

# **GoldHunter - Project Report**

## **M.C.A Lateral IV Semester**

**Amrita School of Arts & Sciences**  
**Mysore Campus**

Amrita Vishwa Vidyapeetham  
#114, 7th Cross, Bogadi II Stage  
Mysore, Karnataka  
India - 570026



श्रद्धावान् लभते ज्ञानम्

Submitted by:

**Nagaraj Poti**  
**(MY.SC.P2MCA14078)**

Under Supervision of:

Mr. Shivaramu K N  
Asst. Professor  
Mr. Manishankar S  
Asst. Professor

## Abstract

Machine learning is a scientific discipline that explores the construction and study of algorithms that can learn from data. Such algorithms operate by building a model from example inputs and using that to make predictions or decisions, rather than following strictly static program instructions. **GoldHunter** is an open source tool that aims at allowing students and other users at understanding the working of various algorithms for data pre-processing, classification, regression, clustering, association rules, and visualization. The following report looks at some of the disadvantages of the currently existing solutions that are already available and lists the advantages of this new implementation.

## Introduction

Machine learning is a sub field of computer science stemming from research into artificial intelligence. It has strong ties to statistics and mathematical optimization, which deliver methods, theory and application domains to the field. Machine learning is employed in a range of computing tasks where designing and programming explicit, rule-based algorithms is infeasible.

Machine learning tasks are typically classified into three broad categories, depending on the nature of the learning "signal" or "feedback" available to a learning system. These are:

- i. Supervised learning - The computer is presented with example inputs and their desired outputs, given by a "teacher", and the goal is to learn a general rule that maps inputs to outputs.
- ii. Unsupervised learning - no labels are given to the learning algorithm, leaving it on its own to find structure in its input. Unsupervised learning can be a goal in itself (discovering hidden patterns in data) or a means towards an end.
- iii. In reinforcement learning, a computer program interacts with a dynamic environment in which it must perform a certain goal (such as driving a vehicle), without a teacher explicitly telling it whether it has come close to its goal or not.

**GoldHunter** software tool is an initiative to address the inherent difficulty of implementing machine learning algorithms. The inspiration to develop this software came from a desire to improve the quality of material available to academic scholars in the field of machine learning and data mining. Although there is one other software solution known as "Weka" that is available, the software is not designed to focus on specific needs of institutions and students; rather it solves the needs of advanced scholars and researchers. **GoldHunter** is aimed at providing features that would be of significant use to educational institutes.

## Objective

The objectives of **GoldHunter** are -

- ✓ Reduced operational algorithm search time
- ✓ Increased operational efficiency
- ✓ Provide an online and offline service
- ✓ Allow efficient comparison of the different machine learning algorithms
- ✓ Upload documents to an globally accessible online or locally accessible offline database
- ✓ Accept documents in xml, arff,xls formats
- ✓ Search databases and files for interesting patterns and classify them
- ✓ Use the tool for professional purposes of understanding customer buying patterns

This software package can be readily used by non-programming personnel such as:

- Students
- Research Scholars
- Professors/Teachers

## Existing Implementation

There are some software tools available for both commercial and non-commercial purposes. However, all the available solutions have one or more of the following drawbacks:

- Some of the more effective solutions are paid services whereas the free solutions are bare-bones.
- Database searches cannot be personalised based on certain parameters that a user may want to enforce
- Some of the current solutions do not provide an offline based service
- The existing solutions do not have added functionality like algorithm runtime comparison
- Algorithms implemented in current systems are not as optimised as they could be.
- Weka is a very advanced tool that is not suitable for beginners
- **GoldHunter** is the first effort at implementing such as a software tool using the .NET platform

## Proposed Modification

The proposed software, **GoldHunter**, aims for the ulterior goal of giving the student both a clear sense of how machine learning algorithms work, as well as a renewed respect for their complexity and power. The software aims to the following improvements over existing solutions:

- Easy to use GUI
- Provide option of curated database search to institutions
- Allow multiple formats to be accepted for processing

- Quicken algorithm run time
- Allow the uploading and running of algorithms over dynamic input parameters
- Be a tool that goes hand-in-hand with the study of machine learning at a basic level
- Make machine learning techniques generally available
- Provide a tool that can be used by businesses and individuals for data mining purposes

