Roll number	your name	your department	Your verification points(please give numbers for different points and also add title and duration if you want)					
180020091	S Anand Natarajan	Chemical Engineering	Music Genre Classifier | Course Project – Introduction to Machine Learning [Jan '21 – Apr '21] Guide: Prof. Amit Sethi 1) Implemented KNN, SVM, and Neural Network Classifiers to classify raw audio samples into 10 music genres 2) Achieved 75% accuracy and 96.2% AUC-ROC score using audio features like MFCCs, Chroma Energy Normalized etc					
203070095	Mrigank Srivastava	Electrical engineering	Trained SVM on classification data from the ISLR Library and UCI ML Repository with a prediction accuracy (min) of 86% on test data, discussed kernels on structured data.					
203190024	Abijith P Y	IEOR	Unsupervised Clustering for Indian Railway Optimization (Jan 2021 - Apr 2021) 1. Cleaned the raw data to obtained relevant information using pandas and matplotlib libraries 2. Implemented unsupervised clustering using K-Means algorithm and authenticated the solutions using silhouette scores and Elbow knee method 3. Developed a new algorithm to segregate trains on non-complementary days to identify least utilized tracks					
203190026	Zubeen Kishore Borkar	IEOR	Title: Grouping of Non-Daily Trains into Daily Trains Objective: To form a compact and optimized timetable for both daily and non-daily trains running on the same route and use clustering techniques to form group of trains. •Cleaned the raw data to obtained relevant information using pandas and matplotlib libraries •Implemented unsupervised clustering using K-Means algorithm and authenticated the solutions using silhouette scores •Developed a new algorithm to segregate trains on non-complementary days to identify least utilized tracks					
203190025	Sajith Menon	IEOR	Title :Grouping of Non-Daily Trains into Daily Trains. Objective : To form a compact and optimized timetable for both daily and non-daily trains running on the same route and use clustering techniques to form group of trains. 1.Cleaned the raw data to obtain relevant information using pandas and matplotlib libraries 2. Implemented unsupervised clustering using K-Means algorithm and Verified the solutions using Silhouette scores 3 Developed a new algorithm to segregate trains on non-complementary days to identify least utilized tracks					
203190022	Amit Gadekar	IEOR	Title: Human Activity Recognition using Machine Learning points: 1. Implemented Logistic Regression and Decision Tree from scratch in python to detect human activities. 2. Applied Blending ensemble technique on HAR dataset and compared accuracy with different ML techniques.					
203190023	Amit kumar seth	Ieor	Classified 60000 labelled Stack Overflow posts into 3 classes to identify the low-quality questions using NLP Compared and analyzed the performance of various models with respect to TFIDF and Global Vector embeddings Achieved an accuracy of 85% with Logistic Regression, 81% with SGD-C and 77% with Random Forest model					
203010013	Bhagyasree M	Aerospace	Building Machine Learning models from scratch 1) The three ML models namely Random forest, SVM and Neural networks were built from scratch using python without built-in libraries 2) The accuracies of the three models were compared with each other and with corresponding sci-kit libraries 3) Concluded that Random forest had the highest accuracy than SVM and the two layer Neural network Duration - Jan 21 - May 21					
180040025	Bhushan Misal	Civil	Music-genre detection (Mar'21-May'21) 1. Classified raw audio samples from various sources into 10 different genre using audio features like MFCCs, Chroma energy, etc. 2. Engineered a machine learning Model using PCA, Neural Networks and KNN to detect the genre of given audio with 75%+ Accuracy and 96.2% AUC-ROC score.					
193709008	Alen Rujis	Electrical	Content Based Image Retrival Jan'21-May'21 1. Extracted 12 images similar to that of a Query Image from the given dataset (Corel-5K) 2. Applied Transfer Learning approach for feature extraction of images using ResNet50					
193355001	Vishnu Jayan	CTARA	1. Extracted data of last 100 years from various websites and created a dataset for using Excel for predicting the yield of sugar cane at Maharashtra. 2. Modified the datasets by eliminating insignificant data points for applying machine learning algorithms 3. Applied linear regression, svm, forest algorithms to predict the yield and analyzed the accuracy of each methods					
19307R001	Shivali Bhatnagar	Electrical Department	1. SVM Classifier from scratch | Machine Learning Guide: Prof. Amit Sethi, EE Dept., IIT Bombay | Python (Jan'20 - Feb'20) a) Implemented and trained SVM Classifier using gradient descent method and plotted it in Python b) Visualized the decision boundary of the classifier trained on linearly non-separable data too 2. Attrition Classification | Machine Learning Guide: Prof. Amit Sethi, EE Dept., IIT Bombay | Python (Feb'20 - Mar'20) a) Performed classification on Kaggle dataset of Employee Attrition Problem using various ML Techniques b) Obtained the best accuracy of 86% on test data and 90% accuracy on validation data using XGB Classifier 3. Gender Recognition using Voice/Speech Analysis and ML Techniques | Machine Learning Guide: Prof. Amit Sethi, EE Dept., IIT Bombay | Python (Mar'20 - Apr'20) a) Implemented Statistical and Signal Processing approaches for feature extraction on different datasets of voice samples and trained different classifiers based on ML Techniques on training data b) Obtained test accuracy of 87% with XGBoost Classifier on real-time voice samples, using Signal Processing Approach and almost 100% accuracy on validation data					
193079004	Goutham A P	Electrical Engineering	These projects and assignments are from 'EE 769, Spring 2019' 1. Gender Recognition using Voice/Speech Analysis and ML Techniques| Machine Learning Guide: Prof. Amit Sethi, EE Dept., IIT Bombay | Python (Mar´20 - Apr´20) a) Proposed two different methods for feature extraction from voice samples - Extracted statistical features of dominant and fundamental frequencies from the voice sample Spectrogram - Implemented Signal Processing approch involving extraction of various spectral and cepstral features, MFCC, ZCR etc. b) Trained and tested various classifiers: KNN, Random Forest, Support Vector Machine, AdaBoost, XGBoost etc. c) Achieved almost 100% accuracy on validation data & 87% accuracy on real-time voice samples using Signal Processing Approach \& XGBoost Classifier. 2. IBM HR Analytics Employee Attrition Classification on Kaggle| Machine Learning Guide: Prof. Amit Sethi, EE Dept., IIT Bombay | Python (Feb'20 - Mar'20) a) Performed various data preprocessing and visualisation techniques on the attrition data. b) Implemented different feature extraction and ML classification methods involving PCA, SVM, Random Forest, AdaBoost etc. b) Ranked 11 among 195 participants by achieving the best accuracy of 90% on test data 3. SVM Classifier from Scratch |Introduction to ML [Jan'20 - Feb'20] a) Implemented and trained an SVM Classifier using gradient descent method and visualised it in Python b) Visualised the decision boundary implemented by classifier trained on linearly non-seperable data 4. SHALA (Stay at Home And Learn AI) [Jun'20] a) Completed assignments from below modules - Data Science - Machine Learning					
203350016	MD AQUIB	CTARA	TITLE: Prediction of Sugarcane yield using Machine Learning Modified the datasets by eliminating insignificant data points for applying machine learning algorithms Applied linear regression, SVMs, random forest algorithms to predict the yield and analyzed the accuracy of each methods					

Roll number	your name	your department	Your verification points(please give numbers for different points and also add title and duration if you want)					
203190013	Utkarsh Konge	IEOR	Title - AI-based Moderator for Online Forums Duration - Jan-May'21 Points - Classified 60000 labelled Stack Overflow posts into 3 classes to identify the low-quality questions using NLP Compared and analyzed the performance of various models with respect to TFIDF and Global Vector embeddings Achieved an accuracy of 84% with Logistic Regression, 81% with SGD-C and 77% with Random Forest model Identified low-quality questions through multi-class classification of 60K Stack Overflow posts using Natural Language Processing Reviewed and compared performance of TFIDF and Glove embeddings	Achieved an accuracy of 84% using Logistic Regression				
203190002	Shubham Bhasin	IEOR	AI Based Moderator for Online forums Jan - May'21 Course Project	Course: Introduction to Machine Learning	Guide: Prof. Amit Sethi	Tool: Python, NLTK •Identified the low-quality questions by classifying 60000 labeled Stack Overflow posts into three classes using NLP •Compared and analyzed the performance of various models with respect to TF-IDF and Global Vector embeddings •Achieved an accuracy of 84% with Logistic Regression, 81% with SGD-C, and 77% with Random Forest model					
193236001	Harsha Priyanka Guntaka	Systems and Control	1. Employee Attrition Prediction [Mar'20]: (a) Performed classification on Kaggle dataset for employee attrition using various machine learning algorithms (b) Identified significant features in the dataset and obtained 88.88%accuracy using Support Vector Machine(SVM) 2. Support Vector Machine [Feb'20]: (a) Implemented Support Vector Machine using gradient descent to classify both linear and non-linear datasets (b) Applied polynomial kernel for the given non-linear dataset and visualized the decision boundary obtained					
19307R019	Raghuwanshi Ajinkya Sangra	Electrical Engineering	•Road Line Lane Detector Jan'20 - Jun'20 Course: Introduction to Machine Learning	Prof. Amit Sethi •Implemented the design using Hough's Transform method to detect the lane markings of a road •Tested and verified the working of the model using a video data file which can be extended in real-time system •Attrition Classification Jan'20 - Jun'20 Course: Introduction to Machine Learning	Prof. Amit Sethi •Designed the ML model with various algorithms to get the best result for the provided data set •Compared all the incorporated models and took the most promising of all to give the best output •Pre-processing (one-hot encoding, feature scaling) was performed (accuracy - 89.3% & class highest - 92.59%) •SVM Classifier from scratch Jan'20 - Jun'20 Course: Introduction to Machine Learning	Prof. Amit Sethi •Implemented and trained SVM Classifier using gradient descent method and developed the system in Python •Visualized the decision boundary implemented by classifier trained on linearly inseparable data to					
203070028	Vyomkesh Chaudhary	Electrical Engineering	Project Name : Image Segmentation Duration : Jan'21-May'21 – Developed a model to perform automatic segmentation on various images of animals and the background. – Obtained 0.7 IOU (Intersection over Union) Score on the Oxford-IIIT Pet Dataset using U-Net architecture. (Assignment 2 as project) Project Name : Prediction of non-functional water pumps Duration : Mar'21-Apr'21 – Classification on dataset using various ML Techniques and obtained best accuracy using Random Forest – Finding the best accuracy was through constant iteration of feature selection, feature elimination and hyperparameter tuning					
193109014	Manthan Dhisale	Mechanical department	1. Development of Convolution Neural Network Speed-up with Sparsity [Industry Sponsored Project, in association with Sentry AI, USA] [Prof. Amit Sethi	Introduction to Machine Learning: Course Project] •Highlights: The project achieved a boost in the inference speed of the existing YOLOv5 model by pruning and sparsity over the neural nets. The optimal sparsity obtained was 30% for a YOLOv5 with a 3.57% Precision drop and 3.125% Recall drop 2. Development of Convolution Neural Network Speed-up with Sparsity [Industry Sponsored Project, In association with Sentry AI, USA] [Prof. Amit Sethi	Introduction to Machine Learning: Course Project] Results: The optimal sparsity in YOLOv5 obtained was 30% for a YOLOv5 with a 3.57% Precision drop and 3.125% Recall drop	[Jan '21-May '21] [Jan '21-May '21]				

