```
#Alexander Buchholz
#CPE695
#Final Project, MovieNNregression
#This program uses the Keras Python library built on Tensorflow and
IMDB movie data to train a Neural Network to predict
#movie Return On Investment (ROI) given genre, director, ratings, and
actors. In this case, a regression algorithm is used to
#estimate the %ROI, with mean and standard deviation as metrics of
success.
#The code written here was written by myself, but also used a
combination of the methodologies shown in the following
#Machine Learning Master tutorials:
#https://machinelearningmastery.com/regression-tutorial-keras-deep-
learning-library-python/
#https://machinelearningmastery.com/multi-class-classification-
tutorial-keras-deep-learning-library/
from pandas import read csv
from keras.models import Sequential
from keras.layers import Dense
from keras.wrappers.scikit learn import KerasRegressor
from keras.utils import np utils
from sklearn.model selection import cross val score
from sklearn.model selection import KFold
from sklearn.preprocessing import StandardScaler
from sklearn.pipeline import Pipeline
from sklearn.preprocessing import LabelEncoder
import numpy as np
import pandas as pd
# load CSV dataset
dataframe = read csv("CPE 695 Group Project Clean Data - 2760
Movies.csv", delim whitespace=False, header=None)
dataset = dataframe.values
# split into input (X) and output (Y) variables. Ignore the first
element (the label) for each column.
genre = dataset[1:,4]
rating = dataset[1:,5].astype(float)
director = dataset[1:,6]
actor1 = dataset[1:,7]
actor2 = dataset[1:,8]
#Convert to np.array and remove commas in the ROI values
Y = np.array(dataset[1:,12])
Y = np.array([x.replace(',', '') for x in Y])
Y = Y.astype(float)
```

```
#Encode genre, director, actor1, and actor2 as integer values so that
they can be correctly processed
encoder genre = LabelEncoder()
encoder director = LabelEncoder()
encoder actor1 = LabelEncoder()
encoder actor2 = LabelEncoder()
encoder genre.fit(genre)
encoder director.fit(director)
encoder actor1.fit(actor1)
encoder actor2.fit(actor2)
encoded genre = encoder genre.transform(genre)
encoded director = encoder director.transform(director)
encoded actor1 = encoder actor1.transform(actor1)
encoded actor2 = encoder actor2.transform(actor2)
#Combine the newly encoded columns with the rating to form the input
matrix.
X = np.stack((encoded genre, rating, encoded director, encoded actor1,
encoded actor2),axis=-1)
# define base model
def baseline model():
    # create model
    model = Sequential()
    model.add(Dense(10, input dim=5, kernel initializer='normal',
activation='relu'))
    model.add(Dense(1, kernel initializer='normal'))
    # Compile model
    model.compile(loss='mean squared logarithmic error',
optimizer='adam')
    return model
# evaluate model with standardized dataset
build fn=baseline model
estimators = []
estimators.append(('standardize', StandardScaler()))
estimators.append(('mlp', KerasRegressor(build fn, epochs=50,
batch size=5, verbose=1)))
pipeline = Pipeline(estimators)
kfold = KFold(n splits=10)
results = cross val score(pipeline, X, Y, cv=kfold)
#since mean squared logarithmic error was used as the loss function,
exponentiate results to get stromy.
