Which of the following can NOT be taken as a feature for an "image" based dataset?

○ Sum of pixel values

○ N-grams

○ Number of black pixels surrounded by 4 non-black pixels

◉ The entire image taken as a 28*28 vector

Clear selection

Data normalization is a method in which we transform the features so that they have similar scales. (TRUE / FALSE)

◉ TRUE

○ FALSE

Clear selection

Which of the following are data augmentation techniques for image based datasets?

○ Shear

○ Rotation

◉ Both Shear and Rotation

○ None of the above

Clear selection

Which technique can be used to reduce overfitting in machine learning models?

○ Feature extraction

○ Feature scaling

◉ Cross-validation

○ Imputation

Clear selection

Which of the following statements about one-hot encoding is true?

◉ One-hot encoding is a technique for converting categorical data into numerical format.

○ One-hot encoding is used for reducing the dimensionality of the data.

○ One-hot encoding assigns a continuous value to each category.

○ One-hot encoding is mainly used for regression tasks.

Clear selection

Which of the following ML paradigms involves an ML agent trained to act based on a reward function or a policy?

○ Supervised learning

○ Self supervised learning

◉ Reinforcement learning

○ Unsupervised learning

Clear selection

Which Python library is commonly used for implementing K-Nearest Neighbors (K-NN) algorithm?

🔘 Scikit-learn

⚪ Pandas

⚪ Matplotlib

⚪ Tensorflow

Clear selection

Which metric we used to evaluate KNN based model in our lab?

🔘 Mean Squared Error (MSE)

⚪ R-squared (R²)

⚪ Accuracy

⚪ Root Mean Squared Error (RMSE)

Clear selection

Which of the following transformations is not considered a linear transformation?

○ Uniform scaling

○ Non-uniform scaling

⦿ Shear

○ None of the above

Clear selection

Which of the following are NOT properties of clipping in data preprocessing?

○ Clipping helps prevent data loss by scaling values within a specified range.

○ It can effectively handle outlier influence in the dataset.

⦿ Clipping guarantees the transformation of data into a normal distribution.

○ Clipping is a good method for dealing with extreme values in the data.

Clear selection

Which of the following features can be extracted from the text?

○ Unigram

○ Bigram

○ N-gram

⦿ All of the above

Clear selection

What is the purpose of a validation dataset?

○ To provide additional training data

⦿ To test the model's performance on unseen data during training

○ To measure the accuracy of the training data

○ To train the model without any supervision

Clear selection

Which of the following is not used for text featurizing ?

○ N-Grams

◉ MFCCs

○ Bag of words

○ One-Hot Encoding

Clear selection

What is the purpose of the "fit" method when using scikit-learn models?

○ To make predictions on new data

○ To visualize the data

○ To evaluate the model's performance

◉ To train the model on the given dataset

Clear selection

How many character Bigrams are possible in a text written using only english lowercase alphabets?

🔘 26*26

⚪ 26*25

⚪ 25*25

⚪ 26*27

Clear selection

The accuracy on the training set for a Nearest Neighbor (NN) based classification is

⚪ Depends on the dataset

⚪ Less than 1

🔘 Always equal to 1

⚪ 0.5

Clear selection

Which of the following are NOT disadvantages of the nearest neighbor classifier? (select the most correct option)

◉ High computational complexity during prediction

◯ Sensitivity to irrelevant features

◯ Difficulty handling datasets with high dimensionality

◯ Always requires a large amount of training data

Clear selection

Which is NOT needed for unsupervised learning?

◉ Hyperparameter tuning

◯ Labelled dataset

◯ Feature extraction

◯ All of the above

Clear selection

Which of the following is NOT a classifier?

◯ Decision Trees

◯ Support Vector Machines

◯ KNN

◉ None of the above

Clear selection

Accuracy is the ratio of the number of correctly classified samples to the total number of samples. (TRUE / FALSE)

◉ TRUE

◯ FALSE

Clear selection

Which Python library is commonly used for implementing machine learning algorithms and data preprocessing?

○ Numpy

◉ Pandas

○ Matplotlib

○ Tensorflow

Clear selection

What value of rotation angle (in degrees) was giving best accuracy for different shear angle values for the MNIST dataset?

○ 30°

○ 90°

○ 0°

◉ 45°

Clear selection

Which of these is a popular normalization technique?

○ Z-score

○ Clipping

○ Min-Max

◉ All of the above

Clear selection

What does the term "matrix rank" refer to in linear algebra?

○ The number of rows in the matrix

○ The number of columns in the matrix

○ The maximum absolute value in the matrix

◉ The maximum number of linearly independent rows or columns in the matrix

Clear selection

Which scikit-learn function is used to split a dataset into training and testing sets?

◉ train_test_split()

◯ split_dataset()

◯ test_train_split()

◯ dataset_split()

Clear selection

Back   Submit   Clear form

Never submit passwords through Google Forms.

This form was created outside of your domain. Report Abuse - Terms of Service - Privacy Policy

Google Forms