ÇANKAYA UNIVERSITY
FACULTY OF ENGINEERING
COMPUTER ENGINEERING DEPARTMENT

Project Report
Version 1

CENG 407
Innovative System Design and Development I

202002
AIRBNB PRICE DETERMINANT

201627015 Melisa YILDIZ
201611033 Selin KARA
201514039 Merve KARAKAYA
201512019 Rana Gürcü DELİÖMEROĞLU

Advisor: *Dr. Roya CHOUPANI*

Table of Contents

Abstract

Airbnb is a worldwide house and property rental platform that provides short-term accommodation services to those who prefer houses to hotels. Airbnb is getting more popular day by day.The platform puts the responsibility of determining the price of the hosts. However, the prices are not determined randomly.As a matter of fact, many features such as location, availability of public transportation, size of the rented house, the existence of shopping areas in the neighborhood and many more play a role in determining the price.In this project, the real data from the official site of Airbnb will be used to determine which attributes are the most decisive ones in determining the price.The model to be developed with this project is to provide an auxiliary service for homeowners to estimate the prices more accurately.In order to work more efficiently within the scope of this project, we are conducting extensive searches and investigations on machine learning,deep learning and data mining, which are the applications of artificial intelligence.

Key words: airbnb , determining the price , machine learning , deep learning , data mining , artificial intelligence

Özet:

Airbnb, evi otele tercih edenlere kısa süreli konaklama hizmetleri sunan dünya çapında bir ev ve mülk kiralama platformudur. Airbnb gün geçtikçe daha popüler hale geliyor.Platform, ev sahiplerinin fiyatını belirleme sorumluluğunu üstlenir. Ancak fiyatlar tesadüfi olarak belirlenmez, aslında fiyatın belirlenmesinde konum, toplu taşıma kullanılabilirliği, kiralanan evin büyüklüğü, mahallede alışveriş alanlarının varlığı ve daha birçoğu gibi birçok özellik rol oynamaktadır. Bu projede Airbnb'nin resmi sitesinden alınan gerçek veriler, fiyatın belirlenmesinde hangi özelliklerin en belirleyici olduğunun belirlenmesinde kullanılacak. Bu proje ile geliştirilecek model, ev sahiplerinin tahmin etmeleri için yardımcı bir hizmet sağlamaktır. daha doğru fiyatlar. Bu proje kapsamında daha verimli çalışabilmek için yapay zekanın uygulamaları olan makine öğrenmesi, derin öğrenme ve veri madenciliği konularında kapsamlı araştırmalar ve araştırmalar yapıyoruz

Anahtar Kelimeler: airbnb, fiyat belirleme, makine öğrenmesi, derin öğrenme, veri madenciliği , yapay zeka

1. Introduction

In recent years, people prefer to use platforms such as Airbnb, which determine the price according to the characteristics of the place they will stay, instead of fixed and generally high-priced hotels during their touristic trips and holidays. Although hosts can compare prices manually with other houses and they can determine a price for their house but this is not a healthy method. In order to get a more accurate result when determining the price, a prediction algorithm should be established according to the characteristics of the house. Features such as number of bedrooms, number of bathrooms, neighborhood, house type (detached house / apartment / room), if it is a room, it can be shared with another person or not are determined for more accurate price prediction . This article aims to develop an accurate price development model using machine learning, deep learning so that hosts can more easily price according to the features above.

1.1 Problem Statement

The Airbnb platform does not have a price determinant for homeowners based on the characteristics of their home.Through this project, we aim to provide a platform where homeowners can determine the prices of their homes.It sets an appropriate rental price using the properties found in the house.Such a platform has not been implemented before, but there are many resources on the subject.

1.2 Background or Related Work

In the past years, multiple price setting projects have been carried out for hosts using Airbnb. Various machine learning algorithms and models were used in all of these projects. Machine learning is an area where we can find a lot of work today, so we were able to examine many literature that could be a reference to our project.In addition, the fact that Airbnb is a widely used application today has been effective. While we are researching, We saw that as the demand for Airbnb has increased, the number of projects done over the years has increased. These projects helped us a lot at the idea stage.

1.3 Solution Statement

After to the researches, We examined machine learning algorithms which are Logistic Regression, KNN, Naive Bayes, Random Forest,Decision Tree. Also, Regression ve Classification models are examined. These models and algorithms will be used in our project. Some of them will be eliminated according to their efficiency and results.

1.4 Contribution

We will try to synthesize all our research and all the articles we read and transfer it to our own project. In order to get the best results in our own project, we separated the algorithms and models that worked best in previous projects among algorithms and models. We decided that regression and classification were the best models for our project..

2. Literature Search

In December 2018, V. Raul Perez-Sanchez and his team published an article about this subject. In their paper, they adopt a hedonic price model, thanks to this model, they determine the relation between Airbnb accommodation attributes and price.The model uses OLS to estimate the relationship between features and pricing [1].

OLS model can be explained as follows:

$$\ln \ln \; (P_i) \;= \alpha + \sum_{j=1}^{n} \beta_j X_{ij} + \sum_{k=1}^{m} \gamma_k + D_{ik} + \varepsilon_i \quad [1]$$

$\ln \ln \; (P_i)$: the neperian logarithm of *daily price* (" *i* ")
α : the fixed component
β_j : estimated parameter which is related to characteristic "*j*"
X_{ij} :continuous variable which considers the characteristic "*j*" of observation "*i*"
γ_k : estimated parameter which is related to characteristic "*k*"
D_{ik} : the fictitious variable which considers the characteristic "*k*" of observation "*i*"
ε_i : error associated with the observation "*i*". [1]

QRM overcomes some of the limitations of the OLS model, allowing the modeling of different sizes of the dependent variable.
QRM can be explained as follows:[1]

$$\gamma_i = x_i \beta_\theta + u\theta_i \; [1]$$

where
γ_i : the dependent variable[
x_i : the matrix of independent variables
β_θ : the vector of parameters to be estimated for quantile θ
$u\theta_i$: the aleatory perturbation that corresponds to quantile . [1]

Studies based on the hedonic regression model in 2019 examined pricing determinants from an algorithmic perspective.

For Airbnb pricing studies, researchers have examined past studies and defined some explanatory variables. The researchers used the semilogarithmic hedonic price model to observe the effect of these explanatory variables on the price [2] .

