

Dany Haddad

danyhaddad@utexas.edu
github.com/dmh43
(512) 589-2287

WORK AND RESEARCH EXPERIENCE

Cognitive Scale

Machine Learning Applied Scientist

Austin, TX

January 2018 — Present

- Entity linking for information retrieval
 - Developing **entity based search** and **recommendation** for use in financial, medical and retail verticals.
 - Developed an entity linking model using deep learning, matching state of the art performance.
- Responsible AI
 - Built a web-service for **explaining predictions** of complex classification models.
 - Built a recommendation engine which incorporates user feedback as well as discrimination and privacy concerns.
 - Developed an interpretable anomaly detection system in the context of **noisy labels** and severe **class imbalance**.
 - Developed techniques for extracting rules as explanations from predictive models.
- Implemented **approximate message passing** for classification in sequence models.
- Implemented **approximate soft-max techniques** for deep NLP models.
- Determined trends and seasonalities in clickstream data to provide recommendations for **SEO**.
- Built a streaming data pipeline for training machine learning models using content from Wikipedia.

University of Texas

Graduate Research Assistant — Intelligent Data Exploration and Analysis Lab

Austin, TX

August 2017 — May 2019

- Introduced two novel contributions for training neural networks using **weak supervision** for document retrieval (published in SIGIR 2019).
- Developed novel techniques that identify training examples that explain the bias learned by a model.
- Extended techniques for unbiased recommendation and **fair ranking** algorithms.
- Improved state of the art defense mechanisms against **adversarial examples** for deep network classifiers.
- Improved the robustness of a deep learning based commonsense-reasoning model using adversarial training.
- Applied **robust statistics** techniques to achieve performance guarantees for time series forecasting in the context of arbitrary noise.

Originate

Software Engineer, Technical Lead

Los Angeles, CA

March 2016 — June 2017

- Transitioned the data model of a 20 kloc codebase from a schema-less Firebase data-store to a PostgreSQL database utilizing a GraphQL API.
- Developed a scalable and reliable data aggregation backend for a commercial real estate search tool.
- Front-end and back-end development for web-apps.

PUBLICATIONS

Learning More From Less: Towards Strengthening Weak Supervision for Ad-Hoc Retrieval

Published in: SIGIR, 2019

EDUCATION

University of Texas at Austin: *Master of Science, Electrical Engineering*

May 2019

Thesis Title: Strengthening Weak Supervision for Information Retrieval

Courses: Bayesian Methods, Responsible AI, Data Mining, Deep Unsupervised Learning, Time Series Analysis, Large-Scale Optimization, Genomic Data-Science and Bioinformatics

SKILLS

- Model selection and evaluation for predictive and prescriptive applications.
- Data exploration, feature extraction and selection.
- Proficient in Python, Clojure, and Javascript.
- Proficient SQL and No-SQL user and database maintainer.
- Working knowledge of Apache Spark.
- Experience with OCaml.
- Strong theoretical understanding of trade-offs of SVM, random forest, GLMs and deep learning.
- Deep network training and architecture selection methodology.
- Proficient in pytorch, numpy and pandas.
- Experience writing CUDA GPU code.

INDEPENDENT PROJECTS

- Developing a tool for expediting and streamlining the development of ML models ensure the reproducibility of results: `rabbit-ml`
 - Automatically handles creation, loading and invalidation of caches.
 - Records and manages model hyperparameters and performance metrics.
- Developed a symbolic electronic circuit solver for teaching introductory electrical engineering.
- Built a melody extraction tool for transcribing polyphonic music.
- Contributor to `StumpWM`, written in Common Lisp.
- Contributor to `dumpster-dive`, a tool for parsing Wikipedia markup and dumping it to a local datastore.
- Contributor to `Tertestrial`, a software testing framework.
- Contributor to `core.matrix.complex`, a Clojure library for working with complex valued matrices.
- Contributor to multiple open source Emacs Lisp tools.

AWARDS

- Received the Charles W. and Margaret A. Tolbert Endowed Scholarships and Fellowships in Electrical and Computer Engineering.
- Received the Nicholas and Maria Weber Electrical Power Endowed Scholarship, 2014
- Named an Engineering Scholar, Cockrell School of Engineering, 2013
- Received the C.W. Cook Endowed Presidential Scholarship, 2013
- Received the Gulf Coast Power Association Program in Electric Power Scholarship, 2012