ISSN XXXX XXXX © 2017 IJESC

Research Article Volume 7 Issue No.7

Evaluation of Web Metrics

Pooja Mittal,
M.Tech Student
Department of Computer Science & Engineering
Dr. A.P.J. Abdul Kalam Technical University, Uttar Pradesh, India

Abstract:
Websites have become an integral part of over day to day life. They act as a source of information, a place to hang out with friends and also as a medium of communication. With such important characteristic development of web sites should be done carefully. The quality of web sites is typically concerned with performance and usability and is measured using web metrics. Web metrics are the measure of attributes of a web page. Collecting, analyzing and interpreting web metrics is referred as web analytics. Quality assessments of websites made by experts are not cost-effective and cannot be fair and also regular quality assessment of websites have to carried out which cannot be possible manually. Therefore we have made an attempt to automate the website quality evaluation process. In this study we categorized websites into three categories which are collected from pixel awards website and then analyze these categories using web metrics. For analysis we have created a web scrapper tool which evaluates web sites using web page metrics and applied different machine learning algorithm on these web page metrics to predict goodness of websites and also classified them within a particular category. The result of this paper provides an empirical foundation for web site designing.

I. INTRODUCTION

Improving the effectiveness and quality of websites is a question of significant importance for the successes of e-commerce, e-health, entertainment and other aspects of life that use the Internet technology as an effective channel to reach the target audience. Various website evaluation tools are now available to assess the effectiveness and quality of the websites associated to these industries. Normally most of these tools are designed for use by website users to provide their subjective opinions on the value and usefulness of information for which they are targeted. Subsequently, user's feedback and comments form a basis of website improvement where changes are deemed necessary. Specifically, there exists no established mechanism or relevant studies on assessing the quality and the effectiveness of online resources. However, few websites like webby awards, pixel awards etc. have made an attempt to provide a good rated websites which can be the basis of evaluating a website from quality perceptive. These websites provide the list of top rated websites of different categories like Entertainment, music, shopping websites etc. which are close to perfect in terms of their quality and popularity. Quality of any web site is represented in terms of how easily an individual navigate through web site i.e. user –friendliness, amount of information present on one page, visibility, web traffic. Website popularity is defined in terms of the number of hits on a web page generated by a website visitor. Web designing is the field of designing the web site. Web designing has been moved from the modest beginning of a text page to the high end of animated designs on the web pages. When developing website assessment tools, the technical and content aspects of the website should be taken into consideration. While content aspect refers to those features related to the content of a website (e.g. accuracy, objectivity and relevancy), the technical aspect refers to those features related to the design and usability of a website (e.g. navigation, interactivity and accessibility). The two aspects make up what we call website quality dimensions or simply website metrics. In this study, we define website metrics as sets of indicators to take into account when judging the perceived quality of the website material.

1. AUTOMATIC EVALUATION TOOLS

There are many criteria to evaluate a web site. Those may include: usability, authority, currency, objectivity, coverage, performance, traffic ranking, link popularity, accessibility, security, design patterns, HTML syntax analysis, and browser compatibility. Output data of traffic based and time-based analysis must be interpreted in order to identify usability problems. Server logs are problematic because they only track unique navigational events (e.g. do not capture use of back button). In HTML syntax analysis it inspect the static HTML for pre-determined guidelines, such as number of words in link, links which are image links, all images contain an ALT attribute[5].These guidelines may cover universally accepted guidelines or guidelines accepted in a specific society. A list of Automatic Evaluation tools which depends on the characteristics of HTML.

1.1 Web XM

WebXM is used to automate inspection of some page defects. These defects include broken links, spelling errors, slow loading pages, poor search and navigation to help improve usability of the web site. WebXM automates more than170 accessibility checks, namely appropriate text and background color contrast or the presence of text equivalent alt tags on images.

1.2 Booby

Bobby is a web accessibility testing tool. It is designed to remove barriers on accessibility issues. It also encourages compliance with existing accessibility guidelines, including Section 508 of the US Rehabilitation Act and the W3C's Web Content Accessibility Guidelines (WCAG) [28]. Bobby examines every page of a website and tests every page of web site individually. Then it checks the web site for several accessibility requirements.

2. NIST WEB METRICS

The US National Institute of Standards and Technology (NIST) have developed prototype tools. These tools aim to evaluate web site usability [19]. There tools are WebSAT, WebCAT.

2.1 WebSAT

The Web Static Analyzer Tool (WebSAT) is a prototype tool that inspects the HTML code of web pages for usability problems. WebSAT allows the webmaster to investigate these problems. Then webmaster can remove these problems from the web page design. WebSAT not only applies its own set of usability rules but also applies the IEEE Std.2000-1999 (NIST 2001b). [18] Likewise Bobby, accessibility is measured in accordance with the three priority levels suggested by WAI recommendations [28].

