

تحلیل رضایت کاربر مبتنی بر رفتار ضمنی در حین جستجو

معصومه عظیم‌زاده^۱، نوید فرهادی^۱، محمد مهدی اثنی عشری^۱

^۱ پژوهشکده فناوری اطلاعات، پژوهشگاه ارتباطات و فناوری اطلاعات، تهران
{ azim_ma, farhadi, esnaashari }@itrc.ac.ir

چکیده

تا کنون روشهای مختلفی برای ارزیابی میزان رضایتمندی کاربران از موتورهای جستجو توسعه داده شده است. روشهای صریح از طریق ارزیابی انسانی و تحلیل مشاهدات کاربران به ارزیابی موتور جستجو می‌پردازند. انجام این روش‌ها به دلیل محدودیت تعداد افراد شرکت کننده در آزمایش و همچنین هزینه زیاد در مقیاس بالا امکان‌پذیر نمی‌باشد. در مقابل، روشهای ضمنی رضایت کاربران را به کمک بازخوردهای ضمنی آنها در حین استفاده از موتور جستجو تخمین می‌زنند. ایده اصلی این روش‌ها آن است که رضایت کاربر از یک نتیجه جستجو با رفتار او بعد از کلیک روی آن نتیجه مشخص می‌شود. در این پژوهش به منظور ارزیابی میزان رضایتمندی کاربران از نتایج موتور جستجوی پارسی‌جو، رفتارهای ضمنی مجموعه‌ای از کاربران این موتور جستجو، شامل کلیک، لغزش در صفحه، مشاهده و توجه به محتوای صفحه و ...، ثبت گردیده و سپس به کمک تکنیک‌های یادگیری ماشین و با استفاده از داده‌های آموزشی تولید شده توسط همین مجموعه از کاربران، وزن مناسبی برای هر یک از این رفتارها به دست آمده است. نتایج حاصله نشان‌دهنده اهمیت رفتارهای میزان لغزش در صفحه و مدت زمان فعال بودن کاربر در صفحه نتایج در تعیین میزان رضایت کاربر می‌باشد. مبتنی بر میانگین وزن دار رفتارهای ضمنی کاربران، میزان رضایتمندی از موتور جستجوی پارسی‌جو ۳۰ درصد بوده است.

کلمات کلیدی

رضایت کاربر، موتور جستجو، بازخورد ضمنی، افزونه مرورگر

۱- مقدمه

بازخورد صریح از کاربران دارای مشکلات و محدودیتهایی است به عنوان نمونه کاربر از فرایند دریافت اطلاعات آگاهی داشته و ممکن است بازخوردهای معمول خود را نداشته باشد. همچنین، داشتن کاربرانی که جستجوهای سفارشی انجام بدهند و نتایج ارزیابی را به صورت صریح اعلام

کاربران در ارزیابی کارایی سامانه‌های بازیابی اطلاعات نقش مهمی دارند چرا که رضایت کاربران مهمترین معیار ارزیابی موتور جستجو می‌باشد. به این منظور از بازخوردهای صریح یا ضمنی کاربران استفاده می‌گردد. دریافت

۲- کارهای مرتبط

با توجه به اهمیت رضایت کاربر از نتایج موتور جستجو فعالیتهای مختلفی برای بهره‌گیری از رفتار و بازخوردهای ضمنی کاربران انجام شده است. بخش عمده‌ای از آنها به تحلیل لاگ موتور جستجو [7-13] پرداخته‌اند. دسته دیگر از فعالیتهای از نرخ کلیک کاربر برای بهبود الگوریتمهای رتبه‌بندی و یافتن اهداف کاربران از جستجو استفاده نموده‌اند [14-16]. اما فعالیتهای محدودی به تحلیل عملکرد کاربران بعد از کلیک روی نتایج پرداخته‌اند.

در [17] ضمن اینکه تلاش شده وابستگی بین اخذ رضایت کاربر به صورت صریح و ضمنی اثبات شود، موثرترین پارامترهای تعیین کننده رضایت کاربر به صورت ضمنی محاسبه گردیده است. در انتها اثبات شده بهترین معیارهای پیش بینی رضایت شامل نرخ کلیک کاربر، زمان سپری شده در صفحه نتایج و روش خروج از صفحه (بستن مرورگر، ایجاد پرس و جوی جدید و...) می‌باشد.

