

This code performs several steps to merge, normalize, train, and evaluate a 1D Convolutional Neural Network (CNN) model. Here's an explanation of the different components and their relationships:

1. Merging Data:

- The code merges two dataframes, `data1` and `data2`, on the 'sid' column using the `pd.merge()` function.
- The merged data is saved as a CSV file at the path `'/home/mahinur/Desktop/merged_data.csv'` using the `to_csv()` method.

2. Selecting Numeric Columns:

- The code selects the numeric columns from the merged data using the `select_dtypes()` method with the `include` parameter set to `[float, int]`.
- These numeric columns are assigned to the variable `numeric_columns`.

3. Normalizing the Merged Data:

- The code uses Min-Max scaling to normalize the merged data.
- It creates an instance of the `MinMaxScaler` and applies it to the merged data using the `fit_transform()` method.
- The normalized data is stored in the `normalized_data` dataframe.

4. Saving Normalized Data:

- The normalized data is saved as a CSV file at the path `'/home/mahinur/Desktop/normalized_data.csv'` using the `to_csv()` method.

5. Loading Normalized Data:

- The normalized data is loaded from the CSV file at `'/home/mahinur/Desktop/normalized_data.csv'` using the `pd.read_csv()` function.
- The loaded data is assigned to the `normalized_data` dataframe.

6. Preparing Data for Training:

- The features (X) are extracted from the normalized data by dropping the 'output1' column using the `drop()` method. The remaining columns are converted to a NumPy array using the `values` attribute.
- The target (y) is extracted from the 'output1' column of the normalized data using the `values` attribute.

7. Reshaping Data:

- The feature data (X) is reshaped to match the expected input shape of the 1D CNN model.
- It is reshaped into a 3D array with dimensions (number of samples, number of time steps, number of features), where number of time steps is the number of columns in X (`X.shape[1]`) and number of features is 1.
- The reshaped X is stored back in the X variable.

8. Splitting Data:

- The data is split into training and testing sets using the `train_test_split()` function from scikit-learn.
- The training set contains 80% of the data, while the testing set contains the remaining 20%.
- The training and testing sets are assigned to `X_train`, `X_test`, `y_train`, and `y_test` variables.

9. Creating the 1D CNN Model:

- A sequential model is created using the `Sequential()` class from Keras.
- The model consists of several layers:
 - A 1D convolutional layer (`Conv1D`) with 32 filters, a kernel size of 3, and 'relu' activation.
 - A max pooling layer (`MaxPooling1D`) with a pool size of 2.
 - A flattening layer (`Flatten`) to convert the 2D output of the convolutional layer to a 1D vector.
- Two fully connected (`Dense`) layers with 64 and 1 neuron(s) respectively. The first layer uses 'relu' activation, and the last layer uses 'sigmoid' activation.

10. Compiling the Model:

- The model is compiled with the 'binary_crossentropy' loss function, 'adam' optimizer, and 'accuracy' metric.

11. Training the Model:

- The model is trained using the `fit()` method on the training data.
- The training is performed for 10 epochs with a batch size of 16.
- The validation data (`X_test` and `y_test`) is provided to monitor the performance during training.

12. Evaluating the Model:

- The trained model is evaluated using the testing data (`X_test` and `y_test`) using the `evaluate()` method.
- The calculated loss and accuracy are stored in the variables 'loss' and 'accuracy'.
- The test loss and test accuracy are printed to the console.