dl-practical-2

February 17, 2024

[1]:
```
import pandas as pd
import numpy as np
import seaborn as sns
import matplotlib.pyplot as plt
```

[3]:
```
df = pd.read_csv("letter-recognition.data")
df
```

[3]:

	T	2	8	3	5	1	8.1	13	0	6	6.1	10	8.2	0.1	8.3	0.2	8.4
0	I	5	12	3	7	2	10	5	5	4	13	3	9	2	8	4	10
1	D	4	11	6	8	6	10	6	2	6	10	3	7	3	7	3	9
2	N	7	11	6	6	3	5	9	4	6	4	4	10	6	10	2	8
3	G	2	1	3	1	1	8	6	6	6	6	5	9	1	7	5	10
4	S	4	11	5	8	3	8	8	6	9	5	6	6	0	8	9	7
…	..	..	..	..	..	..	…	..	..	..	…	..	…	…	…	…	…
19994	D	2	2	3	3	2	7	7	7	6	6	6	4	2	8	3	7
19995	C	7	10	8	8	4	4	8	6	9	12	9	13	2	9	3	7
19996	T	6	9	6	7	5	6	11	3	7	11	9	5	2	12	2	4
19997	S	2	3	4	2	1	8	7	2	6	10	6	8	1	9	5	8
19998	A	4	9	6	6	2	9	5	3	1	8	1	8	2	7	2	8

[19999 rows x 17 columns]

[4]:
```
df.shape
```

[4]:
```
(19999, 17)
```

[5]:
```
df.sample(30)
```

[5]:

	T	2	8	3	5	1	8.1	13	0	6	6.1	10	8.2	0.1	8.3	0.2	8.4
17273	Y	3	5	4	4	2	4	10	2	7	11	11	6	1	11	2	5
2022	K	5	10	6	7	2	4	7	9	2	7	6	12	3	8	3	11
19002	L	2	6	2	4	1	0	1	5	6	0	0	6	0	8	0	8
10078	J	2	4	3	7	1	12	3	9	4	13	5	12	1	6	0	8
4561	R	5	11	7	8	6	7	7	4	8	7	6	6	3	8	5	8
922	F	2	1	3	2	1	5	11	4	5	10	9	5	1	9	3	6
2834	Z	3	6	6	4	3	7	7	2	10	12	6	8	1	8	6	8

3362	N	4	6	6	5	5	7	7	4	4	6	5	8	6	8	4	6
6389	Y	6	11	6	8	3	3	10	3	7	12	12	7	1	11	3	5
16363	Y	1	0	2	0	0	7	10	1	3	7	12	8	1	11	0	8
15330	E	5	11	5	8	6	3	7	5	9	7	7	14	0	8	6	8
14443	R	2	4	4	3	2	8	7	4	5	8	5	7	2	7	4	11
16959	M	2	1	2	1	1	7	6	10	0	7	8	8	6	6	0	8
18184	C	4	10	5	8	3	5	8	7	8	7	8	15	1	8	4	9
7066	E	4	8	6	6	7	6	7	3	6	6	7	11	3	10	8	7
18355	Z	4	7	6	5	4	7	9	2	9	11	8	5	1	8	6	5
3527	E	6	9	8	7	5	9	7	2	9	11	4	9	2	8	5	10
12634	W	6	5	8	8	4	5	8	5	2	7	8	8	9	9	0	8
175	E	3	4	5	2	2	7	7	2	8	11	7	9	2	8	4	8
13112	Z	6	9	8	7	5	8	6	2	9	12	5	10	3	7	7	9
7154	E	4	6	6	4	4	6	8	2	9	11	7	10	2	9	4	8
15703	E	5	9	7	6	5	7	8	1	8	11	6	9	2	8	4	9
2542	H	4	8	6	6	7	8	7	4	3	6	6	7	7	8	7	6
18670	B	3	5	4	4	3	8	7	6	6	6	6	5	2	8	7	8
3207	E	3	7	4	5	3	3	7	5	9	7	6	14	0	8	6	9
15500	I	5	8	6	9	6	8	9	5	6	6	6	7	3	9	9	8
13951	S	5	10	6	7	4	7	8	3	7	10	7	7	3	7	5	6
1950	N	4	10	4	7	3	7	7	14	2	4	6	8	6	8	0	8
15230	O	3	1	4	3	2	7	7	7	5	7	6	8	2	8	3	8
6941	D	4	7	4	5	2	6	7	10	9	7	7	6	3	8	4	8