برخی فعالیتهای با محوریت سنجش عدم رضایت کاربران از اطلاعات رفتاری کاربر بعد از کلیک روی نتایج استفاده نموده است. در [18] به منظور حفظ کاربران و جلوگیری از تمایل آنها به سمت سایر موتورهای جستجو میزان ناراضیاتی کاربر محاسبه شده است. به این منظور از لاگ کاربر که با استفاده از افزونه فایرفاکس جمع‌آوری شده استفاده گردیده است. افزونه داده‌هایی نظیر میزان تمرکز کاربر بر صفحه، رویداد کلیک، رویدادهای پیمایشی نظیر حرکات موس، کپی و چسباندن^۳ اطلاعات و لغزش در طول صفحه را ثبت می‌کند.

پژوهشهایی نیز مانند [20] وجود دارد که به موتور جستجو محدود نبوده و به حوزه‌های عمومی‌تر مانند مرور و پیمایش صفحات وب و سنجش علاقه‌مندی به آنها می‌پردازند. در این پژوهش رفتار مرورگری کاربر را شامل مواردی نظیر کلیک کردن، پرینت^۴ و انتخاب متن می‌داند. یکی از مزایای استفاده از این گونه روشها را مستقل از زبان بودن آنها می‌داند.

در [21] به کاربرد بازخوردهای ضمنی برای فیلتر صفحات پرداخته شده است. به این صورت که از طریق مرورگر وب با دنبال کردن اولویتهای کاربر به فیلتر صفحات می‌پردازد. در واقع برای هر کاربر مبتنی بر عکس‌العمل او در برابر صفحات فیلتر شده پروفایلی مبتنی بر یادگیری ایجاد می‌شود. در این پژوهش برای هر سند خوانده شده توسط کاربر اطلاعاتی نظیر زمان خواندن، زمان لغزش، نشانه‌گذاری و دنبال کردن پیوندها جمع‌آوری می‌گردد.

۳- راه‌حل پیشنهادی

رفتار کاربران در حین جستجو می‌تواند میزان رضایت آنان را از نتیجه ارائه شده توسط موتور جستجو نشان دهد. به عنوان نمونه، وقتی کاربر صفحه ارائه شده توسط موتور جستجو را باز کرده و بلافاصله بعد از چند ثانیه می‌بندد، نشان از این دارد که نتیجه مورد نظر مورد علاقه کاربر نبوده و مرتبط با پرس-وجوی وی نمی‌باشد. از طرفی دیگر زمانی که برای مدت قابل توجهی روی صفحه نتیجه باز شده تمرکز می‌کند و بخش‌های مختلف صفحه را مشاهده

کنند هزینه‌بر است. این در حالی است که، کاربران بسیاری هم‌روزه نیازهای اطلاعاتی خود را با وارد نمودن پرس و جوهای خود در جویسگرهای اینترنتی مرتفع می‌کنند و نتایج ارائه شده توسط موتور جستجو را مشاهده می‌کنند. با توجه به حجم بالای استفاده کنندگان از سیستمهای بازیابی اطلاعات و موتور جستجو استفاده از بازخورد ضمنی کاربران منجر به تسهیل و تسریع ارزیابی-های مورد نیاز و اعمال تغییرات متناسب با آنها می‌شود. ضمن آنکه کاربران از نقش و تاثیر خود بر این ارزیابی آگاه نیستند بنابراین قضاوت بی‌طرفانه‌ای از این طریق در اختیار ناظرین و توسعه‌دهندگان این سامانه‌ها قرار خواهد گرفت. بازخوردهای ضمنی کاربران در بسیاری از فعالیتهای ارزیابی سامانه‌های بازیابی اطلاعات و موتورهای جستجو بکاررفته است اما بیشتر این فعالیتهای به استفاده از اطلاعات لاگ موتور جستجو (سمت سرور) و نرخ کلیک^۱ کاربران پرداخته‌اند [7-16]. آمار نرخ کلیک کاربران زمانی که موتور جستجوی دارای کاربری بسیار بالایی باشد کارایی بیشتری در استخراج نتایج دارد. هر چند در برخی پژوهشها نرخ کلیک کافی دانسته نشده و از ترکیبی از آن با سایر فیدبکهای ضمنی کاربران استفاده شده است [17]. بازخوردهای ضمنی کاربر همچنین در سیستمهای پیشنهاد دهنده و سیستمهای فیلترینگ کاربرد زیادی دارد.

