E-Valuer

(An Intelligent tool to assist in making smarter property related decisions)

System Requirement Specification

Mrs. M.P.A.W. Gamage

Mrs. Pasangi Rathnayake

Rodrigo U.S.D.   IT 16 15 4490

Project ID: 19-010

Sri Lanka Institute of Information Technology

Special Honours Degree of Bachelor of Science in Information Technology

May 2018

E-Valuer

AN INTELLIGENT TOOL TO ASSIST IN MAKING SMARTER PROPERTY DECISIONS

Project ID: 19-010

System Requirement Specification

(System requirement specification documentation submitted in partial fulfilment of the requirement for the Degree of Bachelor of Science Special (Honors)

In Information Technology)

Bachelor of Science (Honors) in Information Technology Specialized in Information Technology/ Software Engineering

Department of Information Technology

Sri Lanka Institute of Information Technology

Sri Lanka

May 2019

DECLARATION

I declare that this is my own work and dissertation does not incorporate without acknowledgment of any material previously submitted for a Degree or Diploma in SLIIT or any other university or institute of higher learning. To the best of my knowledge and belief the document does not contain any material previously published or written by another person except where the acknowledgment is made in the text

Project ID : 19-010

Project Team Members:

Student Name	Registration No	Signature
Rodrigo U.S.D.	IT16154490	

The above candidates are carrying out research for the undergraduate dissertation under my supervision.

Signature of the supervisor : ....

Date : ....

Signature of the co-supervisor : ....

Date : ....

TABLE OF CONTENTS

List of Figures

List of Tables

1 INTRODUCTION

Software requirements specification is the document which provides all of the information of the requirements of the system as expected by the end users and provides clarification of requirements for other stakeholders of the system. This is supposed to contain well detailed description of the functional and non-functional requirements of the proposed system, which will be used in the process of software development, quality assurance, and maintenance of the system.

1.1 Purpose

The purpose of the software requirement specification (SRS) is to present a detailed description of Current Value Predictor with Artificial Intelligence model which makes predictions based on a time series machine learning algorithm, of the project 'E-Valuer', a web application that predicts the current and future value of a land and factors influencing the same to provide the users with accurate computerized prediction system along with a detailed report of the land selected. The SRS document gives a detailed description on the functional and non-functional requirements as expected by the user. This document will demonstrate what the main requirements are and interfaces that the system will have after the implementation and the interconnections. The SRS is the basic document to refer when developing the system and help to understand how developers are going to develop the proposed system as well as the quality assurance(QA) engineers to design the test cases and other testimonies of QA.

1.2 Scope

This document covers the requirements for the first release of E-Valuer a land value prediction system which will be assist the users in making smart property related decisions. The purpose of this is to guide developers in selecting a design that will be able to accommodate the full-scale application expected in this release. It provides the clear overall design of the system which will provide the foundation for the development of the final end product.

• Functional requirements- specific functionalities, tasks or behaviors of the system

• Non-functional requirements - constraints, performance factors, etc...

The purpose of this application is to provide its users with accurate prediction of the value of the land which is similar to an estimation given by a professional valuer. Users can get the detailed report of the land with minimal effort needed which includes current value prediction based on the correct analyzing of the significance of the properties that can be computationally analyzed to help users for deciding the suitability of the land for their intended purposes. There are numerous factors which affect the value of a land such as physical factors , economic factors, and social factors. The details of the physical factors and economical factors along with recent valuation details will be considered in this application when deciding the ultimate output , prediction of the current value. In this application, the effect of social factors will not be taken into consideration as much as above mentioned because of the unavailability of proper information infrastructure to analyze the same factors. But we assume it will not affect the accuracy of the output since recent valuations are taken into consideration.

1.2.1 Main benefits of the system

This project is aimed to develop a portable application which assists people by providing them with accurate valuation, facts about how the land is going to be affected by various means of development projects, ultimately to decide whether it would be useful for their expected purpose.

1.2.2 Objectives of the system

● Identifying the most accurate algorithm from non-conventional time based approach of prediction for example Recurrent Neural networks (RNN) and non-conventional non-time based approach for such as Artificial Neural Networks (ANN) in the domain of prediction of values in the domain of current value prediction

● Identifying methods to optimize the values provided as current value prediction so that the procedure could be used as a reliable option.

● Identifying method to predict future value based on the fluctuation rates and records of weather conditions.

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

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

1.2.3 Main goal of the system

The main objective of our research is to develop a portable application which can provide instant report of a selected land parcel that gives the users an insight of the land with current value and future value.

1.3 Definitions, Acronyms, and Abbreviations

Acronyms

SRS	System Requirement Specification
ROI	Return of Investment
RNN	Recurrent Neural Networks
ANN	Artificial Neural Networks
AI	Artificial Intelligence
QA	Quality Assurance

Table 1.1 : Abbreviations

1.5 Overview

The document is intended to describe the requirements, both functional and non-functional requirements of developing a land value prediction system. It is a web application which is compatible to use with a desktop computer and a mobile device that help the users to get an accurate insight of the land they are going to buy.

Chapter 1: Brief the purpose of preparing the SRS document. Scope consist of a description that clearly mention the tasks that the project team will do and will not do, it also describes the benefits, objectives, and goals of the particular software. Overview section explains how the SRS is organized and describes what the rest of this document contains in a brief manner.

Chapter 2: Contains overall description of SRS in user's point of view that is, it is described in non-technical terms which can be understood by the users. It includes Product perspective, Product functions, User Characteristics, Constraints, Assumptions and Dependencies and Apportioning of requirements. Main target of Product perspective is to compare its similarity and differences to other systems in the marketplace and find whether the existing system is available in regard for the developing application. Product functions are mentioned by summarizing major functions of the application. In User Characteristics, it Indicate what kind of people the typical user is likely to be. In Constraints sub section describes all conditions that may limit developer's options they can be originating from many sources. Assumptions and dependencies sub section describes any assumptions being made during the development of the application.