1 Converting categorical data to numerical data

[6]:
```
from sklearn.preprocessing import LabelEncoder
```

[7]:
```
lab = LabelEncoder()
```

[8]:
```
df['T'] = lab.fit_transform(df['T'])
```

[9]:
```
df
```

[9]:

	T	2	8	3	5	1	8.1	13	0	6	6.1	10	8.2	0.1	8.3	0.2	8.4
0	8	5	12	3	7	2	10	5	5	4	13	3	9	2	8	4	10
1	3	4	11	6	8	6	10	6	2	6	10	3	7	3	7	3	9
2	13	7	11	6	6	3	5	9	4	6	4	4	10	6	10	2	8
3	6	2	1	3	1	1	8	6	6	6	6	5	9	1	7	5	10
4	18	4	11	5	8	3	8	8	6	9	5	6	6	0	8	9	7
…	..	..	..	..	..	..	…	..	..	..	…	..	…	…	…	…	…
19994	3	2	2	3	3	2	7	7	7	6	6	6	4	2	8	3	7
19995	2	7	10	8	8	4	4	8	6	9	12	9	13	2	9	3	7
19996	19	6	9	6	7	5	6	11	3	7	11	9	5	2	12	2	4
19997	18	2	3	4	2	1	8	7	2	6	10	6	8	1	9	5	8
19998	0	4	9	6	6	2	9	5	3	1	8	1	8	2	7	2	8

```
[19999 rows x 17 columns]
```

[10]:
```
x = df.drop(["T"], axis=1)
```

[11]:
```
x
```

[11]:
```
        2   8  3  5  1  8.1  13  0  6  6.1  10  8.2  0.1  8.3  0.2  8.4
0       5  12  3  7  2   10   5  5  4   13   3    9    2    8    4   10
1       4  11  6  8  6   10   6  2  6   10   3    7    3    7    3    9
2       7  11  6  6  3    5   9  4  6    4   4   10    6   10    2    8
3       2   1  3  1  1    8   6  6  6    6   5    9    1    7    5   10
4       4  11  5  8  3    8   8  6  9    5   6    6    0    8    9    7
…      ..  .. .. .. ..  …   .. .. ..  …   ..  …   …   …   …   …
19994   2   2  3  3  2    7   7  7  6    6   6    4    2    8    3    7
19995   7  10  8  8  4    4   8  6  9   12   9   13    2    9    3    7
19996   6   9  6  7  5    6  11  3  7   11   9    5    2   12    2    4
19997   2   3  4  2  1    8   7  2  6   10   6    8    1    9    5    8
19998   4   9  6  6  2    9   5  3  1    8   1    8    2    7    2    8

[19999 rows x 16 columns]
```

