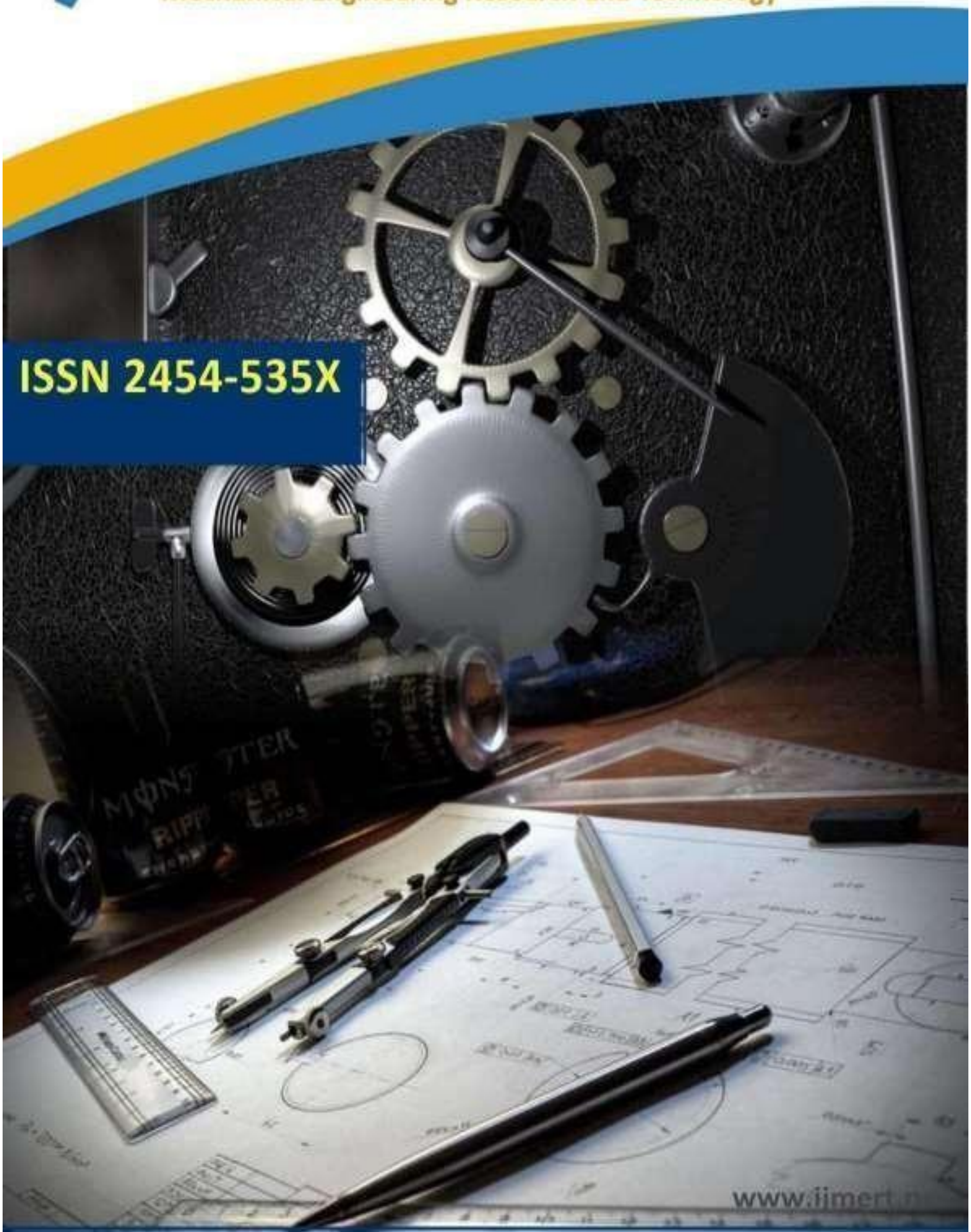
International Journal of
Mechanical Engineering Research and Technology

ISSN 2454-535X



Email ID: info.ijmert@gmail.com or editor@ijmert.net

Int. J. Mech. Eng. Res. & Tech 2017

ISSN 2454 – 5X

http://www.ijmert.com
Vol. 15 Issue 1, Jan 2023

VISUALVALUE RNN: REAL ESTATE APPRAISAL THROUGH IMAGE ANALYSIS

Mahaboob subhani dudekula[1],Satheesh yarramada[2],Yasani nikhitha[3]

[1]Assistant Professor,M.Tech.,BRILLIANT GRAMMAR SCHOOL EDUCATIONAL SOCIETY'S GROUP OF INSTITUTIONS-INTEGRATED CAMPUS
Abdullapurmet (v), hayath nagar (m), r.r dt. Hyderabad

[2]Assistant Professor,M.Tech.,BRILLIANT GRAMMAR SCHOOL EDUCATIONAL SOCIETY'S GROUP OF INSTITUTIONS-INTEGRATED CAMPUS
Abdullapurmet (V), Hayath Nagar (M), R.R Dt. Hyderabad
Department of CSE (Networks),