193079002	Akshay Bajpai	Electrical Engineering - El	1 . Medical Image segmentation
			Guide: Prof. Amit Sethi, EE Dept., IIT Bombay| Advanced Machine Learning (Aug 20' - Dec 20')
			   ◦ Objective: Segmentation of the right ventricle from cardiac MRI using UNet Based Deep Learning models
			   ◦ Performed data augmentation on RVSC-MICCAI 2012 dataset using using several standard techniques.
			   ◦ Experimented with various loss functions like BCE loss, Focal Loss, BCE+Dice loss,
			     BCE+Dice+InverseDiceloss, Switching loss and obtained an accuracy of 72% with
			     BCE+Dice+InverseDice loss.
			
			2. Sentiment Analysis on IMDB dataset
			Guide: Prof. Amit Sethi, EE Dept., IIT Bombay
				Advanced Machine Learning (Aug 20' - Dec 20')
				   ◦ Trained LSTM, GRU, Bi-LSTM & Bi-GRU to study performance changes with number of units and layers and the effects of using different learning-rate schedulers like Cosine Annealing scheduler, Exponential rate scheduler, Step scheduler and Reduce Learning Rate(LR) on Plateau scheduler.
				   ◦ Achieved an accuracy of 87% using BiGRU and 86% using LSTM with Reduce LR on Plateau Scheduler.
				
				3. PyQT based UI to train ML models
				Guide: Prof. Amit Sethi, EE Dept., IIT Bombay
					Introduction to Machine Learning (Jan 20' - May 20')
					   ◦ Developed UI using PyQT to enable quick data visualisation and ML model training.
					   ◦ Implemented several models (Linear regression,SVM,Logisitc regression, Neural networks,Decision tress, Random forests) that could be tuned, trained and saved on the input data.
					
					4. Attrition Classification
					Guide: Prof. Amit Sethi, EE Dept., IIT Bombay| Introduction to Machine Learning (Jan 20' - Jun 20')
					   ◦ Achieved accuracy of 89.30% on IBM attrition dataset by training SVM classifier on Kaggle.
					   ◦ Extracted relevant and less correlated features & applied "One Hot Coding" for features with multi-classes.					

Roll number	your name	your department	Your verification points(please give numbers for different points and also add title and duration if you want)					
193079003	Soham Naha	Electrical Engineering	1. Sentiment Analysis Guide: Prof. Amit Sethi, EE Dept., IIT Bombay | Advanced Machine Learning (Aug 2020 - Dec 2020) ◦ Trained various neural networks LSTM, GRU, Bi-LSTM and Bi-GRU on IMDB dataset to study their performance on changing the number of units and layers and achieved an accuracy of 87% using Bi-GRU ◦ Studied the effects of using different learning-rate schedulers like Cosine Annealing, Exponential & Step 2. Medical Image segmentation Guide: Prof. Amit Sethi, EE Dept., IIT Bombay | Advanced Machine Learning (Aug 2020 - Dec 2020) ◦ Performed data augmentation on RVSC-MICCAI 2012 dataset using color jitter, random horizontal flip, and random vertical flip to include defects that might occur in real image acquisitions ◦ Experimented with various loss functions like BCE loss, Focal Loss, BCE+Dice loss, BCE+Dice+InverseDice loss, Switching loss and obtained an accuracy of 72% with BCE+Dice+InverseDice loss 3. Attrition Classification Guide: Prof. Amit Sethi, EE Dept., IIT Bombay | Introduction to Machine Learning (Jan 2020 - Jun 2020) ◦ Achieved accuracy of 89.3% by training SVM classifier on Kaggle IBM Attrition dataset. ◦ Extracted relevant and less correlated features and applied One-Hot Encoding for features with multi-classes 4. Support Vector Machine Guide: Prof. Amit Sethi, EE Dept., IIT Bombay | Introduction to Machine Learning (Jan 2020 - June 2020) ◦Objective: Implement the modified SVM algorithm in the paper titled Pegasos:Primal Estimated sub-GrAdient SOlver for SVM using NumPy. ◦Use SVM classifier on linear data and kernelized-SVM on non-linear data					
203350022	Sourabh Jhariya	CTARA	1. Implemented Linear Regression, SVM & Random Forest algorithms by analyzing the accuracy of each algorithm to predict the sugarcane yield in Maharashtra 2. Modified the datasets by eliminating the insignificant data points for applying in machine learning algorithms Title: Prediction of Sugarcane Crop Yield using Machine learning algorithms					
19307R002	Samarth Nigam	EE	AML-- 1. Image Segmentation of Right Ventricle of Human Heart | Advanced Machine Learning Instructor : Prof. Amit Sethi , Electrical Engineering, IIT Bombay (Oct'20 - Nov'20) –Engineered Pixel-Wise Segmentation of Right Ventricle of Heart from MRI Images for RVSC dataset by implementing U-Net Architecture as a baseline model to get a dice score of 0.2761 –Performed ablation studies along with hyperparameter tuning to achieve the best dice score of 0.2426 – Performed ablation studies along with hyperparameter tuning to achieve an accuracy of 75.74% 2. Sentiment Analysis in NLP using RNNs | Advanced Machine Learning Instructor : Prof. Amit Sethi, Electrical Engineering, IIT Bombay (Nov'20) –Designed an LSTM based sentiment classifier for IMDB movie database using Glove library –Incorporated models using Bidirectional LSTMs and GRUs to compare the accuracies and achieved a test accuracy of 87.51% for a 2 layer Bidirectional LSTM model with BCE as loss function 3. Emotion and Gender Recognition from Facial Image Data| Advanced Machine Learning Instructor : Prof. Amit Sethi, Electrical Engineering, IIT Bombay (Nov'20-Dec'20) –Employed D-CNN based automatic Facial Emotion and Gender detection system, trained on fer2013 dataset and deployed it over a GUI software created using python's tkinter module –Extracted faces from self-generated test images videos using Harr Cascade Classifiers, trained a D-CNN from scratch and exploited transfer learning by using VGG-16to achieve a decisive accuracy of 75% IML -- 1. House Pricing Prediction | Introduction to Machine Learning Instructor : Prof. Amit Sethi, Electrical Engineering, IIT Bombay (Apr'20-May'20) –Diversified various tree-based regression models and performed pre-processing, feature engineering and training on Ames Housing Dataset to determine house prices on the basis of 79 features –Secured a rank of around 200 by implementing a StackingCVRegressor as an ensemble of 3 regressors, namely, GradientBoostingRegressor, LightGBMRegressor and XGBosstRegressor 2. Attrition Classification | Introduction to Machine Learning Instructor : Prof. Amit Sethi, Electrical Engineering, IIT Bombay (Mar'20) – Facilitated classification on Employee Attrition Problem dataset of Kaggle using various ML Techniques and obtained the best accuracy using Logistic Regression – Analysed dataset and structured pre-processing, including one-hot encoding and feature scaling, before training to achieve an accuracy of 87.65% on test set where class highest was 92.59% 3. SVM Classifier from scratch | Introduction to Machine Learning (Feb'20) – Implemented and trained SVM Classifier using gradient descent method and visualized it in Python – Visualised the decision boundary implemented by a classifier trained on linearly inseparable data too					
193109018	Chinmay Vilasrao Gandhshre	Mechanical Engineering	Title - AI Bot to predict stock Market Points- 1) Created Intelligent Bot to maximize the profit by using Q-Learning which is model-free Reinforcement Learning 2) Trained more than 11k parameters using Deep Q-Network which is build using Tensorflow& Data reader library of python. 3) Apple stock market dataset which is collected since year 2010, till today (it gets updated) is used for training purpose.					
193109013	Supriyo Roy	Mechanical Engineering	Product Attribute Extraction from Text Course Project | Introduction of Machine Learning | Guide: Prof. Amit Sethi [Jan-May'21] Objective: To extract brand name from product title using NLP Techniques 1. Converted words to vectors using Glove Embedding and Experimented Bi-LSTM & CNN for attributes extraction 2. Attained data validation accuracy of 90.73% & 98.01% using Bi-LSTM & CNN Algorithm respectively					
193079005	Varsha S	EE	1) Droplet Detection on Camera Lens | Introduction to ML | Prof. Amit Sethi, Saqib Shamsi (Whirlpool), (Spring'20) a) Designed a deep learning model to detect whether the droplets in an image are distortion on the lens or area part of the scene captured, with ResNet-18, VGG and DenseNet as various backbone architectures. b) Experimented with Label smoothing and Adversarial Learning on the dataset compiled from Youtube videos, obtained a test accuracy of 69.46% using weighted cross entropy loss on ResNet-18 backbone. 2) Image Segmentation of Right Heart Ventricle | Advanced Topics in ML | Prof. Amit Sethi a) Engineered Pixel-wise Segmentation of Right Ventricle of Heart in MRI Images from RVSC-MICCAI 2012 dataset by implementing a U-Net Architecture as baseline model to get a Dice-score of 0.2761. b) Performed ablation studies along with hyperparameter tuning on learning rate, optimizers, dropout and varied loss functions to achieve the best accuracy of 76.74%. 3) Sentiment Analysis in NLP using RNNs | Advanced Topics in ML | Prof. Amit Sethi, (Autumn'20) a) Implemented LSTM based sentiment classifier for IMDB movie database using GloVe. b) Achieved a test accuracy of 87.51% for a 2 layer Bidirectional LSTM model with Binary Cross Entropy loss. 4) Emotion and Gender Recognition from Faces | Advanced Topics in ML | Prof. Amit Sethi, (Autumn'20) a) Trained a D-CNN based automatic Facial Emotion and Gender detection system, on FER2013 dataset. b) Extracted faces from self-generated test images and real-time videos using Harr Cascade Classifiers, trained a D-CNN from scratch and exploited transfer learning by using VGG-16 to achieve a decisive accuracy of 75%. 5) Attrition Classification | Introduction to Machine Learning | Prof. Amit Sethi, (Autumn'20) a) Classified on Employee Attrition Kaggle dataset by conducting data analysis and various mix of classifiers. b) Obtained the best test accuracy of 90.09% on test set with GradientBooster classifier. 6) SVM Classifier from scratch | Introduction to Machine Learning | Prof. Amit Sethi, (Autumn'20) a) Implemented and trained SVM Classifier using gradient descent method and visualised it in Python. b) Visualised the output decision boundary of the model trained on linearly separable and non-separable data.					