Hedonic price model was improved by Rosen (1974) in order to measure the marginal effects of the properties of a product on the price. The function is defined below:

P	Observed price of a product
Z	Vector of the product's attributes or utilities
ε	Error Term

Figure 3. The Original Hedonic Price Function

The semilog form of hedonic price function was used in the studies, which is indicated as:

ln P	log form of the observed price of a sample
β0	Constant term
Xᵢ	ith explanatory variable
βt	Semi Elasticity of P with respect to Xi
ε	Error Term

Figure 4. Semilog Form of Hedonic Price Function

As a result, it has been determined from the application of this model that some of the explanatory variables have a great effect on the price, and some do not. On the other hand, some variables on the price are expected to be clarified in future studies as their results are not reached.

In this study covering the year 2020, a new approach type, Multiscale Geographically Weighted Regression (MGWR), was used in the regression model. The main idea in this model is to determine the Airbnb price by considering the spatial differences. Scales should have differences such as local and global depending on the spatial differences so that the pricing study can be done more accurately. According to MGWR, which is the model chosen to analyze in the study, a different scale should be applied to estimate the effect of each feature on the price, namely its coefficient. The mathematical approach of this model is as follows: The here shows the different scale process applied to each feature (j) [3].

As a result of the application of this model, it has been observed that the scales applied to the features selected to examine the effect on pricing vary considerably according to the applied places, even if the same feature is applied.

2.1 Dataset and Features

2.1.1 Dataset

A dataset is a structure that holds more than one data information in it. The data contained in the dataset are taken from one or more databases. Each column in the table reflects the name of a data variable and each row reflects the data of that variable. In our project, we chose to use data sets for Airbnb listings in Istanbul. insideairbnb.com maintains datasets for users to access Airbnb data collectively. Airbnb data of each city is available on the site. We split the Istanbul dataset into trains: test at a ratio of 7: 3 for the dataset.

2.1.2 Features

Features are the main factor for price prediction. It should be informative to the user and relevant to the price. Because of that, we selected our features based on the important requirements for a guest. . In order for the host to get the most accurate price estimation, we determined the following features according to the guest needs:

- House type: Determining the type of house for pricing is the main step. Whether the house is an apartment, detached or room type changes the pricing.

- Room type: After choosing the home type, the user should also indicate whether the room is shared or not, if it is a home-type room, this criterion also affects the price.

- Neighborhood: Which neighborhood the house is located in is also an important factor. For example, the rental price of a house in Beşiktaş would be higher than Avcılar.

- Number of bedrooms: The number of bedrooms in the house should also be specified. The large number of bedrooms is a factor that increases the price.

- Number of bathrooms: The number of bathrooms in the home is also an important factor in price determination. More than one bathroom increases the price.

- Extra features: The features here will be extracted from the dataset with machine learning algorithms and will be determined later.

2.2   Model and Algorithms

Machine learning tries to estimate the output values as closely as possible according to the given inputs. Analyzes the input sets entered for this forecast. It establishes relationships between input values. The algorithms it uses optimize operations over time to improve performance and develop artificial intelligence.

In Supervised Machine Learning, the machine learning algorithm is built on a data set according to the given inputs and a desired output. The algorithm makes predictions and the algorithm works until the closest result is reached. Classification, Regression and Prediction are models that work with controlled learning.

The classification model is used to create outputs and classify new observations with what algorithms learn from inputs. Data sets are used to determine target classes and to create conditions for boundaries. After the conditions are determined, the next step is to predict the target class, or output.

Regression model focuses on one dependent and more than one independent variable. Makes output estimates based on these variables. The general purpose of the regression model is to think like a mathematical equation and make a y estimate that gives results according to x variables.

We can easily use Classification and Regression models for a price determinant system. Algorithms we can use depending on these methods:

2.2.1 Naive Bayes Algorithm

The Naive Bayes algorithm is one of the most efficient and most used algorithms in machine learning studies.Its logic is based on Bayes' theorem. The algorithm's attributes are independent and easy to implement. Big data entries can be scaled more easily in this algorithm.

2.2.2 Logistic Regression Algorithm

Logistic Regression algorithm is a very useful algorithm to understand the effect of multiple inputs on a single output. It focuses on binary classification. If the problem is multi-categorized, sequential logistic regression is applied. It is especially used in problems such as purchasing and credit analysis.

2.2.3 K Nearest Neighbor (KNN) Algorithm

The KNN algorithm predicts the probability that an input is a member of another set after the inputs are received and divided into sets. Someone sets data for reference which name is k and compares other data with this reference data k.

2.2.4 Decision Tree Algorithm

In the Decision Tree algorithm, the tree sets rules for classifying data. Divides the inputs into homogeneous clusters according to differentiers. It then divides the data until the leaves reach the maximum depth and repeats this process, selecting the one with the highest probability of accuracy.

2.2.5 Random Forest Algorithm

It is a classification algorithm derived from Random Forest decision trees. This algorithm works effectively even with large data sets. It can be used easily for Regression and

Classification models. The algorithm starts with a tree structure, after the first entry is entered, other entries are entered down. It is divided into small clusters according to the relationships among the entries.

3. Summary

3.1 Summary of Conceptual Solution

As a result of this study, we examined similar studies conducted in 2017, 2018, 2019 and 2020 for our Airbnb price determinant project. We grasped why, when and for what Airbnb, a Sharing economy business model, is used. We conclude that a price determinant system is required for these platform users. We have observed that regression and classification were the method with the most results in price determinant projects that previously worked for Airbnb and projects in the field of sharing economy. We researched the algorithms which are used in classification and regression. According to the researches, We examined Logistic Regression, KNN, Naive Bayes, Random Forest and Decision Tree algorithms. These models and algorithms can be used easily in a price determinant system. Because of that, We will use these algorithms and models in our project.

3.2 Technology Used

For the coding part of this project, TensorFlow and Pandas libraries will be used for machine learning in Python programming language. For Python we will also use Anaconda's Jupyter editors. Also, Visual Studio and Php will be used for the web part of the project. MSSQL will also be used for the database part in order to be comfortable to use with Python and Visual studio.

