JITESH RAWAT

DATA ENGINEER

CONTACT

+91 9324979351

jitezzrawat99@gmail.com

Linkedin

Github

SKILLS

- **Python**
- **SQL**
- **Data Science**
- **ETLs**
- **APIs**
- **Spark**
- **Google Cloud Platform**
- **AWS (Redshift)**
- **Big Query**
- **Machine Learning**

EDUCATION

Master of Science (M.Sc)
Data Science and Computing

Sri Sathya Sai Institute Of Higher Learning

2020 - 2022
GPA: 7.8

Bachelor of Computer Applications (BCA)

Sri Sathya Sai Institute Of Higher Learning

2017-2020

GPA: 6.9

COURSE

Data Engineering on Google Cloud Agilitics - Virtual Classroom

This course covered hands-on practical designing and building data processing systems on Google Cloud. From setting up Cloud SQL to ingesting data in BigQuery for data analytics. The course covered various solutions of Google Cloud like Data Proc, Pub/Sub, Data Flow, BigTable, AutoML.

PAPER PUBLICATION

Published a paper titled “Comparative Analysis of Stacked Models and other Machine Learning Algorithms in Estimating Claim Cost for Automobile Insurance” in the International Journal of Emerging Technologies and Innovative Research.

WORK EXPERIENCE

Data Engineer

Atidiv

October 2022 - current

- Building data pipelines in ETL or ELT to fetch the data from various APIs and store them into data warehouses either in batch or streams for analytical purpose.
- Structuring Tap - Target oriented data pipelines (ELT) used to fetch data from sources like API's, data warehouse generally scheduled on airflow deployed on a cloud VM.
- Built customized pipelines for Reverse ETL tasks.
- Handled data pipeline issues and also worked rebuilding pipelines into robust and standardized manner.

Data science Intern

Tech Actuarial

Jan 2022 - Apr 2022

- Industry project under Tech Actuarial research department. The project work included working in the Pyspark framework for custom implementation of stacked models for estimation of claim cost for auto insurance data.

PROJECTS

Estimating claims cost using Stacked models and Comparative analysis with individual Machine learning techniques.

- Research based and Actuarial domain specific project where regression analysis and comparison based on individual machine learning models and stacked models in Pyspark for estimation of auto insurance claims.
- Technology used : Python, Big data, Pyspark. Dataset : Crash Analysis System (New Zealand)

Pose Estimation Controller

Link : https://github.com/JiteshRawat/Pose_Estimation_Controller

A Convolutional neural network is trained on data of 7 different poses to predict them and control game moves in real time.

Sustainable Development Goals Classifier

Link : https://github.com/JiteshRawat/SDG-Classifier

NLP based multiclass classification model to classify UN Sustainable Development Goals (SDGs).

Hostel Store Recommendation system in R

Link : https://rawatjitzz.shinyapps.io/Store_Items_Recommender

Recommendation system and market basket analysis used in a web interface using R shiny, R and Mongodb for my Hostel Store's Data.