Josef Affourtit

Cambridge, MA

617.417.9864 | jaffourt@mit.edu | jaffourt | josefaffourtit

Professional Experience

Massachusetts Institute of Technology

May 2019 - Present

Research Associate

Task Representations in Neural Networks

- Led a research project examining task representations of **natural language** and memory in recurrent neural networks (RNN) using multi-task learning; this project addressed previously unexplained functional organization of RNNs in **machine learning** and **AI** research
- Built generative natural language processing models and trained for representation learning in a **PyTorch** codebase
- Implemented statistical analyses on recurrent units in **R** and **Python**, with visualizations using **matplotlib**, **pyplot** and **ggplot**

Compositional Semantics in Natural Language Processing and Human Cognition

- Led a research project investigating compositional semantics of **computational language models** and human semantic representations, leveraging high dimensional **neural data** (fMRI); this project investigated unexplained divergence between human cognition and computational language models, providing an empirical framework for benchmarking
- Tested **natural language processing** model predictions by running interactive online experiments on Amazon **MTurk**, using HTML and JavaScript
- Implemented **signal preprocessing** using gaussian mixture models, artifact detection, and multivariate statistics on neural **time series** data in **MATLAB**
- Created software to perform statistical tests computing complex similarity rankings between semantic classes in **Python** (NumPy, scikit-learn)

Probabilistic Language Atlas & Data Release Web App

- Led a research project investigating the probabilistic hemodynamic activations across human brains during a language task; this project created a foundational tool for researcher's investigations into the neurocognitive correlates of language
- Created software to standardize and preprocess over 300,000 fMRI brain images in two separate pipelines, and applying **linear models** and **time-series** analyses in **Matlab**, **Python**, and **R**
- Designed and deployed RESTful web app on **Amazon Web Services** for interactive data visualization, and open source data distribution in Javascript (**React**), **SQL**, and Python (**Django**)
- Created documentation of data analyses, data processing pipelines, and web app for team use, and maintained distributed version control systems in **Git**

Riverside Research

May 2018 - May 2019

Machine Learning Researcher

Pathway Estimation Using Remotely Sensed Spectral-Terrain Data

- Led a research project which processed high dimensional satellite imagery, and predicted optimal routes to navigate the terrain; this project's novel algorithm was presented at Military Operations Research Symposium, 2019
- Applied novel **signal processing** and **machine learning** methods to high dimensional imagery, such as **Principal Component Analysis**, **Orthogonal Matching Pursuit**, **Automatic Target Generation Procedure**, and **Non-negative Matrix Factorization** in MATLAB
- Implemented search algorithms and **reinforcement learning** to solve an optimization problem for agent-based path estimation within a weighted graph in MATLAB
- Designed and created a **Graphical User Interface** for interactive 3D visualizations at each step of the algorithm in MATLAB

Temperature Dependent Tissue Characterization Using VNIR Imagery

- Assisted in a research project which characterized temperature dependent tissue samples using Hyper Spectral Imagery and machine learning; this project investigated a machine learning approach to automate the Maillard reaction in a mass production environment
- Built and trained a **deep learning** model using **competitive leaky learning** to classify clusters of distinct spectral vectors
- Assessed model accuracy by computing the **Kullback–Leibler divergence**, and Bhattacharyya distance between spectral clusters

Education

Massachusetts Institute of Technology

Cambridge, MA

Professional Certificate Program in Machine Learning & Artificial Intelligence — *2020*

Ohio State University

Columbus, OH

B.S. Physics — *2019*

Technical Skills

Programming **Python** (PyTorch, TensorFlow, scikit-learn, NumPy, SciPy, Pandas, Pyplot), **R** (dplyr, ggplot2), **MATLAB**, **SQL**, **JavaScript** (Node, React), **HTML**, Unix/Linux (Bash)

Other Machine learning and statistics (classification, regression, clustering, hypothesis testing, model fitting), high performance computing (slurm), data analysis, data visualization, REST API development

Honors & Awards

2021 **Spot Award,** Massachusetts Institute of Technology — *Cambridge, MA*

2020 **Spot Award,** Massachusetts Institute of Technology — *Cambridge, MA*

2019 **Sharpe Innovation Commons Seed Grant Award,** Ohio State University — *Columbus, OH*