HOUSE PRICE PREDICTION USING RANDOM FOREST ALGORITHM AND PERSONALISED HOUSE RECOMMENDATION WITH ANALYTICAL HEIRARCHY PROCESS METHODOLOGY

Sairam Kumaran & Binny Kaur

Odette School of Business, University of Windsor

Abstract

House price prediction and personalized house recommendation are important tasks in the real estate industry, aiding both buyers and sellers in making informed decision. Many real estate firms have long made decisions based on a combination of intuition and traditional, retrospective data. Today a host of new variables make it possible to paint more vivid pictures of location's future risk and opportunities (Mckinsey,2018).

MagicBricks is a Real Estate company in Canada. Company's online platform provides deep coverage of the real estate market and property trends in major cities of Canada. The platform provides users insights on Tax planning and home loans. The company is expanding its business in Windsor city. Company recently acquired 110 new properties and will the predicting the price of those properties. As analysts at the company MagicBricks, we will be further working for a client to select their dream home. As the client is looking to make the purchase of house. We will be providing them with through consultation first with the price prediction and further knowing their preferences or subjective criteria's we will narrow down the options of houses that will best suit their requirement.

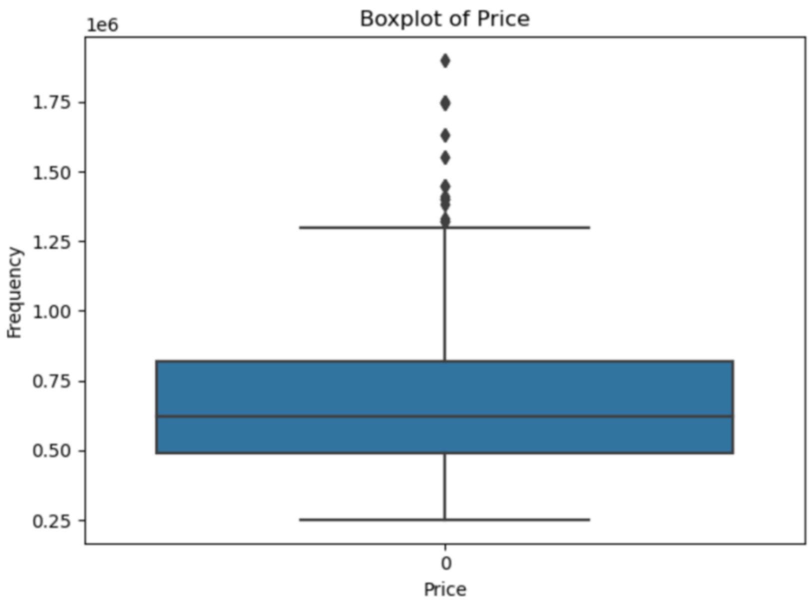
Introduction

Real Estate valuation is process to determine the market price of the property. From data driven decision making to enhanced risk management, increased efficiency, improved customer service, optimal resource allocation, predictive maintenance and sustainable practices, Artificial Intelligence algorithms such as Random Forest transforms how real estate projects are planned and executed. There are various factors that influence the valuation of a property. (LinkedIn, 2023). Some of the factors include location, size, condition, amenities, and recent sales of comparable properties in the area. Our data set specific to the area of study includes total of 546 properties with attributes price, lot size, bedrooms, bathrooms, stories, driveway, recreation, full base, gas heating, air conditioner and garage.