Roll number	your name	your department	Your verification points(please give numbers for different points and also add title and duration if you want)
193109018	Chinmay Vilasrao Gandhshre	Mechanical Engineering	AI Bot to predict stock Market| Prof. Amit Sethi [Jan'21 – May'21] •Created Intelligent Bot to maximize the profit by using Q-Learning which is model-free Reinforcement Learning. •Trained more than 11k parameters using Deep Q-Network which is build using Tensorflow & Data reader library. •Apple stock market dataset which is collected since year 2010, till today (it gets updated) is used for training purpose.
193079015	Nimish Dharamshi	Electrical Department	1) Employee Attrition Classification | Introduction to Machine Learning (Spring'19) ▫Performed pre-processing steps such as data cleaning, backward, forward & correlation based feature selection,encoding categorical data fields, tackling class imbalance using class resampling & loss function weighting,standard scaling and Principal Component Analysis. ▫Trained and tuned four different models: Logistic regression, random forest classifier, MLP and SVM, Adaboost and Stack classifier for the binary classification. Evaluation metric studied were confusion matrix,accuracy and achieved best accuracy of 90.29%. 2) Kaggle Competition "House Price Advanced Regression" | Introduction to Machine Learning (May'20 - Jun'20) ▫Predicted house price values of79-dimensional house features, achieving best performance of0.1174RMSE. ▫Handled outliers & highly skewed features, features dimensionality reduction, meta model stacking. Extensively experimented with regression models like SVR, Random Forest, gradient boosting. 3) Sentiment Analysis in NLP using RNNs | Advanced Machine Learning (Sept'20) ▫Analyzed sentiments on IMDB movie database, using RNNs like bidirectional LSTM, GRU. ▫Experimented with hyperparameters such as word embeddings and its dimensionality, number of RNN layers, dropout percentage to achieve the best performance of 90.09% accuracy on IMDB database. 4) Handwritten digit generation and classification | Advanced Machine Learning (Nov'20 - Dec'20) ▫Generated digits using Generative Avdserial Networks for data augmentation applications. ▫Compared Vanilla GAN, Deep Convolutional GAN, Conditional GAN results qualitatively for mode collapse, class imbalance, model and time complexity, and convergence problems
193079001	Suhasinee Jain	Electrical Engineering	1)House Price Prediction | Introduction to Machine Learning (May'20-Jun'20) ▫Implemented ensemble based regression model on kaggle house prediction dataset and achieved 0.12 RMSE. ▫Performed data cleaning on79 dimensional dataset to remove null values and outliers followed by feature selection, and experimented with models such as random forest, adaptive and gradient boosting, etc. 2)Handwritten Digit Generation | Advanced Topics in Machine Learning (Nov'20-Dec'20) ▫Implemented and trained three different GANs: VanillaGAN, DCGAN and cGANon MNIST dataset. ▫Compared the GANs based on model and time complexity, mode collapse and quality of image generated. 3)Pixel-wise Image Segmentation of Right Ventricle of Heart | Advanced topics in machine learning (Sep'20) ▫Experimented UNet model with different loss functions like BCE, dice, inverse dice and switching loss. ▫Tuned regularization coefficient and schedulers to achieve dice loss of 0.12 on RVSC-MICCAI 2012 dataset. 4)Sentiment Analysis in NLP | Advanced Topics in Machine Learning (Sep'20) ▫Trained RNN models with an embedding layer for the task and achieved 85.16% accuracy on IMDB dataset. ▫Experimented with uni and bidirectional LSTM and GRUs and different dimensions of word embeddings. 5)Classification for Employee Attrition | Introduction to Machine Learning (Feb'20-Mar'20) ▫Performed pre-processing steps such as data cleaning, feature selection using correlation and recursive feature elimination, encoding categorical data fields, upsampling, standard scaling and PCA. ▫Trained and tuned four different models: Logistic regression, random forest classifier, MLP and SVM for the binary classification and achieved best accuracy of 91.414% with SVM model
180020058	Neelkamal Bhuyan	Electrical Engineering	EE769 (Spring 2021) Project Title - Dimensionality Reduction Algorithms Applied to Hyperspectral Images 1. Evaluated evolutionary algorithms' performance in dimensionality reduction of satellite images. 2. Compared feature reduction by Genetic Algorithms (GA) and Ant Colony Optimization (ACO) with PCA. 3. GA and ACO were implemented in python and were evaluated with F1-score in pixel classification by SVM
203100068	Mudit Sand	Mechanical Engineering	Product Attribute Extraction From Text [Jan-May'21] Course Project | Introduction to Machine Learning | Guide: Prof. Amit Sethi Objective: Extract essential features from the given text •Used Bi-LSTM and CNN Layer architecture to fit the models for Natural Language Processing. •TensorFlow is used for processing the data and building the models. •Achieved a Data validation accuracy of more than 90%.
18D110017	Harsh Tantway	Metallurgical Engineering	TITLE - Multi-class poverty level classification and prediction of households DURATION - January to April, 2021 1. Implemented a poverty prediction model in order to segregate households into four levels of poverty (supervised multi-class classification machine learning model) 2. The aim was to utilize individual and household socio-economic parameters and try to correctly classify families on the basis of their poverty level. (*) The project followed a process flow as mentioned below - 3. Extracted the data of Inter-American Development Bank, cleaned its improper variable values and reassigned/filled the null values of the dataset 4. Exploratory data analysis and visualizations: Plotted KDE(Kernel density estimation) plots of the key features such as number of persons per room, monthly payment, etc to analyze the dependence and variation of variables 5. The Spearman and Pearson correlation values were found to be 1.0 and 0.96 between two variables which were related in a quadratic fashion and hence used feature reduction in this case 6. The imbalanced classification problem was tackled by using the weighted F1 score in place of the normal F1 score. Empty and erratic labels in the dataset have been duly corrected. 7. Applied various ML classification models such as Random Forest Classifier, Logistic Regression, Support Vector Classifier and K-Neighbors Classifier with number of neighbors as 5, 10 and 20 and compared these models on the basis of their F1 scores and inferred that the Random Forest Classifier outperf
18D100013	Om R Ingole	Mechanical Engineering(t	TITLE - Multi-class poverty level classification and prediction of households DURATION - January to April, 2021 1. Implemented a poverty prediction model in order to segregate households into four levels of poverty (supervised multi-class classification machine learning model) 2. The aim was to utilize individual and household socio-economic parameters and try to correctly classify families on the basis of their poverty level. (*) The project followed a process flow as mentioned below - 3. Extracted the data of Inter-American Development Bank, cleaned its improper variable values and reassigned/filled the null values of the dataset 4. Exploratory data analysis and visualizations: Plotted KDE(Kernel density estimation) plots of the key features such as number of persons per room, monthly payment, etc to analyze the dependence and variation of variables 5. The Spearman and Pearson correlation values were found to be 1.0 and 0.96 between two variables which were related in a quadratic fashion and hence used feature reduction in this case 6. The imbalanced classification problem was tackled by using the weighted F1 score in place of the normal F1 score. Empty and erratic labels in the dataset have been duly corrected. 7. Applied various ML classification models such as Random Forest Classifier, Logistic Regression, Support Vector Classifier and K-Neighbors Classifier with number of neighbors as 5, 10 and 20 and compared these models on the basis of their F1 scores and inferred that the Random Forest Classifier outperf
180040012	Ankit Kumar	Mechanical engineering	Twitter Sentiment Extraction | Course Project(Feb '21 - April '21) EE769 Course Project | Guide: Prof Amit Sethi, IIT Bombay •Automated the detection of the word of the tweet that decides its sentiment through transfer learning •Achieved a Jaccard score of 70 percent on the test Data set implementing Roberta architecture •Increased the score by 10 percent through the labeling issue removal by processing the selected text of tweets
17d070051	Botcha Ritesh sadwik	Electrical engineering	Conditional generative adversarial networks (mentor: Aritomo Shinozaki, Duration: feb-may 2021, Prof: Amit sethi, Course: ee769) 1. Worked on an industry project for Sentry AI under Aritomo Shinozaki in fields of cGANs 2. Compared and analyzed SBGAN image generation performance with other SOTA methods like BigGAN, ProGAN 3. Synthesized semantically segmented images upon which generated complex image scenes 4. Worked on datasets like Cityscapes-25K, ADE_Indoor containing over 30000+ images
203020005	Mishra Digvijay Sanjaybhai	Chemical Engineering	Social Distance Detection in Video Using Deep Learning | Introduction to ML | Prof. Amit Sethi, IIT Bombay [Jan'21-Apr'21] 1)Loaded the COCO dataset, which contains around 0.12 million images; loaded weights which are trained using just human label 2)Implemented the regression-based model YOLO-V3 from Computer Vision (DNN module) for object detection using Python 3)Used Euclidean distance for human identification & achieved 37% Mean Average Precision (MAP) on 25 FPS video
18d110015	Anshul Kumar	Mems	Title: prediction of graduate admissions chance A three months long project including data collection, secondary research on available similar projects and then coming up with my own solution with repor. 1) performed EDA on the collected data set to clean the data set and Prepared the data sets for the ml model 2) used various models(including LR, SVM, K- NN, Random forest and gradient booster) and then got the best model to get to the result 3) used mae as measuring unit
180260043	Yash Dixit	Electrical Engineering	1)Reviewed numerous research papers in the field of NLP and Sentiment Analysis 2)Manipulated textual data from Amazon's customer review database by means of lemmatization and vectorization to make it suitable for classification 3)Used various ML models and fine tuned their hyperparameters to classify the reviews into 5 categories with the highest accuracy
203020048	Vishal Saini	Chemical engineering	Title - Social distance detection by yolo object detection [jan'21-may'21] * Worked on object detection model named YoLo , using COCO data set containing 121408 images as training set * Utilized Euclidean distance for tracking of object, further measuring distance between them to conclude social distancing

