

S.NO	Titile	Allocated	Description	Technology	Dataset	Timeline
1	Applying Machine Learning techniques to Predict the Relative Performance of CPU.	Heba	<p>With technology, increased productivity goals, faster internet, and more devices, we've created a need for speed wherever we go. We're used to getting results instantaneously and expect our devices to keep up with our requests as we multi-task our way through life. Computer processors and their clock speed are two features we most commonly associate with high-performing, fast technology.</p> <p>The marketplace for general-purpose microprocessors offers hundreds of functionally similar models, differing by traits like frequency, core count, cache size, memory bandwidth, and power consumption. Their performance depends not only on microarchitecture, but also on the nature of the workloads being executed. Given a set of intended workloads, the consumer needs both performance and price information to make rational buying decisions. Many benchmark suites have been developed to measure processor performance, and their results for large collections of CPUs are often publicly available. However, repositories of benchmark results are not always helpful when consumers need performance data for new processors or new workloads. Moreover, the aggregate scores for benchmark suites designed to cover a broad spectrum of workload types can be misleading.</p> <p>Approach: To address these problems, We will be using Regression algorithms such as Linear Regression, Decision tree, Random forest, AdaBoost, GradientBoost and XGBoost. We will train and test the data with these algorithms. From this the best model is selected and saved in pkl format. We will also be deploying our model locally using Flask.</p>	ML	https://archive.ics.uci.edu/ml/datasets/Computer+Hardware	13 July
2	ANN for Evolving Efficient Classification Patterns in Lymphography	sai charan	<p>A lymph node is an oval-shaped organ of the immune system, distributed widely throughout the body. They tend to expand in size for diverse reasons, indicating health complications that scale from trivial, to life-threatening ailments such as cancers. In the latter, the condition of lymph nodes is so significant that it is used to accurately sense the stage in Cancer progression, which decides the treatment to be adopted. Lymphography is a medical imaging technique in which a radio contrast agent is injected, and then an X-ray picture is taken to visualize structures of the lymphatic system, including lymph nodes, lymph ducts, lymphatic tissues, lymph capillaries and lymph vessels. This data is necessary to decide on whether the clinical details acquired from a Lymphograph pertains to a normal or abnormal finding. Additionally the existing state of the lymph nodes could also suggest the possibility of occurrence of cancer. Though the procedure for performing Lymphography involves potential hurdles, the data from the images facilitate accurate and precise determination of the state of the lymph nodes, ducts and capillaries. Hence proper classification and determination of credential attributes could simplify the process of disease prediction and evoke deterrent measures.</p> <p>Approach: To address these problems, We will be using ANN to classify of patient records obtained from clinical data. Feature selection is a supervised method that attempts to select a subset of the predictor features based on the information gain.</p>	DL	https://archive.ics.uci.edu/ml/datasets/Lymphography	13 July
3	Predicting the success of Marketing Campaigns using Machine Learning.	sanjana	<p>Campaign Analytics is a powerful tool that allows you to gain visibility into the behavioral attributes that make up campaign interactions, and learn how those attributes influence the consumer's path to conversion. Whether you're focused on clicks, conversions, or another KPI, Campaign Analytics shows the overlap of specific audience behaviors and the action you are tracking for success.</p> <p>Campaign analytics help you assess the effectiveness of your marketing efforts by systematically analyzing campaign success metrics. Now, marketers can justify the decisions and adjustments they make during campaigns because they have numbers to back up their actions. Metrics indicate how targets and customers are responding to marketing communications and messaging.</p> <p>Approach: To address these problems, We will be using Classification algorithms such as Logistic Regression, Decision tree, Random forest, AdaBoost, GradientBoost and XGBoost. We will train and test the data with these algorithms. From this the best model is selected and saved in pkl format. We will also be deploying our model locally using Flask.</p>	ML	https://archive.ics.uci.edu/ml/datasets/Bank+Marketing	13 July
4	Predicting Forest Cover Types with the Machine Learning Workflow	mamatha	<p>This study area includes four wilderness areas located in the Roosevelt National Forest of northern Colorado. These areas represent forests with minimal human-caused disturbances, so that existing forest cover types are more a result of ecological processes rather than forest management practices.</p> <p>Approach: To address these problems, We will be using Classification algorithms such as Logistic Regression, Decision tree, Random forest, AdaBoost, GradientBoost and XGBoost. We will train and test the data with these algorithms. From this the best model is selected and saved in pkl format. We will also be deploying our model locally using Flask.</p>	ML	https://www.kaggle.com/competitions/forest-cover-type-prediction/data	13 July