Chapter 3: This section describes developer's point of view of the system. Use technical words/phases understand by the software engineers, developers, and testers. External interface requirements, performance requirements, design constraints, application attributes and other requirements also explained in advance.

2. OVERALL DESCRIPTIONS

Real Property are the most valuable possession of most of the common people. Getting the proper valuation for these real properties is very much important. This system is an innovative solution proposed to facilitate land valuation based on recent sales, prediction of future price and the effect of proposed development work on the land, so that real-estate customers and owners of real estate companies can be benefitted and make smarter property related decisions.

This intelligent tool can help people to identify the land they are willing to buy. This current value predictor component is responsible for providing accurate current valuation of the land since it is difficult to find out the reliable average prices of the land around manually, especially if the area is unfamiliar. We can get the service of a professional valuer, but that process is known to be subjective to the person and time

Land valuation is the process of assessing the characteristics of a given piece of land based on experience and judgment.[1] The determination of a land parcel value depends on a number of physical and economic characteristics which must be taken into consideration very carefully in a land valuation procedure.[1] These values can be affected by various social factors too. For example, if there is a crime happened in that land, it can cause a negative effect on the value..

Hence, real estate appraisal it is a challenging multidimensional problem that involves estimating many facets of a property, its neighborhood, and its city.[2]

Since, Sri Lanka is lacking a good data platform to gather all these data, considering all these factors can take ages to do proper valuation considering all these factors.

The manual process is a time-consuming slow task which needs to be done by an experienced professional valuer. The valuation approaches used by those professionals are limited due to the lack of digital data in Sri Lanka. Also, it is a known fact that the valuation process can be so subjective to the person.

All above mentioned problems can be solved with the implementation of the system which will save time, cost, and with improved fairness of the predictions which will be done.

2.1 Product perspective

Existing solutions

The use of AI for residential value forecasting has been suggested in the literature from 1990s. [4]. Although Sri Lanka is lacking an automated land valuation system, many up and running, reliable solutions have been implemented in developed countries like New Zealand, England and Wales, USA etc. It is obvious with the well-structured digital data infrastructure of those countries, they can implement very accurate systems. Our intention is to identify the ways to use their underlying methodology in a suitable manner in Sri Lankan context.

1.Zillow Zestimate

Zillow is an online real estate database company that was founded in 2006, and was created by Rich Barton and Lloyd Frink, former Microsoft executives and founders of Microsoft spin-off Expedia. [5] Zillow.com supports United States of America (USA) and Canadian property listing. Zillow compliments that Zestimate provides forecast for 12 months with below accuracy rates.

Model	Average Absolute % Error	Improvement over Naïve
Naïve Forecast	7.35%	0%
County Forecast	6.47%	11.9%
Zestimate Forecast	5.84%	20.5%

Table 2.1 : Average absolute percentage error of the 12-month forecast.

Source: *https://www.zillow.com/research/zestimate-forecast-methodology/*

Features :

● Estimates for 12 months

Zestimate determines an estimation for 12 months for a house based on neighbourhood comparable houses. Accuracy of zestimate depends on the amount of data used as the underlying approach is Hedonic regression analysis based proprietary algorithm [6] which

analyses of several features of the house. The forecasted value is interpolated using cubic spline to connect to current value. [6]

2.Trulia

Trulia is also a product offered in USA, which offers a range of services for real estate sector. The price estimates are based on publicly available information the home's physical characteristics (e.g. location, number of bedrooms, etc.), Property tax information, Recent sales of similar nearby homes.

It involves more community interaction, for example, Trulia Neighbourhoods provide photographs, drone footage, etc. so that who are interested about the neighbourhood can refer. Trulia provides price using public data which shows the price fluctuation of a house, comparative to the other homes with same ZIP code.

Below is the accuracy report of Trulia estimates.

National	Within 5% of Sale Price	Within 10% of Sale Price	Within 20% of Sale Price	Median Error
United States	48.2%	67.7%	82.3%	5.3%

Table 2.2: Trulia accuracy report

Source:*www.trulia.com/info/trulia-estimates/*

Features -

● Crime map - Crime map data is sourced from CrimeReports.com and SpotCrime.com, which aggregate crime data from law enforcement agencies and news reports.

● Local schools with schools rating - Data of the schools around the premises with details such as Grades taught, GreatSchool Score.

● Commute times at a glance - Using data from OpenStreetMaps and General Transit Feed Specification (GTFS) feeds, the user can get an idea of commute times at a glance.[7]

3. QV.co.nz - QV homeguide

Quotable Value (QV) provides independent and authoritative information on any home in New Zealand on or off the market [8] QV.co.nz and their mobile App QV homeguide is known to be providing more accurate values of real estate property and key details to assist people to make instant decisions regarding property. QV with CoreLogic, a company which analyzes information assets and data to provide clients with analytics and customized data services provide a range of reports valuable to the user.

Features - QV homeguide app

● Online Value Estimation - Provides the likely selling price of a property during that particular time

● Sales activity - Sales activity specific property found on the app

● Suburb Demographics - Median price data, Demographic data, Current listings, and latest auction results [9]

● E-Valuer Report - Subjected to a fee complete valuation report of the property can be downloaded.