با توجه به اینکه یک موتور جستجوی نوپا مانند پارسا جو که در [1-2] و [5-6] ارزیابی شده علی‌رغم قابل قبول بودن نتایج با توجه به محدودیت منابع و زیرساختهای در دسترس نرخ کلیک خیلی بالایی ندارد. لذا استفاده از این معیار برای سنجش کیفیت نتایج کافی نیست. لذا در این پژوهش سعی شده از رفتار ضمنی کاربر بعد از بازشدن صفحه نتایج یا سربرگها به منظور ارزیابی رضایت کاربر از نتایج ارائه شده توسط موتور جستجو استفاده شود.

کاربران مختلف در حین مشاهده نتایج، رفتار متفاوتی با نتایج مرتبط یا نامرتب با پرس‌وجو دارند که از طریق ثبت آنها و استخراج الگوهای مشترک می‌توان بصورت خودکار رضایت کاربر را استخراج نمود. به عنوان نمونه کاربران در صورت نامرتب بودن سربرگ باز شده بلافاصله آن را می‌بندند و یا در صورت مرتبط بودن مدت زمان بیشتری در صفحه مانده و نتیجه را مرور می‌کنند. در این پژوهش برای جمع‌آوری اطلاعات مربوط به رفتار کاربران در حین جستجو یک افزونه برای مرورگر فایرفاکس توسعه داده شده است که بدون دخالت و توجه کاربر، می‌تواند رفتار وی را ثبت کند. رضایت کاربران از موتور جستجو از طریق میانگین وزن‌دار رفتارهای ضمنی کاربر تخمین زده می‌شود. تحلیل نتایج حاصله نشان می‌دهد که مهم‌ترین فعالیت کاربران پارسا جو شامل لغزش^۲ و کلیک روی نتایج بوده است. همچنین میزان رضایت کاربران از نتایج پارسا جو ۳۰ درصد می‌باشد.

مطالب این مقاله در ادامه به این صورت سازماندهی شده است. بخش دوم به مروری بر کارهای مرتبط اختصاص دارد. در بخش سوم راه‌حل پیشنهادی ارائه گردیده است. بخش چهارم به ارزیابی روش پیشنهادی و تحلیل نتایج اختصاص داده شده و در انتها در بخش پنجم به جمع‌بندی و نتیجه‌گیری مطالب ارائه شده پرداخته خواهد شد.

³ paste

⁴ print

¹ click

² scroll

در هنگام ثبت تمام رویدادها پارامترهایی به صورت کلی ذخیره می‌شوند که نشان‌دهنده اطلاعات مربوط به رویداد است. این پارامترها در جدول ۲ معرفی شده‌اند

جدول (۲): پارامترهای مخصوص هر رویداد ارسال شده از طریق افزونه

کد رویداد	توضیحات
آدرس URL صفحه	آدرس صفحه‌ای که رویداد در آن انجام شده است. به عنوان کلید مشخص کننده صفحه در نظر گرفته می‌شود
تاریخ	تاریخ وقوع رویداد
زمان	ساعت وقوع رویداد با دقت میلی‌ثانیه
کد بازدیدکننده	یک کد منحصر به فرد ($GUID^9$) به ازای هر کاربر است
کد نشست	یک کد منحصر به فرد که مشخص کننده نشست می‌باشد
کد نوع رویداد	مشخص کننده نوع رویداد اتفاق افتاده است

علاوه بر پارامترهای کلی که به ازای هر رویداد ارسال می‌شود، پارامترهایی نیز برای هر نوع رویداد وجود دارد که مقدار مشخصی را برای رویداد مورد نظر ارسال می‌کند. این پارامترها را در جدول ۳ ملاحظه می‌کنید.

