



## بخشی از ترجمه مقاله

عنوان فارسی مقاله :

کرنل معنایی مجموعه ای از نوشته ها و داده ها برای دسته بندی  
متون از طریق استفاده از ارزش واقعی واژگان

عنوان انگلیسی مقاله :

A corpus-based semantic kernel for text classification by using  
meaning values of terms



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

### 6. Conclusions and future work

We introduce a new semantic kernel for SVM called *Class Meanings Kernel* (CMK). The CMK is based on meaning values of terms in the context of classes in the training set. The meaning values are calculated according to the Helmholtz Principle which is mainly based on Gestalt theory and has previously been applied to several text mining problems including document summarization, and feature extraction (Balinsky et al., 2010, 2011a, 2011b, 2011c). Gestalt theory points out that meaningful features and interesting events appears in large deviations from randomness. The meaning calculations attempt to define meaningfulness of terms in text by using the human perceptual model of the Helmholtz principle from Gestalt Theory. In the context of text mining, the textual data consist of natural structures in the form of sentences, paragraphs, documents, topics and in our case classes of documents. In our semantic kernel setting, we compute meaning values of terms, obtained using the Helmholtz principle in the context of classes where these terms appear. We use these meaning values to smoothen document term vectors. As a result our approach can be considered as a supervised semantic smoothing kernel which makes use of the class information. This is one of the important novelties of our approach since the previous studies of semantic smoothing kernels does not incorporate class specific information.

### نتایج و کارهای آینده

ما کرنل معنایی جدیدی برای SVM، به نام CMK (کرنل معنایی طبقات)، تعریف می کنیم. CMK، بر مبنای ارزش معنایی عبارات، در زمینه طبقات مربوط به مجموعه های آموزشی، می باشد. ارزش های معنای مطابق با اصل Helmholtz، محاسبه شده که بر مبنای نظریه Gestalt می باشند که قبلا برای مسائل داده کاوی مثل خلاصه بندی اسناد و بیرون کشیدن ویژگی ها، مورد استفاده قرار گرفتند. (Balinsky و همکاران 2010، 2011). نظریه Gestalt، بی این نکته توجه می کند که این ویژگی های معنایی و وقایع جالب، به صورت تصادفی و بسیار متغیر ظاهر می شوند. محاسبات معنایی تلاش می کنند تا معنادار بودن عبارات در متون را با استفاده از مدل رویکرد ادراکی انسان تعریف کنند. این مدل بر مبنای اصل Helmholtz برگرفته شده از نظریه Gestalt می باشد. در زمینه داده کاوی، داده های متنی شامل ساختارهای طبیعی به شکل جمله، پاراگراف، اسناد، موضوعات و دسته بندی اسناد می باشد. ما در محیط کرنل معنایی، ما ارزش معنایی عبارات را در متون دسته بندی شده، با استفاده از اصل Helmholtz، محاسبه می کنیم. ما از ارزش معنایی، برای هموار کردن بردار عبارت در اسناد استفاده می کنیم.



### 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.

برای جستجوی جدیدترین مقالات ترجمه شده، [اینجا](#) کلیک نمایید.