

要 旨

SPAM メールの判別に適した機械学習

藤森 夏輝

現在, SPAM メールか否かを判別する手法として, 機械学習の一つであるナイーブベイズ分類器 (ベイジアンネットワーク) があり, 迷惑メールフィルタ製品として広く用いられている. しかし, その他の多くの機械学習手法を SPAM メールの特徴量に適用し, 定量的に比較された研究はあまりない. そこで本研究では, SPAM メールの判別に, ナイーブベイズ分類器, ニューラルネットワーク, サポートベクターマシン (SVM), バギング, AdaBoost, RandomForest をそれぞれ適用し, 英語と日本語の SPAM メール判別を行い, 各手法の学習性能を明らかにする. 英文 SPAM メール判別のデータには, UCI Machine Learning Repository 提供のデータセット「Spambase」を用いる. 総数 4601 通のうち, ランダムに選出された 500 ~ 4000 通をそれぞれ訓練データとし, それぞれの残りをテストデータとして判別を行う. 日本語 SPAM メール判別のデータセットは, 独自に作成したコーパスを用いる. 全 1400 通のうち, ランダムに選出された 1000 通を訓練データとし, 残りをテストデータとして判別を行う. 英語 SPAM メールにおいて同様の条件のもとで判別を行い, 判別結果を比較する. 結果として, 英語 SPAM メール判別では, Bayes モデル以外の 5 手法が 90.6 ~ 94.9 %の判別率となることを示す. また, 日本語の SPAM メールの判別率は, Bayes モデルが 42.8 %, そのほかの手法は 77.0 ~ 79.5 %の判別率となることを示す.

キーワード SPAM, 機械学習, ナイーブベイズ分類器, ニューラルネットワーク, サポートベクターマシン, バギング, AdaBoost, Random Forest

Abstract

Performance Comparison of Machine Learning Algorithms for SPAM Discrimination

Fujimori Natsuki

Many other machine learning techniques are able to applied for classification, and quantitative comparison is required. In this research, Naive Bayes Classifier, Neural Network, Support Vector Machine (SVM), Bagging, AdaBoost, and Random Forest are applied to classify e-mail written in both English and Japanese in order to filter SPAM mail out. Those algorithms are compared with each other from a viewpoint of classification precision. For English e-mail classification, the dataset "Spambase" of UCI Machine Learning Repository is used. Total number of the data is 4601, and training data is from 500 to 4000, which are randomly selected, and the rest are the test data. For Japanese e-mail classification, original corpus is created and used. Total number of the data is 1400 and training data are randomly selected to 1000. SPAM email in English discrimination is performed under the same conditions to compare the result of precision. As a result, for English SPAM distinction, all algorithms except Naive Bayes Classifier achieves the precision exceeding 90 %. Moreover, for Japanese SPAM, Naive Bayes Classifier became a distinction rate of 42.8 %, and 77-79 % obtained by other algorithms.

key words SPAM, Machine Learning, Naive Bayes Classifier, Neural Network, Support Vector Machine, Bagging, AdaBoost, Random Forest