Imports and Data Load

In [1]:
```
import pickle
import pandas as pd
import matplotlib.pyplot as plt
%matplotlib inline
import numpy as np

from sklearn.model_selection import train_test_split
from sklearn.metrics import mean_absolute_error
from sklearn.metrics import mean_squared_error
from keras import models
from keras import layers
from keras import optimizers
from keras import regularizers
from keras.models import load_model
from keras.callbacks import ModelCheckpoint

from sklearn.feature_extraction.text import TfidfVectorizer

from Functions import *
```

In [2]:
```
# Printing Errors
def errors(model,error_func,X_tr,X_te,y_tr,y_te,squared=False):
    train_error = error_func(y_tr,model.predict(X_tr))
    test_error = error_func(y_te,model.predict(X_te))
    if squared:
        train_error = train_error**.5
        test_error = test_error**.5
    print("Train Error:",round(train_error))
    print("Test Error:",round(test_error))
```

In [3]:
```
df = pickle.load(open(r"Data\players_cleaned_df.pickle","rb"))
```

Base Model

Train Test Validation Split

- Using 80% of the data to train

In [4]:
```
X = df['title']
y = df['views']
X_train_pre, X_test_pre, y_train, y_test = train_test_split(X,y,test_size=150,random_state=4521)
X_train_f_pre,X_val_pre,y_train_f,y_val = train_test_split(X_train_pre,y_train,test_size=100,random_state=4521)
```

TF IDF Vectorizing

In [5]:
```
tf = TfidfVectorizer(preprocessor=splitter, lowercase=False)
fit_tf = tf.fit(X_train_f_pre)
```

In [6]:
```
X_train_f = tf.transform(X_train_f_pre).todense()
```

```
X_val = tf.transform(X_val_pre).todense()
X_test = tf.transform(X_test_pre).todense()
```

Modeling

Loss on RMSE

In [7]:
```
# Building Model
model_mse = models.Sequential()
model_mse.add(layers.Dense(units=100,activation='relu',input_shape=(X_train_f.shape[1],)))
model_mse.add(layers.Dense(units=50,activation='relu'))
model_mse.add(layers.Dropout(rate=0.3))
model_mse.add(layers.Dense(units=50,activation='relu'))
model_mse.add(layers.Dropout(rate=0.3))
model_mse.add(layers.Dense(units=1,activation='linear'))

# Compilation Step
model_mse.compile(optimizer='adam',
             loss='mse',
             metrics=['mse'])
```

In [8]:
```
# Callbacks to save best model and weights
mse_callback = ModelCheckpoint(r'Data/model_mse.h5',monitor='val_mse',mode='min',save_best_only=True)
```

In [9]:
```
# Training Step
history_mse = model_mse.fit(X_train_f, y_train_f, batch_size=5,epochs=50,validation_data=(X_val,y_val),
                            callbacks=[mse_callback])
```

```
Epoch 1/50
201/201 [==============================] - 0s 2ms/step - loss: 22864877568.0000 - mse: 22864877568.0000 - val_loss: 22608545792.0000 - val_mse: 22608545792.0000
Epoch 2/50
201/201 [==============================] - 0s 1ms/step - loss: 21012891648.0000 - mse: 21012891648.0000 - val_loss: 17454555136.0000 - val_mse: 17454555136.0000
Epoch 3/50
201/201 [==============================] - 0s 1ms/step - loss: 14194554880.0000 - mse: 14194554880.0000 - val_loss: 11084909568.0000 - val_mse: 11084909568.0000
Epoch 4/50
201/201 [==============================] - 0s 1ms/step - loss: 11371836416.0000 - mse: 11371836416.0000 - val_loss: 10426193920.0000 - val_mse: 10426193920.0000
Epoch 5/50
201/201 [==============================] - 0s 1ms/step - loss: 10849937408.0000 - mse: 10849937408.0000 - val_loss: 10297724928.0000 - val_mse: 10297724928.0000
Epoch 6/50
201/201 [==============================] - 0s 2ms/step - loss: 10489520128.0000 - mse: 10489520128.0000 - val_loss: 10192214016.0000 - val_mse: 10192214016.0000
Epoch 7/50
201/201 [==============================] - 0s 1ms/step - loss: 10188888064.0000 - mse: 10188888064.0000 - val_loss: 10096762880.0000 - val_mse: 10096762880.0000
Epoch 8/50
201/201 [==============================] - 0s 1ms/step - loss: 10152669184.0000 - mse: 10152669184.0000 - val_loss: 10010775552.0000 - val_mse: 10010775552.0000
Epoch 9/50
201/201 [==============================] - 0s 1ms/step - loss: 10008129536.0000 - mse: 10008129536.0000 - val_loss: 9987360768.0000 - val_mse: 9987360768.0000
Epoch 10/50
201/201 [==============================] - 0s 1ms/step - loss: 9552846848.0000 - mse: 9552847872.0000 - val_loss: 9956941824.0000 - val_mse: 9956941824.0000
Epoch 11/50
201/201 [==============================] - 0s 1ms/step - loss: 9161172992.0000 - mse: 9161172992.0000 - val_loss: 9869966336.0000 - val_mse: 9869966336.0000
Epoch 12/50
201/201 [==============================] - 0s 806us/step - loss: 9258063872.0000 - mse: 9258063872.0000 - val_loss: 9872061440.0000 - val_mse: 9872061440.0000
Epoch 13/50
201/201 [==============================] - 0s 1ms/step - loss: 9315786752.0000 - mse: 9315786752.0000 - val_loss: 9854706688.0000 - val_mse: 9854706688.0000
Epoch 14/50
201/201 [==============================] - 0s 833us/step - loss: 8742283264.0000 - mse: 8742283264.0000 - val_loss: 9880892416.0000 - val_mse: 9880892416.0000
Epoch 15/50
201/201 [==============================] - 0s 811us/step - loss: 8788494336.0000 - mse: 8788495360.0000 - val_loss: 9911536640.0000 - val_mse: 9911536640.0000
Epoch 16/50
201/201 [==============================] - 0s 866us/step - loss: 8809339904.0000 - mse: 8809339904.0000 - val_loss: 9928307712.0000 - val_mse: 9928307712.0000
```

```
Epoch 17/50
201/201 [==============================] - 0s 871us/step - loss: 8456592384.0000 - mse: 8456592384.0000 - val_loss: 9974156288.0000 - val_mse: 9974156288.0000
Epoch 18/50
201/201 [==============================] - 0s 786us/step - loss: 8088896000.0000 - mse: 8088896000.0000 - val_loss: 10027202560.0000 - val_mse: 10027202560.0000
Epoch 19/50
201/201 [==============================] - 0s 791us/step - loss: 7946066432.0000 - mse: 7946066432.0000 - val_loss: 10075773952.0000 - val_mse: 10075773952.0000
Epoch 20/50
201/201 [==============================] - 0s 784us/step - loss: 8402590720.0000 - mse: 8402591232.0000 - val_loss: 9986622464.0000 - val_mse: 9986622464.0000
Epoch 21/50
201/201 [==============================] - 0s 794us/step - loss: 7432691712.0000 - mse: 7432691712.0000 - val_loss: 10104501248.0000 - val_mse: 10104501248.0000
Epoch 22/50
201/201 [==============================] - 0s 811us/step - loss: 7418897408.0000 - mse: 7418897408.0000 - val_loss: 10254196736.0000 - val_mse: 10254196736.0000
Epoch 23/50
201/201 [==============================] - 0s 806us/step - loss: 7511144960.0000 - mse: 7511144960.0000 - val_loss: 10359079936.0000 - val_mse: 10359079936.0000
Epoch 24/50
201/201 [==============================] - 0s 811us/step - loss: 7054074368.0000 - mse: 7054074368.0000 - val_loss: 10450403328.0000 - val_mse: 10450403328.0000
Epoch 25/50
201/201 [==============================] - 0s 851us/step - loss: 6811890688.0000 - mse: 6811890688.0000 - val_loss: 10523816960.0000 - val_mse: 10523816960.0000
Epoch 26/50
201/201 [==============================] - 0s 776us/step - loss: 6609806848.0000 - mse: 6609806848.0000 - val_loss: 10584741888.0000 - val_mse: 10584741888.0000
Epoch 27/50
201/201 [==============================] - 0s 915us/step - loss: 6278964224.0000 - mse: 6278964224.0000 - val_loss: 10814289920.0000 - val_mse: 10814290944.0000
Epoch 28/50
201/201 [==============================] - 0s 803us/step - loss: 6228643840.0000 - mse: 6228643840.0000 - val_loss: 10831601664.0000 - val_mse: 10831601664.0000
Epoch 29/50
201/201 [==============================] - 0s 826us/step - loss: 6382892544.0000 - mse: 6382893056.0000 - val_loss: 10973137920.0000 - val_mse: 10973137920.0000
Epoch 30/50
201/201 [==============================] - 0s 794us/step - loss: 6168603136.0000 - mse: 6168603136.0000 - val_loss: 11144339456.0000 - val_mse: 11144339456.0000
Epoch 31/50
201/201 [==============================] - 0s 826us/step - loss: 5850513920.0000 - mse: 5850513920.0000 - val_loss: 11192217600.0000 - val_mse: 11192217600.0000
Epoch 32/50
201/201 [==============================] - 0s 791us/step - loss: 5858704896.0000 - mse: 5858704896.0000 - val_loss: 11398353920.0000 - val_mse: 11398353920.0000
Epoch 33/50
201/201 [==============================] - 0s 786us/step - loss: 5905655808.0000 - mse: 5905655808.0000 - val_loss: 11580657664.0000 - val_mse: 11580657664.0000
Epoch 34/50
201/201 [==============================] - 0s 858us/step - loss: 5691936768.0000 - mse: 5691936768.0000 - val_loss: 11785588736.0000 - val_mse: 11785588736.0000
Epoch 35/50
201/201 [==============================] - 0s 796us/step - loss: 5059318272.0000 - mse: 5059318784.0000 - val_loss: 11862248448.0000 - val_mse: 11862248448.0000
Epoch 36/50
201/201 [==============================] - 0s 803us/step - loss: 5471619072.0000 - mse: 5471619072.0000 - val_loss: 12058333184.0000 - val_mse: 12058332160.0000
Epoch 37/50
201/201 [==============================] - 0s 818us/step - loss: 5588750336.0000 - mse: 5588750336.0000 - val_loss: 12300024832.0000 - val_mse: 12300024832.0000
Epoch 38/50
201/201 [==============================] - 0s 789us/step - loss: 5049052672.0000 - mse: 5049052672.0000 - val_loss: 12417171456.0000 - val_mse: 12417171456.0000
Epoch 39/50
201/201 [==============================] - 0s 784us/step - loss: 5125270016.0000 - mse: 5125270016.0000 - val_loss: 12548323328.0000 - val_mse: 12548323328.0000
Epoch 40/50
201/201 [==============================] - 0s 851us/step - loss: 4869039616.0000 - mse: 4869039616.0000 - val_loss: 12732727296.0000 - val_mse: 12732727296.0000
Epoch 41/50
201/201 [==============================] - 0s 831us/step - loss: 4914767872.0000 - mse: 4914767872.0000 - val_loss: 12838292480.0000 - val_mse: 12838292480.0000
Epoch 42/50
201/201 [==============================] - 0s 866us/step - loss: 4585613312.0000 - mse: 4585613312.0000 - val_loss: 12899891200.0000 - val_mse: 12899891200.0000
Epoch 43/50
201/201 [==============================] - 0s 808us/step - loss: 4261107200.0000 - mse: 4261107456.0000 - val_loss: 13349902336.0000 - val_mse: 13349902336.0000
Epoch 44/50
201/201 [==============================] - 0s 779us/step - loss: 4801070080.0000 - mse: 4801070080.0000 - val_loss: 13468987392.0000 - val_mse: 13468987392.0000
Epoch 45/50
201/201 [==============================] - 0s 791us/step - loss: 4417151488.0000 - mse: 4417151488.0000 - val_loss: 13740407808.0000 - val_mse: 13740407808.0000
Epoch 46/50
201/201 [==============================] - 0s 789us/step - loss: 4595155968.0000 - mse: 4595155968.0000 - val_loss: 13663938560.0000 - val_mse: 13663938560.0000
Epoch 47/50
201/201 [==============================] - 0s 791us/step - loss: 4085115648.0000 - mse: 4085115648.0000 - val_loss: 14013236224.0000 - val_mse: 14013236224.0000
Epoch 48/50
201/201 [==============================] - 0s 794us/step - loss: 4431241728.0000 - mse: 4431241728.0000 - val_loss: 14421834752.0000 - val_mse: 14421834752.0000
```

```
Epoch 49/50
201/201 [==============================] - 0s 786us/step - loss: 4279117312.0000 - mse: 4279117568.0000 - val_loss: 14184458240.0000 - val_mse: 14184458240.0000
Epoch 50/50
201/201 [==============================] - 0s 774us/step - loss: 3892425728.0000 - mse: 3892425984.0000 - val_loss: 14483440640.0000 - val_mse: 14483440640.0000
```

