

Reventh Sharma

+91 799-065-0706
resharma@ucsd.edu

in reventh-sharma
Reventh-Sharma
reventh-sharma.github.io

Education

Master of Science in Data Science

University of California, San Diego
Incoming Sep-2023

Bachelor of Engineering in Chemical Engineering Minor in Data Science

Birla Institute of Technology & Science, Pilani
Aug 2017 – July 2021

CGPA: 8.74

Relevant Coursework: Intro to Computer Programming, Foundation of Data Science, Probability & Statistics, Applied Statistical Methods, Optimization, Machine Learning, Neural Networks & Fuzzy Logic, Data Mining

MOOC Specializations: [Data Structures & Algorithms](#), [Object Oriented Programming](#)

Technical skills

Programming Languages

C/C+ Python Scala Java MySQL JavaScript

Python Frameworks & ML Libraries

Airflow Tensorflow PyTorch OpenCV Transformer
Detectron2 PySpark SparkML Gensim NLTK SpaCy
FastAPI

Tools & Technologies

Docker Azure Hadoop HDFS Spark

Work Experience

Couture.ai

Data Science Engineer

Jul 2021 – Jul 2023
Bangalore, India

- **Image Search :** Developed a real-time Image Search ML API for Ajio, an eCommerce Fashion Platform, facilitating the segmentation of fashion items from user-uploaded images and retrieval of similar fashion items from the store's product catalog. The API successfully drove a transition of 13,000 users from traditional text search to image-based search within one month of launch.

CV Segmentation Vision Transformer Faiss

- **Similar Style Recommendation System:** Designed top-N similar items retrieval system for store's catalog products by fine-tuning text and vision transformer models on the inHouse fashion dataset. The implementation of this transformer-based system led to a notable 5.3% increase in the click-through rate of users.

NLP CV Transformers Faiss

- **Personalized Recommendation System:** Developed a user-cluster level relevance ranking algorithm based on collaborative filtering and XgBoost Learning to Rank model. In AB tests, the developed model demonstrated 7-12% increase in conversion percentage (Number of buys/ Number of clicks), validating the enhanced recommendations.

Matrix Factorization XgBoost Unsupervised Clustering

Couture.ai

Data Science Intern

Jan 2021 – Jun 2021
Bangalore, India

- **Recommendation Systems:** Built automated data processing pipelines to perform real-time ETL operation on user interactions data and catalog data. Designed metrics for offline evaluation of relevance ranking algorithms.

Personalized Recommendation Data Engineering Big Data

- **Process Optimization:** Developed a multivariate constrained optimization technique using Adam Optimizer to maximize acetylene conversion in a GCU Plant.

Constrained Optimization

Big Data

Research Experience

Image Captioning System Using ResNet-18 and Transformer: Term Paper

Jan 2020 – May 2020

Department of Computer Science, BITS-Pilani

- Developed an image captioning system utilizing ResNet18 as the encoder and incorporating a Transformer model as the decoder.
- The model outperformed ResNet18+LSTM architecture within literature, achieving a BLEU-4 score of 25.3 on the MS-COCO dataset.

ResNet18

Transformers

Image Captioning

Development of Dynamic Gas Concentration Predictor from a Multi-Sensor Array

Aug 2019 – Dec 2019

Department of Chemical Engineering, BITS-Pilani

- Built an LSTM-based model to predict concentration of gases in mixtures based on TGS sensor array output within a fluctuating environment.
- Created an anomaly detection algorithm that effectively identified and trimmed anomalous sensor values by applying four-sigma rule to chi-squared distributed sensor data within moving windows.
- The end-to-end prediction pipeline demonstrated a significant 10% improvement in prediction accuracy compared to RNN-based reservoir computing algorithms and ML-based time series regressors in literature.

Time Series Prediction

Anomaly Detection

RNN

Wheat Type Classification Using Active Learning

Mar 2020 – Jun 2020

Department of Computer Science, BITS-Pilani

- Designed an active learning model utilizing random forest architecture to predict the type of wheat by analyzing its dimensional and visual characteristics.
- Used pool based uncertainty method to train model on a selected subset, comprising 20% of total train data volume. This method increased the robustness of the training, resulting in final test accuracy of 93%.

Active Learning

Ensemble Learning Method

Statistical Analysis and Forecasting of Solar Energy in the State of Rajasthan, India

Aug 2020 – Dec 2020

Department of Mathematics, BITS-Pilani

- Performed statistical tests to validate seasonality within the solar energy data.
- Predicted next week and month's expected radiation intensity using ARIMA and SARIMA-based statistical models and LSTM-based deep learning model.

Time Series Forecasting

ARIMA

SARIMA

Statistical Tests

LSTM

A Review on Bottom-Hole Pressure Estimation in Oil Wells using Computational Intelligence Techniques: Term Paper

Aug 2020 – Dec 2020

Department of Chemical Engineering, BITS-Pilani

- Authored a comprehensive review article outlining recommended approaches for data collection, model selection, and optimization methods to achieve low-cost and accurate prediction of bottom-hole pressure in oil wells.

Data Mining

Feature Selection

Optimization

Predictive Analytics