2.2 WebCAT

The Web Category Analysis Tool (WebCAT) allows webmaster to conduct a simple category analysis in the web quickly [12]. This is based on traditional card sorting techniques. The webmaster creates a set of categories and a number of items which are to be assigned by test subjects to the categories. Then the Webmaster can compare the real assignments with intended assignments which will meet user needs.

3. LIMITATIONS OF EXISTING WEB EVALUATION TOOLS

A number of existing website evaluation methods generally requires the evaluator who has IT background to assess the qualities in a website. It is difficult to apply if the people do not have any IT skills. Many new website software technologies and rules are not considered in existing website quality evaluation methods. The web developer is confused by the overall picture of the evaluation criteria. A new website evaluation methods need to involve the all identified new software technologies as the numbers of new criteria. The specific quality criteria for a website's reputation are clarified in many existing website evaluation methods, however most creditable criteria are immeasurable. The strengths and weaknesses of the web evaluation results should be applied to the user's expectations, and ease of understanding.

3.2 Web Page Metrics

Metrics, as we know, refer to standards of measurement. Therefore, web metrics are standardized ways of measuring something that relates to the Web. Web metrics helps organizations to understand, manage and improve their web systems and hence enhance the quality of their online presence. Depending on the ISO 9126 quality model different metrics are

3.2.1 Efficiency web metrics [20] [9]

Efficiency metrics include related to size of a web page and the load time of a website/webpage.

Table.1. Efficiency web metrics list

Metric	Meaning
efficiency_css_size	Css size per page
efficiency_homepage_load_time	Homepage load time
Efficiency_image_size	Image size
efficiency_javascript	Script size per page
efficiency_page_load_time	page load time
efficiency_page_size	Page size

3.2.2 Functionality web metrics

It includes navigation, forms, identity and other aspects related to the functionality offered by the site.

Table .2. Functionality web metrics list

Metric	Meaning
forms_form_info_request [10],[23]	presence of contacts/info form
forms_labels[23]	number of label tags
Identity_auther[9]	Average presence of author
Identity_logo[9]	presence of site name in title
Identity_sitename_title[23]	Presence of navigation bar
Navigation_bar[20]	Presence of navigation bar
Navigation_bread_crums[20]	Presence of bread_crums

3.2.3 Maintainability web metrics

Includes aspects related to the number of items to maintain (e.g. scripts, styles used, tables).

Table.3. Maintainability web metrics list

Metric	Meaning
Maintenance_num_script[23]	Script files no per page
Maintenance_num_styles[23]	Css file number per page
Maintenance_num_tables[1]	Tables number per page

3.2.4 Portability web metrics

Includes aspects related to page layout, use of html standards, etc.

3.2.5 Reliability web metrics

Includes aspects related to the validation and links status

3.2.6 Usability web metrics

Includes aspects related to accessibility, multimedia and textual contents.

4. INDEPENDENT AND DEPENDENT VARIABLES

The dataset comprises of 22 measures to be used for web pages, one dependent and twenty one independent variables. These variables cover those attributes that can be computed automatically. We developed a Web Scrapper tool developed in

PYTHON technology to compute these metrics. We have used attribute selection technique for reducing data dimensionality provided in WEKA tool [32].

5. WEB METRICS SELECTED FOR STUDY

1. Meta tag
Total number of Meta tag on a page. This attribute is calculated by counting total number of Meta tag on a page. The text in meta tags is not displayed by browser.

2. Meta Keywords
Total number of meta keywords on a page. This attribute is calculated by counting total number of words in meta tag where name attribute is keywords. Meta Keywords are separated with comma (,) therefore words between two commas is considered as one keyword.

3. Minimum Meta Keyword length
Attribute is calculated by identifying a meta keyword from Meta Keywords which contain minimum number of characters. No spaces, commas, new line and tab are considered while counting characters in a keyword.

4.Maximum Meta Keyword length
Attribute is calculated by identifying a meta keyword from Meta Keywords which contain maximum number of characters No spaces, commas, new line and tab are considered while counting characters in a keyword.

5. Meta Descriptor words
Total number of words in meta tag where name attribute is descriptors. No spaces, commas,|,",\n,\t are considered while counting descriptor words.

6.Total links
Total number of links in a web page. The links which are in comments are not counted.

7. Image links
Total number of links which are images in a web page i.e. clicking on image leads to a new page.

8.Average number of words in a link
Attribute is calculated by dividing sum of words in all text links by total text links.

9.Total images
Attribute is calculated by counting number of images in a page.