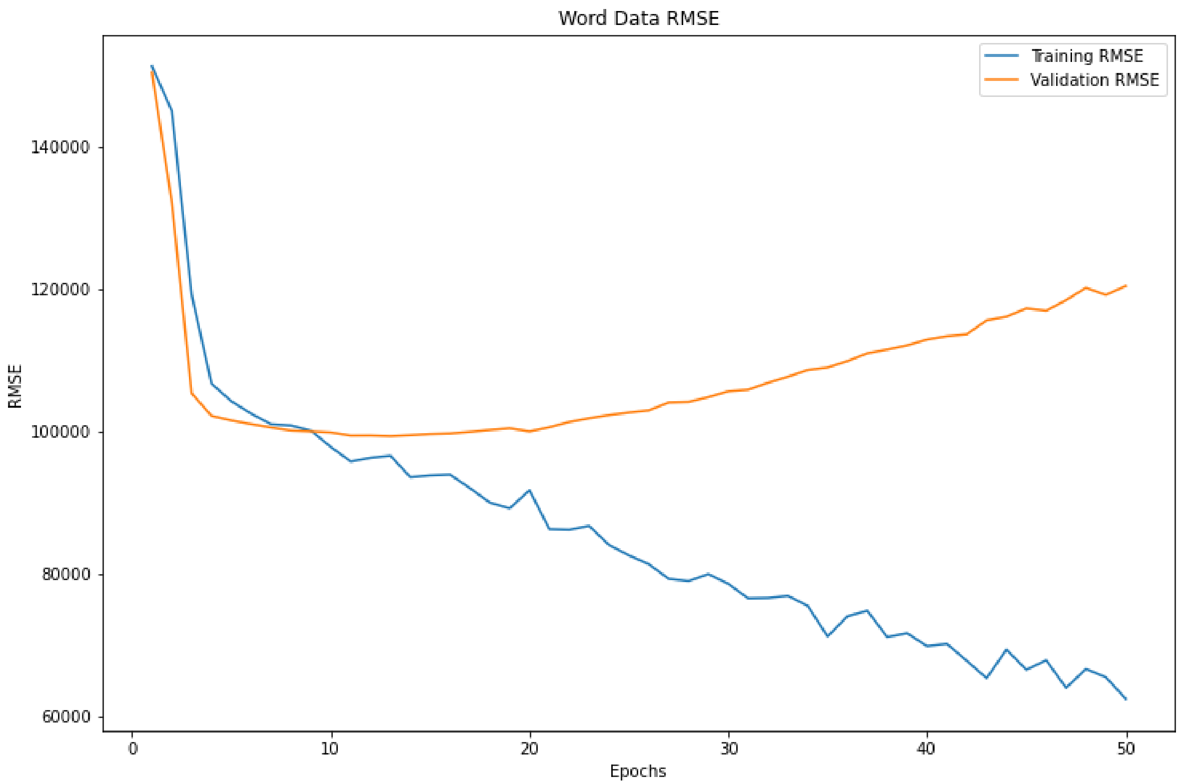
In [10]:

```
# Plotting Loss
fig, ax = plt.subplots(figsize=(12, 8))

model_dict = history_mse.history

rmse_values = np.sqrt(model_dict['mse'])
val_rmse_values = np.sqrt(model_dict['val_mse'])

epochs = range(1, len(rmse_values) + 1)
ax.plot(epochs, rmse_values, label='Training RMSE')
ax.plot(epochs, val_rmse_values, label='Validation RMSE')
plt.legend()
plt.title('Word Data RMSE')
plt.xlabel('Epochs')
plt.ylabel('RMSE')
plt.show()
```



