FABRICE HAREL-CANADA

SOFTWARE DEVELOPER

fabricehc@cs.ucla.edu | 954-805-5735 | Los Angeles, CA

EXPERIENCE

Graduate Student Researcher
University of California, Los Angeles

Jun 2019 – Present | Los Angeles, CA

- Evaluated the reliability of testing metrics for deep learning
- Developed extensions to adversarial attack algorithms

Software Development Consultant
Axenzi Consultants

Oct 2018 – Present | Los Angeles, CA

- Helping clients to plan and implement technology projects - specs, devops, and app / database development

Course Reader
University of California, Los Angeles

Jan 2019 – Mar 2019 | Los Angeles, CA

- CS 143: Database Systems | Developed test cases for a Spark project in Scala and actively addressed student questions

Project Manager + Developer Team Lead
Verox Tech

July 2015 – Sept 2018 | Los Angeles, CA

- Designed and successfully launched a smarter insurance app by building, motivating, and guiding a team of 23
- Contributed 100s of check-ins, primarily consisting of C# Web APIs and database tables, triggers, sprocs, and T-SQL scripts
- Designed and implemented a 3-system legacy data migration
- Responsible for producing, coordinating and reviewing nearly every aspect of the SDLC:
  - code reviews and code optimization
  - requirements and spec generation
  - database design + SSIS + SSRS
  - technical documentation
  - UI design standards
  - product design
  - server setup / config
  - build and deployment scripts

Business + Data Analyst
People's Trust Insurance

Dec 2012 – Sept 2018 | Deerfield Beach, FL

- Go-to data expert for executives to produce hundreds of reports and analyses for pivotal decision making at the executive level
- Developed numerous front-end / back-end application features that were successfully deployed into production to automate tedious and time-intensive workflows
- Performed continuous requirements gathering & demos with business stakeholders to support 60+ sprints of development for both applications and business intelligence teams, including the transition to a new policy and claim management system

OBJECTIVE

Highly-motivated to add value as a data scientist, build useful tools, and leverage + expand my database development, machine learning, and leadership skills.

EDUCATION

Computer Science PhD
University of California, Los Angeles

Jan '20 – Dec '23 | 3.92 / 4.0 GPA

- Neural Networks and Deep Learning
- ML Testing and Debugging
- Natural Language Processing
- Advanced Data + Knowledge Bases

Computer Science MS
University of California, Los Angeles

Sept '18 – Dec '19 | 3.9 / 4.0 GPA

BS Information Systems
University of Florida

Aug '08 – May '12 | 3.85 / 4.0 GPA

- Magna cum Laude + Thesis

SKILLS

t-sql
python
pytorch
tensorflow + keras
spark
java
c#

numpy | pandas | scipy | sklearn | opencv | matplotlib | plotly | seaborn | linq | html | css | js

TOOLS

ssms + ssis + ssrs
tfs + azure devops
visual studio
excel

jupyter | office | LaTeX | aws | gcp | exchange server | moqups | wordpress

RESEARCH + PROJECTS

Relational Stock Price Prediction using GNNs

University of California, Los Angeles

3 mos

- Approximating company-to-company relationships from stock data
- Initialize relations with 30-day rolling correlations; train them using GNNs
- Predictions within 75% of optimal for NYSE; 35% of optimal for NASDAQ

Test Metrics for Deep Neural Networks (DNNs)

University of California, Los Angeles

5 mos FSE 2020

- Formalized properties expected of useful test suites for DNN testing
- Devised a new regularizer to induce higher network activations
- Conducted a massive evaluation across 1500+ experimental configurations to determine the reliability of DNN test metrics. One of the most popular metrics, *neuron coverage*, does not correlate with defect detection, input realism, or output diversity.

Machine Learning on Source Code

University of California, Los Angeles

3 mos

- Trained a Google Transformer network to perform two tasks on Java and Python code bases:
  - Intake a function and output a plain English summary of what it does.
  - Intake a code snippet and output a syntactically valid function name.

Insurance Claim Predictions

University of California, Los Angeles

2 mos

- Data extraction from 4 sources with extensive preprocessing and imputations performed via SQL and Python
- Overcome class imbalance issues via SMOTE and oversampling
- Developed a non-blocking implementation of a Naive Bayes Classifier (NBC) for Spark Streaming, directly calculating frequencies and conditional probabilities using PySpark and SparkSQL to make real-time predictions for claims in a high-volume environment
- Implementation and tuning of 11 other models resulted in the top performer being a XGBoost classifier with a 93% F1 score

Differential State Analysis and Targeted Input Selection

University of California, Los Angeles

2 mos

- Explored how to identify layers with the highest proportion of underperforming neurons and then using their weights to generate heat maps that can be used to identify subsequent training batches for overfitting / underfitting issues on underperforming classes
- Reduces average training time by 90% and improves performance by 10%

CNN & RNN Applications in EEG Decoding

University of California, Los Angeles

1 mos

- Achieved state-of-the-art prediction accuracy on BCI Competition classification of user action from neural signals by implementing and extending upon architectures from several leading papers.

INTERESTS

machine learning | statistics | logic | software engineering | board games | video games | philosophy | cooking | design | travel | self-improvement

REFERENCES

Prof. Miryung Kim
@ UCLA
miryung@cs.ucla.edu

Prof. Quanquan Gu
@ UCLA
qgu@cs.ucla.edu

Michael Simhai
@ Verox Tech
michael@veroxtech.com

Lisa Branon
@ People's Trust Insurance
lbranon@upcinsurance.com