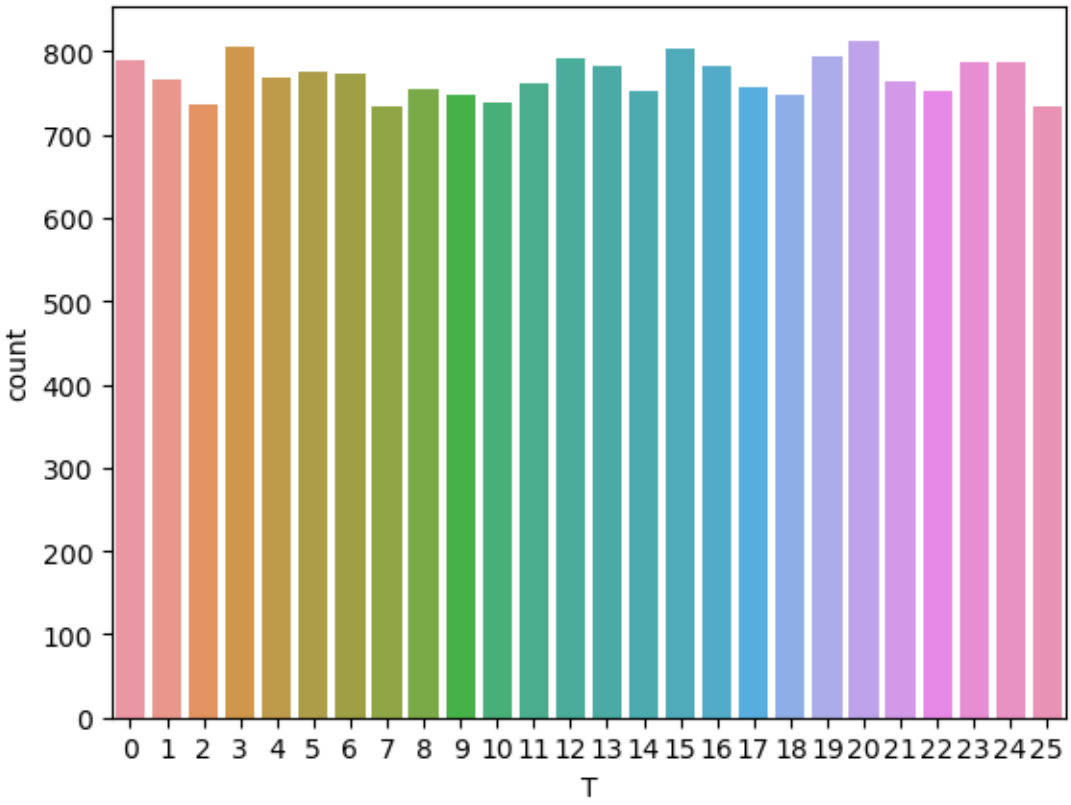
[12]:
```
y= df["T"]
```

[13]:
```
y
```

[13]:
```
0         8
1         3
2        13
3         6
4        18
         ..
19994     3
19995     2
19996    19
19997    18
19998     0
Name: T, Length: 19999, dtype: int32
```

[14]:
```
sns.countplot(x=y)
```

[14]:
```
<Axes: xlabel='T', ylabel='count'>
```



[15]:
```
y.value_counts()
```

[15]:
```
T
20    813
3     805
15    803
19    795
12    792
0     789
23    787
24    786
16    783
13    783
5     775
6     773
4     768
1     766
21    764
11    761
17    758
8     755
```

```
14    753
22    752
18    748
9     747
10    739
2     736
7     734
25    734
Name: count, dtype: int64
```

[16]:
```
from sklearn.preprocessing import MinMaxScaler
sc=MinMaxScaler()
```

[17]:
```
x_scale = sc.fit_transform(x)
x_scale
```

[17]:
```
array([[0.33333333, 0.8       , 0.2       , …, 0.53333333, 0.26666667,
        0.66666667],
       [0.26666667, 0.73333333, 0.4       , …, 0.46666667, 0.2       ,
        0.6       ],
       [0.46666667, 0.73333333, 0.4       , …, 0.66666667, 0.13333333,
        0.53333333],
       …,
       [0.4       , 0.6       , 0.4       , …, 0.8       , 0.13333333,
        0.26666667],
       [0.13333333, 0.2       , 0.26666667, …, 0.6       , 0.33333333,
        0.53333333],
       [0.26666667, 0.6       , 0.4       , …, 0.46666667, 0.13333333,
        0.53333333]])
```

[18]:
```
from sklearn.model_selection import train_test_split
x_train,x_test,y_train,y_test= train_test_split(x_scale, y,random_state=42,␣
 ↪test_size=0.5)
```

