AGILE TEST AUTOMATION FOR WEB APPLICATION USING TESTNG FRAMEWORK WITH RANDOM INTEGRATION ALGORITHM IN MACHINE LEARNING TO PREDICT ACCURACY AND RESPONSE TIME ON AUTOMATED TEST RESULTS

V. VAMSI KRISHNA¹, G. GOPINATH²

¹Research Scholar, School of Computer Science, Engineering & Applications, Bharathidasan University, Tiruchirappalli, Tamil Nadu, India.
²Professor, School of Computer Science, Engineering & Applications, Bharathidasan University, Tiruchirappalli, Tamil Nadu, India.

Correspondence Email: vamsikrishnavit@gmail.com

ABSTRACT
"Testing" must be a part of any software engineering approach that intends to build high-quality apps. By executing it with input values, testing seeks to find faults in the tested-object and develop confidence in its proper functioning. Web apps take first place in development and testing, according to everyday usage. By automating the entire software development testing process, testing automation saves time and money for developers and testers. Our proposed solution would use the "TestNG framework," an automated testing framework, to test a public website and save the test results in the format of a ".csv" or ".xls" file to a given directory. The Support-Vector-Machine Algorithm (SVM), Random-Forest Algorithm & other machine-learning algorithms have been used to analyses the output file. The outcomes of all of the different ways will be compared and displayed on a graph. By Automating The Testing Framework Manual Testing Task Will Become Easy.

Keywords: Test Automation, Web Applications, TestNGFramework, SVM, Random Forest Algorithm, Random Integration Algorithm.

1. INTRODUCTION

The practice of building a programmed (test-script) in any programming/scripting language [1] that uses external automation helper-tools to duplicate manual-test-case methods. [2] stands for software-testing automation. It refers to the process of automating manual testing tasks. In order to test Software-Under-Test (SUT), software-test engineers should create & run a programmed [3]. To put it another way, toolkits are being created to test already-implemented source code. Its goal is to build test methods that can be automated [4]. Both writing test scripts and writing software are development activities; one is for the programmed itself, and the other is for the test scripts. Automating execution phase of software-testing cycle is quite familiar in automation business. “Software-Testing” is entirely automated by test script authoring is no longer a possibility; new toolkits are required to fully automate testing stages. [5].

Web-based apps are more dynamic and interesting than traditional programmed. As a result, traditional online app testing approaches and technologies are ineffective. A Web-application is an internet-based system that includes a database (or back-end) and interactive Webpages for users (or front-end)[6]. Web-applications are complicated software systems that are frequently updated and modified. Testing is both challenging and vital for them. It's tough because traditional testing methods and tools ignore the particular characteristics of web-based programmed, rendering them worthless. Testing web-based apps is dangerous because failure can be costly.
The agile approach is a software development project management style that involves self-organizing, cross-functional teams and their clients collaborating to create requests and solutions. Agile is a collection of ideas that emphasize adaptability and flexibility. In order to respond quickly to changing business needs, Agile places a premium on teams' ability to deliver in small, digestible chunks.

Based on the TestNG framework, this paper proposes a Web-application automation-testing approach. We'd put in place a system for automated Web-application testing. The format of '.csv' or '.xls' files was used in this proposed approach to automatically test the automation-testing framework. Each page encountered is analyzed using machine-learning algorithms to extract the necessary information. The data gathered will now be used to investigate the errors discovered while testing the web application.

2. RELATED WORK

Those who have contributed to this study include Z. Sun et al. [7] proposed a web-application testing platform based on hybrid automation software-testing framework that merge technical advantages of data & keyword driven software-test frameworks is proposed to address the issue of testing data differentiation & test case explosion in software-testing procedure of web-application system, as well as to lower the cost of automation testing maintenance.