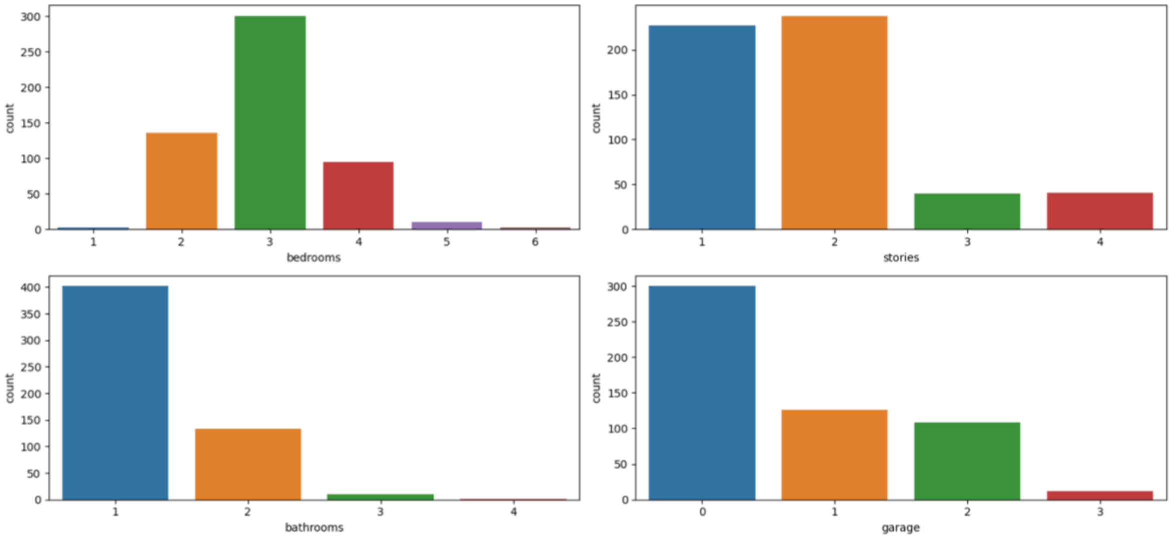
Method/ Process

Descriptive Analysis of Data

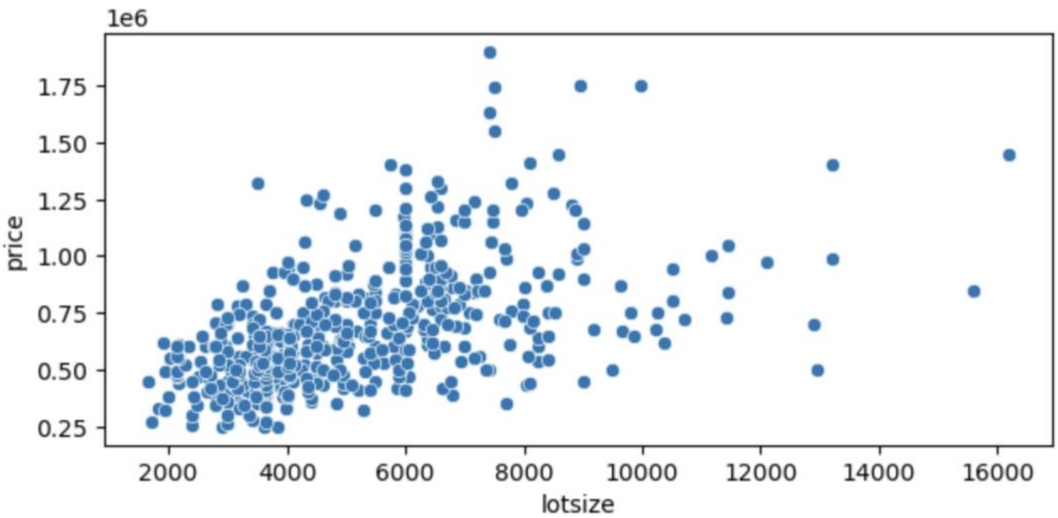
As observed from summary statistics of data, housing prices in the City of Windsor, ranges between 25000 and 190000, with an average of 68120. The distribution of sales price is shown below, which indicates a higher frequency of the houses valued at 50000 and it quickly decreases after 65000. Some outliers were also observed where the houses were valued at a higher price than 150000. As the data is from the year 2016, it felt right to improvise the values by multiplying with 10.



Visualizing different attributes spread across entire property data.



Determining the correlation between lot size and price variables.



Forecasting the price of houses

Random Forest is a supervised machine learning algorithm that is constructed on decision tress algorithm. It can produce reasonable prediction without hyper-parameter tuning. It can effectively handle both numerical and categorical features. It has the capabilities to handle missing values and outliers, which contributes to its robustness.

We used 436 properties as training set and 110 properties as training set. The calculated model score accuracy is 82%. Please refer to Python workbook for code and statistics.

This study aims to build a predictive model for house prices in the city of Windsor, employing Random Forest Regression. The dataset includes features such as price, lot size, bedrooms, bathrooms, stories, driveway, recreation, full basement, gas heating, air conditioning, garage, and client preference. The following steps outline the process of building the prediction model.

Data Collection and Preprocessing:

- Gather real estate data for the city of Windsor, including relevant features and target variable (house prices).
- Check for missing values, outliers, and data inconsistencies.
- Perform data cleaning and imputation as needed.
- Encode categorical variables (e.g., driveway, recreation, full basement) into numerical representations.

Exploratory Data Analysis (EDA):

- Conduct a comprehensive analysis of the dataset to identify patterns and correlations between variables.
- Visualize the distribution of house prices and other features.
- Evaluate the impact of different features on house prices using scatter plots, histograms, and correlation matrices.

Feature Engineering:

- Create new features if relevant to enhance the prediction model's performance.
- Select the most informative features based on correlation analysis and domain knowledge.

Data Splitting:

- Divide the dataset into training and testing sets to evaluate the model's generalization performance accurately.
- Utilize a common ratio, such as 80/20, for the split.

Random Forest Regression Model:

- Implement the Random Forest Regression algorithm, a powerful ensemble learning technique, to predict house prices.
- Fine-tune hyperparameters, including the number of trees and maximum depth, using techniques like cross-validation to optimize the model's performance.

Model Training and Evaluation:

- Train the Random Forest Regression model using the training dataset.
- Evaluate the model's performance on the testing dataset using metrics such as Mean Squared Error (MSE), R-squared (R2), and Mean Absolute Error (MAE).

Model Interpretation:

- Analyze the importance of different features in the prediction process.
- Visualize feature importance using plots, such as bar charts or heatmaps.

Prediction and Recommendation:

- Utilize the final trained model to predict house prices for newly acquired properties in the city of Windsor.
- Present the predictions along with confidence intervals for clients' consideration.