Best Results

In [11]:

```
best_mse = load_model(r'Data/model_mse.h5')
errors(best_mse,mean_squared_error,X_train_f, X_test,y_train_f,y_test,squared=True)
```

```
Train Error: 92474
Test Error: 80133
```

In [42]:
```python
print("Average Views: ", round(y_train_f.mean()))
```

```
Average Views:  107394
```

Loss on MAE

In [12]:
```python
# Building Model
model_mae = models.Sequential()
model_mae.add(layers.Dense(units=100,activation='relu',input_shape=(X_train_f.shape[1],)))
model_mae.add(layers.Dense(units=50,activation='relu'))
model_mae.add(layers.Dropout(rate=0.3))
model_mae.add(layers.Dense(units=50,activation='relu'))
model_mae.add(layers.Dropout(rate=0.3))
model_mae.add(layers.Dense(units=1,activation='linear'))

# Compilation Step
model_mae.compile(optimizer='adam',
             loss='mae',
             metrics=['mae'])
```

In [13]:
```python
# Callbacks to save best model and weights
mae_callback = ModelCheckpoint(r'Data/model_mae.h5',monitor='val_mae',mode='min',save_best_only=True)
```

In [14]:
```python
# Training Step
history_mae = model_mae.fit(X_train_f, y_train_f, batch_size=5,epochs=50,validation_data=(X_val,y_val),
                            callbacks=[mae_callback])
```

```
Epoch 1/50
201/201 [==============================] - 0s 2ms/step - loss: 107011.4453 - mae: 107011.4453 - val_loss: 104764.5234 - val_mae: 104764.5234
Epoch 2/50
201/201 [==============================] - 0s 1ms/step - loss: 86345.1406 - mae: 86345.1406 - val_loss: 57888.5391 - val_mae: 57888.5391
Epoch 3/50
201/201 [==============================] - 0s 1ms/step - loss: 54574.5352 - mae: 54574.5352 - val_loss: 51394.2266 - val_mae: 51394.2266
Epoch 4/50
201/201 [==============================] - 0s 1ms/step - loss: 51534.9492 - mae: 51534.9492 - val_loss: 50059.7500 - val_mae: 50059.7500
Epoch 5/50
201/201 [==============================] - 0s 2ms/step - loss: 49361.3359 - mae: 49361.3359 - val_loss: 48721.4688 - val_mae: 48721.4688
Epoch 6/50
201/201 [==============================] - 0s 1ms/step - loss: 47231.3672 - mae: 47231.3672 - val_loss: 48354.5039 - val_mae: 48354.5039
Epoch 7/50
201/201 [==============================] - 0s 2ms/step - loss: 46755.5898 - mae: 46755.5898 - val_loss: 47161.7891 - val_mae: 47161.7891
Epoch 8/50
201/201 [==============================] - 0s 1ms/step - loss: 45464.2852 - mae: 45464.2852 - val_loss: 47056.0781 - val_mae: 47056.0781
Epoch 9/50
201/201 [==============================] - 0s 1ms/step - loss: 44402.3125 - mae: 44402.3125 - val_loss: 46692.6992 - val_mae: 46692.6992
Epoch 10/50
201/201 [==============================] - 0s 788us/step - loss: 43394.0508 - mae: 43394.0508 - val_loss: 46802.6250 - val_mae: 46802.6250
Epoch 11/50
201/201 [==============================] - 0s 1ms/step - loss: 41698.8281 - mae: 41698.8281 - val_loss: 46638.0117 - val_mae: 46638.0117
Epoch 12/50
201/201 [==============================] - 0s 793us/step - loss: 40573.7734 - mae: 40573.7734 - val_loss: 47569.9062 - val_mae: 47569.9062
Epoch 13/50
201/201 [==============================] - 0s 853us/step - loss: 39210.9609 - mae: 39210.9609 - val_loss: 47104.0469 - val_mae: 47104.0469
Epoch 14/50
201/201 [==============================] - 0s 813us/step - loss: 37762.3828 - mae: 37762.3828 - val_loss: 48053.7969 - val_mae: 48053.7969
Epoch 15/50
201/201 [==============================] - 0s 786us/step - loss: 38326.9922 - mae: 38326.9922 - val_loss: 47839.0742 - val_mae: 47839.0742
Epoch 16/50
201/201 [==============================] - 0s 781us/step - loss: 37075.1992 - mae: 37075.1992 - val_loss: 48363.8164 - val_mae: 48363.8164
Epoch 17/50
201/201 [==============================] - 0s 789us/step - loss: 36390.5547 - mae: 36390.5547 - val_loss: 48390.9609 - val_mae: 48390.9609
```

```
Epoch 18/50
201/201 [==============================] - 0s 794us/step - loss: 36131.7969 - mae: 36131.7969 - val_loss: 48962.1289 - val_mae: 48962.1289
Epoch 19/50
201/201 [==============================] - 0s 806us/step - loss: 35132.2578 - mae: 35132.2617 - val_loss: 48851.9844 - val_mae: 48851.9844
Epoch 20/50
201/201 [==============================] - 0s 781us/step - loss: 35110.4102 - mae: 35110.4102 - val_loss: 48773.9219 - val_mae: 48773.9219
Epoch 21/50
201/201 [==============================] - 0s 794us/step - loss: 33489.8125 - mae: 33489.8125 - val_loss: 48902.8086 - val_mae: 48902.8086
Epoch 22/50
201/201 [==============================] - 0s 769us/step - loss: 34056.5625 - mae: 34056.5625 - val_loss: 49159.5781 - val_mae: 49159.5781
Epoch 23/50
201/201 [==============================] - 0s 799us/step - loss: 33200.8945 - mae: 33200.8906 - val_loss: 49952.7031 - val_mae: 49952.7031
Epoch 24/50
201/201 [==============================] - 0s 801us/step - loss: 33130.9297 - mae: 33130.9297 - val_loss: 49907.0547 - val_mae: 49907.0547
Epoch 25/50
201/201 [==============================] - 0s 774us/step - loss: 32135.2695 - mae: 32135.2695 - val_loss: 49784.7109 - val_mae: 49784.7109
Epoch 26/50
201/201 [==============================] - 0s 774us/step - loss: 31463.3008 - mae: 31463.3008 - val_loss: 50212.3008 - val_mae: 50212.3008
Epoch 27/50
201/201 [==============================] - 0s 806us/step - loss: 31575.9102 - mae: 31575.9102 - val_loss: 50355.8945 - val_mae: 50355.8945
Epoch 28/50
201/201 [==============================] - 0s 821us/step - loss: 31666.7969 - mae: 31666.7969 - val_loss: 51314.6367 - val_mae: 51314.6367
Epoch 29/50
201/201 [==============================] - 0s 771us/step - loss: 30925.7598 - mae: 30925.7598 - val_loss: 51230.0469 - val_mae: 51230.0469
Epoch 30/50
201/201 [==============================] - 0s 781us/step - loss: 30831.8984 - mae: 30831.8984 - val_loss: 51784.1719 - val_mae: 51784.1719
Epoch 31/50
201/201 [==============================] - 0s 789us/step - loss: 30097.2676 - mae: 30097.2676 - val_loss: 51212.1797 - val_mae: 51212.1797
Epoch 32/50
201/201 [==============================] - 0s 784us/step - loss: 29891.1953 - mae: 29891.1914 - val_loss: 51281.8164 - val_mae: 51281.8164
Epoch 33/50
201/201 [==============================] - 0s 789us/step - loss: 29686.7871 - mae: 29686.7871 - val_loss: 51840.7812 - val_mae: 51840.7812
Epoch 34/50
201/201 [==============================] - 0s 836us/step - loss: 29048.4629 - mae: 29048.4629 - val_loss: 51725.4258 - val_mae: 51725.4258
Epoch 35/50
201/201 [==============================] - 0s 774us/step - loss: 28326.9355 - mae: 28326.9355 - val_loss: 53296.2695 - val_mae: 53296.2695
Epoch 36/50
201/201 [==============================] - 0s 789us/step - loss: 28548.0176 - mae: 28548.0176 - val_loss: 52490.6953 - val_mae: 52490.6953
Epoch 37/50
201/201 [==============================] - 0s 813us/step - loss: 28404.0000 - mae: 28404.0000 - val_loss: 52686.0586 - val_mae: 52686.0586
Epoch 38/50
201/201 [==============================] - 0s 776us/step - loss: 28144.8262 - mae: 28144.8262 - val_loss: 53218.9062 - val_mae: 53218.9062
Epoch 39/50
201/201 [==============================] - 0s 801us/step - loss: 27857.2793 - mae: 27857.2793 - val_loss: 53390.2734 - val_mae: 53390.2734
Epoch 40/50
201/201 [==============================] - 0s 858us/step - loss: 28127.9512 - mae: 28127.9512 - val_loss: 53219.7383 - val_mae: 53219.7383
Epoch 41/50
201/201 [==============================] - 0s 779us/step - loss: 28088.2812 - mae: 28088.2812 - val_loss: 53781.9453 - val_mae: 53781.9453
Epoch 42/50
201/201 [==============================] - 0s 799us/step - loss: 28166.3809 - mae: 28166.3809 - val_loss: 54238.2617 - val_mae: 54238.2617
Epoch 43/50
201/201 [==============================] - 0s 781us/step - loss: 27677.0371 - mae: 27677.0371 - val_loss: 54818.6562 - val_mae: 54818.6562
Epoch 44/50
201/201 [==============================] - 0s 915us/step - loss: 27511.9082 - mae: 27511.9082 - val_loss: 54264.6445 - val_mae: 54264.6445
Epoch 45/50
201/201 [==============================] - 0s 786us/step - loss: 27040.6895 - mae: 27040.6895 - val_loss: 55468.1133 - val_mae: 55468.1133
Epoch 46/50
201/201 [==============================] - 0s 806us/step - loss: 26729.5742 - mae: 26729.5742 - val_loss: 56815.9336 - val_mae: 56815.9336
Epoch 47/50
201/201 [==============================] - 0s 771us/step - loss: 25381.4980 - mae: 25381.4980 - val_loss: 55713.6484 - val_mae: 55713.6484
Epoch 48/50
201/201 [==============================] - 0s 781us/step - loss: 25906.0332 - mae: 25906.0332 - val_loss: 55635.4062 - val_mae: 55635.4062
Epoch 49/50
201/201 [==============================] - 0s 801us/step - loss: 25913.2793 - mae: 25913.2793 - val_loss: 56609.4805 - val_mae: 56609.4805
```