4. Software Requirements Specification

4.1 Introduction

4.1.1 Purpose

The purpose of this document is to explain the Airbnb pricing platform. This platform will prepare an accurate pricing policy for Airbnb hosts. In this document, the necessary explanations and models will be presented in detail for a better understanding of the project. Additionally, in these models, the features of the users who can access the platform will be explained.

4.1.2 Scope of Project

Airbnb is a platform where homes can be shared for travelers who prefer short-term accommodation and home to hotel. Airbnb is different from other types of accommodation due to its structure. This is why Airbnb pricing strategy requires different approaches. Different methods are used in this pricing strategy. This situation differs from country to country, from city to city and even in regions within the same city.

There are many factors for pricing policy. When we look at the literature studies on this subject, different effects of the factors on pricing policies are observed. For example, some factors have a great effect on price determination, while some do not have a positive or negative effect. Some are still up for debate.

The aim of this project is to develop a necessary subsidiary platform for Airbnb that can help with the difficulties that arise in the price determination policy mentioned above. Therefore, this platform can be easily used by Airbnb hosts.

4.1.3 Glossary

Host	People sharing their homes on the Airbnb platform
MSSQL	Microsoft SQL Server
PHP	Hypertext Preprocessor

4.1.4   References

[1] IEEE Recommended Practice for Software Design Descriptions

[2] IEEE Recommended Practice for Software Requirements Specifications

[3] IEEE Standard for Software Project Management Plans

4.1.5   Overview of Document

This document is arranged according to the IEEE Recommended Practice for Software Requirements Specifications [2].

Second part contains a general description of the project. Also, product perspective, user characteristic, constraints, assumptions and dependencies, and apportioning of requirements sections are included.

In the third part, there are external interface requirements, functional requirements, performance requirements, software system attributes and design constraints.

4.2   Overall Description

In this section, the main factors and requirements affecting the project will be described. In order to understand these requirements easily, this section of the SRS will provide the necessary background. More detailed definitions will be included in the third section of the SRS.

4.2.1   Product Perspective

First of all, the necessary data will be obtained from Airbnb's official site for the Airbnb price determination platform that we will create. Later, this data will be classified according to price determination factors and an auxiliary platform will be developed that will determine the price through an appropriate model.

4.2.2   Development Methodology

While developing our project, we tried to use Scrum, an agile software development methodology. Basically, our work in this direction is compatible with the Waterfall model, which is one of the software development models. For this, we first conducted a literature review. Then, we will carry out the Software Requirements Specification and Software Design Document processes that provide the documentation required for the development of

the project. After these stages, we will make the project a usable platform, so that the operation of the project will proceed more systematically.

Thanks to the regular meetings held every week with our team and our advisor, we draw up plans and ensure the systematic progress of our project by taking the deadline into consideration.

4.2.3 User Characteristics

4.2.3.1 *Host*

- This type of user is the person who will benefit from the platform.
- Host will be able to filter according to the price determination features on the platform.
- Host will be able to do price research in accordance with price determination features.
- Host will be able to get the most optimized Airbnb price.

4.2.3.2 *Admin*

- Admin will be able to enter price determination features on the platform.
- Admin will be able to access the database.
- The admin user type will be able to change the interface of the platform.

4.2.4 Constraints

- There are two types of users on the platform, host and admin. Host is the person who can use this platform and the admin is the person who can manage the platform.
- The types of housing where this platform can be used are detached house, apartment or just a room. When the single room option is selected among the housing types, the issue of whether it can be shared or not is limited.
- There are also restrictions on price determination, Restrictions are information of neighborhood, the number of bathrooms and bedrooms, and extra features in the house.

4.2.5 Assumptions and Dependencies

Any user with the appropriate hardware and software specified in the requirements specification which is the third section can use the platform.

4.2.6 Apportioning of Requirements

The database used by the Airbnb price determination platform we have created is dynamic and can be updated. Meanly, the platform can be a source of inspiration for future studies.

4.3 Requirements Specification

4.3.1 External Interface Requirements

4.3.1.1 User interfaces

This system will be able to work actively on all platforms with python 3.6 installed. In the interface of this system, which has two types of users, Host who is the first user, can select the room type, if it is a home-type room, it can also select the room type, add the distance information to the hospital / public transport / supermarket / city center, and finally can see the estimated house rental price. Admins who are the second user can edit the system.

4.3.1.2 Hardware interfaces

The computer to be used for this system should be on a minimum 32 bit based. Also, this system is usable for all operating systems.

4.3.1.3 Software interfaces

The computer to be used for this system must contain Anaconda for Python and MSSQL.

4.3.1.4 Communications interfaces

Internet connection is required to operate the system.
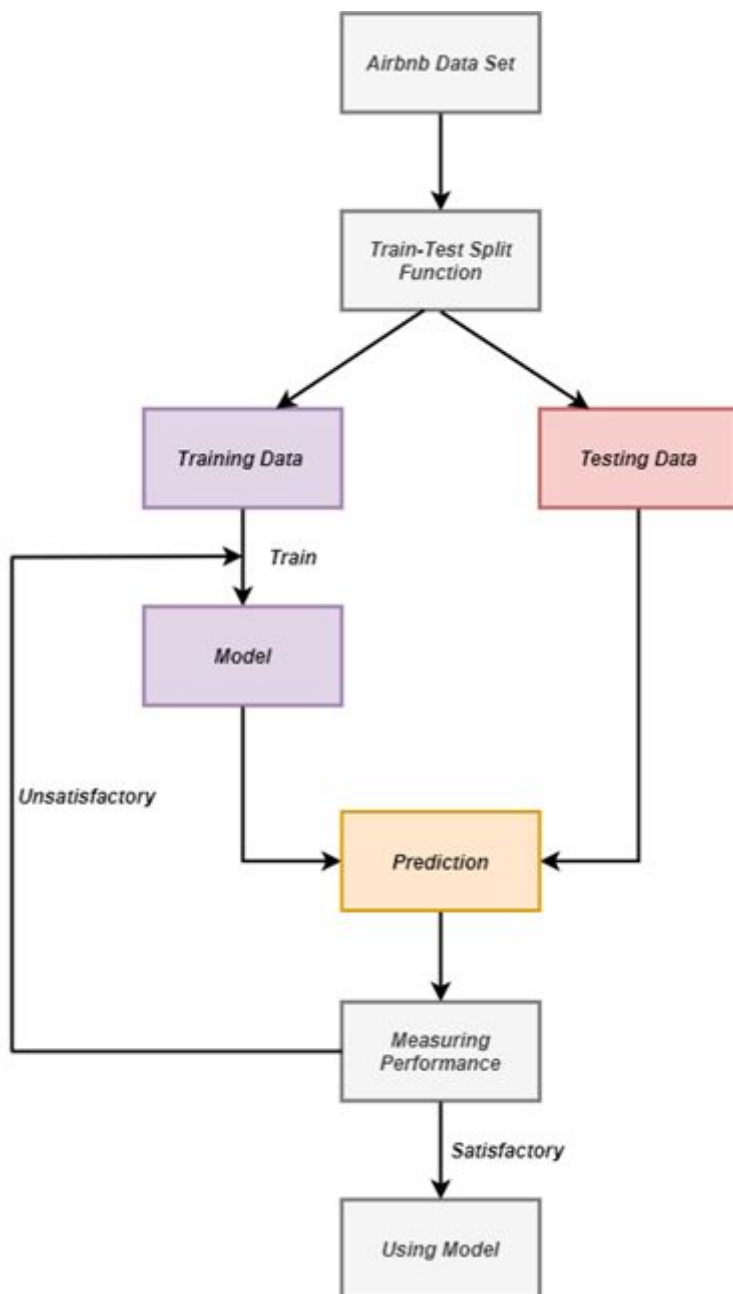
4.3.2 Functional Requirements

