DEAP LEARNING:

I have used KERAS to perform this deap learning exersise. I have used NHL Data and dropped some columns "id", "PlayerName", "sum_7yr_TOI", "Overall", "GP_greater_than_0","Country", and "DraftYear".

"sum_7yr_GP" is a target column.

I have standardized our training and test sets. Training data was chosen "DraftYear" 2004,2005,2006. And test data 2007.

Linear activation function was used for the output node. We cannot use sigmoid function for the output node because we are not having a classification problem.

Model was using to be model = Sequential().

There are settings I tried:

1) I used one hidden layer and an activation function RELU - rectified linear units for hidden layer. Plus. Loss function was calculates as MSE (Mean_Squared Error),

Optimiser was used Adam and metrics to evaluate accuracy was R^2.

```
adam = keras.optimizers.Adam()
rmsprop = keras.optimizers.RMSprop(lr=0.001)
model.compile(loss='mean_squared_error',
              optimizer=adam,
              metrics=[rsquared])
```

I have chosen 2 nodes for the second layer and Relu activation function.

```
model.add(Dense(units=2, activation='relu', input_shape=(22,)))
model.add(Dense(units=1, activation='linear'))
```

MSE was about 12000

I have tried 1 layer without any hidden layers:

```
model.add(Dense(units=1, activation='linear', input_shape=(22,)))
```

MSE was around 15000.

I have tried 3 layers (2 hidden layers with 4 nodes):

```
model.add(Dense(units=4, activation='relu', input_shape=(22,)))
model.add(Dense(units=4, activation='linear'))
model.add(Dense(units=1, activation='linear'))
```

The result was much better. Squared error was less, aroud 10000 than the previous times.

The Best result was acheived with the following metrics:

```
model.add(Dense(units=10, activation='relu', input_shape=(22,)))
model.add(Dense(units=50, activation='relu'))
model.add(Dense(units=50, activation='relu'))
model.add(Dense(units=1, activation='linear'))
```

The Squared Error was less than 9000. THerefore we can conclude that for our problem it's better to have more hidden notes and more layers.