Real Estate Analytics With Respect To Andhra Pradesh: Machine Learning Algorithm Using R-Programming

Dr.A.M.Mahaboob Basha, B.Ankaiah, J.Srivani, U.Dadakalander

Abstract: The aim of this empirical research study is to estimate the “value of land” in the study area with respect to eight different factors. The researcher has administered a structure questionnaire with 1100 respondents to analyse the opinion of respondents with respect to “land value” in Andhra Pradesh. The predictive analytics being applied to predict the future growth rate of real estate position in the Andhra Pradesh. The R-Programming concern, multiple linear regression analysis algorithm is applied to test the data in different aspects. The outcome of the research explains about there are certain factors like: security, market, schools, colleges, hospitals, sea, ports, registration value, environment, pollution, facilities will influence the land value in the Andhra Pradesh apart from the decision of present government regarding capital establishment.

Keywords: Real Estate, trends in real estate, public opinion on real estate, real estate analytics, machine learning.

1 INTRODUCTION

The planning commission of India, has defines “Real Estate”, as land, including the air above on it and the ground below it, and any building or structure on it, any building or structure on it Mr. Ashish Mittal&Ms. Khusboo Bhargava in 2014 [1]. It can also be termed as reality in nature. The place contains residential housing, commercial offices, trading spaces such as theatres, hotels and restaurants, retail outlets, industrial buildings and etc. We buy the land at a price and we have the best builders in the competitive market to construct under bond at a fixed cost. As per Selim in 2008 [2] explained about various factors which will impact on real estate housing prices are type of house, type of building, number of rooms available.

2 LITERATURE REVIEW

As per kubat in 2009 [3] explained about various factors which will influence the land value: distance to the sea, the distance to the central business area, distance to the colleges and universities, distance to the health care facilities which will influence the price of the land. Ranjan in 2013[9] witnessed that the transformational change in Indian real estate sector that there is considerable growth in the last few decades, which has potentiality to grow in future. Real estate development (Development services, SDC Companies 2012[4] is a multifaceted that will range from leasing, re-leasing of building, purchase of factors of production, sale of land etc. It is all about bring new thoughts into real property. The real estate regulatory Act (RERA) came into the force across India on 1st July, 2017.

There are numerous state level governing bodies were developed to control real estate business in India. The real estate business basically fall under three folds. They are 1.residential houses, industrial lands, manufacturing buildings, shopping complexes etc. In very recent days, in Andhra Pradesh real estate played a very significant role. The major crucial factor and other factors like: AP State Capital place decision, location, nearness to market, security, registration value, pollution, facilities, environment, plays a significant role while estimating the value of the land. In multiple party systems like India, facing many consequences especially in the real estate sector. The existing government is not able to supply sufficient information to the people related to real property Prashant Das et al, 2013[6]. Even as the Indian real estate required more transparency related to Indian real property Karsten Lieser in 2011[5]. The expansion of Indian real estate sector will strengthen economy of the nation. Considerably, residential sector is under more pressure Ramprakash et al, 2016 in [7]. It is essential to discuss in this context kunal wadhwani in 2009 [8] in his thesis “opportunities and challenges in Indian Real Estate”, identifies the inflow of foreign direct investment in Indian real estate will give positive energy to strengthen Indian economy. Around from the 2009 onwards due to the issues of state division and AP capital, the land value in the AP is stagnated, Aju Thomas in 2015[10].

Objectives of the study:-
1. To study the opinion of the respondents priorities with respect to real estate in Andhra Pradesh.
2. To study the factors which impact on real estate growth in Andhra Pradesh.
3. To suggest best factors which impact on development of real estate in Andhra Pradesh.

3 RESEARCH METHODOLOGY

The research methodology for this research paper is descriptive research design. It is purely depends up on first hand information of the respondents selected in the study area. Data Sources: Both Primary and secondary data sources used for the data collection. Majorly primary data sources are used for data collection. Sample Size: The

---

*Dr.A.M.Mahaboob Basha, Associate Professor, Department of Management studies, Audisankara College of Engineering & Technology, Gudur, drbasha3@gmail.com*

*Mr.B.Ankaiah, Research Scholar, Department of Business Management, Yogi Vemana University, Kadapa, ankaiah.mba@gmail.com*

*Mrs.J.Srivani, Associate Professor, Department of Management Studies, Audisankara College of Engineering & Technology, Gudur, srivani45ree@gmail.com*

*Mr.U.Dadakalander, Research Scholar, Vikarma simhapuri University, Nellore, dadakalander12@gmail.com*
research study majorly depends up on primary data sources. The primary data is collected from various districts of the Andhra Pradesh.