[19]:
```
x_train.shape
```

[19]:
```
(9999, 16)
```

[20]:
```
x_test.shape
```

[20]:
```
(10000, 16)
```

[21]:
```
y_train.shape
```

[21]:
```
(9999,)
```

[22]:
```
y_test.shape
```

[22]:
```
(10000,)
```

[23]:
```
from sklearn.neighbors import KNeighborsClassifier
knn=KNeighborsClassifier()
```

[24]:
```
knn.fit(x_train, y_train)
```

[24]:
```
KNeighborsClassifier()
```

[25]:
```
y_train= knn.predict(x_test)
```

[26]:
```
y_pred=knn.predict(x_test)
y_pred
```

[26]:
```
array([23, 17, 15, …,  9, 16, 12])
```

[27]:
```
from sklearn.metrics import accuracy_score,classification_report
```

[28]:
```
accuracy_score(y_test,y_pred)
```

[28]:
```
0.9318
```

[29]:
```
print(classification_report(y_test, y_train))
```

```
              precision    recall  f1-score   support

           0       0.96      0.98      0.97       389
           1       0.81      0.94      0.87       378
           2       0.92      0.96      0.94       344
           3       0.85      0.95      0.90       405
           4       0.91      0.89      0.90       402
           5       0.89      0.90      0.90       389
           6       0.91      0.93      0.92       381
           7       0.85      0.83      0.84       374
           8       0.93      0.97      0.95       349
           9       0.97      0.92      0.95       384
          10       0.89      0.87      0.88       359
          11       0.99      0.96      0.97       380
          12       0.98      0.95      0.97       386
          13       0.97      0.93      0.95       392
          14       0.88      0.91      0.90       398
          15       0.94      0.91      0.92       405
          16       0.95      0.92      0.94       387
          17       0.90      0.93      0.92       378
          18       0.98      0.92      0.95       383
          19       0.96      0.91      0.93       385
          20       0.96      0.97      0.97       401
```

```
           21       0.97      0.95      0.96       404
           22       0.97      0.95      0.96       396
           23       0.97      0.93      0.95       390
           24       0.94      0.97      0.95       380
           25       0.97      0.97      0.97       381

     accuracy                           0.93     10000
    macro avg       0.93      0.93      0.93     10000
 weighted avg       0.93      0.93      0.93     10000
```