[3]UG Students, BRILLIANT GRAMMAR SCHOOL EDUCATIONAL SOCIETY'S GROUP OF INSTITUTIONS-INTEGRATED CAMPUS
Abdullapurmet (V), Hayath Nagar (M), R.R Dt. Hyderabad

ABSTRACT

Real estate appraisal, which is the process of estimating the price for real estate properties, is crucial for both buys and sellers as the basis for negotiation and transaction. Traditionally, the repeat sales model has been widely adopted to estimate real estate price. However, it depends the design and calculation of a complex economic related index, which is challenging to estimate accurately. Today, real estate brokers provide easy access to detailed online information on real estate properties to their clients. We are interested in estimating the real estate price from these large amounts of easily accessed data. In particular, we analyze the prediction power of online house pictures, which is one of the key factors for online users to make a potential visiting decision. The development of robust computer vision algorithms makes the analysis of visual content possible. In this work, we employ a Recurrent Neural Network (RNN) to predict real estate price using the state-of-the-art visual features. The experimental results indicate that our model outperforms several of other state-of-the-art baseline algorithms in terms of both mean absolute error (MAE) and mean absolute percentage error (MAPE).

I. INTRODUCTION

Real estate appraisal, which is the process of estimating the price for real

International Journal of Mechanical Engineering Research and Technology
ISSN 2454 – 5X
http://www.ijmert.com
Vol. 15 Issue 1, Jan 2023

estate properties, is crucial for both buys and sellers as the basis for negotiation and transaction. Real estate plays a vital role in all aspects of our contemporary society. In a report published by the European Public Real Estate Association

(EPRA http://alturl.com/7snxx), it was shown that real estate in all its forms accounts for nearly 20% of the

economic activity. Therefore, accurate prediction of real estate prices or the trends of real estate prices help governments and companies make informed decisions. On the other hand, for most of the working class, housing has been one of the largest expenses. A right decision on a house, which heavily depends on their judgment on the value of the property, can possibly help them save money or even make profits from their investment in their homes. From this perspective, real estate appraisal is also closely related to people's lives. Current research from both estate industry and academia has reached the conclusion that real estate value is closely related to property infrastructure, traffic, online user reviews and so on. Generally speaking, there are several different types of appraisal values. In particular, we are interested in the market value, which refers to the trade price in a competitive Walrasian auction setting. Today, people are likely to trade through real estate brokers, who provide easy access online websites for browsing real estate property in an interactive and convenient way. Fig. 1 shows an example of house listing from Realtor (http://www.realtor.com/), which is the largest real estate broker in North America. From the figure, we see that a typical piece of listing on a real estate property will introduce the infrastructure data in text for the house along with some pictures of the house. Typically, a buyer will look at those pictures to obtain a general idea of the overall property in a selected area before making his next move. Traditionally, both real estate industry professionals and researchers have relied on a number of factors, such as economic index, house age, history trade and neighborhood environment [5] and so on to estimate the price. Indeed, these factors have been proved to be related to the house price, which is quite difficult to estimate and sensitive to many different human activities. Therefore, researchers have devoted much effort in

International Journal of Mechanical Engineering Research and Technology
ISSN 2454 – 5X
http://www.ijmert.com
Vol. 15 Issue 1, Jan 2023

building a robust house price index. In addition, quantitative features including Area, Year, Stores, Rooms and Centre are also employed to build neural network models for estimating house prices. However, pictures, which is probably the most important factor on a buyer's initial decision making process, have been ignored in this process. This is partially due to the fact that visual content is very difficult to interpret or quantify by computers compared with human beings. A picture is worth a thousand words. One advantage with images and videos is that they act like universal languages. People with different backgrounds can easily understand the main content of an image or video. In the real estate industry, pictures can easily tell people exactly how the house looks like, which is impossible to be described in many ways using language. For the given house pictures, people can easily have an overall feeling of the house, e.g. what is the overall construction style, how the neighboring environment looks like. These high-level attributes are difficult to be quantitatively described. On the other hand, today's computational infrastructure is also much cheaper and more powerful to make the analysis of computationally intensive visual content analysis feasible. Indeed, there are existing works on focusing the analysis of visual content for tasks such as prediction [13], [14], and online user profiling [15]. Due to the recently developed deep learning, computers have become smart enough to interpret visual content in a way similar to human beings. Recently, deep learning has enabled robust and accurate feature learning, which in turn produces the state-of-the-art performance on many computer vision related tasks, e.g. digit recognition, image classification, aesthetics estimation and scene recognition. These systems suggest that deep learning is very effective in learning robust features in a supervised or unsupervised fashion. Even though deep neural networks may be trapped in local optima, using different optimization techniques, one can achieve the state-of-the-art performance on many challenging tasks mentioned above. Inspired by the recent successes of deep learning, in this work we are interested in solving the challenging real estate appraisal problem using deep visual features. In particular, for images

International Journal of Mechanical Engineering Research and Technology
ISSN 2454 – 5X
http://www.ijmert.com
Vol. 15 Issue 1, Jan 2023

related tasks, Convolutional Neural Network (CNN) are widely used due to the usage of convolutional layers. It takes into consideration the locations and neighbors of image pixels, which are important to capture useful features for visual tasks. Convolutional Neural Networks have been proved very powerful in solving computer vision related tasks.

