

**PROJECT – DATA SCRAPING USING BLUE PRISM**

**MENTOR -**

**HEMANT KUMAR GAHLOT**

**BY-**

**ANILA VIPPARTHI**

**BTECH 3RD YEAR EEE**

**SRM INSTITUTE OF SCIENCE AND TECHNOLOGY**

**CHENNAI**

1. **INTRODUCTION-** 
   1. **OVERVIEW**

This data scraping project focuses on the concept of Robotic process automation (RPA) which is a software technology that makes it easy to build, deploy, and manage software robots that emulate human actions interacting with digital systems and software. The usage of software robots makes the process faster and more consistent than people, without human intervention. In this data scraping project our motive is to extract the data from the flipkart website by auto launching it. This project is expected to perform the following

* Auto Launch the website
* Extract the Data
* Store the Data Extracted

To automate in such a way that it reads the name, price and ratings of all the mobile models of the brand that we have selected. This project has been performed using the blue prism software with various tools that help ease the process.

* 1. **PURPOSE**

The purpose of this project is to automatically launch the shopping website according to the commands that we have given. Read the name, price and ratings of the mobile models that we have specified in the instructions.

1. **LITERATURE SURVEY**
   1. **EXISTING PROBLEM**

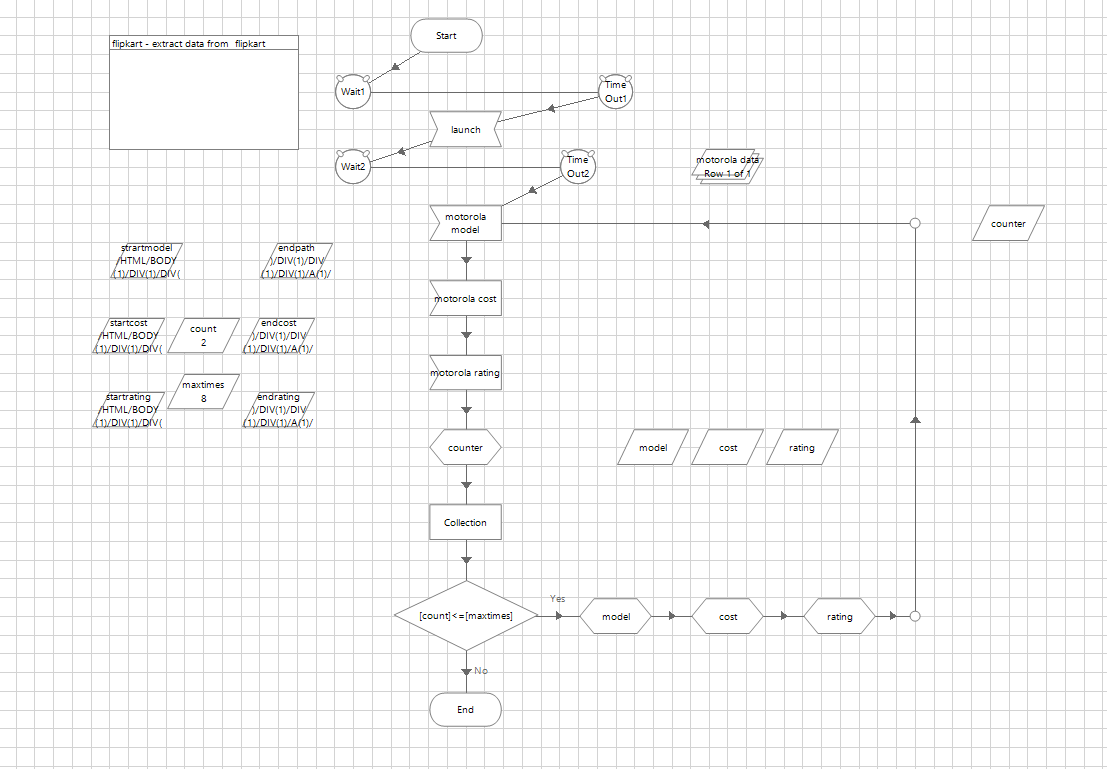
The traditional and manual method that exists is that we visit the website in the browser by searching the specific model of any product that we wish to purchase. It is quite a tedious process to manually search for a particular product and compare the price range and ratings of the same.

* 1. **PROPOSED SOLUTION**

This problem can be solved by using the blue prism software that uses the concept of robotic process automation to ease the process of looking for product in a browser and compare its price range and ratings etc.

This not only eases the process but also does not require human intervention once the instructions have been given in the software. It is performed at a much faster pace with much more accuracy than a human performing the task. It simultaneously compares various parameters and gives an output within a few seconds.

