



SDI Review Form 1.6

Journal Name:	British Journal of Mathematics & Computer Science
Manuscript Number:	Ms_BJMCS_34128
Title of the Manuscript:	Web Document Classification Using Naïve Bayes
Type of the Article	

General guideline for Peer Review process:

This journal's peer review policy states that **NO** manuscript should be rejected only on the basis of '**lack of Novelty**', provided the manuscript is scientifically robust and technically sound.

To know the complete guideline for Peer Review process, reviewers are requested to visit this link:

(<http://www.sciencedomain.org/page.php?id=sdi-general-editorial-policy#Peer-Review-Guideline>)



SDI Review Form 1.6

PART 1: Review Comments

	Reviewer's comment	Author's comment (if agreed with reviewer, correct the manuscript and highlight that part in the manuscript. It is mandatory that authors should write his/her feedback here)
<u>Compulsory</u> REVISION comments	Under Introduction and Review of related work, only recent papers from 2007 to till date, references are to be listed and compared. All 26 reference papers are not necessary. Fig 2,3,4 are not necessary since they are screen dump. Under conclusion, large amount of data means , how much in size ? The result should be compared with existing methods and graphical chart will be of importance to exhibit relative performance. PLAGIARISM ISSUE: As per quetext.com ===== The Naïve Bayes classifier is based on the simplifying assumption that the 104 attribute values are conditionally independent given the target value. https://www.researchgate.net/profile/Q_M_Jonath... RESEARCH METHODOLOGY 112 The research methodology is in three major steps which are Data collection, Data preparation and the 113 Machine Learning. https://www.slideshare.net/NoopurSahu1/research... 114 4.1 Data Collection 115 Generally, the data	



SDI Review Form 1.6

	collection step involves gathering text (or web documents) from LAUTECH 116 website, Figure 1 below illustrates a web page sample from which our data were collected. https://issuu.com/archiblog/docs/ggis_1.8_user ... Since our focus is on the texts alone, in the data 118 preparation stage all the pictures and html tags will be removed because the HTML structure can 119 carry some semantics associated with the document class, the HTML structure is ignored so as to 120 simplify the document processing and document representation. http://dl.acm.org/citation.cfm?id=632805 Firstly, a 127 concatenation of all the text or web documents (text corpus) obtained from the data collection step will 128 be done and saved in a single text file, where each document is represented on a separate line in 129 plain text format. https://www.scribd.com/document/99080631/ASP-Ma...	
<u>Minor</u> REVISION comments		
<u>Optional/General</u> comments		

Reviewer Details:

Name:	<i>S.Sridhar</i>
Department, University & Country	<i>R.V.College of Engineering, India</i>