scientists, epidemiologists, and software engineers to utilize deep learning models (LSTMs/transformers) and transfer learning to forecast COVID-19 spread at county levels.

Flow Forecast Open Source Project August-2020 to present

- Developed transformer, LSTM, and regression models to forecast river flows around the country.
- Created a deep learning for time series forecasting library to effectively track hyper-parameters/experiments.

Skills

- **Languages**: Python (PyTorch, Tensorflow, Keras, Bokeh, Pandas, and scikit-learn), SQL, and Scala
- **Technologies/Platforms**: Docker, Kubernetes, Terraform, BiqQuery, GCP, Wandb
- **Specialties:** Transfer learning, attention, transformers (HuggingFace), NLP (Spacy)

Education

B.A. Hispanic Studies with Minors in Computer Science and NEJS Brandeis University Aug-2013 to May-2017