1. **THEORETICAL ANALYSIS**
   1. **BLOCK DIAGRAM**



* 1. **HARDWARE/ SOFTWARE DESIGNING**

The software that has been used to design the above flow diagram is the blue prism software.

Blue Prism is an RPA Tool which holds the capability of virtual workforce powered by software robots. This helps the enterprises to automate the business operations in an agile and cost-effective manner. The tool is based on [Java Programming Language](https://www.edureka.co/blog/java-tutorial/) and offers a visual designer with drag and drop functionalities.

In this section of the RPA Blue Prism blog, I will introduce you to the various components/essentials of Blue prism.

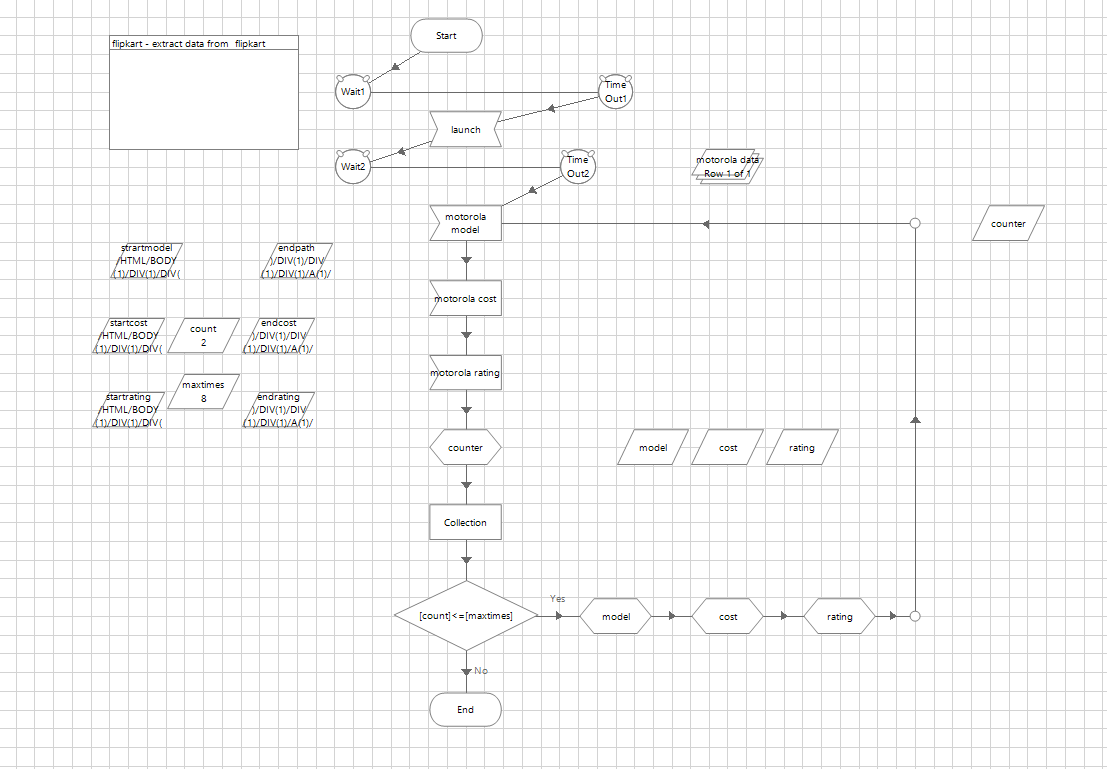
The four main components of Blue Prism are:

* + - Process Diagram
    - Process Studio
    - Object Studio
    - Application Modeller