Figure 1. System Flow Diagram

The system flow diagram of the machine learning model that will work for price determination on this platform is as in Figure 1. Accordingly, data from Airbnb is primarily divided into train and test. It then uses the data reserved for training to develop the model. Test data are used to observe the prediction performance of this model. The model can be used if it gives satisfactory estimates, and if it does not, the model should be developed again.

4.3.2.1   *Profile Management Use Case*

Use Case:

- Start
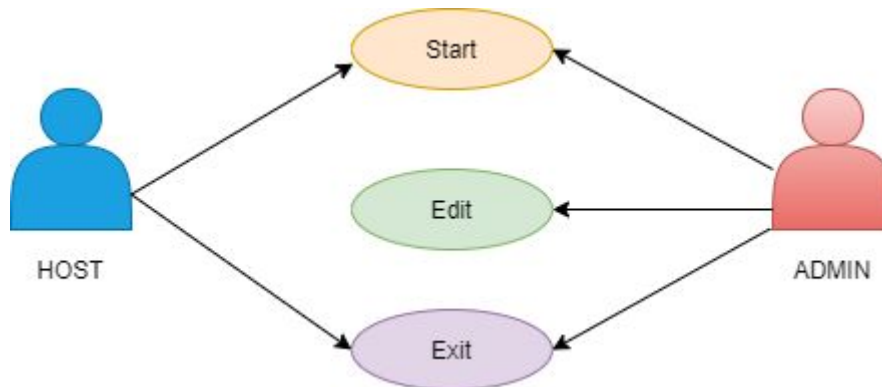- Edit
- Exit

Diagram:

Figure 1 Profile Management Use Case

Brief Description:

The Profile Management Diagram describes the basic operations for Host and Admin that they will use in the system. It uses common Exit operation as Host and Admin. The host and admin can enter the system without Login. Also, Admin has Edit authority.

Initial Step-by-Step Description:

- Host can enter the system without Login.
- Admin can enter the system without Login.
- Admin can edit the system.
- Admin and Host can exit the system unconditionally.

Use Case:

- Start
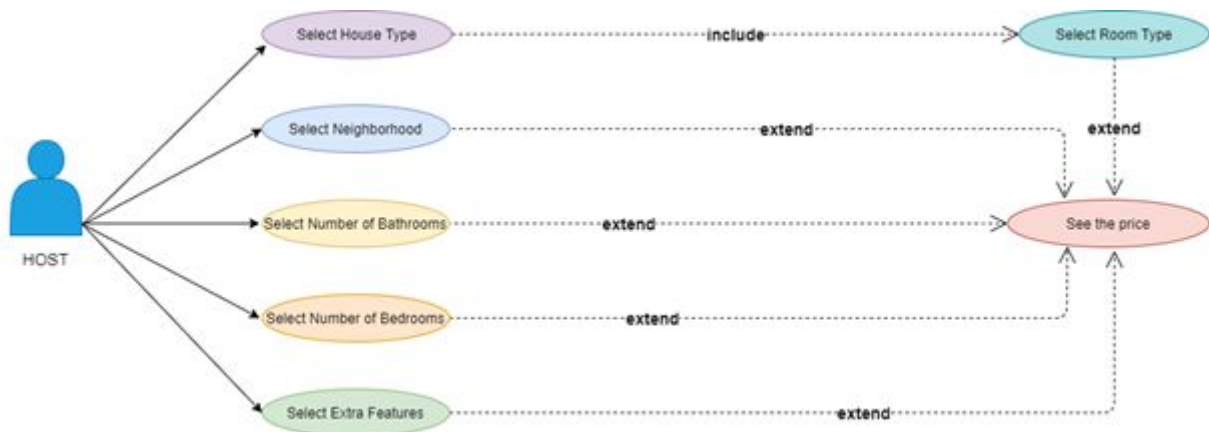- Edit
- Exit

Diagram:

Figure 2 Host Use Case

Brief Description:

The Host Use Case Diagram explains the operations that hosts can do. First of all, the hosts should reach this created platform via any internet browser. Hosts should select the house type, room type, neighborhood, number of bathrooms, number of bedrooms, and extra features to get an estimate of the price on my site. These selecting operations in the system start with the house type. After selecting the house type, if it is a room, the room type should also be selected. After the house type, the neighborhood where the house is located, the number of bathrooms and bedrooms in the house should be selected. Besides these, the host can select the extra features (eg tv, wifi, etc.) for his home. Extra features that can be selected for the house in this system will be determined later by the machine learning algorithm according to the data in the Airbnb Istanbul dataset. After entering this information, the host will see the price, which is a numerical value that the Airbnb pricing platform will give it.

According to the researches in the literature, Airbnb price determinants vary considerably according to many situations. However, the information entered by the host in the use case for this platform is chosen in this way especially because it is the basic price determination criteria that can fit many situations.