Roll number	your name	your department	Your verification points(please give numbers for different points and also add title and duration if you want)						
203190027	Arvind Kumar	IEOR	Title : Handwritten Equation Solver Using CNN Points :1.Evaluate and implement code handwritten solver 2.Compare Hand written equation solver using CNN 3.Develop a GUI interface to take input from user. 4.Successfully got accuracy of 98% on MIST dataset vs 80% on Kaggle dataset						
203070019	Anurag Dixit	Electrical Engineering (EE	Project Name: Automated Snake Game Agent using DQN Algorithm (March 2021 to May 2021) 1. Environment Initialization and customization for Snake Playground. 2. Selection of best training algorithm to maximize reward for our agent. 3. Training of the neural-network and hyper-parameter tuning using DQN Algorithm. 4 . Presenting results of our DQN greedy agent with respect to default greedy agent.						
18B090004	Dipanshu Sharma	Mathematics	Bank Note Authentication End to End ML Project (Mar '21 - Apr '21) • Developed an ML model to check the authenticity of the banknotes and created a front-end to maximize the reach of the model • Implemented model selection, hyperparameter tuning and model evaluation methods and selected a model with 99.7% accuracy • Used NumPy, pandas, matplotlib, sklearn for the backend and also used Flask, Flassger, Swagger and Pickle for the front-end API • Deployed the model on the web using Heroku Cloud-based platform and used Postman for visualization on the hosts' system						
203070010	Satyam Keshri	Electrical Engineering	Title - Music Genre classification using Spotify API Date - May 2021 1. Implemented a custom CNN based classification algorithm for genre classification on Spotify songs extracted from Spotify Web API. 2. Handled the class imbalance for 8 genre classes with more than 1500 data samples and achieved an accuracy of 71.3%. 3. Extracted pitch and timbre vectors from the API to create a 2D spectrogram like input features. 4. Used Keras with Tensorflow backend along with spotipy, numpy, seaborn libraries						
203195001	Anand Siyote	IEOR	Title : Handwritten equation solver using CNN 1. Comparison of handwritten equation solver using CNN with other methods. 2. Developed the code to include a GUI input screen for taking input. 3. Successfully got an accuracy of 98% on MNIST dataset versus 80% on Kaggle dataset.						
180100040	Divyansh Srivastava	Mechanical Engineering	Deep Image Prior	Course Project (Mar'21 - Apr'21) Instructor – Prof. Amit Sethi, Dept. Of EE, IIT Bombay	Course – EE 769 1. Implemented paper Deep Image Prior in python using keras & tensorflow libraries for image restoration 2. Performed image denoising & super resolution from noise using structural information present in the NNs						
19307R021	Yash Pratap Singh Tomar	Electrical Engineering	Optical Character Recognition for KYC through PAN cards March 2020 - June 2020 ◦ Build an OCR system for PAN card images using YoloV3 algorithm and PyTesseract in collaboration with Arthimpact Digital Loan Pvt, Ltd. Mumbai. ◦ The model could extract Name, Father's Name, DOB, PAN number, signature and photograph.						
193079033	Sachin Doifode	Electrical Engineering	1) Emotion and Gender Recognition from Facial Image Data using Deep Learning [Oct'20 - Dec'20] Course: Advance Topics in Machine Learning, Guide: Prof. Amit Sethi, EE, IIT Bombay • Implemented D-CNN based automatic Facial Emotion and Gender detection system, trained on Fer2013 dataset • Extracted faces from self-generated test images and real-time videos using Harr Cascade Classifiers with accuracy of 75% 2) Image Segmentation of Right Ventricle of Human Heart [Jul'20 - Oct'20] Course: Advance Topics in Machine Learning, Guide: Prof. Amit Sethi, EE, IIT Bombay • Engineered Pixel-Wise Segmentation for RVSC-MICCAI 2012 dataset by using U-Net to get a dice score of 0.2761 • Performed ablation studies with hyperparameter tuning, learning rate, optimizers & dropout, achieving accuracy of 76.74% 3)Sentiment Analysis in NLP using RNNs [Oct'20 - Nov'20] Course: Advance Topics in Machine Learning, Guide: Prof. Amit Sethi, EE Dept., IIT Bombay • Designed and developed LSTM based sentiment classifier for IMDB movie database using Glove word embedding • Implemented models using Bidirectional LSTMs and GRUs and achieved an accuracy of 87.51% with BCE loss function						
203190021	Kshitij Kushwaha	IEOR	Project title - Vehicle Detection for Self driving cars Tool: Sklearn, Python, OpenCV2 Course: Introduction to Machine Learning Guide: Prof. Amit Sethi Project duration: Mar'21 - May'21 Objective: To train a machine learning model which can detect a car image from the image captured by a car's dash camera 1. Studied and applied Histogram of Oriented Gradients (HOG) feature descriptor for feature engineering. 2. Trained and compared different ML algorithms like SVC, Random Forest, Perceptron, and Adaboost classifier. 3. Achieved an accuracy of 8/9 on test images using the SVC model and performing appropriate hyperparameter tuning. 4. Reduced total number of features to 1764 (by 85.67%) using Histogram of Oriented Gradients feature descriptor. 5. Performed hyperparameter tuning and cross-validation using GridSearch for SVC, Adaboost, and Random Forest. 6. Attained ROC-AUC score of 0.9951 for SVC and obtained 8/9 correct classification on the test image set 7. Reduced number of features from 12288 to 1764 (by 85.67%) using Histogram of Oriented Gradients feature descriptor.						
203190030	Sandesh Bhaskar Gaikwad	IEOR	Title : Vehicle Detection for Self Driving Cars Duration : Mar'21-May'21 Tools : OpenCV , Scikit-learn ,Python Points : 1)Studied and applied Histogram of Oriented Gradients (HOG) feature descriptor for feature engineering 2)Reduced total number of features in an image from 12288 (i.e.64×64×3 image pixel) to 1764 (i.e. 85.67%) using Histogram of Oriented Gradients feature descriptor 3)Reduced total number of features to 1764 (by 85.67%) using Histogram of Oriented Gradients feature descriptor 4)Performed hyperparameter tuning and cross-validation using GridSearch for SVC, Adaboost, and Random Forest 5)Trained and compared different ML algorithms like SVC, Random Forest, Perceptron and Adaboost classifier 6)Attained ROC-AUC score of 0.9951 for SVC and obtained 8/9 correct classification on the test image set						
203190005	Rishabh Kumar	IEOR	Title : Image Segmentation Duration : Jan'21 - may'21 Points: 1. Developed a model to perform automatic segmentation on various images of animals and the background. 2. Obtained 0.7 IOU (Intersection over Union) Score on the Oxford-IIIT Pet Dataset using U-Net architecture.						

