

Srishti Sehgal

MACHINE LEARNING ENGINEER



SRISHTI.SEHGAL@MAIL.UTORONTO.CA



647-861-2126



WWW.LINKEDIN.COM/IN/SRISHTISEHGAL



EDUCATION

BASc in Engineering Science

University of Toronto 2015-2020

MINOR: Robotics engineering and Engineering Business

RELEVANT COURSES:

Data Structures and Algorithms, Machine Learning and Data Mining, Control Systems, Statistics, Dynamics/Classical Mechanics, Scientific Computing and Signal Processing, State Estimation and Optimization, Image Processing for Computer Vision

</> RELEVANT SKILLS

PROGRAMMING LANGUAGES/SOFTWARES

Python	Matlab
Java	Simulink
C	Arduino
C++	HTML/CSS
3D Printing	Solidworks
Tensorflow	Scikit-Learn
PyTorch	TFLearn
Caffe	Keras
Git	R
Solidworks	LaTeX
Arduino	

SOFTWARE ENGINEERING

Object-oriented Programming, Test Driven Development, Scalable Architecture Design, Code Review, Agile Development, Refactorization of Legacy Code, Revision Control, Continuous Integration



PROFESSIONAL EXPERIENCE

STRUCTURES AND FULL-SCALE TESTING ENGINEERING INTERN

NATIONAL RESEARCH COUNCIL CANADA OTTAWA, ON | MAY '18 – PRESENT

- Wrote modern, performant, maintainable code (Python, Java, C, C++) for a diverse array of client and internal projects
- Optimized and trained deep learning (DNN) models for acceleration on GPU (CUDA) by reducing elapsed time by 85% and improved results by 30%
- Applied machine learning techniques using Tensorflow, PyTorch and Caffe in system health monitoring to generate empirical models and achieve 85% accuracy
- Published four papers in reputable peer-reviewed machine learning journals and wrote several technical reports for internal use and for clients like Department of Defence
- Delivered high-quality presentations in weekly group meetings and at the CASI conference
- Used predictive analytics (supervised, unsupervised, semi-supervised machine learning models and genetic programming with parallel computing) to provide insights and in-depth analysis on the status equipment health
- Used Python and Arduino to control positioning sensors and worked with LiDAR data remotely
- Used test-driven development techniques and software development paradigm (Agile) to develop code to automate data collection and load prediction from full-scale testing projects. This method proved to be a faster development cycle (time decrease by 80%) and contributed to an easier codebase to refactor and maintain
- Developed a white-box machine learning model for load prediction using strain values with errors less than 5%
- Developed reusable object-oriented software to generate shear, bending moment and torsion models for fighter jet parts subjected to full-scale testing
- Supervising junior co-op on new and old machine learning projects to improve work efficiency by 50%

RESEARCH ASSISTANT

UNIVERSITY OF TORONTO TORONTO, ON | MAY '17 – JANUARY '18

- Developed custom image processing, data preprocessing and visualization algorithms in Matlab, Python, and R to perform image segmentation and analyze dynamic properties like velocity of a complex fluid for an independent research project with high accuracy on 300 datasets
- Organized large amounts of information and drew conclusions from research to author and publish a paper on top of a full-time course load

BACK-END DEVELOPER

INKSPIRE

REMOTE | JANUARY '17 – MAY '17

- In a team of four students, worked as a developer in an Agile environment to develop a unique online platform, similar to Google Docs, for the stakeholder with HTML/CSS, JavaScript and Angular while learning and successfully implementing new languages like JSON and Node.js with no prior knowledge
- Designed a new and improved user-first interface resulting in increased web traffic by 25%

LEADERSHIP AND

VOLUNTEER WORK

NSight Mentorship

- VP Finance and Mentor for Engineering Mentorship
- July '16 – May '18

Galbraith Society

- VP Finance for the Galbraith Undergraduate Research Journal
- May '17 – May '18

University of Toronto Space Design Competition

- VP Competitions of UofT Space Design Competition
- Jan '17 – Present

Women in Stem Outreach

- Leading member for Logistics and Mentorship for Women in Research @ NRC (workplace of current internship)
- May '18 – Present

INTERESTS

PARTICIPATING IN HACKATHONS,
BHARATANATYAM (DANCING), PLAYING THE
GUITAR, BEING A FOODIE



SAMPLE MACHINE LEARNING PROJECTS

AD BLOCKER CHROME EXTENSION WITH A MACHINE LEARNING APPROACH

- Developed object-oriented code to detect spam ads on a webpage using deep learning-based models in Tensorflow with natural language processing
- Hid the spam ads with 88% accuracy
- Easily developed a functional prototype within a week
- Constructed a chrome extension with test-driven development

CONSTRUCTING A NOVEL ARCHITECTURE FOR DEEP LEARNING

- Developed code for a senior data scientist for designing chemical syntheses from scratch using PyTorch and Tensorflow and will soon be published for Natural Resources Canada
- Developed novel network architecture from a paper describing the model
- Developed code to train networks on multiple GPUs (CUDA) on a super computer. This reduced the training time by 75%
- Developed a robust evolutionary algorithm wrapper around the machine learning model for hyperparameter optimization to achieve better results faster

ANALYSIS OF AUDIO FEATURES OF SPOTIFY'S HITS USING MACHINE LEARNING

- Implemented cross industry standard procedure for data mining to develop a modular data pipeline to preprocess input data into model
- Used feature selection techniques to unravel importance of music features
- Used Spotify's WebAPI to extract features of music like loudness, artists...
- Used MySQL to write to and read to database. This improved time by 20%
- Compared explicit regression equations (containing variables like loudness, length of song...) to equations developed by regression trees

COMPUTER VISION PROGRAM FOR RECYCLABLE MATERIAL DETECTION

UoTHACKS VI

- Optimized a deep learning architecture in PyTorch (DNN) and developed a computer vision algorithm using Python to segment recyclable materials from non-recyclable material in order to decide which items can be recycled (e.g. metal) in a given pile of household waste image
- Achieved 78% accuracy while working under pressure under 36 hours

STOCK MARKET PREDICTION USING MACHINE LEARNING

QHACKS

- Developed a Python program in a team of two other students that predicts the movement of a user-defined stock/tweet
- Intensive pre-processing was performed on twitter data and financial open/close price data to ensure that missing terms were accounted for and that noise terms like punctuation and language stop words were removed from the dataset
- Tuned BERT model, after reading its pseudo-code from publication, in PyTorch with Tensorflow backend to achieve an accuracy of 62% all in the span of 36 hours

CAR LOAN DEFAULT PREDICTION

LTFS DATA SCIENCE FINHACK

- Identified attributes having strong correlations with default target data using dimension reduction techniques (like PCA and autoencoders)
- Developed ETL pipelines using agile development in Python and MySQL
- Performed research on features to remove erroneous rows of data, which improved classification results by 24%
- Normalized all continuous data with z-scores and encoded ordinal/categorical features to improve model performance by 40%
- Explored network-based models (DNN), trees, clustering algorithms using PyTorch and scikit-learn with custom Python code to automatically tune hyperparameters. Achieved ranks within top 14% of all competitors