**E-Veluer (Intelligent Tool to Assist in Making
Smarter Property Related Decisions)**

K.Thayalini

IT15156884

Bachelor of Science Special (honors) Degree in Information
Technology

Department of Information Technology

Sri Lanka Institute of Information Technology

Sri Lanka

2019

E-Veluer (Intelligent Tool to Assist in Making Smarter Property Related Decisions)

K.Thayalini

IT15156884

The dissertation was submitted in partial fulfilment of the requirements for
the for the B.Sc. Special Honors degree in Information Technology

Department of Information Technology

Sri Lanka Institute of Information Technology

2019

Declaration

I declare that this is my own work and this Software Requirement Specification entitled Intelligent Tool to Assist in Making Smarter Property Related Decisions, submitted to Sri Lanka Institute of Information Technology is a record of an original work done by me, under the guidance of our supervisor Mrs. M.P.A.W. Gamage. This document does not incorporate without acknowledgement any material previously submitted for a Degree or Diploma in any other University or institute of higher learning and to the best of our knowledge and belief it does not contain any material previously published or written by another person except where the acknowledgement is made in the text.

Name	Student ID	Signature
K.Thayalini	IT15156884	

The above candidate   is   carrying out research for the undergraduate Dissertation under my supervision.

……………………………….   ……………………………..

Signature of the supervisor: Date

(Mrs. M.P.A.W. Gamage)

ABSTRACT

Real Property are the most valuable possession of most of the common people. Getting the proper valuation for these real properties is very much important. This document analyses an innovative solution proposed to facilitate land valuation based on recent sales, prediction of future price and the effect of proposed development work on the land, so that real-estate customers and owners of real estate companies can be benefitted and make smarter property related decisions. This intelligent tool can help people to identify the land they are going to buy in terms of current value and future value. Deep learning, and optimization are the main research components of this system. The system utilizes LSTM model as well as KNN and MLR model in making predictions. LSTM model can make predictions with an accuracy of over 0.75 in current value prediction and also future value predictions with reasonable accuracy. This paper discusses about the research methodology we have used in identifying the most suitable algorithms which can serve our intended purpose.

ACKNOWLEDGEMENT

Contents

LIST OF FIGURES

LIST OF TABLE

LIST OF ABBREVIATINS

Abbreviation	Description
SRS	**System Requirement Specification**
ROI	**Return of Investment**
RNN	**Recurrent Neural Networks**
ANN	**Artificial Neural Networks**
AI	**Artificial Intelligence**

1 INTRODUCTION

1.1 Background Context (Literature Survey)

Real Property are the most valuable possession of most of the common people Getting the best possible valuation for these genuine properties is especially significant. This framework is an imaginative arrangement proposed to encourage land valuation dependent on past deals, expectation of future cost and the impact of proposed advancement deal with the land, so land clients and proprietors of land organizations can be profited and settle on more astute property related choices.

land esteems can be influenced by different social factors as well. For instance, if there is a wrongdoing occurred in that land, it can cause a negative impact on the worth. Hence, real estate appraisal it is a challenging multidimensional problem that involves estimating many facets of a property, its neighborhood, and its city.[1] Since, Sri Lanka is deficient with regards to a decent information stage to assemble every one of these information, considering every one of these variables can take ages to do appropriate valuation thinking about every one of these elements. The manual procedure is a tedious moderate undertaking which should be finished by an accomplished proficient valuer. The valuation methodologies utilized by those experts are restricted because of the absence of computerized information in Sri Lanka.

Analyzing the previous land sale details and trends in those fluctuations and considering those data to predict the valuation is called the sales comparison approach.[3] When the area conservative worth is joined with impact of neighborhood factors, for example, walkability and so on we trust it is conceivable to give an exact, reasonable forecast of the estimation of the land.

The impact of innovation on day by day life of the Sri Lankans has expanded colossally. Individuals will in general use traffic information, web based shopping like never before. Since the manual procedure is excessively moderate and ward to settle on a snappy better choice of the value of the land and appropriateness of it with the end goal of the client, our endeavour is to carefully help the individuals in property

related basic leadership by giving them precise forecasts of the qualities and future investigations of the land.