جدول (۳): پارامترهای مخصوص هر رویداد ارسال شده از طریق

افزونه

نوع رویداد	پارامترها
بارگذاری محتوای صفحه	ندارد
مخفی شدن سربرگ	مدت زمانی که صفحه فعال بوده است (به ثانیه)
تمرکز و مشاهده سربرگ	ندارد
لغزش در طول صفحه	میزان لغزش صفحه (به درصد)
پرس و جوی کاربر	متن پرس و جو
باز شدن نتیجه جستجو	پارامتر اول: متن پرس و جو پارامتر دوم: رتبه نتیجه جستجو (عدد صحیح)
پرینت	ندارد
کلیک	ندارد
زمان اولین لغزش	مدت زمان از بارگذاری صفحه تا اولین مکان لغزش (به ثانیه)
زمان اولین کلیک	مدت زمان از بارگذاری صفحه تا اولین کلیک (به ثانیه)
زمان بسته شدن سربرگ	کل مدت زمانی که صفحه فعال بوده است (به ثانیه)
زمان اولین مشاهده	مدت زمان از بارگذاری صفحه تا اولین مشاهده صفحه (به ثانیه)
افزوده شدن به علاقه‌مندی‌ها	ندارد

افزونه در هنگام وقوع هر رویداد، یک درخواست به سمت سرور ارسال می‌کند. سرور با دریافت درخواست، اطلاعات مربوط به آن را ذخیره می‌کند. به همین منظور امکانی در سمت سرور ایجاد شده است که قابلیت دریافت و ذخیره رویدادها را داشته باشد. خلاصه‌ای از اطلاعات دریافتی توسط سرور و ذخیره شده در بانک اطلاعاتی در جدول ۴ ارائه گردیده است.

می‌کند می‌توان نتیجه گرفت که صفحه موردنظر محتوای مرتبط بیشتری در مورد پرس‌وجو برای ارائه به کاربر داشته است. بنابراین، چنین جایگزین‌هایی برای دریافت ضمنی بازخورد درخصوص مرتبط بودن نتایج ارائه شده می‌تواند بسیار دقیق‌تر از روش‌های مشابه باشد.

برای انجام اینگونه تحلیل‌ها، نیاز داریم تا رویدادهای ایجاد شده توسط کاربر را دریافت کرده و بر روی آن‌ها تحلیل‌های موردنظر را انجام دهیم. دریافت چنین رویدادهایی برای تمام صفحات و سربرگ‌های باز شده توسط کاربر، تنها از طریق خود مرورگر امکان‌پذیر است. زیرا صفحه مربوط به موتور جستجو، دسترسی لازم برای ثبت رویدادهای اتفاق افتاده در صفحه دیگر مربوط به نتایج جستجو را ندارد و تنها می‌تواند بخشی از رویدادهای داخل خود را ثبت کند که مد نظر ما نیست. مرورگر به تمامی سربرگ‌های باز شده توسط کاربر و صفحه‌های آن‌ها دسترسی دارد و می‌تواند وقایع آن‌ها را ثبت کند. بنابراین، برای جمع‌آوری اطلاعات مربوط به رفتار کاربران در حین جستجو یک افزونه برای مرورگر فایرفاکس توسعه داده شده است که بدون دخالت و توجه کاربر، می‌تواند رفتار وی را ثبت کند.

افزونه توسعه داده شده مبتنی بر رویداد عمل می‌کند. به این معنی که وقوع رویدادهای خاصی را به همراه زمان دقیق، مشخصات کاربر ایجاد کننده رویداد و پارامترهای رویداد ثبت می‌کند. رویدادهایی که توسط افزونه ثبت می‌شوند در جدول ۱ معرفی شده‌اند.

جدول (۱): رویدادهای قابل ثبت توسط افزونه

کد رویداد	عنوان رویداد	توضیحات
۱	بارگذاری ^۵ محتوای صفحه	پس از بارگذاری تمام محتوای صفحه رخ می‌دهد
۲	مخفی شدن ^۶ سربرگ	هنگامی که سربرگ به هر دلیلی از دید کاربر مخفی شود
۳	تمرکز ^۷ و مشاهده سربرگ	هنگامی که سربرگ دوباره توسط کاربر مشاهده شود
۴	لغزش در طول صفحه	در زمان لغزش در طول صفحه
۵	پرس و جوی ^۸ کاربر	هنگامی که کاربر توسط موتور جستجو یک عبارت را جستجو می‌کند
۶	باز شدن نتیجه جستجو	در زمان باز شدن نتیجه جستجو توسط کاربر
۷	پرینت	در زمان پرینت صفحه توسط کاربر
۸	کلیک	رویداد کلیک کردن بر روی صفحه
۹	زمان اولین لغزش (FirstScrollTime)	زمانی که اولین لغزش در صفحه اتفاق می‌افتد
۱۰	زمان اولین کلیک (FirstClickTime)	زمانی که اولین کلیک در صفحه اتفاق می‌افتد
۱۱	زمان بسته شدن سربرگ	در زمان بسته شدن سربرگ توسط کاربر
۱۲	زمان اولین مشاهده (FirstFocusTime)	زمانی که اولین بار کاربر صفحه را مشاهده می‌کند (به این معنی که سربرگ آن را باز می‌کند)
۱۳	افزوده شدن به علاقه‌مندی‌ها ^۹	رویداد ذخیره آدرس URL صفحه در لیست علاقه‌مندی‌ها

