

# Dheeraj Singh

Senior Project Associate  
Department of Computer Science & Engineering  
Indian Institute of Technology, Kanpur  
Kanpur, Uttar Pradesh, India, 208016

Email: [dheeraj2444@gmail.com](mailto:dheeraj2444@gmail.com)  
[dheeraj2444@outlook.com](mailto:dheeraj2444@outlook.com)  
Phone: +91-9916410277/+91-9987580804  
Homepage: <https://dheeraj2444.github.io/>

---

INTERESTS	Data Analysis, Data Mining, Predictive Analytics, Computer Vision, Natural Language Processing, Machine Learning
EDUCATION	<i>Bachelor of Technology (Honours) Chemical Engineering, 7.06/10.00</i> 2009 - 2013 <b>Indian Institute of Technology, Kharagpur, India</b>
EXPERIENCE	<div><div><i>Senior Project Associate</i> April 2016 - Present <b>Indian Institute of Technology, Kanpur, India</b> <b>Vehicle Recognition System</b> <ul style="list-style-type: none"><li>Developed a python-based software system to identify Indian vehicle license plate using template matching framework</li><li>Employed OpenCV library for image processing: Morphological transformations, Gaussian filtering, Adaptive histogram equalization, Contour formation, Character segmentation</li><li>Formulated rules to build an OCR system for identification of characters and differentiation of similar looking characters based on the pixel arrangement</li><li>Performed threshold scaling for binarization of gray scaled images to enhance the matching of segmented characters with pre-defined templates</li><li>New algorithm improved the recognition success rate by 7% as compared to the existing one( 83% vs. 76% ; Sample space = 1000)</li></ul></div><div><i>Data Visualization Application</i> Guided by: Prof. Gaurav Pandey (EE) <ul style="list-style-type: none"><li>Developed a web-based user interactive application in PHP for real-time management and visualization of data stored in MySQL database</li><li>Defined the complete database schema, configured, and deployed the same using phpMyAdmin</li><li>Implemented device responsiveness and interoperability using the Bootstrap framework</li><li>Integrated the Google chart API to visualize the variability of data parameters in terms of distribution, trend, correlation, deviation, ratio, and frequency</li><li>Automated the process to remotely refresh application backend data stored in server at another location every 24 hours</li></ul></div></div> <div><i>Senior Business Analyst</i> May 2015 - February 2016 <b>Tinyowl Technologies, Mumbai, India</b> Food-tech start-up <ul style="list-style-type: none"><li>Built a logistic regression model to predict the probability of users from different clusters returning back to the platform for targeted &amp; channelized marketing using <i>glm</i> package in R</li><li>Built an automated system to collate and quantify user sentiments on Twitter using <i>tm</i> package in R, thereby improving customer service and engagement strategies. [Link]</li><li>Developed an algorithm to identify top performing restaurants within each locality using scoring algorithm incorporating Gini coefficients and centroid method for weightage allocation</li><li>Built a internal dashboard to track various pre-defined business metrics and trends using <i>shiny</i> package in R</li><li>Recommended the marketing team to use targeted strategies by segmenting the user-base based on various attributes i.e. spending behaviour, ordering pattern using k-means clustering</li></ul></div> <div><i>Business Analyst</i> January 2015 - April 2015 <b>Embibe (Individual Learning), Mumbai, India</b> Education-tech start-up <ul style="list-style-type: none"><li>Tracked and analyzed day-to-day performance of metrics to gauge user traction, marketing impacts, user engagement, and product performance for making better business decisions</li><li>Performed cohort analysis to quantify user behaviour. Resulting inferences were used in targeted customer acquisition resulting in 23% increase in revenue within a span of 2 months</li></ul></div>

- Extracted raw data, analyzed and derived insights for pitching the organizational growth to investors
- Maintained Management Information System (MIS) to track all cash flows & other business activities

*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 like retail, consumer packaged goods, pharmaceuticals, and restaurant chains
- Quantified return on investments (ROI) from marketing expenditure by Regression modeling via tactics, brands, and geographies; traditional tactics (TV, Print, Display and Trade Promotions); non-traditional tactics (Paid Search, Emails, Social Media Ads)
- Necessitated Secondary Analysis of effectiveness and efficiency of marketing sub tactics to reach highly targeted customers
- Gained experience of the entire project life-cycle from conceptualization to the deck delivery on client-side by working on the entirety of a project
- Coordinated on a day to day basis with a global front desk team managing clients at New York and Connecticut
- Awarded Spot performer of Q3-2014 for enthused performance in analytics division

## PERSONAL PROJECTS

*Movie Recommendation System*

- Built a Movie Recommendation System in Python based on collaborative filtering algorithm that uses Euclidean Distance or Pearson Coefficient to find similar users and returns the list of top recommended movies for a given user. Takes method to be used and a number of movies to be recommended as command line argument. Used MovieLens movies rating dataset from GroupLens