We intend to employ the pictures for the task of real estate price estimation. We want to know whether visual features, which are a reflection of a real estate property, can help estimate the real estate price. Intuitively, if visual features can characterize a property in a way similar to human beings, we should be able to quantify the house features using those visual responses. Meanwhile, real estate properties are closely related to the neighborhood. In this work, we develop algorithms which only rely on 1) the neighbor information and 2) the attributes from pictures to estimate real estate property price. To preserve the local relation among properties we employ a novel approach, which employs random walks to generate house sequences. In building the random walk graph, only the locations of houses are utilized. In this way, the problem of real estate appraisal has been transformed into a sequence learning problem. Recurrent Neural Network (RNN) is particularly designed to solve sequence related problems. Recently, RNNs have been successfully applied to challenging tasks including machine translation image captioning [26], and speech recognition. Inspired by the success of RNN, we deploy RNN to learn regression models on the transformed problem. The main contributions of our work are as follows:

To the best of our knowledge, we are the first to quantify the impact of visual content on real estate price estimation. We attribute the possibility of our work to the newly designed computer vision algorithms, in particular Convolutional Neural Networks (CNNs). We employ random walks to generate house sequences according to the locations of each house. In this way, we are able to transform the problem into a novel sequence prediction problem, which is able to preserve the relation among houses. We employ the novel Recurrent Neural Networks (RNNs) to predict real estate properties and achieve accurate results.

International Journal of Mechanical Engineering Research and Technology
ISSN 2454 – 5X
http://www.ijmert.com
Vol. 15 Issue 1, Jan 2023

II. LITERATURE SURVEY

Image Based Appraisal of Real Estate Properties, Quanzeng You, Ran Pang, and Jiebo Luo,Fellow, IEEE, Real estate appraisal, which is the process of estimating the price for real estate properties, is crucial for both buyers and sellers as the basis for negotiation and transaction. Traditionally, the repeat sales model has been widely adopted to estimate real estate prices. However, it depends on the design and calculation of a complex economic-related index, which is challenging to estimate accurately. Today, real estate brokers provide easy access to detailed online information on real estate properties to their clients. We are interested in estimating the real estate price from these large amounts of easily accessed data. In particular, we analyze the prediction power of online house pictures, which is one of the key factors for online users to make a potential visiting decision. The development of robust computer vision algorithms makes the analysis of visual content possible. In this paper, we employ a recurrent neural network to predict real estate prices using the state-of-the-art visual features. The experimental results indicate that our model outperforms several other state-of-the-art baseline algorithms in terms of both mean absolute error and mean absolute percentage error.

PROBLEM STATEMENT

The problem of real estate appraisal has been transformed into a sequence learning problem. Recurrent Neural Network (RNN) is particularly designed to solve sequence related problems. Recently, RNNs have been successfully applied to challenging tasks including machine translation, image captioning, and speech recognition Inspired by the success of RNN, we deploy RNN to learn regression models on the transformed problem.

III. EXISTING SYSTEM

Current research from both estate industry and academia has reached the conclusion that real estate value is closely related to property infrastructure, traffic, online user reviews and so on. Generally speaking, there are several different types of appraisal values. In particular, we are interested in the

ISSN 2454 – 5X
http://www.ijmert.com
Vol. 15 Issue 1, Jan 2023

market value, which refers to the trade price in a competitive walrasian auction setting Traditionally, both real estate industry professionals and researchers have relied on a number of factors, such as economic index, house age, history trade and neighborhood environment and so on to estimate the price. Indeed, these factors have been proved to be related to the house price, which is quite difficult to estimate and sensitive to many different human activities. The current algorithms are 1). Regression Models and 2). Deep Walk. Regression model has been employed to analyze real estate price index. Recently, the results in Fu et al. show that sparse regularization can obtain better performance in real estate ranking. Thus, we choose to use LASSO which is an l1-constrained regression model, as one of our baseline algorithms. Deep Walk is another way of employing random walks for unsupervised feature learning of graphs. The main approach is inspired by distributed word representation learning. In using Deep Walk, we also use _-neighborhood graph with the same settings with the graph we built for generating sequences for B-LSTM. The learned features are also fed into a LASSO model for learning the regression weights. Indeed, deep walk can be thought as a simpler version of our algorithm, where only the graph structures are employed to learn features. Our framework can employ both the graph structure and other features, i.e. visual attributes, for building regression model.