1.HousePrice.ai

Creating a methodology that would bring more sophisticated information, greater accuracy and analytical rigor to the United Kingdom (UK) residential property market is the motivation behind HousePrice.ai. Their proprietary model provides a combination of multi-disciplinary experiences of AI and Big Data to provide most accurate estimations. HousePrice.ai has Horizon app, which calculates capital, rental and gross development values for a single property or an entire portfolio. [2]

Features-

● Current and Future value prediction - Produces accurate property valuations both in the present time and can offer future predictions. Valuations are based on objective measurable values, creating a fact-based result as opposed to a subjective one [3]. This tool allows the user to adjust, add and remove factors within the surrounding areas to determine how external changes will affect property prices

● Distance to Schools, commutes etc.
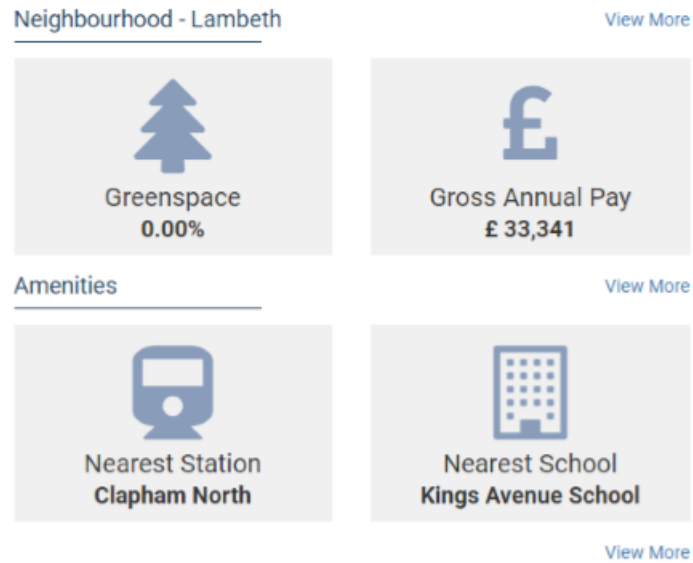
Figure 1 Brief Neighbourhood analysis

Source : Sample Valuation Report - HousePrice.ai , Horizon
https://myhorizon.io/valueReport?id=59ddcdc7a699d278745b81e1

☐ Historical data relevant to location

Figure 2 Historical Sales analysis

Source : Sample Valuation Report - HousePrice.ai , Horizon
https://myhorizon.io/valueReport?id=59ddcdc7a699d278745b81e1

Comparison of Existing Systems

Table 1Comparison of existing systems

	Zillow	Trulia	QV-CoreLogic	HousePrice.ai	Our Product
Current Value Prediction	Yes	Yes	Yes	Yes	Yes
Use of AI/ Machine Learning	Yes	Unknown	No	Yes	Yes
Future Value forecasting	For 12 months	No	Yes, with E-Valuer report	For 3 years	For 5 years
Future Development effects prediction	No	No	Yes	No	Yes
Available for Sri Lanka	No	No	No	No	Yes

1.2 Research Gap

There are such huge numbers of frameworks like our framework in abroad. We will likely distinguish the approaches to utilize their hidden strategy in a reasonable way in Sri Lankan setting. we utilize numerous wellsprings of information to entrap the monetary commitment of the area's qualities, for example, walkability and security discernment. We likewise create and discharge a structure ready to now-cast lodging costs from Open information, without the requirement for recorded exchanges.

There are examines that have been led to foresee the Stock costs of Sri Lanka with the utilization of Artificial Intelligence and Machine Learning draws near, tilted An intermittent neural system approach in anticipating day by day stock costs an application to the Sri Lankan securities exchange [4], there is no data with respect to use of AI innovation in land esteem expectation in Sri Lankan setting. Our proposed framework utilizes AI model to do the estimations and gauges. We structure the application as much as economical and practical to the client, rearranged UIs, with precise outcomes which gives a succinct however complete report of the examination of the land.I provide all the estimations and predictions for a fair price which is beneficial in terms of time and cost over the manual method of valuation.

 User Interface should be appealing to the users despite of their educational or social level for an application to be useful. We believe our application will provide services to all categories of users, mostly the buyers, then surveyors, valuation officers etc. We can provide a simple user interface to input the location of the land and ultimately produce the report which includes current estimation, future value prediction, and details of possible development projects of the area which can be referred and understood by almost anyone.