Y. Kravchenko et al. [8] offer a dynamic Web-app verification technique primarily based totally at the shortest time criterion, in addition to a device evaluate. Dynamic verification techniques depend on the effects of real-global paintings as verified with the aid of using the software program machine or its prototypes to make sure that the effects are in compliance with the necessities and layout decisions. The consciousness of the evaluate could be on dynamic-verification-checking out and -tracking methodologies. The software program-verification aids jUnit (checking out) and AppDynamics (tracking) are assessed.

Reinforcement-Learning-Driven and Adaptive Testing (RAT) is defined by M. Amouei et al. [9] As an automatic black field checking out approach for coming across injection vulnerabilities in WAFs. SQL injection and cross-web website online scripting are of unique hobby to us due to the fact they’ve continuously been most of the pinnacle twenty vulnerabilities over the past decade. RATs, in unique, gather assault samples. It then employs a reinforcement mastering approach along with a ground-breaking adaptive seek set of rules to speedy locate definitely all circumventing attack patterns.

A. Sedaghatbaf et al. [10] describe ACTA, an automatic check introduction technique for black-container overall performance testing. ACTA does now no longer require a widespread amount of historic check information to examine the overall performance traits of the gadget below check due to the fact it's far primarily based totally on lively learning. Instead, uncertainty sampling is used to perceive which assessments ought to be finished at the fly. An ACTA permits for the specification of overall performance standards in phrases of circumstances, in addition to the improvement of assessments to fulfil the ones conditions. It is primarily based totally on an unsure sort of generative antagonistic networks.

S. Karlsson et al. [11] provided a technique for mechanically growing GraphQL queries inorder to check GraphQL API’s. This has been done with the aid of employing a property-primarily based totally on totally approach to create question generator which is primarily based totally on total GraphQL illustration gadget beneath neath test. We are checking out a real-international software program gadget which suggests that this output is powerful & green at detecting actual flaws, overlaying a whole schema in seconds.

For evaluating app UIs, S. Yu et al. [12] advocate a widget feature matching & layout characterization matching-primarily based totally image-pushed mobile-app checking out device. We utilize pc vision (CV) technology which is used to carry out UI function contrast and format hierarchy extraction on mobile-application screenshots to reap UI systems that include wealthy contextual facts approximately app widgets which includes positions, relative relationship, and so on. Based at the UI systems we’ve uncovered, we are able to create a platform-unbiased take a look at script and function the goal widgets beneath neath take a look at. As a result, the proposed device employs a singular platform-unbiased take a look at-script encoding to non-intrusively replay take a look at scripts.

To achieve adaptive exploration of web-applications, Y. Zheng et al [13] gift Web
Explore, an automated give-up-to-give up web-checking out framework. Web Explore makes use of curiosity-pushed reinforcement mastering to construct extremely good motion sequences (check cases) with sequential logical relationships. In addition, for the duration of the online-checking out process, Web Explore generates an automaton that serves as high-stage steerage to sell and enhance checking out efficiency. Six real-global projects, a industrial SaaS web-app, and a survey of the global’s pinnacle 500 web-apps had been all evaluated in depth.

A. Martin-Lopez et al. shows how specify & automate dependencies of inter parameter in web-application API’s in their study [14]. We’ll start with Inter-parameter Dependency Language (IDL), which is a domain specific language specifying relationship between input parameters in web-application services. To enable automatic IDL specification analysis, we offer a mapping that converts IDL documentation into Constraint-Satisfaction-Problem (CSP).

We had proposed collection of 9 IDL document investigation methods that can be used to see if a given request, for example, meets all of the service’s dependencies. Finally, for IDL specifications and analyses, we propose editor, parser, OAS extension, and constraint programming library.

T. Rangnau et al. [15] showsus to integrate 3 distinct automated dynamic software-testing approaches into CI/CD pipeline while still allowing for empirical measurement of the overhead introduced. Then we categorise and suggest preliminary answers to the DevSecOps community's various research/technology challenges. Our findings will aid us in making informed judgments as we deploy DevSecOps approaches in agile enterprise application engineering and corporate security.

