Chhaya Choudhary

chhayc@uw.edu | linkedin.com/in/chhayachoudhary | GitHub: chhayac
Redmond, WA | (984) 244-8065 | VISA - H4 EAD (No immediate sponsorship required)
Web: https://chhayac.github.io

Education

M.S. in Computer Science & Systems *Sept 2017 - Mar 2019 (Expected)*
University of Washington, Tacoma, GPA: 3.97/4.0
B.Tech. in Information Technology *July 2007 - May 2011*
Banasthali University, Niwai, India

Skills

Programming & Tools: Java, Python, C, SQL, Machine Learning and Deep Learning libraries (Tensorflow, Keras, Scikit-Learn, Pandas, Numpy), LaTeX, AWS (EC2, S3, Lambda, EMR, SageMaker), Jupyter, Matplotlib, OpenCV, Git, Statistical Techniques (Regression, Classification, Time Series Modeling, Clustering, Tree-based methods, Support Vector Machines), MS-Excel

Publications

Chhaya Choudhary, Raaghavi Sivaguru, Mayana Pereira, Bin Yu, Martine De Cock, and Anderson C. Nascimento. **Algorithmically Generated Domain Detection and Malware Family Classification.** *The Sixth International Symposium on Security in Computing and Communications (SSCC' 18)* (Accepted)

Bin Yu, Jie Pan, Daniel Gray, Jiaming Hu, Chhaya Choudhary, Mayana Pereira, Anderson Nascimento, and Martine De Cock. **Weakly Supervised Deep Learning for the Detection of Domain Generation Algorithms.** *IEEE Transactions on Information Forensics & Security, 2018* (Under Review)

Data Science Experience

Data Scientist Intern, Infoblox *June 2018 - present*

- Developing machine learning and deep learning models to detect and categorize the Domain Generation Algorithm (DGA) generated domain names to proactively help customers detect malware threats with 99% accuracy.
- Working on data exploration, data cleaning, feature engineering, feature selection, visualization and performance evaluation of models with large real traffic datasets.
- Ranked 1st in the Detecting Malicious Domain names (DMD 2018) competition, a workshop co-located with ICACCI'18 and SSCC'18.

Graduate Research Assistant, University of Washington *Jan 2018 - present*

- Working with Dr. Martine De Cock on detection of DGA using machine learning and deep learning techniques like Random forest, Convolutional Neural Networks(CNNs) and Recurrent Neural Networks(RNN-LSTMs).
- Working on evaluating state-of-the-art DGA classifiers against adversarial examples using Autoencoders and Generative Adversarial Networks (GANs).

Professional Experience

Software Engineer ***(Noida, India)*** *July 2011 - Dec 2015*

- Developed modules using Java to integrate costing and staffing system with a new web-based product(Maintenix) for Qantas engineering. It ensured high application availability and improved airworthiness of aircraft.
- Implemented automated solutions for daily report generation tasks in engineering applications to reduce time and effort expended using Java and SQL, thereby saving 50 Person days/year.

Technical Proficiency Demonstrated: Java, SQL, IBM Mainframe, Python, PL/1

Major Academic Projects

Age, Gender, and Personality Trait prediction of social media users *Fall 2017*

- Designed and developed machine earning and deep learning models to predict age, gender and personality traits of social media users using their status updates, profile pictures, and page likes with 87% accuracy.

Recommendation Engine from Scratch *Fall 2017*

- Developed a movie recommendation engine using k-nearest neighbors algorithm.

Ranking lyrics for Online Search *Winter 2018*

- Implemented natural language processing ranking models(Query Likelihood, Vector space, word2vec pre-trained embeddings from Google) for ranking the lyrics of songs according to a search query using million song dataset.

Traffic Sign Recognition Using Deep Learning *Winter 2018*

- Designed and developed a Convolutional Neural Network(CNN) to recognize traffic signs from images with an accuracy of 98.4% using data augmentation techniques.

Empirical Study of Network Flow Algorithms *Spring 2017*

- Implemented Ford-Fulkerson, Scaling Ford-Fulkerson, and Preflow-push algorithms in Java to find maximum flow in a network. Conducted an empirical study depending on factors like choice of data structures, cycles in graphs, graph size, and graph type.

Serverless Computing Microservices Composition *Spring 2018*

- Researched, analyzed and evaluated performance and cost metrics for various function compositions of microservices on AWS Lambda.

Technical Proficiency Demonstrated: Python, Java 8, Convolutional Neural Networks, Keras, Tensorflow, OpenCV, Scikit-Learn, Numpy, Pandas, Jupyter, Matplotlib, Algorithms, Data Structures, NLTK

AWARDS AND HONORS

GHC'18 Scholar, Awarded scholarship to attend the Grace Hopper Celebration 2018, Houston *May 2018*
Student Grant Usenix Security'18, Awarded to attend 27th Usenix Security Symposium, Baltimore *July 2018*
Upsilon Pi Epsilon (UPE), Awarded for academic excellence at the University of Washington *April 2018*