

SNEHA NAGPAUL

snehanagpaul@me.com

571-365-6300

sneha.co

EDUCATION

Master of Science - Computer Science - George Mason University

Dec 2017
(Current GPA: 4.0)

Bachelor of Science - Computer Science - Delhi University, India

May 2010

SKILLS

Languages	Python, SQL, R, C, C++, Java
Libraries	Keras, TensorFlow, PyTorch, Theano, Pandas, Scikit-Learn, NumPy
Tools	Tableau, Git, Jupyter Notebooks, PyCharm, RStudio, Xcode
Operating Systems	Unix/Linux, MacOS, Windows
Natural Languages	English (fluent) , Hindi (fluent) , French (working)

TECHNICAL WORK EXPERIENCE

Data Science Intern - All Traffic Solutions, Herndon, VA

July 2017 - August 2017

Maintained the production environment of an **IOT** stack and contributed to the deployment of machine learning models in a big data environment with technologies such as **Apache Hadoop** and **Apache Spark** for developing online learning models that send real time alerts to traffic and parking monitors for immediate action.

Graduate Research Assistant - George Mason University, Fairfax, VA

Jan 2017 - June 2017

As part of a graduate research assistantship at George Mason University, mined massive data sets consisting of raw news articles and cross referenced with police datasets to help **predict crime** instances in major cities of Colombia.

Graduate Assistant - George Mason University, Fairfax, VA

Sep 2016 - Dec 2016

Assisted with evaluating exams, quizzes and assignments for a discrete mathematics and formal logic graduate level class at GMU.

Founder - actuarialscience.in, New Delhi, India

Jan 2013- Aug-2015

Managed a team that covered important technical Actuarial Science conferences as part of a survey on the latest developments in the Indian insurance scenario. The non profit organization also prepared reports on the demographics of actuaries in India with data collected from the Institute of Actuaries India (IAI). These reports were then used to predict the future requirements of actuarial talent and were geared towards policy development at IAI.

ACADEMIC PROJECTS

Geographic Information Retrieval: Deep Learning, Natural Language Processing (2017)

Discovering efficient deep learning architectures that use text to predict latitude/longitude. Explored RNNs (**LSTMs**) by using word embeddings from pre-trained models like GloVe and Word2vec and adapting them to the GIR use case by **Transfer Learning** in Keras, Tensorflow and PyTorch.

Cervix Classification: Deep Learning, Computer Vision (2016)

Practically applied concepts of Transfer Learning in a Deep Learning context for image classification to identify type of cervix for subsequent cancer diagnosis. Explored architectures like **Convolutional Neural Networks** and **ResNets** using Deep Learning libraries such as Keras, Theano and Tensor Flow in Python3 optimized for running on GPU CUDA cores.

Loss Prediction for Insurance: Regression Analysis (2016)

Analyzed a mix of categorical and continuous variables to get real valued predictions for a loss generated from a skewed distribution based on experience (claims) data. Regression techniques such as boosted trees and random forest were employed after careful feature selection and feature engineering.

Sentiment Analysis for Raw Text Reviews: Natural Language Processing (2016)

After preprocessing raw text data and employing dimensionality reduction techniques, data were classified as positive or negative using the K-nearest Neighbors classification algorithm.

Grouping Text Data into Similar Categories: Cluster Analysis (2016)

Given a bag of words representation of articles in newspapers, dimensionality reduction techniques such as feature weighing and Singular Value Decomposition were applied. Following which, K Means clustering techniques were performed to group articles by type of sport covered.

Recommendation System for Movie Rating Prediction: Collaborative Filtering (2016)

Given a set of ratings and reviews, a collaborative filtering technique was devised following feature reduction using Truncated SVDs from Scikit-Learn. Content based approaches were used to augment predictions of movie ratings given a user.

Evolution of Face Recognition: Multidisciplinary colloquium (2016)

Reviewed and charted (on a poster) the research conducted on Face recognition through the 1990's to present day in order to study the patterns of cross-disciplinary borrowing in the field of Computer Vision.

Analysis of P2P networks using Markovian Random Walks (2010)

Random Walk methods, using Markov Chain Monte Carlo simulation techniques, applied to network analysis of Peer to Peer networks in order to improve fidelity and help construct a more robust network structure.

PROFESSIONAL AFFILIATIONS

IFoA — Institute and Faculty of Actuaries – Part- Qualified Actuary

since 2012

CAS — Casualty Actuarial Society - Member

since 2015

EXTRA CURRICULAR ACTIVITIES AND CAUSES

Women of Color in STEM(GMU)

Member since 2015

Society of Women Engineers(GMU)

Member since 2015

American Statistical Association(GMU chapter)

Former Secretary and Member since 2015

Organizing Hack-a-thons in collaboration with the George Washington University ASA chapter to invite other local schools and technology startups to participate and sponsor data centric events.

Hack for Change: Women's empowerment in India (2014): Statistically approached data to produce an analysis of interviews that were used to draw inference on the financial and social standing of rural women in India.

Actionable results were then forwarded to political organizations as part of a community development project.