VEDANT MEHTA

DATA ANALYST

CONTACT

vedantmehta08@tamu.edu
(979) 422-2501
College Station, TX
linkedin.com/in/vedantmehta08
github.com/vedantmehta2808

PROFILE

Data enthusiast looking for a co-op / internship opportunity in the field of Data Analytics

EDUCATION

DEC-2019
TEXAS A&M UNIVERSITY
[COLLEGE STATION, TX]
MS in Industrial Engineering
(Recipient of Competitive Grad Scholarship)

2016
PANDIT DEENDAYAL PETROLEUM UNIVERSITY
[GUJARAT, INDIA]
B. Tech. in Industrial Engineering

METHODOLOGIES

- Data Preprocessing
- Exploratory Data Analysis
- Supervised learning
- Unsupervised learning
- Neural network
- Experimental design

TECHNICAL SKILLS

- Statistics and Probability
- Python, R
- SQL
- Tableau
- Git, Spark
- Excel, Word, Powerpoint

PROJECTS

Predictive modeling for wind power estimation

- Data preprocessing including handling missing values, **feature scaling**, and handling outliers
- **Preparing the data** for predictive modelling
- Applied algorithms like **Support Vector Machine, Decision trees, Random forest** to build the predictive model
- Libraries used for building the machine learning model include **pandas, scipy, numpy, scikit-learn, seaborn, keras**
- Predicted wind power from the historical data and achieved a **RMSE loss of 0.035**

'vonMises' – Statistical package developed in Python

- Developed a package on the basis of **vonMises distribution** in circular statistics
- Consists of **4 main functions**, namely, **random number generator**, probability density, **distribution function**, quantiles
- Available to the whole **Python community for download** via pypi.org
- Achieved **100% accurate results** when compared to those in R

Human Activity Recognition (HAR) using Smartphones data

- Applied **feature engineering** on the dataset including finding the important features on the basis of **Pareto analysis**
- Predictive analysis on dataset contained **561 features** with **10799 instances** using R libraries like **dplyr, ggplot, gbm, caret, randomforest, neuralnet**
- Predictive model built to classify the type of activity from among six classes with an **accuracy of 95%**
- Applied algorithms like SVM, **Ensemble technique**, Random forest, PCA, LDA, QDA, **Logistic regression and KNN**

EXPERIENCE

2017
Process Analyst Intern | Cosmos Impex Pvt. Ltd.

- **Led a team of four interns** to complete the cycle time improvement project
- Manually **gathering data** and identifying **KPIs** pertaining to the production line
- Exploratory data analysis using **pivot table and pivot chart**
- Data analysis in Excel for **improving the cycle time**
- Improved the **cycle time by 23%** which led to **increase in production by 20%**