H. Zhu et al. [16] provide a fascinating experiment for evaluating automation of oracle software. In common, software-testing is a process of inductive inference which tester tries to infer general qualities of a software system from its behavior on small number of test-cases. This debate looks at the theoretical foundations of software testing via the prism of computational machine-learning theories.

DARIO is a test oracle proposed by A. Arrieta et al. [17] that uses regression-learning methods to forecast the system's Quality-of-Service. The regression-learning algorithm in Oracle is trained on test data from earlier tested versions. Empirical estimation based on industry case study demonstrates that applicability of our technique. The regression tree method outperformed the other four regression learning algorithms that were tested. The regression tree technique's accuracy in predicting DARIO verdicts ranged from 79 to 87 percent.

Based on current approaches such as the keyword-driven and page-object models, B. Swathi et al. [18] present a methodology for constructing a test automation framework. The issue of test case generation and optimization has been addressed. Various encoding approaches regard test-case representation as a difficulty in soft computing techniques. It covers test case representation as well as test case generation using soft computing algorithms.

J. Kahles et al. [19] show how machine learning can be used in agile software testing to automate root cause investigation. After interviewing testing engineers, [20] It methodically extract essential information from raw log data (human experts). Despite finding only minimal correlations among specific clusters & failure underlying causes, remaining clusters' uncertainty necessitates labelling. Following new set of interviews with software-testing engineers, five ground-truth categories have been developed. It trained artificial neural networks [21] to undertake data categorization or pre-processing via clustering using manually labelled data [22].

3. EXISTING SYSTEM

"Manual-Testing" is a software testing framework in which a tester conducts all test cases by hand rather than utilizing automated technologies [23]. An approach for detecting bugs, flaws, and weaknesses in software is manual testing. Manual software-testing is most basic of all testing methods& it aids in detection of [24] major faults in software-based applications.

Many small businesses continue to rely on manual testing in the belief that it will provide them with greater accuracy [25]. Because they require a larger number of efficient resources to work on the automated testing process of their own software products, but they never anticipated it being a one-time event. We intend to address the existing issues in our proposed approach.
4. PROPOSED SYSTEM

The term "Automation Testing" refers to a software-testing technique that compares actual result to expected result. This is accomplished by using test scripts/any automated testing tool. "Automation-Testing" utilized to automate manual testing tasks that are time-consuming and difficult to complete. Here proposed work will test a public website using the "TestNG framework," an automated testing framework, and save the test results to a specified path in the form of a .csv or .xls file. Mixtures of machine-learning algorithms were utilized to analyse the output file, including the Support-Vector-Machine (SVM) Algorithm, Random Forest Algorithm, and Random Integration Algorithm. All of the algorithms will be compared, and the results will be plotted in a graph.

Pseudo code of SVM Algorithm

```
Commence
Initialised SVM parameter & structure
Produce an initial number of birthing lairs Bi = (q = 1, 2, 3, 4…n)
While (Stopping condition)
  If noise = false
    Look for proximity for a new lair by means of a Brownian walk
  Else
    Apply exploration for a way for a new layer by means of levy walk
  End if
  Calculate fitness of each new lair and evaluate with previous
  If
    B best, t > B best, p+1
    Select new lair
    B best = B best, p
  Else
    Go to 4
  End if
Rank solutions;
Return most excellent lair
The overall greatest lair is fed to SVM classifier for training
Training SVM classifier
End while
End
```

The distance of any line, ax + by + c = 0 from a given point say, (x0, y0) is given by d. In the same way, the distance of a hyper-plane equation: wTΦ(x) + b = 0 from a given point vector Φ(x0) can be simply written as:

\[ d_H(\phi(x_0)) = \frac{|w^T\phi(x_0) + b|}{||w||_2} \]

Here ||w||_2 is the Euclidean-norm for the length of w given by:

\[ ||w||_2 = \sqrt{w_1^2 + w_2^2 + w_3^2 + \ldots w_n^2} \]
Random Forest Algorithm

