Esteban Ariza
Johan Giraldo
Mateo Valdés
Juan Restrepo

Engineering Method

Problem description:

Alfonso is a genius biologist who loves to eat various kinds of mushrooms every day. One day he was eating mushrooms while reading the news and he found an article in the Smithsonian Magazine website titled "You May Have Been Eating Mushrooms That Were Unknown to Science" (link to the article: https://www.smithsonianmag.com/smart-news/you-may-have-been-eating-mushrooms-unknown-science-180951974/. For more information about the article, please consult the bibliography included at the end of this document). According to the article, many mushrooms, including some we eat every day, are still unknown to science and may be poisonous or harmful to our health. Since he eats many mushrooms every day, Alfonso got scared and shared the article with his biologist colleagues. His biologist colleagues then asked our team, TWICE Spiderman (a team made up of Esteban Ariza, Johan Sebastián Giraldo, Juan José Restrepo and Mateo Valdés), for help.

The biologists say that we must first read the data already available regarding the mushrooms. Furthermore, the biologists advise us to make graphs and charts about all the available data in order to get a better grasp of the information. They advise us to make graphs of edible vs. poisonous mushroom quantity, the number of mushrooms with a certain odor, number of rings in each of the mushrooms, number of bruises in each of the mushrooms, and the number of mushrooms with a certain cap color. Finally, they want us to use all of the information and determine, given any mushroom's attributes, whether the mushroom is edible or poisonous. They say that if we do this Alfonso will be able to enjoy mushrooms again without feeling scared and crying.

Identifying the problem:

The biologists want a way to determine if a mushroom is edible or poisonous given its attributes.

We must also create the graphs to represent the following quantities:

- Number of edible vs. poisonous mushrooms
- Number of mushrooms with a certain odor
- Number of mushrooms with a certain amount of rings
- Number of mushrooms with a certain amount of bruises
- Number of mushrooms with a certain cap color

Functional Requirements:

Name	FR1: Load data
Summary	The systems loads the data file of the mushrooms with the following attributes: ● Type ● Cap shape ● Cap surface ● Cap color ● Bruises ● Odor ● Gill attachment, ● Gill spacing ● Gill size ● Gill color ● Stalk shape ● Stalk root ● Stalk surface above ring ● Stalk surface below ring ● Stalk color above ring ● Stalk color below ring ● Veil type ● Veil color ● Ring number ● Ring type ● Spore print color ● Population ● Habitat
Input	The data file with all the mushroom attribute data
Output	The file has been loaded successfully

Name	FR2: Create the graphs
Summary	The system creates five graphs to represent the following quantities: ● Number of edible vs. poisonous mushrooms ● Number of mushrooms with a certain odor

	● Number of mushrooms with a certain amount of rings ● Number of mushrooms with a certain amount of bruises ● Number of mushrooms with a certain cap color
Input	The mushroom attribute data
Output	The graphs have been generated successfully

Name	FR3: Determine the type of a mushroom
Summary	The system determines the type (possible type: edible and poisonous) of a mushroom based on the following attributes about the mushroom: ● Cap shape ● Cap surface ● Cap color ● Bruises ● Odor ● Gill attachment, ● Gill spacing ● Gill size ● Gill color ● Stalk shape ● Stalk root ● Stalk surface above ring ● Stalk surface below ring ● Stalk color above ring ● Stalk color below ring ● Veil type ● Veil color ● Ring number ● Ring type ● Spore print color ● Population ● Habitat
Input	The attributes of the mushroom
Output	The type of the mushroom

Research:

Decision trees:

Decision trees is "a predictive model that divides the predictor space by grouping observations with similar values for the response or dependent variable" (Merayo, 2020).

The process of these is dividing the space of the predictors, grouping observations with similar values for a response or dependent variable, thus dividing the sample space into subregions, of which a series of rules or decisions applies, so that each subregion has the equal greater quantity of the individuals of one of the populations. If one of the sub-regions contains data different from the classes, it is divided into even smaller sub-regions until they are separated into smaller sub-regions that integrate data from the same class (Merayo, 2020).

Fungi:

Living beings that are part of the Fungi Kingdom, one of the groups in which life is classified, are commonly called fungi. There are more than 144,000 different species of fungi (caracteristicas.co, 2020). In turn, fungi are classified into four large groups: **Saprophytes** (They feed on the decomposition of organic matter), **Mycorrhizae** (Proliferates in plants, exchanging nutrients), **Lichenized** (Product of union with a cyanobacteria), **Parasite** (Poriferan inside the body of other living beings).(Máxima Uriarte,2020). We recognize when a mushroom is edible given its smell or taste, the shape of the hat. One of the edible mushrooms are the white cap mushrooms, characterized by their strong smell, short and thick feet, likewise truffles are another type of edible mushrooms that grow on the roots of plants. Finally we have the poisonous mushrooms, which their species are: Amanita Phalloides, Amanita Muscaria, Cortinarius Orellanus, Lactarius Scrobiculutus, etc. Many species of mushrooms, including some we eat every day, have not been discovered and classified by scientists yet, which, according to scientists, can be a problem because they could be poisonous or harmful to our overall health (Schultz, 2014).