Through the systematic implementation of these steps, the built Random Forest Regression model will provide valuable insights into house prices in Windsor, supporting MagicBricks in making informed decisions and empowering clients with accurate real estate predictions.

	lotsize	bedrooms	bathrooms	stories	garage	driveway_yes	recreation_yes	fullbase_yes	gasheat_yes	aircon_yes	prefer_yes	Predicted_Prices
390	9620	3	1	1	2	1	0	1	0	0	1	869510.0
398	7000	3	1	1	2	1	0	1	0	0	1	829705.0
210	2640	2	1	1	1	0	0	0	0	0	0	407570.0
534	5000	3	1	3	0	1	0	0	0	1	0	827490.0
468	2176	2	1	2	0	1	1	0	0	0	1	541337.5
...	...	...	...	...	...	...	...	...	...	...	...	...
485	4880	4	2	2	1	1	0	0	0	1	1	782860.0
436	7800	3	2	2	0	1	0	0	0	0	1	831360.0
284	9000	3	1	2	2	1	0	0	0	0	0	978280.0
501	8100	4	1	4	2	1	0	1	0	1	0	1311440.0
377	7420	4	2	3	2	1	0	0	0	1	1	1245444.2

Client's requirement: MagicBrick's client wanted a house which costed no more than 900,000 and is a one-story house. Out the 110 properties only 5 properties qualified this requirement.

	lotsize	bedrooms	bathrooms	stories	garage	driveway_yes	recreation_yes	fullbase_yes	gasheat_yes	aircon_yes	prefer_yes	Predicted_Prices
390	9620	3	1	1	2	1	0	1	0	0	1	869510.000000
398	7000	3	1	1	2	1	0	1	0	0	1	829705.000000
452	6100	3	1	1	0	1	0	1	0	1	1	835720.000000
394	6600	3	1	1	0	1	1	1	0	0	1	828750.000000
432	6040	3	1	1	2	1	0	0	0	0	1	810090.000000

Personalized Home Recommendation

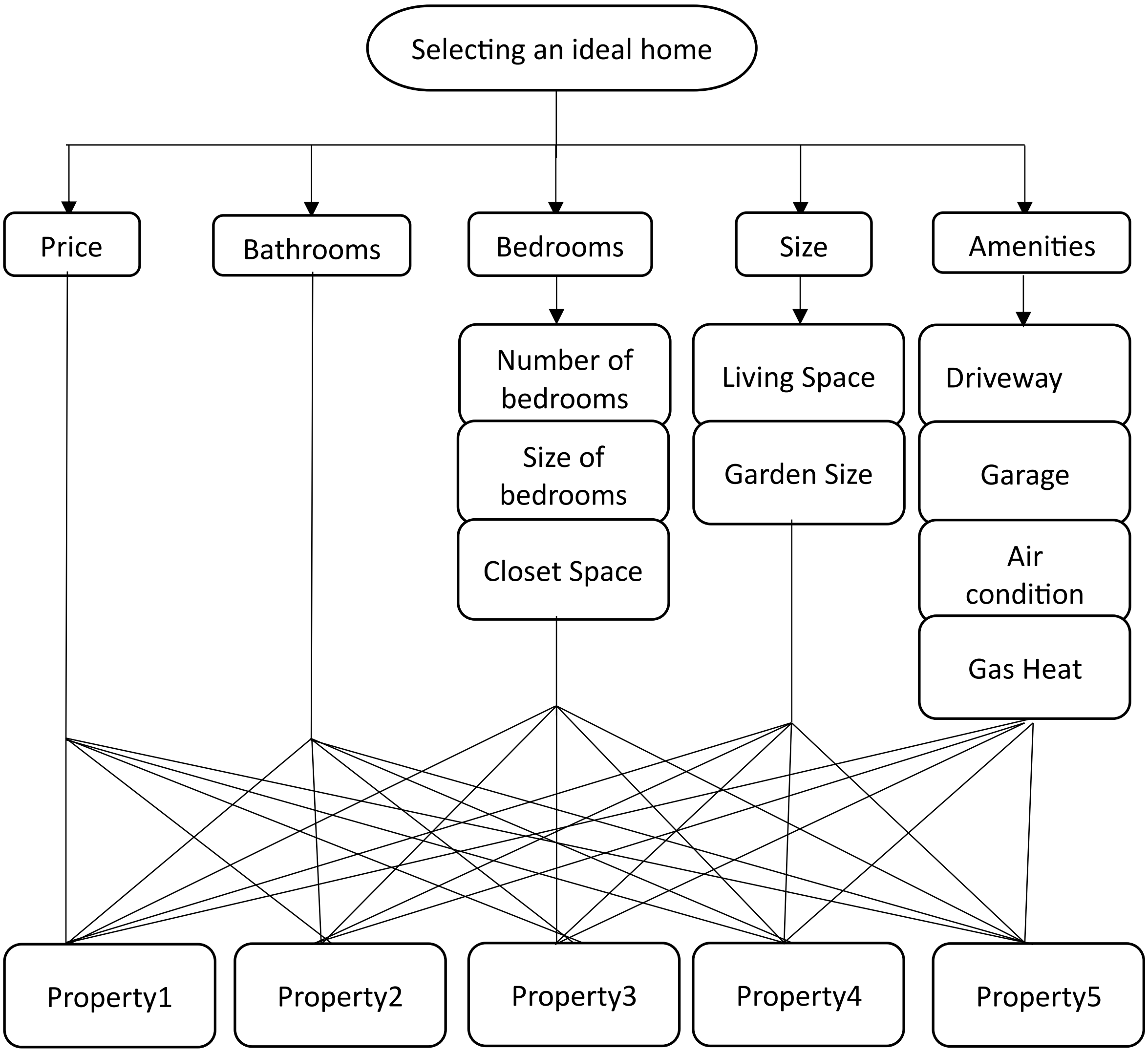
All 5 properties have different attributes. It is important to understand what attributes carried more weight or preference for our client. Selecting one best property out of 5 evolves multicriteria driven decision making. Hence, we will be using Analytical Hierarchy Process to achieve this goal.