Initial Step-by-Step Description:

- Host should select their house type. There are three options for this, an entire house, an entire apartment, and a room.
- If the Host selects the room, They should choose their room type. The room which is selected can be a shared room or private room.
- The Host should select the neighborhood of the house. Options that are listed here will be include string values which are neighborhoods in İstanbul.
- The Host should select the number of bathrooms in the house. Options that are listed will be integer values.
- The Host should select the number of bedrooms in the house. Options that are listed will be integer values.
- The Host should select the extra features for the house like tv, wifi, kitchen, etc. Options that are listed will be string values but these features are determined clearly after the dataset is examined by machine learning algorithms.
- After selecting this information, They can see their estimated house rent price value.

4.3.3   Performance Requirements

Airbnb price is a determinant web application. There is no need for a specific system to run this application, minimum conditions will be sufficient. For example, the processor must be at least Intel® Core ™ i3 processor or Amd Phenom X4. The operating system used must be Linux, MacOS or Windows.

4.3.4   Design constraints

In addition to the Constraints section, This platform can be used by all web browsers. The data to be used for the price setting platform will be accurate data from Airbnb's official site. It can be used by all operating systems.

4.3.5 Software system attributes

There are a number of attributes of software that can serve as requirements. It is important that required attributes be specified so that their achievement can be objectively verified. Subsections provide a partial list of examples.

4.3.5.1 *Usability*

When using this application, the host should select the house type, room type (if the house type is a single room), neighborhood, number of bathrooms, number of bedrooms, and extra features to determine the price of the house.

4.3.5.2 *Adaptability*

This platform works in integration with data on Airbnb. Therefore, as data on Airbnb may change over time, the prices set by the platform may change accordingly.

4.3.5.3 *Scalability*

The criteria used by the system for the price determinant are sufficient for now, but these criteria can be increased according to the demands of the users.

4.3.5.4 *Maintainability*

Since Airbnb is a platform that is constantly open to new data entry, the data that this application will use is updated depending on the data on Airbnb. There will be no change in the interface of the platform that will fundamentally affect the flow of the project, but the data used for price determination in the background should constantly change

4.3.5.5 *Portability*

Data for this web application will be pulled from Airbnb via Python. The Anaconda resource containing IDEs that will allow us to easily develop Python will be used. Also, PHP and MSSQL will be used for the web development part.

5. Software Design Description

5.1 Introduction

5.1.1 Purpose

The purpose of the Software Design Description document is to make the explanations of the steps to design the software required for a system and the details such as what will be built and how to design these steps.

5.1.2 Scope

The scope of the Software Design Description document is system oriented. It allows the information and documents required for this system to be gathered, edited and changed when necessary. In short, it helps to create the basics of the software to be developed before the system becomes a product with explanations and visuals.

5.1.3 Glossary

Table 2 Glossary of SDD

Term	Definition
SDD	Software Design Description
SRS	Software Requirements Specification
Host	People sharing their home on the Airbnb Platform

5.1.4 References

[1] IEEE Recommended Practice for Software Design Descriptions.

[2] IEEE Recommended Practice for Software Requirements Specifications.

[3] https://wireframe.cc

5.1.5 Overview of document

The overview of the Software Design Description document will be as follows: introduction, design explanations, necessary software and hardware architectures, interfaces and designs

5.2 Design Consideration

This section contains the terms and concepts required for design in the context of Software Design Description. Accordingly, this part of the document contains approaches, tools, assumptions, dependencies and restrictions.

5.2.1 Approach

Our basic approach is to show in detail the patterns, structural styles, frame templates required for design.

5.2.2 Tools Used

The data required for this web platform will be extracted from Airbnb's official site. This data will be processed with Python. Anaconda source will be used for this. In addition, PHP will be used for the web part and MSSQL will be used for the database.

5.2.3 Constraints

The restrictions for the system are as follows: There are two types of users. One of them, the host, will use this platform for price estimation and Admin is the person who develops and manages the platform. Housing types are entire house, entire apartment or just one room. If there is only one room, it should be stated that this room will be shared or non-shared. There are also restrictions on the features that determine the price. These features are selected as follows: Selecting number of bathrooms and bedrooms. In addition, the remaining features will be determined by the algorithm to be used according to the data later.

5.2.4 Assumptions and Dependencies

There is no need for a special login to use the platform. Users with basic software and hardware requirements can access and use this platform.

5.3 Architecture Design

5.3.1 Software Architecture

Airbnb Price Determinants platform has a data-based software architecture. Thanks to this database, users can take advantage of the platform and access the value corresponding to the price up to date. The representation of the software architecture is as follows:
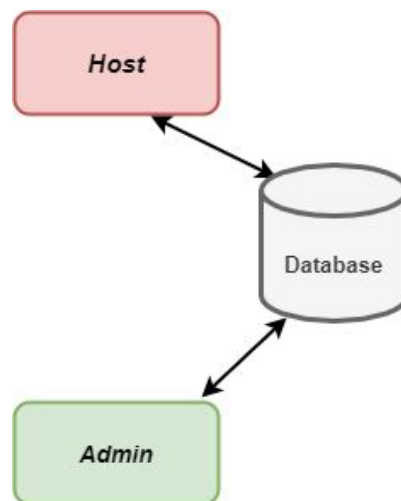
Figure 1: Software Architecture of Platform

