1.

You are given a dataset on movie reviews with a 1,000 labeled reviews. The labels are one of five movie genres: Action, Comedy, Drama, Horror, and Sci-Fi. The dataset has roughly 200 movie reviews for each movie genre.

**Your first task** is to learn a supervised classifier to identify just the reviews for Comedy movies from the dataset. Such a task is:

Single-class classification

Two-class (Binary) classification

Multi-class classification

Multi-label classification

1  
point

2.

The dataset available for the first task is:

Balanced

Insufficient

Skewed

Unlabeled

1  
point

3.

Suppose you decide to train a support vector machine classifier for this first task. The methodology you will employ will be a:

A. One vs One classifier

B. One vs Rest classifier

C. Single binary classifier

Either A or B

Classifier cannot be trained

1  
point

4.

You are given a dataset on movie reviews with a 1,000 labeled reviews. The labels are one of five movie genres: Action, Comedy, Drama, Horror, and Sci-Fi. The dataset has roughly 200 movie reviews for each movie genre.

**Your second task** is to learn to identify all five movie genres. Such a task is:

Single-class classification

Two-class (Binary) classification

Multi-class classification

Multi-label classification

1  
point

5.

The dataset available for the second task is:

Balanced

Insufficient

Skewed

Unbalanced

1  
point

6.

Suppose you decide to train a support vector machine classifier for the second task. The methodology you will employ will be a:

A. One vs One classifier

B. One vs Rest classifier

C. Single five-class classifier

Either A or B

Classifier cannot be trained

1  
point

7.

How many binary classifiers will you need to train for the second task using the one-vs-one classification approach?

1

5

10

25

Assume you have N different classes. One vs all will train one classifier per class in total N classifiers. For class ii it will assume ii-labels as positive and the rest as negative. This often leads to imbalanced datasets meaning generic SVM might not work, but still there are some workarounds.

In one vs one you have to train a separate classifier for each different pair of labels. This leads to N\* (N−1)/2 classifiers. This is much less sensitive to the problems of imbalanced datasets but is much more computationally expensive.