

Rahul Pande

linkedin.com/in/pande-rahul

Worcester, MA
rspande@wpi.edu | 508-333-9986

OBJECTIVE

Seeking full-time opportunities from May 2020 in the field of Data Science with a focus in Machine Learning

EDUCATION

- **Worcester Polytechnic Institute (WPI)** Worcester, MA
Master of Science, Data Science GPA: 3.88/4.0 May 2020
- **Vishwakarma Institute of Technology (VIT, Pune)** Pune, India
Bachelor of Technology, Electronics and Telecommunication GPA: 8.59/10 May 2015

COURSES

WPI: Deep Learning, Big Data Management, Statistical Learning **VIT:** Pattern Recognition, Linear Algebra

SKILLS

Programming: Python, R, SQL, Java **Technologies:** Hadoop, Spark, Lucene, Tableau, MongoDB, D3
Packages: scikit-learn, numpy, scipy, pandas, matplotlib, ggplot2, dplyr, caret, Dask, Tensorflow

EXPERIENCE

- **Data Science Intern** Quincy, MA
National Fire Protection Association June 2019 - December 2019
 - **Geospatial Analytics:** Geocoding 150M+ addresses using Spark, Lucene and libpostal. Analyzing incident trends and defining response boundaries for Fire Stations
 - **Software Engineering:** Deployed geocoding/reverse-geocoding APIs using Docker to facilitate consumption of services. Achieved 70% increase in speed with median distance <1 mile compared to Google Geocoding API
- **Graduate Researcher (Independent Study)** Cambridge, MA
Takeda Pharmaceutical January 2019 - May 2019
 - Large scale deep learning on UK Biobank actigraphy and phenotype dataset (500K subjects)
 - Achieved 34.8 rmse and f1 score of .78 in predicting age and insomnia respectively from the actigraphy data using multitask learning with 1D CNN, LSTM, GRU
- **Assistant at Academic and Research Computing** Worcester, MA
WPI August 2018 - June 2019
Developing and maintaining the website for Northeast Cyberteam (NSF Award Number: 1659377) to improve its ease of use (necyberteam.org)
- **Product Analyst** Pune, India
Vuclip India Pvt. Ltd. April 2016 - June 2018
Set up analytic processes and frameworks, and helped improve user engagement on the platform through effective use of machine learning.
 - **Complaints Prediction:** Developed a Random Forest model to predict if a user will complain (cross-validation test recall of 0.79). Decreased complaints from 11% to 7% avoiding hefty penalties.
 - **Content Recommendation:** Formed user clusters (k-prototype) based on the type and amount of content consumption and then recommended content based on collaborative filtering.
 - **Analytics:** Holistic content performance analytics related to Original Series content at Vuclip. Mining market data to improve content production/acquisition strategy
 - **Data Engineering:** Authored numerous Apache Airflow data pipelines to replace cron scripts which significantly reduced database locks, failures and execution time. Improved workload management on multiple servers

PROJECTS

- **Music Generation with Neural Nets** (WPI | March 2019): Generated synthetic music sequences (MIDI notes) using LSTM and GAN and outperformed \sqrt{n} randomized real sequences
- **Blind Source Separation** (WPI | December 2018): Separating mixed audio signals using Non-negative Matrix Factorization (NMF).
- **Website Insights** (Vuclip | 2017): Get insights on a web forum using Scrappy, NLTK and, D3.
- **Android Development** (Personal | 2013): Developed the official Android Application for VIT, Pune. Second rank in on-spot android app development competition.