Non-Linear Regression with One dimensional Convolutional Neural Network

Akshaykumar Patel

Department of Computer Science

Lakehead University

Thunder Bay,ON,Canada

apatel67@lakeheadu.ca

Abstract - For Non-Linear Regression,we prepared this paper for predicting median house value on the basis on One dimensional convolutional neural network.Main goal was to determine the house value on the basis of a dataset given to us of California Housing dataset which has several feature in it.Using the execution environment i.e.,Google Colab I was trying to get R2 score using Relu(Rectified Linear Unit) activation function on different layers of model.Pytorch libraries are defines the CNN model for the experiment. On the part of results are obtained by changing the hyperparameters and the features(layers) of the model.It is the R2 score that defines the performance of the model.

I. INTRODUCTION

Regression analysis in which non linear parameters are used for modelling oberservational data & depends on more than one independent variable can be termed as Non-linear regression.Convolutional neural network when applied to get visual imagery results in R2 score and L1 loss for a particular model of one dimensional.Convolutional neural networks (CNNs) are considered as machine learning tools based on learning data models.A 1D CNN is very effective when you expect to derive interesting features from shorter (fixed-length) segments of the overall data set and where the location of the feature within the segment is not of high relevance.This applies well to the analysis of time sequences of sensor data . It also applies to the analysis of any kind of signal data over a fixed-length period (such as audio signals). Another application is NLP is the work given to us that is house value prediction for a particular area such as California(given dataset).

II. LITERATURE REVIEW

Basically our main goal for implementing is to discover pattern in Natural Language Processing.For discovering patterns our algorithms are classified on the basis on the way of Learning i.e.,Supervised & Unsupervised learning.Regression particularly comes under the category of supervised.Regression mainly focuses on analysis to create a mathematical model that can be used to predict the values of a dependent variable based upon the values of an independent variable. In other words, we use the model to predict the value of Y when we know the value of X[1].This paper includes the steps formed for predicting house value in terms of R2 score and L1 loss.

III. PROPOSED MODEL

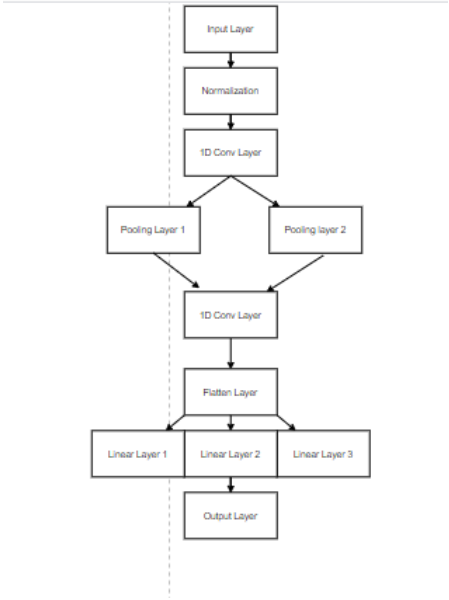
Model is capable of giving the results on the basis of the dataset provided as the input.Dataset used in the model is an Housing Dataset[3].This dataset has various features are longitude,latitude,housing median age,total number of rooms & total number of bedrooms, number of household & median income.Using this feature we extracted the first 10 records of the database as shown in Figure 1.

The first ten records of the dataset are :

	longitude	latitude	housing_median_age	total_rooms	total_bedrooms	population	households	median_income	median_house_value	ocean_proximity
0	-122.23	37.88	41.0	880.0	129.0	322.0	126.0	8.3252	452600.0	NEAR BAY
1	-122.22	37.86	21.0	7099.0	1106.0	2401.0	1138.0	8.3014	358500.0	NEAR BAY
2	-122.24	37.85	52.0	1467.0	190.0	496.0	177.0	7.2574	352100.0	NEAR BAY
3	-122.25	37.85	52.0	1274.0	235.0	558.0	219.0	5.6431	341300.0	NEAR BAY
4	-122.25	37.85	52.0	1627.0	280.0	565.0	259.0	3.8462	342200.0	NEAR BAY
5	-122.25	37.85	52.0	919.0	213.0	413.0	193.0	4.0368	269700.0	NEAR BAY
6	-122.25	37.84	52.0	2535.0	489.0	1094.0	514.0	3.6591	299200.0	NEAR BAY
7	-122.25	37.84	52.0	3104.0	687.0	1157.0	647.0	3.1200	241400.0	NEAR BAY
8	-122.26	37.84	42.0	2555.0	665.0	1206.0	595.0	2.0804	226700.0	NEAR BAY
9	-122.25	37.84	52.0	3549.0	707.0	1551.0	714.0	3.6912	261100.0	NEAR BAY

