Dwarkanath (DK) Prabhu

503 Cherry Street, Apt 120, College Station, TX77840 | Mobile: (979) 709-2312
dwarkanath.prabhu@gmail.com | http://dwarkanath.xyz

EDUCATION

Texas A&M University, College Station, TX   [Dec 2018]
M.Eng., Industrial Engineering, CGPA: 3.50
Courses: Machine Learning, Data Analysis for Engineers, Probability for Engineering
Decisions, Distribution Theory, Simulation, Advanced Quality Control

IIT Bombay, Mumbai, India   [Aug 2012]
B. Tech, Mechanical Engineering

SKILLS

- Languages: R, Python, SQL, C++, VBA, MATLAB, HTML, CSS
- Libraries: NumPy, Pandas, Scikit-learn, Keras, TensorFlow, NLTK
- Visualization: ggplot2, Shiny, Matplotlib, Seaborn, Tableau
- Machine Learning and Data Analysis: Linear Regression, Logistic Regression, Decision Tree, Random Forest, SVM, Boosting, Neural Networks, K-Means, kNN
- Statistics: Markov Chains, Monte Carlo Simulation, Queuing Theory, SIMIO, Minitab

DATA ANALYSIS EXPERIENCE

Melanoma Detection Using Deep Learning   [Aug 2018 – Present]

- Analyzing a labeled database of 10,000+ skin images to classify into 7 classes of lesions
- Augmented existing data using image transformations such as mirroring and random cropping
- Achieved validation accuracy of 83% using ResNet50 with weights pretrained on ImageNet

Fake News Detection Using Machine Learning (Class Project)   [Mar 2018 – May 2018]

- Trained machine learning models over a dataset of almost 50,000 labeled news items
- Encoded the headline and body text in two separate bi-directional LSTM and concatenated their output to combine with features extracted from headline and body
- Detected stance of the body text relative to the claim made in the headline to identify whether given headline-article pair agree, disagree, discusses the same topic, or are unrelated
- Improved accuracy by 11% over baseline model using a Softmax for classification

Classification of Radar Signals from Ionosphere (Class Project)   [Mar 2017 – May 2017]

- Loaded and cleaned data of 300+ pulses received from the ionosphere by 17 different antennas in a radar system with 34 attributes (2 per antenna)
- Applied classification methods in R such as logistic regression, SVM, trees, LDA etc. on the dataset to predict the signal quality as "good" or "bad"
- Achieved the best 5-fold cross validation accuracy of 92% using the Support Vector Machine method with radial kernel

Anomaly Detection (Quality Control) on Multivariate Data (Class Project)   [Oct 2016 – Dec 2016]

- Analyzed a dataset from a manufacturing plant with 209 dimensions and 552 samples in Python using numpy
- Conducted Principal Component Analysis (PCA) in Python to reduce dimensions from 209 to 4 using a scree plot with matplotlib
- Applied T^2 Phase I analysis with a Type I error of 5% recursively to remove out-of-control data till only in-control data was left

Drayage Vehicle Requirement with Demand Variability (Class Project)   [Mar 2017 – May 2017]

- Simulated a port using SIMIO for variable demand for loading and unloading of freight
- Determined the requirement of drayage vehicles using the modeled simulation to achieve 99% service level

PROFESSIONAL EXPERIENCE

Texas A&M University, College Station, TX   [Sep 2017 – May 2018]
Graduate Research Assistant

Automation of production line for a chemical manufacturer *(8-month project)*
- Conducted time study onsite to get at least 3 observations each of 15 processes
- Identified steps in processes that can be eliminated using automation
- Used triangular distribution to run 100 simulation runs after eliminating automated steps to estimate process times mean and standard deviation by central limit theorem
- Run Monte Carlo simulation over 1000 shifts using calculated post-automation process times to estimate work force requirements that can cover 99.7% (3 standard deviations) of shifts
- Recommended automation that cut mean process times by 12% and workforce requirement by 33%

Power Generation Capacity Planning *(conference paper)*
- Estimated water availability for power generation in all 254 counties of Texas using Texas Water Development Board data for the next 25 years
- Modeled a linear programming problem to minimize water utilization and cost for power generation in Texas using CPLEX Concert for C++
- Wrote C++ code to implement epsilon-constraint method to solve the linear programming problem and write solutions to a text file
- Read in solutions from text file using R and cleaned the data to make it suitable for visualization
- Loaded GIS data for Texas to plot solutions on the map using ggplot2
- Built an app using Shiny framework for R to display user-selected maps from a dropdown

Teaching CPLEX
- Wrote a manual on using CPLEX Concert Technology for C++ ([Download])
- Mentored a team from Universidad Autónoma de Yucatán to solve a multi-objective linear programming problem using CPLEX

Aasaanjobs.com (startup), Mumbai, India   [Jan 2015 – Jan 2016]
Marketing Manager

- Joined the core team of an early stage startup to head the marketing team
- Recruited and led a team of 6 content writers and paid marketing experts
- Investigated sources of user acquisition using SQL to optimize marketing spending
- Achieved 20,000 app downloads in three months since launch at an average cost of $1 per install

Deloitte, Gurgaon, India   [Jul 2012 – Dec 2014]
Strategy & Operations Analyst

Supply Chain Transformation for a leading footwear manufacturer *(10-month long project)*
- Identified quick-hit opportunities to reduce existing inventory
- Mapped current state value stream using Microsoft Visio for facilities in 2 cities and drafted a future state map to eliminate waste
- Used Excel Solver to map manufacturing location to demand center that minimized logistics cost
- Classified data of 100+ manufacturing facilities from SAP of finished goods, raw material and semi-finished goods inventory into 3 cost-based classes
- Formulated aggregate plans for each class of goods
- Created dashboards in Excel for production teams to generate schedule and procurement plan
- Improved service levels from 65% to 90% and increased rated production capacity by 20%

Mid to long term strategy in India for a leading automobile manufacturer *(4-month long project)*
- Forecasted demand in urban and rural markets for all vehicle segments for the next 5 years
- Created buyer personas for all car models using customer survey data from SPSS
- Developed business plan to acquire at least 15% market share in each segment