*Speaker Recognition System*

- Built a speaker recognition system to identify time points of speaker change for a given conversation (used YouTube videos) using K-means clustering and MFCC & delta coefficients as features vector. Used *tuneR* package in R

*LinkedIn Job Scraper*

- Built a LinkedIn Scraper that takes job title e.g. Business Analyst, Machine Learning, or any combination of words and the number of pages to crawl as a command line argument from user and returns the list of all jobs along with job portal, job title, location, company name matching those keywords. Used *BeautifulSoup* module in Python

*Kaggle Competitions*

- Trained and tested numerous prediction and classification models using supervised and unsupervised learning, namely, Decision Trees, Random Forest, SVM, Xgboost. Performed K-fold Cross Validation to avoid over-fitting. Performed data manipulation techniques like missing value imputation, outliers removal, feature engineering
  - San Francisco Crime Classification
  - Walmart Trip Type Classification
  - Prudential Life Insurance Assessment
  - Airbnb New User Bookings

## ACADEMIC PROJECT & INTERNSHIP

*Undergraduate Dissertation*

August 2012 - March 2013

**Indian Institute of Technology, Kharagpur, India**

**Enhanced settling of particles in baffled settling tank**

*Guided by: Prof. B.Meikap (CH)*

- Compared and analyzed factors affecting particle settling viz, water temperature, particle size, particle type, currents type, etc.
- Proposed a new method by using equidistant inclined parallel plates placed along the length of the tank to enhance settling rates of various particles in waste water slurry emanating from steel plants
- Findings improved the efficiency of steel plants by reducing settling time and thereby facilitating the recovery and recycle of important metals

**Modeling of refinery for Optimization and MS blending**

- Conceptualized and designed a flow diagram of the refinery configuration based on mass balance
- Optimized naphtha flows for MS blending to maximize revenue within the constraints of public demand & quality specifications for a given simulation period using linear programming
- Implementation of model was successfully carried out for tank inventory management, product blending, final dispatch and refinery scheduling activities
- Achieved results similar to SIMPTO, software used by HMEI for scheduling
- Model was selected by the company to guide new trainees to perform simulations for monitoring day to day operations
- Only student in the department to be awarded a perfect score (10/10) for industrial training

**RELEVANT COURSEWORK** Mathematics-I, Programming and Data Structure, Mathematics-II, Transform Calculus, Numerical Solution of Ordinary and Partial Differential Equations, Advanced Mathematical Techniques in Chemical Engineering, Computer Aided Process Engineering

**TEST SCORES** GRE: Total- 310/340, Verbal- 150/170, Quantitative- 160/170, Analytical writing- 3/6  
TOEFL: Total- 104/120, Reading- 24/30, Listening- 28/30, Speaking- 26/30, Writing- 26/30

**TECHNOLOGY SKILLS** *Programming Languages:* C, R, Python  
*Web Development:* PHP, HTML, CSS, Javascript, Bootstrap  
*Libraries and Software:* NumPy, Pandas, Scikit-Learn, Matplotlib, OpenCV, MATLAB, L<sup>A</sup>T<sub>E</sub>X, Git, Vim  
*Databases:* MySQL, SQLite, PostgreSQL, MongoDB  
*Operating System:* Linux, Windows

**EXTRA-CURRICULAR ACTIVITIES**

- Kshitij, Techo-Management fest, IIT Kharagpur, India (2011-2013)
  - *Publicity Head*
    - \* Kshitij is Asia's largest Techo-Management fest, with a total participation of over 60,000 and prize money of over INR 5 million
    - \* Formulated an entire country-wide publicity campaign for the event, witnessing an increment of 8.3% in participation
    - \* Managed a budget of INR 1 million for the entire publications and Merchandise
    - \* Led a team of 25 members, 5 design team heads, and 10 design team members
  - *Core Team Member*
    - \* Raised funds close to INR 100 thousand by setting up meetings with alumni and companies in Hyderabad, India
    - \* Organized a workshop on Data Base Management System and Designing of Terra-drifters in association with Directi in Osmania University, Hyderabad with a footfall of 2500
    - \* Successfully carried out the publicity of fest in 10 colleges of Hyderabad, which involved giving presentations to around 400 students and media publication
    - \* Solely handled the end to end publications before and during the event
- Part of inter-hall gold winning Ad-design team of Patel Hall of Residence, IIT Kharagpur
- Completed one-year training in National Sports Organisation (NSO)- Hockey at IIT Kharagpur
- Completed one-year training in National Cadet Corps (NCC) at IIT Kharagpur
- Won district gold medal in boxing in Military School, Ajmer, India

**REFERENCES** **Prof. Arnab Bhattacharya** — arnabb@iitk.ac.in  
*Associate Professor, Computer Science & Engineering, IIT Kanpur, India*  
**Prof. Gaurav Pandey** — gpandey@iitk.ac.in  
*Assistant Professor, Electrical Engineering, IIT Kanpur, India*  
**Arnab Chatterjee** — arnab.chatterjee@ipsos.com  
*Senior Manager, Ipsos Research, New York, USA*