1.3 Research Problem

The main research problem is to develop an automated system to evaluate the land based on its neighbourhood economic value and identify the possible effects of development work on the value of the land in the future. This requirement of a solution to predict the future value came from an expertise. To predict the effect of future development work on a particular land, since Sri Lanka is a developing country, although the rate of development may vary, infrastructure development projects are carried out frequently

Our intention is to provide people with fair accurate prediction of the land they are going to buy, so that they can decide the investment is fruitful for them. We believe this is an area improvement is needed because we can assist people in making decisions related to property, which would be the largest investment most probably in many people's lives.

1.4 Research Objectives

1.4.1 Main objective

The main objective of our research is to develop a portable application which can provide instant report of a selected land parcel which can provide the users with an insight of the land with future value base on future infrastructure development projects.

1.4.2 Specific objectives

Identifying the most accurate algorithm from conventional Multiple Regression Analysis (MRA) and non-conventional Artificial Neural Networks (ANN) in the domain of providing values in the domain of current value prediction following the Sales Comparison Approach

Identifying the effect of proposed development plans on the future price of the selected land plot

● creating a concise yet complete report based on the selected land plot which can be used to assist in making smarter property related decisions.

2 METHODOLOGY

2.1 Pre-processing

The research component is the prediction of future commercial value of a land based on the infrastructure facilities. The prediction is made with relative to the future development projects in the area. This includes the infrastructure facilities such as school, hospitals, highway, and apartment. The commercial value of a land in future is calculated based on the percentage ratio. Each of these infrastructure facilities are given a specific percentage value based on studies made on the effect on land value with emerging facilities. The dataset collection plays a major role here which involves the percentage calculation. The machine learning field is to be used in order to predict the land value. In order to achieve these objectives the machine is trained and tested with the dataset to predict the future commercial value of the land. Now land price 2k after five year near to the land have school, hospital so land value is school 20%+ hospital40% so future land value 2k+20%+40%   likewise. So data pre-processing get the percentage
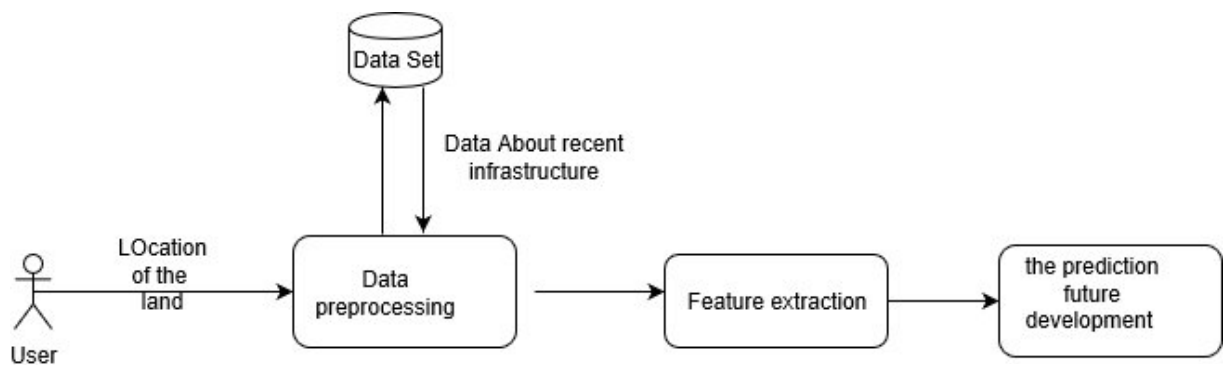

Figure 3 System Diagram for feature prediction

2.2 Prediction

2.2.1 Type of Prediction

In the prediction system there mainly use one algorithm to predict the data that is Neural Network .Here I am using Sequential modal. the purpose of the using one algorithms is to provide a most accreted point. In doing that thing, first of all, need to train the algorithm using future infrastructure dataset using the past information scraped, trained required algorithms.

I use the nearby infrastructure info to estimate the contribution of each structures to the future price. I used the number of schools, hospitals, etc. And their distance from the point of interest to estimate the increase rate of schools, hospitals etc. using machine learning. I used the number of schools and their average distance from the current location. That means I need 10 entries as input and the model gives me 5 entries as output, these outputs are the increase rate of each structures in the future

Prodict the future land commercial value. It is an exceptionally amazing method and can be utilized to comprehend the variables that impact productivity. It very well may be utilized to estimate deals in the coming a long time by examining the business information for earlier months. It can likewise be utilized to increase different experiences about client conduct.

