Pulkit Jain

Bryan, Texas - 77801 | (979) 985-8735 | pulkit.jain@tamu.edu
http://people.tamu.edu/~pulkit.jain, www.linkedin.com/in/pulkitjain1004, github.com/pulkitjain1004

Masters fellow in industrial engineering and member of population informatics research group. With 2 years of experience as analyst, providing business solutions by leveraging analytical and management skills.

EDUCATION

M. Eng. (Industrial Engineering), Texas A&M University, College Station, TX — **GPA 4.0/4.0**

Coursework: Data Analysis, Time-Series Analysis, Simulation Methods, Statistical Computations, Engineering Economy, Survey of Optimization — Aug 2016 – May 2018

B. Tech. (Mechanical Engineering), G.B. Pant University of Agri. and Tech., Pantnagar, India — **GPA 7.56/10**

Coursework: Numerical Methods, Linear Programming, Simulation, Probability & Statistics — Aug 2009 – Jun 2013

CORE SKILLS

- Statistical Modelling
- Machine Learning
- Time Series / Forecasting Analysis
- Linear Programming
- Optimization
- Regression, Hypothesis Testing
- Project Management
- Simulation Methods
- PCA, Six-Sigma Statistics

COMPUTER SKILLS

- R, Shiny (expert)
- Python, SQL, SAS certified
- C & MATLAB (prior experience)
- Apache Hadoop, Hive, Linux
- Tableau, Visual Basic
- Minitab, Simio
- Google Analytics, Git, HPC servers
- MS Access, Excel, Word, PowerPoint
- HTML, CSS, Bootstrap

CERTIFICATIONS

- Lynda certifications in Analyzing Big Data with Apache Hadoop, Hive & Pig. — Feb 2017
- Coursera certification in Machine Learning & Game Theory hosted by Stanford University. — Nov 2015
- Certificate In Project Management (CIPM), Institute of Projects & Program Management. — July 2015

WORK EXPERIENCE

Texas A&M Engineering Experiment Station — College Station, TX
Graduate Student Researcher – Data Analytics & Population Informatics — May 2017 – Present

- Using Predictive analytics to link multiple datasets with hospital and patient information.
- Combined Machine Learning algorithms with rule-based approaches to increase efficiency.
- Used **SQL** to extract data & **R** to build prediction models for individuals; through SVM, Random Forests & Neural Nets. Used **SAS** to link hospital databases in American Hospital Association.
- Used **Git** to synchronize local machine with TAMU's HPC server and enhance workflow tracking.

Steel Authority of India Limited — Bhilai, India
Project Analyst – Modernization / Expansion — Feb 2014 – July 2016

- Part of planning team responsible for over $3.5 million expansion projects.
- Resolved shortcomings in original execution plan by identifying redundancies in project commissioning. Mitigated company loss by inspecting project compliance with contract including CPM delay analysis of project.
- Used **SQL** to design the database & checked for spare inventory by cross-linking multiple projects. Performed cash flow analysis of new plan & created material requirements plan (MRP).

PROJECTS

Analyze Uber Data for Emerging Markets in NYC (2014 - 2015), Citadel — Feb 2018

- Used **Tableau** and **R** to account that Uber is servicing areas that were previously neglected by taxis. Analyzed geography, demography and correlation with other Taxi services.

Develop Strategic Insights about Home Loans Market using HMDA data, TAMU — Oct 2017

- Used **R-shiny** to build an interactive map, capable of stratifying loans and reveal market size / trends for entire US. Recorded defects / outliers in loan applications, with ability to update log registers and provide back-end support.

Time-Series Analysis for Forecasting Rice Prices, TAMU — May– July 2017

- Created ARIMA models by examining correlation and variance for last 40 years. Explained Price Shock of 2008 through curbs in international supply chain and speculations in futures markets.
- Used **R** to identify various price determinants of rice and create a multivariate time-series model.

Trends & Patterns in Cardiac Patients, TAMU — Jan– May 2017

- Used **R** to detect patterns in diagnostics of CABG surgery patients through "ApproxMAP". Wrote **SQL** queries to manipulate large data sets & created a "**Shiny**" web application for public use.
- Exhibited a poster presentation at Texas A&M Research Computing Week – 2017.

Prediction Model of Bike Demand for Capital Bikeshare, TAMU — Jan – Mar 2017

- Used **Python** (Pandas, SciPy, NumPy, Matplotlib) to develop various regression models.
- Used feature selection and achieved over 80% accuracy by classifying days and weather.