^۵ Load

^۶ Blur

^۷ Focus

^۸ Query

^۹ bookmark

^{۱۰} Globally Unique Identifier

جدول (۴) : پارامترهای دریافتی سرور

پارامترها	نوع رویداد
تاریخ رویداد	Date
ساعت رویداد	Time
میلی ثانیه زمان رویداد	Milliseconds
آدرس url صفحه که هنگام ارسال با استفاده از تابع <code>encodeURIComponent</code> کد شده است	Referurl
کد رویداد که مشخص کننده نوع رویداد است	eventId
مقدار پارامتر اول رویداد	Value
مقدار پارامتر دوم رویداد	value2
کد بازدید کننده	visitorId
کد نشست	sessionId

بعد از باز شدن صفحه، اولین مشاهده اتفاق افتاده است. مقدار خروجی این تابع نیز با در نظر گرفتن سقف نرمال شده است.

• زمان اولین لغزش:

این معیار به معنی فاصله زمانی بارگذاری صفحه تا زمانی که برای اولین بار لغزش اتفاق افتاده می‌باشد. این معیار به واحد ثانیه گزارش می‌شود و با در نظر گرفتن سقف نرمال شده است.

بیشترین مقداری که کاربر صفحه را پیمایش کرده است به درصد توسط این تابع برگردانده می‌شود. به عنوان نمونه عدد ۱۰۰ نشانگر این است که کاربر تا انتهای صفحه پیش رفته است. مقدار خروجی این تابع نیز به صورت عددی بین صفر تا یک نرمال می‌شود.

• تعداد دفعات لغزش:

تعداد دفعاتی که عملیات لغزش توسط کاربر اتفاق افتاده است. هر چه بیشتر کاربر صفحه را بالا و پایین رفته باشد، این تعداد عدد بزرگتری را نشان می‌دهد. مقدار خروجی آن با در نظر گرفتن یک سقف بین صفر تا یک نرمال شده است.

• چاپ شدن صفحه:

این تابع عدد صفر به معنی چاپ نشدن صفحه و عدد یک به معنی چاپ شدن صفحه توسط کاربر را برمی‌گرداند.

• افزوده شدن به علاقه‌مندی‌ها:

این تابع در صورت افزوده شدن صفحه به علاقه‌مندی‌های کاربر عدد ۱ و در غیر اینصورت عدد صفر را بر می‌گرداند.

نتایج معیارهای محاسبه شده توسط توابع فوق ترکیب شده و تحت عنوان یک عدد نهایی به نام میزان مرتبط بودن سند با نتیجه پرس‌وجو گزارش می‌شود. هر مقدار در یک وزن مشخص ضرب شده و مجموع آن‌ها مقدار امتیاز سند را نتیجه می‌دهد. البته لازم به ذکر است که در صورتی که سند به علاقه‌مندی‌های کاربر افزوده شده باشد و یا چاپ شده باشد، بدون در نظر گرفتن معیارهای فوق به عنوان سند مرتبط امتیاز ۱ می‌گیرد و در غیر اینصورت رابطه فوق محاسبه می‌شود. چرا که در صورت وقوع این رویدادها، با احتمال خوبی سند به طور کامل با پرس‌وجو مرتبط بوده است.

باید مقادیر وزن‌هایی که برای هر معیار در نظر گرفته شده است به درستی تنظیم شود. مجموع وزن‌ها یک است به این معنی که میزان مرتبط بودن سند و عدد نهایی آن در بازه $[0, 1]$ قرار می‌گیرد. بعد از محاسبه معیارها باید آنها را با وزن‌هایی با یکدیگر ترکیب نمود تا بتوان میزان مرتبط بودن را محاسبه نمود. از طرفی دیگر، امکان ارزیابی صریح، نتایج نیز برای دسته‌ای از کاربران فراهم شده است که در آن کاربر میزان مرتبط بودن سند با پرس‌وجو را به صورت صریح گزارش کرده است. میزان مرتبط بودن به صورت صریح در سه مقدار مرتبط، تقریباً مرتبط و نامرتبط از کاربر دریافت می‌گردد. سپس با استفاده از تکنیک یادگیری ماشینی، از داده‌های نتایج صریح برای بدست آوردن مقدار مناسبی برای وزن‌های معیارهای مختلف استفاده می‌گردد.