[30]:
```
pip install tensorflow
```

```
Requirement already satisfied: tensorflow in
c:\users\harsh\documents\anaconda\lib\site-packages (2.15.0)
Requirement already satisfied: tensorflow-intel==2.15.0 in
c:\users\harsh\documents\anaconda\lib\site-packages (from tensorflow) (2.15.0)
Requirement already satisfied: absl-py>=1.0.0 in
c:\users\harsh\documents\anaconda\lib\site-packages (from tensorflow-
intel==2.15.0->tensorflow) (2.1.0)
Requirement already satisfied: astunparse>=1.6.0 in
c:\users\harsh\documents\anaconda\lib\site-packages (from tensorflow-
intel==2.15.0->tensorflow) (1.6.3)
Requirement already satisfied: flatbuffers>=23.5.26 in
c:\users\harsh\documents\anaconda\lib\site-packages (from tensorflow-
intel==2.15.0->tensorflow) (23.5.26)
Requirement already satisfied: gast!=0.5.0,!=0.5.1,!=0.5.2,>=0.2.1 in
c:\users\harsh\documents\anaconda\lib\site-packages (from tensorflow-
intel==2.15.0->tensorflow) (0.5.4)
Requirement already satisfied: google-pasta>=0.1.1 in
c:\users\harsh\documents\anaconda\lib\site-packages (from tensorflow-
intel==2.15.0->tensorflow) (0.2.0)
Requirement already satisfied: h5py>=2.9.0 in
c:\users\harsh\documents\anaconda\lib\site-packages (from tensorflow-
intel==2.15.0->tensorflow) (3.9.0)
Requirement already satisfied: libclang>=13.0.0 in
c:\users\harsh\documents\anaconda\lib\site-packages (from tensorflow-
intel==2.15.0->tensorflow) (16.0.6)
Requirement already satisfied: ml-dtypes~=0.2.0 in
c:\users\harsh\documents\anaconda\lib\site-packages (from tensorflow-
intel==2.15.0->tensorflow) (0.2.0)
Requirement already satisfied: numpy<2.0.0,>=1.23.5 in
c:\users\harsh\documents\anaconda\lib\site-packages (from tensorflow-
intel==2.15.0->tensorflow) (1.24.3)
Requirement already satisfied: opt-einsum>=2.3.2 in
c:\users\harsh\documents\anaconda\lib\site-packages (from tensorflow-
intel==2.15.0->tensorflow) (3.3.0)
Requirement already satisfied: packaging in
```

```
c:\users\harsh\documents\anaconda\lib\site-packages (from tensorflow-
intel==2.15.0->tensorflow) (23.1)
Requirement already satisfied:
protobuf!=4.21.0,!=4.21.1,!=4.21.2,!=4.21.3,!=4.21.4,!=4.21.5,<5.0.0dev,>=3.20.3
in c:\users\harsh\documents\anaconda\lib\site-packages (from tensorflow-
intel==2.15.0->tensorflow) (4.25.2)
Requirement already satisfied: setuptools in
c:\users\harsh\documents\anaconda\lib\site-packages (from tensorflow-
intel==2.15.0->tensorflow) (68.0.0)
Requirement already satisfied: six>=1.12.0 in
c:\users\harsh\documents\anaconda\lib\site-packages (from tensorflow-
intel==2.15.0->tensorflow) (1.16.0)
Requirement already satisfied: termcolor>=1.1.0 in
c:\users\harsh\documents\anaconda\lib\site-packages (from tensorflow-
intel==2.15.0->tensorflow) (2.4.0)
Requirement already satisfied: typing-extensions>=3.6.6 in
c:\users\harsh\documents\anaconda\lib\site-packages (from tensorflow-
intel==2.15.0->tensorflow) (4.7.1)
Requirement already satisfied: wrapt<1.15,>=1.11.0 in
c:\users\harsh\documents\anaconda\lib\site-packages (from tensorflow-
intel==2.15.0->tensorflow) (1.14.1)
Requirement already satisfied: tensorflow-io-gcs-filesystem>=0.23.1 in
c:\users\harsh\documents\anaconda\lib\site-packages (from tensorflow-
intel==2.15.0->tensorflow) (0.31.0)
Requirement already satisfied: grpcio<2.0,>=1.24.3 in
c:\users\harsh\documents\anaconda\lib\site-packages (from tensorflow-
intel==2.15.0->tensorflow) (1.60.1)
Requirement already satisfied: tensorboard<2.16,>=2.15 in
c:\users\harsh\documents\anaconda\lib\site-packages (from tensorflow-
intel==2.15.0->tensorflow) (2.15.2)
Requirement already satisfied: tensorflow-estimator<2.16,>=2.15.0 in
c:\users\harsh\documents\anaconda\lib\site-packages (from tensorflow-
intel==2.15.0->tensorflow) (2.15.0)
Requirement already satisfied: keras<2.16,>=2.15.0 in