10. ALT Images
Total number of images which contain an alt attribute. When a particular image is not loaded by a browser then the text present in alt attribute is displayed in place of the image.

11.Words in alt images
Attribute is calculated by summing all words present in all alt images.

12.Division tag
Total number of divisions of a web page i.e. in how many section a web page is divided.

13. Paragraph This attribute is calculated by counting total paragraph in a web page.

14. Scripts This attribute is calculated by counting total number of java scripts used in a web page.

15. Size
Attribute refers to number of bytes requires to store a web page on a system. We assume that only one server request is send at one time.

16. Body Word Count
this attribute is calculated by counting total words present in a body tag of a web page. We discard all words which are in comments and script tags.

17. Title Length
Total number of words present in web page title. Special characters are also considered.

18. Tables Total number of tables in a web page. Tables are also used sometime for division of a web page.

19. Load Time
Attribute refers to time taken by a web page to load i.e. difference between the first request and first response time.

20. Total Headings
Total number of headings in a web page. We consider all six types of headings. Headings font size is bigger than rest of the word i.e. they are more visible than other word.

21.Link Headings
Total number of headings which are link i.e. on clicking that heading we move to a new page.

6. EMPIRICAL DATA COLLECTION

We analyze the web pages collected from pixel awards website. The pixel awards web site was established by Erick & Laubach in 2006.The Pixel Awards judges are proven innovators in their fields with broad web expertise and a knack for spotting extraordinary talent with fairness and accuracy [22]. Each website is evaluated on the basis of innovation, content, navigation, visual design, functionality and overall site experience. The Pixel Awards judges are proven innovators in their respective fields with broad web expertise and a knack for spotting extraordinary talent with fairness and accuracy as described in Pixel Awards [22]. The websites placed in 24 categories are judged on the basis of creative and technical blend of impeccable graphic design, artistry, technological expertise, and a powerful, stimulating user experience [22]. These sites are the best of the web, thus each site for its respective category is evaluated for innovation, content, navigation, visual design, functionality and overall site experience. For over study, we collected data from 7 categories of pixel awards for each year from 2006 to 2012. The categories we selected are TV, Movies, Blogs, Community, Food & Beverage, Travel and Commerce. We have merged the categories which have some properties in common and created three models.

Model 1: In this model we have collected data for Blogs and Community website. Blogs also sometime referred as community and people often build a community over a blog. Data set is created from 119 web pages and some level-1 pages are also included.

Model 2: In this model we have collected data for TV and Movies website as both have similar structure and purpose. Data set is created from 51 web pages

Model 3: In this model we have collected data for Food & Beverage, Travel and Commerce websites. All three behave as e-commerce websites. Data set is created from 129 web pages and some level-1 pages are also included. The data points used to predict good- bad class for each model is tabulated in table III.

Table .4. Data points for good or bad classification

Model	Good data points	Bad data points	Total Points
Model 1	35	84	119
Model 2	12	39	51
Model 3	28	101	129

The data points used to predict the class of each website within the particular model is tabulated in table IV.

Table.5. Data points for class predication

Model	Websites	Good data points	Total points
Model 1	Blog	62	119
	Community	57	
Model 2	TV	25	51
	Movies	26	
Model 3	Food & Beverages	39	129
	Travel	41	
	Commerce	49	

7. RESEARCH METHODOLOGY

7.1 Methodology

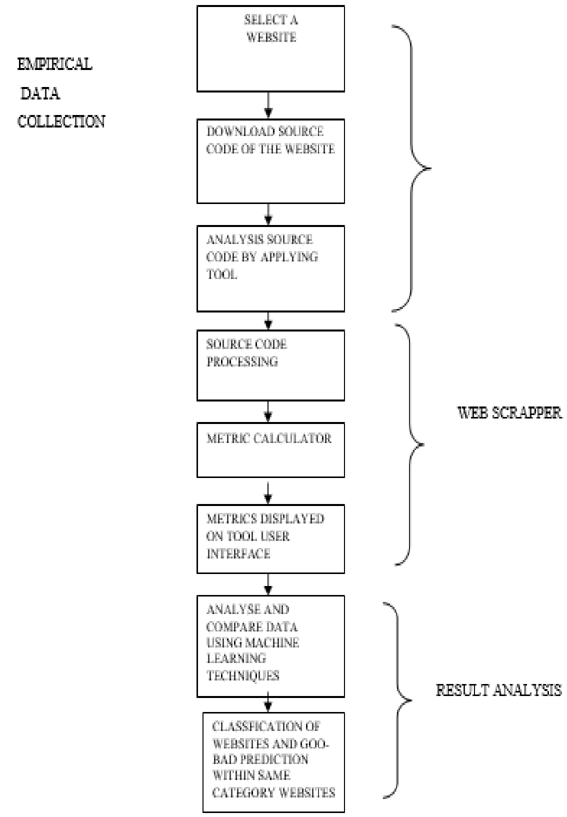
We employ quantitative web-page attributes in our methodology to classify websites belongs to same domain (TV, movies website belong to entertainment domain) and in order to these we construct a model. Figure 4.1 shows the flowchart of methodology.

