SNEHA NAGPAUL

San Francisco,CA | snehanagpaul@me.com | 571-365-6300

PROFILE

Self driven entrepreneurial software engineer and data scientist with a background in applied financial statistics, looking to make a social impact with technology.

SKILLS

Languages	Python, SQL, R, C, C++, Java
Libraries	Keras, TensorFlow, PyTorch, Theano, Pandas, Scikit-Learn, NumPy
Tools	Tableau, Git, Jupyter Notebooks, PyCharm
Spoken Languages	English (fluent) , Hindi (fluent) , French (working)

TECHNICAL WORK EXPERIENCE

Data Science Intern - Vinsol US Inc., San Francisco, CA April 2018 - Present

- Developing socially aware data solutions in health care, consumer products and e-commerce.

Data Science Intern - All Traffic Solutions, Herndon, VA July 2017 - August 2017

- Developed time series forecasting models for traffic engineers using Big Data pipelines constructed with Apache Spark.
- Exposed model endpoints using Flask in Python3 using API design standards.
- Increased revenue by working with Sales and Marketing teams and building a UI report after analyzing client activity on the internal software. Used Tableau, Matplotlib and Jupyter notebooks.

Graduate Research Assistant - George Mason University, Fairfax, VA January 2017 - June 2017

- Mined massive data sets consisting of raw unstructured articles and cross referenced findings with police data.
- Created a framework for local law enforcement for recommending crime hotspots using spatio temporal analysis.

Founder - actuarialscience.in, New Delhi, India January 2013- August 2015

- Reported on Actuarial Science conferences and latest developments in the Indian insurance landscape.
- Suggested governmental policy changes by cross referencing datasets from the Indian Actuarial Body and other data sources.

PUBLICATIONS AND ACADEMIC PROJECTS

From Language to Location using Multiple Instance Neural Networks - SBP, Conference Proceedings, Washington DC (2018)

- Used neural networks for finding the geographic location of short text data using semi-labelled information.
- Improved prior semi supervised learning approaches by using deep learning for multiple instance learning instead of kernels.

Cervix Classification: Deep Learning, Computer Vision - Kaggle Competition (2016)

- Practically applied concepts of Transfer Learning in a Deep Learning context for image classification to identify type of cervix for subsequent cancer diagnosis.
- Reached high classification accuracy (top third of the leaderboard) using architectures like Convolutional Neural Networks and ResNets with Deep Learning libraries such as Keras, Theano and Tensor Flow in Python3 which are optimized for running on GPU CUDA cores with AWS.

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

- Preprocessed IMDB movie review dataset and implemented K-nearest Neighbors classification algorithm in Python3 for sentiment analysis.

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.

EDUCATION

Master of Science - Computer Science - George Mason University (GPA: 3.96) May 2018

Course Work: Formal Logic and Discrete Mathematics, Systems Programming, Analysis of Algorithms 1& 2, Data Mining, Time Series Mining, Object Oriented Software Specifications in Java, User Interface Design

Bachelor of Science - Computer Science - Delhi University, India May 2010

AFFILIATIONS AND EXTRA CURRICULAR ACTIVITIES

Women Who Code Since 2016

Casualty Actuarial Society Since 2015