Siddharth Arora

in linkedin.com/in/sidd-arora   github.com/sarora   sarora.github.io
+1 (438) 342 8552   @ sidd.arora.05@gmail.com
Canada

A data scientist with 2-years experience, and a background in software development, skilled in programming and machine learning.

My interest lies in developing robust, scalable data pipelines, and using analytics to extract key insights from large datasets to solve real-life business problems. I excel at succinctly communicating these insights to broader, non-technical audiences.

TECHNICAL SKILLS

Software and Programming Languages	Python (numpy,pandas, scikit-learn, keras), R (tidyverse, ggplot,forecast, Shiny)
Machine Learning	Classification, regression, clustering, feature engineering, hyper-parameter tuning
DataBases	Oracle, PostgreSQL, MS SQL, MySQL
Statistical Methods	Hypothesis testing, confidence intervals, survival analysis, principal component analysis and dimensionality reduction, Bayesian inference
Cloud Services	AWS (EC2, S3, EMR), Docker
Workflow Tools	Agile, Jira, Team Foundation Server, GitHub, Trello, RStudio, Visual Studio, PyCharm

EDUCATION

Aug 2017 - Jun 2018	**Master of Data Science** University of British Columbia, Vancouver, Canada
Aug 2009 - Aug 2013	**Bachelor of Engineering with First Class Honors** - Software Engineering University of Waikato, New Zealand

EXPERIENCE

Nov 2018 Present	Data Scientist | Stradigi AI, Montreal, Canada **Stradigi AI** provides state of the art AI platform that empowers users to build and execute solutions to complex business challenges faster and in a secure way.

In a team of 9:

- Create data pre-processing and transformation libraries for data cleaning and manipulation using python for Stradigi's AI platform.
- Benchmark and analyze Stradigi's Machine Learning algorithms against the current state of the art to ensure cutting edge technology.
- Build Machine Learning (ML) pipelines for Straidgi's AI platform, from data ingest through to solutions for specific use cases like sentiment analysis, document classification, name-entity recognition, image classification, image segmentation, etc.
- Work closely with Researchers and AI Developers in an Agile setting to ensure production quality software, and efficacy of Machine Learning models being used.

`Python` `Pandas` `Numpy` `Tensor Flow` `Docker` `BitBucket` `Scikit` `AWS`

Apr 2018 Jun 2018	Data Science Intern(Industry Capstone Project) | Unbounce, Vancouver, Canada **Unbounce** provides an online platform that enables marketers or digital agencies to create, publish, and test landing pages, pop-ups or sticky bars to improve conversion rates.

In a team of 4:

- Applied survival analysis methods and techniques to determine key factors in customer attrition.
- Automated the data pipeline infrastructure to extract and transform data into usable components and integrous components.
- Utilized easily decipherable visualizations to report findings to people with little statistical knowledge or non-technical background.

`Python` `Pandas` `R` `Survival Analysis` `Trello` `Github`

Siddharth Arora

in linkedin.com/in/sidd-arora github.com/sarora sarora.github.io
+1 (438) 342 8552 @ sidd.arora.05@gmail.com
Canada

May 2016
Dec 2016

Data Scientist (Contract) | Brewbound.co.nz , Auckland, New Zealand

Brewbound is an e-commerce startup providing fast online liquor delivery service in New Zealand.

In a team of 2:

> Cleaned and wrangled data on a regular basis to generate meaningful datasets from marketing data.
> Identified trends, patterns, and relationships in marketing data to provide answers to business questions, particularly regarding the effectiveness of various advertising channels.
> Communicated findings to key decision makers to help facilitate the decision making process.

R | Shiny | MS SQL | Trello | Team Foundation Server

Feb 2013
Sep 2016

Software Developer | IMS Health , Auckland, New Zealand

IMS Health (now IQVIA) is a leading global company providing clients in the healthcare industry with comprehensive solutions to measure and improve their performance.

> Developed and updated functionalities for IMS Health's Nexxus Mobile Intelligence - a CRM solution, that helps drive customer engagement for life sciences industry using C#, JavaScript, XML.
> Performed maintenance and bug fixing on the existing code base to ensure production quality software.
> Participated in daily stand-ups and bi-weekly sprint meetings in an Agile setup and developed interpersonal skills.

C# | Oracle | XML | JavaScript | Angular.js | Agile | Jira

✚ PROJECTS

Ongoing

Twitter Sentiment Analysis | Personal Project.
A project to analyze tweets and classify them as either a positive or a negative sentiment.

> Queried the Twitter API via Tweepy - a python library, to get the most recent tweets.
> Applied machine learning classification algorithms via the Natural Language Tool kit (NLTK) to classify the tweets as either positive or negative.

Python | Pandas | GitHub | Tweepy | Machine Learning

Apr 2018

Social Standards and Salary Inclinations Study | Study conducted as part of Masters course.
A hypothesis to determine whether a person's social standards is correlated with a person's drive for financial wealth.

In a team of 3:

> Created, and conducted a survey on SurveyMonkey collecting responses from over 100 participants.
> Cleaned and wrangled the data into a usable format, checking for anomalies, duplication, and missing values using R. Particular attention was paid to anonymizing personal identifiable information.
> Applied logistic regression to model the data. The study showed that there was not enough evidence to prove a correlation between person's social standards and drive for financial wealth .

R | Statistical Modeling | GitHub

For more projects and blog posts please visit sarora.github.io