5	Machine Learning Model for Occupancy Rates and Demand in the Hospitality Industry	jansi	<p>Hotels in the modern form emerged a long time ago. In 1794 the first hotel (City Hotel) was constructed in New York City and it contained 73 rooms. In 1829, Tremont Hotel, the largest and most luxurious hotel in the world, was also created in the United States in Boston, where hotel specialists at that time considered that the interior design, building design and the equipments of a hotel the evolution of a the modern hotel industry in America and spread out to the rest of the world. Ever since, hotels have begun growing and constantly changing from all aspects and in a gradual form. Professional theoretical references suggest that hotels are considered one of the key elements that underpin tourism industry and the transport of all forms, as well as other activities.</p> <p>Not making a full use of the hotel's absorptive capacity results in the depletion of that hotel's profit, since ideal profitability rate is reached when the occupancy rates is equal to the hotel's absorptive capacity. The main problem for the hotels managements is the lack of knowledge of the key factors that control the occupancy rates.</p> <p>Approach: To address these problems, We will be using Regression algorithms such as Linear Regression, Decision tree, Random forest, AdaBoost, GradientBoost and XGBoost. We will train and test the data with these algorithms. From this the best model is selected and saved in pkl format. We will also be deploying our model locally using Flask.</p>	ML	https://archive.ics.uci.edu/ml/datasets/Occupancy+Detection+	13 July
6	Identification of Methodology Used in Real Estate Property Valuation	sai shashank	<p>Real estate prices are observed when properties change hands. The high costs and large average volume of a typical transaction (as compared to equity markets for example), lead to infrequent observations for the same asset. In between transactions, real estate professionals and investors need to rely on valuations — the most likely price to be obtained in the market, had the property been put up for sale. It's a hypothetical value, not the actual registered price. The AI wave brings unique business and ethical challenges, challenges to which real estate will not remain immune. More importantly, it indicates the increased role played by high purity timely data. Blockchain-stored transaction data becomes a valuable factor of production, one with progressively important value in an AI-driven economy. Future business models should therefore design new products and services not by obfuscating the value of this input to avoid compensating users for their data contribution, but by formally recognizing its worth and organizing the proper market needed to transact it.</p> <p>Approach: To address these problems, We will be using Regression algorithms such as Linear Regression, Decision tree, Random forest, AdaBoost, GradientBoost and XGBoost. We will train and test the data with these algorithms. From this the best model is selected and saved in pkl format. We will also be deploying our model locally using Flask.</p>	ML	https://archive.ics.uci.edu/ml/datasets/Real+estate+valuation+data+set	13 July
7	Intelligent air quality sensing system	Neha	<p>Air pollution consists of harmful gases and fine Particulate Matter (PM2.5) which affect the quality of air. This has not only become the key issues in scientific research but also turned to be an important social issues of the public's life. Therefore, many experts and scholars at different R&Ds, universities, and abroad are involved in lot of research on PM2.5 pollutant predictions. Nowadays, accurate air pollution prediction and forecast become a challenging and significant task due to increased air pollution which acts as a fundamental problem in many parts of the world. Generally, the pollution is divided into two types: natural pollution because of volcanic eruptions and forest fires resulting in emission of SO2, CO2, CO, NO2, and sulfate as air pollutants and man-made pollution because of some human activities such as burning of oils, discharges from industrial production processes, and transportation emissions that have PM2.5 as its major air pollutant which has received much attention due to their destructive effects on human health, other kinds of creatures, and environment</p> <p>Various studies testify that air pollution leads to respiratory and cardiovascular disease leading to death of animals and plants, acid rain, climate change, global warming, etc. thus making economic losses and the human life of a society difficult to survive in the world. Regarding the effects of PM2.5 investigated over the last 25 years using the comparative analysis of ML techniques</p>	ML	https://www.kaggle.com/datasets/calebreigada/us-air-quality-1980present	13 July