4. HousePrice.ai

Creating a methodology that would bring more sophisticated information, greater accuracy and analytical rigor to the United Kingdom (UK) residential property market is the motivation behind HousePrice.ai. Their proprietary model provides a combination of multi-disciplinary experiences of AI and Big Data to provide most accurate estimations. HousePrice.ai has Horizon app, which calculates capital, rental and gross development values for a single property or an entire portfolio. [10]

Features-

●   Current and Future value prediction - Produces accurate property valuations both in the present time and can offer future predictions. Valuations are based on objective measurable values, creating a fact-based result as opposed to a subjective one [11]. This tool allows the user to adjust, add and remove factors within the surrounding areas to determine how external changes will affect property prices
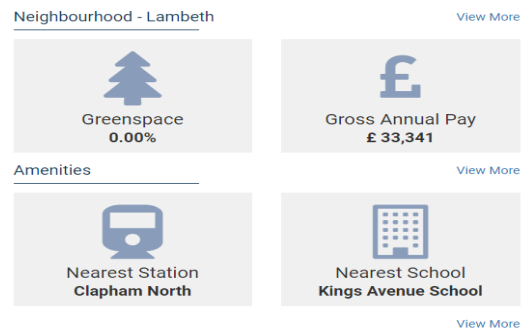
●   Distance to Schools, commutes etc.

Figure 2.1 : Brief Neighbourhood analysis

Source : *Sample Valuation Report - HousePrice.ai , Horizon, https://myhorizon.io/valueReport*

•   Historical data relevant to location
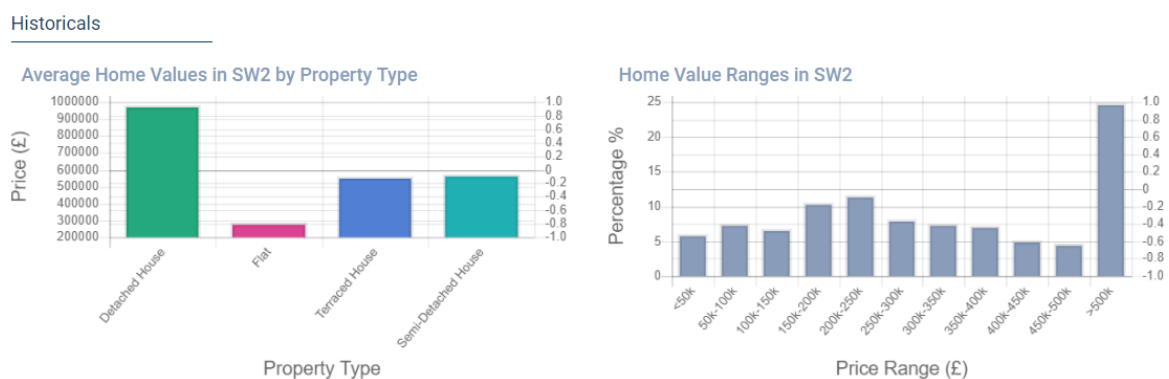
Figure 2.2: Historical Sales analysis

Source : *Sample Valuation Report - HousePrice.ai , Horizon https://myhorizon.io/valueReport*

Comparison of Existing Systems for the features of Current Value predictor module

	Zillow	Trulia	QV-CoreLogic	HousePrice.ai	Our Product
Current Value Prediction	Yes	Yes	Yes	Yes	Yes
Provision of possible natural hazards of the area	No	Yes	No	Yes	Yes
Use of machine learning algorithm	Yes	Unknown	No	Yes	Yes
Data used for predictions	Statistical data	Publicaly available data	Statistical data	Publicaly available data	Data provided by Valuers, and relevant departments
Mobile / Desktop / Web Application	Web	Web	Web	Web	Web
Available for Sri Lanka	No	No	No	No	Yes

Table 2.3: Comparison of existing systems

2.1.1 System interfaces

• The Application

The application must be hosted to be accessible by public. Users might access the system via desktop computer or a mobile device. Web application commonly use a combination of server-side script (python) and client-side script (Angular 6) to develop the application. Accessible for a range of devices, Adaptable to increased workload are some benefits of using web application.

• Database Access

Proposed database is intended to store , retrieve , update and manipulate information to system which includes,

o Store historical data relevant to predictions provided by valuers

o Retrieve data and feed the Artificial intelligence model

• GPS access

Acquire the current location

• Machine Learning with time-series algorithm (RNN)

