

وزن دهی به اقلام بر اساس محتوا جهت ارتقاء سیستم‌های فیلترینگ اشتراکی

نیلوفر راستین^۱، منصور ذوالقدری جهرمی^۲

^۱ دانشجوی کارشناسی ارشد، دانشکده برق و کامپیوتر، دانشگاه شیراز، شیراز،
niloofar.rastin@gmail.com

^۲ استاد، دانشکده برق و کامپیوتر، دانشگاه شیراز، شیراز،
z.zjahromi@gmail.com

چکیده

روش‌های محتوا محور و فیلترینگ اشتراکی^۱ از راهکارهای موفق در سیستم‌های پیشنهادگر می‌باشند. روش محتوا محور بر اساس ویژگی‌های اقلام تعریف می‌شود. این روش بررسی می‌کند که اقلام مورد علاقه کاربر دارای چه ویژگی‌هایی بوده‌اند، سپس اقلام دارای ویژگی‌های مشابه را به او پیشنهاد می‌کند. روش فیلترینگ اشتراکی بر اساس تعیین اقلام مشابه یا کاربران مشابه کار می‌کند که به ترتیب فیلترینگ اشتراکی مبتنی بر اقلام و مبتنی بر کاربران نامیده می‌شود.

روش پایه فیلترینگ اشتراکی مبتنی بر کاربر، به منظور پیش‌بینی امتیاز قلم هدف، هیچ تمایزی بین اقلام قائل نمی‌شود. به عبارت دیگر امتیازهای تمامی اقلام به طور یکسان در انتخاب همسایگی (کاربران مشابه) و پیش‌بینی تاثیر می‌گذارند. در این مقاله یک سیستم پیشنهادگر فیلترینگ اشتراکی مبتنی بر کاربر، مجهز به مکانیزم تخصیص پویای وزن به اقلام، ارائه شده است. مبنای این مکانیزم، تخصیص وزن به اقلام بر اساس میزان شباهت آنها با قلم هدف می‌باشد. میزان شباهت اقلام توسط یک روش محتوا محور سنجیده می‌شود. از آنجا که پایگاه داده مورد استفاده در این مقاله مربوط به فیلم است، برای بالا بردن کارایی این روش علاوه بر استفاده از ویژگی‌های ژانرها، از ویژگی‌های دیگری از جمله کارگردانان و بازیگران به عنوان داده‌های مکمل استفاده شده است. سیستم ارائه شده توسط معیار استاندارد میانگین خطای مطلق، روی پایگاه داده‌های MovieLens و EachMovie سنجیده شده است. نتایج به دست آمده، نشانگر عملکرد قابل ملاحظه روش پیشنهادی در مقایسه با سایر روش‌های ارائه شده می‌باشد.

کلمات کلیدی

سیستم‌های پیشنهادگر، روش فیلترینگ اشتراکی، روش محتوا محور، انتخاب همسایگی، پیش‌بینی.

۱- مقدمه

روند کار سیستم‌های محتوا محور به این صورت است که ابتدا بر اساس نحوه امتیازدهی کاربر به اقلام مختلف، نمایه‌ای از علائق او ساخته می‌شود. سپس، بر اساس میزان تطابق خصوصیات اقلام با نمایه ساخته شده از کاربر، پیشنهادها به کاربر ارائه می‌شود.

روش فیلترینگ اشتراکی به دو گروه اصلی فیلترینگ اشتراکی مبتنی بر حافظه و مبتنی بر مدل تقسیم می‌شود. در فیلترینگ اشتراکی مبتنی بر حافظه سرتاسر ماتریس کاربران - اقلام بررسی می‌شود. در این ماتریس امتیازهایی که کاربران به اقلام مختلف داده‌اند وجود دارد. در راهکار مبتنی بر حافظه، پیش‌بینی براساس کاربران یا اقلام انجام می‌پذیرد. فیلترینگ اشتراکی مبتنی بر مدل از امتیازهای موجود برای یادگیری یک مدل پیش‌بینی استفاده می‌کند. این مدل بر

با افزایش ناگهانی اطلاعات، نیاز به وجود ابزاری ضرورت یافت تا کاربر توسط آن بتواند به اطلاعات مورد نیازش دسترسی پیدا کند. در این زمینه، سیستم‌های پیشنهادگر نقش قابل توجهی را ایفا کرده‌اند. این سیستم‌ها ابزارهای نرم افزاری و تکنیک‌هایی هستند که اقلام را مطابق با نیاز کاربر به او معرفی می‌کنند. اقلام می‌تواند فیلم، موزیک، صفحه وب، کالا یا سایر موارد باشد. دو راهکار کلیدی سیستم‌های پیشنهادگر، روش فیلترینگ اشتراکی و روش محتوا محور می‌باشد. البته راهکار سومی نیز از تلفیق دو روش مذکور به وجود می‌آید که منجر به ایجاد سیستم پیشنهادگر تلفیقی می‌شود.