The flowchart of methodology is divided into three sections.

a. Empirical Data Collection: - In this section we first select the website from different nominated categories in 6 years (2006-2012) from Pixel awards for which metrics estimations are to be calculated. Secondly we download the source code of the website and then apply a web scrapper to calculate the different metrics for these websites.

b. Web Scrapper: - Web scrapper will automate the process of web metric extraction from a web page. Firstly it will preprocess the source files to remove all unwanted things and then metrics are calculated.

c. Result Analysis: - In this section we use different machine learning techniques to analyze and compare data. Comparison shows which algorithm gives better results compared to others.



7.2 Description of Tool

We have developed a web scrapper in python language that will automate the process of web metric extraction from a web page. Web scraping (web harvesting or web data extraction) [30] is a computer software technique of extracting information from websites. Usually, such software programs simulate human exploration of the World Wide Web by either implementing low-level Hypertext Transfer Protocol (HTTP), or embedding a fully-fledged web browser, such as Internet Explorer or Mozilla Firefox [8prev].

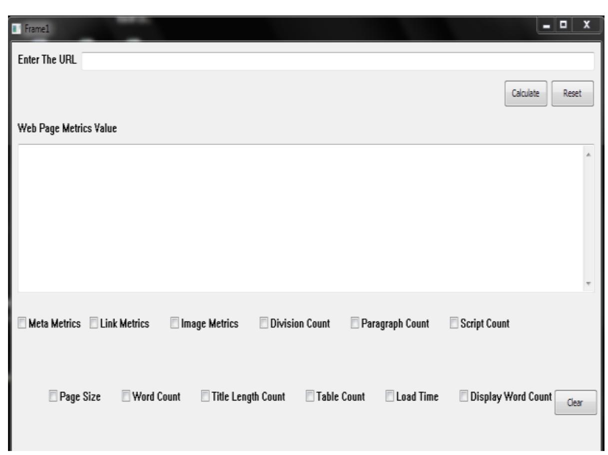


Figure.1.GUIof Tool

7.2.1 Algorithm of Web Scrapper

INPUT - URL of a Web Page

OUTPUT-Web page Metrics (only those whose checkbox are on)

STEPS:-

1. Enter the URL of the web page. The format of the URL http:// www.example.com
2. Store the Source Code of the page in a Variable called "soup" using method urllib2 and bs4(Beautiful Soup)
3. Variable "soup" is passed as a parameter to different methods in order to calculate the web page metrics.
4. The Page Metrics which we calculate are

a. Metrics related to meta tag:

- Number of meta tag:- Calculated by counting all the meta tag in a page.
- Meta keywords:- Calculate all the keywords which are present in a meta tag attribute keywords. Keywords are separated with comma (,) therefore words between two commas is considered as one keyword.
- Maximum and Minimum length of a meta keyword:- The length of each meta keyword is find out . Length here represents number of words in a keyword. Out of these length minimum and maximum length are find out.
- Meta Descriptors:- Calculate number of words in a meta tag attribute descriptor. While calculating number of words we also consider special symbols like '|', '"', ',' , ',' etc but neglect the spaces between the words.

b. Metrics related to Links:

- Total links in a page:- Search all anchor tag in a page as we know anchor tag are used to
- Image links :-

If between <a> and an <img> tag is present Then it is an image link Else not an image link

- Average number of words in a link: - for all links which are not image link there is string between <a> and .

Link name= sum of words between <a> and

Average number of words in a link = sum of Link name of each non image link/total links.

c. Metrics related to Images:

- Total images: - total <img> tag which are present in a web page.
- Alt images: - total images which has an alt attribute.
- Words in alt images:- Sum of the words in a string of an alt attribute for all Alt images.

d. Number of division of a web page:

- Search all <div> tag in a web page. This searching is done using the help of Beautiful Soup.

e. Total paragraph in a web page:

- Search all<p> tag in a web page.

f. Number of Scripts in a web page:

- Search all <script> tag with attribute type = "text/javascript"

g. Size of a web page

- Calculating web page size we retrieve the entire resource and calculate its length .This is done using the urlopen method of urrllib2 module.

h. Body Word Count:

- Word count represents total number of words on a web page when we load it. We calculate word count by converting an html page into a text page.
- HTML to TEXT conversion
- Remove all comments. In html comments are thing which are between <!- and -- >. Eg. <!—this is a comment -- >
- Remove all <script> tag as they contain definitions of function which are not displayed by browsers.
- The words which are in alt attribute of an image are not included in word count. Title length not included in word count.

i. Title length:

- Number of words which are present <title> ……</title>
- Commas are not included in title length.
- Semi-colons are not included in title length.
- Newline and Tab are not counted in title length.

j. Load Time:

- Number of time took by urlopen module to read the web page. Time module of python is also used to calculate load.