۴- تحلیل نتایج

امکان بازخورد صریح در صفحه نتایج موتور جستجو مطابق با شکل ۱ فراهم شده است. کاربر می‌تواند بعد از مشاهده نتایج نظر خود را درباره آن‌ها صریحاً

با استفاده از داده‌های مربوط به رویدادهایی که کاربر در حین جستجو ایجاد کرده است، می‌توان مرتبط بودن و یا نامرتبط بودن نتیجه جستجو را با توجه به نحوه رفتار کاربر محاسبه کرد. به همین منظور، معیارهای زیر برای هر نتیجه جستجوی مشاهده شده توسط کاربر محاسبه می‌شوند:

• میانگین میزان لغزش:

میزان لغزش به صورت عددی بین صفر تا صد به معنی درصدی از ارتفاع صفحه که اسکرول در آن قرار دارد گزارش می‌شود. این معیار میانگین لغزش-های گزارش شده را به معنی اینکه کاربر بیشتر در چه ارتفاعی از صفحه قرار داشته است، گزارش می‌کند.

این تابع با دریافت آدرس صفحه، کد بازدیدکننده و کد نشست، صفحه مورد نظر را که توسط بازدیدکننده مورد نظر در نشست مشخص شده باز شده است را یافته و میانگین لغزش آن را گزارش می‌کند. این سه پارامتر در تمام توابعی که در ادامه بررسی می‌شوند حضور دارند. این سه پارامتر می‌توانند به عنوان کلیدی برای مشخص کردن یک صفحه مورد نظر که توسط کاربری در یک نشست خاص باز شده تلقی شوند.

لازم به توضیح است که عدد خروجی این تابع به صورت نرمال شده بین صفر تا یک برگردانده می‌شود. این تابع با دریافت مقدار موردنظر و سقف مقدار ممکن، مقدار نرمال شده آن را برمی‌گرداند.

• تعداد کلیک‌های روی صفحه:

این تابع تعداد کلیک‌های انجام شده روی صفحه مورد نظر را برمی‌گرداند. با در نظر گرفتن یک سقف برای تعداد کلیک‌ها، این مقدار نیز نرمال شده و به صورت عددی بین صفر تا یک گزارش می‌شود.

• زمان اولین کلیک:

زمانی که کاربر اولین کلیک را روی صفحه انجام داده است توسط این تابع گزارش می‌شود. این تابع نیز با در نظر گرفتن یک سقف تمام مقادیر خروجی خود را نرمال می‌کند.

• مدت زمان فعال بودن صفحه:

زمان بارگذاری صفحه تا بسته شدن نشان‌دهنده مدت زمان فعال بودن صفحه می‌باشد. این تابع مجموع زمان‌هایی را که صفحه فعال بوده است برمی‌گرداند. مقدار خروجی این تابع نیز نرمال می‌شود.

• زمان اولین مشاهده:

زمانی که برای اولین بار محتوای صفحه توسط کاربر مشاهده شود. این مقدار به ثانیه است. به عنوان نمونه عدد ۳,۵ نشانگر این است که سه و نیم ثانیه

خبر وب تصویر آوا ترجمه نقشه (جدید) بیشتر >

پارسی جو

موتور جستجوی بومی

صفحه ۱ | ۸۲,۱۵۴ نتیجه (۰,۳۳ ثانیه)

تایخ جستجو

وب

خبر

تصویر

آوا

همه زمان ها >

همه فایل ها >

جستجوی اینترنت

فضا | عرفان

ارتقای موتور جستجوی بومی "پارسی جو" - آی سی تی پرس...
 ارتقای موتور جستجوی بومی پارسی جو - Ictpress - سرویس جستجوی تصویری با پوشش ده میلیون صفحه به نخستین موتور جستجوی بومی که با نام تجاری پارسی جو در حال سرویس دهی است افزوده شد. به گزارش
www.ictpress.ir/Default,fa-IR,ICTPress,Content,NewsDetail,Key...
 بیشتر

