

Rahul Singh

+91 7065921135 | rahulsingh50.github.io | connect2rsingh8@gmail.com | rahulsingh50 | @LinkedIn

RESEARCH INTERESTS

Data Science, Big Data, Machine Learning, Statistical Modeling, NLP, Deep Learning, Data Analysis, Business Intelligence

ROLE AND RESPONSIBILITY

- Having 5+ years 'experience (3+ year's relevant and 2+ years self-work due to covid) **Data Scientist and BI** focused on deriving useful business insights from large datasets and helping organizations generate business value and revenue using an agile and data-driven approach.
- Good exposure in develops Business Analytics, Predictive Modelling, Statistical Modelling, Data Mining, NLP, Machine Learning and Deep Learning solutions for Text, Image and Video data.
- Hands on experience model development cycle (supervised and unsupervised) for data at the scaling level. Solid understanding of explanatory data analysis (EDA), data pipe-line (ETL), information extraction, data cleaning, Model building, Optimization, visualization and deployment using the agile method in Investment, Financial Services, Marketing and IOT domain.
- Translated complex analytical requirements into technical design including data models, ETLs and Dashboards Reports. Experience deploying dashboards and self-service analytics solutions on both relational and non-relational databases.
- Strong Skills in Data management including data governance, data quality, data integration, processing and provisioning with Familiarity with data security and privacy regulations such as GDPR and CCPA.

WORK EXPERIENCE

Self-Work (Due to Covid)

July 2020 - Present

Data Scientist

Noida, India

- Working as Quantitative Research Analyst on Financial and Investment data using Machine learning, Deep Learning, Statistical modelling, financial modelling, Forecasting, Investment Analysis, Derivative Market (Options, Futures), Algo Trading, Visualization etc.
- Good understanding of Blockchain, Cryptocurrencies, Smart contracts.
- Responsible for the end-to-end development cycle for Equity Investment return of clients, Risk management, Portfolio optimization. Some notable work is given below.
 - Portfolio diversification using Correlation Matrix
 - Ratio, Regression, Performance and Risk Analysis in Investment
 - Portfolio Optimization using Markowitz Model
 - Time Series Modelling using EST, ARIMA, EWMA model etc.
 - Stock Market Data Analysis Using ChatGPT & Python.
 - Model Volatility with GARCH Class Models.
 - Monte Carlo Simulation.
 - Stock Sentiment analysis.
 - Crawling data different sites using API.
 - Implementing Trading Algorithms using Technical Analysis Python and Dashboarding.
 - Understand fundamental and Technical Analysis and Trading Automation
 - Generating various KPI and Reports.

Environment: Python, Pandas, NumPy, Tableau, Power BI, R, Excel, Stats model, Prophet, Scikit Learn, Tensor flow, Keras, PySpark, DataBricks, Azure Spark SQL PyFolio, ChatGPT, MLOps, MySQL etc.

Worked on for **KocharTech IOT Solution** for Video, IOT and Telecom data.

Real Time Video Analytics Platform: Responsible for the end-to-end development cycle of for Object Detection and Classification problems for Real Time Video Analytics Platform KocharTech IOT division

- **BPO Use Cases:** Mobile Detection, ID card wearing or not.
- **Retail Use Cases:** People counting various shops (No of clients coming and out particular day, Desk cleaning, Guard detection. All activity done on edge and server real time basis and reporting higher authorities using automatic mailing system).
Implemented, trained, and deployed the model as TensorFlow Lite model supported using TFServing for on-device inference and the Object Detection model as a web API on AWS for an Android app.. Trained a custom YOLO model to detect ~150 objects with a mAP of 67%.
Environment: Python, TensorFlow, Keras, OpenCV, Pandas Custom YOLO, Pytorch, Rest API, Flask ,MLOps etc.
- **Multilingual Chatbot:** Worked on Multilingual Chat bot for 2nd largest gulf telecom client for various services using Text classification Models, Spacy and NLTK. It has custom intent, entities and utterances. It also integrates for website, WhatsApp and Facebook messenger can use social media for lead generation and boost sales.
Environment: Python, PHP, NLTK, Spacy, Tensor flow, Keras, RASA, Docker, Deployment, Amazon Web services (AWS), MLOps etc.
- **Router Fault Detection:** Find out the fault various label using the custom object detection model using the Selfie link app by developed by KocharTech Android Team.
- **IOT Data Analysis:** Worked on telecome data 2nd largest gulf telecom client for various Analysis the Hadoop server storage IOT and telecom. Make Hadoop pipeline using pyspark, MLLib, PysparkSQL, and generating reports
- **Telecom Data Analysis:** Develop end to end life cycle for analytics ETL (PostgreSQL to Big Query), Big Query ML, Dashboarding and Reporting.

Working with team of Marketing Analytics on various project some are given below.

- **Social Media Analyzer:** Developed a program in Python to gather, analyze and find trends in the Constalytics unstructured user's comments on Twitter & Facebook, gave 10 tips on how Constalytics can use Social Media for lead generation and boost sales.
Environment: Python, Pandas, Numpy, Skit-learn, NLTK, Google Studio, Google Analytics & Google Trends, SQL
- **Digital Ad Analytics:**
 - Impressions and reach
 - Click-through-rate (CTR)
 - Cost-per-click (CPC)
 - Cost-per-lead & Key takeaways
- **Customer Analytics:**
 - Customer Life time value
 - Customer acquisition cost**Environment:** Excel, Python, Web scraping, Google Trends, SQL etc

Working as computer teacher teaches computer subject.

EDUCATION

ABES Engineering College
Masters of Computer Application (MCA)

Ghaziabad, India
Aug. 2006 – June 2009

Chhatrapati Shahu Ji Maharaj University
Bachelors of Computer Application (B.Sc. Computer)

Kanpur, UP, India
Aug. 2002 – June 2005

RELEVANT COURSEWORK

Certified MOOCs

- [Introduction to TensorFlow for Artificial Intelligence Machine Learning and Deep Learning](#)-Coursera
- [The Data Science Toolbox](#) – Coursera
- [Big Data](#) - Coursera

Other Coursework: Object Oriented Programming, Data Analysis using R, Unix Programming, Data Structure and Algorithms, Statistics, Cloud Computing, MLOps.

TECHNICAL SKILLS

Languages: Python, R, JavaScript's, SQL

Frameworks/Databases: TensorFlow, PyTorch, Pandas, OpenCV, NumPy, Scikit-learn, NodeJS, Flask, FastAPI
MongoDB, MySQL, SQL Server, PostgreSQL etc.

Developer Tools: Git, Docker, Jira, PyTest etc.

BI Tool: Excel, Tableau, PowerBI, Looker, Snowflake, Synapse

Big Data: Hadoop, Big Query, Big Query ML, Pyspark, Spark SQL, Mlib

OS: Unix/Linux, Windows

Cloud: Google Cloud, Azure Cloud, Databricks.

** Links in blue are hyperlinks