Start time = when reading of a page started

End time = when reading of a page ended.

Load time = End time – Start time (in sec)

k. Metrics related to display word count:

- Total headings in a web page:- There are 6 types of heading in html h1,h2,h3,h4,h5,h,6

Total headings = total <h1>tag + total<h2>tag + total<h3>tag total<h4>tag + total<h5>tag + total<h6>tag.

- Link headings: - Heading which are also links.

IF <a> between <h1> and </h1>

Then it is link heading.

ELSE

Not a link heading.

5. Web metrics which are calculated are displayed on tool user interface. This estimation can be saved for the further references.

7.3 Machine Learning Algorithms For Data Analysis

7.3.1 Naive Bayes Classifier

Naive bayes classifier is a supervised machine learning algorithm (needs to be trained) based on the Bayesian theorem. Naive bayes is also called idiot bayes and it assumes that all features are conditionally independent given the class label [7].

To demonstrate the concept of naive bayes classifier, consider an example

Let there be a set of variables X ={x1, x2, x3……xn} and variables are classified as C1 and C2. Initially n1 belong to C1 class and n2 belong to C2 class. Consider a new variable z which is to be classified

According to Bayes theorem

$$\text{Posterior} = \frac{\text{Prior *likelihood}}{\text{Evidence}}$$

Therefore

Posterior probability of z being C1 = prior probability of C1 * likelihood of z in C1

Prior probability of C1 = n1/X

$$\text{Likelihood of } z \text{ in } C1 = \frac{\text{number of } C1 \text{ in vicinity of } z}{n1}$$

Similarly cases for C2 class and the z belong to class which has higher posterior probability

Advantages

It is quite accurate and very fast. Out performs more sophisticated classifiers on many datasets, achieving impressive results [2].

7.3.2 Bagging

Bagging [bootstrap aggregating] was proposed by Leo Breiman [8] in 1966 to reduce the variance of predictor. It improves the classification by combining classifications of randomly generated training sets.

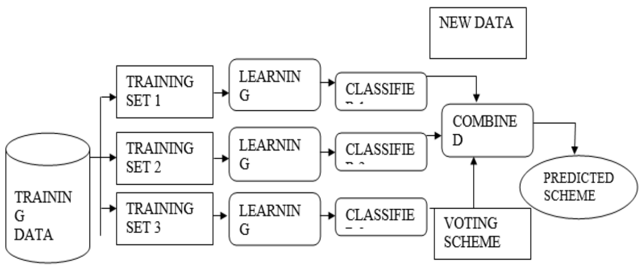


Figure.2. Flowchart of Bagging Classifier

Classification: - Voting scheme.
Prediction: - Averaging scheme.

The effect of combining different classifiers (hypotheses) can be explained with the theory of bias-variance decomposition.

• Bias – an error due to a learning algorithm
• Variance – an error due to the learned model (data set related)
• The total expected error of a classifier = Bias + Variance

Bagging provides a substantial reduction in prediction error for regression as well as classification methods. Since the method employs averaging of several predictors, it is not useful for improving linear models.

8. NAIVE BAYES ANALYSIS

Table 8.1 and 8.2 represent website predication and 10-cross validation results for goodness of a webpage for all the 3 models by naïve bayes classifier.

Observation which are made from Table 5.1 and 5.2:

- In model 1, 32 website out of 35 are correctly predicted as good and 39 websites out of 84 are correctly predicted as bad.
- In model 2, 7 websites out of 12 are correctly predicted as good and 35 websites out of 39 are correctly predicted as bad.
- In model 3, 19 websites out of 28 are correctly predicted as good and 74 websites out of 101 are correctly predicted as bad.

Table .6. Goodness Of Websites Using Naïve Bayes Classifier For Model 1, 2 And 3

Parameter	Model 1	Model 2	Model 3
Number of good websites correctly predicted	32	7	19
Number of bad websites correctly predicted	39	35	74

Table.7. 10-Cross Validation Results for Models Using Naïve Bayes Classifier for Goodness.