AHP Process: Analytical Hierarchy Process is multi criteria driven process invented by Thomas L. Satty in 1970s. It is structured methodology that works through pairwise comparison of qualitative and quantitative evaluation criteria and uses the priority scales driven by expert judgement for comparison of intangible factors in the evaluation criteria.

Following steps outline the procedure we used AHP to reach our goal to select the best home for our client.

1. Determine the criteria (and sub-criteria) if any to evaluate.
2. Develop the decision hierarchy with the decision goal (find a best home) at the top, various alternatives (Property 1-5) at the bottom and various evaluation criteria (Price, bedrooms, bathroom, etc.) in the middle.
3. Perform the analysis.
    - Perform pairwise comparison of the alternatives (Property 1 -5) based on their strengths in meeting the evaluation criteria and determine priorities among them.
    - Perform pairwise comparison of the criteria(Price, bedrooms, bathroom, etc.) and sub-criteria's (size of bedroom, garage etc.) based on their importance in achieving the goal of the decision-making and determine priorities among them.
4. Synthesize the priorities from steps 2 and 3 to find the overall priority for each of the alternatives and assign a rank to each of the alternatives on the basis of its overall priority.
5. Make a decision by selecting the highest ranking alternative.

Proposed AHP Model



Analysis

To perform the analysis of strength of various alternatives we gave a survey form to the client to rate each alternative with respect to each other on scale of 1 to 9. We used Saaty' Scale to determine the importance of each rating.

The Fundamental Scale for Pairwise Comparisons		
Intensity of Importance	Definition	Explanation
1	Equal importance	Two elements contribute equally to the objective
3	Moderate importance	Experience and judgment slightly favor one element over another
5	Strong importance	Experience and judgment strongly favor one element over another
7	Very strong importance	One element is favored very strongly over another; its dominance is demonstrated in practice
9	Extreme importance	The evidence favoring one element over another is of the highest possible order of affirmation
Intensities of 2, 4, 6, and 8 can be used to express intermediate values. Intensities 1.1, 1.2, 1.3, etc. can be used for elements that are very close in importance.		

Consistency Ratio

The Analytical Hierarchy Process (AHP) utilizes the consistency ratio (CR) to assess the reliability of pairwise comparisons made by decision-makers. The CR is calculated using the following formulas:

Consistency Index (CI): CI = (λ_{max} - n) / (n - 1)

(where λ_{max} is the principal eigenvalue of the pairwise comparison matrix, and n is the number of elements being compared)

Random Index (RI): Pre-determined values based on the number of elements being compared.

CR = (CI) / (RI)

If CR exceeds a predefined threshold (typically 0.10), it indicates inconsistency in judgments, requiring a reassessment to ensure more dependable and accurate decision-making. Consistency in AHP is critical for producing robust and reliable results, guiding decision-makers to make well-informed choices based on consistent priorities.

Synthesis and Ranking

We used Python libraries to solve our problem statement. Please refer to Python Workbook for detailed code. Here, we are explaining each step of AHP process with result snippets from our Python Workbook.

Goal: To select the ideal home for our client out of 5 alternative properties.

Step 1. Defining criteria and sub-criteria.

Criteria	Sub-Criteria
Price	None
Bathrooms	None

Bedrooms	Number of Bedrooms
	Size of Bedrooms
	Closet Space
Size	Living Space
	Garden Area
Amenities	Driveway
	Garage
	Air Conditioning
	Gas Heat

Step 2. Normalized Pairwise comparison of criteria & sub-criteria and priority eigen values.

CRITERIAS	Price	Size	Bedrooms	Bathrooms	Ammenities
Price	1.000000	0.333333	3	0.500000	0.250000
Size	3.000000	1.000000	5	0.500000	0.333333
Bedrooms	0.333333	0.200000	1	0.250000	0.166667
Bathrooms	2.000000	2.000000	4	1.000000	0.500000
Ammenities	4.000000	3.000000	6	2.000000	1.000000

```
NORMALIZED CRITERIA MATRIX
------------------
```

CRITERIAS	Price	Size	Bedrooms	Bathrooms	Ammenities
Price	0.096774	0.051020	0.157895	0.117647	0.111111
Size	0.290323	0.153061	0.263158	0.117647	0.148148
Bedrooms	0.032258	0.030612	0.052632	0.058824	0.074074
Bathrooms	0.193548	0.306122	0.210526	0.235294	0.222222
Ammenities	0.387097	0.459184	0.315789	0.470588	0.444444

CRITERIAS	priority index
Price	0.106890
Size	0.194467
Bedrooms	0.049680
Bathrooms	0.233543
Ammenities	0.415421

```
CRITERIA WISE PRIORITY EIGEN VECTOR
------------------------------------
```

Step 3: Compute the consistency ratio of the normalized matrix to ensure consistency.

