Rajat Aghi

rajataghi@gmail.com | 240-467-8864 | rajataghi.github.io

Education

MASTER OF INFORMATION MANAGEMENT, DATA ANALYTICS **EXP. MAY 2018**
UNIVERSITY OF MARYLAND, COLLEGE PARK | GPA: 3.85/4.0

BACHELOR OF ENGINEERING, COMPUTER SCIENCE **JUNE 2015**
GURU GOBIND SINGH INDRAPRASTHA UNIVERSITY, INDIA | GPA: 3.6/4.0

Work Experience

DATA ANALYST INTERN | USDA – AGRICULTURAL RESEARCH SERVICE **JUN '17 – AUG '17**

- Compared nutrient compositions on food labels with lab values, for multiple food products using various **hypothesis testing** methods such as t-tests and ANOVA.
- Implemented **regression modeling** to calculate missing nutrient values for food items by correlating them to other foods with similar nutrient composition.

BUSINESS INTELLIGENCE ANALYST | ASPIRING MINDS ASSESSMENTS LTD. **JUL '15 – JUL '16**

- Initiated and managed the sales pipeline of 15+ sales managers in Excel. Applied regression analysis to predict manager wise expected revenue.
- Analyzed hiring data **via SQL** and generated business intelligence dashboards based on client requirements. Helped reduce hiring costs by up to $12,000.
- Formulated a **predictive model** to identify the optimal time to air an advertisement for maximizing hits on the website. Leveraged Google Analytics to collect website data.

Academic Experience

PREDICTING CONSUMER COMPLAINTS AGAINST FINANCIAL INSTITUTIONS

- Utilized CFPB's consumer complaints data set to develop predictive models using neural networks, decision trees and SVMs.
- Trained the models on around 200,000 data points to get prediction **accuracy of up to 80%**.

COMPARE NEWS SOURCES TO IDENTIFY INCIDENCES OF COLLUSION

- Parsed 1.2 million articles from 4 news sources via the Diffbot API. Stemmed & cleaned data for analysis.
- Applied **n-gram analysis** and fuzzy matching to compare articles. Created **Tableau dashboards** to summarize data and identify trends. Presented the analysis to reporters from Washington Post.

IDENTIFYING NEIGHBORHOOD BOUNDARIES USING MOBILITY DATA

- Scraped 5 million records from Car2go and Capital Bikeshare on AWS via Pandas and BeautifulSoup in Python. Applied DBIndex and k-means clustering to devise neighborhood boundaries for the D.C. region.
- **Presented results** at the pre-conference symposium for the International Cartographic Conference 2017.

Technical Skills

- **Programming Languages:** Python, R, SQL, HTML, JavaScript
- **Tools:** Tableau, AWS, Jupyter Notebook, Excel VBA, Google Analytics
- **Databases:** MySQL, MongoDB