RNN can solve complex problems by chaining together lots of simple neurons. Here will be created an RNN to estimate the value of a land parcel and to analyze the performance of time series algorithm in the domain of land value prediction. Machine learning is used because we have adequate sample data provided by professional valuers.
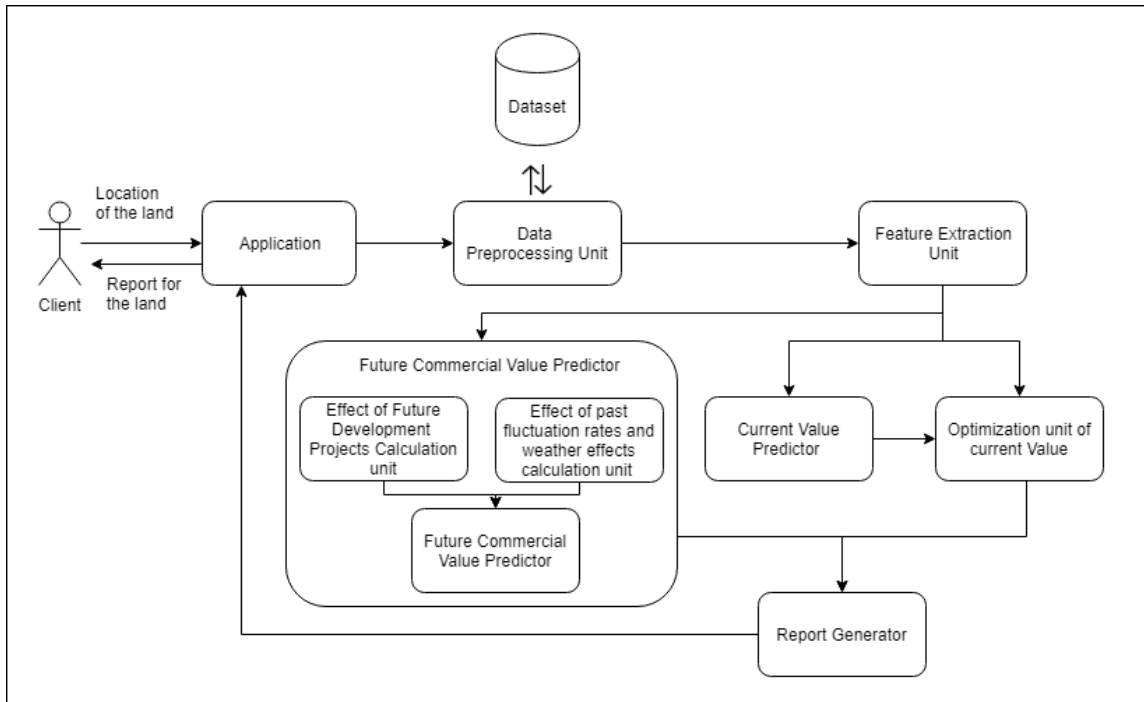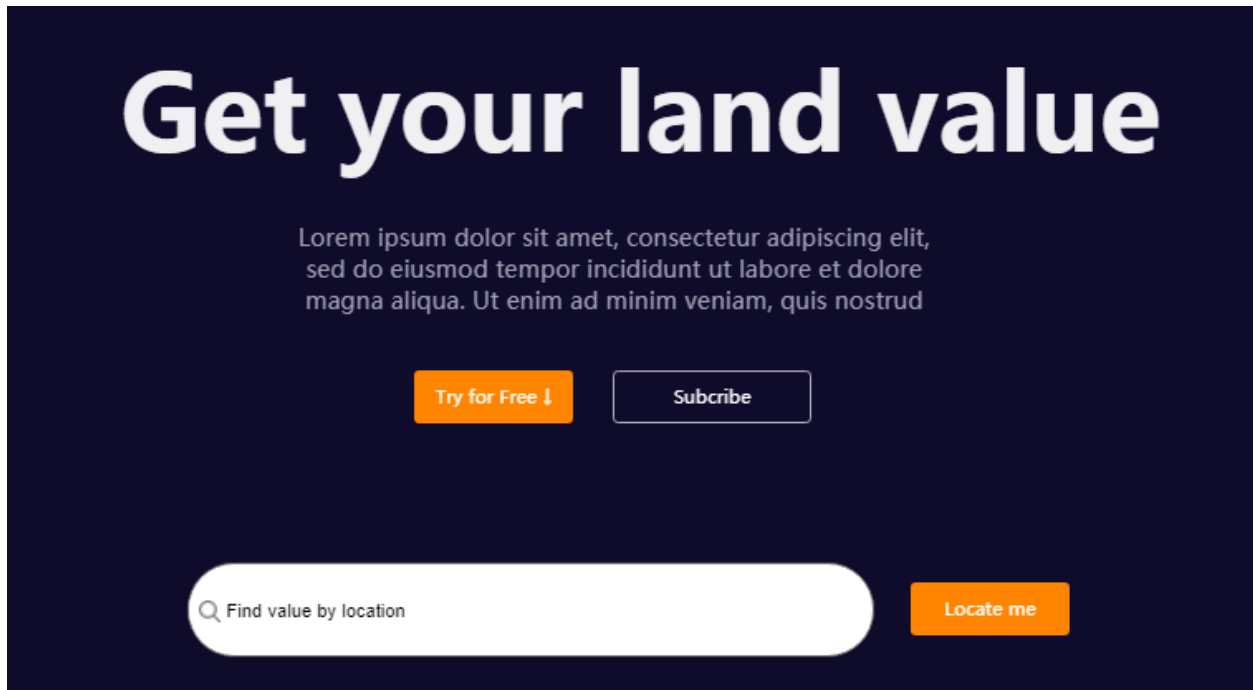
Figure 2.3: Architecture Diagram

