225 East 95th St
New York NY – 10128

Pawandeep Singh

singh@alumni.emory.edu
(678)-557-0528

Education

Emory University   Atlanta, GA
M.S. In Computer Science (Machine Learning Concentration)   Aug 2013-May 2015
- Completed courses: Data Mining, Machine Learning, Natural Language Processing, Numerical Analysis, Applied Linear Models.

Guru Tegh Bahadur Institute of Technology   New Delhi, India
Bachelor of Technology in Computer Science   Aug 2007-May 2011

Work Experience

Hudson Data   New York,NY
Data Scientist   May 2016-Current
- Performed benchmarking of machine learning libraries such as H20.ai, Skytree and Scikit-learn for various machine learning models for classification and clustering tasks. Further, integreated them with hadoop and hdfs for scaling. Created data science training materials for data science bootcamp.

Youth on Move   New York,NY
Software Developer   Nov 2015-April 2016
- Performed refactoring of the existing codebase for the website and establish best engineering practices. Created a knowledge repository using project wiki to share all the learnings.

Emory University   Atlanta, GA
Research Assistant   June 2014-Nov 2015
- Performed complexity analysis of law personnel data over a period of time, measuring entropy and Flesch score. Measured relative complexity of personnel of different countries and relative increase in rate of complexity.
- Performed data wrangling, munging to obtain and clean data from different countries government sites using numpy, pandas. Further, used scikit-learn for implementing the models and data visualizations.

Infosys Limited   Hyderabad, India
Systems Engineer   Sep 2011-June 2013
- Worked as part of the offshore team for a major aircraft manufacturer on their Data Delivery Program to deliver relevant data according to individual airline plane configuration. Developed a file splitting module to split transaction files for parallel processing of transactions and hence reducing the transaction time by over 50%.
- Performed unit testing of individual components as new features were added and automated them using *JUnit*.

Academic Projects

CraigCarDeals   Jan 2016-Feb 2016
- Developed a web app that finds the best car deals on craigslist by scraping car listings and building a random forest pricing model. **Backend**: python, pandas, sklearn, Postgres, Heroku **Frontend**: JS, Html, CSS

Predictive Modeling of Drug Sensitivity   Oct 2014-Dec 2014
- Developed elastic nets and lasso regression models to predict drug sensitivity for different cancer cell lines to customize treatment for individual patients. Further used PCA for dimensionality reduction as the data contained 1639 Mutations, 23316 Copy Number and 18988 Gene Expressions. **Framework**: R

Early Alzheimer's Detection   Nov 2013-Dec 2013
- Developed machine learning models using Naive Bayes, SVM and Neural Networks to classify people based on their performance on a web based test using mouse tracking and clicking data. **Framework** : Matlab
- Further optimized the model using convex optimization, feature normalization and bias-variance optimization.

Technical Skills

- **Languages:** *Python, Java, R, Matlab, C, HTML/CSS, SQL, XML, Bash scripting*
- **Libraries & Frameworks:** *pandas, numpy, scikit-learn, Mallet, OpenMPI, H2o, Kafka*
- **Tools & Platforms**: Git, MongoDB, MySQL, Hive, Hadoop, Heroku, AWS, Postgres
- **Domain:** Machine Learning, Data Mining, Predictive Analytics, Natural Language Processing

Github: github.com/singhpawan   **LinkedIn:** linkedin.com/in/mepawandeep   **Website:** www.singhpawan.com