DESIGN

By: Divya Shrivastava & Kirubel Tadesse

1.1 Summary of our Visual Analytics tool

Our Visual Analytics tool enables users to interactively change the web page parameters like speed, load time, etc to visualize the improvement of the performance of the websites. This would enable them to identify and analyze the user-friendliness of their website. Moreover, our visual analytics tool will give them a performance feedback scale on 7 parameters. The user will be able to see how each parameter affects the overall performance of their website. In addition, the tool also presents a predictive model in which a change in the 7 parameters affects their present performance.

1.2 Design Principle

The design process has two parts:

1. Development
2. User Interface design

1.2.1. Development

Development process involves the following steps:

1. Data Collection of top 100 websites from Alexa

2. Development of an automated script to run Webpagetest and design a model based on the performance of the top Alexa website.

3. The script will run a test on the top 100 alexa websites and save the result of the parameters such as on-load time, the first byte, start to render, speed index, document complete, and fully loaded for each of the 100 websites.

4. Once we have the data for each website from webpage test, we develop our standard Model against which we would be evaluating the results other websites.

5. The Model will be developed by calculating the minimum, mean and maximum value of each of the parameters from data we have collected.

6. For example – we compute minimum, maximum and mean on-load time for each of the 100 websites.

7. Figure 1 below shows our design sketch.

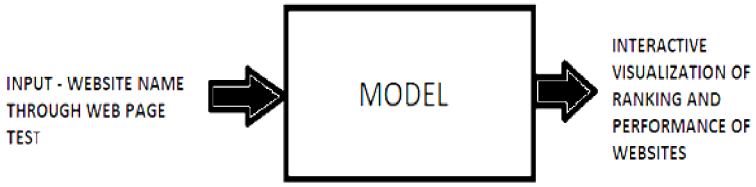
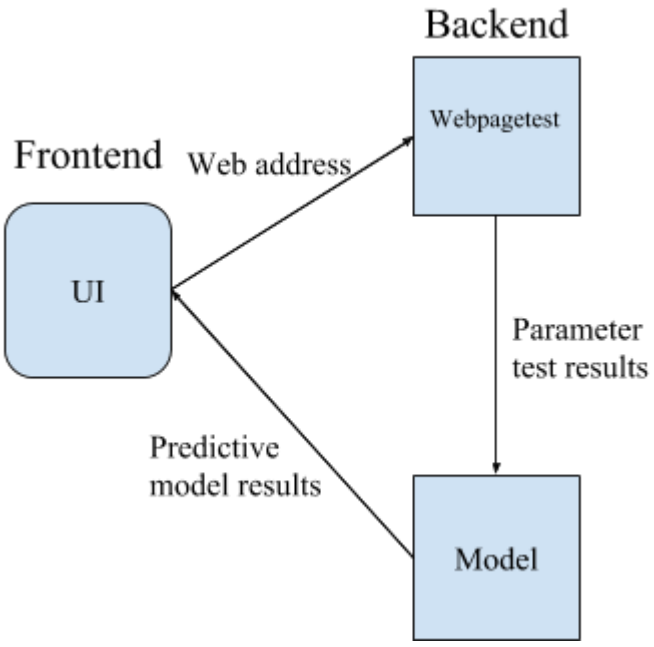


Figure 1



1.2.2. Visualization

1. After the development of our Standard Model, we design the User Interface.

2. UI page would be divided into 3 parts: Ranking, Performance and Console.

3. The ranking section would visualize the ranks of the websites through interactive charts.

4. Performance would visualize the performance of the websites against the 7 parameters.

5. The console would have interactive scales for each of the 7 parameters for the User to modify their values and visualize the performance.

6. The first step would be the UI design of the scale for each of the 7 parameters on the basis of the model developed having minimum, maximum and mean values.

7. The UI would be interactive, and the user would be able to change the scale of each of the parameters to improve the performance of the website.

8. We would be using Bokeh - a powerful plotting library in Python. It would help to create interactive plots to visualize the rankings and the performance of the websites.