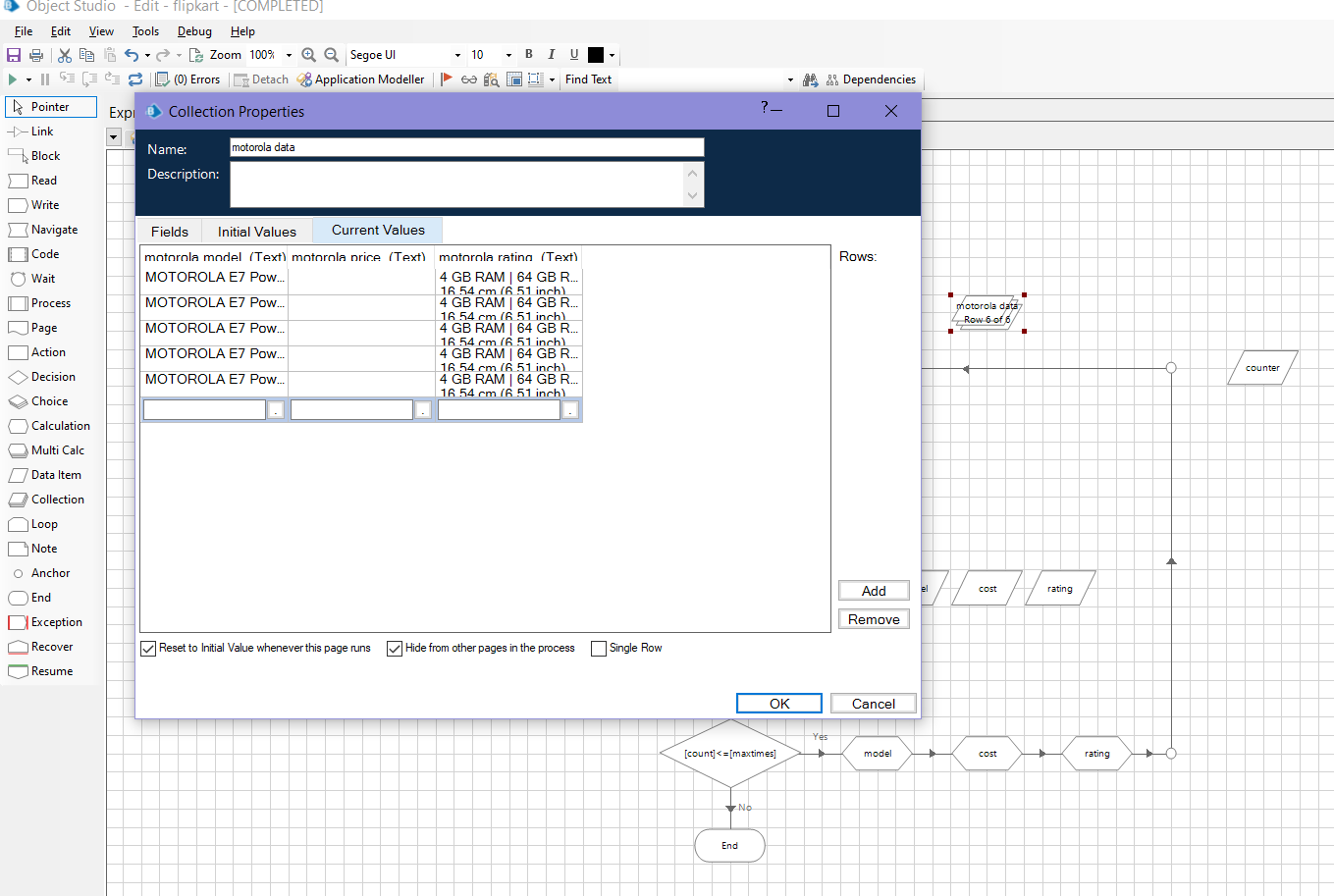
1. **EXPERIMENTAL INVESTIGATIONS**

I analysed about blue prism tools in the software and investigated the process of datascraping.

1. **FLOWCHART**



1. **RESULT**



1. **ADVANTAGES AND DISADVANTAGES**

**The major advantages of using the proposed solution which is robotic process automation are –**

* **Code free** - RPA doesn't require any coding or programming knowledge. The modern RPA tools are used to automate applications in any department where the clerical work is performed across an enterprise.

It gives an advantage over the traditional methods of automation and enables accelerated delivery of business applications. Besides, this platform reduces the initial cost of installation, training, and deployment.