### Hardware requirements

#### Server side -

Processor: Intel Pentium IV or above

RAM: 1GB or more

Hard Disk: 40 GB or more

Display: EGA/VGA Colour Monitor

Key Board: 108 keys

Mouse: 3 Button scroll

#### Client side -

A computer enabled with http web browser functionality and .NET Framework 3.5 and above

### Software requirements

#### Server side -

Windows Operating system

.NET Framework 4.0 or above

Database server SQL Server

#### Development Tools

Microsoft Asp.NET 2012

Database-Microsoft SQL Server 2005 or above

Adobe Photo Shop

#### Documentation & Presentation Tools

Microsoft Word 2007

Microsoft Power Point 2007

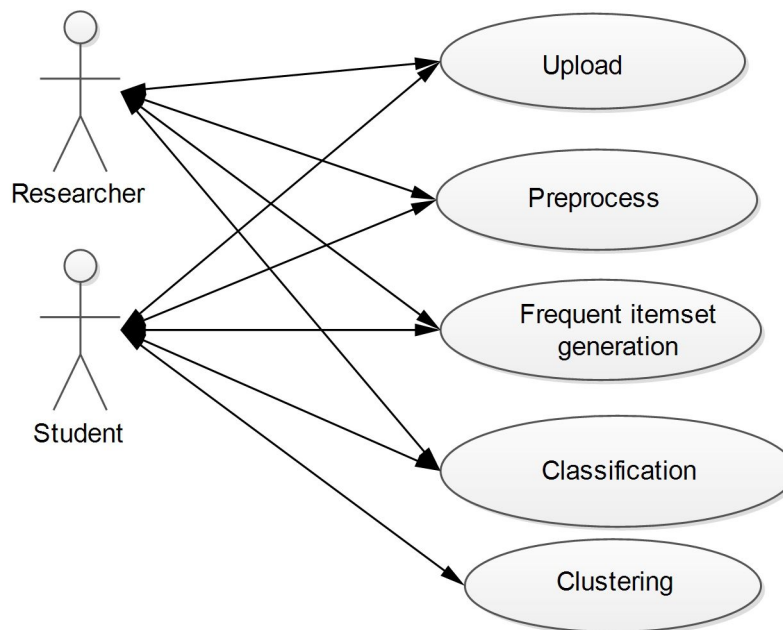
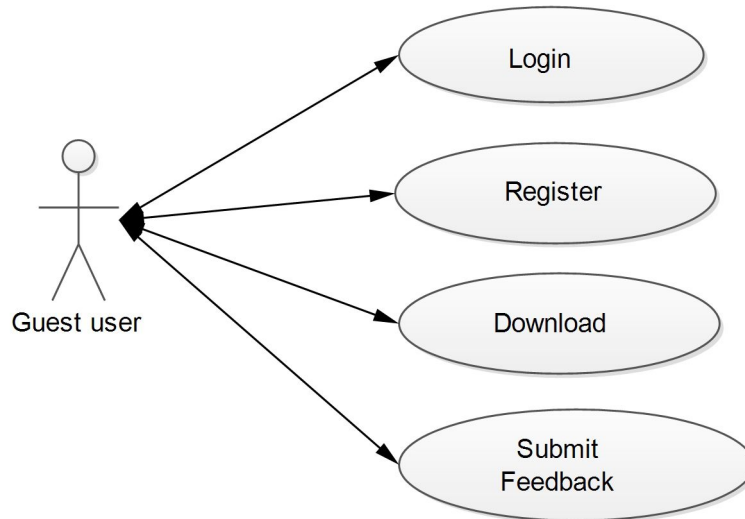
#### Client side -

