

تلفیق خوشه بندی و مدل مارکوف در یک چارچوب جدید برای پیشبینی صفحه بعدی انتخابی توسط کاربر

مهدی مقیمی¹، حسن صفری نادری²، مهرداد جلالی³

¹ دانشجوی کارشناسی ارشد دانشگاه امام رضا (ع) - مشهد
Research.moghimii@gmail.com

² دانشجوی کارشناسی ارشد دانشگاه امام رضا (ع) - مشهد
h_safarinadery@yahoo.com

³ عضو هیئت علمی دانشگاه آزاد اسلامی واحد مشهد-مشهد
dr_mehrdadjalali@yahoo.com

چکیده

کاوش استفاده از وب که از شاخه های وب کاوی می باشد به پیشبینی صفحه بعدی کاربر و همچنین شناسایی رفتار کاربر می پردازد. یکی از چالش های این حوزه، تشخیص شیوه رفتاری کاربر جهت پیشبینی دقیقتر صفحه بعدی مورد نظر وی است. آنچه کار مهم است، صحت پیش بینی به همراه کاهش زمان مورد نیاز برای پیش بینی می باشد. ما در این مقاله از مدل خوشه بندی کامیانه استفاده کرده و توسط یک چارچوب جدید قابل گسترش، به پیش بینی صفحه بعدی کاربر می پردازیم. هدف ما در این مقاله ارائه راهکاری برای افزایش صحت و کاهش زمان پیش بینی برخط صفحه بعدی کاربر می باشد. نتایج حاکی از افزایش صحت پیش بینی و کاهش زمانی پیش بینی برخط به میزان قابل ملاحظه ای می باشد.

کلمات کلیدی

کاوش استفاده از وب، خوشه بندی، رتبه، کلاس بندی، درهم سازی

1- مقدمه

نشست³ (یا تراکنش) می نامیم و حاوی صفحاتی می باشد که هر کاربر از تارنمای مورد نظر بازدید کرده است. کار اصلی ما در این مقاله ارائه یک چارچوب ترکیبی⁴ مناسب می باشد که علاوه بر افزایش صحت پیش بینی، کاهش زمان پیش بینی را موجب می گردد. ما از [3] برای ارائه این چارچوب الهام گرفته ایم. همچنین هدف دیگر ما بهبود صحت مقاله شماره [1] می باشد. این مقاله از روش خوشه بندی بهینه شده کا میانه [15] برای خوشه بندی تراکنش های کاربران در فاز برون خط استفاده کرده است. آنها همچنین در فاز برخط کار خود، از یک روش ابداعی برای سوق دادن کاربران جدیدی که به سیستم وارد می شوند به بهترین خوشه استفاده می کنند و از این روش برای پیش بینی صفحه بعدی کاربر بجای استفاده از مدل های معمولی مانند مدل مارکوف استفاده می کنند. چارچوب مورد نظر طوری طراحی شده است که می توان آن را گسترش داد و در فاز پیش بینی، روش پیش بینی دیگری را جایگزین روش موجود نموده و یا به روش موجود روشی جدید افزود.

در حوزه های زیادی مساله وب کاوی مورد کاربرد می باشد. هدف اصلی، پیشنهاد صفحه یا صفحات جدید بر اساس صفحات پیموده شده قبلی و سلیقه کاربر است. کاوش استفاده از وب¹ زیر مجموعه ای از این حوزه می باشد که به کمک فایل رخداد⁸ سرویس دهنده؛ به پیش بینی صفحه بعدی درخواستی توسط کاربران می پردازد. این کار، کاربردهای زیادی دارد، به عنوان مثال می توان به افزایش سرعت سرویس دهی سرویس دهنده های وب به جهت بهبود فرایند کش آنها اشاره کرد. پیشنهاد بهترین صفحه باعث کاهش ترافیک به دلیل کاهش جستجو در صفحات، افزایش سرعت نمایش صفحات، خریدهای برخط که نتیجه آن افزایش رضایت کاربران خواهد بود. یکی از ویژگی های کار با این فایل ها، توانایی استخراج جریان کلیک² کاربران می باشد. جریان کلیک صفحات کاربران معیاری جهت شناسایی کاربران بر اساس سلاقی آنها است. به وسیله انجام تکنیک هایی به نام پیش پردازش، از فایل رخداد سرویس دهنده، فایلی جدید بدست می آید که هر سطر آن را یک