Disadvantage:

- The existing system is quite difficult to estimate and sensitive to many different human activities. There are lot of difficult works have been done with the existing systems to measure the number of factors such as economic index, house age, history trade and neighborhood environment.
- Current research from both estate industry and academia has reached the conclusion that real estate value is closely related to property infrastructure, traffic online user Reviews and so on.

IV. PROPOSED SYSTEM:

We intend to employ the pictures for the task of real estate price estimation. We want to know whether visual features, which are a reflection of

International Journal of Mechanical Engineering Research and Technology
ISSN 2454 – 5X
http://www.ijmert.com
Vol. 15 Issue 1, Jan 2023

a real estate property, can help estimate the real estate price. Intuitively, if visual features can characterize a property in a way similar to human beings, we should be able to quantify the house features using those visual responses. Meanwhile, real estate properties are closely related to the neighborhood. In this work, we develop algorithms which only rely on 1) the neighbor information and 2) the attributes from pictures to estimate real estate property price To preserve the local relation among properties we employ a novel approach, which employs random walks to generate house sequences. In building the random walk graph, only the locations of houses are utilized. In this way, the problem of real estate appraisal has been transformed into a sequence learning problem. Recurrent Neural Network (RNN) is particularly designed to solve sequence related problems. Recently, RNNs have been successfully applied to challenging tasks including machine translation, image captioning, and speech recognition. Inspired by the success of RNN, we deploy RNN to learn regression models on the transformed problem. The main contributions of our work are as follows:

To the best of our knowledge, we are the first to quantify the impact of visual content on real estate price estimation. We attribute the possibility of our work to the newly designed computer vision algorithms, in particular Convolutional Neural Networks (CNNs). We employ random walks to generate house sequences according to the locations of each house. In this way, we are able to transform the problem into a novel sequence prediction problem, which is able to preserve the relation among houses. We employ the novel Recurrent Neural Networks (RNNs) to predict real estate properties and achieve accurate results.

Advantage:

- A picture is worth a thousand words. One advantage with images and videos is that they act like universal languages. For the given house pictures, people can easily have an overall feeling of the house, e.g. what is the overall construction style, how the neighboring environment looks like. These high-level attributes are difficult to be quantitatively described
- Map Based Location information are most commonly effective than

International Journal of Mechanical Engineering Research and Technology
ISSN 2454 – 5X
http://www.ijmert.com
Vol. 15 Issue 1, Jan 2023

the viewing in raw details. The most accurate details can be viewed in simple steps

