**System ‘Spamfree’**

**Version:** 0.1

**Title**: Spamfree Interface

**Description**: Uses data of spam and non-spam emails, detecting various words in email content. Full Knowledge Discovery in Databases process is executed by collecting the .arff data file, selecting attributes, preprocessing, scaling (transforming), using six classification techniques (data mining), and evaluating those classifiers using 10-fold cross validation (pattern evaluation). A function is designed which uses an email in the format of a string, detects which spam keywords are in the text and ultimately predicts if the email is spam or not spam.

**Depends**: R(>=3.3.1)

**Required Library**: Foreign, RWeka and rvest

**Author**: Josephine Michalek 385809

**Date**: 2016-10-05 10:55:49

R topics documented

detect\_spam *detecting whether the email is spam or non-spam* 2

## detect\_spam *detecting whether the email is spam or non-spam*

**Description**

This function helps the user by analyzing emails and classifying whether they might be spam or not spam. You can use detect\_spam by inputting a string of the email in question, and the answer will be output.

**Usage**

detect\_spam(emailstring)

**Argument**

emailstring A string.

**RWeka, rvest**

This function utilizes tools from the RWeka and rvest packages, so you’ll need to make sure to have both installed.

**Examples**

**#**A spam email will need to be available for you to use as a string

emailstring = “Subject: zero down internet opportunity !

$ 0 down internet opportunity \* complete computer system \* e - com web site package \* tutorials - we ' ll teach you \* earn big while you learn for more info : 520-615 – 3253”

detect\_spam(emailstring)

[1] Spam