December 5, 2016

Somin Wadhwa

Contact Information

Undergraduate Student
Block 1, Computer Science & Engineering
Maharaja Agrasen Institute of Technology.
Rohini, Delhi, India.

Phone: (+91) 9312349897
E-mail: sominwadhwa@gmail.com
sominwadhwa@iCloud.com
GitHub: sominwadhwa

Interests

Exploratory Data Analysis, Supervised Learning

Education

B.Tech in Computer Science & Engineering — July 2014 – present
Maharaja Agrasen Institute of Technology (Overall Percentile: **80.6%** July 2016)
Guru Gobind Singh Indraprastha University, Delhi, India

High School: Bal Bharati Public School, Pitampura, Delhi — March 2012 – April 2014
All India Senior School Certificate Examination, CBSE (Percentile: **93.8%**)
Secondary School: Bal Bharati Public School, Pitampura, Delhi — March 2000 – April 2012
CBSE (GPA: **8.8**)

Experience

Internship: CodingNinjas — June, 2016 - July, 2016
6 Weeks of work in with the technical team where I worked on building machine learning models and assignments for their upcoming programs. Models and Assignments included regression models, SVMs, Neural Nets and several Unsupervised Models.

Technical Skills

Strongest Areas: Supervised Learning, Algorithms, Data Analysis
Languages: Python, C++
Tools & Frameworks: Matlab, LaTeX, Conda, iWork Suite, iPython Notebooks
Database Tools: SQL

Relevant Courses Taken

Data Structures & Algorithms, Databases, Machine Learning, Automata Theory, Probability, Differential & Inferential Statistics, Software Engineering

Research Work

Somin Wadhwa, Study of Random Numbers & their applications in computational physics using Monte-Carlo method", *XXVII IUPAP Conference on Computational Physics, IIT Guwahati*, December 2-5 2015 (Abstract)(Here)

Selected Projects

Image Apportionor
A simple clustering based image processing. In its nascent stages, the algorithm was just a test implementation of the K-Means clustering but ultimately evolved into something much more. The project was done in Summer '16 as a part of the mandatory training project after my second year in undergraduate studies.

Kaggle/Housing Prices (Advanced Regularization Techniques)*
With 79 explanatory variables describing (almost) every aspect of residential homes. This short term work explores real-life data from The Ames Housing dataset which was compiled by Dean De Cock for use in data science education. It's an incredible alternative for data scientists looking for a modernized and expanded version of the often cited Boston Housing dataset. The data given was "skewed" in nature and thus needed to be normalized first. Then advanced regularization techniques (Lasso & Ridge) were applied and compared using CV scores. Accuracy achieved was nearly 94%. Still working to improve the results using gradient boosting.

UdStudentData
Rudimentary data analysis of some student data obtained from www.udacity.com. Analysis was solely done in python and feature (analysed & plotted) variations in Time, Lessons Completed & Number of Days of student visits in a particular course were studies. Entire analysis is based on three parameters Enrollments, Daily Engagements & Project Submissions.

Web Crawler*
Some standard python scripts that use the beautifulSoup library to crawl through various websites related to major sports leagues and fetch real-time standings of the respective teams. This project was done primarily to fetch the league standings for National Basketball Association (NBA) but the functionality can be extended as well.

Radioactive-Decay Simulator
This project was done as a part of the paper written for the XXVII IUPAP CCP (2015) in which various Monte-Carlo simulations were obtained for the radioactive-decay phenomenon. This project in its initial stages was purely implemented in C++ and plots were obtained using a seperate spreadsheet software. Later studies suggested that the entire process can automated via a numerical-computation tool such as Scilab. The project can further be implemented in Python with the help of relevant libraries (matplotlib, NumPy)

Kaggle/RMS Titanic
This was my first attempt at some real feature engineering and application of Machine Learning models to make real predictions. In this given data set, nearly 7-10 features were analysed/visualized according to the chances of survival of the passengers. Then using the relevant features (as seen through the visualizations and mathematical analysis) a logistic regression classifer was trained. Accuracy achieved was nearly 80% which was further improved using RandomForest Classifier.

**Ongoing Projects*
All the projects (including the above-mentioned) are/will be available on GitHub.

Training

Algorithms: Design & Analysis by Stanford University (MOOC), July 2016
6 Weeks of training in Algorithms Design & Analysis. Certificate earned on October 29, 2016

Summer Training: Machine Learning (MOOC), July 2016
11 Weeks of training in Machine Learning (Supervised & Unsupervised) on Octave. Certificate earned on August 23, 2016

Other Achievements

- **Secretary**(2015-present) 'Association of Computing Machinery' at MAIT.
- **Interned** at a national NGO 'Umeed - A drop of Hope' (NGO Reg: S/792/DIST.SOUTH/201) and jointly participated in Project- Knowledge for All (KFA).
- **Rotaractor** (2014-2015) Member of 'Rotaract Club of Delhi Akash' where I jointly organized several large scale events like 'CanSupport's Walk of Life (8th Feb 2015) - Fight against cancer.', 'Patrika - A paper recycling drive.'
- **Joint Convernor** Computer Science Events, Techsurge & Mridang Annual - technical & cultural fest of M.A.I.T

Hobbies & Interests

Reading about Economics, Basketball, Watching Documentaries, Quora, HackerRank

References

Available upon request.