Model	Sensitivity	Specificity	Cutoff	AUC
Model 1	.714	.714	.834	.776
Model 2	.833	.795	.2485	.793
Model 3	.714	.723	.4095	.805

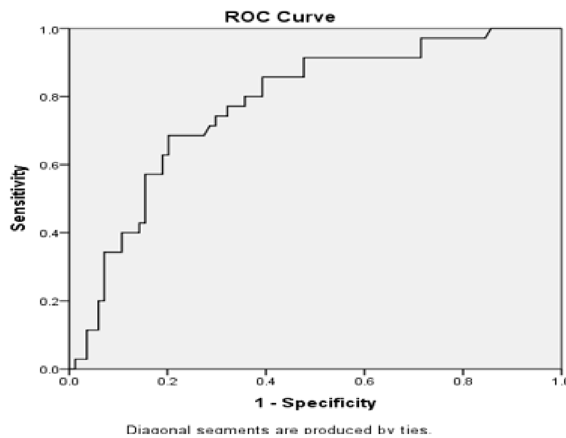


Figure.3. ROC Curve for Model 1 using Naïve Bayes Classifier for goodness

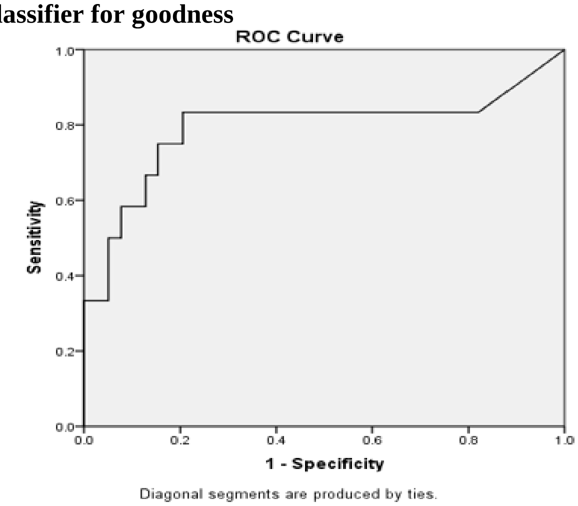


Figure.4.ROC Curve for Model 2 using Naïve Bayes Classifier for goodness

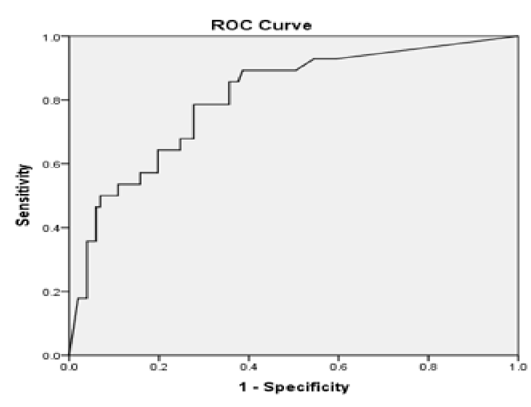


Figure.5. ROC Curve for Model 3 using Naïve Bayes Classifier for goodness

Table 8.3 and 8.4 represent web page predication and 10-cross validation results for classification of web page of same category of all the 3 models by naïve bayes classifier.

Observation which are made from Table 5.6 and 5.7:

- In model 1, 20 websites out of 62 are correctly predicted as Blog and 47 websites out of 57 are correctly predicted as Community.
- In model 2, 13 websites out of 25 are correctly predicted as TV and 20 websites out of are correctly predicted as 26 Movies.
- In model 3, 22 websites out of 39 are correctly predicted as Food & Beverages, 38 websites out of 49 are correctly predicted as Commerce and 22 websites out of 41 are correctly predicted as Travel.

Table.8.class prediction of websites using naïve bayes classifier for model 1, 2 and 3

Model	Websites	data points
Model 1	Blog	20
	Community	47
Model 2	TV	13
	Movies	20
Model 3	Food & Beverages	22
	Commerce	38
	Travel	22

Table.9. 10-cross validation results for models using naïve bayes classifier for class prediction

Model		Sensitivity	Specificity	Cutoff	AUC
Model 1		.661	.614	.190	.657
Model 2		.750	.652	.767	.674
Model 3	Food & Beverages	.61	.817		.855
	Commerce	.58	.828		
	Travel	.78	.81		

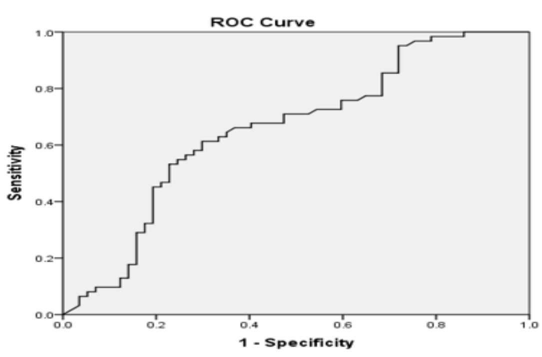


Figure.6. ROC Curve for Model 1 using Naïve Bayes Classifier For Class Prediction