2.1.2 User interfaces
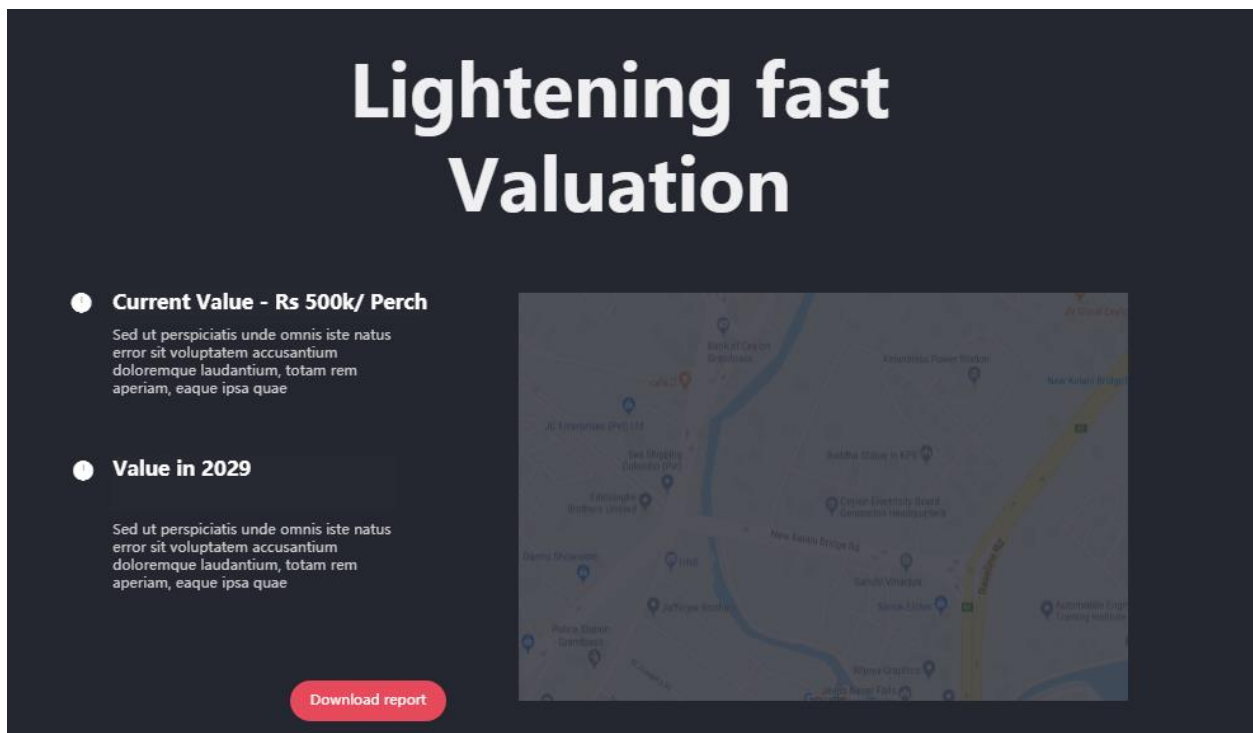

Figure 2.4 : Location input screen

Figure 2.5 : Basic valuation output and download option

Figure 2.6 : Input screen for data collection

2.1.3 Hardware interfaces

For clients we expect a smart mobile phone / tablet / laptop / desktop computer with minimum RAM of 2GB with web browser

For the developer end, we expect a computer with

- CPU: Quadcore Processor

- RAM: At least 8 GB

- Storage: 1 TB

2.1.4 Software interfaces

• Application Tool

We use Android studio for develop our mobile application

• Database

For create our application we use MongoDB.

• Anaconda Navigator

We use Anaconda Navigator as the application launcher. It allows us to launch applications and easily manage conda packages, environments and channels without the need to use command line commands.

• Jupyter

Jupyter is an IDE we have used to develop our machine learning models and it is powerful interactive development environment for the Python language with advanced editing, interactive testing, debugging and introspection features.

2.1.5 Communication interfaces

Internet connectivity is important, other than that this system does not use any communicational interfaces.

To create an internet connection between the client and server the device needs SIM card used for accessing the broad band network works using

• 3G standards like CDMA, HSPA

• 4G, 4G LTE, WCDMA are available with data transmission rate around 144 KB/per second and above.

2.1.6 Memory constraints

The management application will not be needed internal or external memory space in device to install application since it is web hosted. But, it will be need some memory space for cache data and other data. Hence, it is expected to use no more than 16 MB of Ram and 20 MB of external storage.

2.1.7 Operations

• Predict the current price of the selected land parcel.

This component is responsible for developing an AI model which predicts the current price of a land parcel upon submission of the location based on data gathered by analyzing those submitted by valuers and other key factors identified as significant to the area, following a time-series based algorithm. One of our specific objectives is to test the accuracy of time-series algorithms and vice

versa in the domain of land value prediction. Here an RNN will be built to analyze data and make predictions.

•   Data gathering module for the professional valuers who provide data

We expect to gather data from professional valuers as the training data for our AI model, the data gathering and storing module will also be developed under this component.

2.1.8 Site adaptation requirements

We consider that the content displayed regarding valuer data to be customized according to the location since it is a significant factor which can give user an impression of the accuracy of the system at a glance.

2.2 Product functions
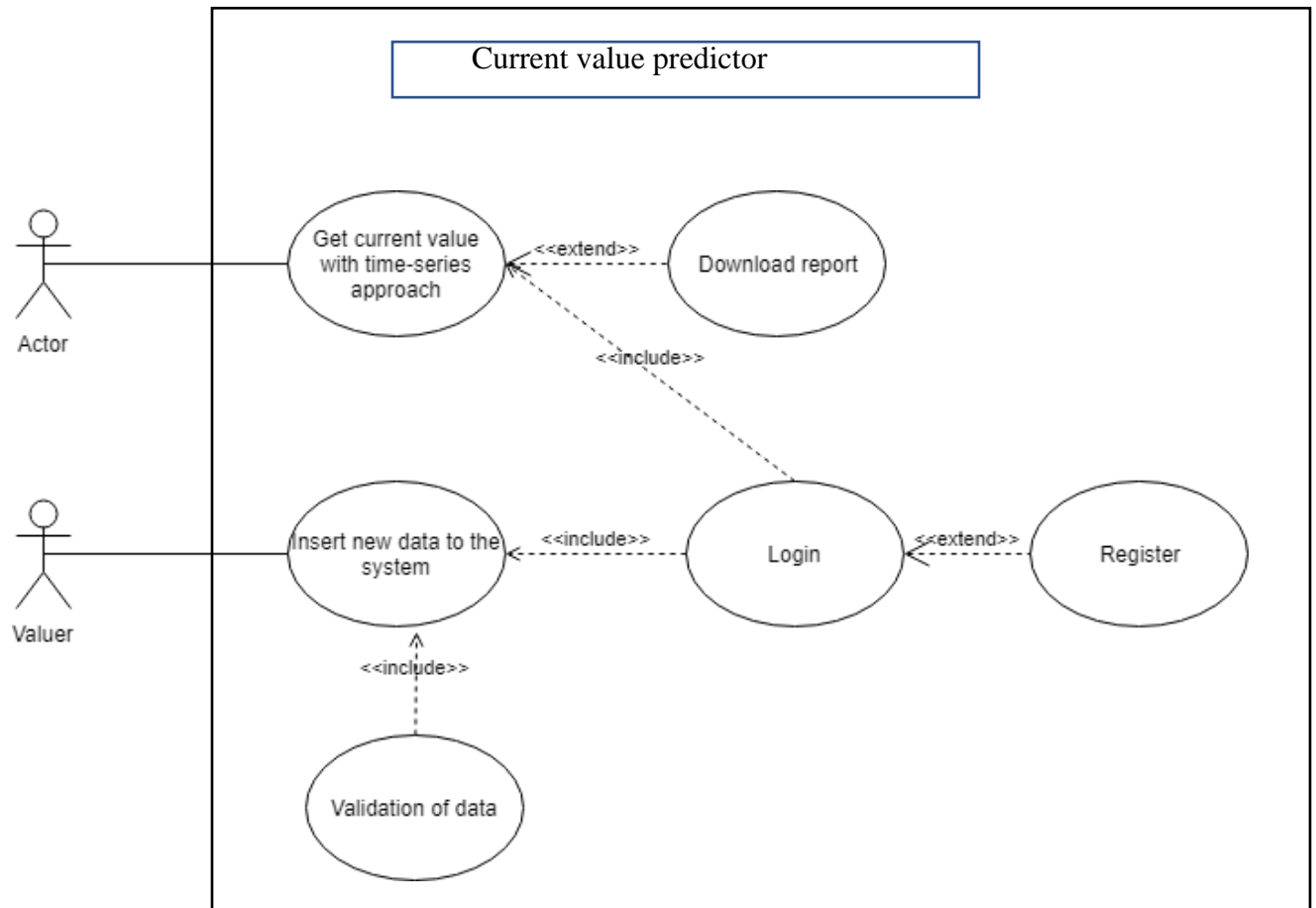
2.2.1 Use case diagram for Current Value Predictor component