c:\users\harsh\documents\anaconda\lib\site-packages (from tensorflow-
intel==2.15.0->tensorflow) (2.15.0)
Requirement already satisfied: wheel<1.0,>=0.23.0 in
c:\users\harsh\documents\anaconda\lib\site-packages (from
astunparse>=1.6.0->tensorflow-intel==2.15.0->tensorflow) (0.38.4)
Requirement already satisfied: google-auth<3,>=1.6.3 in
c:\users\harsh\documents\anaconda\lib\site-packages (from
tensorboard<2.16,>=2.15->tensorflow-intel==2.15.0->tensorflow) (2.27.0)
Requirement already satisfied: google-auth-oauthlib<2,>=0.5 in
c:\users\harsh\documents\anaconda\lib\site-packages (from
tensorboard<2.16,>=2.15->tensorflow-intel==2.15.0->tensorflow) (1.2.0)
Requirement already satisfied: markdown>=2.6.8 in
c:\users\harsh\documents\anaconda\lib\site-packages (from
tensorboard<2.16,>=2.15->tensorflow-intel==2.15.0->tensorflow) (3.4.1)
```

```
Requirement already satisfied: requests<3,>=2.21.0 in
c:\users\harsh\documents\anaconda\lib\site-packages (from
tensorboard<2.16,>=2.15->tensorflow-intel==2.15.0->tensorflow) (2.31.0)
Requirement already satisfied: tensorboard-data-server<0.8.0,>=0.7.0 in
c:\users\harsh\documents\anaconda\lib\site-packages (from
tensorboard<2.16,>=2.15->tensorflow-intel==2.15.0->tensorflow) (0.7.2)
Requirement already satisfied: werkzeug>=1.0.1 in
c:\users\harsh\documents\anaconda\lib\site-packages (from
tensorboard<2.16,>=2.15->tensorflow-intel==2.15.0->tensorflow) (2.2.3)
Requirement already satisfied: cachetools<6.0,>=2.0.0 in
c:\users\harsh\documents\anaconda\lib\site-packages (from google-
auth<3,>=1.6.3->tensorboard<2.16,>=2.15->tensorflow-intel==2.15.0->tensorflow)
(5.3.2)
Requirement already satisfied: pyasn1-modules>=0.2.1 in
c:\users\harsh\documents\anaconda\lib\site-packages (from google-
auth<3,>=1.6.3->tensorboard<2.16,>=2.15->tensorflow-intel==2.15.0->tensorflow)
(0.2.8)
Requirement already satisfied: rsa<5,>=3.1.4 in
c:\users\harsh\documents\anaconda\lib\site-packages (from google-
auth<3,>=1.6.3->tensorboard<2.16,>=2.15->tensorflow-intel==2.15.0->tensorflow)
(4.9)
Requirement already satisfied: requests-oauthlib>=0.7.0 in
c:\users\harsh\documents\anaconda\lib\site-packages (from google-auth-
oauthlib<2,>=0.5->tensorboard<2.16,>=2.15->tensorflow-intel==2.15.0->tensorflow)
(1.3.1)
Requirement already satisfied: charset-normalizer<4,>=2 in
c:\users\harsh\documents\anaconda\lib\site-packages (from
requests<3,>=2.21.0->tensorboard<2.16,>=2.15->tensorflow-
intel==2.15.0->tensorflow) (2.0.4)
Requirement already satisfied: idna<4,>=2.5 in
c:\users\harsh\documents\anaconda\lib\site-packages (from
requests<3,>=2.21.0->tensorboard<2.16,>=2.15->tensorflow-
intel==2.15.0->tensorflow) (3.4)
Requirement already satisfied: urllib3<3,>=1.21.1 in
c:\users\harsh\documents\anaconda\lib\site-packages (from
requests<3,>=2.21.0->tensorboard<2.16,>=2.15->tensorflow-
intel==2.15.0->tensorflow) (1.26.16)
Requirement already satisfied: certifi>=2017.4.17 in
c:\users\harsh\documents\anaconda\lib\site-packages (from
requests<3,>=2.21.0->tensorboard<2.16,>=2.15->tensorflow-
intel==2.15.0->tensorflow) (2023.7.22)
Requirement already satisfied: MarkupSafe>=2.1.1 in
c:\users\harsh\documents\anaconda\lib\site-packages (from
werkzeug>=1.0.1->tensorboard<2.16,>=2.15->tensorflow-intel==2.15.0->tensorflow)
(2.1.1)
Requirement already satisfied: pyasn1<0.5.0,>=0.4.6 in
c:\users\harsh\documents\anaconda\lib\site-packages (from
pyasn1-modules>=0.2.1->google-
```

```
auth<3,>=1.6.3->tensorboard<2.16,>=2.15->tensorflow-intel==2.15.0->tensorflow)
(0.4.8)
Requirement already satisfied: oauthlib>=3.0.0 in
c:\users\harsh\documents\anaconda\lib\site-packages (from requests-
oauthlib>=0.7.0->google-auth-
oauthlib<2,>=0.5->tensorboard<2.16,>=2.15->tensorflow-intel==2.15.0->tensorflow)
(3.2.2)
Note: you may need to restart the kernel to use updated packages.
```