5.3.2 Hardware Architecture

A user layer, a server layer and a database layer will be needed for this platform because users need a user layer to access the platform, while a server layer is required to use the platform. Also, a database will be required to keep the platform up to date. In short, although there is no restriction in the user-side, the access to the server-side is limited to the admin type user
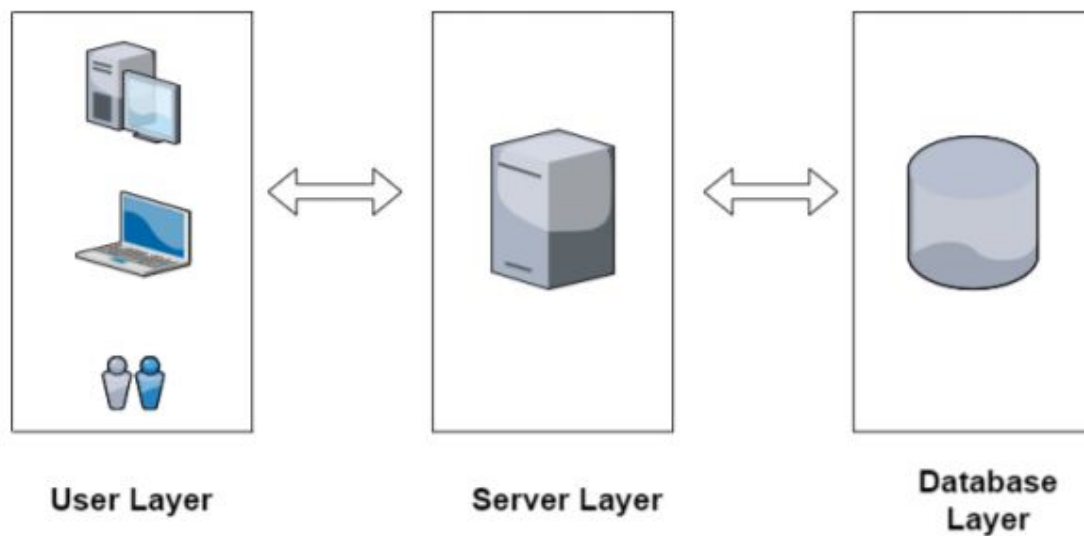
Figure 2: Hardware Architecture of Platform

5.4   System Interfaces

5.4.1 Software Interfaces

There is no need for a specially defined operating system for the computer to use this platform.

5.4.2 Hardware Interfaces

There is no need for a specially defined system to run this platform. This platform can be easily run on any computer with basic requirements.

5.5   User Interface Design

The interface perspective provides a visual design of the product to be created for designers, programmers and testers and provides a better understanding. The interface design presented in this part of the document is designed in accordance with the use cases and scenarios in the SRS document.
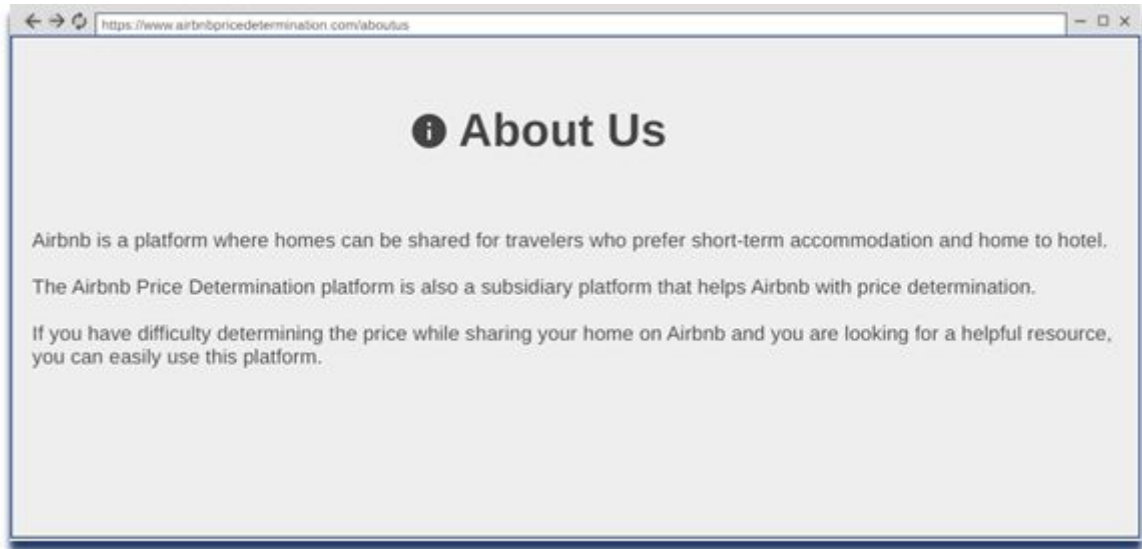
5.5.1 Home Page
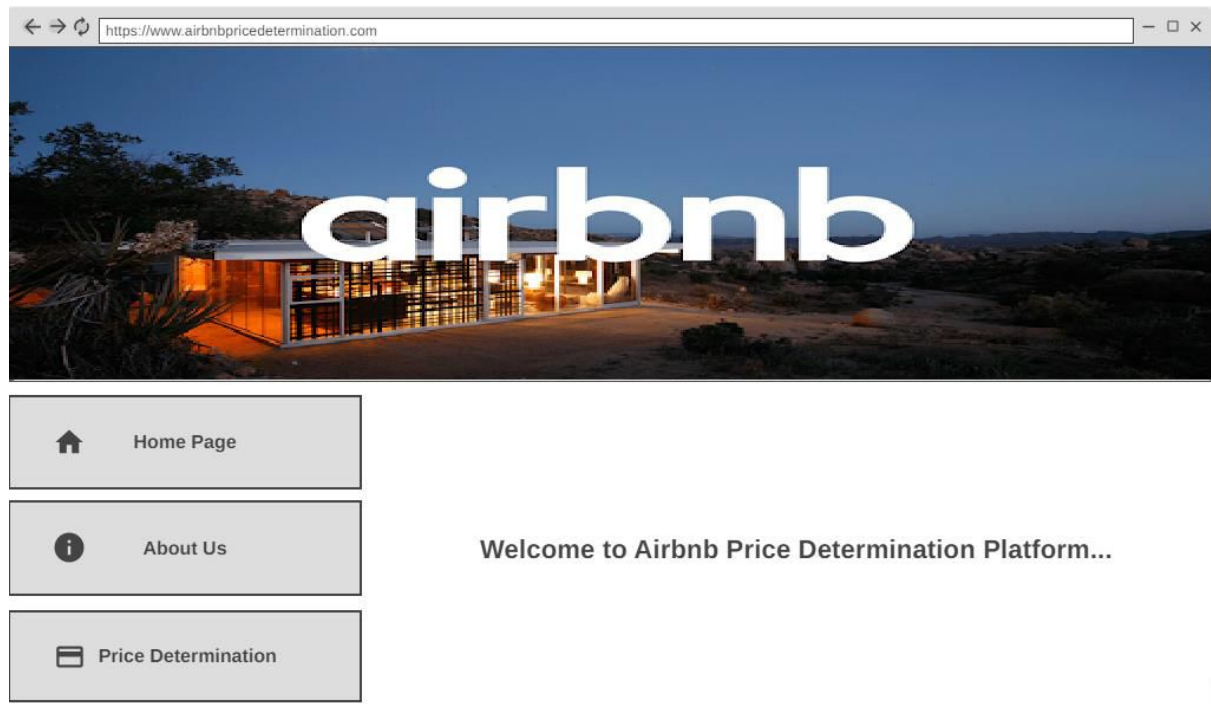
Figure 3: Home Page

5.5.2 About Us Page

When the Airbnb Price Determination platform is opened, this page is accessed when the About Us button is clicked on the main page.