**Software:** R-Programming, machine learning algorithm, multiple linear regression analysis algorithm used to test the data in various aspects.

### 4 DATA ANALYSIS AND INTERPRETATION

**Mapping Graph 1: Train (Vs) Test Data**

The collected total data divide into test data and train data, in the relation of 20:80. To test the value of land in AP, we have taken various parameters like: land registration value, security, market and environment, location, facilities, infrastructure facilities, polution and other factors. The trained data trained with 20% and after tested with 80%. The results of the analysis, showed that, both trained data and test data have shown equal results. Our analysis has received 95% accuracy.

**Graph 2: Correlation plot matrix plot**

Land value = f (location, facilities, infrastructure, polution)

The above scatter Plot Matrix 2 reveals the positive or negative relationship among the various variables. The all variables like: location, facilities, infrastructure, has shown positive and high relationship with the land value in Andhra Pradesh whereas the polution shows negative relationship with the land value. If the polution is high, the landvalue will be lesser vice-versa.

**Correlation Matrix: 3Land value= Registration value+security+market+Environment**

The above scatter Plot Matrix 3 shows relationship among the variables. Where a line which moves from left to right, shows a positive significant relationship with the value of land. The all variables like: registration, security, market, has shown positive and high relationship with the land value in Andhra Pradesh. It depict that, higher the security, nearness to the market, environment conditions shows a positive significant relationship with the land value. There is bit difference with respect to registration value, if higher the registration value, lower the demand of the land.

**Graph 4: Corrgram Graph**

From the above corrgram graph 4 reveals about relationship among the variables. The dark blue color shows highest positive relationship among the variables followed by the sky blue color and light blue colors show
the weak relationship among the variables. The dark red color shows high negative relationship among the variables. In the above diagram environment, market, security, registration value, infrastructure, facilities and locations showing the positive relationship with land value. The pollution only shows the negative relationship with land value. If higher the pollution, lesser the land value vice versa.

**Graph 5: Correlation Matrix Grid**

<table>
<thead>
<tr>
<th></th>
<th>Location</th>
<th>facilities</th>
<th>Infrastructure</th>
<th>pollution</th>
<th>registration value</th>
<th>security</th>
<th>market</th>
<th>environment</th>
<th>land value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Location</td>
<td>1</td>
<td>0.81</td>
<td>0.89</td>
<td>-0.13</td>
<td>0.87</td>
<td>0.87</td>
<td>0.87</td>
<td>0.7</td>
<td>0.98</td>
</tr>
<tr>
<td>facilities</td>
<td>0.81</td>
<td>1</td>
<td>0.82</td>
<td>-0.16</td>
<td>0.76</td>
<td>0.61</td>
<td>0.82</td>
<td>0.98</td>
<td>0.82</td>
</tr>
<tr>
<td>Infrastructure</td>
<td>0.89</td>
<td>0.82</td>
<td>1</td>
<td>-0.13</td>
<td>0.97</td>
<td>0.7</td>
<td>0.99</td>
<td>0.98</td>
<td>0.98</td>
</tr>
<tr>
<td>pollution</td>
<td>0.16</td>
<td>-0.16</td>
<td>-0.13</td>
<td>1</td>
<td>-0.15</td>
<td>-0.07</td>
<td>-0.12</td>
<td>-0.14</td>
<td>-0.13</td>
</tr>
<tr>
<td>registration value</td>
<td>0.87</td>
<td>0.76</td>
<td>0.97</td>
<td>-0.15</td>
<td>1</td>
<td>0.68</td>
<td>0.93</td>
<td>0.97</td>
<td>0.96</td>
</tr>
<tr>
<td>security</td>
<td>0.62</td>
<td>0.61</td>
<td>0.7</td>
<td>-0.09</td>
<td>0.68</td>
<td>1</td>
<td>0.68</td>
<td>0.7</td>
<td>0.68</td>
</tr>
<tr>
<td>market</td>
<td>0.87</td>
<td>0.8</td>
<td>0.96</td>
<td>-0.13</td>
<td>0.93</td>
<td>0.68</td>
<td>1</td>
<td>0.96</td>
<td>0.95</td>
</tr>
<tr>
<td>environment</td>
<td>0.89</td>
<td>0.82</td>
<td>0.99</td>
<td>-0.14</td>
<td>0.97</td>
<td>0.7</td>
<td>0.96</td>
<td>1</td>
<td>0.99</td>
</tr>
<tr>
<td>land value</td>
<td>0.87</td>
<td>0.82</td>
<td>0.98</td>
<td>-0.13</td>
<td>0.96</td>
<td>0.68</td>
<td>0.95</td>
<td>0.99</td>
<td>1</td>
</tr>
</tbody>
</table>