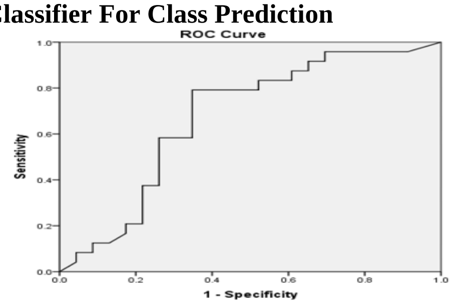


Figure.7.ROC Curve for Model 2 using Naïve Bayes Classifier for class prediction

9.CONCLUSION AND FUTURE WORK

The goal of this research is to find the effect of web page measures on the categorization of web sites into good or bad and also their effect on classifying websites of same category/model. Different machine learning techniques have been applied for classifying and categorization of websites and also analyzed their performance. The main contribution of this report is summarized as follows: First, we collected 3 sets of data of Pixel Awards for each category we created from 2006 to 2012, considering 0-level and some 1-level web pages for each website. Second, we computed 21 web page metrics for these web pages using a PYTHON based tool. Third, we applied machine learning methods such as Naïve Bayes, to predict the effect of web page metrics on the classification of web pages into good or bad classes and predict the class of web site of same category. Although, this research is conducted for three categories only, in which two categories have two web sites and one have three websites, this study can be repeated for more categories. Our main results are summarized as follows: The most significant metrics for categorization of web sites into good or bad for Model 1 are Meta Keywords, Meta Descriptor, Division, Script, and Title Length. Paragraph, Script, Load Time for Model 2 and Meta Keywords, Meta Descriptor, Total Link, ALT words, Paragraph , Body Word Count, Title Length for Model 3. This signifies that for different categories, the various attributes were included as important metrics for web site development. The most significant metrics for categorization of web sites for Model 1 are Meta tag, Meta Descriptor, Total Link, Image Link, Script. Meta tag, Total Link, Avg. word in Link, Division for Model 2 and Meta tag, max keyword, Meta descriptor, Avg. word in link, ALT images, script, size, and Title length for Model 3. This signifies that for different categories, the various attributes were included as important metrics for web site development. Random Forest outperformed the other models although all models predicted good area under ROC analysis.

9.1 Application of Work

In our work we are able to establish two relationships. Firstly, web metrics and quality of the website. Secondly, web metrics and class of the web site. In order to establish these relationships we have collected data sets from Pixel awards website which honors web site on different criteria's. A website designer not able to understand these criteria's and cannot use them to improve development process. Thus, over work able to provide designers with important metrics that can be helpfully in web site design and also the model for verifying the quality of website. Website quality can easily be estimated by computing the values of values of web metrics and then applying the Random Forest model which is more effective than all the models. The websites which are classified as bad need more attention. We also identify the class of website within same categories of websites. This will help designer to be careful with web metrics values which are helpful in predicting the class so that websites can be distinguishable.

9.2 Future Work

Our study confirms that web metrics can be helpful in predicting the goodness and class of the websites of same category with the help of machine learning methods. In future we can do similar study on different data set and also consider more web page metrics. We plan to carry our research for all the levels of web pages in the website.

II. REFERENCES

[1].C. Calero, J. Ruiz, and M. Piattini, "Classifying web metrics using the web quality model", Online Information Review, vol. 29, n. 3, pp. 227- 248, Emerald Group Publishing, 2005.

[2].I. H. Witten, E. Frank, and M.A. Hall, “Data Mining: Practical Machine Learning Tools and Techniques,” Morgan Kaufmann, San Francisco, 3 edition, 2011.

[3].ISO 9126-1, "Software Engineering-Product Quality - Part 1: Quality Model",October 2001.

[4]. J. Jugini, S.Laskowski, “Design of a File Format for Logging Website Interaction”, NIST Special Publication, 2001.

[5].J. Scholtz, S. Laskowski and L. Downey, “Developing usability tools and techniques for designing and testing web sites,” In Proceedings of the 4th Conference on Human Factors & the Web, 1998.

[6].K. M. Khan, “Assessing Quality of Web Based System,” IEEE/ACS International Conference on Computer Systems and Applications, AICCSA, 2008.

[7].K. P. Murphy, “Naive Bayes Classifiers,” Technical Report, October 2006.

[8].L. Breiman, “Bagging Predictors,” In Machine Learning Journal 26(2), 123-140, 1996.

