



## بخشی از ترجمه مقاله

عنوان فارسی مقاله :

شناسایی Spam در بین اطلاعات به دست آمده ایمیل-  
یک آنالیز عملکرد نسبی

عنوان انگلیسی مقاله :

Spam Mail Detection through Data Mining – A Comparative  
Performance Analysis



توجه !

این فایل تنها قسمتی از ترجمه میباشد. برای تهیه مقاله ترجمه شده کامل  
با فرمت ورد (قابل ویرایش) همراه با نسخه انگلیسی مقاله، [اینجا](#) کلیک نمایید.



## بخشی از ترجمه مقاله

### VI. CONCLUSION

In Spam mail classification is major area of concern these days as it helps in the detection of unwanted e-mails and threats. So now a day's most of the researchers are working in this area in order to find out the best classifier for detecting the spam mails. So a filter is required with high accuracy to filter the unwanted mails or spam mails. In this paper we focussed on finding the best classifier for spam mail classification using Data Mining techniques. So we applied various classification algorithms on the given input data set and check the results. From this study we analyze that classifiers works well when we embed feature selection approach in the classification process that is the accuracy improved drastically when classifiers are applied on the reduced data set instead of the entire data set. The results gained were promising Accuracy of the classifier Random Tree is 99.715% with best-first feature selection algorithm and accuracy is 90.93% only when we don't apply this subset selection algorithm. So, here in this study we achieve highest accuracy = 99.715%. As we all know that it is very difficult to achieve 100% accuracy but Random Tree and Random Forest (accuracy>99%) is very nearby to that. Therefore it is find that tree like classifiers works well in spam mail detection and accuracy improved incredibly when we first apply feature selection algorithm into the entire process.

### نتیجه‌گیری

در طبقه‌بندی ایمیل‌های Spam امروزه کارهای زیادی برای شناخت این ایمیل‌های نامطلوب بررسی شده است. امروزه خیلی از محققین برای پیدا کردن بهترین طبقه‌بندی کننده برای یافتن ایمیل‌های Spam تلاش زیادی می‌کنند. بنابراین فیلترهای با دقت بالا نیاز است تا ایمیل‌های نامطلوب را حذف کند. در این مقاله ما به بررسی و یافتن طبقه‌بندی کننده‌هایی برای یافتن بهترین دسته برای ایمیل‌های Spam با استفاده از ساختار جمع‌آوری اطلاعات پرداختیم. بنابراین ما انواع مختلفی از الگوریتم‌های طبقه‌بندی با ورودی‌های مختلف آزمایش کردیم. با آنالیز طبقه‌بندی کننده‌ها به این نتیجه رسیدیم که موقعی که ما روش انتخاب ویژگی را در مراحل طبقه‌بندی به کار بردیم، دقت کار به طرز قابل توجهی افزایش پیدا کرد. با توجه به نتایج به دست آمده دقت روش درخت رندم هنگامی که الگوریتم انتخاب ویژگی را به کار برده‌ایم ۹۹/۷۱۵ درصد است و هنگامی که الگوریتم انتخاب ویژگی را به کار نبرده‌ایم دقت آن ۹۰/۹۳ درصد است. بنابراین در این مقاله ما به بیشترین دقت که در حدود ۹۹/۷۱۵ درصد است رسیدیم. همان‌طور که می‌دانیم رسیدن به دقت ۱۰۰ درصد تقریباً مشکل می‌باشد اما روش درخت رندم و جنگل رندم به دقتی بیشتر از ۹۹ درصد رسیده‌اند. بنابراین واضح است که روش درخت رندم همانند طبقه‌بندی کننده‌ها به خوبی در تشخیص Spam و دقت آن کار می‌کند. هنگامی که الگوریتم انتخاب ویژگی را برای کل مراحل به کار بردیم به دقت خیلی خوبی رسیدیم.



### توجه!

این فایل تنها قسمتی از ترجمه می‌باشد. برای تهیه مقاله ترجمه شده کامل با فرمت

ورد (قابل ویرایش) همراه با نسخه انگلیسی مقاله، [اینجا](#) کلیک نمایید.

همچنین برای مشاهده سایر مقالات این رشته [اینجا](#) کلیک نمایید.