Ming-Chang (Eric) Chiu

3572 S Budlong Ave, Los Angeles, CA, 90007 • http://charismaticchiu.github.io/ • +1 702-209-6629
mingchac@usc.edu • https://www.linkedin.com/in/eric-chiu • https://github.com/charismaticchiu

EDUCATION

University of Southern California (USC) — Los Angeles, CA
Master (M.S.) in Electrical Engineering, Data Science track — Aug 2016 – May 2018 (Expected)
- Analysis of Algorithms; Machine Learning; Pattern Recognition; Natural Language Processing
- GPA: 3.78 / 4.0

National Tsing Hua University (NTHU) — Hsinchu, Taiwan
Bachelor of Science (B.S.) in Computer Science & Electrical Engineering — Sep 2011 – Jun 2015
- Selected as honorary member of Phi Tau Phi Scholastic Society (only 1 in department)
- Cloud Programming; Scientific Computing; Numerical Analysis; Operating System
- Last 60 GPA: 4.05 / 4.30 Cumulative GPA: 3.82 / 4.30

University of Minnesota — Twin-Cities, MN
Exchange Student in Computer Science; GPA: 4.0 / 4.0, 15 credits — Fall 2014
- Introduction to Intelligent Robotics; Advanced Programming Principles

SKILLS

Languages: Python, C/C++, MATLAB, Ocaml, Java, SQL, BASH, Javascript
Technologies: Hadoop, Docker, TensorFlow, AWS, Git, jQuery, Linux, Spark

WORK EXPERIENCE

Information Retrieval and Data Science Group, USC — Los Angeles, CA
Researcher [Project: TensorFlow trained Byte Histograms for Better MIME Detection] — Sep 2017 – Present
- Developing TensorFlow CNN, GAN models to extract fingerprints of file types for network security in TREC-DD data and evaluating neural network models with Apache Tika default
- Generating byte frequency analysis signatures for particular MIME types and accumulating this signature for 93 file types in the TREC-DD polar dataset
- Integrating a command-line interface that can be run on the TREC-DD-Polar data

Institute for Creative Technology — Los Angeles, CA
Researcher [Project: Conversation Quality Assessment] — Jan 2017 – Present
- "Learning Dyadic Attention Networks to Predict Outcomes of Motivational Interviewing" X. Huang, L. Liu, M.-C. Chiu, J. Woolley, S. Scherer and B. Borsari. Submitted to NAACL 2018
- Trained Long Short Term Memory (LSTM) deep learning model and word embeddings for behavioral modeling based on Fisher and alcoholism treatment data
- Predicted new conversation and treatment session quality reaching 75% accuracy
- Generated conversation snapshots on top of Fisher dataset and created Amazon Mechanical Turk jobs to collect objective assessments from people

Illumina, Inc. — San Diego, CA
DevOps Applications Intern — May 2017 – Aug 2017
- Built a machine learning Cron for analyzing jobs in the High Performance Computing cluster to identify "destined to fail" jobs using Sci-kit learn, reaching 95% accuracy
- Developed a deep learning Daemon that constantly loads new samples from database (Hive) to train a classification model using Mini-batch update technique with TensorFlow, achieving 78% accuracy
- Maintained enterprise Atlassian Jira and Confluence; created customer workspaces for internal clients

Vision Science Lab, NTHU — Hsinchu, Taiwan
Research Assistant [Project: The World is Changing: Finding Changes on the Street] — Feb 2015 – Sep 2015
- Constructed image change detection model in MATLAB, successfully detected street view mismatches in Dash camera images with respect to preprocessed Google Street View (GFV) to provide updated information
- Applied RANSAC to re-outline the areas of mismatches in the original GFV images with accuracy outperforming baseline by 46%
- Devised a reusable manual labeling software and data types that recorded ground truth mismatch areas to help data collection

PROJECTS

Kaggle Competition: Porto Seguro's Safe Driver Prediction GitHub — Nov 2017
- Won Silver medal (top 4% out of 5,332 teams) in this biggest Kaggle competition in history
- Implemented XGBoost, Neural Network, lightGBM algorithms for predicting the probability that a driver will initiate an auto insurance claim in the next year
- Incorporated hierarchical interpolation, and boosting techniques to combine then produce better models

Movie Recommender — Mar 2015 – Jun 2015
- Implemented and tested 3 collaborative filtering algorithms in Python and utilized MovieLens dataset to recommend movies
- Back-end analysis system deployed on AWS EC2, enabling the recommender to regularly update recommendations by checking new user preferences
- Created front-end webpage using jQuery, AJAX and Bootstrap for visual effect

Dictionary Search Engine Github — Feb 2015 – Apr 2015
- Implemented PageRank algorithm for Apache Hadoop in JAVA and constructed a search engine which prioritizes relevant links
- Coded under scalable MapReduce framework on 8-node distributed computers allowing massive dataset to be processed
- Devised file system database for dictionary content retrieval by applying Apache Hbase and Hive