Figure 2.7 : Use case diagram

2.2.2 Use case scenarios

Use case name	Predicting current price using machine learning model
Actor	User (Buyer/ Land sale owner)
Precondition	User must be logged in
Flow of events	1.User input the location of the land by manually searching for landmarks of google map or locate me option to input current location 2. Get the valuations provided by the system 3. Download the report
Exceptions	If the user cannot log on to the application, user cannot see the predicted prices.
Post Condition	-

| Use case name	Inserting new valuation data to the system
Actor	Professional Valuer
Precondition	User must be logged in
Flow of events	1.User inputs the data he/she has regarding a valuation done 2.Inserted data is validated with the predefined constraints 3.Save the data in private database 4.Update valuer profile
Exceptions	If the user cannot log on to the application, user cannot insert new data If the data is not validated those data will not be saved
Post Condition	-

2.2.3. System diagram for current value predictor
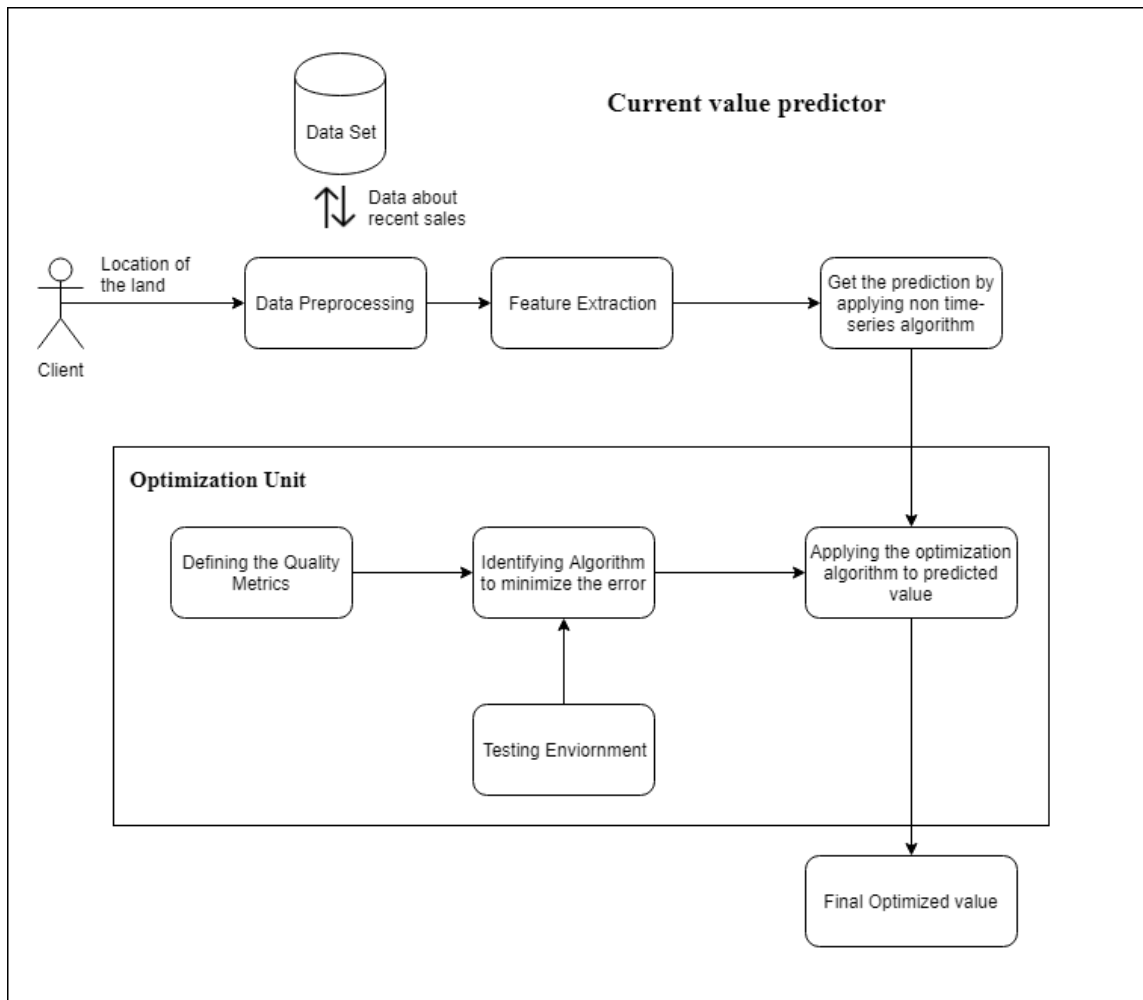
Figure 2.8 : System architecture diagram

2.3 User characteristics

Real Estate sector comprise of different types of stakeholders. Hence, the system is expected to be used by the listed users below.