[31]:
```
import tensorflow as tf
```

```
WARNING:tensorflow:From C:\Users\harsh\Documents\anaconda\Lib\site-
packages\keras\src\losses.py:2976: The name
tf.losses.sparse_softmax_cross_entropy is deprecated. Please use
tf.compat.v1.losses.sparse_softmax_cross_entropy instead.
```

[42]:
```
model = tf.keras.Sequential([
    tf.keras.layers.Dense(64, activation = "relu", input_shape = (x_train.
 ↪shape[1],)),
    tf.keras.layers.Dense(32, activation = "relu"),
    tf.keras.layers.Dense(32, activation = "relu"),
    tf.keras.layers.Dense(16, activation = "relu"),
    tf.keras.layers.Dense(16, activation = "relu"),
    tf.keras.layers.Dense(16, activation = "relu"),
    tf.keras.layers.Dense(16, activation = "relu"),
    tf.keras.layers.Dense(4, activation = "relu"),
    tf.keras.layers.Dense(4, activation = "relu"),
    tf.keras.layers.Dense(4, activation = "relu"),
    tf.keras.layers.Dense(4, activation = "relu"),
    tf.keras.layers.Dense(1)
])

#compile the model
# model.compile(optimize = "adam", loss = "mean_squared_error")
model.compile(
        optimizer="adam", loss="mean_squared_error",
    )

#training the model
model.fit(x_train, y_train, epochs = 100, batch_size = 32, validation_data =␣
 ↪(x_test, y_test))
```

```
---------------------------------------------------------------------------
ValueError                                Traceback (most recent call last)
Cell In[42], line 23
     18 model.compile(
```

```
     19         optimizer="adam", loss="mean_squared_error",
     20     )
     22 #training the model
---> 23 model.fit(x_train, y_train, epochs = 100, batch_size = 32,␣
 ↪validation_data = (x_test, y_test))

File ~\Documents\anaconda\Lib\site-packages\keras\src\utils\traceback_utils.py:
 ↪70, in filter_traceback.<locals>.error_handler(*args, **kwargs)
     67     filtered_tb = _process_traceback_frames(e.__traceback__)
     68     # To get the full stack trace, call:
     69     # `tf.debugging.disable_traceback_filtering()`
---> 70     raise e.with_traceback(filtered_tb) from None
     71 finally:
     72     del filtered_tb

File ~\Documents\anaconda\Lib\site-packages\keras\src\engine\data_adapter.py:
 ↪1960, in _check_data_cardinality(data)
   1953     msg += "  {} sizes: {}\n".format(
   1954         label,
   1955         ", ".join(
   1956             str(i.shape[0]) for i in tf.nest.flatten(single_data)
   1957         ),
   1958     )
   1959 msg += "Make sure all arrays contain the same number of samples."
-> 1960 raise ValueError(msg)

ValueError: Data cardinality is ambiguous:
  x sizes: 9999
  y sizes: 10000
Make sure all arrays contain the same number of samples.
```

[43]:
```
#evaluate the model
loss = model.evaluate(x_test, y_test)
print(f"MSE of test data = {loss}")
```

```
WARNING:tensorflow:From C:\Users\harsh\Documents\anaconda\Lib\site-
packages\keras\src\utils\tf_utils.py:492: The name tf.ragged.RaggedTensorValue
is deprecated. Please use tf.compat.v1.ragged.RaggedTensorValue instead.

313/313 [==============================] - 3s 5ms/step - loss: 214.5534
MSE of test data = 214.55340576171875
```

[ ]: