Logistic Regression Rapport :

Written by :

-Abdellaoui Djouairia

-Meddah Baya

-El moukrfi Mounia

-Tamaoust Nermine

Written on 31/01/2024

Class L2 Group A

Table of Contents :

-introduction: - what is logistic regression

- **The problem**

**- Methodology: -tools**

**-steps**

**-Findings and Analysis :- About our project**

**- What the project contains ?**

**- For Raylib functions**

**-Discussion :- Challenges encountered**

**- Lessons learned**

**-Conclusion**

**Introduction :**

**What is Logistic Regression ?:**

**Logistic regression, also known as a logit model, is a statistical analysis method to predict a binary outcome, such as yes or no, based on prior observations of a data set.**

**A logistic regression model predicts a dependent data variable by analyzing the relationship between one or more existing independent variables. For example, logistic regression could be used to predict whether a political candidate will win or lose an election or whether a high school student will be admitted to a particular college. These binary outcomes enable straightforward decisions between two alternatives.**

**The problem :**

**We have a website that has multiple adds on it, the owner wants to see if the costumers will click on the adds or not based on the time they spent on it and their age. The goal is to predict whether a customer will click on the ad or not based on these two features(the time spent on the website,the costumer’s age).**

**Methodology:**

**Tools:**

**\*VSCode: Visual Studio Code, commonly referred to as VS Code, is an integrated development environment developed by Microsoft for Windows, Linux,macOC web browsers. Features include support for debugging, syntax highlighing, intelligent code completion, snippets, code refactoring, and embedded version control with Git. Users can change the theme, keyboard shortcuts, preferences, and install extensions that add functionality.**

**\*Raylib:** **is a cross-platform open-source software development library. The library was made to create graphical applications and games.**

**\*Git: is a distributed version control system that tracks versions of files. It is often used to control source code by programmers who are developing software collaboratively. Design goals of Git include speed, data integrity, and support for distributed, non-linear workflows — thousands of parallel branches running on different computers.**

**\*Github:** **GitHub is a proprietary developer platform that allows developers to create, store, manage, and share their code. It uses Git to provide distributed version control and GitHub itself provides access control, bug tracking, software feature requests, task management, continuous integration, and wikis for every project.**

**\*Notepad++: (sometimes npp or NPP), is a text and source code editor for use with Microsoft Windows. It supports tabbed editing, which allows working with multiple open files in one window. The program's name comes from the C postfix increment operator.**

**Steps:**

**1-We download all the tools**

**2-we create a folder in our desktop and we name it “project algo” then we click on it and with the right click on the mouse we open Git Bash and push the files into Github**

**4-for Raylib Open files then> "Disque Local" where the raylib folder is and then click on it > raylib > examples > core > click on core\_basic\_window.c**

**5-the file appears in Notepad++, then go to plugins > admin > download NppExec and NppExecScript**

**6-After that, you can modify the file "core\_basic\_window.c" and write the code you want there.**

**Findings and Analysis :**

**About our project :**

**We work on this project to learn how to make a machine learning functions on c and c++, learn how to read a file on c/c++, learn how to visualize our code with new tools such as Raylib/SDL, and how to use Git/Github (make commits on, make a copy of our files, learn how to push the files into Github…).**

**What the project contains ? :**

**The project contains files c/c++ of logistic regression functions,the code with ,csv that took a sample from it to calculate, a files that reads some lines of the csv on c/c++**

**Our goals is to solve any classification problem since logistic works with it,and visualize the solution to see where it classifies**

**For our code it contains 4 principale functions which are :**

**-linear regression: we use this function to help us later for the logistic regression,it works by taking two variables of each data of our sample table and multply the first value by x which is the number that we want to calculate with for example if we want to know if the costumer will click after 3 min the x then will equal to 3.**

**Logistic regression : or sigmoid function this is the most important one in the code cause it’s the one that classifie our data in range between 0 and 1, when the value it’s greater than 0.5 means it’s true and the opposite is correct**

**The odd function: this function predect how much is the ratio of the probability of the event occurring to the probability of it not occurring, it took the logistic regression(which is the probability value in this function) and calculate with it, it’s used to describe the relationship between a linear model and probabilities**

**The log odd : this one take the odd value to calculate, we use this function to check if the logistic regression that we calculate before is correct or not, but we can skip this function because it’s not that important just to make sure that our calculation is correct and all.**

**For Raylib functions :**

**screenWidth and screenHeight to declare the dimensions of our window screen and must be constant.**

SetConfigFlags **It must be called before initializing the window because it’s rendering the screen quality.**

**InitWindow to make the screen shown up with the dimension that we choose before.**

**Vector startpoint and endPoint/** **dynamicEndPoint to make the start and the end point of the white and the red line.**

**SetTargetFPS and speed so we can make the red line move from the left into right.**

**the loop while (!WindowShouldClose()) to make sure that since the window screen didn’t close it continue to draw what inside the loop.**

**DrawLineBezier to draw a line from the startpoint to the end point.**

**ClearBackground to choose the background color.**

**BeginDrawing from the name means the drawing start.**

**dynamicEndPoint.x it took the speed value to make the red line move in the next condition of it.**

**DrawText to draw text like the plot title and the value**

**DrawLine to draw x-axis and y-axis**

**DrawCircle to draw the value as a circle in the plot**

**EndDrawing to end the drawing of the loop**

**CloseWindow use it when the loop ends after the window closed by clicking on close**

**Visualization :**

**(check the repositery)**

**Discussion :**

**Challenges encountered :**

**-we fond a big challenge on how to install SDL and Raylib but SDL was pretty bad experience.**

**-the same with Raylib but we manage to fix it with notepad++.**

**-we fond another dificulty on understand how should the logistic regression visualization be but we fond the solution.**

**-we fond some dificulty on raylib functions but we manage to understand the important one from their official website.**

**Lessons learned :**

**How to make a functions that are mostly used on python or R for machine learning that mostly works by only calling it and give it the value.**

**How to read a file csv on C or C++.**

**How to manage to work with Git and Github by commits.**

**Conclusion :**

**We learn a lot on this journey of making a visualization of a problem from a code on c to a code with Raylib library (work from scratch).**

**In the end of this rapport i want to said that this is my first time writing it i hope that’s the way that the rapport must be and please don’t mind the autograph error and spelling error.**