*In a****classification****problem, the outputs are categorical or discrete.*

Some of the most common types of classification problems include:

* *Classification on tabular data:* The data is available in the form of rows and columns, potentially originating from a wide variety of data sources.
* *Classification on image or sound data:* The training data consists of images or sounds whose categories are already known.
* *Classification on text data:* The training data consists of texts whose categories are already known.
* *-*2 class SVM
* The classifier chooses from multiple categories; each output can belong to one or more categories.
* Multi-class multi-label classification
* The classifier chooses from multiple categories; each output belongs to single category only.
* Multi-class single-label classification
* The classifier choose from only two categories; each output belongs to one or the other.
* Binary classification
* SUBMIT

Which of the following charts or metrics are used when evaluating results of a classification algorithm?

-ROC corve

Confusion matrix

Precision

Recall

*In a****regression****problem, the output is numerical or continuous.*

Common machine learning algorithms for regression problems include:

* Linear Regression
  + Fast training, linear model
* Decision Forest Regression
  + Accurate, fast training times
* Neural Net Regression
  + Accurate, long training times

*As the name suggests,****clustering****is the problem of organizing entities from the input data into a finite number of subsets or clusters; the goal is to maximize both intra-cluster similarity and inter-cluster differences.*

**ESCRIPTION**

**TYPE OF CLUSTERING**

Groups members based on how closely they are packed together; can learn clusters of arbitrary shape.

Density-based clustering

Builds a tree of clusters.

Hierarchical clustering

Groups members based on their distance from the center of the cluster.

Centroid-based clustering

Groups members based on the probability of a member belonging to a particular distribution.

Distribution-based clustering

SUBMIT

<https://google.qwiklabs.com/public_profiles/8f761f1b-98e9-4a64-bd60-cb29c337864>

The main aim of a recommendation system is to recommend one or more items to users of the system. Examples of an item to be recommended, might be a movie, restaurant, book, or song. In general, the user is an entity with item preferences such as a person, a group of persons, or any other type of entity you can imagine.

There are two principal approaches to recommender systems:

* The content-based approach, which makes use of features for both users and items. Users can be described by properties such as age or gender. Items can be described by properties such as the author or the manufacturer. Typical examples of content-based recommendation systems can be found on social matchmaking sites.
* The Collaborative filtering approach, which uses only identifiers of the users and the items. It is based on a matrix of ratings given by the users to the items. The main source of information about a user is the list the items they’ve rated and the similarity with other users who have rated the same items.