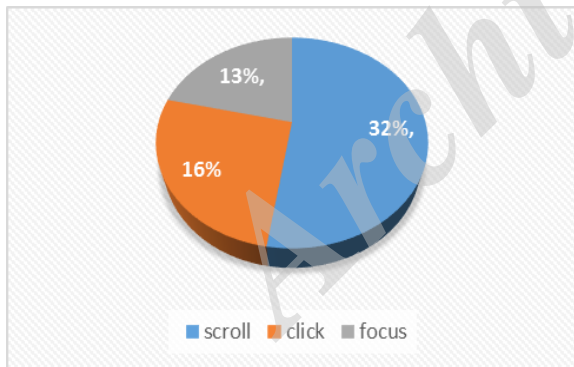
راه اندازی موتور جستجوی بومی در هفته دولت - Gerdab.ir | کرداب...
 موتور جستجوی بومی در هفته دولت در هفته دولت یکی از موتورهای جستجو افتتاح می شود و امیدواریم با حمایت بیشتری که صورت می گیرد قبل از پایان امسال بتواند تعداد مخاطبان بیشتری را جذب کند. به گزارش کرداب،
www.gerdab.ir/fa/news/11999
 بیشتر

همشهری آنلاین: جزئیات موتور جستجوی "پارسی جو" میزان استفاده ایرانی‌ها از گوگل...
 مهر نوشت: راه اندازی موتور جستجوی بومی از اواسط سال 89 در دستور کار وزارت فناوری اطلاعات قرار گرفت که در آن زمان طرحی مطرح شد که براساس آن موتور جستجوی ملی با
www.hamshahronline.ir/details/189899
 بیشتر

جزئیات موتور جستجوی "پارسی جو" - فرازو
 به گزارش خبرنگار مهر، راه اندازی موتور جستجوی بومی از اواسط سال 89 در دستور کار وزارت ارتباطات و فناوری اطلاعات قرار گرفت که در آن زمان طرحی مطرح شد که براساس آن موتور جستجوی ملی بر سبتر شبکه ملی
fararu.com/fa/news/131139
 بیشتر

شکل (۱): ارزیابی صریح نتایج جستجو از طریق افزونه

در شکل ۲ رویدادهای پرتکرار برگرفته از بازخوردهای کاربران ارائه گردیده است. همانگونه که شکل نشان می‌دهد، بیشترین فعالیت کاربران لغزش در صفحه نتایج است که ۳۲ درصد از کل رویدادها بوده است. همچنین کلیک کاربران با سهم ۱۶ درصدی از رویدادها بعد از لغزش بیشترین سهم فعالیت کاربران را به خود اختصاص داده است.



شکل (۲): رویدادهای پرتکرار حاصل از بازخوردهای کاربران

با استفاده از ترکیب وزن دار معیارهای فوق، میزان رضایتمندی از نتایج ارائه شده توسط موتور جستجوی پارسی جو، در مدت ۲ ماه مهر و آبان ۱۳۹۳، برابر با ۳۰٪ تخمین زده شده است.

۵- جمع بندی و نتیجه گیری

در این مقاله به تحلیل رضایت کاربران از موتور جستجوی پارسی جو مبتنی بر بازخوردهای ضمنی کاربران پرداخته شد. به این منظور رویدادهای به وقوع پیوسته توسط کاربران روی نتایج موتور جستجو از طریق یک افزونه توسعه

اعلام کند. در کنار هر نتیجه جستجو منویی ظاهر می‌شود که آیت‌های زیر را دارد:

- **مرتبط:** در صورتی که نتیجه موردنظر با پرس‌وجو کاملاً مرتبط بود، باید از این گزینه برای نتیجه ارزیابی استفاده شود.
 - **تقریباً مرتبط:** برای نتایجی که ارتباط کمتری با پرس‌وجو دارند این گزینه باید علامت زده شود.
 - **نامرتب:** اگر نتیجه نمایش داده شده هیچ ارتباطی با پرس‌وجوی موردنظر نداشت، باید این گزینه انتخاب شود.
- به عنوان نمونه در شکل ۱ منوهای ارزیابی و چند مورد ارزیابی شده از نتایج جستجو مشاهده می‌شود. با انتخاب هر مورد، منوی ارزیابی محو شده و با رنگ مناسب عنوان ارزیابی کاربر نمایش داده خواهد شد. نتایج هر بازخورد نیز به سمت سرور ارسال می‌شود.

