

# Harman Shah Singh

510-779-4848 • [harman\\_shahsingh@berkeley.edu](mailto:harman_shahsingh@berkeley.edu) • <https://www.linkedin.com/in/harman-shah-singh>

## EDUCATION

---

### University of California, Berkeley

*Master of Information Management and Systems* (GPA - 3.71/4.0)

Courses : Applied Machine Learning, Applied NLP, Data Mining & Analytics, Quantitative Research Methods, Haas@Work

Berkeley, CA

August 2016 - May 2018 (Expected)

### Guru Gobind Singh Indraprastha University

*Bachelor of Technology, Computer Science and Engineering* (Percentage - 78.46%)

New Delhi, India

August 2010 - June 2014

## EXPERIENCE

---

### Autodesk, Inc.

*Intern Data Scientist - Product Intelligence*

- Harness the power of unsupervised machine learning techniques to identify user clusters based on the product usage patterns, and use the insights gathered from these clusters to provide recommendations to product teams for increased user adoption/retention.
- Develop custom product metric aggregations in Spark tables and design Oozie workflows to automate the data onboarding processes.
- Design dashboards in Looker to analyze user activity and product performance over time and across different geographies.

San Francisco, CA

May 2017 - Present

### Intel Corporation

*Business Technology Analyst (Part Time)*

- Conducted a market-landscape analysis of the global IoT industry to identify discontinuities, potential areas of market immersion and sources of incremental opportunities to strategically nurture a robust IoT developer network and accelerate IoT commercialization.

Santa Clara, CA

January 2017 - May 2017

### University of California, Berkeley

*Graduate Student Instructor, Department of Statistics*

- Work with faculty members to conduct weekly sessions and create information manuals for assisting students with the fundamental concepts of probability and statistics, such as regression, random variables, sampling, and discrete and multinomial distribution.

Berkeley, CA

August 2016 - Present

### Publicis.Sapient (Formerly Sapient Corporation)

*Data Science & Analytics Associate*

- Performed regression analysis to predict an employee's exit probability, reducing employee attrition rates by 17% from 2015 to 2016.
- Devised the capacity planning and revenue maximization strategy by creating novel predictive analytics models using R.
- Led the development of an intensive business reporting system using Hadoop, SQL Server 2012 and Tableau. The system resulted in an 11% increase in the number of new businesses generated, with significant improvements in performance and reporting flexibility.
- Provided solutions for data extraction, organization and retrieval problems using NoSQL and MapReduce, resulting in a 40% improvement in the process turnaround time and a cost reduction of US \$100,000 in the client's organization and retrieval processes.

Gurugram, India

June 2014 - July 2016

## PROJECTS

---

### Deconstructing The 2016 U.S. Presidential Elections using Tweet Analysis | *Natural Language Processing*

- Used NLP techniques like hashtag parsing, sentiment analysis and named-entity recognition, along with machine learning techniques to analyze how tweets from different parts of the country represented viewpoints and shaped the 2016 US Presidential Election results.

### Predicting Student Dropouts in MOOCs Using RNNs and Ensemble Methods | *Machine Learning*

- Used machine learning techniques such as Recurrent Neural Networks and Ensemble Methods to conduct time-series analysis and predict a student's likelihood of dropping out of a MOOC on a weekly basis, based on his activities and interaction within the MOOC.

### Predicting Medication Change and Hospital Readmissions in Diabetic Patients | *Machine Learning*

- Used data mining techniques such as feature engineering, data balancing and interactions, along with robust machine learning models to predict medication change and hospital readmission in diabetic patients, and interpret these models to devise corrective measures.

### Boston AirBNB Listings Activity Analytics | *Data Mining, Predictive Modeling & Representation Learning (Word2Vec)*

- Used data mining and machine learning techniques like feature engineering and ensemble methods to predict the price and availability of Airbnb listings in Boston, and provide vibe-based neighborhood recommendations for customers using clustering and Word2Vec.

## SKILLS AND AWARDS

---

**Technologies :** Machine Learning, Natural Language Processing, Spark, Big Data using Hadoop (HDFS, Hive, Sqoop, Pig, MapReduce, Oozie), NoSQL Data Stores, Data Analytics, Data Visualization

**Programming :** Python, R, C, C++, Java, HTML

**Databases :** Microsoft SQL Server 2012, PostgreSQL, Oracle 10g

**Tools :** Jupyter Notebooks, SQL Workbench, Qubole, RStudio, Tableau, Looker, Microsoft Power BI, Toad Data Point

**Awards :** Rookie Of The Year ( December 2015), Emerging Data Scientist (October 2015) - Publicis.Sapient