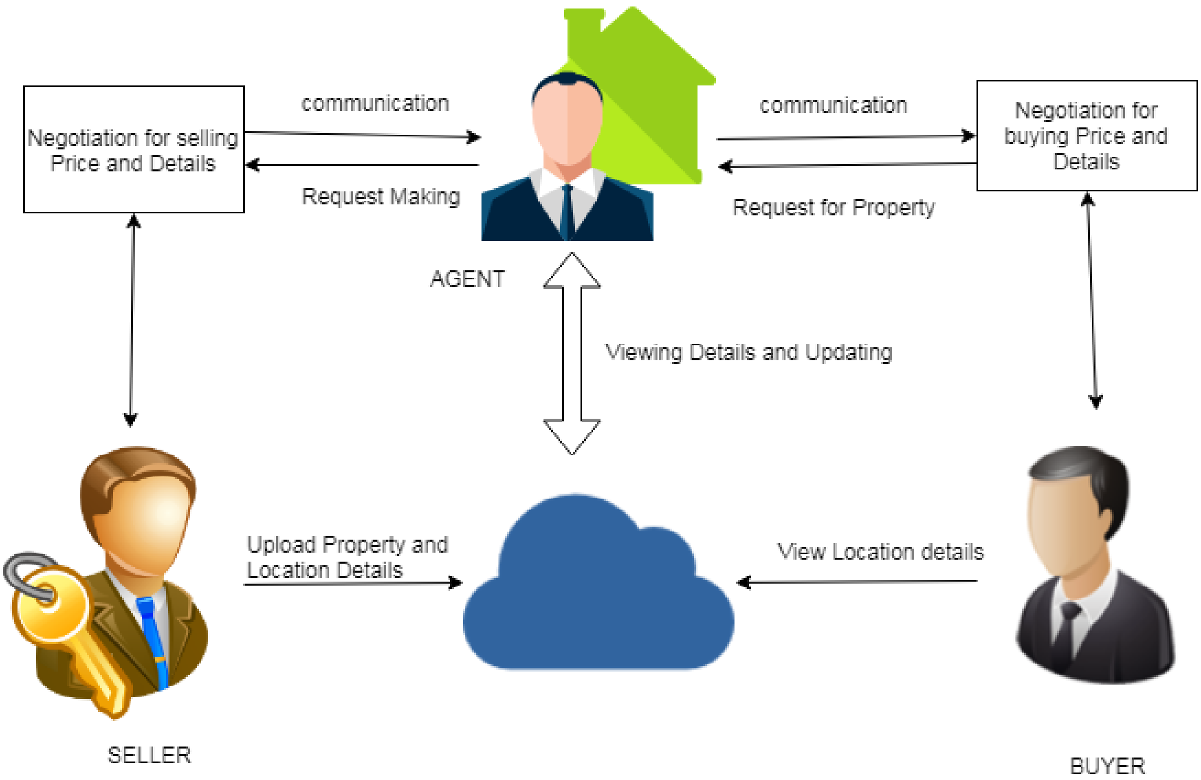


Fig1: Architecture diagram

V. MODULES:

In this project there are four modules present as listed in the below

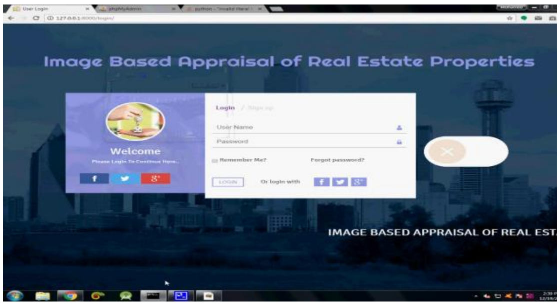
1. Property Addition
2. Adding Location Details
3. Price Negotiation
4. Geometrical Analysis

MODULES DESCRIPTION:

PROPERTY ADDITION

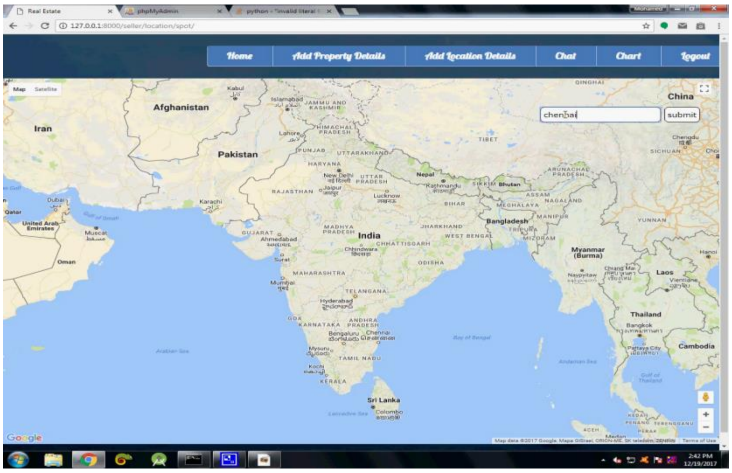
The property addition is the main initiative module for the project. Once authorized user login into the system, they can perform their activity as per their wish. In this module, User must have interested in selling the property which they own. The Property details such as Location, Address, and Facilities that the households are need to add to the cloud where everything that seller uploads can viewable to buyer and agent.

International Journal of Mechanical Engineering Research and Technology
ISSN 2454 – 5X
http://www.ijmert.com
Vol. 15 Issue 1, Jan 2023



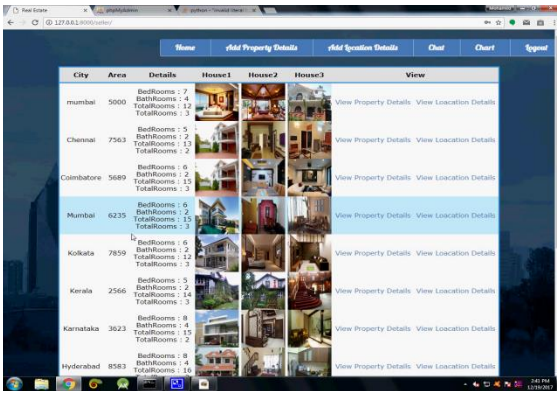
ADDING LOCATION DETAILS

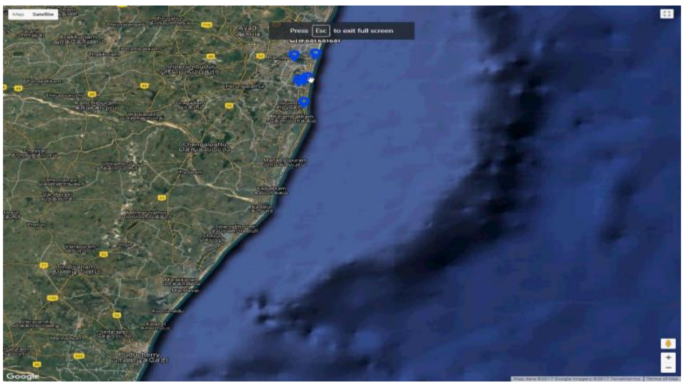
In this module user that is seller need to upload the details of their location as well as their neighboring facility location such as schools, colleges and medical etc., In previous modules also user need to add the location that are into the raw typed format but here in this module we can upload the location details in maps and map formats. Spotting these locations can be very handy for agents or users to get to know about the details of property and neighboring details.



PRICE NEGOTIATION

This module is mainly designed for buyers and agents. Firstly, buyer sends the request to agents along with the cost of expectations and other query details about property. Once agents view the request from the buyer, Agent can decide the price according to the merit of location and both the buyer and seller. This module designed like chat. Dual way communication can be accomplished among the various users.

