# Kiran Rawat

100 Western Battery Road, Toronto, Canada, M6K3S2 | (236) 866-1611| rawat.kiran111@gmail.com

## CAREER PROFILE

Master's graduate in Big Data with 2 years of experience implementing data-driven solutions and 3 years of experience developing software for finance and mining domains, including team sizes from 5-20 people. Developed data pipelines, analyzed complex large volumes of data and built predictive models to gather impactful insights, which helped business stakeholders in planning out a part of their marketing and business strategy.

## CAREER HIGHLIGHTS

- Successfully delivered a gold mining prediction project, helping the company identify the top 15 potential mining locations with 85% f1-score. Also, featuring the data science impact on gold mining in the SSR Mining's quarterly newsletter.
- Developed an application test pipeline to increase efficiency and reduce 30% of work hours in regression testing per sprint.
- Created an infrastructure dashboard website for monitoring internal data centres to help ease fault detection and highlight potential issues.

## PROFESSIONAL EXPERIENCE

**SSR Mining Inc.,** Vancouver, BC, Canada

### Data Scientist Intern (May 2019 – Feb 2020)

- Worked with a team of 2 geologists to understand the business requirements and collected historical mining datasets from 4 different data sources (size ~30 GB).
- Designed and engineered the test data pipeline on AWS for the mining team, which included data extraction, transformation and loading (ETL), training, evaluation and inference.
- Performed data cleaning, transformation, feature engineering, statistical analysis and data sampling on the data using Pandas and Scikit-Learn.
- Worked on Gradient Boosting Algorithms, Random Forest, SVM and Feedforward Artificial Neural Network to develop the gold prediction models.
- Boosted deep learning neural network's performance by 10% when compared with the gradient boosting algorithms and reached 0.85 F1-score.
- Communicated design decisions and complex predictive analytics results to the non-technical stakeholders effectively through creative data visualization using Tableau and Matplotlib.
- **Technologies Used:** Python, Matplotlib, Tableau, Pandas, Amazon S3, Amazon EC2, Scikit-Learn, NumPy, TensorFlow

**Infosys Ltd., Chandigarh, India**

### Senior Systems Engineer (Data Analyst) (Sep 2016 – Aug 2018)

- Developed a data analytics tool to generate valuable business insights, which helped in product marketing and improving sales figures by 15%.
- Worked on test pipeline automation to reduce manual test and documentation efforts resulting in increased efficiency and a 25% reduction in issues.
- Designed 3+ intuitive UI interfaces to present the real-time findings in the front-end, working in agile environment.
- Worked with docker on cloud (AWS) to enhance load handling and deploying applications in production systems.
- Collaborated with the DB team and worked on data modeling using databases such as MySQL and Oracle.
- **Technologies Used** - Python, Java, Django, MySQL, JavaScript, HTML5, CSS3, Git, SVN

**Infosys Ltd., Bhubaneswar, India**

**Systems Engineer (Feb 2015 – Aug 2016)**
- Developed a Django based highly scalable python website using various emerging technologies to enhance internal server management.
- Attended technical meetings with clients to understand business architecture, building high-quality design documents and robust algorithms.
- Created PL/SQL DML procedures, queries and triggers to encapsulate the business rules in the dashboard application.
- **Technologies Used** - Python, Django, MySQL, JavaScript, HTML5, CSS3, Git, SVN

## PROJECTS

**Simon Fraser University, BC, Canada**
**Predicting Stable Portfolios using Machine Learning and Deep Learning (Jan 2019 – Apr 2019)**
- Scraped, cleaned and parsed 4600 SEC (Security and Exchange Commissions) filings and collected the OHLC data of S&P500 companies for the past 10 years using Quandl API.
- Designed, tested data pipelines and analyzed the effect of SEC filings on stock prices of S&P500 companies using natural language processing library (NLTK).
- Created 90 days windows to map OHLC data to quarters and trained LSTM time-series forecasting model on sentiments of SEC filings and OHLC data.
- Constructed stable portfolios using these quarterly predictions on historical stock prices.

**Simon Fraser University, BC, Canada**
**NYC Taxi Analysis (Oct 2018 – Nov 2018)**
- Analyzed data to provide meaningful business insights on improving NYC taxi business value using data analytics tools (R and Tableau).
- Extracted, transformed and cleaned the data and developed the pipeline on google cloud using PySpark and SQL programming.

## TECHNICAL SKILLS

- **Machine Learning Libraries:** PyTorch, Keras, TensorFlow, Scikit-Learn, Pandas, NumPy, SciPy
- **Languages:** Python, SQL, R, Java, Scala, Bash Scripting, JavaScript, Angular
- **Frameworks:** Django, Flask, Apache Spark, Anaconda, Spyder, Maven, Spring
- **Development and Deployment Tools:** OpenShift, Docker, Git, Jupyter Notebook, RStudio
- **Databases:** SQL Server, Oracle 10g, PostgreSQL, Hive, NoSQL Cassandra
- **Hadoop Ecosystems:** HDFS, Kafka, Map Reduce, Hadoop Streaming
- **Spark Ecosystems:** RDD, DataFrame, Spark-SQL, Spark MLib
- **Cloud Computing Services:** AWS (EC2, S3), GCP, Azure
- **Visualization Tools and Libraries:** Tableau, D3.js, Matplotlib, Plotly, Seaborn

## EDUCATION

**Simon Fraser University, Burnaby, Canada (Sep 2018 – Apr 2020)**
**Master's in Computer Science (Big Data), GPA of 3.85 out of 4.33**
- **Key Subjects:** Machine Learning, Statistics, Programming in Big Data, Visualization, Algorithms for Big Data

**Uttarakhand Technical University, Dehradun, India (Aug 2011 – Jun 2014)**
**Bachelor's in Engineering, Computer Science, GPA of 3.3 out of 4**
- **Key Subjects:** Applied Mathematics, Data Mining, Data Structures, Database Management Systems, Software Engineering