Http web browser

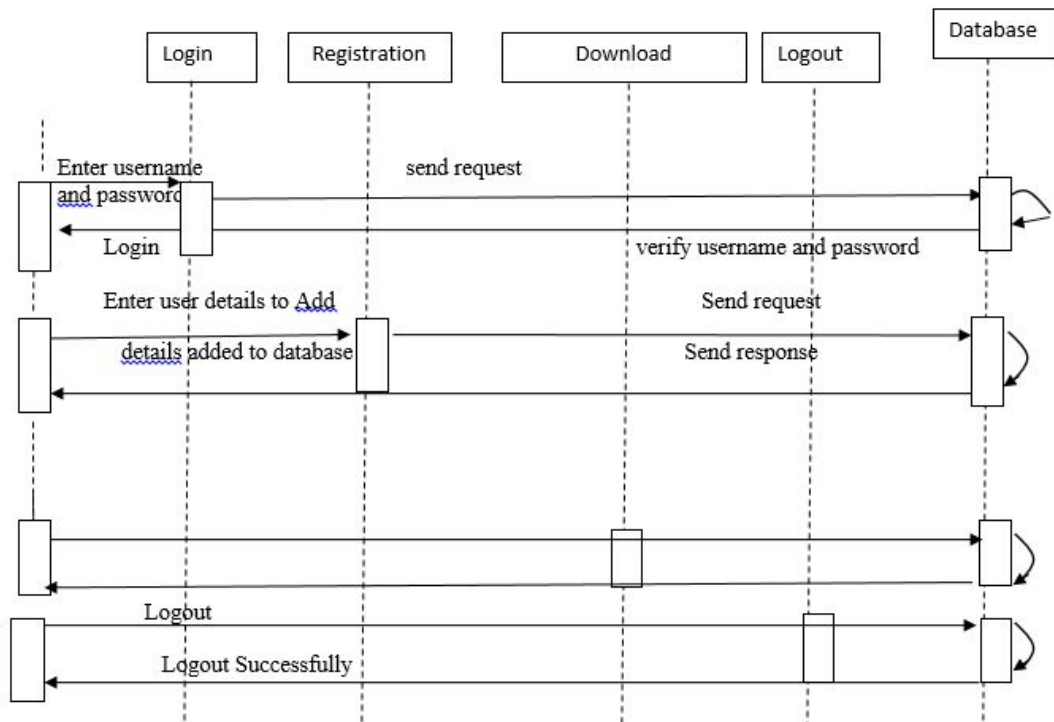
.NET Framework 3.5 and above

## Project Diagrams

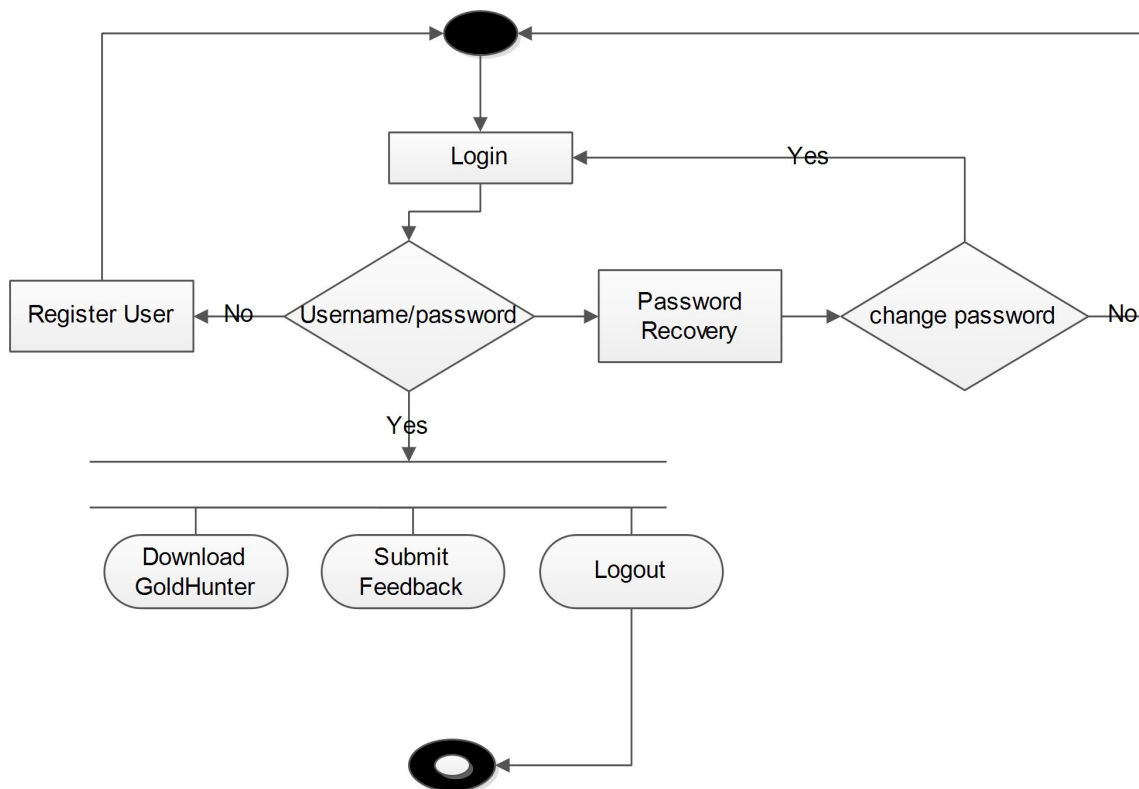
Use case diagrams:



