JITESH RAWAT

DATA ENGINEER

CONTACT

+91 9324979351

jitezzrawat99@gmail.com

Linkedin

Portfolio

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

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

- Orchestrated ETL and ELT pipelines, seamlessly extracting diverse API data for storage in data warehouses, optimizing for both batch and streaming analytics.
- Architected Tap-Target data pipelines with a focus on ELT methodology, leveraging Airflow on a cloud VM for efficient scheduling.
- Spearheaded the development of tailored Reverse ETL pipelines.
- Proactively addressed data pipeline challenges, demonstrating a knack for troubleshooting, and implemented comprehensive rebuilds, ensuring robust and standardized workflows for enhanced efficiency and reliability.

Data science Intern

Tech Actuarial
Jan 2022 - Apr 2022

- Led a transformative initiative in Tech Actuarial Research, utilizing advanced Pyspark proficiency to implement stacked models, markedly improving precision in estimating claim costs for extensive auto insurance datasets, showcasing strong analytical and technical capabilities in driving impactful data science solutions.

PROJECTS

Amazon Product Review Analysis

https://amazon-reviews-dashboard-11.streamlit.app/

Search for any keyword and the data for that is scraped from Amazon. Designed and executed a data engineering project focused on the systematic extraction and analysis of product reviews and ratings from the Amazon e-commerce platform, culminating in the creation of an interactive sample product review dashboard.

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.