here, Score Model segment that attempts to anticipate the test information. The train model gives a prescient calculation that a score model employments. The Scored Labels segment characterizes the expectation of Sales Amount

```python
import pandas
import matplotlib.pyplot as plt
from tensorflow.python.keras.models import Sequential
from tensorflow.python.keras.layers import Dense, Dropout
from keras import metrics

# load dataset
data_frame = pandas.read_csv("dataset.csv", header=0, sep=",")
dataset = data_frame.values
# split into input (X) and output (Y) variables
X = dataset[:,1:6]
Y = dataset[:,6:]

# Keras Model Configuration
model = Sequential()
model.add(Dense(20, input_dim=X.shape[1], kernel_initializer='normal', activation='relu'))
model.add(Dropout(0.5))
model.add(Dense(20, kernel_initializer='normal', activation='relu'))
model.add(Dense(Y.shape[1], kernel_initializer='normal'))
model.summary()


model.compile(loss='mean_squared_error', optimizer='adam'  , metrics=['accuracy'])


# Neural Network Output
history = model.fit(X, Y, epochs=600, batch_size=50, verbose=1, validation_split=0.2)

print(history.history.keys())
plt.plot(history.history['loss'])
plt.plot(history.history['val_loss'])
plt.title('model loss')
plt.ylabel('loss')
plt.xlabel('epoch')
plt.legend(['train', 'validation'], loc='upper left')
plt.show()

model.save('percent_model.h5')
```

Figure 4 train model

Figure 5 Results

2.2.2 Data collection

Primary data has been collected through interviews and personal visits to land area to know the present situation of the market and the secondary data is collected mainly through various survey department, land estate agents. The data is usefully for assessing the performance of property as a key to future land price.

The study focuses on Colombo which experienced relatively high infrastructure development, which made it possible to become the study area in this study. It aims to determine the condition of the field and facilitate the preparation of observation strategies in the surrounding area which is expected to have a relatively large influence and the most appropriate analysis method to be developed in decision making in this study.

2.3 Design

When a customer goes to a land he is willing to buy, they can input the current location through the application. Based on that location, the suitable recent sales data are selected. Then those data will be analysed by the AI model to predict the future value. That predicted value is optimized to produce the most accurate future value. Then the future value will be predicted by calculates the effect of proposed development projects in the area. All these units generate a report which depicts these two types of data with relevant other data in a simpler way anyone can understand. The application of machine learning, and deep learning algorithms have been tested in of the components

2.4 Finding the Service providers available Geo-location

1) Identifying a device's location

Geo-location for the Web version of our application implemented using Google's Location APIs. It can intelligently manage underlying location technology while meeting various development needs when implementing location-based features.

2) Providing directions

To display point-to-point directions on a map within the application, we used Google Maps API. Registering the app as a routing app then makes exact locations and infra-structure

3) Integrating with mapping software

To implemented location-based service we use Google Maps. Geo-location is the most important technology.

2.5 Choose the service providers

Our application is mainly targeted for real-estate customers and owners of real estate companies. But any person can use this application. This intelligent tool can help people to identify the land they are willing to buy.

This Component forecast of the Future value. Our customer service provides forecasts of the future value of the land, as well as a report that indicates the effect of the development projects proposed in the area.

This proposed system would be of great help in making better property decisions, which is a huge investment and a very important decision in everyone's life.

2.6 Testing and Implementation

2.6.1 Testing

After the implementation procedure the testing of the component is compulsory. Testing eliminate most of the errors in the developed component. The developed system is tested for the selected news domain.

There are 3 types in testing.

2.6.1.1 Unit Testing

Unit testing is performed regarding to each and every individual component. Unit testing is performed under white –box testing approach. Unit testing helps developers to decide that the individual units of the program are working as per requirement and are error free, thus making the overall system error free.

2.6.1.2 Integration Testing

Integration testing is to test errors in the integrated modules. Even if the units of the software are working fine individually, there is a need to find out if the units integrated together would also work without errors. In here the testing will be done in the Black –Box testing approach.

2.6.1.3 System Testing

System testing will be done finally. The software product is compiled as product and then it tested as a whole. This can be accomplished by using one or more of the following tests.

- Functionality Testing:
  - Testing of all the functionalities of the software against the requirement.
- Performance Testing:
  - To estimate how efficient the product in performance testing can be used. The system will be pushed to higher and lower limits and the

defects will be identified. It measures the effectiveness and average time taken by the software to do desired tasks.

- Security and Portability Testing:
    - o This test is done when the software is meant for work on various platforms and accessed by number of persons.