results = np.exp(results)
#Evaluate and print mean and standard deviation of the results
print("Standardized: %.2f (%.2f) MSE" % (results.mean(),
results.std()))
```

```
Epoch 1/50
25.2793
Epoch 2/50
14.9188
Epoch 3/50
10.6877
Epoch 4/50
8.3299
Epoch 5/50
6.7670
Epoch 6/50
5.6376
Epoch 7/50
4.7792
Epoch 8/50
4.1075
Epoch 9/50
3.5718
Epoch 10/50
3.1385
Epoch 11/50
2.7874
Epoch 12/50
2.5025
Epoch 13/50
2.2725
Epoch 14/50
2.0887
Epoch 15/50
1.9449
Epoch 16/50
1.8336
Epoch 17/50
```

```
1.7493
Epoch 18/50
1.6874
Epoch 19/50
1.6428
Epoch 20/50
1.6121
Epoch 21/50
1.5916
Epoch 22/50
1.5783
Epoch 23/50
1.5700
Epoch 24/50
1.5647
Epoch 25/50
1.5615
Epoch 26/50
1.5596
Epoch 27/50
1.5584
Epoch 28/50
1.5577
Epoch 29/50
1.5573
Epoch 30/50
1.5569
Epoch 31/50
1.5567
Epoch 32/50
1.5566
Epoch 33/50
1.5566
Epoch 34/50
```

```
1.5564
Epoch 35/50
1.5563
Epoch 36/50
1.5563
Epoch 37/50
1.5563
Epoch 38/50
1.5561
Epoch 39/50
1.5561
Epoch 40/50
1.5561
Epoch 41/50
1.5560
Epoch 42/50
1.5560
Epoch 43/50
1.5559
Epoch 44/50
1.5559
Epoch 45/50
1.5558
Epoch 46/50
1.5558
Epoch 47/50
1.5558
Epoch 48/50
1.5557
Epoch 49/50
1.5557
Epoch 50/50
1.5557
```

```
276/276 [============ ] - 0s 159us/step
Epoch 1/50
25.2469
Epoch 2/50
14.4726
Epoch 3/50
10.2564
Epoch 4/50
7.9481
Epoch 5/50
6.4298
Epoch 6/50
Epoch 7/50
4.5079
Epoch 8/50
3.8595
Epoch 9/50
3.3432
Epoch 10/50
2.9277
Epoch 11/50
2.5930
Epoch 12/50
2.3231
Epoch 13/50
2.1077
Epoch 14/50
1.9367
Epoch 15/50
1.8041
Epoch 16/50
1.7025
Epoch 17/50
```

```
1.6263
Epoch 18/50
1.5708
Epoch 19/50
1.5322
Epoch 20/50
1.5062
Epoch 21/50
1.4896
Epoch 22/50
1.4790
Epoch 23/50
1.4723
Epoch 24/50
1.4685
Epoch 25/50
1.4660
Epoch 26/50
1.4647
Epoch 27/50
1.4638
Epoch 28/50
1.4633
Epoch 29/50
1.4630
Epoch 30/50
1.4630
Epoch 31/50
1.4627
Epoch 32/50
1.4626
Epoch 33/50
1.4624
```

```
Epoch 34/50
1.4623
Epoch 35/50
1.4622
Epoch 36/50
1.4622
Epoch 37/50
1.4622
Epoch 38/50
1.4622
Epoch 39/50
1.4621
Epoch 40/50
1.4620
Epoch 41/50
1.4619
Epoch 42/50
1.4618
Epoch 43/50
1.4619
Epoch 44/50
1.4619
Epoch 45/50
1.4617
Epoch 46/50
1.4617
Epoch 47/50
1.4617
Epoch 48/50
1.4617
Epoch 49/50
1.4617
Epoch 50/50
```

```
1.4615
276/276 [============ ] - 0s 149us/step
Epoch 1/50
25.1589
Epoch 2/50
15.1857
Epoch 3/50
10.9972
Epoch 4/50
8.6036
Epoch 5/50
7.0021
Epoch 6/50
5.8372
Epoch 7/50
4.9472
Epoch 8/50
4.2485
Epoch 9/50
3.6892
Epoch 10/50
3.2365
Epoch 11/50
2.8687
Epoch 12/50
2.5692
Epoch 13/50
2.3270
Epoch 14/50
2.1325
Epoch 15/50
1.9788
Epoch 16/50
1.8588
```

```
Epoch 17/50
1.7671
Epoch 18/50
1.6986
Epoch 19/50
1.6490
Epoch 20/50
1.6138
Epoch 21/50
1.5897
Epoch 22/50
1.5741
Epoch 23/50
1.5642
Epoch 24/50
1.5577
Epoch 25/50
