

**ACADEMIC DETAILS**

\*upto 2<sup>nd</sup> sem

YEAR	DEGREE	INSTITUTE	CPI/%
2021-2023	M.Tech. (Industrial & Management Engineering)	Indian Institute of Technology, Kanpur	8.15*
2017-21	B.Tech. (Textile Technology)	U.P. Textile Technology Institute, Kanpur	8.31
2017	Class XII (CBSE)	Delhi Public School Kalyanpur,kanpur	80.60%
2014	Class X (CBSE)	Delhi Public School Kalyanpur,kanpur	8.8 CGPA

**INTERNSHIP**

**Axtria-Ingenious Insight**

(May'22 – July'22)

**Data Quality Visualization on Power BI & Microstrategy**

- Analyzed Business Requirement Document of Novartis Brand to understand the objective of project and to drive the sales performance and execution.
- Dataset Included Data quality data of various Metrics, Geography, Field Forces, brands & their competitors having weekly/monthly sales, % variation b/w sales, Hierarchy wise sales etc.
- Transformed the excel sheets by pivoting down the tables so as to visualize them accordingly on Power BI & Microstrategy.
- Created KPI's and different Visualizations in Power BI desktop and similarly dossiers in Microstrategy to get insights, which would help Novartis & Axtria in making data driven decisions

**ACADEMIC PROJECTS**

**Sentiment Analysis on IMDB Movie Review** (Applied Machine Learning)

(Mar'22 – Apr'22)

- The IMDB dataset consists of 50,000 Movie reviews that have been pre-labeled with "good" and "negative" sentiment class labels.
- Classified sentiment based on review text, performed data cleaning and pre-processing by Stop-word removal, tokenization, **Stemming**.
- Executed Feature Extraction Techniques – **Bag of Words, TF-IDF, and Word2Vec**.
- Applied Models –**Multinomial Naïve Bayes, Logistic Regression, and Support Vector Machine**.
- Used **Accuracy, Precision, Recall and F1-Score** as evaluation metrics for comparing models.

**Analysis of the Factors Affecting Sales Price of Houses in King County, USA** (Statistical Modelling for Business Analytics)

(Aug'21 – Sept'21)

- Carried out **Exploratory Data Analysis & Data Visualization**.
- Start building model using **uni-variate Linear Regression** model, and then using **multi-variate Linear Regression** model.
- Calculated measure of fit, **correlation matrix**, performed Breusch-Pagan test for **heteroskedasticity**, used **VIF** (Variance Inflation factor) to check **multi-collinearity** and looked for omitted variable bias by performing backward elimination and eliminating insignificant variables.

**Telecom Customer Churn Prediction** (Statistical Modelling for Business Analytics)

(Sept'21 – Oct'21)

- Dataset contained 7043 rows (customers) and 21 features such as "customer account", "customer demographic info", "paperless etc.
- Performed **EDA**, applied **SMOTE** to balance the data and **RFE (Recursive Feature Elimination)** to select the 15 significant features.
- **Logit** and **Probit** models were used for classifying the churn class; features were dropped based on p-value and VIF.
- Logit showed better results as the accuracy was about 79%, precision of 73.8% and a recall of 62.4%, AUC of ROC curve was 0.83.

**Analysis of Hurdles in Evolution of EV's in India** (Marketing Research)

(Feb'22 – Apr'22)

- Conducted Survey on EV's in India to to know peoples perspectives for EV's and to know their requirement or expectations from electric vehicles.
- Data collected using online survey, focus groups and conducting 12+ personal Interview with people from different states.
- Designed cross-sectional case study dynamic survey form using Scaling techniques, pretesting to control internal & external validity.
- Conducted exploratory descriptive research & analysis data using t -test in SPSS with 95% confidence level to test the hypothesis.

**SELF PROJECTS**

**Air-Passengers-Forecasting-using-Time-series**

- Predicted Air Passengers travel for next 10 years analyzing previous month wise data from 1949 to 1960 using time series techniques
- Decomposed Time series into its component to analyze Trend, Seasonality, and noise
- Checked Stationarity using Rolling Statistics test & ADF-test (Augmented Dickey-Fuller) and stationarised time series by Differencing(d)
- Plotted PACF (Partial Autocorrelation function) and ACF (Autocorrelation function) to find optimal parameters p, d, q
- AR, MA, ARIMA, time series models applied and used RSS as evaluation metric

**Image Classification using CNN**

- Loaded cifar10 Dataset consisting of 60000 32 X 32 color images in 10 classes from tensor flow library having various objects like ship, frog, cat, horse etc.
- Splitted the data in training and testing with 50000 & 10000 samples respectively and the Normalized our data.
- Builed ANN with 5 epochs having relu and softmax as an activation function and sparse categorical cross entropy as loss function giving accuracy at around 49%.
- Modeled CNN having 5 epochs and using convolution & max pooling layer with relu and softmax as an activation function giving accuracy around 70% and at the end of 15 epochs as 83% which is a significant improvement over ANN.

**Credit Risk and Fraud detection**

- Dataset is highly imbalance contain 284,807 transactions with 35 features out of 492 are Fraud's
- Perform Explanatory Data Analysis (EDA), Data Preprocessing and Data visualization, Class imbalanced data handled by SMOTE and Used Accuracy Precision, Recall and F1-Score as metrics to compare model.
- Applied Logistic Regression, Random Forest and Decision tree with sampling Best Model is Random Forest.

## COURSEWORK AND SKILLS

<b>Academic Courses</b>	<b>Applied Machine Learning</b>   <b>Data Mining</b> and Knowledge Discovery   Probability and Statistics   Introduction to Computing (Java)   Operation Research for Management   Statistical Modelling For Business Analytics   Marketing Research   Business Management using cloud
<b>Online Courses</b>	Python and Data Handling Libraries Fully Diploma 2021   Power BI Zero to Hero   Statistics for Data Science and Business Analysis   Excel Skills for Business: essentials
<b>Technical Skills</b>	<b>Machine Learning</b>   <b>Statistical Analysis</b>   <b>Natural Language Processing</b>   <b>Python</b> (Numpy, Pandas, Scikit-Learn, Matplotlib, Seaborn)   <b>Java</b>   <b>HTML</b>   <b>Power BI</b>   <b>Microstrategy</b>   <b>MS Excel</b>   <b>SQL</b>
<b>Soft Skills</b>	Critical Thinking, Problem Solving, Teamwork, Effective Communication, Adaptability

## POSITION OF RESPONSIBILITY , ACHIEVEMENT AND EXTRA CURRICULAR ACTIVITIES

- Organized 25+ webinars & Alumni meet and managing the logistics of the webinar as a Senior ACR coordinator at IME, MTech IITK
- Teaching Assistant for the Coarse MBA 666 (Project Management).
- Event Head in the Annual Fest(PANORAMA) 2015 in my school.
- Secured AIR 14 in GATE(TF) -2021
- Awarded 5-star Gold Badge in SQL under specialized skills at [Hacker Rank](#)
- Appreciation certificate from SBI LIFE INSURANCE for awaring the people from CORONA-VIRUS
- Organized Young Innovators Summit 2018 held at UPTTI Kanpur
- Participated in 2 days Robotics & IOT workshop conducted by Technophilia solution at IIT BHU powered by MICROSOFT.