Artificial Intelligence in Real Estate

Abhiroop Ghosh[1], Manan Mer[2], Yash Panchal[3], Silviya D'monte[4]

1,2,3,4Department of Computer Engineering, Universal College of Engineering , Vasai, India

---***---

Abstract— Real estate is one of the most profitable industries on the planet. According to projections, the global value of business acquisitions is expected to reach $9.6 trillion in the next years. However, the real estate industry is still in the early stages of implementing AI and services. AI in real estate will become more visible as technology advances. In the real estate industry, it is critical to consider artificial intelligence and machine learning. In fact, in the last few years, during the COVID pandemic, almost every aspect of real estate was possible through the internet and Artificial Intelligence. This project provides more accuracy and uses feature engineering for faster and more detailed analysis of any estate. A.I. in real estate projects is a website where the investor gets a better evaluation and analysis.

Keywords— Real estate, Artificial Intelligence, Feature engineering, Estate analysis

I. INTRODUCTION

Machine learning is a branch of Artificial Intelligence (AI) that employs algorithms and technologies to extract useful data from large amounts of data. Global pandemic limitations have had a direct influence on traditional real estate processes – and for the better, unexpectedly. Thousands of businesses, realtors, appraisers, mortgage lenders and other businesses have been forced to manage the crisis by incorporating rapidly emerging PropTech, and with good cause. Real estate AI apps can manage predetermined data flows, learn user behaviour streamline, and speed operations, and allow more accurate assessments and market forecasts in the short term. Homeowners, potential renters, and purchasers are embracing real estate AI apps, and investors are aware that real estate is the world's most valuable asset class.

There have been multiple attempts to benefit real estate by using Machine Learning and A.I., but there have been constant issues with the accuracy as well as resulting in the loss of jobs in this sector. A.I. in real estate, this project aims to improve the accuracy whilst also providing a platform for real estate agents does not only make it safer for people to analyze their requirements but also safeguards customers from fake data. This project also integrates multiple modules which takes the accuracy and customer satisfaction to the next level.

II. LITERATURE SURVEY

The research presented in the paper by Adyan Nur Alfiyatin, Hilman Taufiq, Ruth Ema Febrita, and Wayan Firdaus Mahmudy aims to predict the house prices in Malang city based on the prices of NJOP using regression analysis and particle swarm optimization(PSO). PSO is used to select the effected variables, while regression analysis is used to determine the best coefficient for prediction. Several tests have been conducted to predict house prices using linear regression and particle swarm optimization methods. The system is modelling house price predictions into multiple models each of them representing one area. The best parameter values achieved are 1800 particles, 700 iterations, and inertia weights of 0.4 and 0.8, resulting in an RMSE of IDR 14.186 as the minimal prediction error. The error prediction values for the other model remain high.

In the paper presented by R Manjula, Shubham Jain, Sharad Srivastava, and Pranav Rajiv Kher, in this work the models are created using various features such as square feet of the house, the number of bedrooms, ambience, etc. So, each of the features in their model is given a certain weight and it determines how important is that feature to our model prediction. This is called feature engineering. Companies such as "Zillow.com" and "Magicbricks.com" frequently have a vast dataset of housing values that they use to estimate values using machine learning. For each of the models, we calculated the root mean square error (RMS value) as proposed. With this, we have applied the following methodologies: Simple Regression Model, Multivariate Regression model, Polynomial Regression.

In the paper presented by Alisha Kuvalekar, Shivan Manchewar, Sidhika Mahadik, and Shila Jawale, a Decision tree machine learning algorithm is used to construct a prediction model to predict potential selling prices for any real estate property. To help estimate prices even better, additional characteristics such as air quality and crime rate were included in the dataset. These traits are not commonly seen in the datasets of other prediction systems, which distinguishes this system. The Flask Framework is used to integrate the trained model with the User Interface. The system provides 89% accuracy while predicting the prices of real estate prices.

In the paper presented by Quang Truong, Minh Nguyen, Hy Dang, Bo Mei House Price Index (HPI) is used to measure price changes in residential housing in many countries. The HPI is a weighted repeat sales index, which means it tracks average price changes in repeat sales or refinancing on the same properties. It enables housing economists to estimate changes in the rates of mortgage defaults, prepayments, and housing affordability in specific geographic areas using some analytical approaches.

Data Preprocessing, Data Analysis, Model Selection, Random Forest, XGBoost, Light GBM, Hybrid Regression, and Stacked Generalization. The Random Forest method has the lowest error on the training set, but it is susceptible to overfitting. Because the dataset must be fit multiple times, its time complexity is high. When it comes to accuracy, both XGBoost and LightGBM perform well, but their time complexities, particularly LightGBM, are superior. Because of the generalisation, the Hybrid Regression method outperforms the three previous methods. Finally, while the Stacked Generalization Regression method has a complex architecture, it is the best option when accuracy is the most important factor.