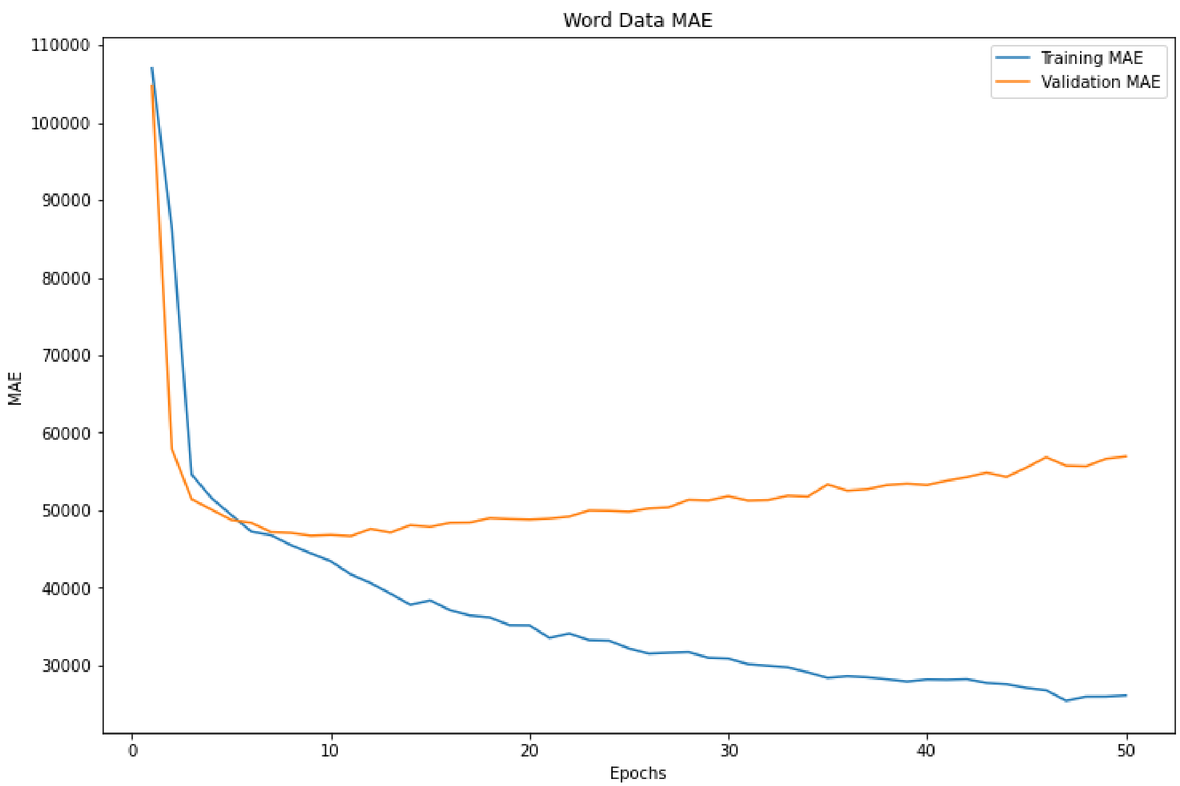
```
Epoch 50/50
201/201 [==============================] - 0s 771us/step - loss: 26062.0176 - mae: 26062.0176 - val_loss: 56917.1914 - val_mae: 56917.1914
```

In [15]:
```python
# Plotting Loss
fig, ax = plt.subplots(figsize=(12, 8))

model_dict = history_mae.history

mae_values = model_dict['mae']
val_mae_values = model_dict['val_mae']

epochs = range(1, len(mae_values) + 1)
ax.plot(epochs, mae_values, label='Training MAE')
ax.plot(epochs, val_mae_values, label='Validation MAE')
plt.legend()
plt.title('Word Data MAE')
plt.xlabel('Epochs')
plt.ylabel('MAE')
plt.show()
```



Best Results

In [16]:
```python
best_mae = load_model(r'Data/model_mae.h5')
errors(best_mae,mean_absolute_error,X_train_f, X_test,y_train_f,y_test,squared=False)
```

```
Train Error: 38081
Test Error: 46656
```

Model After Removing Outliers

Outlier Removal using IQR along with Train Test Validation Split

In [17]:
```
round(y_train.describe(),2)
```

Out[17]:
```
count       1102.00
mean      107397.09
std       106641.20
min         8088.00
25%        54421.00
50%        81922.00
75%       127335.75
max      1783323.00
Name: views, dtype: float64
```

In [18]:
```
iqr_df = pickle.load(open(r"Data\players_cleaned_df.pickle","rb"))
```

In [19]:
```
X_iqr = iqr_df[['title']]
y_iqr = iqr_df[['views']]

X_train_iqr_pre,X_test_iqr_pre,y_train_iqr,y_test_iqr = train_test_split(X_iqr,y_iqr,test_size=150,random_state=4521)
```

In [20]:
```
# Obtaining Lower and Upper Limits using IQR
q25 = y_train_iqr['views'].describe()['25%']
q75 = y_train_iqr['views'].describe()['75%']
iqr = q75-q25
low_lim = q25-1.5*iqr
upp_lim = q75+1.5*iqr
```

In [21]:
```
# Resetting Index to use to help filtering out locations of outliers
X_train_iqr_pre.reset_index(inplace=True)
X_test_iqr_pre.reset_index(inplace=True)
y_train_iqr.reset_index(inplace=True)
y_test_iqr.reset_index(inplace=True)
```

