Write the set (1,2,3) as a row vector and a column vector.

These vectors are _____ of each other.

What is an instance in ML?

What is a prediction in ML?

What is a target variable in ML?

We use _____ to optimize an ML classifier. This is also referred to as _____ because the predicted values of the target variables are known.

The known values of these target variables are also referred to as _____.

Each type of data we have about an instance is called a _____ and we use these to base determine the predicted value for an instance.

We are observing dice rolls. Roll R can be any number from the set (1,2,3,4,5,6). The variable R is considered a _____.

What are distance and similarity in ML?

When a website recommends you a product based off of your purchase history, this is referred to as _____.

The process of picking the recommended product or products is referred to as _____. What are the two versions of this process?

If a matrix contains mostly zeros, it is referred to as a _____.

In the problem 2^n = 16, n could be referred to as the _____ or _____.

In that same problem, 2 would be referred to as the _____.

What is the difference between supervised and unsupervised machine learning?

Define the ML terms class, classifier, and classification.

_____ is the probability that two events, A and B, both occur and is denoted by _____.

The probability of these two events occurring at the same time might be _____ if there joint occurrence is related, otherwise the two events might be _____.

A collection of large texts (or documents) is known as a _____. A unit of text within one of these documents is known as a _____.

When analyzing text, we often omit words such as 'the' or 'and' because they do not contain useful information. We refer to these words as _____.

When analyzing text, we often want to condense different tenses of a word (ex. 'run', 'runs', and 'running') into a single feature. To do this, we use a process called _____.

One problem you might encounter when training a ML classifier is when a floating point calculation returns an extremely small number. This problem is referred to as _____.

A variable in a ML model that is set before the model is trained is known as a _____.

What is the difference between a positive and negative instance in ML?

Define the ML terms true positive, true negative, false positive, and false negative.

Write the formula for accuracy using the variables TP, TN, FP, FN to represent the terms in the previous question. What is accuracy in simple terms?

In order to determine which model best represents a set of data, we must determine the model's _____.

What is the formula for sample standard deviation?

What does ddof stand for?

What is the difference between sample and population standard deviation?

The variable used to represent a model is _____, the variable used to represent a vector of labels is _____, and the variable used to represent a vector of target variables is _____.

Use the variables mentioned above to write the formula for finding residuals.

What do SSE, MSE, and RMSE stand for? What are they used for when fitting a line/curve to some data?

Y = ax + b is a function for a _____ regression, while y = ax^2 + bx + c is a function for a _____ regression.

What does OLS stand for?

What is a statistic?

When analyzing climate data, the _____ would be that all variability in the trends in the data are simply caused by natural variability, and that there is no climate change.

_____ tells us whether the probability of a statistic would be as or more extreme if the null hypothesis is true (or, in other words, that the calculation is statistically significant).

_____ tells us whether there is a correlation between two variables in a set of data.

_____ is a measure of goodness of fit that, in a linear regression, represents the percent of variability among target variables that is explained by the predictor variables.

A function or functions that describes some data is referred to as a _____.

_____ is the task of finding the best solution from a set of possible solutions.

In ML, the task of finding the smallest distance between a model and a sample is known as _____.

Observe the following functions and label them with the name of the value that they calculate:

$$P(A \mid B) = \frac{P(A \cap B)}{P(B)}$$

$$|A| = \begin{vmatrix} a & b & c \\ d & e & f \\ g & h & i \end{vmatrix} = a \begin{vmatrix} \square & \square & \square \\ \square & e & f \\ \square & h & i \end{vmatrix} - b \begin{vmatrix} \square & \square & \square \\ d & \square & f \\ g & \square & i \end{vmatrix} + c \begin{vmatrix} \square & \square & \square \\ d & e & \square \\ g & h & \square \end{vmatrix}$$

$$= a \begin{vmatrix} e & f \\ h & i \end{vmatrix} - b \begin{vmatrix} d & f \\ g & i \end{vmatrix} + c \begin{vmatrix} d & e \\ g & h \end{vmatrix}$$

$$= aei + bfg + cdh - ceg - bdi - afh.$$

$$\sum_{i=1}^{n} a_i b_i = a_1 b_1 + a_2 b_2 + \cdots + a_n b_n$$

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix}^{-1} = \frac{1}{\det \mathbf{A}} \begin{bmatrix} d & -b \\ -c & a \end{bmatrix}$$

$$\sqrt{\frac{\sum_{t=1}^{T} (\hat{y}_t - y_t)^2}{T}}.$$

$$\sqrt{\sum_i (v_i - 0_i)^2} = \sqrt{\sum_i v_i^2}$$

$$\boldsymbol{u} \cdot \boldsymbol{v} = \| \boldsymbol{u} \| \| \boldsymbol{v} \| cos\theta$$

For that last function, what do ||u|| and ||v|| represent?

For the following sets of conditions, identify the special matrix they are describing:

N = m _____

a[i][j] == 0 and i != j _____

a[i][j] == 0 and i > j _____

a[i][j] == 0 and i < j _____

(a[i][j] == 0 and i != j) and (a[i][j] == 1 and i == j) _____

a[i][j] == a[j][i] _____