Roll number	your name	your department	Your verification points(please give numbers for different points and also add title and duration if you want)
203350013	Kaustubh Patil	Technology & Developme	Course Project Title: Machine failure forecasting using sensor data on cloud Duration: Apr'20-May'20 Course: EE 769 – Introduction to Machine Learning	Prof. Amit Sethi, IIT Bombay 1. Implemented Logistic Regression and Random Forest Regressor on machine sensor cloud data 2. Created an API in Google Apps Script editor to insert data into Google Spreadsheet 3. Evaluated these algorithm by Accuracy, Recall, MAE and Specificity; LR identified anomalies with 5.3% false positives Title: Implementation of Machine Learning Regression and classification models - LASSO, Elastic Net, Support Vector Machine, and Random Forest 1. Performed exploratory data analysis using box plot and heatmaps; Trained and tested wine quality dataset (12 features) 2. Trained and tested Mice Protein Expression data set (80+ features) and validated from K-fold Cross Validation 3. Improving the prediction of model using hyper parameter tuning and achieved accuracy of 55% with RF model Title: Prediction of Functionality of Pumps from African Pump data set using ML Classification models 1. Applying Data Visualization to get the outliers, to check for data imbalance and performing Data preprocessing by checking for missing values, scaling and encoding on 40+ features 2. Trained and tested Support Vector Machine, Random Forest, and K-Nearest Neighbor classification models [1]
203110001	Avdhesh Kumar Sharma	MEMS	Title- Content Based Image Retrieval Using SVM model.(Jan- May'21) 1)Extracted 12 similar images with 98% accuracy from the given dataset (Corel-5K) of Query Image. 2)Applied Transfer Learning approach (SVM Model) for feature extraction of images using ResNet50.
180040093	Saurabh Khandelwal	Chemical Engineering	SKETCH TO COLORED IMAGE GENERATION USING GAN	Prof. Amit Sethi	Course Project [JAN'21 - MAY'21] ▢ Implemented Generative Adversarial Network using TensorFlow to automate the colorization of greyscale sketches ▢ Optimized hyperparameters including batch size and buffer size to minimize the generator and discriminator loss functions
193079019	Saloni Shah	Electrical	ML accelerated Full Waveform Inversion of Ground Penetrating Radar Data	Introduction to ML(May '21) oGuide : Prof Amit Sethi oImplemented two types of machine learning architectures - Bidirectional recurrent neural networks (BiRNN) with long short-term memory (LSTM) and Temporal Convolutional Network (TCN) oTested both these ML architectures on MNIST Data oGenerated data using gprMax using 2000 random variations of parameters such as radius of bar, cover depth and water content of the concrete used and tried implementing both the architectures on GPR data
203110061	Rajrishi Sarkar	MEMS	1)Using LSTM in stock analysis|Deep learning Course Project (Feb'21-May'21) Guide: Prof Amit Sethi •Extracted the company data from NY finance, preprocessed it & performed EDA. •Applied Long Short-Term Memory (LSTM) model with 3 added layers and dropout regularization to predict the closing stock-prices and visualized the training and validation loss of the model and achieved an R2-score of 0.94 and MSE of 0.05. 2)Classifying mice protein expressions on the basis of genotypes|ML assignment (March'21) Guide: Prof Amit Sethi •Preprocessed the data, imputed missing values and obtained the most important variables w.r.t the mice protein classes. •Identified best classification models on the basis of cross validation scores and performed hyperparameter tuning on random forest classifier and support vector classifier. Obtained an accuracy of 100%. Visualized the model performances by plotting their learning curves.
180100123	Vankala N Sai Krishna Kumar	Mechanical Engineering	Snake Game Playing Agent Using ML April 2021Course Project under the guidence of Prof. Amit Sethi •Created a ML agent which can play snake game based on explore-exploit algorithm. •Lead a group of 3 members to create the above agent using Deep Q-Learning Algorithm. •This Agent clearly surpasses the agent which is created by using traditional algorithm.
193079008	Alen Rujis	Electrical	Prediction of Defective water pumps using given dataset Classification on dataset using various ML Techniques and obtained best accuracy using Random Forest Finding the best accuracy through constant iteration of feature selection, feature elimination, hyper parameter tuning.
203070003	Vedant Kandoi	Electrical	Jan 2021 - May 2021 Introduction to Machine Learning 1. Designed an ML model for speech emotion recognition using Convolutional Neural Networks in a team of 3. **Main Project** 2. Classification of Down Syndrome using various models like Lasso, Ridge and Support Vector Classification. **Assignment** 3. Determining quality of wine using regression with Lasso, Ridge and Support Vector Regression. **Assignment** 4. Designed a model to determine quality of water pumps using Support Vector Machines. **Assignment**
203110002	Aaryan Bagani	MEMS	Course Project: Content Based Image RetrievalJan'21 - May'21 Introduction to Machine Learning	Prof. Amit Sethi Extracted 12 images similar to that of a Query Image from the given dataset (Corel-5K) Applied Transfer Learning approach for feature extraction of images using ResNet50
203110049	Ankush nag	MEMS	Title:- Multi Level question prediction in e-commerce conversation Duration:- Feb'21-Apr'21 Points:- 1) Used Requests and Beautiful Soup for scraping and parsing data from the Websites. 2) Trained word2vec and BERT model on Q&A from ecom websites, extracted using python, RegEx and NLP. 3) Applied Genism & Cosine similarity package to find the similarity between the trained and user query. Title:- Predict the quality of red and white wine using classification based model Duration:- Feb'21 Points:- 1) Test accuracy of 94% after Extra tree classifier model training, hyperparameter tuning using RandomizedSearchCV and data pre-processing (including handle imbalanced Data Set using RandomOverSampler). Title:- Predict which pumps are functional, non- functional and which need some repairs. Duration:- Mar'21 Points:- 1) Performed EDA followed by Feature engineering including handle imbalanced Dataset, RFECV. 2) Hyperparameter tuning and Training on XGB, Randomforest and Extratree,followed by visualization using TSNE.
203100049	Modi Harsh Jashvantbhai	Mechanical Engineering	Detection of American Sign Language Letters using Machine Learning; Dr. Amit Sethi [Mar-May 2021] •Self-created/collected training and testing data from various sources and applied pre-processing methods. •Trained Convoluted Neural Network to recognize sign language letters using webcam and analyzed the correctness/errors.
180020002	Aakash Kumar Singh	Chemical Engineering	Duration (Feb'21-May'21) Course Project Under Prof Amit Sethi (IIT Bombay) • Automated the detection of the word of the tweet that decides its sentiment through transfer learning • Achieved a Jaccard score of70 percent on the test Data set implementing Roberta architecture

