



دانشگاه علوم پزشکی کرمان

دانشکده مدیریت و اطلاع‌رسانی پزشکی

پایان‌نامه مقطع کارشناسی ارشد کتابداری و اطلاع‌رسانی پزشکی

عنوان:

متن‌کاوی مقالات کتابداری و اطلاع‌رسانی نمایه شده در پایگاه اطلاعاتی وب‌آوساینس

به‌منظور کشف ساختار درونی رشته و اعتبارسنجی آن

توسط: عاطفه ذوالفقار نسب حاجی‌زاده

استاد راهنما: دکتر فاطمه معتمدی

استاد مشاور: دکتر محمد اعظمی

سال تحصیلی: ۱۳۹۵ - ۱۳۹۶

چکیده

مقدمه: مقالات علمی به دلیل ویژگی‌های خاصی که دارند، از اهمیت بسزایی برخوردارند. به‌منظور بازیابی مطلوب محتوای اطلاعاتی آن‌ها و جلوگیری از بروز مشکلاتی از جمله ریزش کاذب یا بازیابی مدارک نامرتب و همچنین پیشگیری از اتلاف وقت کاربران استفاده از روش‌ها و فنون تحلیلی مناسب جهت کاوش و پردازش متون ذخیره‌شده، بسیار ضروری به نظر می‌رسد. متن‌کاوی شاخه‌ای از داده‌کاوی است که به دلیل تلاش در جهت دست آوردن دانش مفید از میان داده‌های متنی، اهمیتی دوچندان یافته است. لذا این مطالعه با هدف متن‌کاوی مقالات کتابداری و اطلاع‌رسانی نمایه شده در پایگاه اطلاعاتی وب آو ساینس به‌منظور کشف ساختار درونی رشته و اعتبارسنجی آن انجام گرفته است.

روش کار: این پژوهش توصیفی - مقطعی و از نوع کاربردی است که با استفاده از الگوریتم‌های خوشه‌بندی در متن‌کاوی صورت گرفته است. 51980 رکورد بازیابی شده با استفاده از روش سرشماری از پایگاه «وب آو ساینس» در بازه زمانی سال‌های 1983 تا 2015 با استفاده از نرم‌افزار هیست سایت جمع‌آوری گردید. به‌منظور انجام عملیات متن‌کاوی، در مرحله اول داده‌های جمع‌آوری شده ابتدا مورد پیش‌پردازش قرار گرفتند و در انتهای این مرحله، فرهنگ لغتی از واژگان رشته کتابداری و اطلاع‌رسانی ایجاد گردید. در مرحله بعد با به‌کارگیری نرم‌افزار رپیدماینر، الگوهای حاوی اطلاعات با استفاده از الگوریتم‌های خوشه‌بندی k-means به دست آمدند و در مرحله پایانی، نتایج اطلاعات استخراج‌شده و ویژگی‌های آن‌ها با استفاده از دو معیار دقت و بازیابی مورد ارزیابی و بررسی قرار گرفتند.

یافته‌ها: نتایج حاصل از اعمال الگوریتم خوشه‌بندی k-means در بستر متن‌کاوی ایجاد 7 خوشه ابتدایی در این حوزه بود که پس از بررسی و ارزیابی لغات تشکیل‌دهنده هر خوشه با عنوان مقاله و کلیدواژه‌های آن، 5 خوشه با نام‌های «اینترنت و فناوری»، «ژورنالیسم و علم‌سنجی»، «بازیابی اطلاعات»، «علم کتابداری و اطلاع‌رسانی» و «خدمات کتابخانه‌ای» تشکیل شد. به‌منظور ارزیابی خوشه‌های تشکیل شده نیز از دو معیار دقت و بازیابی که از معیارهای بازیابی اطلاعات هستند، استفاده شد که معیار دقت 0.85 و معیار بازیابی 0.71 به دست آمد که نشان‌دهنده مطلوبیت خوشه‌های به‌دست‌آمده بود.

نتیجه‌گیری: بررسی‌ها بر روی خوشه‌های تشکیل شده، نشان‌دهنده این بود که حوزه‌های مرتبط با فناوری و اینترنت بیشتر مورد توجه و حوزه‌های مرتبط با دانش نظری این رشته در حاشیه قرار گرفته‌اند. همچنین کلماتی در هر زیر گروه حوزه رشته کتابداری و اطلاع‌رسانی، کاربرد بیشتری دارند، مشخص شده‌اند و با توجه به همین موضوع، پایگاه‌های اطلاعاتی بزرگ، با بهره‌گیری از این کلمات، می‌توانند کاربران را در دستیابی به اطلاعات دقیق و مورد نیاز در مدت زمانی بسیار کوتاه، یاری کنند. علاوه بر

این، سیاست‌گذاران و محققین حوزه کتابداری و اطلاع‌رسانی ضمن شناسایی مباحث پرکاربرد و همچنین حوزه‌های کمتر مورد توجه در این رشته، می‌توانند در اعتلا و ارتقای این حوزه تأثیرگذار باشند.

کلمات کلیدی: متن‌کاوی، خوشه‌بندی، الگوریتم k-means، داده‌کاوی، مقالات کتابداری و اطلاع‌رسانی، بازیابی اطلاعات، اعتبارسنجی، پایگاه وب آو ساینس