Abstract

Website categorization is one of the challenging tasks in the world of ever increasing web technologies. There are different ways to categorization of web pages using different features and approach. Website contains lot of information like text, images, animation, video and links. So this information is call as features of website. For the website categorization purpose all Feature have most important role. The web has a lot of information in the form of images, video, animation and text etc present in the document. In proposed System uses number of feature of website and use three different classifier for website classification are naive bays classifier, linear classifier- perceptron and stochastic classifier. Here eight major categories of
website have been selected for categorization; these are business & economy, job search, and science, education, sports, news & media, government, entertainment. Proposed system gives ranking to website. It will be more helpful for software developer or website designer for evolution of their site using our system so that they can judge that their website belongs to respective category or not.

References

- J. B. Leela devi, 2dr. A. Sankar - Improvements in Neural Network for Classification of Web Pages 1j. B. Leela devi, 2dr. A. Sankar 1University research scholar, Anna University, Tamil Nadu, India 2. Associate Professor, PSG College of Technology, Coimbatore, India.
- Brian D. Davison and Xiaoguang Qi. Web Page Classification: Features and Algorithms. In ACM Computing Surveys, Article 12, Publication date: February 2009
- H. Yu, J. Han, and K. C. C. Chang. Pebl - Positive example based learning for web page categorization. In Canada, KDD, Edmon-ton, Canada, 2002,
- In Iran University of Science Technology (IUST), Tehran, Iran. Arash Rezaei, and Behrouz Minaei-Bidgoli - Comparison between the Classification Methods using Type of Attributes and Sample Size.
- Sundus Hassan and Muhammad Shahid in Computer Science Department NUCES-FAST, Karachi Campus, Comparing NB Classifier SVM for Text Classification.
- Komal Kumar, Pikakshi Manchanda and Sonali Gupta. The Automated Classification of Web Pages Using Artificial Neural Network. Department of Computer Science, Faridabad, India.

Index Terms

Computer Science

Information Sciences
Keywords
Linear Classifier- Perceptron  Machine Learning  Naive Bayes  Stochastic Classifier