The Consistency Index is: 0.048

The Consistency Ratio is: 0.043

The model is consistent

SUB_CRITERIAS	Living Space	Garden Size
Living Space	1.000000	6
Garden Size	0.166667	1

SUB_CRITERIAS	Number of Bedrooms	Size_of_bedroom	Closet Space
Number of Bedrooms	1.000000	0.333333	4
Size_of_bedroom	3.000000	1.000000	7
Closet Space	0.250000	0.142857	1

SUB_CRITERIAS	Driveway	Garage	Air Conditioning	Gas Heat
Driveway	1.000000	0.333333	4.000000	4
Garage	3.000000	1.000000	7.000000	5
Air Conditioning	0.250000	0.142857	1.000000	2
Gas Heat	0.250000	0.200000	0.500000	1

```
CRITERIA WISE PRIORITY EIGEN VECTOR
-----------------------------------

NORMALIZED CRITERIA MATRIX
-----------------
```

SUB_CRITERIAS	Living Space	Garden Size
Living Space	0.857143	0.857143
Garden Size	0.142857	0.142857

SUB_CRITERIAS	priority index
Living Space	0.857143
Garden Size	0.142857

```
CRITERIA WISE PRIORITY EIGEN VECTOR
-----------------------------------

NORMALIZED CRITERIA MATRIX
-----------------
```

SUB_CRITERIAS	Number of Bedrooms	Size_of_bedroom	Closet Space
Number of Bedrooms	0.235294	0.225806	0.333333
Size_of_bedroom	0.705882	0.677420	0.583333
Closet Space	0.058824	0.096774	0.083333

SUB_CRITERIAS	priority index
Number of Bedrooms	0.264811
Size_of_bedroom	0.655545
Closet Space	0.079644

```
CRITERIA WISE PRIORITY EIGEN VECTOR
-----------------------------------

NORMALIZED CRITERIA MATRIX
-----------------
```

SUB_CRITERIAS	Driveway	Garage	Air Conditioning	Gas Heat
Driveway	0.222222	0.198863	0.320000	0.333333
Garage	0.666667	0.596591	0.560000	0.416667
Air Conditioning	0.055556	0.085227	0.080000	0.166667
Gas Heat	0.055556	0.119318	0.040000	0.083333

SUB_CRITERIAS	priority index
Driveway	0.268605
Garage	0.559981
Air Conditioning	0.096862
Gas Heat	0.074552

Step 4: Calculate the global weights of the criteria and sub-criteria as per the local weights.

	CRITERIAS	RELATIVE_WEIGHT	FACTORS	LOCAL_WEIGHT	GLOBAL_WEIGHT
0	Price	0.106890	None		0.106890
1	Size	0.194467	Living Space	0.857143	0.166700
2	Size	0.194467	Garden Size	0.142857	0.027800
3	Bedrooms	0.049680	Number of Bedrooms	0.264811	0.013200
4	Bedrooms	0.049680	Size_of_bedroom	0.655545	0.032600
5	Bedrooms	0.049680	Closet Space	0.079644	0.004000
6	Bathrooms	0.233543	None		0.233543
7	Ammenities	0.415421	Driveway	0.268605	0.111600
8	Ammenities	0.415421	Garage	0.559981	0.232600
9	Ammenities	0.415421	Air Conditioning	0.096862	0.040200
10	Ammenities	0.415421	Gas Heat	0.074552	0.031000

Step 5: Normalized Pairwise comparison of alternatives with respect to criteria and sub-criteria and priority eigen values.

Living_Space	Property 1	Proeprty 2	Property 3	Property 4	Property 5
Property 1	1	0.500000	0.250000	0.500000	1
Proeprty 2	2	1.000000	0.500000	0.250000	2
Property 3	4	2.000000	1.000000	2.000000	4
Property 4	2	4.000000	0.500000	1.000000	1
Property 5	1	0.500000	0.250000	1.000000	1

Garden_Size	Property 1	Proeprty 2	Property 3	Property 4	Property 5
Property 1	1	0.500000	0.333333	0.250000	0.333333
Proeprty 2	2	1.000000	0.500000	0.250000	1.000000
Property 3	3	2.000000	1.000000	2.000000	4.000000
Property 4	4	4.000000	0.500000	1.000000	2.000000
Property 5	3	1.000000	0.250000	0.500000	1.000000

NO_of_bedrooms	Property 1	Proeprty 2	Property 3	Property 4	Property 5
Property 1	1	0.200000	0.250000	0.500000	0.500000
Proeprty 2	5	1.000000	0.500000	1.000000	2.000000
Property 3	4	2.000000	1.000000	3.000000	2.000000
Property 4	2	1.000000	0.333333	1.000000	4.000000
Property 5	2	0.500000	0.500000	0.250000	1.000000

Size_of_bedroom	Property 1	Proeprty 2	Property 3	Property 4	Property 5
Property 1	1	0.250000	0.166667	0.250000	0.500000
Proeprty 2	4	1.000000	0.333333	0.200000	3.000000
Property 3	6	3.000000	1.000000	2.000000	3.000000
Property 4	4	5.000000	0.500000	1.000000	4.000000
Property 5	2	0.333333	0.333333	0.250000	1.000000

Closet_Size	Property 1	Proeprty 2	Property 3	Property 4	Property 5
Property 1	1	0.500000	0.250000	1.000000	1
Proeprty 2	2	1.000000	0.333333	0.333333	3
Property 3	4	3.000000	1.000000	2.000000	3
Property 4	1	3.000000	0.500000	1.000000	4
Property 5	1	0.333333	0.333333	0.250000	1