2.6.2 Implementation

System implementation is a major role of Systems Development Life Cycle. Implementation is developing the system according to our plan or design. For implementing the Web application we chose Angular 6 version and to make the models we utilized JUPYTER Note book Machine Learning device.We selected Mongodb as our database since it supports cloud based database activities. By sending the database in a cloud domain we can without much of a stretch put updates cutting-edge in the associated framework. Utilizing those strategies and apparatuses we actualized the framework.

The complex component of Web application is testing. Its critical to have full-developed testing framework so by processing and accessing the whether the functionalities of system satisfies with the actual requirement or not, so recognizing necessities programming testing procedure will work.

```
import { Component, ElementRef, NgZone, OnInit, ViewChild } from "@angular/core";
import {isBoolean} from "util";
import { HttpClient } from "@angular/common/http";
import { HttpHeaders } from "@angular/common/http";
const httpOptions = {
  headers: new HttpHeaders({ 'Content-Type': 'application/json' ,
    'Authorization': 'eyJhbGciOiJIUzI1NiIsInR5cCI6IkpXVCJ9.eyJfaWQiOiI1ZDVhMmYzNTU2M2UyMjBhOTQ5YjBmOTIiLCJlbb
  })
};

declare var google: any;

function test(pl) {
  LandPriceComponent.credentials.lat = pl.geometry.location.lat();
  LandPriceComponent.credentials.lng = pl.geometry.location.lng();
}

function future(av, nu) {
  LandPriceComponent.credentials.av_list.push(av)
  LandPriceComponent.credentials.nu_list.push(nu)
  console.log(LandPriceComponent.credentials.av_list);
  console.log(LandPriceComponent.credentials.nu_list);
}

@Component({
  selector: 'app-land-price',
  templateUrl: './land-price.component.html',
  styleUrls: ['./land-price.component.css']
})
export class LandPriceComponent implements OnInit {

  public static credentials = {
    lat: '',
    lng: '',
    av_list: new Array(),
    nu_list: new Array()
  };
```

Figure 6 Sample code

```
send_current_position(json) {
  // alert(LandPriceComponent.credentials.lat);
  var data = {
    lat: '',
    lng: ''
  };
  data.lat = LandPriceComponent.credentials.lat;
  // alert(data.lat);
  data.lng = LandPriceComponent.credentials.lng;
  // alert(data.lng);
  return this.http.post('users/send_current', data, httpOptions).subscribe((current_price)=>
  {console.log(current_price), this.prices.current_price = parseInt(current_price[0]).toString();});

}
```

Figure 7 Send current position

```
send_data(){
  // alert('OK');
  var data = {
    av: [],
    nu: [],
    current: ''
  };
  data.av = LandPriceComponent.credentials.av_list;
  // alert(data.av[4]);
  data.nu = LandPriceComponent.credentials.nu_list;
  data.current = this.prices.current_price;
  return this.http.post('users/send_future', data, httpOptions).subscribe((future_price)=>
  {console.log(future_price), this.prices.future_price = parseInt(future_price[5]).toString() ,this.prices.school_percent = (future
    this.prices.hospital_percent = (future_price[1]*100).toString().slice(0,5), this.prices.supermarket_percent = (future_price[2]*
    this.prices.apartment_percent = (future_price[4]*100).toString().slice(0,5)
  ;});
}
```

Figure 9 Send data future price

```python
import os
import tensorflow as tf
import sys
import numpy as np
from tensorflow.python.keras.models import load_model

os.environ['TF_CPP_MIN_LOG_LEVEL'] = '3'  # or any {'0', '1', '2'}
tf.compat.v1.logging.set_verbosity(tf.compat.v1.logging.ERROR)
# distance need to be divided by 1000 for normalization.
# distance does not exceed 1000m.
n1 = float(sys.argv[1])
d1 = float(sys.argv[2])/1000
n2 = float(sys.argv[3])
d2 = float(sys.argv[4])/1000
n3 = float(sys.argv[5])
d3 = float(sys.argv[6])/1000
n4 = float(sys.argv[7])
d4 = float(sys.argv[8])/1000
n5 = float(sys.argv[9])
d5 = float(sys.argv[10])/1000
current_price = float(sys.argv[11])

model = load_model('percent_model.h5')
increase_rate = model.predict(np.array([d1, d2, d3, d4, d5]).reshape(1,5))

increase_rate1 = 0.2 + n1*(increase_rate[0,0]-0.2)  # increase rate by school
increase_rate2 = 0.4 + n2*(increase_rate[0,1]-0.4)  # increase rate by hospital
increase_rate3 = 0.3 + n3*(increase_rate[0,2]-0.3)  # increase rate by shop
increase_rate4 = 0.1 + n4*(increase_rate[0,3]-0.1)  # increase rate by road
increase_rate5 = 0.3 + n5*(increase_rate[0,4]-0.3)  # increase rate by apartment

if n1 == 0: increase_rate1 = 0
if n2 == 0: increase_rate2 = 0
if n3 == 0: increase_rate3 = 0
if n4 == 0: increase_rate4 = 0
if n5 == 0: increase_rate5 = 0

total_increase_rate = increase_rate1 + increase_rate2 + increase_rate3 + increase_rate4 + increase_rate5
future_price = current_price * (1 + total_increase_rate)

print(increase_rate1)
print(increase_rate2)
print(increase_rate3)
```