In [22]:
```
# Finding Indexes within range
train_keep_in = y_train_iqr['views'].index[(low_lim < y_train_iqr['views']) & (y_train_iqr['views'] < upp_lim)]
test_keep_in = y_test_iqr['views'].index[(low_lim < y_test_iqr['views']) & (y_test_iqr['views'] < upp_lim)]
```

In [23]:
```
# New Train and Test data after removing outliers based on train
X_train_iqrf_pre = X_train_iqr_pre.iloc[train_keep_in].drop('index',axis=1)['title']
y_train_iqrf = np.array(y_train_iqr.iloc[train_keep_in].drop('index',axis=1))
X_test_iqrf_pre = X_test_iqr_pre.iloc[test_keep_in].drop('index',axis=1)['title']
y_test_iqrf = np.array(y_test_iqr.iloc[test_keep_in].drop('index',axis=1))
```

In [24]:
```
# Validation Split
X_train_val_iqr_pre,X_val_iqr_pre,y_train_val_iqr,y_val_iqr = train_test_split(X_train_iqrf_pre,
                                                                               y_train_iqrf,
                                                                               test_size=100,random_state=4521)
```

Refitting TFIDF Vectorizer

In [25]:
```
tf_o = TfidfVectorizer(preprocessor=splitter, lowercase=False)
```

```
fit_tf_o = tf_o.fit(X_train_val_iqr_pre)
```

In [26]:
```
X_train_val_iqr = tf_o.transform(X_train_val_iqr_pre).todense()
X_val_iqr = tf_o.transform(X_val_iqr_pre).todense()
X_test_iqrf = tf_o.transform(X_test_iqrf_pre).todense()
```

Modeling P2

Loss on RMSE

In [27]:
```
# Building Model
model_mse_iqr = models.Sequential()
model_mse_iqr.add(layers.Dense(units=100,activation='relu',input_shape=(X_train_val_iqr.shape[1],)))
model_mse_iqr.add(layers.Dense(units=50,activation='relu'))
model_mse_iqr.add(layers.Dropout(rate=0.3))
model_mse_iqr.add(layers.Dense(units=50,activation='relu'))
model_mse_iqr.add(layers.Dropout(rate=0.3))
model_mse_iqr.add(layers.Dense(units=1,activation='linear'))

# Compilation Step
model_mse_iqr.compile(optimizer='adam',
             loss='mse',
             metrics=['mse'])
```

In [28]:
```
# Callbacks to save best model and weights
mse_iqr_callback = ModelCheckpoint(r'Data/model_mse_iqr.h5',monitor='val_mse',mode='min',save_best_only=True)
```

In [29]:
```
# Training Step
history_mse_iqr = model_mse_iqr.fit(X_train_val_iqr, y_train_val_iqr,
                                    batch_size=5,epochs=50,
                                    validation_data=(X_val_iqr,y_val_iqr),
                                    callbacks=[mse_iqr_callback])
```

```
Epoch 1/50
186/186 [==============================] - 0s 2ms/step - loss: 9669542912.0000 - mse: 9669542912.0000 - val_loss: 9165594624.0000 - val_mse: 9165594624.0000
Epoch 2/50
186/186 [==============================] - 0s 1ms/step - loss: 7269039104.0000 - mse: 7269039104.0000 - val_loss: 3394633728.0000 - val_mse: 3394632960.0000
Epoch 3/50
186/186 [==============================] - 0s 1ms/step - loss: 2552547840.0000 - mse: 2552547840.0000 - val_loss: 1484074368.0000 - val_mse: 1484074368.0000
Epoch 4/50
186/186 [==============================] - 0s 2ms/step - loss: 1995598592.0000 - mse: 1995598592.0000 - val_loss: 1413213696.0000 - val_mse: 1413213568.0000
Epoch 5/50
186/186 [==============================] - 0s 2ms/step - loss: 1826328704.0000 - mse: 1826328704.0000 - val_loss: 1373378560.0000 - val_mse: 1373378560.0000
Epoch 6/50
186/186 [==============================] - 0s 1ms/step - loss: 1632627328.0000 - mse: 1632627328.0000 - val_loss: 1364385280.0000 - val_mse: 1364385280.0000
Epoch 7/50
186/186 [==============================] - 0s 1ms/step - loss: 1577581184.0000 - mse: 1577581184.0000 - val_loss: 1343427712.0000 - val_mse: 1343427712.0000
Epoch 8/50
186/186 [==============================] - 0s 806us/step - loss: 1432932608.0000 - mse: 1432932608.0000 - val_loss: 1349921152.0000 - val_mse: 1349921152.0000
Epoch 9/50
186/186 [==============================] - 0s 820us/step - loss: 1408477952.0000 - mse: 1408477952.0000 - val_loss: 1374935808.0000 - val_mse: 1374935808.0000
Epoch 10/50
186/186 [==============================] - 0s 817us/step - loss: 1323601536.0000 - mse: 1323601536.0000 - val_loss: 1365176960.0000 - val_mse: 1365176960.0000
Epoch 11/50
186/186 [==============================] - 0s 817us/step - loss: 1275370240.0000 - mse: 1275370240.0000 - val_loss: 1371182976.0000 - val_mse: 1371182976.0000
Epoch 12/50
186/186 [==============================] - 0s 823us/step - loss: 1101467904.0000 - mse: 1101467904.0000 - val_loss: 1421867264.0000 - val_mse: 1421867264.0000
Epoch 13/50
186/186 [==============================] - 0s 820us/step - loss: 1169893120.0000 - mse: 1169893120.0000 - val_loss: 1437480576.0000 - val_mse: 1437480576.0000
```

```
Epoch 14/50
186/186 [==============================] - 0s 812us/step - loss: 1065137152.0000 - mse: 1065137152.0000 - val_loss: 1413002752.0000 - val_mse: 1413002752.0000
Epoch 15/50
186/186 [==============================] - 0s 828us/step - loss: 980213952.0000 - mse: 980213952.0000 - val_loss: 1427142400.0000 - val_mse: 1427142400.0000
Epoch 16/50
186/186 [==============================] - 0s 814us/step - loss: 980904960.0000 - mse: 980904960.0000 - val_loss: 1538939904.0000 - val_mse: 1538939904.0000
Epoch 17/50
186/186 [==============================] - 0s 858us/step - loss: 915076672.0000 - mse: 915076672.0000 - val_loss: 1474715520.0000 - val_mse: 1474715520.0000
Epoch 18/50
186/186 [==============================] - 0s 828us/step - loss: 867639296.0000 - mse: 867639296.0000 - val_loss: 1534741376.0000 - val_mse: 1534741376.0000
Epoch 19/50
186/186 [==============================] - 0s 858us/step - loss: 772372864.0000 - mse: 772372864.0000 - val_loss: 1641467904.0000 - val_mse: 1641467904.0000
Epoch 20/50
186/186 [==============================] - 0s 804us/step - loss: 777818496.0000 - mse: 777818496.0000 - val_loss: 1646414720.0000 - val_mse: 1646414720.0000
Epoch 21/50
186/186 [==============================] - 0s 815us/step - loss: 729658624.0000 - mse: 729658624.0000 - val_loss: 1613557632.0000 - val_mse: 1613557632.0000
Epoch 22/50
186/186 [==============================] - 0s 815us/step - loss: 752312576.0000 - mse: 752312576.0000 - val_loss: 1646134912.0000 - val_mse: 1646134912.0000
Epoch 23/50
186/186 [==============================] - 0s 806us/step - loss: 692163648.0000 - mse: 692163648.0000 - val_loss: 1705949696.0000 - val_mse: 1705949696.0000
Epoch 24/50
186/186 [==============================] - 0s 841us/step - loss: 666294784.0000 - mse: 666294784.0000 - val_loss: 1709229568.0000 - val_mse: 1709229568.0000
Epoch 25/50
186/186 [==============================] - 0s 809us/step - loss: 679875392.0000 - mse: 679875392.0000 - val_loss: 1811406336.0000 - val_mse: 1811406336.0000
Epoch 26/50
186/186 [==============================] - 0s 874us/step - loss: 602394944.0000 - mse: 602394944.0000 - val_loss: 1835742464.0000 - val_mse: 1835742464.0000
Epoch 27/50
186/186 [==============================] - 0s 815us/step - loss: 622894592.0000 - mse: 622894592.0000 - val_loss: 1847775104.0000 - val_mse: 1847775104.0000
Epoch 28/50
186/186 [==============================] - 0s 852us/step - loss: 587455168.0000 - mse: 587455168.0000 - val_loss: 1817292416.0000 - val_mse: 1817292416.0000
Epoch 29/50
186/186 [==============================] - 0s 817us/step - loss: 647716608.0000 - mse: 647716608.0000 - val_loss: 1854676224.0000 - val_mse: 1854676224.0000
Epoch 30/50
186/186 [==============================] - 0s 836us/step - loss: 538372416.0000 - mse: 538372416.0000 - val_loss: 1903652096.0000 - val_mse: 1903652096.0000
Epoch 31/50
186/186 [==============================] - 0s 823us/step - loss: 519506688.0000 - mse: 519506688.0000 - val_loss: 1971063296.0000 - val_mse: 1971063296.0000
Epoch 32/50
186/186 [==============================] - 0s 809us/step - loss: 530193280.0000 - mse: 530193280.0000 - val_loss: 1905033984.0000 - val_mse: 1905033984.0000
Epoch 33/50
186/186 [==============================] - 0s 863us/step - loss: 510174496.0000 - mse: 510174496.0000 - val_loss: 1936756736.0000 - val_mse: 1936756736.0000
Epoch 34/50
186/186 [==============================] - 0s 820us/step - loss: 577728960.0000 - mse: 577728960.0000 - val_loss: 1962839040.0000 - val_mse: 1962839040.0000
Epoch 35/50
186/186 [==============================] - 0s 806us/step - loss: 550592960.0000 - mse: 550592960.0000 - val_loss: 1960493568.0000 - val_mse: 1960493568.0000
Epoch 36/50
186/186 [==============================] - 0s 812us/step - loss: 497770528.0000 - mse: 497770528.0000 - val_loss: 2009859840.0000 - val_mse: 2009859840.0000
Epoch 37/50
186/186 [==============================] - 0s 868us/step - loss: 500594720.0000 - mse: 500594720.0000 - val_loss: 2078542592.0000 - val_mse: 2078542592.0000
Epoch 38/50
186/186 [==============================] - 0s 804us/step - loss: 473274528.0000 - mse: 473274528.0000 - val_loss: 2013526272.0000 - val_mse: 2013526272.0000
Epoch 39/50
186/186 [==============================] - 0s 820us/step - loss: 488786848.0000 - mse: 488786848.0000 - val_loss: 2041867904.0000 - val_mse: 2041867904.0000
Epoch 40/50
186/186 [==============================] - 0s 858us/step - loss: 498232736.0000 - mse: 498232736.0000 - val_loss: 2082082432.0000 - val_mse: 2082082432.0000
Epoch 41/50
186/186 [==============================] - 0s 815us/step - loss: 477123616.0000 - mse: 477123616.0000 - val_loss: 2001896192.0000 - val_mse: 2001896192.0000
Epoch 42/50
186/186 [==============================] - 0s 828us/step - loss: 458445664.0000 - mse: 458445664.0000 - val_loss: 2157771264.0000 - val_mse: 2157771264.0000
Epoch 43/50
186/186 [==============================] - 0s 831us/step - loss: 474575968.0000 - mse: 474575968.0000 - val_loss: 2125895936.0000 - val_mse: 2125895936.0000
Epoch 44/50
186/186 [==============================] - 0s 874us/step - loss: 473503616.0000 - mse: 473503616.0000 - val_loss: 2193018624.0000 - val_mse: 2193018624.0000
Epoch 45/50
186/186 [==============================] - 0s 839us/step - loss: 451639328.0000 - mse: 451639328.0000 - val_loss: 2191993344.0000 - val_mse: 2191993344.0000
```

```
Epoch 46/50
186/186 [==============================] - 0s 812us/step - loss: 463224000.0000 - mse: 463224000.0000 - val_loss: 2232480000.0000 - val_mse: 2232480000.0000
Epoch 47/50
186/186 [==============================] - 0s 860us/step - loss: 413991616.0000 - mse: 413991616.0000 - val_loss: 2160750848.0000 - val_mse: 2160750848.0000
Epoch 48/50
186/186 [==============================] - 0s 809us/step - loss: 446111456.0000 - mse: 446111456.0000 - val_loss: 2331602176.0000 - val_mse: 2331602176.0000
Epoch 49/50
186/186 [==============================] - 0s 806us/step - loss: 430186208.0000 - mse: 430186208.0000 - val_loss: 2244182528.0000 - val_mse: 2244182528.0000
Epoch 50/50
186/186 [==============================] - 0s 817us/step - loss: 393826528.0000 - mse: 393826528.0000 - val_loss: 2161483008.0000 - val_mse: 2161483008.0000
```

