

# Dheeraj Singh

🏠 dheeraj2444.github.io/

✉ dhsingh@iu.edu

☎ +1 (812) 361-7212

🐙 Github

🌐 LinkedIn

## EDUCATION

### Indiana University, Bloomington, IN

Master of Science, Data Science, 3.8/4.0

May 2019

Coursework: Artificial Intelligence, Algorithms, Statistics, Data Mining, Exploratory Data Analysis, Machine Learning in Computational Linguistic

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

Bachelor of Technology (Honours), Engineering, 7.1/10.0

May 2013

## INTERESTS

Machine Learning, Deep Learning, Data Mining, Predictive Modeling, Natural Language Processing

## COMPUTING SKILLS

- **Programming Languages:** Python, R, C, SQL, MATLAB
- **Libraries and Software:** NumPy, Pandas, Scikit-Learn, Matplotlib, Jupyter Notebook, Tensorflow, Keras, NLTK, Spacy, ggplot, dplyr, L<sup>A</sup>T<sub>E</sub>X, Vim, Git
- **Databases:** MySQL, SQLite, PostgreSQL, MongoDB

## 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
- 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 such as 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

## SELECTED PROJECTS

- **Deep Learning Specialization:** Implemented forward & backward propagation in deep neural network from scratch, tweaked hyper parameters, optimization techniques, initializations, regularization techniques (Ongoing)
- **Image Orientation Classification:** Trained multiple models, namely, K-Nearest Neighbors, AdaBoost, and Neural Networks over 40K Flickr images to classify the orientation of 1,000 test images
- **Optimal Path Search:** Compared different graph search algorithms (DFS, BFS, Uniform Cost Search, A-Star) to find the optimal path between a given pair of cities for different cost functions
- **Part-of-Speech Tagging:** Developed a model to perform part-of-speech tagging using Hidden Markov Model and Bayesian inferences.
- **Tweets Classification:** Predicted the location (class labels) of set of tweets by training a Naive Bayes Classifier. Implemented multinomial document model using bag-of-words and Laplace Smoothing