Figure 8 Future prediction and percentage calculation

2.7 Tools and Technologies

For the developer end, we expect a computer with

● CPU: Quad core Processor

● RAM: At least 8 GB

● Storage: 1 TB

Developer End:

● Python (Numpy, Pandas, Matplotlib, scikit-learn)

● Anaconda distribution,

● Angular 6

● Google API

● Elasticsearch (Optional)

3 RESULTS AND DISCUSSION

3.1 Results

The results obtained by testing the above models in future value prediction is discussed here. The specific values for each sector are coming as the output. Then the output will be compared with other machine learning algorithms. Extracted structured data are gained by the fellow members and then the data are fed into the algorithm. After that a comparison runs between these data and price with aid of the algorithm. Hereafter, the output which comes from this algorithm and other algorithm is compared. Finally, the decision is made which the most trustworthy prediction is.

Predicting future values is based on infrastructural facilities. The prediction is made with relative to the future development projects in the area. This includes the infrastructure facilities such as school, hospitals, highway, and apartment. The commercial value of a land in future is calculated based on the percentage ratio

In the prediction system there mainly use one algorithm to predict the data that is MSE the purpose of the using algorithms is to provide a most accreted point. In doing that thing, first of all, need to train the algorithm using future infrastructure dataset using the past information collected. Here, Score Model segment that attempts to anticipate the test information.
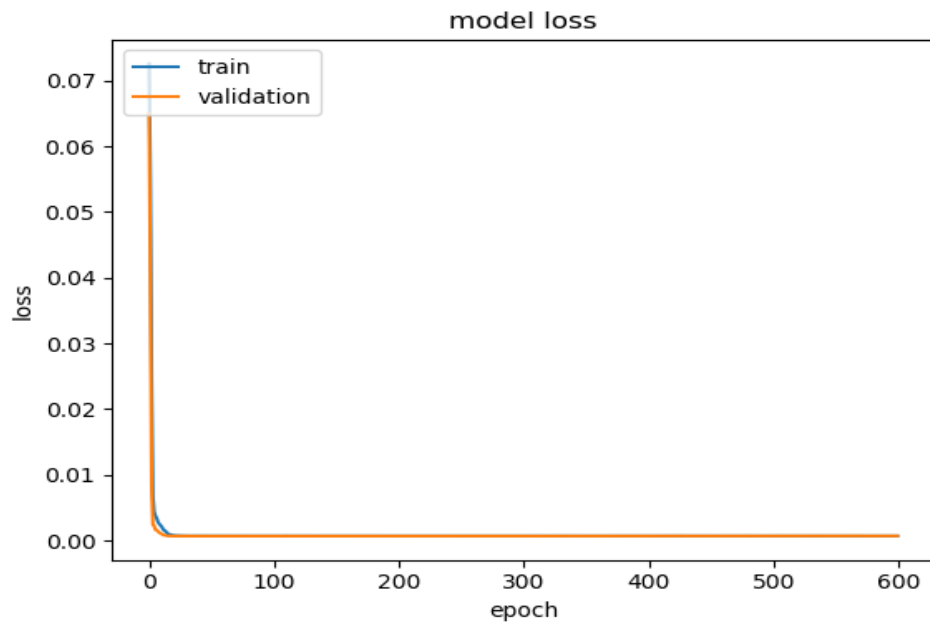