Roll number	your name	your department	Your verification points(please give numbers for different points and also add title and duration if you want)					
203070005	Priyanka Bansal	Electrical	Title: Wine Quality Prediction (Feb' 21) 1. Modeled wine quality predictor on the Wine Quality dataset from UCI Machine Learning Repository using LASSO, Ridge and Random Forest techniques. 2. Compared the influence of features estimated by the respective models with the given reference feature importance and deduced that Random Forest performs better. Title: Predict Down Syndrome in Mice (Feb' 21) 1. Performed classification on UCI Mice Protein Expression dataset usingL1 regularized Logistic Regression, Random forest and Support Vector Classification techniques. 2. Obtained best accuracy of 98.14% with Random Forest Classifier after recursive feature elimination. Title: Tanzanian Water Pumps Problem (Mar' 21) 1. Performed data cleansing, pre-processing and feature engineering on Tanzanian Water pumps dataset ,using SVM, XG Boost, Random Forest and Decision Trees classification techniques. 2. Diligent hyper-parameter tuning was performed for Decision Tree and Random Forests models. Title: Unsupervised Learning(Apr' 21) 1. Performed data transformation to handle skewed features and implemented K-means and DBSCAN clustering techniques on the given dataset. 2. Implemented and studied the effect of dimension reduction using PCA and KPCA techniques. Title: Speech Emotion Recognition Using CNN (Mar' 21 - May'21) 1. Extracted Mel Spectrograms and Mel Frequency Cepstral Coefficients from audio data and were used as features to train CNN(Convolution Neural Network) model. 2. Improved the performance of the existing model by reducing the overfitting using data augmentation.					
203070007	Nidhi Gupta	Electrical Engineering	Title : Wine Quality Prediction (Feb'21) 1. Modeled wine quality predictor on the Wine Quality dataset from UCI Machine Learning Repository using LASSO, Ridge and Support Vector Regression techniques 2. Compared the influence of features estimated by the respective models with the given reference feature importance and deduced that SVR performs better Title: Predict Down Syndrome in Mice (Feb'21) 1. Performed classification on UCI Mice Protein Expression dataset using L1 regularized Logistic Regression, Random forest and Support Vector Classification techniques 2. Obtained best accuracy of 98% with Random Forest Classifier Title: Unsupervised Learning (Apr'21) 1. Performed data transformation to handle skewed features and implemented K- means and DBSCAN clustering techniques on the given dataset 2. Implemented and studied the effect of dimension reduction using PCA and KPCA techniques Title: GAN experiments (Mar'21 - May'21) 1. Studied image translation and resolution enhancement GAN models in generating datasets synthetically 2. Generated realistic city based scenes using pre-trained model trained on Cityscapes dataset and tested the quality of generated scene using detection capability of YOLO object detector on these scenes 3. Demonstrated the feasibility of training a small CNN classifier on synthetically generated datasets for the binary classification of real-life shoes and handbags					
203350023	Rahul Singh	CTARA	TRENDS OF CARBON EMISSIONS PREDICTION USING MACHINE LEARNING ALGORITHMS ◦ Formation and analysis of data collected using seaborn & pandas and remove the outliers from the data. ◦ Implemented Decision Trees and Linear regression algorithms for the prediction and compared their accuracy.					
18B090006	Medhavi	Mathematics	Stock Market Prediction | IIT Bombay (April'21 - May'21) Prof. Amit Sethi, Course Project - Introduction to Machine Learning •Aided to build a logistic regression model using sparse dictionary vector for prediction using sentiment analysis •Implemented a Stacked LSTM(Long-Short Term Memory) model by scraping and using the historic stock price data of S&P 500 for prediction using time series forecasting methods •Integrated above models to increase sentiment analysis accuracy of 76%; achieved reduction in RMSE by 31.4% by using integrated approach over LSTM model					
193109017	Arun Jha	Mechanical	Convolution Neural Net Speed up with Sparsity| Course cum Industrial Project Advisor: Prof. Amit Sethi, IIT Bombay [Aug'19-Dec'19] Industry Mentor: Mr. Aritomo Shinozaki, Sentri AI, USA Objective: Study sparsity techniques applied to YOLO v5 model for security based surveillance applications Points: 1.Reviewed literature to identify pruning schemes including layer-level and channel-wise sparsity for CNN 2.Implemented algorithm for pruning redundant filters in CNN in a data-driven way and enabling the control of the tradeoff between network performance and its scale in pruning 3.Implemented Keras-surgeon library & Tensorflow framework for Sparsity training on LeNeT network 4. Achieved retrained model with 95-96% Pruning on Dense layer in LeNet Network 5.Optimal sparsity obtained was 30% for a YOLOv5 with a 3.57% Precision drop and 3.125% Recall drop					
203110053	Siddharth Ghosh	MEMS Department	Multi-Level Question Prediction in E-Commerce Websites 1. This project sets the scope for the QnA Extraction from Text project. Specifically, the first part of the project focuses on extracting questions and answers from the FAQ section of various websites by a generic code using the BeautifulSoup Python package. The data set so obtained can be used to train a chat 2. The second aspect of the project is training a skip-gram model of Word2vec algorithm and BERT and finally comparing the accuracies of the two models.					
203110059	Uday Shankar	MEMS	Title:- "Multi Level question prediction in e-commerce conversation" 1. Used Requests and Beautiful Soup for scraping and parsing data from the Websites. 2. Trained word2vec and BERT model on Q&A from ecom websites, extracted using python, RegEx and NLP. 3. Applied Genism & Cosine similarity package to find he similarity between the trained and user query. Title:- "Predict which pumps are functional, non- functional and which need some repairs" 1. Performed EDA followed by Feature engineering including handle imbalanced Dataset, RFECV. 2. Hyperparameter tuning and Training on XGB, Randomforest and Extratree,followed by visualization using TSNE.					
193109012	Prakhar Jain	Mechanical	Convolution Neural Net Speed up with Sparsity| Course cum Industrial Project Advisor: Prof. Amit Sethi, IIT Bombay [Aug'19-Dec'19] Industry Mentor: Mr. Aritomo Shinozaki, Sentri AI, USA Objective: Study sparsity techniques applied to YOLO v5 model for security based surveillance applications Approach: ☐Reviewed literature to identify pruning schemes including layer-level and channel-wise sparsity for CNN ☐Implemented algorithm for pruning redundant filters in CNN in a data-driven way and enabling the control of the tradeoff between network performance and its scale in pruning ☐Implemented Keras-surgeon library & Tensorflow framework for Sparsity training on LeNeT network Results: ☐Achieved retrained model with 95-96% Pruning on Dense layer in LeNet Network ☐Optimal sparsity obtained was 30% for a YOLOv5 with a 3.57% Precision drop and 3.125% Recall drop					
203101001	Aabhushan	Mechanical Department	Title-Cassava Leaf Disease Classification | prof amit sethi (Time: Jan-May2021) - Multiclass Image Classification - Trained a CNN models on leaf images datasets using Keras API with TensorFlow as backend. - Also used cv2, os and glob libraries for handling images and files. - Experimented with various models, and Hyperparameters Tuning to improve accuracy - Obtained a validation accuracy between 65-70%.					
203100022	Kshitij Kumar Choudhary	Mechancial Engineering	Title -Cassava Leaf Disease Classification - Multiclass Image Classification - Trained a CNN models on leaf images datasets using Keras API with TensorFlow as backend. - Also used cv2, os and glob libraries for handling images and files. - Experimented with various models, and Hyperparameters Tuning to improve accuracy - Obtained a validation accuracy between 65-70%.					

