Projects Completed in Data Science

Sr .No	Project Title /Domain	Technology Stack Algorithms Used
1	**Identifying Risk factors influencing Diabetes Type- Classification Model** **Domain : Health Care**	**Python** : Machine Learning **Python Libraries** : Numpy, Scikit-learn, pandas, matplotlib, keras, tensor flow **ML Algorithms Used** : In this project, Unsupervised learning namely K Means clustering and Gradient Boosting Regression algorithm are used for obtaining high accuracy percentage. Also Supervised learning algorithms namely Neural Network, LR, LDA, KNN, CART, NB, and SVM,is used to effectively verify Model accuracy of the models.
2	**Data Exploration and Forecasting on Patient Count – Regression Model** **Domain : Health Care**	**Python** : Machine Learning **Python Libraries** : Numpy, Scikit-learn, pandas, matplotlib, keras, tensor flow **ML Algorithms Used**: Unsupervised Learning algorithms namely Arima() Model, Auto Regressor() Model, Stationarity Analysis. Steps Used: Descriptive Analysis and Predictive Analysis and Time Series Analysis for all steps of variable description, exploratory data analysis, data processing and preprocessing, feature selection, model building, and, evaluation of models are done using Python and R.
3	**Infant Mortality Rate Analysis – Regression Model** **Domain : Health Care**	**Python** : Machine Learning **Python Libraries** : Numpy, Scikit-learn, pandas, matplotlib, keras, tensor flow **ML Algorithms Used**: Unsupervised Learning algorithms namely Arima() Model, Auto Regressor() Model, Stationarity Analysis. Steps Used: Descriptive Analysis and Predictive Analysis and Time Series Analysis for all steps of variable description, exploratory data analysis, data processing and

		preprocessing, feature selection, model building, and, evaluation of models are done using Python and R.
4	**Mortality Risk Analysis – Feature selection Analysis** **Domain :Health Care**	**Python** : Machine Learning **Python Libraries** : Numpy, Scikit-learn, pandas, matplotlib, keras, tensor flow **ML Algorithms Used :F Regressor, XG Boost Random Forest Tree Based Models**
5	**Survival Analysis of Unemployment Rate-Classification Model** **Domain : General**	**Python** : Machine Learning **Python Libraries** : Numpy, Scikit-learn, pandas, matplotlib, keras, tensor flow **ML Algorithms Used** :Kaplan Meier Fitter
6	**Loan Eligibility Prediction – Classification Model** **Domain : Banking**	**Python** : Machine Learning **Python Libraries** : Numpy, Scikit-learn, pandas, matplotlib, keras, tensor flow **ML Algorithms Used** : In this project, Unsupervised learning namely K Means clustering and Gradient Boosting Regression algorithm are used for obtaining high accuracy percentage. Also Supervised learning algorithms namely Neural Network, LR, LDA, KNN, CART, NB, and SVM, is used to effectively verify Model accuracy of the models.
7	**Census Income Dataset (using R)**	**R :** Machine Learning **Python Libraries** : Numpy, Scikit-learn, pandas, matplotlib, keras, tensor flow **Technology Stack and Algorithms used:** Exploratory Data Analysis, Data Pre-processing, Data Manipulation, Data Visualisation, Linear Regression, Logistic Regression, Decision Tree, Random Forest
8	**Covid-19 mortality rates.** **Domain :mHealth Care**	**Tableau Desktop**: Machine Learning **Python Libraries** : Numpy, Scikit-learn, pandas, matplotlib, keras, tensor flow
9	**Sales Analysis with respect to** ➢ **Customers personal data (mobile no, Email**	**SQL Server** Querying a large relational database using Adventure work database with SQL server

	Id etc) ➢ **Total Sales made to customers** ➢ **total sales made in the year by month order by increasing sales**	
10	**Analysis of the public review of the products on the Social media** **Domain – E commerce Website**	**Big Data Hadoop** Java, eclipse, neon, javascript, java compiler, apache
11	**Analysis of Taxi App pertaining to number of trips, total number of kilometres travelled, revenue generated.** **Domain : Taxi App**	**Big Data Hadoop Spark**
12	**Analysis of Cosmetic Products on Social Media Platform with respect to likes, dislikes** **Domain :Social Media**	**Spark**
13	**Title :Binary Classification on Customer Churning using Keras** **Domain : Telecom Industry**	**Artifiaial Intelligence** **Keras**
14	**Building an AI-based Chatbot using IBM Watson LAB** **Domain :E Commerce**	**Artificial Intelligence & Deep Learning with Tensorflow**
15	**K-Means cluster analysis on Iris dataset to predict the class of a flower using its petal's dimensions** **Domain: Analytics**	**SAS** Using the famous Iris dataset, predict the class of a flower Perform K-Means cluster analysis
16	**Projects in Kaggle** **Title - Bristol Myers Squibb--Molecular Transition** **Title - BirdCLEF 2021-Birdcall Identification:**	**Python, R** **Libraries used**: Pandas, Scikit-learn, Numpy, Keras, Tensor flow

	Title: RSNA-MICCAI Brain Tumor Radiogenic Classification. **Title: Tabular Playground Series**	

DATA SCIENCE PROJECTS DONE DURING DATA SCIENCE TRAINING PROGRAMS–Edvancer-IIT Kanpur

Machine Learning Projects- Python

Consumer Services- Consumer Complaints Resolutions

- ✓ BFSI Marketing- Understanding Customer Preferences in Insurance Sector
- ✓ Pharma Public Safety- Counterfeit Medicines- Prediction of Sales
- ✓ Manufacturing- Predict Hazard Ratings for a Maintenance Project
- ✓ Real Estate-Flag Junk Property Listings
- ✓ Health Care- To Predict No-shows given the appointment details

Artificial Intelligence Projects:

- ✓ Multiclass Multilabel prediction For stack overflow- Given text for Questions , predict tags associated with them
- ✓ Music Genre Identification- Given audio files for songs , identify which genre they fall in
- ✓ Spam filter for Quora questions
- ✓ Distracted Driver MultiAction Classification Page-Classification of the various distractions of a driver
- ✓ Image Captioning Page- Uploading photos and pixel analysis

Tableau Dashboard Projects:

- ✓ **Sales Analysis**
- ✓ **Customer Analysis**