Apoorba Bibeka

♦ College Station, TX ♦ (979) 721-1612 ♦ abibeka@tamu.edu ♦ www.linkedin.com/in/apoorb ♦ https://github.com/Apoorba-Bibeka

SUMMARY

4 years of experience as a graduate assistant researcher at Texas A&M Transportation Institute. Extensive predictive modeling, data visualization and manipulation experience with R and Python. Seeking full-time opportunities in data science / analytics field starting May 2019.

EDUCATION

Texas A&M University, College Station, TX

M.S. in Industrial Engineering—Data Science (GPA 4/4)   May 2019
M.S. in Civil Engineering—Transportation Engineering (GPA 4/4)   Aug 2016

COURSEWORK

Predictive Modeling | Deep Learning | Design & Analysis of Experiment | Linear Programming | Integer Programming | Non-Linear & Dynamic Programming | Stochastic Processes | Anomaly Detection | Applied Bayesian Methods | Applied Multivariate Analysis | Distribution Theory | Theory of Inference | Applied Categorical Data Analysis | Statistical Computation

SKILLS

Programming Languages: R, Python, C++, JMP, SAS, AMPL, CPLEX, MATLAB
DBMS: PostgreSQL; **Big Data Tools:** Hive; **Cloud Platforms:** Amazon S3 & EC2
R Packages: data.table, ggplot, lattice, boot, glmet, tree, randomForest, gbm, xlsx
Deep Learning: keras; **Python Packages:** numpy, pandas, matplotlib, seaborn, sklearn, scipy, docx
Other Software/Tools: Jupyter Notebook, Spyder, IPython, RStudio, Visual Studio, Git, QGIS

PROJECTS

❖ **Taxi Demand Prediction using Long Short-Term Memory (LSTM) Recurrent Neural Network** *(Python: RNN)*
  o Developed an LSTM RNN to predict hourly taxi demand in a Chicago community area based on the hourly taxi demand for the last seven days.
❖ **Predictive Analysis of Bike Sharing Data** *(R: Regression Models)*
  o Analysed Capital bike-share dataset to improve demand prediction by using supervised learning methods such as random forest and support vector machines. Improved model prediction by using k-fold cross-validation.
  o Identified important demand predictors and developed insights that would help increase revenue.
❖ **Predictive Analysis of Median Chicago Taxi Trips and Fare** *(R: Random Forest, Decision Trees)*
  o Evaluated 110 million Chicago taxi trips between 2013 and 2017 using supervised learning and data visualization to improve Chicago taxi business. Expedited the analysis using High-Performance Computing (HPC) systems.
  o Conducted feature engineering to improve model prediction. Developed heat maps to see how the earning of a typical Chicago taxi vary over time of day and day of the week.
  o Recommended pickup locations for the different time of days for increasing earnings.
❖ **Evaluation of Best Chip Expansion Strategy** *(AMPL & CPLEX: Linear Optimization)*
  o Evaluated different strategies to increase revenue for a snack company by formulating and solving a linear programming problem. Recommended the optimal number of different snacks that need to be shipped from different production sites to customer location to minimize cost.
❖ **Anomaly Detection on a Multivariate Data using PCA and Hotelling T^2 Charts** *(Python: PCA)*
  o Developed a control chart for a manufacturing process to identify and remove out of control observations.
  o Conducted Principal Component Analysis to reduce noise and improve signal for the high dimensional dataset.
❖ **Macro-Level Analysis of Pedestrian and Bike Commuting using Self-Organizing Map** *(R: SOM)*
  o Analysed the correlation between pedestrian and bike commuting and other factors such as crime rate or poverty.

EXPERIENCE

Texas A&M Transportation Institute, College Station, TX   Mar 2015–Present
Graduate Assistant Researcher

❖ Connected Vehicle Applications Evaluation using Trajectory and Weather Data *(Hive, Python, Amazon S3 & EC2)*
❖ Guidelines for Semi-Autonomous Truck Operation on Texas Roadways using Full Factorial Design *(R, C++)*
❖ Assessing Environmental & Safety Impact of Semi-Autonomous Truck using Full Factorial Design *(R, JMP, C++)*
❖ Identifying Optimal Signal Controller Setting for Video and Radar Detectors using Full Factorial Design *(R, C++)*
❖ Evaluating Road User Cost of Construction using Simulated Dataset *(R, Python)*
❖ Exploring the Factors Associated with Near Rear-end Crashes *(R, High Performance Computing System)*

HONOR & AWARDS

Competent Communicator and Leader – Toastmasters International
Scholarship – Intelligent Transportation System (ITS)