

+91-9743614471

Bangalore, India

[Web-Portfolio](#)

NADEEM

nadeem2796@gmail.com

[nadeem-LinkedIn](#)

[nadeem-GitHub](#)

EDUCATION

- | | | |
|--|--|----------------------|
| • International School of Engineering, Data Science | <i>Carnegie Mellon University</i> | Nov 2019– June 2020 |
| • Bachelor of Engineering, Electrical & Electronics | <i>Visvesvaraya Technological University</i> | June 2014 – May 2018 |

TECHNICAL SKILLS

- **Programming Languages:** Python, R, SQL, Apache Spark, Java, C, HTML, CSS.
- **IDE, Applications & Cloud:** AWS, GCP, Git, Tableau, Visual Studio Code, Jupyter Notebook, Google Colab.
- **Data Science Libraries:** Pandas, NumPy, Sci-Kit Learn, TensorFlow, Keras, SciPy, Seaborn, Matplotlib, ggplot2, caret, Spacy, Plotly.

RELATED COURSEWORK

- **Web Technologies:** HTML, CSS, XML, JavaScript, Python, Pyspark, JSON, Flask, Web servers, REST API
- **Data Mining:** MapReduce, Frequent itemset mining, Recommendation systems, Analysis of social networks
- **Natural Language Processing:** Perceptron, BERT, Elmo, Viterbi algorithm, HMM, Neural Nets, NER, TF-IDF, Dependency parsing
- **Database Management:** RDBMS, EERD, Gremlin-Tinker pop, Geospatial data handling, PostgreSQL

CERTIFIED COURSES

- Applied Machine Learning in Python, University of Michigan, Coursera. [More](#)
- Deep Learning, DeepLearning.ai, Coursera. [More](#)
- Data Science, John Hopkins University, Coursera. [More](#)
- Advance Data Science with IBM, IBM, Coursera. [More](#)
- Google Cloud Platform Big Data and Machine Learning Fundamentals, GoogleCloud, Coursera [More](#)

PROFESSIONAL EXPERIENCE

Soothsayer Analytics (Sep 2020 – Present) [More](#)

Data Scientist

Client: SABIC Petro-Chemical Industry

Project: Plant efficiency and yield optimization of the plant

Responsibility:

- Data Extraction in PI Data links of Excel through RDC (VDI).
- Converting the extracted data to hive table using Dataiku.
- Pre-processing and data cleaning using Dataiku.
- Feature engineering and application of a collection of algorithms (Linear Regression, Decision Tree, Random Forest, ANN, ARIMA, and ARIMAX) to build the model.
- Prediction and forecast of the future values.

Tech Stack:

- Dataiku, Pyspark, Python, Tableau.
- PI Data Links (Excel).

PROJECTS

Predict Severity of Airplane Accidents. [Automobiles] [More](#)

- Predict Severity of airplane accident using 'Accident ID'. Required to build Machine Learning models to anticipate and classify the severity of any airplane accident based on past incidents.
- To analyze and implement multiple algorithms and determine which is more appropriate for a problem. Pre-process and clean the data for modelling. Train a classification model using Logistic Regression, SVM, Random Forest, XGBoost, AdaBoost, Decision Tree, Bagging Classifier, Voting Classifier.

Broadband Outage Detection. [Retail Industry] [More](#)

- Indian Broadband Company is Facing Outages duration, Task is to predict the outage duration with the features
- Modelling framework to predict classification of three class (0,1,2), Trained using Decision tree, Random Forest and XGBoost.

News Category Classifier. [Network Media and News] [More](#)

- This dataset contains around 200k news headlines from the year 2012 to 2018 obtained from HuffPost. The model trained on this dataset could be used to identify tags for untracked news articles or to identify the type of language used in different news articles.
- Categorization based Natural Language Processing, that would perform localization and recognition of text and categorize. Train a Classification Model using Text CNN, Bidirectional GRU + Convolution and LSTM.