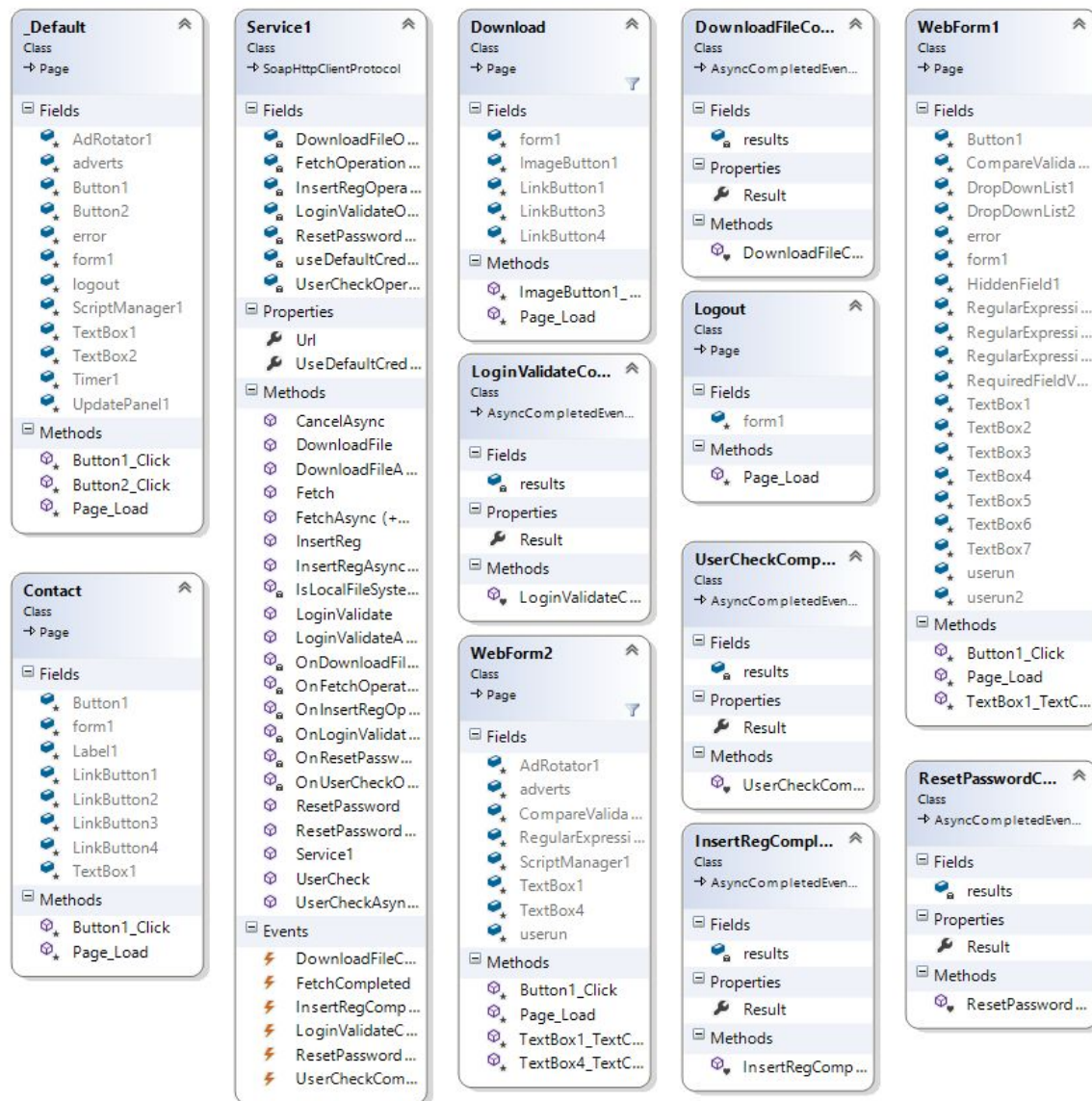
Sequence diagram:



