Rohan Singh

Email : singhrohan@outlook.com
Linkedin: https://www.linkedin.com/in/singhrhn
Mobile : +1-312-885-0056
Github: https://github.com/ygivenx

Experience

- **IBM** — San Francisco, CA
  Senior Data Scientist & Managing Consultant — *July 2020 - Present*
  - **Bias in actuarial models**: Applied several research papers to understand the bias in actuarial data and understand its impact in light of COVID-19 and Black Lives Matter movement. Analyzed for factors such as credit score and traffic police stops. Also, designed a data pipeline to include fairness metrics in current models.
    Technology Used: Python, Deep Learning, GANs, Docker
  - **Covid Recovery Index**: Lead a team of 4 to develop a COVID impact planning strategy for one of the top automobile insurer in the US by predicting the impact of COVID on driving at a county level with a MAPE of 10% for 500 large US counties using ping data from cellphones and other socio-economic and epidemiological variables. The aggregated signal was also fed into a demand sensing models at a state level.
    Technology Used: python, pyspark, fbprophet, Tableau, postgres, IBM Cloud - Object Storage/Analytics Engine, SQL
- **IBM** — San Francisco, CA
  Data Scientist & Senior Developer — *July 2018 - June 2020*
  - **Advanced Automation Platform**: Conceptualized an AWS Serverless architecture to automate claims processing for one of the top insurers in the United States. Lead and worked with a team of developers to implement and roll out to production within 4 months. Improved straight-through processing by 30%.
    Technology Used: python, SQL, sqlalchemy, Terraform, Gitops, AWS lambda, SQS, DynamoDB, Step Functions, ECR, APIGateway
  - **Legal Billing Anomaly Detection**: Developed a proof of concept and later helped implement a full product for a Fortune 10 company to help them find anomalies in legal billing data using NLP and multi-class classification. Helped the client save over 20% in legal bills and reduce the review times by 90% per invoice for a particular subset of the legal billing category.
    Technology Used: python, spacy, NLP, flask
  - **Search Optimization**: Lead the effort to apply machine learning and Lucene to optimize the enterprise search i.e elasticsearch of a leading financial services client.
    Technology Used: python, flask, NLP, RNN, elasticsearch
  - **Natural Language Classification**: Developed a state-of-the-art classification system which classifies queries generated in natural language to a set of defined categories. Also, developed infrastructure using scala and AKKA http to support the machine learning models to be served on the network using a REST API.
    Technology Used: scala, Spark ML, NLP, Akka Http, Java
- **Citadel LLC.** — Chicago, IL/ London, GB
  Software Engineer — *Mar 2014 - Apr 2017*
  - **Data Quality Monitoring System**: Developed an automated monitoring system to check the quality of external data from over 50 vendors using state-of-the-art statistical data analysis, resulting in 200% decrease in turnaround time in data check validations.
    Technology Used: python, perl, bash, SQL Server
  - **Commodity Trading Modelling**: Architect-ed a front-end interface, enabling the trading team to run models on-demand and make quick decisions. Revamped the legacy MS-Excel analysis system to use python along with modern scientific packages, reducing the model development and testing time exponentially.
    Technology Used: Django, python, SQL, HTML
  - **Future Pricing Model**: Implemented a Future Value model to generate trading signals for soft commodity futures.
    Technology Used: python, Excel Solvers

Education

- **Carnegie Mellon University** — Pittsburgh, PA
  Master of Information Systems Management; GPA: 3.87 (Highest Distinction, Dean's List) — *May - 2018*
  Courses: Machine Learning, Distributed Systems, Database Management, NoSQL, Statistics, Data Structures, Optimization, Economic Analysis, Practical Data Science, Managing Disruptive Technologies, Data Mining
- **Birla Institute of Technology** — Mesra, India
  Bachelor of Computer Science; GPA: 3.5 (7.70/10.0 - First Class with Distinction) — *June - 2011*

Skills Summary

- **Languages**: Python, SQL, Java, Scala, Go, R, C++
- **Tools & Skills**: Data Science, Machine Learning, Analytics, Data Visualization, Deep Learning, Statistics, Data Mining, Terraform, Spacy, Spark, Tableau, Keras, TensorFlow, GIT, MS-Excel, IBM Cognos, Pytorch, Data Version Control, MLFlow
- **Databases**: Postgres, DynamoDB, ElasticSearch, MySQL, MS-SQL, Redis
- **Cloud**: AWS, GCP, IBM Cloud