Figure 4: About Us Page

5.5.3 Price Determination Page

When the Airbnb Price Determination platform is opened, this page is accessed when the Price Determination button on the main page is clicked.

On this page, there are features that will determine the price. The user must fill in these features according to the property to be placed on Airbnb. Features such as house type, room type, neighborhood, number of bathrooms and bedrooms are general price determination criteria for Airbnb. However, there will be additions according to the features that will appear as a result of the model to be used in the section specified as extra features.

Figure 5: Price Determination Page

Figure 6: Price Determination Page

5.5.4 Estimated Price

After the features on the Price Determination page are filled, the estimated price value will appear on this page by clicking the See Price button.
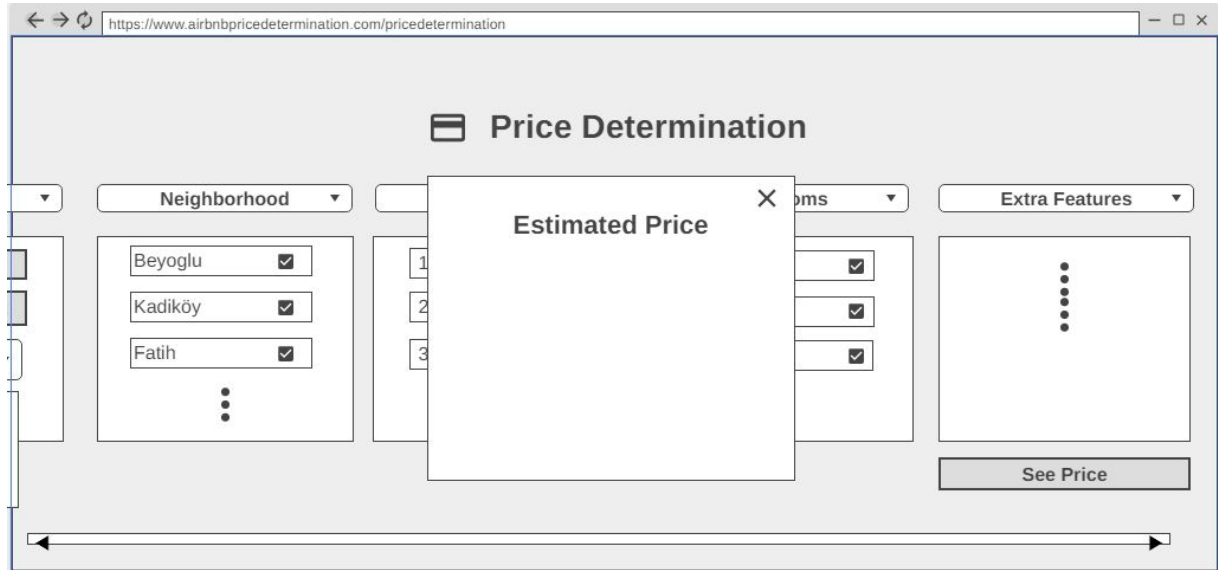

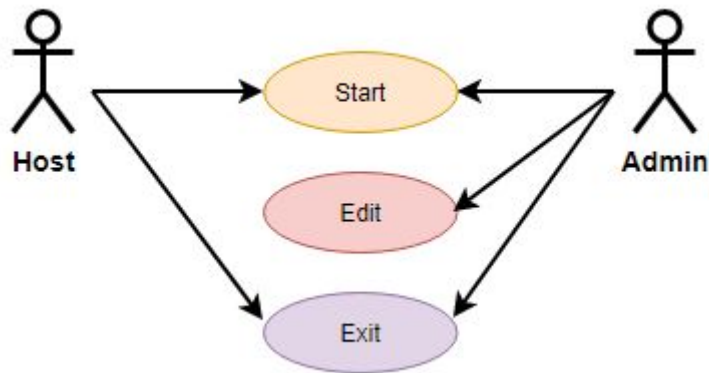*Figure 7: See The Price*

5.6 Process Design

5.6.1 Use Cases

Figure 9: Profile Management Use Case

Admin: This type of user can develop and manage the platform.

Host: This type of user is people who share their homes on the Airbnb platform. Meanly, he/she is the owner.

Table 1: Actions for Admin User Type

Actor	Action	
Admin	Start	Indicates access to the platform
	Edit	It is the act of performing operations such as the database of the platform and interface design
	Exit	Indicates leaving the platform

Actor	Action	
Host	Start	Indicates access to the platform
	Exit	Indicates leaving the platform

Figure 10: Host Use Case

Actor	Action	
Host	Select House Type	There are two types of house types, entire house and entire apartment
	Select Room Type	There are two types of room types, shared and private.
	Select Neighborhood	It specifies the selection from the list of neighborhood according to the location of Airbnb
	Select Number of Bathrooms	Indicates the selection of the number of bathrooms in house types
	Select Number of	Indicates the selection of the number

	Bedrooms	of bedrooms in house types
	Select Extra Features	It is the selection of the criteria of the price determination features to be added according to the result of the model developed according to the data.