III. PROPOSED SYSTEM

In Fig 1., we show the detailed system architecture of Artificial Intelligence in Real Estate. The presented system will be a website in which algorithms can go through millions of documents in seconds, looking through property values, location, home renovations, and even some of a homeowner's personal information. With this hybrid approach of using multiple algorithms as well as regressions, much more accurate data is achieved. The web system supports the buyers as well as the estate agents which ensures there is no loss of human resources and jobs. Investors or buyers have their own unique needs which they can give as inputs to the portal, after which the requirements are compared with the database as well as estate agent inputs to determine the most suitable property. To keep the model efficient feature engineering
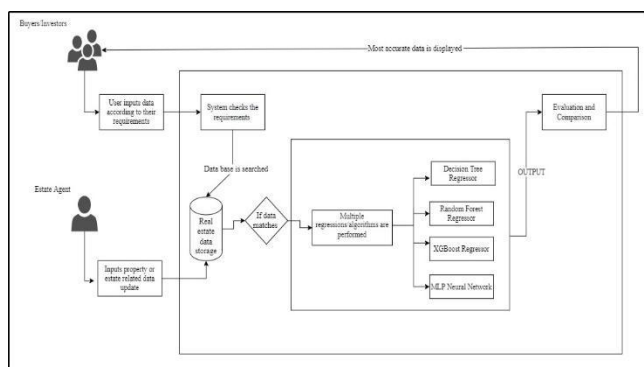

Fig.1 System architecture of A.I. in real estate

Has been used along with AI and ML. Inputs such as bedroom, hall, kitchen, bathroom, square foot, etc are considered as features that are then searched in the data storage. After finding the appropriate and cleaned data multiple algorithms are used such as Decision Tree Regression, Random Forest Regression, XGBoost Regressor, and MLP Neural Network. On receiving the output or results the accuracy is evaluated and compared to find the best possible result, once the evaluation is complete most accurate data is provided to the buyer. Various modules have been integrated to improve accuracy drastically. The modules of this web-based portal are as follows:

A. Requirement/ Feature Input:

Initially, the buyers are required to fill in the required aspects that the property should have. Since feature engineering is used the buyer can be more specific about their needs which can be location, history, facilities, etc.

B. Data searching and processing:

The input features are analyzed and searched within the database for updates or availability. After the data has been parsed it is sent further to undergo multiple regressions and algorithms are applied. As shown in Fig.2, we can see the mechanism of how the user gets the property detail.
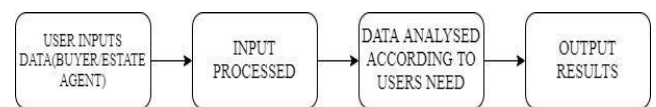

Fig.2 Data searching and application of algorithms

C. Evaluation and Comparison:

Once all the results have been found, the output from each algorithm is compared to check the accuracy and mean absolute error to evaluate the most appropriate data. Since the amount of feature input is more hence various methods have been used to give better results. Additionally, various methods can also be integrated which even further increases the accuracy.

IV. RESULT AND DISCUSSION

The proposed system gives all the features provided by the traditional existing systems, but instead of working only with non-spatial databases, the system also works with spatial data. The following prominent features will be included in the system:

Specification-based searching. This feature displays relevant information to users based on the specifications they have provided to the website. The System will allow the user to search for a property quickly and easily for

buying and selling. The registered user can list his or her property for sale or rent. This system safeguards customers as well as brokers from false advertisements and data.
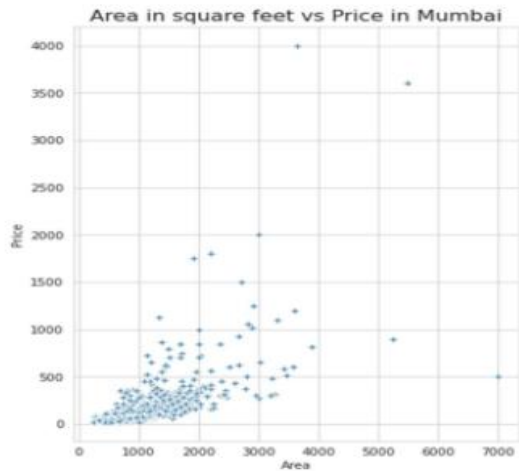