Driveway	Property 1	Proeprty 2	Property 3	Property 4	Property 5
Property 1	1	0.250000	0.166667	0.333333	1
Proeprty 2	4	1.000000	0.200000	1.000000	1
Property 3	6	5.000000	1.000000	5.000000	4
Property 4	3	1.000000	0.200000	1.000000	2
Property 5	1	1.000000	0.250000	0.500000	1

Garage	Property 1	Proeprty 2	Property 3	Property 4	Property 5
Property 1	1	0.333333	0.250000	0.333333	0.333333
Proeprty 2	3	1.000000	0.500000	1.000000	4.000000
Property 3	4	2.000000	1.000000	1.000000	2.000000
Property 4	3	1.000000	1.000000	1.000000	1.000000
Property 5	3	0.250000	0.500000	1.000000	1.000000

AIR_CONDITIONING	Property 1	Proeprty 2	Property 3	Property 4	Property 5
Property 1	1	0.200000	0.333333	0.200000	0.500000
Proeprty 2	5	1.000000	0.333333	0.500000	3.000000
Property 3	3	3.000000	1.000000	2.000000	2.000000
Property 4	5	2.000000	0.500000	1.000000	3.000000
Property 5	2	0.333333	0.500000	0.333333	1.000000

Gas_Heat	Property 1	Proeprty 2	Property 3	Property 4	Property 5
Property 1	1	1	0.250000	0.250000	1
Proeprty 2	1	1	0.250000	0.333333	1
Property 3	4	4	1.000000	3.000000	4
Property 4	4	3	0.333333	1.000000	3
Property 5	1	1	0.250000	0.333333	1

Price	Property 1	Proeprty 2	Property 3	Property 4	Property 5
Property 1	1.000000	0.500000	2.000000	2	0.500000
Proeprty 2	2.000000	1.000000	1.000000	1	0.333333
Property 3	0.500000	1.000000	1.000000	2	0.333333
Property 4	0.500000	1.000000	0.500000	1	0.250000
Property 5	2.000000	3.000000	3.000000	4	1.000000

bathrooms	Property 1	Proeprty 2	Property 3	Property 4	Property 5
Property 1	1	1.000000	1.000000	1.000000	1
Proeprty 2	1	1.000000	1.000000	1.000000	2
Property 3	1	1.000000	1.000000	2.000000	2
Property 4	1	1.000000	0.500000	1.000000	2
Property 5	1	0.500000	0.500000	0.500000	1

```
NORMALIZED CRITERIA MATRIX
------------------
```

Living_Space	Property 1	Proeprty 2	Property 3	Property 4	Property 5
Property 1	0.100000	0.062500	0.100000	0.105263	0.111111
Proeprty 2	0.200000	0.125000	0.200000	0.052632	0.222222
Property 3	0.400000	0.250000	0.400000	0.421053	0.444444
Property 4	0.200000	0.500000	0.200000	0.210526	0.111111
Property 5	0.100000	0.062500	0.100000	0.210526	0.111111

Living_Space	priority index
Property 1	0.095775
Proeprty 2	0.159971
Property 3	0.383099
Property 4	0.244327
Property 5	0.116827

```
SUB CRITERIA WISE LIVING SPACE ALTERNATIVES PRIORITY EIGEN VECTOR
------------------------------------
```

```
NORMALIZED CRITERIA MATRIX
------------------
```

Garden_Size	Property 1	Proeprty 2	Property 3	Property 4	Property 5
Property 1	0.076923	0.058824	0.129032	0.062500	0.040000
Proeprty 2	0.153846	0.117647	0.193548	0.062500	0.120000
Property 3	0.230769	0.235294	0.387097	0.500000	0.480000
Property 4	0.307692	0.470588	0.193548	0.250000	0.240000
Property 5	0.230769	0.117647	0.096774	0.125000	0.120000

Garden_Size	priority index
Property 1	0.073456
Proeprty 2	0.129508
Property 3	0.366632
Property 4	0.292366
Property 5	0.138038

```
SUB CRITERIA WISE GARDEN SIZE ALTERNATIVES PRIORITY EIGEN VECTOR
-------------------------------------
```

```
NORMALIZED CRITERIA MATRIX
------------------
```

NO_of_bedrooms	Property 1	Proeprty 2	Property 3	Property 4	Property 5
Property 1	0.071429	0.042553	0.096774	0.086957	0.052632
Proeprty 2	0.357143	0.212766	0.193548	0.173913	0.210526
Property 3	0.285714	0.425532	0.387097	0.521739	0.210526
Property 4	0.142857	0.212766	0.129032	0.173913	0.421053
Property 5	0.142857	0.106383	0.193548	0.043478	0.105263

NO_of_bedrooms	priority index
Property 1	0.070069
Proeprty 2	0.229579
Property 3	0.366122
Property 4	0.215924
Property 5	0.118306

```
SUB CRITERIA WISE NO OF BEDROOMS ALTERNATIVES PRIORITY EIGEN VECTOR
-------------------------------------
```

```
NORMALIZED CRITERIA MATRIX
------------------
```

Size_of_bedroom	Property 1	Proeprty 2	Property 3	Property 4	Property 5
Property 1	0.058824	0.026087	0.071429	0.067568	0.043478
Proeprty 2	0.235294	0.104348	0.142857	0.054054	0.260870
Property 3	0.352941	0.313043	0.428571	0.540541	0.260870
Property 4	0.235294	0.521739	0.214286	0.270270	0.347826
Property 5	0.117647	0.034783	0.142857	0.067568	0.086957