Figure 9: prediction and prediction probability calculation
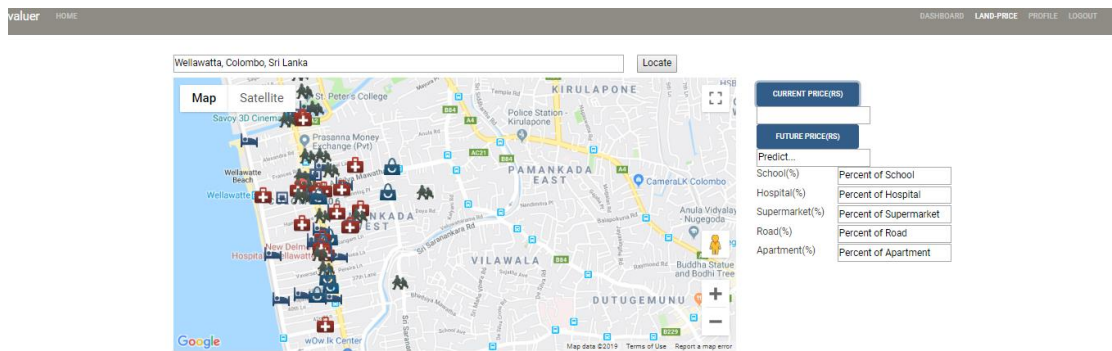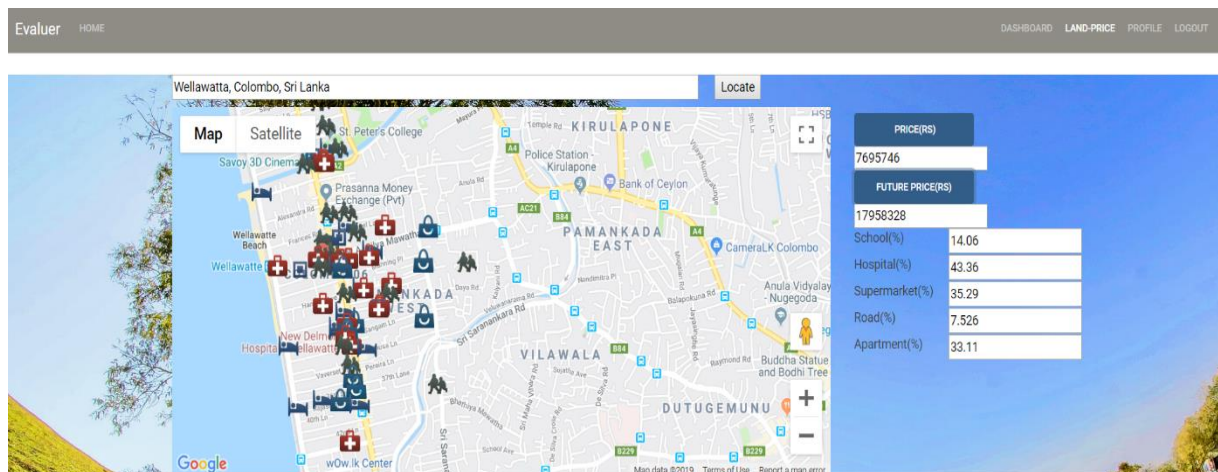
3.1.1 System Interfaces

Figure 10: interface

Figure 11 future price prediction output

3.1.2 Test Cases

Table 2 test case 1

Test Case ID	1
Test case Name	To feed data
Test Input Data	Select the location
Expected Output	Selected location being displayed in to the nearest infrastructural
Actual Output	Selected location being displayed in to the nearest infrastructural

Table 3 test case 2

Test Case ID	2
Test Case Name	To feed data

Test Input Data	Click the Price button
Expected Output	To Display the Land Price
Actual Output	To Display the Land Price

Table 4 test case 3

Test Case ID	3
Test Case Name	Future Price
Test Input Data	Click the future Price button
Expected Output	To Display the future Land Price and Display the infrastructural percentage
Actual Output	To Display the future Land Price and Display the infrastructural percentage

4 CONCLUSION

Based on the observations above, we can conclude that MSE model has the least error among the other tested models and it can achieve an accuracy of around 0.7in predicting future value. Based on the observations above, we can conclude that MSE model has the least error among the other tested models and it can achieve an accuracy of around 0.9 in predicting current value

5 REFERENCE

[1] Nadai, M. D., & Lepri, B. (2018). The Economic Value of Neighborhoods: Predicting Real Estate Prices from the Urban Environment. 2018 IEEE 5th International Conference on Data Science and Advanced Analytics (DSAA). doi:10.1109/dsaa.2018.00043

[2] "ABOUT US." Houseprice.AI-What's the Fair Price ?, www.houseprice.ai/about. [Accessed 24 Feb. 2019].

