

عنوان مقاله:

طبقه بندی متن بر اساس رویکرد ترکیبی از تجزیه مقادیر منفرد و ماشین بردار پشتیبان

محل انتشار:

پنجمین کنفرانس ملی تکنولوژی در مهندسی برق و کامپیوتر (سال: 1399)

تعداد صفحات اصل مقاله: 5

نویسندگان:

سینا دامی - استادیار گروه مهندسی کامپیوتر، واحد تهران غرب، دانشگاه آزاد اسلامی، تهران، ایران

نجمه فرقانی - دانشجوی کارشناسی ارشد IT، واحد تهران غرب، دانشگاه آزاد اسلامی، تهران، ایران

خلاصه مقاله:

در طبقه بندی متن هدف این است که سندهایی را که در اختیار داریم بتوانیم برچسب گذاری موضوعی کنیم. در واقع این موضوع صرفاً یک مسئله باناظر است، یعنی مجموعه ای از اسناد متنی که گروه بندیموضوعی شده اند به عنوان داده ی آموزشی در اختیار سامانه قرار می گیرد تا بتواند با یادگیری از این مجموعه، اسناد جدید ورودی را به یکی از اینگروه های موضوعی ملحق نماید. در طبقه بندی متن معمولاً از کلمات متن بهعنوان ویژگی های آن متن استفاده می شود در نتیجه مسئله طبقه بندی متنبا تعداد بالایی از ویژگی ها مواجه می باشند. در این پژوهش از روش ماشینبردار پشتیبان برای طبقه بندی متنها و تجزیه مقادیر منفرد برای انتخابویژگی های متن استفاده شده تا بتوان طبقه بندی متن را با کیفیت بالاتری بهانجام رساند. برای سنجش نتایج حاصله از مقایسه روش ما با دو روش دیگر مبتنی بر KNN و Naive Bayes استفاده شده است. در نتایجآزمایشگاهی مشاهده شد که روش پیشنهادی به طور میانگین از هر دو روشبهتر است.

کلمات کلیدی:

طبقه بندی متن، ماشین بردار پشتیبان، کاهش ویژگی ها، تجزیه مقادیر منفرد

لینک ثابت مقاله در پایگاه سیویلیکا:

<https://civilica.com/doc/1192705>

