5385 Toscana Way #326
San Diego CA 92122

SAURABH GUPTA

(858) 291-2336
sag043@ucsd.edu
http://saurabhg.me

SUMMARY

Computer Science graduate student at UC San Diego with 2 years of professional experience in developing software products with focus on algorithms, modeling, data mining, machine learning and web development.

EMPLOYMENT

Data Scientist   **The Walt Disney Company India, Mumbai**   **Feb '15 – Jun '15**
- Built a Collaborative Filtering based **Recommender system** to suggest news articles to users
- Developed an "**autocomplete**" algorithm using Tries to give suggestions as users type in a search box
- Constructed a knowledge graph system in Neo4j by mining Wikipedia Infoboxes. Augmented this graph by deriving semantic relationships between named entities in textual data using **Natural Language Processing**

Senior Analytics Specialist   **Opera Solutions, New Delhi**   **Jul '13 – Feb '15**
Demand-driven scheduling of movies in multiplex
- Developed an optimization model to generate optimum movie schedules satisfying all cinema constraints
- Solved the optimization problem using a **Linear Programming** model and **Dynamic Programming** algorithm which made movie scheduling **50% faster** and resulted in **5% more profits** than manual schedules
- Developed an interactive web application in JavaScript and python for viewing/editing the movie schedules

Web content mining to generate competitive advantage for a real estate website
- Designed a web scraping framework in Python to crawl more than **50 websites** every month
- Integrated **edge detection** algorithms with **Random Forests classifier** to detect a pre-defined watermark in images with **95% accuracy** for checking unauthorized use of client's copyrighted images

R&D Intern   **Ecolab Technology Center, Pune**   **May '12 – Jul '12**
- Developed a multivariate statistical model (**nPLS regression**) to predict chemical oxygen demand(COD) of wastewater using fluorescence spectroscopy excitation emission matrices (EEMs)
- Reduced COD measuring time **from 2 hours to 10 minutes** by successful implementation of the model which enabled real-time monitoring of wastewater quality

OTHER PROJECTS

- **Apparel Classifier and Recommender using Deep Learning (2015):** Recognize clothing style and color in apparel images and recommend visually similar products [(Web link)](#)
- **Helpfulness prediction of User reviews ([Kaggle competition](#)):** Predicted whether a user's review of a product will be considered helpful using **Gradient Boosted Trees (Rank: 5/215)**
- **Rating prediction of Amazon reviews([Kaggle competition](#)):** Predicted star ratings for unseen user-item pairs using Matrix Factorization techniques **(Rank: 1/132)**

DEVELOPMENT SKILLS

- *Languages/Tools:* Python, Java, C++, R, MATLAB, Go(*Basic*), Ruby(*Basic*)
- *Web Development:* HTML, CSS, JavaScript, PHP
- *Web Frameworks:* Flask, Django, Ruby on Rails(*Basic*)
- *Databases:* MySQL, NoSQL, Neo4j(Graph DB)

EDUCATION

La Jolla, CA   **University of California San Diego**   **Expected Dec '16**
- M.S. in Computer Science. **GPA: 4.0/4.0**
- Coursework: Learning Algorithms, Neural Networks and Pattern Recognition, Algorithm Design and Analysis, Data Mining, Principles of AI: Probabilistic Reasoning

Mumbai, India   **Indian Institute of Technology Bombay**   **Jul '09 – Apr '13**
- B.Tech in Chemical Engineering. **GPA: 9.0/10**
- Relevant Coursework: Data Structures & Algorithms, Machine Learning, Multivariate Statistics, Optimization