Activity diagram:

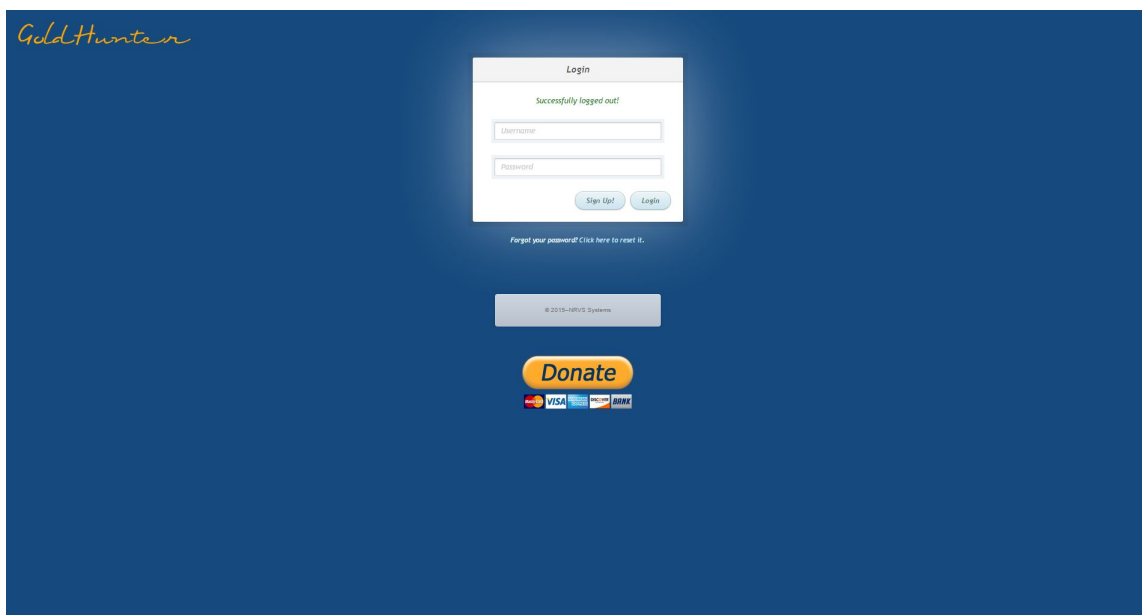
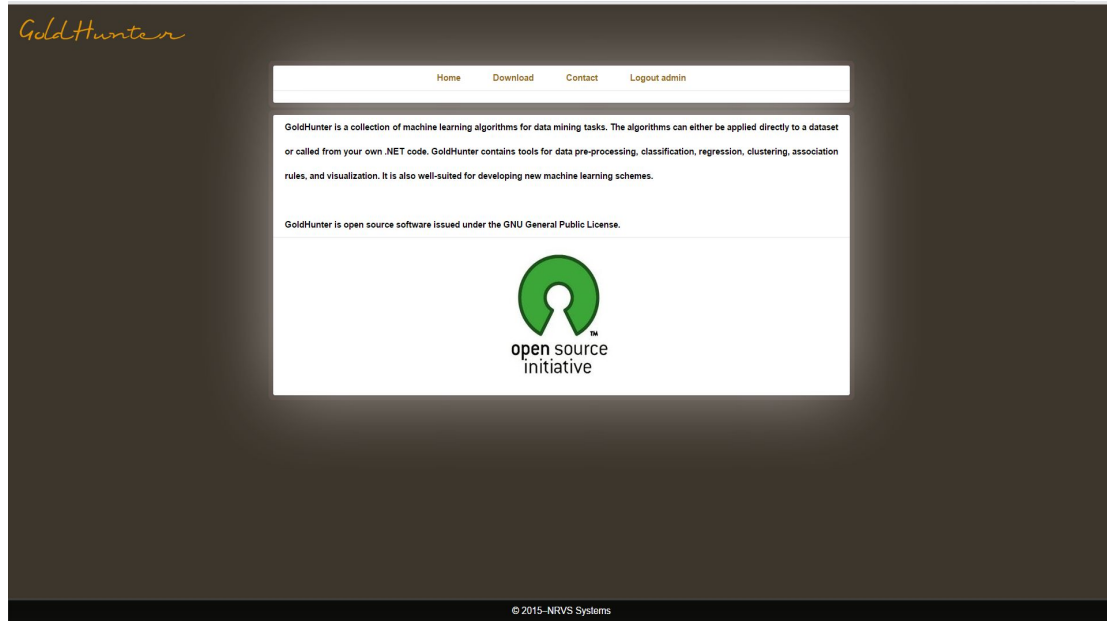


Class diagram:



## Screenshots

Web interface:





## Web service:

### Service1

The following operations are supported. For a formal definition, please review the [Service Description](#).