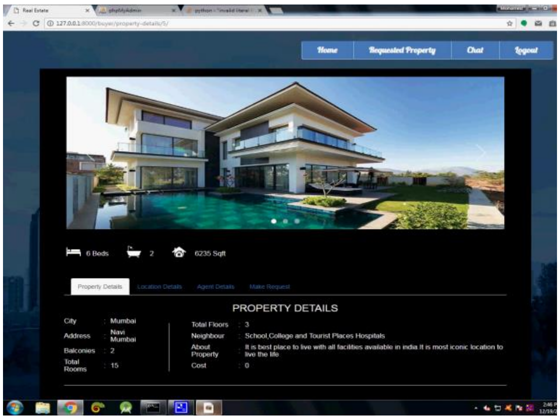




International Journal of Mechanical Engineering Research and Technology
ISSN 2454 – 5X
http://www.ijmert.com
Vol. 15 Issue 1, Jan 2023

GEOMETRICAL ANALYSIS

The Geometrical analysis of given data set is done by charts. Here in this project there are two graphs have been plot between numbers of locations versus city. The pie chart and line charts are established in this project in order to analysis the data effectively.



VI. METHODOLOGY

In this work, all the data are collected from Realtor (http://www.realtor.com/), which is the largest realtor association in North America. We collect data from San Jose, CA, one of the most active cities in U.S., and Rochester, NY, one of the least active cities in U.S., over a period of one year. In the next section, we will discuss the details on how to preprocess the data for further experiments.

Data Preparation

The data collected from Realtor contains description, school information and possible pictures about each real property as shown in Fig. 1 show. We are particularly interested in employing the pictures of each house to conduct the price estimation. We filter out those houses without image in our data set. Since houses located in the same neighborhood seem to have similar price, the location is another important features in our data set. However, after an inspection of the data, we notice that some of the house price is abnormal. Thus, we preprocess the data by filtering out houses with extremely high or low price compared with their neighborhood.

Feature Extraction and Baseline Algorithms

In our implementation, we experimented with Google Net model [43], which is one of the state-of-the-art deep neural architectures. In particular, we use the response from the last avg − pooling layer as the visual features for each image. In this way, we obtain a 1, 024 dimensional feature vector for each image. Each house may have several different pictures on different angles of the same property. We average features

International Journal of Mechanical Engineering Research and Technology
ISSN 2454 – 5X
http://www.ijmert.com
Vol. 15 Issue 1, Jan 2023

of all the images of the same house (also known as average pooling) 2 to obtain the feature representation of the house.

We compare the proposed framework with the following algorithms.

Regression Model (LASSO): Regression model has been employed to analyze real estate price index. Recent show that sparse regularization can obtain better **performance in real estate ranking.**

Deep Walk: Deep walk is another way of employing random walks for unsupervised feature learning of graphs. The main approach is inspired by distributed word representation learning. In using Deep Walk, we also use - neighborhood graph with the same settings with the graph we built for generating sequences for B-LSTM. The learned features are also fed into a LASSO model for learning the regression weights. Indeed, deep walk can be thought as a simpler version of our algorithm, where only the graph structure is employed to learn features. Our framework can employ both the graph structure and other features, i.e. visual attributes, for building regression model.

Training a Multi-layer B-LSTM Model

With the above mentioned similarity graph, we are able to generate sequences using random walks following the steps described in Algorithm 1. For each city, we randomly split the houses into training (80%) and testing set (20%). Next, we generate sequences using random walks on the training houses only to build our **Training sequences for Multi-layer B-LSTM.**

For both cities, we build 200, 000 sequences for training, with a length of 10. Similarly, we also generate testing sequences, where each sequence contain one and only one testing house (see Fig. 4). On the average, we randomly generate 100 sequences for each testing house. The B-LSTM model is trained with a batch size of 1024. In our experimental settings, we set the size of the first hidden layer to be 400 and the size of the second hidden layer to be 200.

Confidence Level

For each testing house, the proposed model can give a group of predictions. We want to know whether or not the proposed model can distinguish the confidence level of its prediction. In particular, we group the testing houses evenly into three groups for each city. The first group has the smallest standard