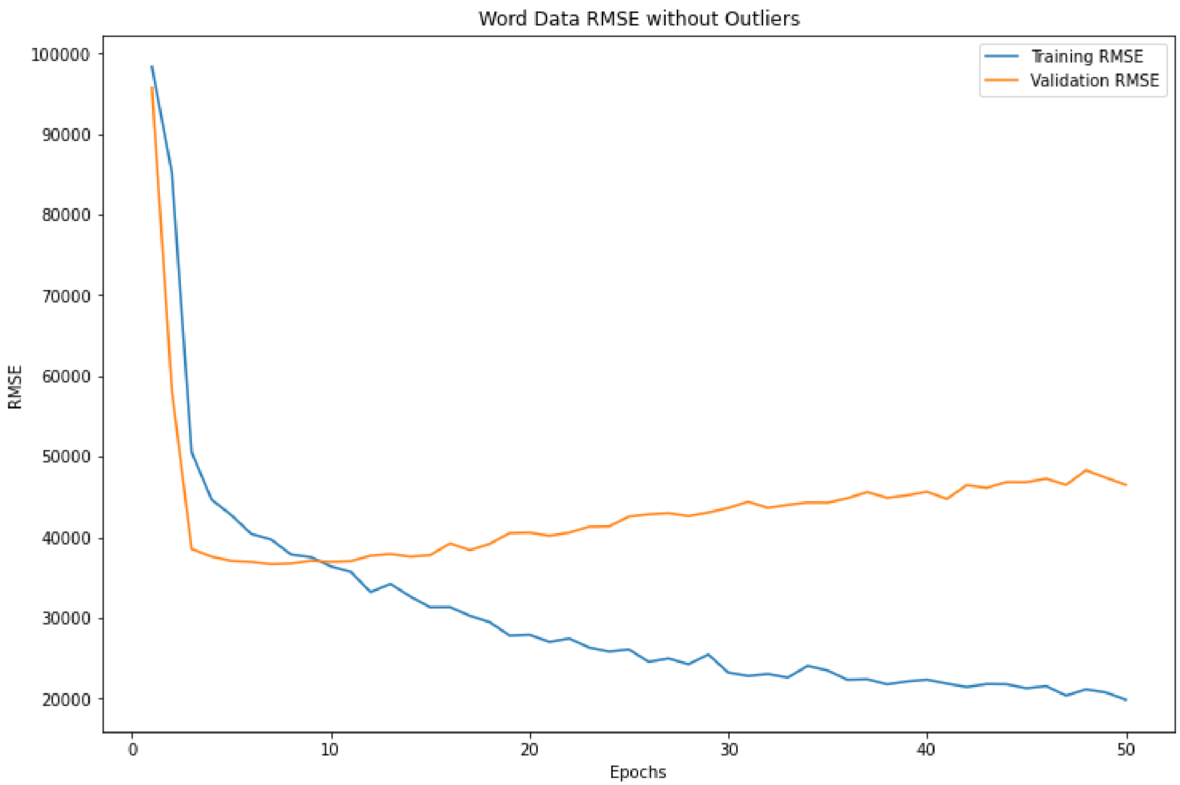
In [30]:

```python
# Plotting Loss
fig, ax = plt.subplots(figsize=(12, 8))

model_dict = history_mse_iqr.history

rmse_values = np.sqrt(model_dict['mse'])
val_rmse_values = np.sqrt(model_dict['val_mse'])

epochs = range(1, len(rmse_values) + 1)
ax.plot(epochs, rmse_values, label='Training RMSE')
ax.plot(epochs, val_rmse_values, label='Validation RMSE')
plt.legend()
plt.title('Word Data RMSE without Outliers')
plt.xlabel('Epochs')
plt.ylabel('RMSE')
plt.show()
```



