



EDUCATION

* till 2nd sem

Qualification	Institute / Board	CGPA / %	Year
MTech (Industrial and Management Engineering)	IIT Kanpur	9.20 *	2021 - Present
BTech (Mechanical Engineering)	Birla Vishvakarma Mahavidyalaya, Anand, Gujarat	7.95	2016 - 2020
Higher Secondary Education	Gujarat Secondary & Higher S. Education Board	89.23 %	2016
Secondary Education	Gujarat Secondary & Higher S. Education Board	87.17 %	2014

PROFESSIONAL EXPERIENCE

Data Science Intern | Ecom Express Limited, Gurgaon

May'22-Jul'22

Topic :- Building a NER based model on " Sub-Locality " address parser

Aim :- To Create a Sublocality dataset from raw addresses and train a custom NER based model for Sublocality tags

- **Examined** address entities and **Normalize** the addresses using regex patterns
- Created different **rules** and various **regex patterns** (one, two and six) for three sub-locality tags : "**ward**", "**phase**" and "**sector**"
- **Prepared** two training datasets from raw datasets having 0.8 and 2.1 million addresses
- Built two **Named Entity Recognition (NER)** models on the training datasets of 32,999 and 1,08,000 instances in **Spacy**
- **Validated** our trained models on a different test dataset with 0.4 million addresses
- Achieved best **Matthews correlation (MCC) : 0.906, Precision : 83.51 %, Recall : 99.26 %** and **F1 score : 90.70 %** for second model

KEY ACADEMIC PROJECTS

House Price Prediction for the city of Ames, USA | Statistical Modelling for Business Analytics | Regression

Aug'21-Sep'21

- The Dataset contains **80** features including the Price (Dependent variable), along with 1460 observations
- Performed **Outlier detection, Feature Engineering & EDA**, used **Correlation** and **VIF** to check **multicollinearity** and to drop features
- Used **RFE** for initial feature selections, identified **heteroskedasticity** using the residual plot and handled it with **log transformation**

Result : Observed **adjusted R² of 0.818** with the final OLS model of only **9** input features

Next day Rain Prediction | Applied Machine Learning | Classification

Jan'22-Feb'22

- The Dataset contains 145460 instances and 22 features of 10 years of daily weather observations across Australia
- Performed **Random imputation & Regression** for **missing value** treatment, **Outlier detection** using **IQR, Feature Engineering, EDA**
- Used **SMOTE** to handle **class imbalance** problem, applied 3 different models : Logistic, Adaboost, and Random Forest

Result : **Random Forest Classifier** with hyperparameter tuning using **RandomizedSearchCV** and **GridSearchCV** performed best

RFM Analysis and Customer Segmentation | Applied Machine Learning | Clustering

Feb'22-Mar'22

- Dataset include 5,41,909 instances with 7 features - InvoiceNo, StockCode, Description, Quantity, InvoiceDate, UnitPrice, CustomerID
- Analyzed **RFM (Recency, Frequency, Monetary)** characteristics using Feature Engineering, used **winsorization** for **outlier treatment**
- Plotted **elbow graph** to identify the optimal number of clusters for the **K-Means** clustering algorithm
- Visualized the **3 clusters** using a 3D plot based on the **amount, number of transactions** and **most recent transaction**

Result : Observed **silhouette score of 0.47** for the three clusters

Forecasting Monthly Champagne Sales | Self Project | Time Series Analysis

Oct'21-Nov'21

- The dataset contains historical sales data for 9 years (1964 to 1972) of Perrin Freres company
- Performed stationarity test using **ADF (Augmented Dickey-Fuller)** and **Differencing** to obtain stationarity
- Plotted **PACF (Partial Autocorrelation function)** and **ACF (Autocorrelation function)** to find optimal parameters p, d, q
- Applied **AR, ARIMA, and SARIMA** models and used **MAPE** as an evaluation metric

Result : Predicted monthly sales for next 2 years based on best tuned model **SARIMA (1,1,1) x (1,1,1,12)** with **MAPE of 7.4 %**

Facial Expression Recognition | Applied Machine Learning | CNN

Mar'22-Apr'22

- Data consists of 48×48 pixel grayscale images of 7 different types of emotions (i.e. Angry, Disgust, Fear, Happy, Neutral, Sad, Surprise)
- Created more images using **Data augmentation** and built a first CNN model with **4 convolution layers** and **3 FC layers**
- Built a second model using **VGG-16**, made a few changes, used **ReduceLRonPlateau & EarlyStopping** callbacks using Keras Library

Result : Obtained Train accuracy of **85.96 %** and Test accuracy of **86.15 %** for second model

COURSEWORK AND SKILLS

* in progress

<i>Academic courses</i>	Statistical Modelling for Business Analytics Applied Machine Learning Data Mining and Knowledge Discovery* Causal Inference for Business Analytics* Probability and Statistics Market Research Introduction to computing
<i>Online Courses</i>	Deep Learning - Simplilearn MySQL for Data Analytics and Business Intelligence - Udemy

<i>Technical Skills</i>	Python SQL PowerBI SPSS MS Excel Data Science and ML : Numpy, Pandas, Scikit-learn, Statsmodel, Regex, Spacy, NLTK, Matplotlib, Seaborn, TensorFlow
-------------------------	--

POSITION OF RESPONSIBILITY

Senator [PG] Y21 - Student's Gymkhana, IIT Kanpur July'22-Present

- Act as the voice of **500+** M.Tech students (2021-23), at the Student's Gymkhana
- Creating awareness among PG students to increase their participation in workshops, events conducted by various clubs

Class Representative - M.Tech IME, IITK

Jan'22-Present

- Representing the batch in both official and informal contexts, keeping track of the activities of each team and its members
- Organizing and coordinating team meetings and webinars

ACR Team Coordinator - M.Tech IME, IITK

Aug'21-Dec'21

- Organized **12** webinars and Alumni meet, prepared MOM

ACHIEVEMENTS & EXTRACURRICULAR ACTIVITIES

- Secured **Educational scholarship** from Elecon Group of Companies for **3 years** of my B.Tech education
- Secured **296th Rank (99.75 percentile)** among 1.2 Lakh+ students in the paper of Mechanical Engineering in **GATE 2021**
- Devoted a Period of **240 hours** of service under regular activities in 4 years as **NSS Volunteer**
 - Organized the **NSS Camp** and led a team of **25** students for a week by coordinating events in the adopted village "Mogri"
- Voluntarily guiding **800+** students through Telegram and WhatsApp in matters pertaining to the GATE Mechanical exam