Size_of_bedroom	priority index
Property 1	0.053477
Proeprty 2	0.159485
Property 3	0.379193
Property 4	0.317883
Property 5	0.089962

```
SUB CRITERIA WISE SIZE OF BEDROOMS ALTERNATIVES PRIORITY EIGEN VECTOR
-------------------------------------

NORMALIZED CRITERIA MATRIX
------------------
```

Closet_Size	Property 1	Proeprty 2	Property 3	Property 4	Property 5
Property 1	0.111111	0.063830	0.103448	0.218182	0.083333
Proeprty 2	0.222222	0.127660	0.137931	0.072727	0.250000
Property 3	0.444444	0.382979	0.413793	0.436364	0.250000
Property 4	0.111111	0.382979	0.206897	0.218182	0.333333
Property 5	0.111111	0.042553	0.137931	0.054545	0.083333

Closet_Size	priority index
Property 1	0.115981
Proeprty 2	0.162108
Property 3	0.385516
Property 4	0.250500
Property 5	0.085895

```
SUB CRITERIA WISE CLOSET SIZE ALTERNATIVES PRIORITY EIGEN VECTOR
-------------------------------------
```

```
NORMALIZED CRITERIA MATRIX
------------------
```

Driveway	Property 1	Proeprty 2	Property 3	Property 4	Property 5
Property 1	0.066667	0.030303	0.091743	0.042553	0.111111
Proeprty 2	0.266667	0.121212	0.110092	0.127660	0.111111
Property 3	0.400000	0.606061	0.550459	0.638298	0.444444
Property 4	0.200000	0.121212	0.110092	0.127660	0.222222
Property 5	0.066667	0.121212	0.137615	0.063830	0.111111

Driveway	priority index
Property 1	0.068475
Proeprty 2	0.147348
Property 3	0.527852
Property 4	0.156237
Property 5	0.100087

```
SUB CRITERIA WISE DRIVEWAY ALTERNATIVES PRIORITY EIGEN VECTOR
-------------------------------------
```

```
NORMALIZED CRITERIA MATRIX
------------------
```

Garage	Property 1	Proeprty 2	Property 3	Property 4	Property 5
Property 1	0.071429	0.072727	0.076923	0.076923	0.040000
Proeprty 2	0.214286	0.218182	0.153846	0.230769	0.480000
Property 3	0.285714	0.436364	0.307692	0.230769	0.240000
Property 4	0.214286	0.218182	0.307692	0.230769	0.120000
Property 5	0.214286	0.054545	0.153846	0.230769	0.120000

Garage	priority index
Property 1	0.067600
Proeprty 2	0.259417
Property 3	0.300108
Property 4	0.218186
Property 5	0.154689

```
SUB CRITERIA WISE GARAGE ALTERNATIVES PRIORITY EIGEN VECTOR
-------------------------------------
```

```
NORMALIZED CRITERIA MATRIX
------------------
```

AIR_CONDITIONING	Property 1	Proeprty 2	Property 3	Property 4	Property 5
Property 1	0.062500	0.030612	0.125000	0.049587	0.052632
Proeprty 2	0.312500	0.153061	0.125000	0.123967	0.315789
Property 3	0.187500	0.459184	0.375000	0.495868	0.210526
Property 4	0.312500	0.306122	0.187500	0.247934	0.315789
Property 5	0.125000	0.051020	0.187500	0.082645	0.105263

AIR_CONDITIONING	priority index
Property 1	0.064066
Proeprty 2	0.206064
Property 3	0.345616
Property 4	0.273969
Property 5	0.110286

```
SUB CRITERIA WISE AIR CONDITIONING ALTERNATIVES PRIORITY EIGEN VECTOR
-------------------------------------
```

```
NORMALIZED CRITERIA MATRIX
------------------
```

Gas_Heat	Property 1	Proeprty 2	Property 3	Property 4	Property 5
Property 1	0.090909	0.100000	0.120000	0.050847	0.100000
Proeprty 2	0.090909	0.100000	0.120000	0.067797	0.100000
Property 3	0.363636	0.400000	0.480000	0.610170	0.400000
Property 4	0.363636	0.300000	0.160000	0.203390	0.300000
Property 5	0.090909	0.100000	0.120000	0.067797	0.100000

Gas_Heat	priority index
Property 1	0.092351
Proeprty 2	0.095741
Property 3	0.450761
Property 4	0.265405
Property 5	0.095741

```
SUB CRITERIA WISE GAS HEAT ALTERNATIVES PRIORITY EIGEN VECTOR
-------------------------------------
```

NORMALIZED CRITERIA MATRIX

Price	Property 1	Proeprty 2	Property 3	Property 4	Property 5
Property 1	0.166667	0.076923	0.266667	0.200000	0.206897
Proeprty 2	0.333333	0.153846	0.133333	0.100000	0.137931
Property 3	0.083333	0.153846	0.133333	0.200000	0.137931
Property 4	0.083333	0.153846	0.066667	0.100000	0.103448
Property 5	0.333333	0.461538	0.400000	0.400000	0.413793

Price	priority index
Property 1	0.183431
Proeprty 2	0.171689
Property 3	0.141689
Property 4	0.101459
Property 5	0.401733

