

pandey.v@northeastern.edu

LinkedIn: [/in/viral-pandey](#)

Viral Pandey

Available: Jan 2021 | [\(617\) 955-0659](tel:(617)955-0659)

viralpandey.github.io

GitHub: [/viralpandey](#)

EDUCATION

Northeastern University, Boston, MA

Sep 2018 - Dec 2020

Khoury College of Computer Sciences | Master of Science in Data Science | GPA: 3.889/4.0

Teaching Assistant: Machine Learning, Data Management and Processing, Programming with Data

Relevant Coursework: Causal Inference, Deep Learning, Foundations of AI, Algorithms

Dhirubhai Ambani Institute of Information and Communication Technology, India

Aug 2014 - May 2018

Bachelor of Technology in Information and Communication Technology

Relevant Coursework: Database Management System, Stochastic Simulation, Models of Computation, OOP, Data Structures

Activities & Societies: Google Developers Group, DAIICT Chapter and YouthRun (Annual techno-cultural fest) organizing committee member

TECHNICAL KNOWLEDGE

Languages: Python, R, SQL, Java, MATLAB, C, C++

Libraries: Pandas, Scikit-Learn, TensorFlow, Keras, Plotly, Matplotlib, Numpy, Pytorch, Pyro, OpenCV, H3, Geopandas

Statistical Methods: Time series forecasting, Hypothesis testing, Classification, Clustering, Regression Analysis

Technologies: Airflow, Git, Hive, RStudio, Jupyter

WORK EXPERIENCE

Tesla, Inc., Palo Alto, CA

Aug - Dec 2019, May 2020 - Present

Data Scientist Intern with Charging Data and Modeling (CDM) Team

- Developed supervised regression models to predict congestion and determine the capacity expansion of Supercharger sites
- Built data pipelines to convert vector data of public roads into Uber's H3 hexagons. This helped me design and put Traffic Coverage and Road Coverage KPIs into production
- Designed a time series forecasting model to estimate quarterly energy usage at sites. This informed the estimation of \$ revenue from the entire Supercharger network for future quarters
- Quantified the population coverage of the world using geo-spatial data of population density per pixel of the world and isochrone coverage (areas within some minutes by driving) of sites
- Identified vehicles that might be involved in potential misuse of the Supercharger network. Proposed false positive scenarios as well as solutions to mitigate such incidents

Khoury College, Boston, MA

Jul 2019

Research Assistant

- Explored Procedure Learning problem of understanding the constituting key actions of complex tasks from instructional video data
- Assembled a Fully Convolutional Sequential Network (FCSN) that produces a compact summary of the procedure steps and their ordering needed to perform a complex task, as well as localization of these steps in videos

Dhirubhai Ambani Institute of Information and Communication Technology, India

Jan - Apr 2018

Data Science Research Intern

- Outperformed other algorithms in forecasting Remaining Useful Life of a jet engine based on NASA's time series dataset by developing a Recurrent Convolutional Neural Network (RCNN) based predictive model

Infoware, India

May - July 2017

Data Engineering Intern

- Implemented Hadoop infrastructure and used Hive on historical data of client's customers across different sales channels

PROJECTS

Quora Insincere Question Classification

Jan - Mar 2019

- Designed a supervised binary classifier to detect insincere questions on the Quora website and compared performances of algorithms such as SVM, CNN and LSTM RNN
- Performed TF-IDF vectorization, Sentiment Analysis using Python NLTK framework for gauging overall sentiment

Bankruptcy Prediction Using Various Classifiers

Jan - Mar 2019

- Tested Logistic Regression, Naive Bayes, LDA, QDA, SVM and feed forward Neural Networks on 5 years of Polish Companies Data of different econometric ratios to determine whether a company is going to go bankrupt
- Handled missing data using Mean value Imputation and SMOTE on training data to handle class imbalance
- Calculated Correlation matrix, performed PCA and cross validation for feature sub-setting