* **Non disruptive** - the transformation process in RPA is very simple and straightforward. The RPA software robots follow the existing security, quality, and data integrity standards to access the end-user system in the same manner as human beings. These software robots also prevent disruption of any kind and maintain functionality and protections.
* **User friendly** - RPA does not require a special kind of knowledge, such as coding, programming, or deep IT skills. RPA software is user-friendly, easy to understand, and easy to use. RPA tools allow users to create bots quickly and effortlessly by capturing mouse clicks and keystrokes with a built-in screen recorder component. Some of the RPA software includes the option to create and edit bots manually using the Task Editor.
* **Rich analytical suite** - RPA software contains an in-built analytical suite that evaluates the performance of the robot workflows. RPA analytical suite also helps in monitoring or managing the automated functions from a central console, which can be accessed from anywhere. It offers basic metrics on robots, workflows, and more. The analysis performed by the analytical suite helps users to track the operations and determine issues. There is no need for any integration since everything is inbuilt and set right out of the box.
* **Security** - When an organization is running on automation, more users will demand access to RPA products. Therefore, it is important to have robust user access management features. RPA tools provide options to assign role-based security capabilities to ensure action specific permissions. Furthermore, the entire automated data, audits, and instructions which can be accessed by bots, are encrypted to avoid any malicious tampering. The enterprise RPA tools also offer detailed statistics of the logging of users, their actions, as well as each executed task. Thus, it ensures the internal security and maintains compliance with industry regulations.
* **Rule based exception handling** - RPA system allows users to deploy bots using rules-based exception handling. This feature proactively handles the exception. For example, RPA robot reports an exception, and then the actions given below are triggered:
* The same process is re-assigned to a different bot by the server.
* The current bot retries the same process and removes the previous bot from production.
* If retry is successful, the server maintains the reassignment and raises an alert to report exception & resolution.
* If retry is unsuccessful, it stops the current bot and raises an alert to report exception as well as failed resolution.
* **Hosting and deployment options** - The RPA system provides deployment options across virtual machines, terminal services, and cloud. Cloud deployment is one of the best among all the other deployment options, which attracts most of the users due to its scalability and flexibility. Therefore, businesses can install RPA tools on desktops and deploy it on servers to access data for completing repetitive tasks. RPA systems can automatically deploy robots in a group of hundreds. Similarly, multiple bots can be used to run different tasks within a single process while processing a high volume of data.
* **Actionable intelligence** - This RPA feature refers to the ability to gain and apply knowledge as skills. Robots first obtain the data and then convert it into information and transform the information into actionable intelligence for the users. Artificial intelligence and cognitive intelligence are the common features of RPA solutions that help bots to improve decision making over the period.
* **Debugging** - One of the biggest advantages of RPA from a development perspective is debugging. Some RPA tools need to be stopped running while making changes and replicating the process. The rest of the RPA tools allow dynamic interaction while debugging. It allows developers to test different scenarios by changing the values of the variable without starting or stopping the running process. This dynamic approach allows easy developments and resolution in a production environment without requiring changes to the process.

**The major disadvantages of using RPA are the following –**

* **Potential job losses** - If a robot can work faster with a more consistent rate, then it is assumed that there will be no need for human input. It is the main concern for the employees, and this results as a major threat to the labor market. However, this thinking is not accurate.
* **Initial investment costs** - RPA is still in the stage of innovation, and so it can present challenges that may result in unwanted outcomes. Therefore, it isn't easy for organizations to decide whether they should invest in robotic automation or wait until its expansion. A comprehensive business case must be developed when considering the implementation of this technology; otherwise, it will be useless if returns are only marginal, which may not worth taking the risk.
* **Hiring skilled staff** - Many organizations believe that to work with RPA, the staff must have significant technical knowledge of automation as robots may require programming skills and an awareness of how to operate them. It further forces organizations to either hire a skilled staff or train existing employees to expand their skills.

An automation company can be a little beneficial during initial installation and set-up. But the skilled staff can only adopt and manage the robots in the long-term.

* **Employee assistance** - People are usually habitual, and any change in the organization may cause stress to the employees. People who are involved in new technology will get new responsibilities, and they will have to learn new concepts of a specific technology. Because everyone may not have the same level of knowledge, it may lead existing employees to resign from their job.
* **Process selection** - It is always best to choose tasks that are repetitive, rules-based, and do not require human judgment. The non-standard processes are difficult to automate, and human interaction is required to complete such processes. So, there are limited tasks that you can automate with RPA.

1. **APPLICATIONS**

**RPA is used applied in the following domains -**

* Common business processes and activities
* Activities in commercial functions
* Sales
* Customer Relationship Management

**Activities in support functions**

* Tech Support
* Technology
* Finance
* HR
* Operations
* Procurement

**Industry-specific activities**

* Banking
* Insurance
* Telecom
* Retail

1. **CONCLUSION –**

Data entry seems simple enough- go to a website, copy paste the data into a spreadsheet and then crunch the numbers. But manually gathering and entering data can add up over time. And unfortunately, humans are prone to errors, leaving you with potentially unreliable information. Luckily there is a way to reclaim your time and eliminate errors – with robotic process automation from automate by using blue prism or ui path or automation anywhere tools.

1. **FUTURE SCOPE**

Robotic Process Automation is quite a new field which is getting popularized so much that the large organizations (mainly in US and Europe) have started investing in this technology and are really focusing on its career opportunities knowing it will do wonders in near future. This is a very fascinating field as something that we could have never imagined a decade back. It is now actually happening.

* It can do multitasking. Like a computer. Or, it can better manage the repeatable tasks.
* It can exponentially increase the accuracy and minimize the chances of error.
* Improves standardization.
* It results in reducing the time taken to finish off the tasks.

1. **BIBLIOGRAPHY**

**https://youtu.be/iVYtNdCYgFo**

1. **APPENDIX**