From the above matrix grid 5 reveals the correlation relationship among various variables with respect to land value in the given area. The variable environment shows 0.99 very high correlation with the land value followed by infrastructure has shown 0.98 strong correlation, registration value 0.96, market 0.95 and facilities and location shows 0.82 and 0.87 strong relationship with the land value. In the correlation matrix dark blue colour value shows strong correlation followed by light blue color shows weak correlation, dark red color shows negatively high correlation.

**Graph:6 Residual Fitted plot**

The residual shows the difference between observed value of dependent variable and predicted value. The model shows relationship between the predicted and actual. In the above case the error seems to be less, because the variation between predicted and actual was very less.

**Graph:7 Normal Q-Q Plot**

In general Q-Q plot used to check the normality of the data. The dependent variable land value is showing then values are normally distributed. Using the above plot we can infer if the data comes from normal distribution or not. If yes, the points shows normal distribution. From the above graph it is almost straight line, which shows the data followed normal distribution. Absence of normality the above curve will be either in right skew or left skew.

**Graph 8: Residual Plot**

From the above graph 8 shows that scale-location plot, whether our residuals are spread equally the predictor range ie. homoscedastic. From the above graph it is evident that the residuals are equally spread at 5% level of significance.
Cook's distance attempts to identify the points which have more influence than other points. Such, influencing points will have a sizable impact of the regression line. In other points, adding or removing such points can change the model statistics.

Even it has shown a positive linear regression relationship with the location of the land with respect to land value. If the land is located to near by highway-NH-5, nearness to the market, hospitals and schools. The location plays a very significant role to decide the land value.

Graph 11 reveals that, land value in Andhra pradesh, it is depends up on the facilities available in that area. If the facilities are good, the land value will be higher vice-versa. The above diagram reveals the positive linear regression line relationship with facilities available with respect to land value.

The above Graph 13 reveals that, the land value is having strong positive association with infrastructure facilities available in that area. Here in AP, ceratin areas like: Vijayawada, Guntur, Nellore and Amaravathi is shwoing a strong posiotive relationship with the infrastructre facilities and landvalue.

The above Graph 14 also, reveals the sample results that, there is a significant positive relationship with the land value with respect to registration value in that area. The line has shown the significant positive relationship. From this it can be evident that, the land value is depends upon certain influencing factors like: registration value of the land.
Even the above Graph 15 also reveals about, the significant positive relationship with the nearness to the market. If the household land is nearness to the market and fed-up with all sorts of facilities, it will show the positive linear regression with the land value.

5 FINDINGS
1. The two scatter plot matrices have shown positive impact on the land value.
2. If the land is nearness to the market value and where all facilities are available, will shows a greater impact on the value of the land.
3. If the household land fed-up with good infrastructure facilities, will shows greater impact on value of the land.
4. If it is with good environment conditions, which will have good impact on value of land.
5. The pollution in the area has shown the negative correlation with the land value. If the pollution is higher, the value will be lesser, vice-versa.
6. If the area is fed-up with good facilities, will show the good land value.
7. The corrgram chart also, shows the positive impact of all factors except with the pollution in that area.
8. The security available in that area is also shown the positive relationship with the land value, higher the land value, greater the security, vice-versa.
9. In the literature also, it is evident that, the place of the capital cities also, decides the value of the land.

6 CONCLUSION
Therefore, it can be conclude that, there are certain factors which influence the growth of real estate in Andhra Pradesh. These positive influencing factors like: security, nearness to the market, registration value, and possibility of highways, schools, colleges, and environment will influence the value of the land in the respective areas of the Andhra Pradesh, which decides the scope of future real estate in Andhra Pradesh.

7 REFERENCES
[4]. SDC Companies, Development services. MGT Radiance “Yes...may luck always shine on you”. 2015.
[9]. Soumya Ranjan, Senior Associate, EY, (2013) - Indian real estate: the year gone by and outlook”.