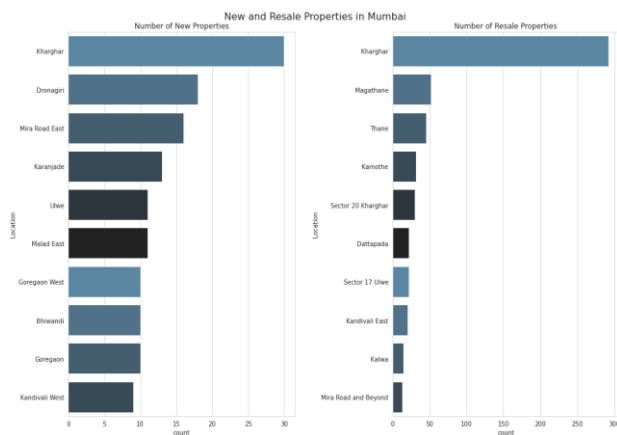Fig.3 Graph of Area in square feet v. Price in the city of Mumbai

Fig.4 Number of new and old properties for sale in areas of Mumbai

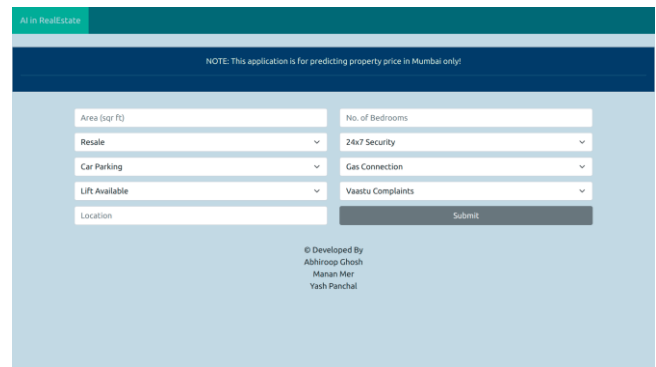Fig.5 Heat map of the city with respective prices
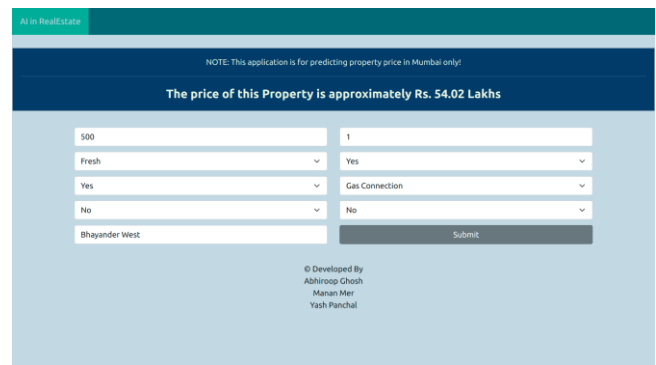

Fig.6 Website - landing page

Fig.5 Website - result page

V. CONCLUSION

After researching AI in Real Estate, we found that skyline AI is the only type of AI used in this sector and this results in a severe reduction in the use of brokers/agents, which is one of the most serious downside of AI. This project will focus on Price prediction using better accuracy using traditional or advanced Machine learning algorithms on real-time data, Fraud detection using days-on-the-market as well as it provides a platform for the brokers/agents to post their updated property price from which the customers can select the most suitable and profitable source. With the rapid digitization of the real estate industry, AI is unavoidable. AI algorithms can find market opportunities for agents to secure more consumers by using predictive models. It may assist clients in predicting the best moment to buy or sell real estate. AI enables you to enter the self-driving dimension, in which AI outsources the hard labour associated with a real estate transaction: complex data, compliance, paperwork, home finding, negotiation, and offers.

VI. REFERENCES

[1] A. Nur Alfiyatin, H. Taufiq, W.F. Mahmudy, Ruth Ema Febrita, "Modeling House Price Prediction using Regression Analysis and Particle Swarm Optimization", IJACSA, 2017.

[2] R Manjula, S Jain, S Srivastava and P.R. Kher, "Real estate value prediction using multivariate regression models", IOP, 2017.

[3] Sayan Putatunda, "ProTech for Proactive Pricing of Houses in classified Advertisement in the Indian Real Estate Market", Research Gate, 2019.

[4] Q. Truong, M. Nguyen, Hy Dang, Bo Mei, "Housing Price Prediction via Improved Machine Learning Techniques", IIKI, 2019.

[5] A. Kuvalekar, S. Manchewar, S. Mahadik, S. Jawale, "House price forecasting using Machine Learning", SSRN, 2020.

[6] R Bhatia, "Housing Prices in Metropolitan Areas of India", Kaggle, 2020, https://www.kaggle.com/ruchi798/housing-prices-eda-and-prediction/data

[7] Shreyas, "House-price-prediction", Github, 2019, https://github.com/Shreyas3108/house-price-prediction