Siddharth Arora

in linkedin.com/in/sidd-arora   ⌗ github.com/sarora   ⌂ sarora.github.io
☐ +1 (438) 342 8552   @ sidd.arora.05@gmail.com
⦿ Toronto, Canada

A seasoned full-stack machine learning engineer skilled in creating robust, scalable data pipelines and utilizing machine learning techniques to derive critical insights.

▤ TECHNICAL SKILLS

Programming Languages	Python (numpy, pandas, scikit-learn, keras, tensorflow, open-cv, spacy, pytorch)
Machine Learning	Classification, regression, forecasting, clustering, feature engineering, hyper-parameter tuning
DataBases	PostgreSQL, SnowFlake
Statistical Methods	Hypothesis testing, confidence intervals, survival analysis, principal component analysis and dimensionality reduction, bayesian inference
Cloud Services	AWS (EC2, ECR, S3, SFN, Athena, CloudWatch)
DevOps/Orchestration	Travis CI, Jenkins, Docker, Docker-Compose, MetaFlow
Workflow Tools	Agile, Jira, Confluence, Notion, BitBucket, Github, Team Foundation Server, Trello, RStudio, Visual Studio, PyCharm, VS Code, Lucid Chart

💼 EXPERIENCE

Jun 2021
Present

Senior Machine Learning Engineer | Realtor.com, Remote (Toronto, Canada)
Realtor.com offers tools and professional expertise to help people move confidently through every step of their home journey.

> Designed, developed and deployed a Machine Learning (ML) pipeline to forecast monthly leads in US zip codes for next year inventory planning decreasing complexity, reducing run time, and improving accuracy.
> Performed ad-hoc analysis tasks and communicated findings to key decision makers to help facilitate the decision making process.
> Performed code and design reviews to ensure a healthy codebase and prevent bugs, all while fostering knowledge sharing and collaboration.

`Python` `Pandas` `Numpy` `PyTorch` `Docker` `Github` `Scikit` `AWS` `MetaFlow`

Jun 2020
Jun 2021

Data Scientist II | Intact Labs, Toronto, Canada
Intact Financial Corporation is the largest provider of property and casualty (P&C) insurance in Canada and a leading provider of specialty insurance in North America.

In a team of 5:
> Built a ML pipeline to predict business retention used to maximize pre-determined objective functions like profit, growth in units, growth in premium, etc.
> Monitored and diagnosed models in production to examine for ML drifts and took measures to ensure pre-determined desired level of performance.
> Reviewed pull requests from team members and provided appropriate and accurate feedback to identify bugs, ship stronger code, improve knowledge sharing and more.

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

Siddharth Arora

in linkedin.com/in/sidd-arora   ⌂ github.com/sarora   🏠 sarora.github.io

📱 +1 (438) 342 8552   @ sidd.arora.05@gmail.com

📍 Toronto, Canada

Nov 2018 May 2020	**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 5:

› Designed and developed data pre-processing and transformation libraries using python for data cleaning and manipulation using python for Stradigi's AI platform.

› Built Machine Learning (ML) pipelines using classical and deep learning algorithms from data ingest through to solutions for specific use cases like sentiment analysis, document classification, time-series forecasting, object detection.

› Devised and developed a benchmarking framework to analyze and optimize Stradigi's Machine Learning pipelines against competitors.

› Worked closely and collaborated with Researchers and AI Developers to ensure production quality software.

Python Pandas Numpy Tensor Flow Docker BitBucket Scikit AWS

Apr 2018 Jun 2018	**Data Science Intern (UBC 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:

› Developed and automated the data pipeline using python to extract and transform data into usable components for modeling.

› Applied survival analysis modeling techniques using R to determine underlying behaviors that impact customer retention at Unbounce.

› Utilized easily decipherable visualizations to report findings to a non-technical audience.

Python Pandas R Survival Analysis Trello Github

Sep 2013 Feb 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

🎓 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

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