- [DownloadFile](#)
- [Fetch](#)
- [InsertReq](#)
- [LoginValidate](#)
- [ResetPassword](#)
- [UserCheck](#)

This web service is using <http://tempuri.org/> as its default namespace.

**Recommendation: Change the default namespace before the XML Web service is made public.**

Each XML Web service needs a unique namespace in order for client applications to distinguish it from other services on the Web. <http://tempuri.org/>

Your XML Web service should be identified by a namespace that you control. For example, you can use your company's Internet domain name as p

For XML Web services creating using ASP.NET, the default namespace can be changed using the WebService attribute's Namespace property. The W

"<http://microsoft.com/webservices/>";

C#

```
[WebService(Namespace="http://microsoft.com/webservices/")]
public class MyWebService {
    // implementation
}
```

Visual Basic

```
<WebService(Namespace="http://microsoft.com/webservices/")> Public Class MyWebService
    ' implementation
End Class
```

C++

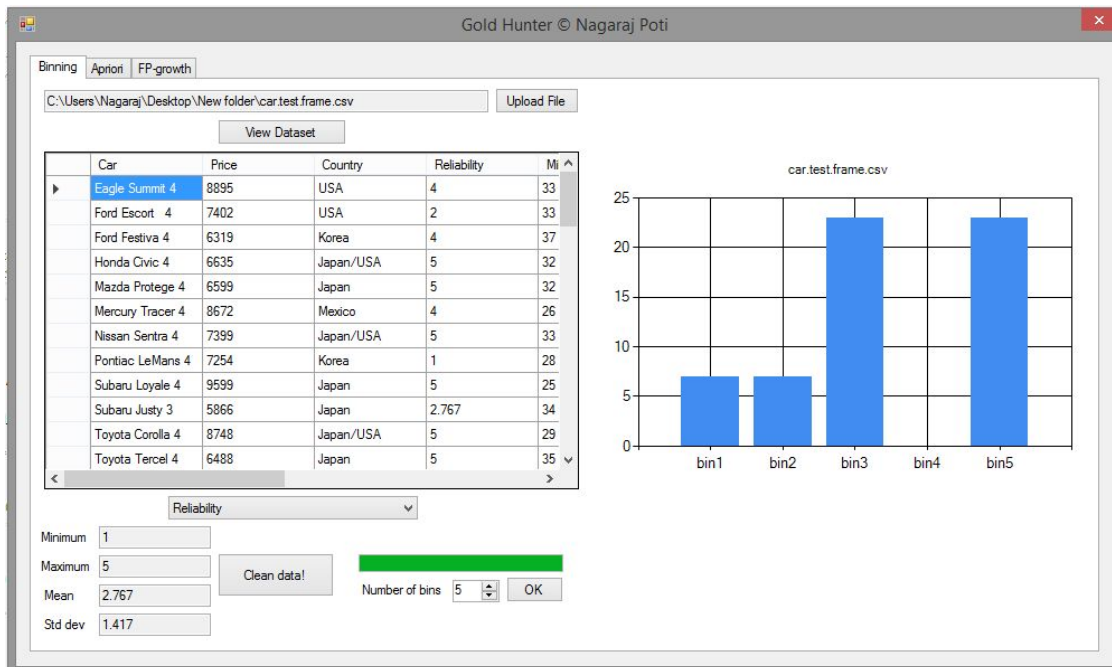
```
[WebService(Namespace="http://microsoft.com/webservices/")]
public ref class MyWebService {
    // implementation
};
```

For more details on XML namespaces, see the W3C recommendation on [Namespaces in XML](#).

For more details on WSDL, see the [WSDL Specification](#).

For more details on URIs, see [RFC 2396](#).

## GoldHunter:



## Conclusion

In this report, a brief summary of the **GoldHunter** software has been presented. The objectives and the system requirements of the software have been discussed. On a concluding note, the project is definitely feasible and more features can be incorporated after the release of early versions of the software.