[9].L. Mich, M. Franch, and L. Gaio, "Evaluating and Designing the Quality of Web Sites", IEEE MultiMedia, vol. 10, n. 1, pp. 34-43, IEEE Computer Society, 2003.

[10].L. Olsina and G. Rossi, "Measuring Web Application Quality with WebQEM", IEEE MultiMedia, vol. 9, n. 4, pp. 20-29, IEEE Computer Society, 2002.

[11].M. A. Hall, “Correlation-based Feature Subset Selection for Machine Learning,” PhD thesis, Department of Computer Science, University of Waikato, Waikato, N.Z., 1999.

[12]. M. Marsico, S. Levialdi "Evaluating web sites: exploiting users expectations", Department of Computer Science, University of Rome, 2003.

[13].M. Nazzal Jamal, M. El-Emary Ibrahim, A. Najim Salam, “Multilayer Perceptron Neural Network for Analyzing the Properties of Jordan oil Shale”, World Applied Sciences Journal 5: 546-552,2008.

[14].M. Stone, “Cross-validatory choice and assessment of statistical predictions,” In Journal of the Royal Statistical Society, Series B (Methodological), 36, 111–147,1974.

[15].M. Y. Ivory Melody, R. RashmiSinha, A. Marti Hearst, “Empirically Validated Web Page Design Metrics”, ACM SIGCHI’01, March 31- April 4,2001.

[16].M. Y. Ivory, R. Sinha, and M. A. Hearst, “Preliminary findings on quantitative measures for distinguishing highly rated information-centric web pages,” In Proceedings of the 6th Conference on Human Factors and the Web, 2000.

[17].M. Zorman, V. Podgorelec, P. Kokol, and S. H. Babic, “Using machine learning techniques for automatic evaluation of Websites,” Proceedings of the Third International Conference on Computational Intelligence and Multimedia Applications ICCIMA '99, pp. 169-173, New Delhi, India, IEEE Computer Society Press, September 1999.

[18].NIST, IEEE Std 2001-1999, from Web Site: http://zing.ncsl.nist.gov/WebTools/WebSAT/ieee guide.html.

[19].NIST, Web Metrics Test bed: Technical Overview, from Web Site: http://zing.ncsl.nist.gov WebTools/tech.html.

[20].O. Signore, "A Comprehensive Model for Web Sites Quality", in proceedings of the Seventh IEEE International Symposium on Web Site Evolution (WSE'2005), pp. 30-38, Budapest, Hungary, 2005.

[21].P. Warren, C. Gaskell, and C. Boldyreff, "Preparing the Ground for Website Metrics Research," Proceedings of the 3rd International Workshop on Web Site Evolution (WSE 2001), Florence, Italy, IEEE Computer Press, 2001.

[22].Pixel Awards, Web Awards Competition (2006), from Web Site, http://www.pixelawards.com/

[23].Rio. Americo “Websites Quality: Does It depend on the application Domain?”International Conference on the quality of Information and Communications Technology, 2010.

[24].S. Koukoulas and G. A. Blackburn, "Introducing new indices for accuracy evaluation of classified images representing semi-natural woodland environments," Photogramm Eng Rem S, vol. 67, pp. 499-510, 2001. Page 76

[25]. SukhpalKaur ,”An Automated Tool For Web site Evalution”, IJCSIT, Vol. 3(3),2012.

[26]. V. M. R Penichet, C.Calero, M.D.Lozano, M.Piattini,” Using WQM for classifying usability metrics”, IADIS International Conference, November 2006

[27]. WAI, Checklist of Checkpoints for Web Content Accessibility Guidelines 1.0, from Web Site: http:// www.w3. org/TR/ WCAG10/full-checklist.html.

[28]. WAI, Web Content Accessibility Guidelines 1.0. W3C Recommendation, from Web Site: http:// www.w3. org/TR/ WCAG10/.

[29]. Watchfire, "Site Quality and Accessibility, from Web Site: http://www.watchfire.com products/webxm/siteusability.aspx.

[30]. Web Scraping, from Web Site, http://en. wikipedia. org/ wiki/ Web_scraping

[31]. WebTango, Automating Web Site Usability Evaluation, from Web Site http://webtango.berkeley.edu/.

[32].Weka3: Data Mining Software in Java. Available from http://www.cs.waikato.ac.nz/ml/ weka/.

[33].Wen Zhu, Nancy Zeng, NingWang, “Sensitivity, Specificity, Accuracy, Associated Confidence Interval and ROC Analysis with Practical SAS® Implementations”, Health Care and Life Sciences, NESUG 2010.

[34]. Y. Freund and R. E. Schapire, “A Decision-Theoretic Generalization of On-Line Learning and an Application to Boosting,” In Journal of computer and system sciences 55, 119_139, 1997.