

# Nilesh Patil

2017903854 | nilesh.patil@rochester.edu | nilesh-patil.github.io | nilesh-patil | Nilesh Patil

## Education

Jun'16 - Dec'17 **University of Rochester**, M.S. in Data Science (Machine Learning and Computer Vision applications)  
Jul'09 - May'13 **Indian Institute of Technology Roorkee (IIT)**, B.Tech. in Metallurgical & Materials Engg.

## Projects

### Computer Vision & Deep Learning



#### Thorax Disease classification from chest X-rays using Deep Learning

- Trained a deep neural network to detect 14 different classes of thorax diseases using chest X-ray scans
- Achieved best AUC of  $0.93 \pm 0.015$  using neural network ensemble

#### Extracting uncertainty information from Deep Neural Networks

- Implemented uncertainty information extraction & confidence interval building for deep neural networks using MC-dropout
- Developed visualization approaches to interpret uncertainty information extracted for different neural network applications

#### Surface construction in live tissue from hyperspectral microscopic images

- 3D Surface and volumetric reconstruction from multiple 2D cross sections of live tissue using Voronoi mesh representations
- Implemented pipeline to be used at the bio-imaging research lab at University of Rochester

#### Machine learning driven Galaxy Morphology prediction

- Collected & processed data from *Sloan Digital Sky Survey* (multi TB image dataset) using a mix of *SQL* & *Python* libraries
- Implemented deep neural networks with skip-connections to predict hierarchical galaxy morphology classes for galaxies
- Implemented large scale distributed prediction for ~12 million galaxies in spark

#### Time series analysis driven Exoplanet Detection

- Collected & processed time-variant photometric flux data from Kepler to build an exoplanet detector based on regional dips
- Achieved classification accuracy of 70% (SVM), 85% (MLPs) & 92% (RNNs) for Kepler's database

#### Extracting text from degraded ancient manuscripts

- Trained a deep neural network for text extraction from heavily degraded ancient manuscripts using FRCNNs
- Implemented a complete ML-pipeline to analyze hyperspectral scans degraded manuscripts

### Network Analysis & Statistical modeling:



#### Analyzing large transportation networks

- Built a large, dense, time-variant geospatial transportation graphs using NYC's transportation data (1 billion+ taxi trips)
- Objective was to analyze travel habits of residents and determining pressure points in the network
- Used RNNs to predict demand at each node (RMSE = 2.01%) which was used to analyze network characteristics

#### Air quality prediction

- Trained Random Forest based ML regression model for predicting Particulate Matter<sub>2.5μ</sub> level in any given locality
- Built multiple possible candidate models to contrast interpretability vs accuracy and achieved RMSE of 6% (linear model), 2.5% (GBM) & 0.98% (Random forest)

## Professional Experience

### AXA, Data Scientist/Sr. Analyst

Dec'14 – Jun'16

- Ideated, championed and contributed in setting up AXA's Hadoop stack and associated toolset like hive, r, python, spark
- Predicting mortality rate using AXA's historical claims dataset
- Promoting Data science skillset and adoption across teams by conducting training sessions, best practices meetups, machine learning & big data tools workshops for AXA-US & AXA-India

### AbsolutData Research & Analytics, Analyst

Jun'13 – Dec'14

#### Sensor & transactional data analytics:

- Developed multi stage semi-supervised machine learning driven models for predicting equipment failure and maintenance
- Worked extensively as R & Analytics resource for the project using *R*, *Hive* & *Hadoop*.
- The data varied from GBs of human entered - work order & maintenance data to 13TBs of Sensor data

## Skills

### Programming

Python, R, SQL, Scala, Julia

### Toolset

Tensorflow, Pytorch, Keras, Opencv, Numpy, Pandas, Scikit-learn, Scikit-image, plyr, dplyr, Pyspark, graphX, HIVE/IMAPALA queries, MapReduce using Python, Tableau, ggplot2, seaborn, matplotlib for visualization

### Machine

### Learning

Experience in framing & solving business problems using machine learning techniques such as Random Forests, CNNs, RNNs, Graph convolutional networks, Support Vector Machines, GBMs, Linear Regression, Logistic Regression, Clustering Techniques (k-means, hierarchical clustering, knn etc), transfer learning