Best Results

In [31]:
```
best_mse_iqr = load_model(r'Data/model_mse_iqr.h5')
errors(best_mse_iqr,mean_squared_error,X_train_val_iqr,X_test_iqrf,
       y_train_val_iqr,y_test_iqrf,squared=True)
```

```
Train Error: 35893
Test Error: 40412
```

Loss on MAE

In [32]:
```
# Building Model
model_mae_iqr = models.Sequential()
model_mae_iqr.add(layers.Dense(units=100,activation='relu',input_shape=(X_train_val_iqr.shape[1],)))
model_mae_iqr.add(layers.Dense(units=50,activation='relu'))
model_mae_iqr.add(layers.Dropout(rate=0.3))
model_mae_iqr.add(layers.Dense(units=50,activation='relu'))
model_mae_iqr.add(layers.Dropout(rate=0.3))
model_mae_iqr.add(layers.Dense(units=1,activation='linear'))

# Compilation Step
model_mae_iqr.compile(optimizer='adam',
             loss='mae',
             metrics=['mae'])
```

In [33]:
```
# Callbacks to save best model and weights
mae_iqr_callback = ModelCheckpoint(r'Data/model_mae_iqr.h5',monitor='val_mae',mode='min',save_best_only=True)
```

In [34]:
```
# Training Step
history_mae_iqr = model_mae_iqr.fit(np.array(X_train_val_iqr), y_train_val_iqr,
                                    batch_size=5,epochs=50,
                                    validation_data=(np.array(X_val_iqr),y_val_iqr),
                                    callbacks=[mae_iqr_callback])
```

```
Epoch 1/50
134/186 [====================>.........] - ETA: 0s - loss: 89491.8750 - mae: 89491.8750WARNING:tensorflow:Callbacks method `on_test_batch_end` is slow compared to the ba
tch time (batch time: 0.0000s vs `on_test_batch_end` time: 0.0005s). Check your callbacks.
186/186 [==============================] - 0s 2ms/step - loss: 86495.9922 - mae: 86495.9922 - val_loss: 87916.8516 - val_mae: 87916.8516
Epoch 2/50
186/186 [==============================] - 0s 2ms/step - loss: 75556.8203 - mae: 75556.8203 - val_loss: 55836.1133 - val_mae: 55836.1133
Epoch 3/50
186/186 [==============================] - 0s 2ms/step - loss: 41165.7852 - mae: 41165.7852 - val_loss: 29477.8281 - val_mae: 29477.8301
Epoch 4/50
186/186 [==============================] - 0s 2ms/step - loss: 34673.3906 - mae: 34673.3906 - val_loss: 28654.9004 - val_mae: 28654.9004
Epoch 5/50
186/186 [==============================] - 0s 2ms/step - loss: 32903.0664 - mae: 32903.0664 - val_loss: 27799.8301 - val_mae: 27799.8301
Epoch 6/50
186/186 [==============================] - 0s 1ms/step - loss: 31716.4629 - mae: 31716.4629 - val_loss: 27242.5273 - val_mae: 27242.5273
Epoch 7/50
186/186 [==============================] - 0s 1ms/step - loss: 30412.2520 - mae: 30412.2520 - val_loss: 27105.2617 - val_mae: 27105.2617
Epoch 8/50
186/186 [==============================] - 0s 1ms/step - loss: 29767.8965 - mae: 29767.8965 - val_loss: 26994.0156 - val_mae: 26994.0156
Epoch 9/50
186/186 [==============================] - 0s 1ms/step - loss: 28266.7305 - mae: 28266.7305 - val_loss: 26533.8516 - val_mae: 26533.8516
Epoch 10/50
186/186 [==============================] - 0s 811us/step - loss: 27163.7559 - mae: 27163.7559 - val_loss: 26768.1641 - val_mae: 26768.1641
Epoch 11/50
186/186 [==============================] - 0s 817us/step - loss: 26788.7832 - mae: 26788.7832 - val_loss: 27010.9316 - val_mae: 27010.9297
Epoch 12/50
186/186 [==============================] - 0s 876us/step - loss: 25233.3105 - mae: 25233.3105 - val_loss: 27389.0918 - val_mae: 27389.0918
Epoch 13/50
186/186 [==============================] - 0s 812us/step - loss: 25205.2305 - mae: 25205.2305 - val_loss: 26945.5293 - val_mae: 26945.5293
```

```
Epoch 14/50
186/186 [==============================] - 0s 815us/step - loss: 24738.4336 - mae: 24738.4336 - val_loss: 27086.1328 - val_mae: 27086.1328
Epoch 15/50
186/186 [==============================] - 0s 817us/step - loss: 23294.9863 - mae: 23294.9863 - val_loss: 27241.1504 - val_mae: 27241.1504
Epoch 16/50
186/186 [==============================] - 0s 820us/step - loss: 23065.3164 - mae: 23065.3164 - val_loss: 27260.5195 - val_mae: 27260.5195
Epoch 17/50
186/186 [==============================] - 0s 809us/step - loss: 22793.9727 - mae: 22793.9727 - val_loss: 27271.8496 - val_mae: 27271.8496
Epoch 18/50
186/186 [==============================] - 0s 817us/step - loss: 21638.5547 - mae: 21638.5547 - val_loss: 27532.5742 - val_mae: 27532.5742
Epoch 19/50
186/186 [==============================] - 0s 804us/step - loss: 22342.0059 - mae: 22342.0039 - val_loss: 27817.6094 - val_mae: 27817.6094
Epoch 20/50
186/186 [==============================] - 0s 850us/step - loss: 21037.4785 - mae: 21037.4785 - val_loss: 27671.8828 - val_mae: 27671.8828
Epoch 21/50
186/186 [==============================] - 0s 817us/step - loss: 21365.1348 - mae: 21365.1348 - val_loss: 27770.2344 - val_mae: 27770.2383
Epoch 22/50
186/186 [==============================] - 0s 825us/step - loss: 19859.0137 - mae: 19859.0137 - val_loss: 28176.1074 - val_mae: 28176.1055
Epoch 23/50
186/186 [==============================] - 0s 809us/step - loss: 21059.2949 - mae: 21059.2949 - val_loss: 28398.5547 - val_mae: 28398.5547
Epoch 24/50
186/186 [==============================] - 0s 828us/step - loss: 19408.8301 - mae: 19408.8301 - val_loss: 28226.6523 - val_mae: 28226.6523
Epoch 25/50
186/186 [==============================] - 0s 976us/step - loss: 19569.7266 - mae: 19569.7266 - val_loss: 28681.8945 - val_mae: 28681.8945
Epoch 26/50
186/186 [==============================] - 0s 831us/step - loss: 18772.6777 - mae: 18772.6777 - val_loss: 28145.8477 - val_mae: 28145.8477
Epoch 27/50
186/186 [==============================] - 0s 817us/step - loss: 19046.9082 - mae: 19046.9082 - val_loss: 28068.9902 - val_mae: 28068.9922
Epoch 28/50
186/186 [==============================] - 0s 815us/step - loss: 18778.9336 - mae: 18778.9336 - val_loss: 28422.7441 - val_mae: 28422.7441
Epoch 29/50
186/186 [==============================] - 0s 809us/step - loss: 18951.2930 - mae: 18951.2930 - val_loss: 28446.5176 - val_mae: 28446.5156
Epoch 30/50
186/186 [==============================] - 0s 874us/step - loss: 18560.9082 - mae: 18560.9082 - val_loss: 28598.6328 - val_mae: 28598.6328
Epoch 31/50
186/186 [==============================] - 0s 839us/step - loss: 18157.3438 - mae: 18157.3438 - val_loss: 28514.8496 - val_mae: 28514.8496
Epoch 32/50
186/186 [==============================] - 0s 820us/step - loss: 18591.2559 - mae: 18591.2559 - val_loss: 29269.2051 - val_mae: 29269.2031
Epoch 33/50
186/186 [==============================] - 0s 809us/step - loss: 18403.0645 - mae: 18403.0645 - val_loss: 28725.0352 - val_mae: 28725.0352
Epoch 34/50
186/186 [==============================] - 0s 844us/step - loss: 18139.2754 - mae: 18139.2754 - val_loss: 28906.2246 - val_mae: 28906.2246
Epoch 35/50
186/186 [==============================] - 0s 809us/step - loss: 17508.7285 - mae: 17508.7285 - val_loss: 28925.7422 - val_mae: 28925.7422
Epoch 36/50
186/186 [==============================] - 0s 804us/step - loss: 17176.2285 - mae: 17176.2285 - val_loss: 29122.6133 - val_mae: 29122.6133
Epoch 37/50
186/186 [==============================] - 0s 866us/step - loss: 17728.2148 - mae: 17728.2148 - val_loss: 29094.0078 - val_mae: 29094.0078
Epoch 38/50
186/186 [==============================] - 0s 815us/step - loss: 16826.3574 - mae: 16826.3574 - val_loss: 29487.0996 - val_mae: 29487.0996
Epoch 39/50
186/186 [==============================] - 0s 815us/step - loss: 17087.8320 - mae: 17087.8320 - val_loss: 29396.5156 - val_mae: 29396.5176
Epoch 40/50
186/186 [==============================] - 0s 796us/step - loss: 17459.8066 - mae: 17459.8066 - val_loss: 29743.4844 - val_mae: 29743.4844
Epoch 41/50
186/186 [==============================] - 0s 828us/step - loss: 16873.1328 - mae: 16873.1328 - val_loss: 29724.2500 - val_mae: 29724.2500
Epoch 42/50
186/186 [==============================] - 0s 825us/step - loss: 16402.9355 - mae: 16402.9355 - val_loss: 29783.7871 - val_mae: 29783.7871
Epoch 43/50
186/186 [==============================] - 0s 812us/step - loss: 17332.8125 - mae: 17332.8125 - val_loss: 29893.6133 - val_mae: 29893.6133
Epoch 44/50
186/186 [==============================] - 0s 903us/step - loss: 16457.3066 - mae: 16457.3066 - val_loss: 30233.7148 - val_mae: 30233.7148
Epoch 45/50
186/186 [==============================] - 0s 849us/step - loss: 16912.3750 - mae: 16912.3750 - val_loss: 30385.6172 - val_mae: 30385.6172
```

```
Epoch 46/50
186/186 [==============================] - 0s 941us/step - loss: 16748.1934 - mae: 16748.1934 - val_loss: 30417.8770 - val_mae: 30417.8770
Epoch 47/50
186/186 [==============================] - 0s 922us/step - loss: 16450.7539 - mae: 16450.7539 - val_loss: 30353.2422 - val_mae: 30353.2422
Epoch 48/50
186/186 [==============================] - 0s 895us/step - loss: 16759.8359 - mae: 16759.8359 - val_loss: 30165.8379 - val_mae: 30165.8379
Epoch 49/50
186/186 [==============================] - 0s 839us/step - loss: 15972.7998 - mae: 15972.7998 - val_loss: 30454.1875 - val_mae: 30454.1875
Epoch 50/50
186/186 [==============================] - 0s 898us/step - loss: 16437.9219 - mae: 16437.9219 - val_loss: 30153.0918 - val_mae: 30153.0918
```