- Customers who are going to buy a land

- Land sale companies

- Brokers

- Valuation officers

2.4 Constraints

1.The application is to be developed for the Colombo district at its initial stage. The scalability of the same to other districts of Sri Lanka is dependent on the availability of data.

2. Internet connectivity is needed to use the application

3. Data collected from professional valuers should be stored securely

4. System should be able to handle data traffic properly

2.5 Assumptions and dependencies

- Wi-Fi or mobile data available in the device used all the time with sufficient bandwidth.

- Stakeholders have knowledge on using our application and familiar with English language.

- Data and information will be limited to Colombo district at this release

2.6 Apportioning of requirements

The requirements described in sections 1 and 2 of this document are referred to as primary specifications; those in section 3 are referred to as requirements (or functional) specifications. The two levels of requirements are intended to be consistent. Inconsistencies are to be logged as defects. In the event that a requirement is stated within both primary and functional specifications, the application will be built from functional specification since it is more detailed.

'Essential requirements' (referred to in section 3) are to be implemented for this version of E valuer. 'Desirable requirements' are to be implemented in this release if possible, but are not committed to by the developers. It is anticipated that they will be part of future release. 'Optional requirements' will be implemented at the discretion of developers.

3. SPECIFIC REQUIREMENTS

3.1 External interface requirements

This segment includes all the functional and quality requirements of the web application

3.1.1 User interfaces

Previous section presents information oriented to the customer/user while section 3 is oriented to the developer.

It contains both content and format as follows:

- Name of item

- Description of purpose

- Source of input or destination of output

- Valid range, accuracy and/or tolerance

- Units of measure

- Timing

- Relationships to other inputs/outputs

- Screen formats/organization

- Window formats/organization

- Data formats

Login page for valuers

This page is the identification and authentication level of the developing mobile application. Here have to consider very sophisticated technology to login the page to identify that user is using a valid valuer login and he can have rights to login the web application. This page has two fields which are username and password. Those fields should be match with database particular user column upon authentication or otherwise should display proper message to alert the user. If a user attempt to login is failed for three times that user should to wait some time to login next time.

Main page

After a successful user login that user will be see the main page with our main input screen in order to input the location, the sole input we expect from the user to give the predictions. In the main page mainly a client oriented one where client does not need to wander through the application to get to the main functionality. Also, the administrator can access all features in the web application. These user interfaces might be subjected to changes depending on the customer requirements and developer concerns.

Insert new values

Once the logged in as a valuer , the above-mentioned link would appear in the same main page. There should be an option to view previous data provided by the user.

View user profile

Once the logged in as a valuer , the above-mentioned link would appear in the same main page. User would be able to view and download complete report of the data fed to the system as the first step.

3.1.2 Hardware interfaces

For the developer end, we expect a computer with

- CPU: Quadcore Processor

- RAM: At least 8 GB

- Storage: 1 TB

3.1.3 Software interfaces

Developer End:

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

- Anaconda distribution,

- Angular 6

- Google API

- Elasticsearch (Optional)

3.1.4 Communication interfaces

Internet connectivity is important, other than that this system does not use any communicational interfaces.

To create an internet connection between the client and server the device needs SIM card used for accessing the broad band network works using

• 3G standards like CDMA, HSPA

• 4G, 4G LTE, WCDMA are available with data transmission rate around 144 KB/per second and above.