1.5537
Epoch 26/50
1.5512
Epoch 27/50
1.5497
Epoch 28/50
1.5488
Epoch 29/50
1.5481
Epoch 30/50
1.5477
Epoch 31/50
1.5475
Epoch 32/50
1.5472
Epoch 33/50
```

```
1.5472
Epoch 34/50
1.5471
Epoch 35/50
1.5471
Epoch 36/50
1.5471
Epoch 37/50
1.5470
Epoch 38/50
1.5469
Epoch 39/50
1.5471
Epoch 40/50
1.5469
Epoch 41/50
1.5470
Epoch 42/50
1.5471
Epoch 43/50
1.5469
Epoch 44/50
1.5471
Epoch 45/50
1.5470
Epoch 46/50
1.5469
Epoch 47/50
1.5468
Epoch 48/50
1.5470
Epoch 49/50
1.5469
Epoch 50/50
```

```
1.5469
Epoch 1/50
26.1283
Epoch 2/50
15.8021
Epoch 3/50
11.4734
Epoch 4/50
9.0400
Epoch 5/50
7.4036
Epoch 6/50
6.2055
Epoch 7/50
5.2868
Epoch 8/50
4.5610
Epoch 9/50
3.9755
Epoch 10/50
3.4993
Epoch 11/50
3.1099
Epoch 12/50
2.7910
Epoch 13/50
2.5292
Epoch 14/50
2.3167
Epoch 15/50
2.1455
Epoch 16/50
```

```
2.0097
Epoch 17/50
1.9032
Epoch 18/50
1.8213
Epoch 19/50
1.7606
Epoch 20/50
1.7163
Epoch 21/50
1.6849
Epoch 22/50
Epoch 23/50
1.6479
Epoch 24/50
1.6383
Epoch 25/50
1.6320
Epoch 26/50
1.6281
Epoch 27/50
1.6256
Epoch 28/50
1.6240
Epoch 29/50
1.6229
Epoch 30/50
1.6223
Epoch 31/50
1.6219
Epoch 32/50
1.6215
Epoch 33/50
```

```
1.6212
Epoch 34/50
1.6210
Epoch 35/50
1.6207
Epoch 36/50
1.6206
Epoch 37/50
1.6205 0s - l
Epoch 38/50
1.6204
Epoch 39/50
1.6204
Epoch 40/50
1.6202
Epoch 41/50
1.6201
Epoch 42/50
1.6201
Epoch 43/50
1.6201
Epoch 44/50
1.6200
Epoch 45/50
1.6200
Epoch 46/50
1.6199
Epoch 47/50
1.6199
Epoch 48/50
1.6198
Epoch 49/50
```

```
1.6197
Epoch 50/50
1.6197
Epoch 1/50
25,4676
Epoch 2/50
14.8023
Epoch 3/50
10.5558
Epoch 4/50
8.2126
Epoch 5/50
6.6653
Epoch 6/50
5.5476
Epoch 7/50
4.7008
Epoch 8/50
4.0394
Epoch 9/50
3.5127
Epoch 10/50
3.0886
Epoch 11/50
2.7454
Epoch 12/50
2.4689
Epoch 13/50
2,2476
Epoch 14/50
2.0718
Epoch 15/50
1.9344
```

```
Epoch 16/50
1.8292
Epoch 17/50
1.7500 0s
Epoch 18/50
1.6914
Epoch 19/50
1.6500
Epoch 20/50
1.6215
Epoch 21/50
1.6028
Epoch 22/50
1.5910
Epoch 23/50
1.5837
Epoch 24/50
1.5790
Epoch 25/50
1.5763
Epoch 26/50
1.5746
Epoch 27/50
1.5736
Epoch 28/50
1.5730
Epoch 29/50
1.5726
Epoch 30/50
1.5724
Epoch 31/50
1.5720
Epoch 32/50
```

```
1.5721
Epoch 33/50
1.5718
Epoch 34/50
1.5717
Epoch 35/50
1.5715
Epoch 36/50
1.5716
Epoch 37/50
1.5714
Epoch 38/50
Epoch 39/50
1.5714
Epoch 40/50
1.5713
Epoch 41/50
1.5714
Epoch 42/50
1.5713
Epoch 43/50
1.5713
Epoch 44/50
1.5713
Epoch 45/50
1.5711
Epoch 46/50
1.5711
Epoch 47/50
1.5710
Epoch 48/50
1.5710
Epoch 49/50
```

```
1.5710
Epoch 50/50
1.5708
Epoch 1/50
25,2980
Epoch 2/50
14.9948
Epoch 3/50
10.8203
Epoch 4/50
8.4576
Epoch 5/50
6.8813
Epoch 6/50
5.7402
Epoch 7/50
4.8735
Epoch 8/50
4.1951
Epoch 9/50
3.6534
Epoch 10/50
3.2172
Epoch 11/50
2.8639
Epoch 12/50