CRITERIA WISE PRICE ALTERNATIVES PRIORITY EIGEN VECTOR

```
NORMALIZED CRITERIA MATRIX
------------------
```

bathrooms	Property 1	Proeprty 2	Property 3	Property 4	Property 5
Property 1	0.200000	0.222222	0.250000	0.181818	0.125000
Proeprty 2	0.200000	0.222222	0.250000	0.181818	0.250000
Property 3	0.200000	0.222222	0.250000	0.363636	0.250000
Property 4	0.200000	0.222222	0.125000	0.181818	0.250000
Property 5	0.200000	0.111111	0.125000	0.090909	0.125000

bathrooms	priority index
Property 1	0.195808
Proeprty 2	0.220808
Property 3	0.257172
Property 4	0.195808
Property 5	0.130404

```
CRITERIA WISE BATHROOMS ALTERNATIVES PRIORITY EIGEN VECTOR
------------------------------------
```

Step 6: Calculating the sum of the global weights of the different alternatives present in the evoked set.

	CRITERIAS	SUB CRITERIAS	GLOBAL WEIGHT	LW_Property1	LW_Property2	LW_Property3	LW_Property4	LW_Property5	GW_Property1	GW_Property2	GW_Property3	GW_Property4	GW_Property5
0	Price	None	0.106890	0.183431	0.171689	0.141689	0.101459	0.401733	0.019600	0.018400	0.015100	0.010800	0.042900
1	Size	Living Space	0.166700	0.095775	0.159971	0.383099	0.244327	0.116827	0.016000	0.026700	0.063900	0.040700	0.019500
2	Size	Garden Size	0.027800	0.073456	0.129508	0.366632	0.292366	0.138038	0.002000	0.003600	0.010200	0.008100	0.003800
3	Bedrooms	Number of Bedrooms	0.013200	0.070069	0.229579	0.366122	0.215924	0.118306	0.000900	0.003000	0.004800	0.002900	0.001600
4	Bedrooms	Size_of_bedroom	0.032600	0.053477	0.159485	0.379193	0.317883	0.089962	0.001700	0.005200	0.012400	0.010400	0.002900
5	Bedrooms	Closet Space	0.004000	0.115981	0.162108	0.385516	0.250500	0.085895	0.000500	0.000600	0.001500	0.001000	0.000300
6	Bathrooms	None	0.233543	0.195808	0.220808	0.257172	0.195808	0.130404	0.045700	0.051600	0.060100	0.045700	0.030500
7	Ammenities	Driveway	0.111600	0.068475	0.147348	0.527852	0.156237	0.100087	0.007600	0.016400	0.058900	0.017400	0.011200
8	Ammenities	Garage	0.232600	0.067600	0.259417	0.300108	0.218186	0.154689	0.015700	0.060300	0.069800	0.050800	0.036000
9	Ammenities	Air Conditioning	0.040200	0.064066	0.206064	0.345616	0.273969	0.110286	0.002600	0.008300	0.013900	0.011000	0.004400
10	Ammenities	Gas Heat	0.031000	0.092351	0.095741	0.450761	0.265405	0.095741	0.002900	0.003000	0.014000	0.008200	0.003000

Step 7: Assign rank to each alternative based upon the priority level.

FINAL RANKING OF THE PROPERTIES

	Property_ID	Eigen Value	Priority_Percent	Rank
2	Property 3	0.324600	32.46 %	1
3	Property 4	0.207000	20.7 %	2
1	Property 2	0.197100	19.71 %	3
4	Property 5	0.156100	15.61 %	4
0	Property 1	0.115200	11.52 %	5

Step 8: As per our decision, Property 3 is best out of 5 other alternatives.

Conclusion: To summarize, the conjunction of random forest algorithm and analytical hierarchy method have been beneficial to address this company case. Whether it is predicting house price accurately or offering personalized house recommendation, this combined approach has unleashed the true potential of real estate analytics.

References

- Mbaabu, O. (2020, Dec 11). Introduction to Random Forest in Machine Learning. https://www.section.io/engineering-education/introduction-to-random-forest-in-machine-learning/
- Ramirez-Rios, D. (2016, Nov 7). Hypothesis Testing: Sales Prices of Houses in the City of Windsor. https://rstudio-pubs-static.s3.amazonaws.com/225821_e7925e3ee76b4b3abeb5d0d2ef83f932.html
- Singh, H. (2016). Project Management Analytics: A data driven approach to making rational and effective project decisions. Pearson Education, Inc.
- Asaftei, G. Doshi, S. Means, J. & Sanghvi, A. (2018, Oct 8). Getting ahead of the market: How big data is transforming real estate. Mckinsey. https://www.mckinsey.com/industries/real-estate/our-insights/getting-ahead-of-the-market-how-big-data-is-transforming-real-estate
- Wappnet Systems Pvt. Ltd. (2023, Jun 16). Leveraging AI for Real Estate Project Management: Unlocking Key Benefits. LinkedIn. https://www.linkedin.com/pulse/leveraging-ai-real-estate-project-management-unlocking/?trk=pulse-article

- Mir, M. (2020, Jul 21). House Prices Prediction Using Deep Learning(Keras-Regression vs Multiple Linear Regression) https://towardsdatascience.com/house-prices-prediction-using-deep-learning-dea265cc3154
- Fang, Y., Li, T., & Zhao, H. S. (2022 Feb 22). *Random Forest model for the house price forecasting*. https://doi.org/10.1109/iccrd54409.2022.9730190