Input: \( X = \text{Train data}, R = \text{Overall Features}, r = \text{members of features}, T = \text{Amount of tress} \)

Output: Input data bagged class tag

Commence
1. While a particular tree in Forest T:
   i. Select a bootstrap sample M with size N out of training data.
   ii. Build tree \( B_t \) through iteratively repeating the different steps regarding tree individual node.
      a. Choose randomly \( r \) from \( R \)
      b. Select optimal between f.
      c. Split node.
2. When the construction of \( T \) trees is executed, examples of Test data will be sent to individual tree & the allocation of class tag on the basis of majority votes will be carried out

End

For each and every decision tree, Random Forest computes nodes significance with Gini Importance, pretentious only 2 child nodes (binary tree):

\[
n_{lj} = w_j C_j - w_{left(j)} C_{left(j)} - w_{right(j)} C_{right(j)}
\]

Significance of each and every feature on decision tree is then deliberated as:

\[
f_{lj} = \frac{\sum_{j: \text{node } j \text{ splits on feature } i} n_{lj}}{\sum_{k: \text{all nodes}} n_{lk}}
\]

These can then be normalized to a value among ‘0’ & ‘1’ by dividing the sum of all feature significance values:

\[
\text{norm}_{f_{lj}} = \frac{f_{lj}}{\sum_{j: \text{feature } i} f_{lj}}
\]

Ultimate feature significance, at Random-Forest level, is its standard over all the trees. The sum of feature’s importance value on each & every tree’s is considered & divided by total number of trees:

\[
\text{RFF}_{f_{lj}} = \frac{\sum_{j: \text{all trees}} \text{norm}_{f_{lj}}}{T}
\]

Random Integration Algorithm

Input: \( X = \text{Training data}, R = \text{Total Features}, r = \text{members of features}, T = \text{Amount of tress} \)

Output: Input data bagged class tag

Start
1. While a distinct tree in Forest T:
   i. Perform Tabu search algorithm to choose a best sample \( M \) with size N out of training data.
   ii. Build the tree \( B_t \) through iteratively repeating the dissimilar steps regarding tree individual node.
      a. Choose randomly \( r \) from \( R \)
      b. Select optimal between f.
      c. Split node.
2. When the creation of \( T \) trees is carried out, examples of test data will be sent to individual tree & allocation of class tag on basis of majority votes will be carried out

End

During the generation of neighbor solutions, dissimilarity among current & fresh neighbor-solution managed by means of a coefficient, \( \alpha \). The change from current point is multiplied by \( \alpha \) during building new neighbors. Coefficient \( \alpha \) is in the form of a sine function

\[
\alpha = 121 + \sin(i\pi N_{\text{neigh}})
\]

Here \( i \) am the index of neighbor, \( N_{\text{neigh}} \) is entire number of neighbor solutions which is being produced at each iteration, & \( \theta \) is a parameter that controls the oscillation period of \( \alpha \).

An objective principle that intended based on sigmoid function given by

\[
S_k = 11 + e^{-\sigma_k - k_{\text{center}}} \times M
\]

\[
f_k = f_k - \Gamma f_k - \Gamma x < \delta
\]

Where \( \delta \) is the ratio of alteration in objective-function value, \( \Gamma = \eta M \), and \( \eta \) is the fraction of the maximum iterations (M) by which the modify in the objective-function is evaluated. As per this stopping measure, if the enhancement over \( \Gamma \) generations is no longer than a threshold
(δ), continuation of further iterations can be ineffective & search should be terminated.

6. EXPERIMENTAL RESULTS

As previously stated, our ambition is to use machine-learning algorithm & TestNG framework to automatically test a Web-based application. We tested several powerful machine-learning models, including Support-Vector-Machine (SVM), Random-Forest & Random Integration, to find the best Machine-Learning algorithm capable of testing more precisely. To aid in successful evaluation of these replica’s, we utilized machine-learning libraries.