3.2 Classes/Objects

This section should list classes pertaining to the domain of the application and are adequate for organizing all of the requirements. These classes are not all of the classes that will be used by the application.
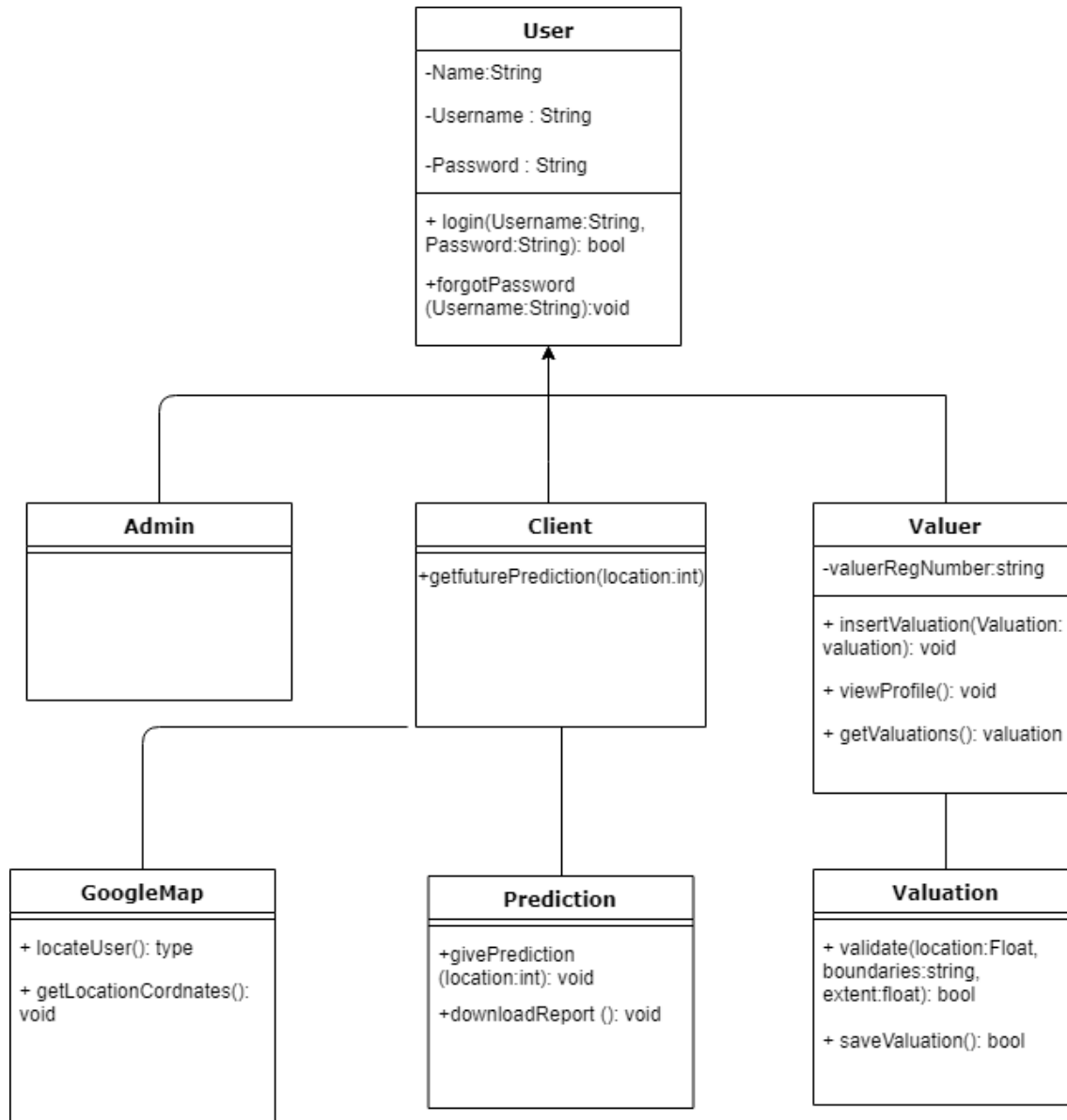

Figure 3.1 : Class diagram

3.3 Performance requirements

Response time

The system is expected to be running at its maximum efficiency upon the availability of sufficient memory and storage as mentioned under 3.1.2 Hardware requirements above.

The system is expected to produce outputs and reports promptly within 1-2 seconds to ensure the efficiency of the user experience. However, the internet connection properties such as bandwidth, connection type might affect the performance of the application.

Concurrent request handling

The system is expected to handle at least 10 requests from different users at a time with a minimal delay of 2 seconds. If the delay exceeds 30seconds the system should restart to give the user best experience.

3.4 Design constraints

Designers are free to create any (good) design that satisfies the requirements as well as to enhance the user experience considering the below mentioned facts.

• More than half of the people leave if the application doesn't load in three seconds

• Prepare your content before the design

• Knowing your target audience can help you decide

• Responsiveness of the design is the most important factor to be concerned regarding E-valuer

3.5 Software system attributes

3.5.1 Reliability

System shall fail not more than once in 60 days since it is expected to have higher number of users once launched and constant failures can create unreliable opinion among users.

Two algorithms of two different approaches are tested to ensure the accuracy and reliability of the output provided. Data is not allowed to be false and corrupted.

3.5.2 Availability

This system we can use 24 hours. The server should be available 98% of the time. The web service should do parallel tasks without crashing 98% of the time.

3.5.3 Security

System shall ask for user-name and password at the beginning. Passwords will be encrypted before saving. Session timeouts might be implemented.

Data provided by professionals will be stored in a private database with strict security measures preventing possible breaches.

3.5.4 Maintainability

Prediction models are expected to change the weights and other factors considered as constant with the development of additional features.

3.6 Other requirements

Inexpensive and Cost effective for user - We provide all the estimations and predictions for a fair price which is beneficial in terms of time and cost over the manual method of valuation.

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

Accuracy - Since AI and related algorithms which were proven to be suitable for Real Estate valuation will be used accuracy of the predictions can be guaranteed

4. REFERENCE LIST

[1] Li, L., Prussella, P.G.R.N.I., Gunathilake, M.D.E.K., Munasinghe, D.S. and Karadana, C.A., 2015. Land Valuation Systems using GIS Technology Case of Matara Urban Council Area, Sri Lanka. Bhumi, The Planning Research Journal, 4(2), pp.7–16.

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

[3] Schulz, R. (2003). Valuation of properties and economic models of real estate markets. Erscheinungsort nicht ermittelbar: Verlag nicht ermittelbar.

[4] Chaphalkar, N.B, & Sayali Sandbhor. (n.d.). Use of Artificial Intelligence in Real Property Valuation. Retrieved from http://www.enggjournals.com/ijet/docs/IJET13-05-03-087.pdf

[5] Zillow. (2019, February 21). Retrieved from https://en.wikipedia.org/wiki/Zillow#Zestimate [Accessed 23 Feb. 2019].

[6] Hagerty, James R. "How Good Are Zillow's Estimates?", The Wall Street Journal, 2007-02-14. Retrieved on 2009-02-25.[Accessed 23 Feb. 2019].

[7] Trulia. (n.d.). Retrieved February 22, 2019, from https://en.wikipedia.org/wiki/Trulia [Accessed 23 Feb. 2019].

[8] "QV Homeguide App Now Available." New Zealand Property Investors Federation, 3 Mar. 2015, www.nzpif.org.nz/news/view/56971. [Accessed 24 Feb. 2019].

[9] "Part 4: Quotable Value Limited's QV Homeguide Application." Office of the Auditor-General New Zealand, www.oag.govt.nz/2018/digital-access/part4.htm. [Accessed 24 Feb. 2019].

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

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