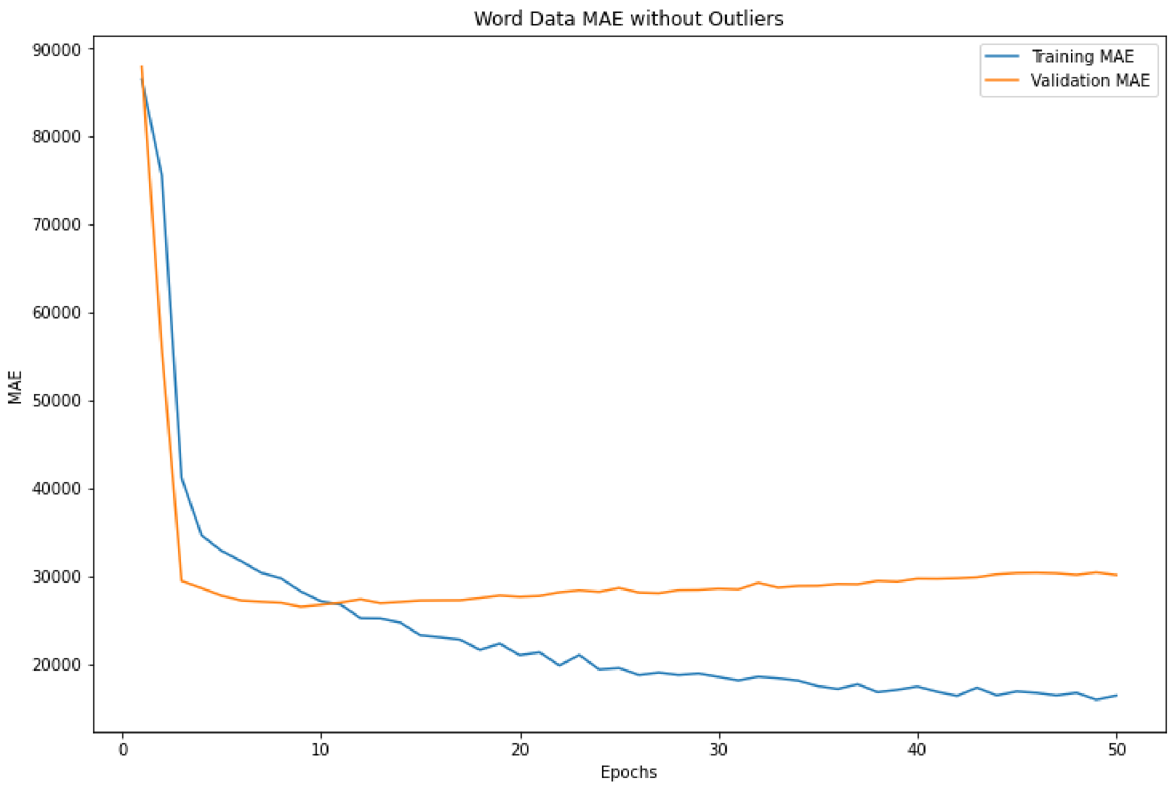
In [35]:

```python
# Plotting Loss
fig, ax = plt.subplots(figsize=(12, 8))

model_dict = history_mae_iqr.history

mae_values = model_dict['mae']
val_mae_values = model_dict['val_mae']

epochs = range(1, len(mae_values) + 1)
ax.plot(epochs, mae_values, label='Training MAE')
ax.plot(epochs, val_mae_values, label='Validation MAE')
plt.legend()
plt.title('Word Data MAE without Outliers')
plt.xlabel('Epochs')
plt.ylabel('MAE')
plt.show()
```



Best Results

In [36]:
```
best_mae_iqr = load_model(r'Data/model_mae_iqr.h5')
errors(best_mae_iqr,mean_absolute_error,X_train_val_iqr,X_test_iqrf,
       y_train_val_iqr,y_test_iqrf,squared=False)
```

```
Train Error: 24838
Test Error: 30935
```

Findings

- Loss on MAE produced the lowest error overall
- Removing outliers improved model performance

Final Stats

In [37]:
```
y_test_preds = best_mae_iqr.predict(np.array(X_test_iqrf))
y_train_preds = best_mae_iqr.predict(np.array(X_train_val_iqr))
```

In [45]:
```
print("Average Views:",round(y_train_val_iqr.mean()))
print("Train MAE:",round(mean_absolute_error(y_train_val_iqr,y_train_preds)))
print("Test MAE:",round(mean_absolute_error(y_test_iqrf,y_test_preds)))
```

```
Average Views: 86653
Train MAE: 24838
Test MAE: 30935
```

Exporting MAE Model without Outliers

In [39]:
```
best_mae_iqr.save(r'Data\model.h5')
best_mae_iqr.save_weights(r'Data\model_weights.h5')
```