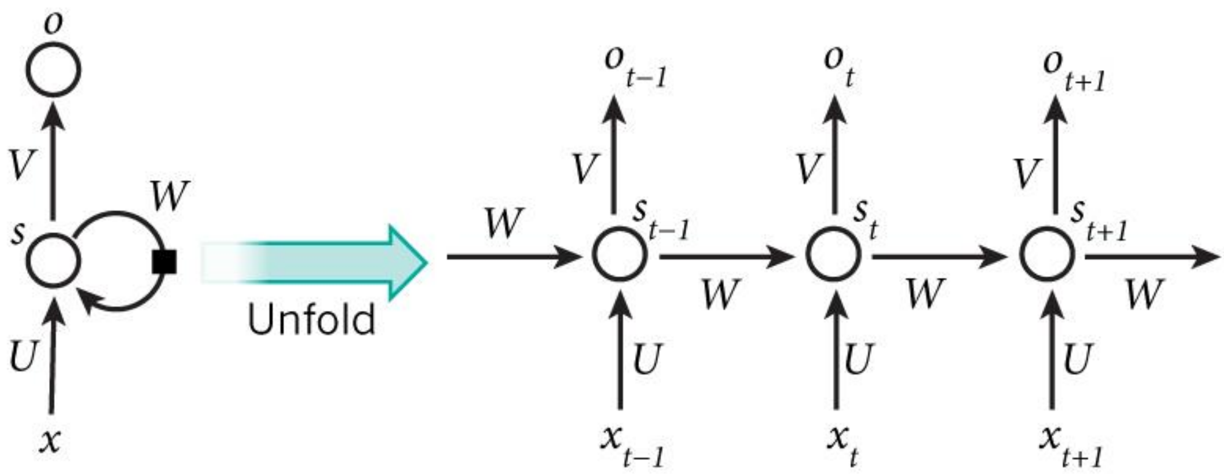
International Journal of Mechanical Engineering Research and Technology
ISSN 2454 – 5X
http://www.ijmert.com
Vol. 15 Issue 1, Jan 2023

deviation of the prediction prices. The second group is the middle one and the last group is the one with the largest standard deviation. The results show that standard deviation can be viewed as a rough measure of the confidence level of the proposed model on the current testing house. Small standard deviation tends to indicate a high confidence of the model and overall it also suggests a smaller prediction error

VII. ALGORITHMS

Recurrent Neural Networks

The idea behind RNNs is to make use of sequential information. In a traditional neural network we assume that all inputs (and outputs) are independent of each other. But for many tasks that's a very bad idea. If you want to predict the next word in a sentence you better know which words came before it. RNNs are called *recurrent* because they perform the same task for every element of a sequence, with the output being depended on the previous computations. Another way to think about RNNs is that they have a "memory" which captures information about what has been calculated so far. In theory RNNs can make use of information in arbitrarily long sequences, but in practice they are limited to looking back only a few steps (more on this later). Here is what a typical RNN looks like:



International Journal of Mechanical Engineering Research and Technology
ISSN 2454 – 5X
http://www.ijmert.com
Vol. 15 Issue 1, Jan 2023

The above diagram shows a RNN being *unrolled* (or unfolded) into a full network. By unrolling we simply mean that we write out the network for the complete sequence. For example, if the sequence we care about is a sentence of 5 words, the network would be unrolled into a 5-layer neural network, one layer for each word.

State Of The Art

State of the art (sometimes **cutting edge**) refers to the highest level of general development, as of a device, technique, or scientific field achieved at a particular time. It also refers to such a level of development reached at any particular time as a result of the common methodologies employed at the time.

VIII. CONCLUSION:

In this work, we propose a novel framework for real estate appraisal. In particular, the proposed framework is able to take both the location and the visual attributes into consideration. The evaluation of the proposed model on two selected cities suggests the effectiveness and flexibility of the model. Indeed, our work has also offered new approaches of applying deep neural networks on graph structured data. We hope our model can not only give insights on real estate appraisal, but also can inspire others on employing deep neural networks on graph structured data.

IX. SCOPE

This website is an easy to view type of a site. It is simply accessible to anyone who has knowledge on browsing websites and this is an example of a simple engineering which is easy and can be used freely without complexity. This website will provide information about houses and he/she can add houses information and no misuse will be conducted as it has some eligible criteria and which has to be taken care off, so the chances are less.

X. REFERENCES

1. Y. Fu et al., "Exploiting geographic dependencies for real estate appraisal: A mutual perspective of ranking and clustering", *Proc. 20th ACM SIGKDD Int. Conf. Knowl. Discovery Data Mining*, pp. 1047-1056, 2014.

2. K. Wardrip, "Public transits impact on housing costs: A review of the literature", *Center for Housing Policy*,

International Journal of Mechanical Engineering Research and Technology
ISSN 2454 – 5X
http://www.ijmert.com
Vol. 15 Issue 1, Jan 2023

Aug. 31 2011, [online] Available: http://www.reconnectingamerica.org/resource-center/browse-research/2011/public-transit-s-impact-on-housing-costs-a-review-of-the-literature/SearchForm/?Search=Keith+Wardrip.

3. Y. Fu et al., "Sparse real estate ranking with online user reviews and offline moving behaviors", *Proc. IEEE Int. Conf. Data Mining*, pp. 120-129, 2014.

4. A. Beja and M. B. Goldman, "On the dynamic behavior of prices in disequilibrium", *J. Finance*, vol. 35, no. 2, pp. 235-248, 1980.

5. M. J. Bailey, R. F. Muth and H. O. Nourse, "A regression method for real estate price index construction", *J. Amer. Statist. Assoc.*, vol. 58, no. 304, pp. 933-942, 1963.