2.5791
Epoch 13/50
2.3497
Epoch 14/50
2.1677
Epoch 15/50
```

```
2.0252
Epoch 16/50
1.9153
Epoch 17/50
1.8321
Epoch 18/50
1.7704
Epoch 19/50
1.7265
Epoch 20/50
1.6964
Epoch 21/50
Epoch 22/50
1.6630
Epoch 23/50
1.6547
Epoch 24/50
1.6500
Epoch 25/50
1.6472
Epoch 26/50
1.6452
Epoch 27/50
1.6440
Epoch 28/50
1.6431
Epoch 29/50
1.6427
Epoch 30/50
1.6424
Epoch 31/50
1.6421
Epoch 32/50
```

```
1.6418
Epoch 33/50
1.6416
Epoch 34/50
1.6415
Epoch 35/50
1.6413
Epoch 36/50
1.6412
Epoch 37/50
1.6410
Epoch 38/50
1.6410
Epoch 39/50
1.6409
Epoch 40/50
1.6406
Epoch 41/50
1.6405
Epoch 42/50
1.6403
Epoch 43/50
1.6403
Epoch 44/50
1.6402
Epoch 45/50
1.6401
Epoch 46/50
1.6399
Epoch 47/50
1.6397
Epoch 48/50
1.6398
```

```
Epoch 49/50
1.6396
Epoch 50/50
Epoch 1/50
24.9859
Epoch 2/50
14.8677
Epoch 3/50
10.7349
Epoch 4/50
Epoch 5/50
6.8155
Epoch 6/50
5.6794
Epoch 7/50
4.8161
Epoch 8/50
4.1409
Epoch 9/50
3.6053
Epoch 10/50
3.1745
Epoch 11/50
2.8262
Epoch 12/50
2.5460
Epoch 13/50
2.3217
Epoch 14/50
2.1435
Epoch 15/50
```

```
2.0042
Epoch 16/50
1.8974
Epoch 17/50
1.8175
Epoch 18/50
1.7582
Epoch 19/50
1.7156
Epoch 20/50
1.6868
Epoch 21/50
1.6676
Epoch 22/50
1.6556
Epoch 23/50
1.6480
Epoch 24/50
1.6431
Epoch 25/50
1.6402
Epoch 26/50
1.6384
Epoch 27/50
1.6374
Epoch 28/50
1.6365
Epoch 29/50
1.6361
Epoch 30/50
1.6356
Epoch 31/50
1.6354
```

```
Epoch 32/50
1.6352
Epoch 33/50
1.6351
Epoch 34/50
1.6350
Epoch 35/50
1.6348
Epoch 36/50
1.6347
Epoch 37/50
1.6347
Epoch 38/50
1.6346
Epoch 39/50
1.6345
Epoch 40/50
1.6345
Epoch 41/50
1.6344
Epoch 42/50
1.6342
Epoch 43/50
1.6343
Epoch 44/50
1.6340
Epoch 45/50
1.6341
Epoch 46/50
1.6340
Epoch 47/50
1.6339
Epoch 48/50
```

```
1.6339
Epoch 49/50
1.6339
Epoch 50/50
1.6338
Epoch 1/50
25,4776
Epoch 2/50
15.7145
Epoch 3/50
11.4615
Epoch 4/50
9.0156
Epoch 5/50
7.3742
Epoch 6/50
6.1785
Epoch 7/50
5.2666
Epoch 8/50
4.5491
Epoch 9/50
3.9724
Epoch 10/50
3.5032
Epoch 11/50
3.1198
Epoch 12/50
2.8069
Epoch 13/50
2.5534
Epoch 14/50
2.3468
```

```
Epoch 15/50
2.1813
Epoch 16/50
Epoch 17/50
1.9478
Epoch 18/50
1.8705
Epoch 19/50
1.8130
Epoch 20/50
1.7715
Epoch 21/50
1.7423
Epoch 22/50
1.7224
Epoch 23/50
1.7090
Epoch 24/50
1.7003
Epoch 25/50
1.6947
Epoch 26/50
1.6914
Epoch 27/50
1.6891
Epoch 28/50
1.6878
Epoch 29/50
1.6870
Epoch 30/50
1.6864
Epoch 31/50
```

```
1.6860
Epoch 32/50
1.6858
Epoch 33/50
1.6856
Epoch 34/50
1.6854
Epoch 35/50
1.6853
Epoch 36/50
1.6854
Epoch 37/50
670/2484 [======>.....] - ETA: 2s - loss: 1.5619
C:\ProgramData\Anaconda3\lib\site-packages\keras\callbacks\
callbacks.py:95: RuntimeWarning: Method (on train batch end) is slow
compared to the batch update (0.449026). Check your callbacks.