Roll number	your name	your department	Your verification points(please give numbers for different points and also add title and duration if you want)						
203350005	Rachna Tripathi	CTARA	1) Intrusion detection for NIR images using YOLOv5 algorithm| Mentored by Sentry AI [Jan'21-May'21] •Used OSU Thermal Pedestrian Database containing 284 NIR images & manually annotated 256 images using bounding box annotation •Trained and tested three models of YOLOv5 object detection algorithm| YOLOv5s, YOLOv5l, YOLOv5x •Analyzed models on the basis of losses and performance using TensorBoard and Utils module of Python •Achieved best performance with YOLOv5s model with of Mean Average Precision of 0.90 2)Implementation of Machine Learning Regression and Classification models| LASSO, Elastic Net, Support Vector Machine, and Random Forest [Feb'21] •Visualized & pre-processed the data| Trained & tested Wine Quality Data sets (12 features) |71% Accuracy with RF •Trained and tested Mice Protein Expression Data set (80+ features) |Best accuracy of 91% with Random Forest 3) Prediction of Functionality of Pumps from African Pump Data set using ML Classification Models [Mar'21] •Data exploration, visualization & feature engineering for 40+ features •Trained and tested Support Vector Machine, Random Forest, and K-Nearest Neighbor classification models •Hyperparameter tuning for all the three frameworks and achieved highest accuracy 0f 53.62 % with KNN model						
203100013	Het Mevada	Mechanical Engineering	1. 1. Cassava Leaf Disease Classification | Prof amit sethi - Objective was to predict image class based on Multiclass Image Classification - Trained a CNN models on leaf images datasets using Keras API with TensorFlow as backend. - Experimented with various models, and Hyperparameters Tuning to improve accuracy and to get an optimum bias-variance trade-off - Obtained a validation accuracy between 65-70%.						
203300007	Nayan Jyoti Das	BSBE	Title: Food Delivery Time Prediction 1.Developed and trained an ML model for estimating order delivery time using location and traffic dataset 2. Used technologies like Python, pandas, scikit-learn, data analysis and visualization						
203350018	Vikas Rathaur	CTARA	Framework for Sentiment Analysis of Customer Reviews |Course Project| Prof. Amit Sethi [Jan'21 – May'21] •Analyzed the Amazon customer review dataset, cleaned the data, and performed EDA to get the characteristic of data •Different ML approaches as NLP and 4+ classification models have been applied to predict the sentiments of words ••Logistic regression model shows the best accuracy as compared to Naive Bayesian, SVM, and Two layer NN models Prediction of quality of the pumps | Ministry of water, Africa | Prof. Amit Sethi [Jan' 21 - May'21] •Predicted weather the pump is functional, nonfunctional, or needs repair with 3+ classifier methods •Random forest classifier shows the best accuracy of 94.6% as compared to Decision trees, and Logistic regression [2]						
203350015	Rushikesh Dilip Patil	CTARA	1) Understanding the Application cGAN methods like Semantic Bottleneck-GAN for Scene Generation [Jan'21 – May'21] •Worked on an industry mentored project by Sentry AI to implement Semantic Bottleneck-GAN model in image synthesis •Compared and analyzed the image generation performance of SB-GAN with other SOTA models like BigGAN & ProGAN •Worked on the datasets from two domains, Cityscapes-25k and ADE Indoor containing over the images of complex scenes 2) Prediction of quality of the water pumps installed by Ministry of water in African countries [Jan'21 –May'21] •Implemented 3+ classifier models to predict if the particular pump installed is functional, non-functional or needs repairing •Achieved the best classification accuracy up to 94.3% for Random Forest Classifier compared to Decision Tree and SVM						
203350017	Milindkumar Ukey	CTARA	Title: Understanding the Application cGAN methods like Semantic Bottleneck-GAN for Scene Generation [Jan'21 – May'21] •Worked on an industry mentored project by Sentry AI to implement Semantic Bottleneck-GAN model in image synthesis •Compared and analyzed the image generation performance of SB-GAN with other SOTA models like BigGAN & ProGAN •Worked on the datasets from two domains, Cityscapes-25k and ADE Indoor containing over the images of complex scenes						
203350009	Kiran Kumar B	CTARA	Prediction of quality of the pumps | Ministry of Water, Africa Jan' 21 - May'21 Introduction to Machine Learning, Prof. Amit Sethi, IIT Bombay •Implemented 3+ classifier methods to predict if the particular installed pump is functional, non functional or needs repair. • Random forest classifier is the best classifier compared to decision trees, SVM with Gaussian kernel on the performance. •Achieved the best classification accuracy of 94.3% for Random forest classifier model. Dimensionality Reduction of Hyperspectral remote sensing data| Mentored Project Jan'21 - June'21 Introduction to Machine Learning, Prof. Amit Sethi, IIT Bombay •Guided and Mentored by Prof.Krishna Mohan from Centre for Studies in Resource Engineering (CSRE), IIT Bombay. •Performed Principle component analysis for dimensionality reduction on 220 spectral bands on a hyperspectral dataset to 30 features. •Built a Convolutional Neural network model on the resulted data to classify the image into 16 classes. •Achieved classification accuracy of 0.897 and cross entropy loss of 0.322 on the deep learning model trained						
203300003	Jatin Nareshbhai Jagani	BSBE	1. Developed and trained an ML model for estimating order delivery time using location and traffic dataset 2. Implemented machine learning techniques to predict food delivery time using Decision tree, Random Forest, XGBoost Regressor, Linear Regression models.						
203300002	Vishalkumar Keshavbhai Sav	BSBE	1. Developed and trained an ML model for estimating order delivery time using location and traffic dataset 2. Implemented machine learning techniques to predict food delivery time using Decision tree, Random Forest, XGBoost Regressor, Linear Regression models.						
203350017	Milindkumar Ukey	CTARA	Title: Prediction of Quality of the Water pumps installed by the Ministry of Water in an African Country [Jan'21-May'21] •Implemented 3 classifier models to predict if the particular pump installed is functional, non-functional or needs repairing •Achieved the best classification accuracy of 72.87% for Random Forest Classifier compared to KNN and SVM models [3]						
180070062	Sai Ganesh Vanapalli	Electrical Engineering	Snake Game Playing Agent Using ML April 2021 Course Project under the guidence of Prof. Amit Sethi •Created a ML agent which can play snake game based on explore-exploit methodology. •Lead a group of 3 members to create the above agent using Deep Q-Learning Algorithm. •This Agent clearly surpasses the agent which is created by using traditional algorithm.						
20335t005	Jayant	CTARA	Sentiment analysis using logistic regression & BiLSTM 1. Deployed logistic regression (LR) & biLSTM models to perform sentiment analysis of 5 million+ Amazon food reviews. 2. Achieved an accuracy of 75% using biLSTM model, 36 times better than LR model by optimizing the hyper-parameter tuning. 3. Pre-processed the data along with NLTK library to create more robust and reliable dataset to improve model's performance Pump repairability (ministry of africa) prediction/classification using 4 ML models 1. Predicted pump repairability in African countries by deploying decision tree, SVC, bagging and gradient boost models on 40K+ dataset. 2. Achieved an highest accuracy 85% by using SVM classifier after comparing the performaces of 4 machine learning model. 3. Deployed t-sne for dimensionality reduction to analyse the most important features and to improve model performance by 30%.						
203190004	Mohammad Sikander Khan	IEOR	Electricity Demand Prediction using LSTM and ARIMA [Jan'21-June'21] Guide : Prof. Amit Sethi | Tools : Keras, AutoArima | (Course Project -Machine Learning) 1.) Implemented LSTM and ARIMA to predict the electricity demand from house-hold consumption dataset 2.) Achieved RMSE-0.155 using ARIMA and RMSE-0.612 using LSTM model on univariate dataset 3.) ARIMA performed better for univariate datasets and short term forecasting as compared to LSTM						
203170006	Rama Khandarkar	Energy Science and Engi	Toxic Comments Classification 1. Developed a model to classify the toxic comments using an open-source dataset of 159000+ training samples 2. Implemented text cleaning pipeline for further classification using Natural Language Processing						

Roll number	your name	your department	Your verification points(please give numbers for different points and also add title and duration if you want)
180020070	Rachit Adlakha	Chemical Engineering	Sketch to Colored image generation | Course Project : Introduction to Machine Learning(Apr''21) Guide: Prof. Amit Sethi 1) Implemented Generative Adversarial Network using TensorFlow to automate colorization of greyscale images 2) Optimized hyperparameters including batch & buffer size to minimize generator and discriminator loss functions
203070023	Borse Siddhesh Suresh	EE	1)Zero-based Time Tabling - Up/Down Classification [Jan '21 – Apr '21] Course: Introduction to Machine Learning Guide: Prof. Amit Sethi, Dept. of EE, IIT Bombay •This project was mentored project under the guidance of Prof. Madhu Belur, Dept. of EE, IIT Bombay and Indian Railways. •The objective was to assign directions (up/down) to the trains and use the trained model for test data. •Performed feature reduction, feature extraction and implemented one hot encoding. •Used various ML techniques such as Decision Tree, Random Forest and AdaBoost. 2)Unsupervised Learning [Mar '21 – Apr '21] Course: Introduction to Machine Learning Guide: Prof. Amit Sethi, Dept. of EE, IIT Bombay Clustering: •Visualised and pre-processed the data from the given data set. •Trained k-means and DBSCAN and analysed performance. PCA (Principal Component Analysis): •Visualised and pre-processed the data from the given data set. •Trained PCA and plotted the variance explained versus PCA dimensions.
203020005	Mishra Digvijay Sanjaybhai	chemical engineering	Social Distance Detection in Video Using Deep Learning | Introduction to ML | Prof. Amit Sethi, IIT Bombay [Jan'21-Apr'21] 1)Loaded the COCO dataset, which contains around 0.12 million images; loaded weights which are trained using just human label 2)Implemented the regression-based model YOLO-V3 from Computer Vision (DNN module) for object detection using Python 3)Achieved 37% Mean Average Precision (MAP) on 25 FPS video Using Euclidean distance for human identification
180070031	Mahankali Gopi Krishna	Electrical Engineering	Digit Recognizer using ML models from scratch April 2021 Course Project under the guidance of Prof. Amit Sethi •Created multiple models(SVM,NN) from scratch without using scikit learn and good accuracy. •Used advanced methods like Ensembling, PCA, Bagging to improve modelling time.
203020030	Shubhi Garg	Chemical	1)Human Activity Recognition Introduction to Machine Learning Prof. Amit Sethi, IIT-B| Machine Learning | Python[Jan-Jun 2021] •Analyzed sensors data to recognize human activity to monitor health metrics •Implemented Logistic Regression, Decision Tree models from scratch to recognize human activity from sensors data •Achieved nearly equal accuracy of 96% in Logistic Regression and 82% in Decision Tree with and without using Sklearn •Analysed application of ensemble techniques and attained 94% accuracy with blending models
180020002	Aakash Kumar Singh	Chemical Engineering	Twitter Sentiment Extraction | IIT Bombay [Feb'21-May'21] Course Project under Prof. Amit Sethi, Professor | Electrical Department | IIT Bombay, India ☐Automated the detection of the word of the tweet that decides its sentiment through transfer learning & 5-fold CV ☐Achieved a Jaccard score of 70 % on the test Data set using Cross-Entropy loss function implementing Roberta architecture [4]
180020002	Aakash Kumar Singh	Chemical Engineering	Twitter Sentiment Extraction | IIT Bombay [Feb'21-May'21] Course Project under Prof. Amit Sethi, Professor | Electrical Department | IIT Bombay, India ☐Automated the detection of the word of the tweet that decides its sentiment through transfer learning & 5-fold Cross validation(CV) ☐Achieved a Jaccard score of 70 % on the test Data set using Cross-Entropy loss function implementing Roberta architecture
180020002	Aakash Kumar Singh	Chemical Engineering	Twitter Sentiment Extraction | IIT Bombay [Feb'21-May'21] Course Project under Prof. Amit Sethi, Professor | Electrical Department | IIT Bombay, India ☐Automated the detection of the word of the tweet that decides its sentiment through transfer learning & 5-fold CV ☐Achieved a Jaccard score of 70% on the test Data set using Cross-Entropy loss function & implementing Roberta architecture
20335t002	Kritika Rai Dwivedi	CTARA- Technology and	1.Intrusion detection for NIR images using YOLOv5, object detection models for vision AI methods| Mentored by Sentry AI [Jan'21-May'21] •Annotated 256 images from 284 NIR images of OSU Thermal Pedestrian database using bounding box •Trained and tested three models of YOLOv5 object detection algorithm| YOLOv5s, YOLOv5l, YOLOv5x •Analyzed models on the basis of losses and performance using TensorBoard and Utils module of Python •Achieved best performance with YOLOv5s model | Mean Average Precision 0.90 and F1 score 0.85 2.Implementation of Machine Learning Regression and Classification models| LASSO, Elastic Net, Support Vector Machine, and Random Forest [Feb'21] •Visualized & pre-processed the data |Trained & tested Wine Quality Data sets| White Wine- 4898 rows and 10+features , Red Wine- 1599 rows and 10+features •Trained and tested Mice Protein Expression Data set of 1080 rows and 80+ features |Best accuracy of 93% with Random Forest | Checked Recursive Feature Elimination to improve model performance 3.Prediction of Functionality of Pumps from African Pump Data set using Machine Learning Classification Models [Mar'21] •Data exploration, visualization & feature engineering for 40+ features of train data with 44550 rows and test data with 14850 rows •Trained and tested Support Vector Machine, Random Forest and AdaBoost classifier •Hyperparameter tuning for all the three frameworks and achieved highest accuracy of 53.69 % with RFC model resulted in 14631 pumps functional and 219 pumps needs repairing
18B030018	Prasanna Vivek Telawane	Mechanical	•Implemented KNN, SVM, and Neural Network Classifiers to classify raw audio samples into 10 music genres •Achieved 75% accuracy and 96.2% AUC-ROC score using audio features like MFCCs, Chroma Energy Normalized etc
203350019	Arshee Rizvi	CTARA	Title: Machine failure forecasting using sensor data on cloud Points: 1) Created an API in Google Apps Script editor to insert data into Google Spreadsheet. 2) Implemented Logistic Regression and Random Forest Regressor on machine sensor data 3) Evaluated these algorithm by Prediction Accuracy, Recall, MAE and Specificity
180040081	Rishabh Sharaff	Civil Engineering	MACHINE FAULT DETECTION | MACHINE LEARNING PROJECT [Mar'21-May'21] 1)- Achieved 98% accuracy in detecting health of machines under 1 second by implementing ML models: SVM, RF and ANN 2)-Computed 11 different statistical features from a single parameter; visualized the data points in time and frequency domain
180070003	Aisha Meena	Electrical	Deep Image Prior 1. Implemented the paper - Deep Image Prior in python using existing libraries for image restoration 2. Performed image denoising and image super resolution image from a degraded version of the image
203100064	Vasava Arunkumar Tarsing	Mechanical Engineering	House Price Prediction 1)Prepared and cleaned data obtained from Kaggle repository to perform linear regression. 2)Obtain Correlation matrix for features, also performed feature generation using data transformation. 3)Computed root mean square error and r2 score to evaluate performance of different models.

