

Year	Qualification	Educational Institution	Percentage
2019-21	M.Tech (Industrial and Management Engineering)	Indian Institute Of Technology, Kanpur	7.73* (CPI)
2013-17	B.Tech (Mechanical Engineering)	Rajkiya Engineering College Azamgarh	80.06%
2013	Class XII	M.S.S. inter college ,Meerut U.P.	90.4%
2011	Class X	M.S.S. inter college ,Meerut U.P.	75.5%

*up to 2nd semester**INTERNSHIP**

Data Science Intern at Harvesting India private Ltd. Chandigarh.

(April'20 -June'20)

- **Title:** Classify the crops cultivated in any specific village, district or state in various categories, with use of **machine learning and deep learning models**.
- **Data exploration & manipulation:** **Sentinel-2 satellite data** was used as input, which consisted **multi-spectral** (5 bands) raster file (image), with spatial resolution of 10m , spanning over 6 months covering rabi season, October 2019 to March 2020 of Bathinda district of Punjab state .
- **Supervised Classification:** The raster files, 6 of them had resolution of 8298 x 8963, of which for only 131 pixels had ground reading . Trained and tuned supervised models such as, **Logistic Regression, Multi-layer Neural Network, Decision Tree , Random Forest, Support Vector Machine, K- nearest neighbours** with the labelled ground data and then used model for classifying the remaining unlabelled pixels resulting to 100% this was due to less training data (labelled ground data).
- **Unsupervised Classification:** Unsupervised algorithm such as **K- means Clustering with NDVI & GNDVI** technique was implemented on the same data (Bathinda) . **Precision, Recall, F1 score and Accuracy were evaluated as 0.818, 0.911, 0.862, 0.821 respectively**. K-means clustering algorithm same as earlier was applied on another district DHAR of Madhya Pradesh state, with **multi crop classification**.

ACADEMIC PROJECTS

Data Mining & Knowledge Discovery	Predict a movie's worldwide box office revenue: <ul style="list-style-type: none"> • Objective was to predict the movie's worldwide box office revenue studying the train data which contained 4398 movies and test data with 3000 movies, with 22 independent variables and 1 (revenue) dependent variable. • Steps include Data pre-processing , Statistical analysis, Data visualization , Model building and predicting test data from trained model using Python. Fitted the training data in model built using regressors like Linear Regression, Random Forest , SVM, Decision Tree , KNN , LGBM and XGBoost . LGBM and XGBoost performed well with RMSE score of 1.89124 and 1.95831 respectively. • Packages used – pandas , seaborn, matplotlib, Wordcloud , Nltk , scikit-learn , eli5 etc 	Sep'2019-Nov'2019
Statistical Modelling for Business Analytics	Graduate Admission Prediction using Statistical Regression Models <ul style="list-style-type: none"> • The data consisted of 500 observations of 9 variables, Determined correlation matrix and checked for Multicollinearity • Breusch-Pagan test showed heteroskedasticity ,hence "heteroskedastic robust errors" were used for regression on significant variables. • Backward elimination technique is used for variable selection . The models were compared and Adjusted R2 of the best model was 0.8136. Predicting Income class using Logistic Regression using Adult data set. <ul style="list-style-type: none"> • Study the effects of various factors to classify whether a person's income is <50K or >=50K (target variable). • Income class was unbalanced with about 75% of data belonging to class <50K. (46,033 observations) • Logit and Probit models were used for classifying the income class. The performance was similar to an accuracy of about 82.3% precision of 67.2% and a recall of 52.8%, AUC of ROC curve was 0.88. 	Jan'2020-Feb'2020 Mar'2020-Apr'2020
Marketing Research	Analyzing loyalty of the viewers towards a TV channel . <ul style="list-style-type: none"> • Conducted the Online Survey and analyzed data in SPSS.Data is collected using online survey , focus group and personal interviews. • Exploratory and Descriptive Research in SPSS using primary data and various statistical test(One Sample T test, Independent t-test, One Way Anova,Chi-Square test) are conducted to test our hypothesis. • This study suggested content and variety plays a vital role while selecting a particular channel as compared to brand value and price. 	Mar'2020-Apr'2020
Applied Machine Learning	Predicting the Impact of COVID19 Lockdown on Air Pollution in New Delhi. <ul style="list-style-type: none"> • From given years (2018 to till may 2020) of pollution data of Delhi , forecasting PM 2.5 level for next 3 months using various Time Series Techniques(ARIMA , ARMA and SARIMAX models) in Python . • Non Stationary behavior of time series was tested by ADF-test (Augmented Dickey-Fuller) and seasonality can also be seen from ACF and PACF plots. • Used Support Vector Regression technique to predict PM 2.5 level based on the other pollutants in atmosphere and visualize the trends. 	Mar'2020-Apr'2020

COURSE WORK AND SKILLS

Relevant Courses	Data Mining and Knowledge Discovery Probability & Statistics Statistical Modelling for Business Analytics Applied Machine Learning Operations Research Marketing Research E- Supply Chain Management Introduction to Computing (JAVA) Accounting & Finance
Technical Skills	Python (numpy, pandas, matplotlib, seaborn , cufflinks ,scikit learn ,nltk , keras) R (ggplot2, plotly, data , table, dplyr, rattle, rmarkdown, knitr, tidyverse,) Machine Learning Deep Learning Java MS Office (Excel, Word, PowerPoint) SQL

*ongoing

POSITION OF RESPONSIBILITY

- PG Core Team, Counselling Service, IIT Kanpur : Organized three day long institute level **Orientation program** for 1500+ students for this Academic Year.
- Senior Executive at Raktarpan, IIT Kanpur : Organized various blood donation as well as awareness camps in IITK campus.
- Volunteer in winter Orientation in which around 300 students took admission , organized by Counselling Service, IIT Kanpur.

ONLINE LEARNING & CERTIFICATIONS

- Python for Data Science and Machine Learning Boot camp by Jose Portilla (Udemy) | Neural Networks and Deep Learning by Andrew NG (Coursera).

AWARDS AND ACHIEVEMENTS

- Volunteer at SHIKSHA SOPAN , non-governmental organization (NGO) ,that imparts value- based education to underprivileged students.
- Secured 2nd position in open article writing competition in IIT Kanpur.
- Secured 2nd position in TECHSRIJAN 2013,the national level Techno management festival of MMMUT Gorakhpur.