5.6.2   Sequence Diagram

In this section, the actions that the host user can do are specified using the diagram. Accordingly, the host must first access the platform, then open the price determination page on the platform and select the features in the appropriate way. Finally, when you click the see price button, it should be able to access the estimated price.
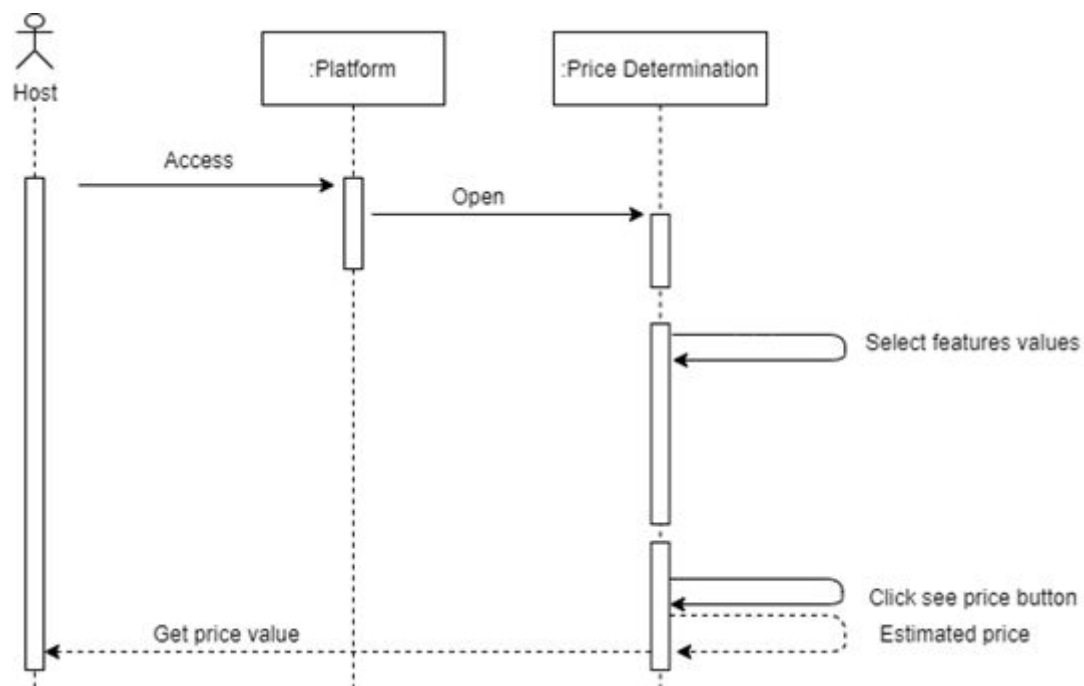

Figure 11: Sequence Diagram of Airbnb Price Determination Platform

5.7   Database Design
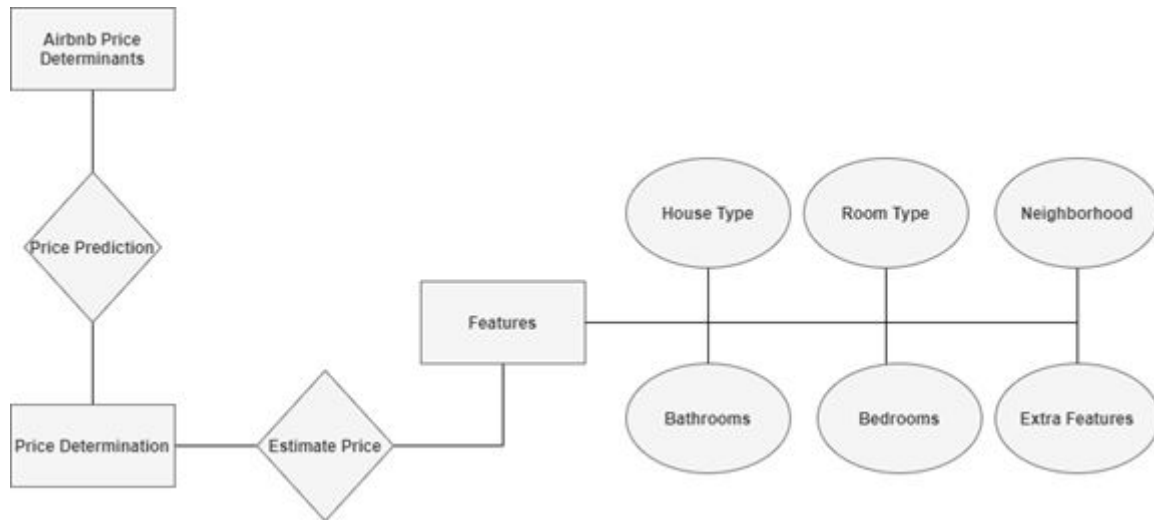
5.7.1 ER Diagram

Figure 12 : ER Diagram

6.   Conclusions

As a result, thanks to this project, we learned some machine learning algorithms and models while examining the literature of similar projects. We had the opportunity to review and analyze Airbnb datasets. After choosing the appropriate dataset, we decided on the algorithms and model we will use for price determination. We have determined the functions and users required for our system in order to create a price estimate. While doing all these, we learned to determine the requirements of a system and to create its design. In this process, before doing any coding, we had difficulties in designing a system, determining the functions and features we will choose to determine the price and choosing the appropriate dataset, but by overcoming all these difficulties, we both learned to design a project and developed ourselves. In our system, house type, room type, number of bedrooms / bathrooms, neighborhood and extra features are used for price determination, but the extra features part is missing in our system. In the coding part, after the results of Machine Learning algorithms, we will add other features in our data.

Acknowledgement

References

[1] V. Raul Perez-Sanchez , Leticia Serrano-Estrada , Pablo Marti , Raul-Tomas Mora-Garcia. The What, Where, and Why of Airbnb Price Determinants. December 2018.

[2] Yuan Caı,Yongbo Zhou, Jıanyu (Jenny) Ma, Noel Scott. Price Determinants of Airbnb Listings: Evidence From Hong Kong. May 2019.

[3] Insu Hong, Changsok Yoo. Analyzing Spatial Variance of Airbnb Pricing Determinants Using Multiscale GWR Approach. June 2020.

[4] Yuanhang Luo (royluo), Xuanyu Zhou (xuanyu98), Yulian Zhou (zhouyl). Predicting Airbnb Listing Price Across Different Cities. December 14, 2019

[5] IEEE Recommended Practice for Software Design Descriptions

[6] IEEE Recommended Practice for Software Requirements Specifications

[7] IEEE Standard for Software Project Management Plans

[8] https://wireframe.cc

[9] http://insideairbnb.com/get-the-data.html

[10] https://www.php.net/

[11] https://www.microsoft.com/tr-tr/sql-server/sql-server-2019

[12] https://www.draw.io