Abstract

In this paper, we present a comprehensive review of recent developments in the application of machine learning algorithms to Spam filtering, focusing on textual approaches. We are trying to introduce various spam filtering methods from Naïve Bias to Hybrid methods for spam filtering, we are also introducing types of filters recently used for spam filtering along with architecture of spam filter and its types. In this paper we are proposing a technique using Local feature classification methods with K mean clustering algorithm in classifier, for spam filtering term selection we are using Document frequency method, for feature extraction we are using bag of words model for classification we are using k-mean clustering method along with local concentration based extraction of content. This method gives good results along with all parameters.

References

Spam Filtering using K mean Clustering with Local Feature Selection Classifier

- Enrique Puertas Sanz, José María Gómez Hidalgo, José Carlos Cortizo Pérez Email Spam Filtering Universidad Europea de Madrid Villaviciosa de Odón, 28670 Madrid, SPAIN.


- Meharn Shami, Susan Dumais, David Hekerman, Eric Horvitz &quot;A Bayesian Approach to Filtering Junk e mail&quot; Microsoft Research.

- Shugang Liu & Kebin Cui &quot;Applications of Support Vector Machine Based on Boolean Kernel to Spam Filtering&quot; Modern Applied Science ccsenet journal Volume 3, No, 10 October 2009

- Andrew Secker, Alex A. Freitas, Jon Timmis &quot;AISEC: an Artificial Immune System for E-mail Classification&quot; Evolutionary Computation, 2003. CEC &apos;03. The 2003 Congress on (Volume:1 ) 2003

- Chris Miller, Group Product Manager Enterprise Email Security &quot;Neural Network-based Antispam Heuristics&quot; Symantec Enterprise Security

- Surendra Kumar Rakse, Sanyam Shukla &quot;Spam Classification using new kernel function in Support Vector Machine&quot; (IJCSE) International Journal on Computer Science and Engineering, Vol. 02, No. 05, 2010, 1819-1823

- Qiang Wang, Yi Guan, Xiaolong Wang &quot;SVM-Based Spam Filter with Active and Online Learning&quot;

- Saadat Nazirova, "Survey on Spam Filtering Techniques" Communications and Network, 2011, 3, 153-160 Published Online August 2011

- Yuanchun Zhu and Ying Tan "A Local- Concentration-Based Feature Extraction Approach for Spam Filtering" IEEE TRANSACTIONS ON INFORMATION FORENSICS AND SECURITY, VOL. 6, NO. 2, JUNE 2011

- Ion Androutsopoulos, Paliouras, G., & Michelakis, E. &quot;Learning to filter unsolicited commercial e-mail.&quot; Tech. rep. 2004/2, NCSR &apos;09.&apos; Demokritos.&quot;


Index Terms

Computer Science

Information Sciences
Keywords
Spam filtering  K mean.