9. The organization of the code for building the visualization tool would be the following:

- Create a parent directory to hold everything
- Within this directory, we have a subdirectory having our data
- Another subdirectory having our scripts
- Main.py file to put everything together and run the application

10. The major step in this process is to visualize how the ranking and performance of the websites get modified when the user changes the values of the 7 parameters below and above mean, or goes to min or goes to max.

11. This visualization would help the user analyze and gather result about the most critical and the least critical parameter which contributes in the performance of the website.

12. Figure 2 below shows the design of our Visual Analytics tool.

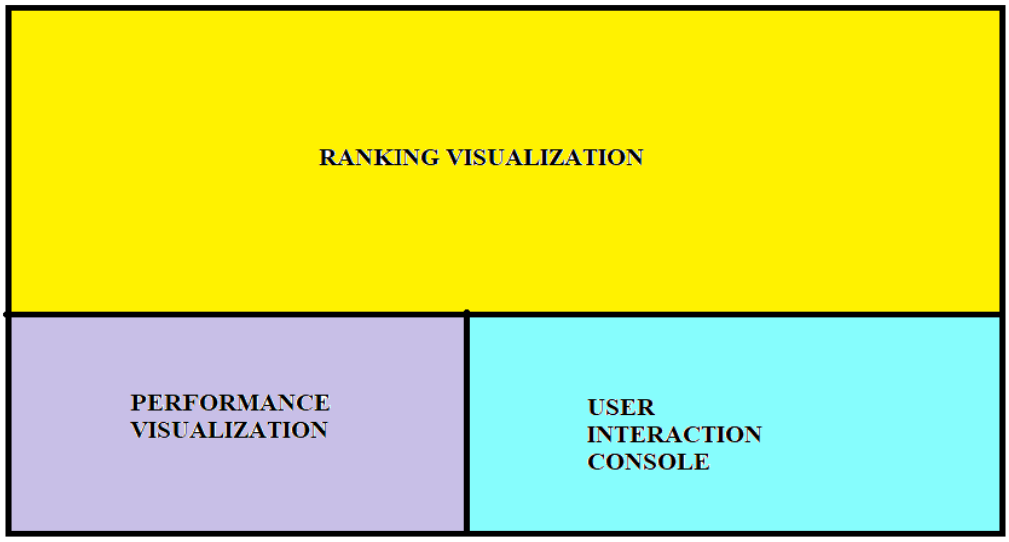


Figure 2.

1.3 Strengths of the design:

1. This is an interactive design helping the user to analyze the performance of the websites when they modify the value of the 7 parameters.

2. The design is well structured into development and UI phases and has clear sequential steps to be followed to achieve the final result.

3. The design can be extended to the future scope of adding machine learning techniques to create a more adaptive model.

1.4 Areas of improvement in the design

Currently, the model is developed using the mean, minimum and maximum values of the 7 parameters. More complex and sophisticated mathematical model could be developed and machine learning could be used to make the model more adaptive and sophisticated.

The current model we developed doesn't consider the network infrastructure and other parameters such as network congestion. Since the location from which the

network is being accessed also play a huge role in the overall performance of the network, it will be very important to take this factor into consideration.

The model is not adaptive to some of the change over time. Once, we develop the fixed values for each parameter and form our model. The model will not change even if there is a change in the parameter. We want the model to be more flexible and adaptive. Therefore, training the model using machine learning line is something to consider in the future.

1.5 Intended Audience:

1. Knowing the exact value of such parameters that have a direct influence on the user might be crucial for competitive business. Being able to know where to improve in relation to their competitive business, will help the business make an informed and tangible changes.

2. This visualization tool could be used by any user to evaluate the performance of the websites and to understand the criticality of each parameter in the performance of the website.

1.6 Earlier ideas which led to this vision

1. So far, web ranking has been done from search engine result using metadata which really cannot explain the user interaction with the website or the webpage user-friendliness.

2. Producing web rankings on the basis of the parameters like speed index, load time is a unique idea developed and will be very useful for competitive business.

1.7 Additional Design sketches

Please find attached additional design sketches for reference along with this document :

1. Development Phase - Design sketch
2. Bokeh Visualization Library - Code Design