% (hook name, delta t median), RuntimeWarning)
C:\ProgramData\Anaconda3\lib\site-packages\keras\callbacks\
callbacks.py:95: RuntimeWarning: Method (on_train_batch_end) is slow
compared to the batch update (0.225013). Check your callbacks.
% (hook name, delta t median), RuntimeWarning)
1.6852
Epoch 38/50
1.6852
Epoch 39/50
1.6852
Epoch 40/50
1.6852
Epoch 41/50
1.6851
Epoch 42/50
1.6852
Epoch 43/50
1.6852
Epoch 44/50
```

```
1.6852
Epoch 45/50
1.6851
Epoch 46/50
1.6851
Epoch 47/50
1.6851
Epoch 48/50
1.6851
Epoch 49/50
1.6852
Epoch 50/50
Epoch 1/50
25.3121
Epoch 2/50
15.5666
Epoch 3/50
11.3659
Epoch 4/50
8.9442
Epoch 5/50
7.3193
Epoch 6/50
6.1381
Epoch 7/50
5.2320
Epoch 8/50
4.5167
Epoch 9/50
3.9421
Epoch 10/50
3.4758
```

```
Epoch 11/50
3.0942
Epoch 12/50
2.7823
Epoch 13/50
2.5288
Epoch 14/50
2.3247
Epoch 15/50
2.1610
Epoch 16/50
2.0319
Epoch 17/50
1.9317
Epoch 18/50
1.8550
Epoch 19/50
1.7986
Epoch 20/50
1.7586
Epoch 21/50
1.7307
Epoch 22/50
1.7113
Epoch 23/50
1.6983
Epoch 24/50
1.6900
Epoch 25/50
1.6850
Epoch 26/50
1.6818
Epoch 27/50
```

```
1.6795
Epoch 28/50
1.6783
Epoch 29/50
1.6773
Epoch 30/50
1.6766
Epoch 31/50
1.6763
Epoch 32/50
1.6760
Epoch 33/50
Epoch 34/50
1.6756
Epoch 35/50
1.6753
Epoch 36/50
1.6752
Epoch 37/50
1.6750
Epoch 38/50
1.6749
Epoch 39/50
1.6747
Epoch 40/50
1.6747
Epoch 41/50
1.6746
Epoch 42/50
1.6746
Epoch 43/50
1.6746
Epoch 44/50
```

```
1.6745
Epoch 45/50
1.6743
Epoch 46/50
1.6743
Epoch 47/50
1.6742
Epoch 48/50
1.6741
Epoch 49/50
1.6740
Epoch 50/50
1.6739
Epoch 1/50
25.6786
Epoch 2/50
15.1120
Epoch 3/50
10.8615
Epoch 4/50
8.4868
Epoch 5/50
6.9140
Epoch 6/50
5.7725
Epoch 7/50
4.9055
Epoch 8/50
4.2277
Epoch 9/50
3.6853
Epoch 10/50
```

```
3.2470
Epoch 11/50
2.8921
Epoch 12/50
2,6043
Epoch 13/50
2.3724
Epoch 14/50
2.1864
Epoch 15/50
2.0396
Epoch 16/50
1.9253
Epoch 17/50
1.8387
Epoch 18/50
1.7746
Epoch 19/50
1.7286
Epoch 20/50
1.6961
Epoch 21/50
1.6739
Epoch 22/50
1.6595
Epoch 23/50
1.6501
Epoch 24/50
1.6443
Epoch 25/50
1.6407
Epoch 26/50
1.6384
Epoch 27/50
```

```
1.6369
Epoch 28/50
1.6361
Epoch 29/50
1.6356
Epoch 30/50
1.6350
Epoch 31/50
1.6347
Epoch 32/50
1.6346
Epoch 33/50
1.6343
Epoch 34/50
1.6343
Epoch 35/50
1.6342
Epoch 36/50
1.6342
Epoch 37/50
1.6342
Epoch 38/50
1.6341
Epoch 39/50
1.6341
Epoch 40/50
1.6340
Epoch 41/50
1.6340
Epoch 42/50
1.6340
Epoch 43/50
1.6340
```

```
Epoch 44/50
1.6339
Epoch 45/50
1.6341
Epoch 46/50
1.6340
Epoch 47/50
1.6339
Epoch 48/50
1.6339
Epoch 49/50
1.6339
Epoch 50/50
1.6339
276/276 [=========== ] - 0s 163us/step
Standardized: 0.23 (0.11) MSE
```