Food and Agriculture Organization of the United Nations (FAO):

FAO is the agency charged with ending hunger. Its objective is to achieve food security for all and at the same time regulate sufficient food (...) with more than 194 member states, the FAO works in more than 130 countries. Finally, Develops international standards to guarantee quality food for all, maintains the largest and

most complete statistical database in the world on food and agriculture (Food and Agriculture Organization, 2020).

.NET Framework:

.NET Framework The .NET Framework is a "software development framework for building and running applications on Windows" (Microsoft Corporation, 2020). There are two major components in the .NET Framework. One component is the Common Language Runtime (CLR), which is an execution engine used to run the applications with features such as garbage collections, threading and exception handling; the other component is the Class Library, which contains various APIs and types to perform actions such as reading and writing files and drawing objects on the screen (Microsoft Corporation, 2020).

Kaggle:

Kaggle is "a crowd-sourced platform to attract, nurture, train and challenge data scientists from all around the world to solve data science, machine learning and predictive analytics problems" (Usmani, 2017). The platform has hundreds of thousands of active members from all over the world and receives more than 100 thousand submissions per month (Usmani, 2017). It contains courses, contests, and a huge collection of datasets.

Windows Forms:

Windows Form is a UI Framework that is part of the .NET Framework. It can be used to create desktop applications and various features to process and visualize data; furthermore, it has a visual designer integrated in Visual Studio (Microsoft Corporation, 2020).

Bibliography:

Microsoft Corporation (2020). What Is The .NET Framework? Retrieved from https://dotnet.microsoft.com/learn/dotnet/what-is-dotnet-framework

Microsoft Corporation (2020). Windows Forms for .NET 5 and .NET Core 3.1. Retrieved from https://docs.microsoft.com/en-us/dotnet/desktop/winforms/?view=netdesktop-5.0

Merayo, P. (2020). What are decision trees and what are they for? Retrieved From
https://www.maximaformacion.es/blog-dat/que-son-los-arboles-de-decision-y-para
-que-sirven/

Máxima, P. (2020). Information and characteristics of fungi. Retrieved from
https://www.caracteristicas.co/hongos/

Food and Agriculture Organization (2020). About FAO. Retrieved from
https://www.fao.org/about/es/

Schultz, C. (2014). You May Have Been Eating Mushrooms That Were Unknown
to   Science.   Retrieved   from
https://www.smithsonianmag.com/smart-news/you-may-have-been-eating-mushro
oms-unknown-science-180951974/

Uzman, Z. (2017). What is Kaggle, Why I Participate, What is the Impact?
Retrieved from https://www.kaggle.com/getting-started/44916

Creative solutions:

Brainstorming:

Alternative A (The expert alternative):

      Hire an expert who specializes in mushrooms to collect all of his knowledge and information available to him into a notebook or database, which he can organize and represent graphically in any way he chooses. Then, when we have a mushroom and need to classify it, we contact (in person or online) the expert and he uses his database, judgement and intuition to tell us whether the mushroom is poisonous or edible.

Alternative B (The lab rat alternative):

      Obtain all the information about the mushrooms from Kaggle and use Excel to create all the tables and charts needed to solve the problem. Then, when we have a mushroom and need to classify it, we check to see if the mushroom is

already in the database. If it is not, we feed the mushroom to a lab rat. If the lab rat dies, we can conclude that the mushroom is poisonous. If the lab rat survives, we can conclude that the mushroom is not poisonous.

Alternative C (Hill Climbing alternative):

Using the information from an online database on the experimentation of edible and toxic mushrooms that is updated every time an experiment is done, using the hill climbing algorithm, given the set of experiments on the experimentation of mushrooms, find a solution on which mushrooms are edible

Alternative D (Java alternative):

Obtain all the information on edible mushrooms from the FAO investigation document, updating it each time new experimentation data appears, classifying them using divide and conquer within an array, whether the mushroom is edible or toxic. Using JavaFX and displaying them in both table and graphs.

Alternative E (SQL alternative):

Using SQL (Server Query Language) as a language, read all the information from an online database about edible and toxic mushrooms, also update it every time an experiment is made to, classify them and connect them to a server to display it on a web page.

Alternative F (C# Decision Tree alternative):

Obtain the database from the University of California Irvine Machine Learning Repository, and then use C# and all available .NET libraries to read and save all the data and display all the graphs. Then create a decision tree algorithm that can determine if a mushroom is poisonous or edible given its attributes.

Alternative G (Human experiment alternative) :

Do an experiment in humans, making them consume the different types of edible and toxic mushrooms, to see their reaction and from there read all the information (updating it each time a new discovery arises in the experimentation) and manually enter the values one by one. Also, classify it given the criterion that the fungus is edible or toxic

Alternative H (Program from scratch alternative):

Obtain the database from Kaggle, then create a program from scratch in any programming language without using any external libraries to create a program that can read the data, display the graphs and use a machine learning algorithm to determine if a mushroom is poisonous or edible. This way, we can have more control over our program and will not have the risk of using external libraries with bugs.