	ML Algorithm	Problem Statement	Project Idea	Public Dataset	Predictor Variable
Supervised	Simple Linear Regression	How does square footage affect house prices?	Predicting house prices based on square footage	https://www.kaggle.com/c/boston-housing	Continuous
	Multiple Linear Regression	How do study hours and attendance affect student performance?	Predicting student performance based on study hours and attendance	https://archive.ics.uci.edu/ml/datasets/Student+Performance	Continuous (2 or more predictors)
	Logistic Regression	Will a customer buy a product based on their characteristics?	Predicting if a customer will buy a product (yes/no)	https://www.kaggle.com/datasets/carrie1/ecommerce-data	Continuous
	Decision Trees	How do leaf measurements predict the species of a plant?	Classifying species of plants based on leaf measurements	https://archive.ics.uci.edu/ml/datasets/Iris	Continuous or Categorical
	Random Forest	How can we detect fraud in credit card transactions?	Fraud detection in credit card transactions	https://www.kaggle.com/mlg-ulb/creditcardfraud	Continuous or Categorical (multiple trees)
	Support Vector Machines	Can we classify images of cats and dogs based on their features?	Image classification (e.g.identifying cats vs dogs)	https://www.cs.toronto.edu/~kriz/cifar.html	Continuous or Categorical
	K-Nearest Neighbors	How does the proximity to stores affect house prices?	Recommender system for movies based on user ratings	https://grouplens.org/datasets/movielens/	Continuous or Categorical
	Gradient Boosting Machines	How can we predict customer churn in a subscription service based on usage patterns?	Predicting customer churn in a subscription service	https://www.kaggle.com/blastchar/telco-customer-churn	Continuous or Categorical
	Naive Bayes	How can we classify emails as spam or not spam based on their content?	Email spam detection	https://spamassassin.apache.org/publiccorpus/	Categorical
	Neural Networks	How can we recognize handwritten digits?	Handwriting recognition (e.g.MNIST dataset)	http://yann.lecun.com/exdb/mnist/	Continuous or Categorical
Unsupervised	K-Means Clustering	How can we segment customers based on their purchasing behavior?	Customer segmentation for targeted marketing	https://www.kaggle.com/vjchoudhary7/customer-segmentation-tutorial	Continuous
	Hierarchical Clustering	How can we group similar articles based on text content?	Grouping similar articles based on text content	http://qwone.com/~jason/20Newsgroups/	Continuous
	Principal Component Analysis (PCA)	How can we reduce the number of features in a dataset while preserving variance?	Reducing features in a high-dimensional dataset for visualization	https://archive.ics.uci.edu/ml/datasets/wine	Continuous
	Independent Component Analysis (ICA)	How can we separate mixed audio signals into their original sources?	Separating audio signals from a mixed recording	http://www.cis.hut.fi/projects/ica/cocktail/cocktail_en.cgi	Continuous
	Apriori Algorithm	What product combinations are commonly bought together?	Market basket analysis for retail stores	https://archive.ics.uci.edu/ml/datasets/online+retail	Categorical
	DBSCAN	How can we identify clusters of crime incidents in geographic data?	Identifying geographic clusters of crime incidents	https://www.kaggle.com/shivamb/real-time-crime	Continuous
	t-Distributed Stochastic Neighbor Embedding (t-SNE)	How can we visualize high-dimensional genetic data?	Visualizing high-dimensional genetic data	https://www.internationalgenome.org/data-portal/data-collection/phase-3	Continuous

Outcome Variable
Continuous
Continuous
Binary
Continuous or Categorical
Continuous or Categorical
Binary or Continuous
Continuous or Categorical
Continuous or Categorical
Categorical
Continuous or Categorical
Clusters
Clusters
Principal Components (Continuous)
Independent Components (Continuous)
Association Rules
Clusters
Lower-Dimensional Embedding (Continuous)