Roll number	your name	your department	Your verification points(please give numbers for different points and also add title and duration if you want)						
203350005	Rachna Tripathi	CTARA	1) Intrusion detection for NIR images using YOLOv5 algorithm| Mentored by Sentry AI [Jan'21-May'21] ·Used OSU Thermal Pedestrian Database containing 284 NIR images & manually annotated 256 images using bounding box annotation ·Trained and tested three models of YOLOv5 object detection algorithm| YOLOv5s, YOLOv5l, YOLOv5x ·Analyzed models on the basis of losses and performance using TensorBoard and Utils module of Python ·Achieved best performance with YOLOv5s model with Mean Average Precision of 0.90 and F1 score of 0.85 2) Implementation of Machine Learning Regression and Classification models| LASSO, Elastic Net, Support Vector Machine, and Random Forest [Feb'21] ·Visualized & pre-processed the data| Trained & tested Wine Quality Data sets (12 features,4.8k+sample points) |best accuracy of 71% with RF ·Trained and tested Mice Protein Expression Data set (80+ features, 1000+ sample points) |Best accuracy of 91% with Random Forest 3)Prediction of Functionality of Pumps from African Pump Data set using ML Classification Models [Mar'21] ·Data exploration, visualization & feature engineering for 40+ features and 44k+ sample points ·Trained and tested Support Vector Machine, Random Forest, and K-Nearest Neighbor classification models						
19307R012	Abbas Taher Barwaniwala	Electrical Engineering	Topic: ML Accelerated Full Waveform Inversion of Ground Penetrating Radar Data May 2021 Course: Introduction to Machine Learning — Instructor: Prof. Amit Sethi 1. Implemented Machine Learning Models - BiRNN with LSTM and TCN on MNIST dataset 2. Generated training and testing data using open source "gprMax" with 2000 random variations of parameters 3. Deployed ML models on GPR dataset to compute the water content of concrete, radius and depth of the rebar						
180020003	Aayam Ayan	Mechanical Engineering	1) Built a regression based sentiment analysis model achieving 76% accuracy on a dataset of 1.6 million tweets 2) Proved a correlation between social media sentiment and market movements with a Pearson coefficient of 0.32 3) Achieved a 31.4% reduction in RMSE over conventional methods using a hybrid NLP-LSTM model 4) Extracted 80k+ topic filtered and uniformly time distributed tweets using the Twitter API						
19307r004	Mohit Agarwala	EE	Employee Attrition Classification | Machine Learning [Aug - Dec'20] ◦ Objective: To predict whether an employee will leave the company or not based on 33 information points ◦ Extracted relevant and less correlated features and applied One-Hot Encoding for multi-classes features. ◦ Achieved accuracy of 88.47% by training SVM (Support vector machine) classifier on Kaggle dataset.						
203350016	Md Aquib	CTARA	Prediction of sugarcane yield | Course Project| EE769 •Analysed the raw data using seaborn and pandas, Pre processed it for applying appropriate machine learning algorithms •Implemented linear regression, SVM, and random forest algorithms for prediction and compared their accuracy						
203070007	Nidhi Gupta	Electrical Department	Topic: Wine Quality Prediction (Jan'21-Feb'21) 1. Modeled wine quality predictor on the Wine Quality dataset from UCI Machine Learning Repository using LASSO, Ridge and Support Vector Regression techniques 2. Compared the influence of features estimated by the respective models with the given reference feature importance and deduced that SVR performs better Topic: Down's Syndrome Prediction from Mice Protein Expression Levels (Jan'21-Feb'21) 1. Implemented classification algorithms such asL1 regularized Logistic Regression, Random Forest and Support Vector Classification for UCI Mice Protein Expression Dataset 2. Obtained best test accuracy of98%withRandom Forest Classifier Topic: Study and Experiments on Generative Adversarial Networks (GANs) (Mar'21-May'21) 1. Studied image translation and resolution enhancement GAN models in generating datasets synthetically. 2. Generated realistic city based scenes using pre-trained model trained on Cityscapes dataset and tested the quality of generated scene using detection capability of YOLO object detector on these scenes 3. Demonstrated the feasibility of training a small Convolutional Neural Network(CNN)classifier on synthetically generated datasets for the binary classification of real-life shoes and handbags.						
18B090004	Dipanshu Sharma	Mathematics	Bank Note Authentication End to End ML Project | Course Project (Mar '21 - Apr '21) •Developed an ML model to check the authenticity of the banknotes and created a front-end to maximize the reach of the model •Implemented model selection, hyperparameter tuning and model evaluation methods and selected a model with 99.7% accuracy •Used NumPy, pandas, matplotlib, sklearn for the backend and also used Flask, Flassger, Swagger and Pickle for the front-end API •Deployed the model on the web using Heroku Cloud-based platform and used Postman for visualization on the hosts' system						
203070005	Priyanka Bansal	Electrical	1.Wine Quality Prediction| Machine Learning Instructor: Prof. Amit Sethi, Electrical Engineering, IIT Bombay |Python ◦Modeled wine quality predictor on the Wine Quality dataset from UCI Machine Learning Repository using LASSO, Ridge and Random Forest techniques. ◦Compared the influence of features estimated by the respective models with the given reference feature importance and deduced that Random Forest performs better. 2.Down Syndrome Prediction from Mice Protein Expression Levels |Machine Learning Instructor: Prof. Amit Sethi, Electrical Engineering, IIT Bombay |Python ◦Implemented the classification model on UCI Mice Protein Expression dataset usingL1 regularized Logistic Regression, Random forest and Support Vector Classification techniques. ◦Obtained test accuracy of 98.14% with Random Forest Classifier after recursive feature elimination. 3.Tanzanian Water Pumps Problem |Machine Learning Instructor: Prof. Amit Sethi, Electrical Engineering, IIT Bombay |Python ◦Performed data cleansing, pre-processing and feature engineering on Tanzanian Water pumps dataset ,using SVM, XG Boost, Random Forest and Decision Trees classification techniques. ◦Diligent hyper-parameter tuning was performed for Decision Tree and Random Forests models. 4Unsupervised Learning |Machine Learning Instructor: Prof. Amit Sethi, Electrical Engineering, IIT Bombay |Python ◦Performed data transformation to handle skewed features and implemented K-means clustering and DBSCAN clustering techniques on the given dataset. ◦Implemented and studied the effect of dimension reduction using PCA and KPCA techniques. 5.Speech Emotion Recognition Using CNN |Machine Learning Instructor: Prof. Amit Sethi, Electrical Engineering, IIT Bombay |Python ◦Extracted Mel Spectrograms and Mel Frequency Cepstral Coefficients from audio data and used these as features to train CNN(Convolution Neural Network) model. ◦Improved the performance of the existing model by reducing the overfitting using data augmentation.						

[1] Responder updated this value.

[2] Responder updated this value.

[3] Responder updated this value.

[4] Responder updated this value.