[3] "Introducing Houseprice.AI: The Must Have Tool for Every Developer." Bridging Loans | Development Loans | AvamoreCapital, 29 May 2018, avamorecapital.com/introducing-houseprice-ai-themust-have-tool-for-every-developer/. [Accessed 24 Feb. 2019].

[4]A. J. P. Samarawickrama and T. G. I. Fernando, "A recurrent neural network approach in predicting daily stock prices an application to the Sri Lankan stock

market," 2017 IEEE International Conference on Industrial and Information Systems (ICIIS), Peradeniya, 2017, pp. 1-6.
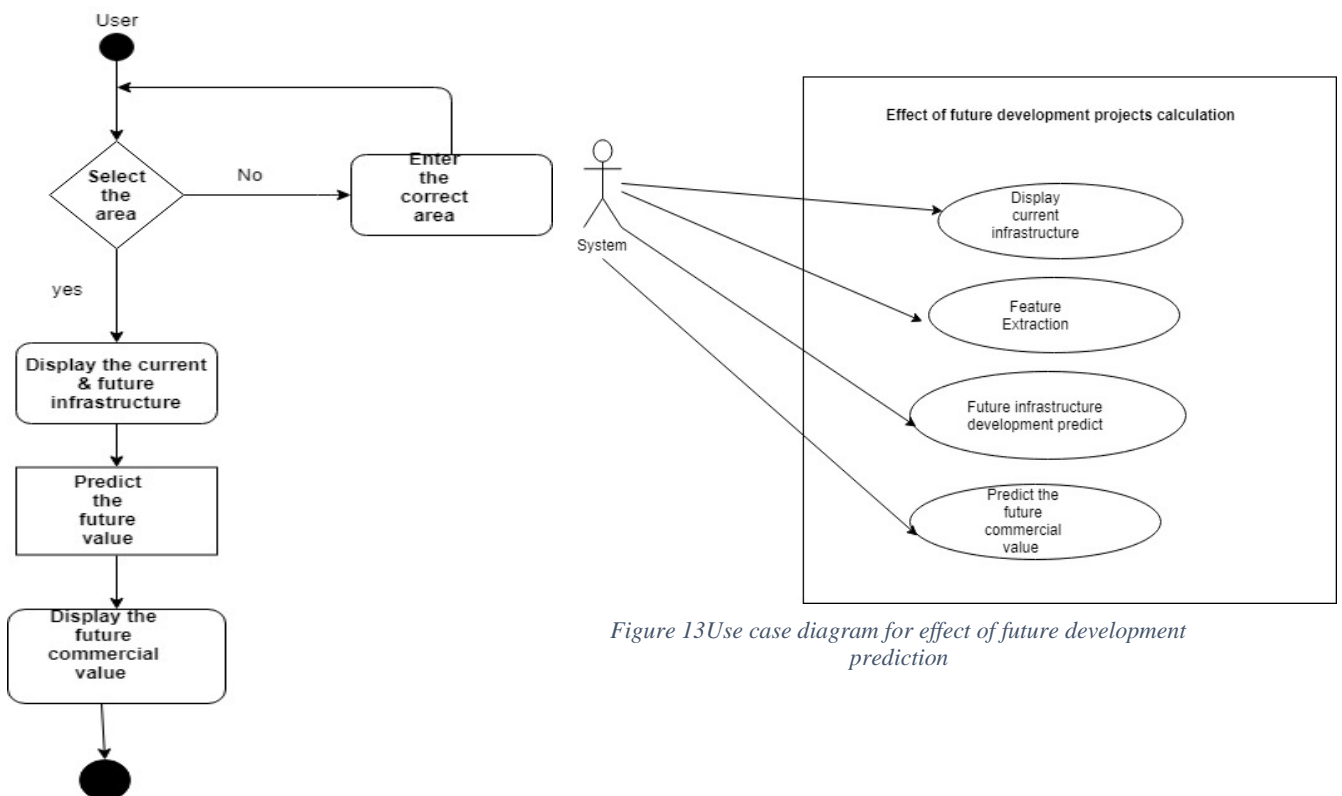

Figure 12 Work flow diagram for effect of future development prediction

Figure 13Use case diagram for effect of future development prediction