6. R. Meese and N. Wallace, "Nonparametric estimation of dynamic hedonic price models and the construction of residential housing price indices", *Real Estate Econ.*, vol. 19, no. 3, pp. 308-332, 1991.

7. S. T. Anderson and S. E. West, "Open space residential property values and spatial context", *Region. Sci. Urban Econ.*, vol. 36, no. 6, pp. 773-789, 2006.

8. C. H. Nagaraja et al., "An autoregressive approach to house price modeling", *Annal. Appl. Statist.*, vol. 5, no. 1, pp. 124-149, 2011.

9. T. Lasota, Z. Telec, G. Trawiński and B. Trawiński, "Empirical comparison of resampling methods using genetic fuzzy systems for a regression problem" in Intelligent Data Engineering and Automated Learning-IDEAL, New York, NY, USA:Springer, pp. 17-24, 2011.

10. O. Kempa, T. Lasota, Z. Telec and B. Trawiński, "Investigation of bagging ensembles of genetic neural networks and fuzzy systems for real estate appraisal" in Intelligent Information and Database Systems, New York, NY, USA:Springer, pp. 323-332, 2011.

11. W. Di, N. Sundaresan, R. Piramuthu and A. Bhardwaj, "Is a picture really worth a thousand words?:-On the role of images in e-commerce", *Proc. 7th ACM Int. Conf. Web Search Data Mining*, pp. 633-642, 2014.

12. X. Jin, A. Gallagher, L. Cao, J. Luo and J. Han, "The wisdom of social multimedia: Using flickr for prediction and forecast", *Proc. Int. Conf. Multimedia*, pp. 1235-1244, 2010.

13. Q. You, L. Cao, Y. Cong, X. Zhang and J. Luo, "A multifaceted approach to social multimedia-based prediction of elections", *IEEE Trans. Multimedia*, vol. 17, no. 12, pp. 2271-2280, Dec. 2015.

14. Q. You, S. Bhatia and J. Luo, "A picture tells a thousand words?About you! user interest profiling from user

International Journal of Mechanical Engineering Research and Technology
ISSN 2454 – 5X
http://www.ijmert.com
Vol. 15 Issue 1, Jan 2023

generated visual content", *Signal Process.*, vol. 124, pp. 45-53, 2016.

15. Y. LeCun et al., "Backpropagation applied to handwritten zip code recognition", *Neural Comput.*, vol. 1, no. 4, pp. 541-551, 1989.

16. G. E. Hinton, S. Osindero and Y.-W. Teh, "A fast learning algorithm for deep belief nets", *Neural Comput.*, vol. 18, no. 7, pp. 1527-1554, 2006.

17. D. C. Cireşan, U. Meier, J. Masci, L. M. Gambardella and J. Schmidhuber, "Flexible high performance convolutional neural networks for image classification", *Proc. 22nd Int. Joint Conf. Artif. Intell.*, pp. 1237-1242, 2011.

18. A. Krizhevsky, I. Sutskever and G. E. Hinton, "ImageNet classification with deep convolutional neural networks", *Proc. 25th Int. Conf. Neural Inf. Process. Syst.*, pp. 1097-1105, 2012.

19. X. Lu, Z. Lin, H. Jin, J. Yang and J. Z. Wang, "Rapid: Rating pictorial aesthetics using deep learning", *Proc. 22nd ACM Int. Conf. Multimedia*, pp. 457-466, 2014.

20. B. Zhou, A. Lapedriza, J. Xiao, A. Torralba and A. Oliva, "Learning deep features for scene recognition using places database", *Proc. 27th Int. Conf. Neural Inf. Process. Syst.*, pp. 487-495, 2014.

21. Y. Bengio, "Practical recommendations for gradient-based training of deep architectures" in Neural Networks: Tricks of the Trade, New York, NY, USA:Springer, pp. 437-478, 2012.O. Vinyals, A. Toshev, S. Bengio and D. Erhan, "Show and tell: A neural image caption generator", *Proc. IEEE Conf. Comput. Vis. Pattern Recog.*, pp. 3156-3164, 2015.

22. Y. LeCun, L. Bottou, Y. Bengio and P. Haffner, "Gradient-based learning applied to document recognition", *Proc. IEEE*, vol. 86, no. 11, pp. 2278-2324, Nov. 1998.

23. D. Bahdanau, K. Cho and Y. Bengio, "Neural machine translation by jointly learning to align and translate", *Proc. Int. Conf. Learn. Represent.*, 2014.

24. O. Vinyals, A. Toshev, S. Bengio and D. Erhan, "Show and tell: A neural image caption generator", *Proc. IEEE Conf. Comput. Vis. Pattern Recog.*, pp. 3156-3164, 2015.

25. A. Graves, A.-R. Mohamed and G. Hinton, "Speech recognition with deep recurrent neural networks", *Proc. IEEE Int. Conf. Acoust. Speech Signal Process.*, pp. 6645-6649, 2013.