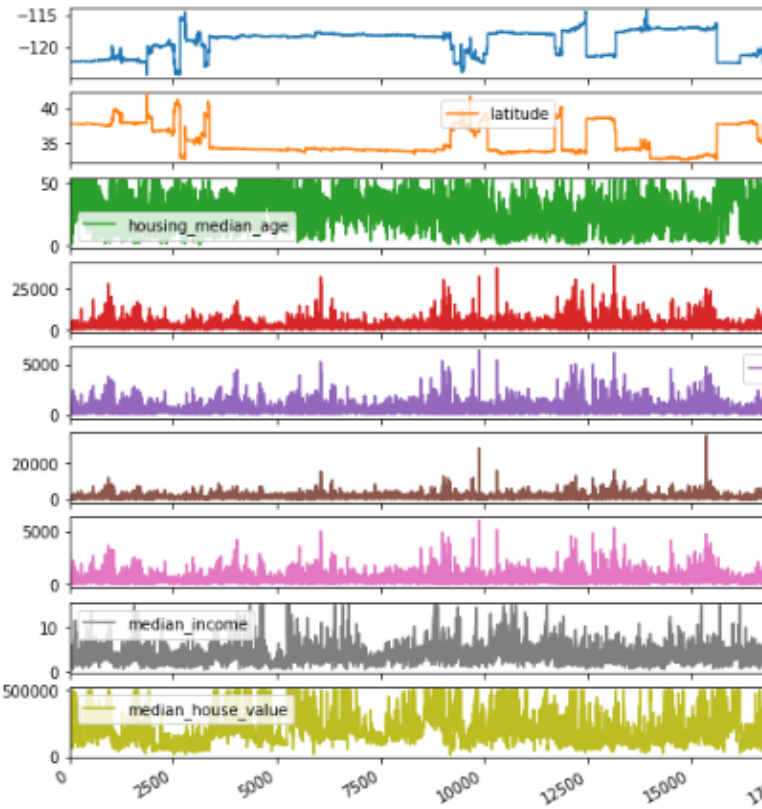
Fig. 1 shows the first ten records for the housing dataset each with different values.

Using this features we selected input and output for the model.Input for the model provided as longitude and median income and output we want as median house value.Our goal was to get the best R2 score & mean squared error(MSE) . Our task was to adjust the learning rate and epochs as well as the amount of training & testing that was told to us is 0.3 & 0.7 respectively.Various layer were used during the development of the model.Flow cart for the layered format of model is shown below.



IV. EXPERIMENTAL RESULTS

On the part of experiment results.Our first step was to show the first 10 records from the housing dataset as shown in Figure 1.And along with that we need to display the separate sub-plots for 8 feature of database.I could not justify whether the shown image size of the graph is the perfect one but I tried my best in getting.Below shown is the graph of sub-plot.



Moving onto the last and next step of our task was predicting the median house value in terms of R2 score and L1 loss with learning rate and epochs using the given features in the dataset.I tools 0.00005 as learning rate and train the model with 500 epochs(Fig.2) and I got the result of R2 score to be 0.57.For the naming convention we were suppose to be named as 1116561_1dconv_reg .For the model,I use relu activation function.On the part of average R2 score as shown in the picture below.

```
Epoch497:
        Loss = 57015.71281179933
         R^2Score = 0.5226078152898249
Epoch498:
        Loss = 57029.99735496076
         R^2Score = 0.5273459757750872
Epoch499:
        Loss = 57012.238176149105
         R^2Score = 0.5293905238334705
Epoch500:
        Loss = 57022.72325182175
         R^2Score = 0.5367648902808259
```

Fig 2. 500 Epochs used for training the model

```
The model's L1loss is: 55893.3365542763
The model's R^2score is: 0.5665852503517
```

V. CONCLUSION

Hence,this paper includes the work of implementing a model based on housing dataset for Non linear regression with 1D Convolutional neural network.Classification and regression are the supervised learning which can help us in many application of Natural language processing.One of the application is House value prediction in the form of Mean Square error and R2 score.

VI. REFERENCES

[1] http://math.hws.edu/javamath/ryan/Regression.html

[2] https://machinelearningmastery.com/supervised-and-unsupervised-machine-learning-algorithms/

[3] https://github.com/ageron/handson-ml/tree/master/datasets/housing

[4] https://cs.stanford.edu/~ermon/cs325/slides/ml_nonlin_reg.pdf