NILESH PATIL

https://github.com/nilesh-patil

npatil4@ur.rochester.edu
+1 201-790-3854

Graduate Projects, University of Rochester

- Analysing large Graphical dataset:
    - Using NYC's transportation data (1 billion+ taxi trips), we built a large time variant network to analyse travel habits of residents and determining pressure points in this network
    - Using RNNs to predict demand at each node on dense time variant, geospatial transportation graphs (use case: NYC transportation network)
- CNN based architecture to build a deep learning classifier for **galaxy classification** from images captured by telescopes
- Random Forest based ML regression model for **predicting pollution level** from open data collected by federal agencies

Professional Experience

Data Scientist / Sr. Analyst, Predictive Analytics   **AXA (New York, US / Pune, India)**   **December '14 to June '16**

- Worked with AXA partners to develop machine learning based predictive analytics frameworks.
- At AXA, we used Hadoop stack for storage + manipulation & R/Python/SPARK for analysis. Aggregation & Exploratory analysis was done using HIVE/IMPALA & R/Python.
- Underwriting pipeline visualization using tableau - developed a dashboard for the client succinctly representing current process & KPIs.
- Mortality rate error minimization using actual historical dataset from AXA's customers & transactions
- Telematics: Identifying driver fingerprint & classifying each ride the driver takes as a function of historical data & GPS position. We used Spark's MLlib library to create a model for identifying driving patterns for each driver & put in place a system which can be referred to while processing accident claims.
- Simulation of US population for agent based modelling using Anylogic. We essentially simulated socio-economic, health & credit conditions for each sub section of US population and effective movement within each group for risk pooling.
- Project Management: Part of my role in Insurance Productivity & Telematics projects was to manage the team of 4 working alongside me, formulate the client & project requirements into business problems and work towards the goals incrementally.
- Promoting Data science community by organizing seminars, best practices sessions, machine learning & big data tools workshops for AXA-US & AXA-India.
- Mentoring & educating colleagues/client in new data science technologies & training in R/Tableau.

Analyst – Big Data/Predictive Analytics   **AbsolutData Research and Analytics**   **June '13 to December '14**

- Sensor & transactional data analytics: Extensive predictive analytics experience for a Mining Equipment Failure Prediction project for one of the world's largest copper and gold mining companies. This was a breakthrough project for the company's venture into Decision and Data Sciences.
- Worked hands-on as R & Analytics resource on data management and data reshaping for the project using R, Hive queries & Hadoop. The data varied from Sensor data (13 TB) to Gigs of manually entered and unclean work order & maintenance data.
- Developed multi stage prediction model for prediction of engine failure (~$850,000 per engine).
- Developed machine learning based predictive models for oil quality in heavy machinery. Used random forests in the final production implementation.
- Implemented an unsupervised learning algorithm which helped reduce false alarms for the mining major based on actual effect of alarms based on historical alarms & maintenance data.
- Trained 47 co-workers on using R as a part of the capability building program.
- Marketing mix modelling: Worked with Mondelez, Japan to analyze the impact of various marketing activities in the year on sales w.r.t investment made.

EDUCATION

Year	Degree	Specialization	Institute
2017	MS	Data Science	University of Rochester, USA
2013	B.Tech.	Metallurgical & Materials Engg.	Indian Institute of Technology - Roorkee, India

Skills

Programming – R, Python, SQL, Scala, Julia

Toolset – Spark, Tensorflow, Keras**,** Tableau for visualization, hands on experience with HIVE/IMAPALA queries, MapReduce using Python, Hadoop

Data Science & Machine Learning – Experience in solving business problems using machine learning techniques such as Random Forests, Neural Networks, Support Vector Machines, Gradient Boosted Regression, Linear Regression, Logistic Regression, Clustering Techniques (k-means, hierarchical clustering, knn etc), CNNs, RNNs