

Abstract: Text classification Algorithms for Sports Injury Prevention

Ho Jang and Hyunju Lee*

*Dept. of Comm. and Info., Gwangju Institute of Science and Technology
{walker83,hyunjulee}@gist.ac.kr*

Abstract

With the increasing number of research papers written about sports injury prevention, it is crucial to find high-quality scientific evidence that can be used to improve care in clinical practice. For this purpose, relevant original research articles should be collected, although this is a labor intensive and time-consuming task. In this study, we develop a text classifier that enables automatic search of original research articles about sports injury prevention. For the positive samples, a human expert collected original research articles about sports injury prevention. Negative samples were collected using PubMed searching. The naive Bayes classifier and Support Vector Machine (SVM) were used to rank documents and their classification performance was compared to that of the PubMed search result for various sports injury topics. The results showed that 127 positive samples were used for the training classifier. In the comparison between two classifiers, the SVM and naive Bayes classifiers outperformed the PubMed search results. Text classifiers based on the naive Bayes classifier and SVM are useful in screening sports injury prevention articles and can reduce human labor.

Acknowledgement

This research was supported by the MKE (The Ministry of Knowledge Economy), Korea and Microsoft Research, under IT/SW Creative research program supervised by the NIPA (National IT Industry Promotion Agency) (NIPA-2011-C1810-1102-0047))