

HIGHLIGHTS OF QUALIFICATIONS

- Experience with interdisciplinary projects at the intersection of Engineering, Statistics and Data Science.
- Possess strong commitment to motivate the team dynamics with the ability to contribute expertise and collaborate efficiently.
- **Programming Skills:** Python, Excel, PostgreSQL, Spark, Scikit-Learn, NumPy, C/C++, Hadoop
- **Analytics Skills:** Statistical Inference, Machine Learning, Deep Learning, Big Data (Basic Knowledge)

EDUCATION

University of Toronto, Canada

2018/09 – 2019/12

Master of Engineering, Electrical and Computer Engineering (3.4/4.0)

SRM Institute of Science and Technology, India

2014/07 – 2018/05

Bachelor of Engineering, Electronics and Instrumentation Engineering (8.2/10.0)

INTERSHIPS

Analyst Intern

2016/06 – 2016/08

Jet Airways, India

- Worked with the Engineering team to develop and maintain the airline data with PySpark ML libraries in UNIX Operating System.
- Implemented the machine learning models to the airline data in order to improve the machine efficiency by 10%.

Doubtnut, Data Analyst Intern, Gurgaon

2017/06 – 2017/09

- Recommendation System for **40,065** student's user history, recommendations driven by **cosine similarity**; applied on decomposed dataset using SVM.
- Developed a traffic/conversion forecasting machine to capture potential customers in advance; Implemented **Decision Trees** on a dataset of **7,500 students** with an AUC score of 0.87.

RELEVANT PROJECTS

Fake News Detection Model – Data Cup Challenge

2019/12

- Performed data cleaning, EDA on the data extracted from the disparate data sources and implemented feature engineering by word embedding and vectorization.
- Implemented the machine learning models (Linear Regression, XGBoost, Linear-SVM and Random Forest) and improved the f1-score to 0.41 by hyper-parameter tuning using GridSearch. [GitHub](#)

2018 Kaggle DS & ML Survey Challenge

2019/10

- The data from Kaggle was cleaned and advance exploratory data analysis was performed in order to identify the patterns and trends. Feature Selection was performed by ANOVA test and PCA was implemented in order to reduce the dimensionality of the features.
- Developed different regression models (Linear, Lasso, Ridge and Elastic Net) for training the data with 10 folds while using R2 and RMSE as metric score to evaluate the performance of the models. Hyperparameter Tuning was also performed in order to tune the above techniques by GridSearch. Elastic Net outperformed the other models by 0.418. [GitHub](#)

Renewable Energy Based Smart Wheelchair

2018/04

- Implemented an Internet-of-things (IoT) Health Monitoring System on a wheelchair.
- The smart wheelchair also had a renewable powered energy source i.e., Solar Panels. The project was also selected for the “**Best Project in Medical Domain**” by the Department of Electronics and Instrumentation. [Report and Working Video](#)

RELEVANT COURSES

Data Camp

- Introduction to Python
- Intermediate Python for Data Science

Udemy

- Introduction to Databases and SQL Querying