TEST CASES

Positive Results
Load the website to be tested into the selenium web driver using the base URL, then write selenium commands to access specific features such as buttons, hyperlinks, test boxes, and so on to execute test cases. Positive outcomes will appear as binaries “1,” as shown in the screenshots below.

Fig-3 Test-Cases For Positive Results

Negative Results
For negative cases the results will be in the form of binaries “0”, kindly look at the below furnished screenshots for more details.

Fig-5 Test-Cases For Negative Results

Fig-6 Test-Cases For Selenium

Table 1 Accuracy Comparison Of SVM, Random Forest And Random Integration Algorithm.

<table>
<thead>
<tr>
<th>ALGORITHM</th>
<th>50 Iteration’s</th>
<th>100 Iteration’s</th>
</tr>
</thead>
<tbody>
<tr>
<td>RANDOM INTEGRATION</td>
<td>80</td>
<td>160</td>
</tr>
<tr>
<td>RANDOM FOREST</td>
<td>78</td>
<td>156</td>
</tr>
<tr>
<td>SVM</td>
<td>75</td>
<td>150</td>
</tr>
</tbody>
</table>
Table 2 Response Time Comparison Of SVM, Random Forest And Random Integration Algorithm.

<table>
<thead>
<tr>
<th>ALGORITHM</th>
<th>50 Iteration’s</th>
<th>100 Iteration’s</th>
</tr>
</thead>
<tbody>
<tr>
<td>SVM</td>
<td>7.2 milliseconds (ms)</td>
<td>17.8 milliseconds (ms)</td>
</tr>
<tr>
<td>RANDOM FOREST</td>
<td>7.0 milliseconds (ms)</td>
<td>15.2 milliseconds (ms)</td>
</tr>
<tr>
<td>RANDOM INTEGRATION</td>
<td>6.5 milliseconds (ms)</td>
<td>13.5 milliseconds (ms)</td>
</tr>
</tbody>
</table>

Table-2 displays the time spent testing an identical test case using the three testing strategies, while Table-1 displays the correctness of the three methods. Only the time spent developing testing scripts in preparation for the execution phase is included in these numbers. The execution time is indicated in milliseconds (ms). Table-2 shows that Random Integration Algorithm requires the least amount of time to run when compared to Support Vector Machine and Random Forest Algorithm. The Random Integration Algorithm takes 6.5 milliseconds in total to apply (ms). The Support-Vector-Machine (SVM) and Random Forest algorithms, respectively, take 7.2 milliseconds (ms) and 7.0 milliseconds (ms).

On the other hand, the accuracy using Random Integration algorithm is 80 percentages at the 50th iteration whereas the accuracy of Random-Forest & Support-Vector-Machine (SVM) algorithm is 78% & 75%. Thus, by using Random Integration algorithm can increase accuracy about approximately 5% more comparing with Random-Forest & Support-Vector-Machine (SVM) algorithm. Tabulated values accuracy and time-consumption is graphically represented in Fig 7 and Fig 8 which gives the clear understanding of all the three algorithms in performance wise.

7. DIFFERENCE FROM PRIOR WORK:

By Automating The Testing Framework Manual Testing Task Will Become Easy. IN Previous work there is no such arrangement.

8. CONCLUSION

The major goal of study is to develop a new software-based automation software-test framework to aid testers in automation software testing process. Projected framework can be utilized successfully to automate web application testing. When the Software-Under-Test (SUT) rapidly changing & regression testing is required to ensure that software version is stable, the TestNG framework is advantageous. Furthermore, software-testers or business end-users capable to do whatever they want while test scripts are running on SUT. Experiments undertaken to assess projected framework's efficiency by applying it to a case study. These tests have yielded promising results, implying that the full software-testing automation process could be automated to achieve 80% accuracy in less time. It can accomplish so much work at
once that manual testing is almost impossible to complete. Furthermore, the suggested solution generates test-cases in an ordinary tabular format, which is both easier & tighter than writing them manually.

REFERENCES


