155 Claremont Avenue, 766
New York, NY 10027

Amla Srivastava

as5196@columbia.edu
(913) 278-7823

EDUCATION

Columbia University, New York **Expected Dec 2017**
M.S. in Data Science GPA: 4.0
Coursework: Machine Learning, Probability, Statistical Inference, Data Visualization, Storytelling with Data, Computer Systems, Natural Language Processing
Involvement: Engineering Graduate Student Council, Dodge Academic Tutor

PEC University of Technology, Chandigarh **May 2015**
B.E. in Computer Science & Engineering GPA: 9.4
Coursework: Algorithms, Databases, Data Structures, Operating Systems, Neural Networks, Artificial Intelligence
Involvement: SPDC & DASA Scholar, Alumni Affairs, Public Relations, Entrepreneurial Cell

WORK EXPERIENCE

NBC Universal, Data Science Intern, New York **Sep 2017 - Present**

- Working with the MSI team to provide data science services to business units within NBC Universal

The Hartford, Data Science Intern, Hartford **May – Aug 2017**

- Developed an unconstrained loss model to enhance Class Plans using H2O and Python; improved Gini score by 4%
- Designed a step-up model for Variable Annuity policyholders with gradient boosting; achieved an AUC score of 0.81
- Built a web application for recommending Increased Limits Factors (ILFs) to determine premium using R and R Shiny

Columbia University, Teaching Assistant, New York **Jan – May 2017**

- Instructed students and designed assignments in Python computing for Applied Sciences under Prof. Daniel Bauer
- Held weekly office hours to answer questions and work with students in a 1-on-1 setting; graded exams

McKinsey & Company, Jr. Research Analyst, Gurgaon **Jun 2015 – Jun 2016**

- Provided research and advisory services for a global Consumer Electronics client
- Analyzed consumer behavior and competitive market trends; developed branding and retail strategies

McKinsey & Company, Intern, Gurgaon **Jan – Jul 2014**

- Studied 'Text Mining', performed resume classification and sentiment analysis using R, RapidMiner and Tropes
- Built a request classification tool in Excel/VBA and RapidMiner to track team performance with an accuracy of 94%

DATA SCIENCE PROJECTS

Traffic Complaints Classification: *Applied Machine Learning* **Apr 2017**

- Built a logistic regression model to classify traffic complaints in Boston using scikit-learn
- Improved model using topic modeling (LDA, NMF) and k-means clustering; achieved a final accuracy of 92%

Vehicular Collisions Visualization: *Data Visualization* **Mar - Apr 2017**

- Created interactive heatmaps, bar charts, etc. using R, Tableau and CartoDB to explore NYPD accident data
- Visualized trends in accident types, seasonality, causes, and risks between 2013-16

Consumer Conversion Prediction: *Applied Machine Learning* **Mar 2017**

- Constructed an ensemble model using logistic regression, SVMs and gradient boosted trees in scikit-learn
- Predicted subscription rates for the bank's direct marketing campaign; achieved an AUC score of 0.83

Character Interactions Visualization: *Storytelling with Data* **Dec 2016**

- Created interactive bubble charts and network graphs using R, D3.js , jQuery and Bootstrap to explore character interactions across scenes in the film 'Love Actually'

Most Popular Product Prediction: *CDSS Microsoft Data Science Hackathon (2nd Place, $3000 prize)* **Oct 2016**

- Built a softmax regression model in Python and Azure ML Studio to predict most popular products in Walmart stores

SKILLS

- **Languages:** Python, R, C, C++, Java, SQL, PHP, HTML, VBA
- **Frameworks and Libraries:** scikit-learn, keras, NumPy, Matplotlib, Hive, Hadoop, H2O, ggplot2, R Shiny, D3.js
- **Tools:** Tableau, RapidMiner, CartoDB, Git, MS Office
- **Status:** U.S. Citizen