

Virtee Parekh

✉ virtee.parekh@gmail.com

☎ +1-585-733-9411

🏠 New York, NY

🌐 virteeparekh

👤 virteep

OBJECTIVE

Data scientist looking for a **full-time** role to employ my skills in **data science** and **machine learning**

WORK EXPERIENCE

Business Intelligence Analyst at Indeed.com

Feb 2019 – Present

- Developing a driver based model using statistical techniques to predict the number of reactivated customers and their corresponding revenue.
- Built a scalable revenue forecasting pipeline using multivariate time series analysis that helps the senior leadership plan investments better.
- Developed a quota pacing dashboard for the Sales team using predictive modeling that saved resources worth millions.
- Reduced the error of a forecasting model across markets from 20% to 7% that has a revenue impact in hundreds of thousands.
- Built a machine learning model to predict the number of requisitions made that facilitate better decision making for Sales.
- Conducted causal impact analysis to determine the impact of accepting of alternate payment methods on revenue.
- Performed survival analysis to develop an interpretable model that predicts lifetime value, affecting significant portions of the marketing budget.
- Led a machine learning study group for a team of 15+ members where I presented tutorials and curated learning resources.
- Winner of Product, Technology, and Engineering award in a company-wide hackathon for developing a resume information extraction tool.
- **Technical Stack:** Python, SQL, Machine Learning, Statistical Analysis, Tableau, AWS

Teaching Assistant at Rochester Institute of Technology

Jan 2018 – Dec 2018

- CSCI 720 - Big Data Analytics
- CSCI 420 - Principles of Data Mining

Business Intelligence Intern at Indeed.com

May 2018 – Aug 2018

- Built a revenue forecasting model with an error rate of 5% that reduced manual effort by 1 hour daily.
- Performed exploratory data analysis to address an unsolved business problem of building logic from complex source data.
- Restructured existing reports and data sources for performance improvements.
- **Technical Stack:** Python, SQL, Prophet, Time Series Analysis

Data Scientist Intern at ITT – Goulds Pumps

Feb 2018 – Apr 2018

- Prepared the dataset, including labeling the data and balancing the classes.
- Developed models that facilitate machine health monitoring and failure prediction using anomaly detection.
- **Technical Stack:** Python, R, Machine Learning, Anomaly Detection

EDUCATION

Master of Science in Computer Science

Rochester Institute of Technology

Dec 2018

Bachelor of Engineering in Computer Engineering

University of Pune

May 2016

PROJECTS

Evaluating Credibility of Web – Based News Articles

Aug 2018 – Dec 2018

- Aggregated a labeled dataset by scraping online news articles.
- Built a model to detect fake news by using feature extraction and deep learning which achieved a test accuracy of 91%.
- Improved the semantic understanding using word embeddings, thus boosting the accuracy of the model by 6%.
- **Technical Stack:** Python, Scikit-Learn, Keras, Deep Learning, Natural Language Processing

Exploring Latent Cognitive Biases in Employer Compensation

Aug 2018 – Dec 2018

- Performed exploratory data analysis to look for observable trends on companies that have a gender pay gap.
- Identified pay gap patterns that differ for various organizations and found a correlation of pay gaps with sectors, locations, and company sizes.
- Analyzed tweets related to the gender pay gap by performing sentiment analysis and generating word clouds.
- **Technical Stack:** Python, R, Tableau

Mood - Based Activity Recommendation System

Aug 2017 – Dec 2017

- Created a chatbot to detect the user's mood and recommend activities to enhance or alleviate the mood based on the emotion.
- **Technical Stack:** Python, IBM Watson

Detection of Abusive Content on Social Media Forums

Jan 2017 – May 2017

- Built a model that uses supervised learning to detect abusive text in social media comments with a resulting test accuracy of 80%.
- **Technical Stack:** Python, Natural Language Processing, Machine Learning

SKILLS

Programming/Scripting Languages

Python, R, Java, C++

Databases/Query Engines

PostgreSQL, Presto, Redash, Hive

Data Science Tools

Pandas, Scikit-learn, NLTK, Keras, PyTorch, TensorFlow

Visualization Tools

Tableau, Python, ggplot

Cloud Computing Platforms

Amazon Web Services