

# Dheeraj Singh

🏠 [dheeraj2444.github.io/](https://dheeraj2444.github.io/)

✉ [dhsingh@iu.edu](mailto:dhsingh@iu.edu)

☎ +1 (812) 361-7212

🐙 [Github](#)

🌐 [LinkedIn](#)

## EDUCATION

**School of Informatics, Computing, and Engineering, Indiana University, Bloomington, IN**

*Master of Science, Data Science*

May 2019

Current Coursework: Elements of Artificial Intelligence, Applied Algorithm, Introduction to Statistics

**Indian Institute of Technology (IIT), Kharagpur, India**

*Bachelor of Technology (Honours), Engineering*

May 2013

## COMPUTING SKILLS

Python, R, C, Shell Scripting, MySQL, PostgreSQL, SQLite, MongoDB, NumPy, Pandas, Scikit-Learn, Matplotlib, OpenCV, MATLAB,  $\LaTeX$ , Git, Vim, PHP, HTML, CSS, Bootstrap, Mac OS X, Linux, Windows

## RESEARCH EXPERIENCE

*Senior Project Associate*

April 2016 - May 2017

**Indian Institute of Technology (IIT), Kanpur, India**

**Vehicle Recognition System**

*Guided by: Prof. Gaurav Pandey (EE)*

- Developed a software system to identify license plate using template matching framework in Python
- Employed OpenCV library for image processing: Morphological transformations, Gaussian filtering, Adaptive histogram equalization, Contour formation, Character segmentation
- Formulated rules to build an OCR for characters identification and differentiation based on the pixel arrangement
- Enhanced character matching with pre-defined templates by performing threshold scaling for image binarization
- Improved recognition success rate by 7% as compared to the existing one( 83% vs. 76% ; Sample space = 1000)

**Data Visualization Application**

*Guided by: Prof. Arnab Bhattacharya (CS)*

- Developed a web-based user interactive application in PHP for real-time management and visualization of data stored in MySQL database; Implemented device responsiveness and interoperability using the Bootstrap framework
- Defined the complete database schema, configured, and deployed the same using phpMyAdmin
- Integrated Google chart API to visualize data variability in terms of distribution, trend, correlation, deviation, ratio

## PROFESSIONAL EXPERIENCE

*Senior Business Analyst*

May 2015 - February 2016

**Tinyowl Technologies, Mumbai, India**

*Food-tech start-up*

- Built a logistic regression model to predict the probability of users from different clusters returning back to the platform for targeted & channelized marketing using a *glm* package in R
- Built internal dashboard to track pre-defined business metrics and trends using *shiny* package in R
- Performed k-means clustering to segment user-base based on attributes like spending behaviour, ordering pattern

*Analyst*

June 2013 - December 2014

**Ipsos Research, Bangalore, India**

*Market Research firm*

- Performed Market Mix Modeling & Pricing Analysis for client specific marketing strategies in a variety of domains
- Quantified return on investments (ROI) from marketing expenditure by Regression modeling
- Awarded Spot performer of Q3-2014 for enthused performance in analytics division

## COURSE PROJECTS

- *Optimal Path Search*: Compared different graph search algorithms (Depth First Search, Breadth First Search, Uniform Cost Search, A-Star) to find the optimal path between a given pair of cities for different cost functions
- *Tweets Classification*: Predicted the city (class labels) of set of tweets by training a Naive Bayes Classifier. Implemented multinomial document model using bag-of-words and Laplace Smoothing
- *N-Queens*: Developed a N-Queen and N-Rook solver for a 8x8 chess board with additional constraints

## PERSONAL PROJECTS

- *Kaggle Competitions*: Built prediction and classification models using Decision Trees, Random Forest, SVM, Xgboost
- *Movie Recommendation System*: Used Collaborative filtering to recommend movies to similar users
- *Sentiment Analysis Tool*: Employed bag-of-words and rule-based approach to classify tweets into positive & negative
- *Speaker Recognition System*: Used K-means; MFCC & delta coefficients as features vector to recognize speaker