لازم به ذکر است به منظور اخذ بازخوردهای کاربران ۵۰ نفر نیروی انسانی افزونه را روی مرورگر فایرفاکس نصب نموده‌اند و برای مدت ۲ ماه بخشی از نیازهای اطلاعاتی خود را از طریق جستجو بر روی موتور جستجوی پارسی جو برطرف نموده‌اند. همچنین، بازخورد صریح خود را نیز در این مدت از طریق افزونه، در اختیار سامانه قرار داده‌اند. نتایج تحلیل صورت گرفته روی آی‌پی‌های کاربران مشارکت کننده در ارزیابی نشان می‌دهد که این کاربران، برای جستجو از ۲۱۷ آی‌پی متفاوت به موتور جستجو متصل شده‌اند.

بعد از ارزیابی صریح نتایج وزن‌های هر معیار یک بار به صورت آزمایشی توسط نرم‌افزار وکا با داده‌های مربوط به بازخورد کاربران بدست آمده‌اند. برای استفاده از این نرم‌افزار داده‌ها در قالب مجموعه‌ای به شکل زیر استخراج شده‌اند.

AvgScrollAmountWeight, ClickCountWeight, FirstClickTimeWeight, FirstFocusTimeWeight, FirstScrollTimeWeight, MaxScrollAmountWeight, ScrollEventCountWeight, TotalActiveTimeWeight, IsPrinted, IsBookmarked, FeedbackType

وزن‌های استخراج شده در قالب یک فایل CSV به نرم‌افزار وارد می‌شوند. لازم به توضیح است که داده‌های مربوط به بازخورد صریح نیز به صورت یک ستون تعریف می‌شوند. خروجی وکا، وزن‌های زیر را برای معیارهای مختلف پیشنهاد نموده است:

- AvgScrollAmountWeight = 0.10*
- ClickCountWeight = 0.15*
- FirstClickTimeWeight = 0.05*
- FirstFocusTimeWeight = 0.05*
- FirstScrollTimeWeight = 0.05*
- MaxScrollAmountWeight = 0.15*
- ScrollEventCountWeight = 0.15*
- TotalActiveTimeWeight = 0.3*

وزن‌های به دست آمده برای معیارها، که از طریق تکنیک‌های یادگیری ماشین تخمین زده شده‌اند، در قالب پرسشنامه‌هایی برای متخصصین امر ارسال شد تا از صحت نسبی آنها اطمینان حاصل گردد.

وزن حاصل شده برای معیارها نشان‌دهنده اهمیت معیار لغزش و مدت زمان فعال بودن کاربر روی صفحه نتایج می‌باشد. که اهمیت این موارد پیش از این نیز در مقاله‌های مرتبط مورد تأیید قرار گرفته است.

engine. *Library and Information Science Research* 27(2) 203-221, 2005

- [9] Zachman, John A., "A Framework for Information Systems Architecture", *IBM Systems Journal*, Vol. 26, No. 3, 1987.
- [10] Silverstein, Craig, et al. "Analysis of a very large web search engine query log." *ACM SIGIR Forum*. Vol. 33. No. 1. ACM, 1999.
- [11] Park, Soyeon, Joon Ho Lee, and Hee Jin Bae. "End user searching: A Web log analysis of NAVER, a Korean Web search engine." *Library & Information Science Research* 27.2 : 203-221, 2005.
- [12] Costa, Miguel, and Mário J. Silva. "A search log analysis of a Portuguese web search engine." *Proc. of the 2nd INForum-Simpósio de Informática* : 525-536, 2010.
- [13] Spink, A., Ozmutlu, S., Ozmutlu, H.C., Jansen, B.J.: *U.S. versus European Web searching trends*. *SIGIR Forum* 36(2) 32-38, 2002.
- [14] T. Joachims. *Optimizing search engines using clickthrough data*. In *Proceedings of the ACM Conference on Knowledge Discovery and Data Mining*, 2002.
- [15] T. Joachims, L. Granka, B. Pang, H. Hembrooke, and G. Gay. *Accurately Interpreting Clickthrough Data as Implicit Feedback*, *Proceedings of the ACM Conference on Research and Development on Information Retrieval (SIGIR)*, 2005.
- [16] T. Joachims, L. Granka, B. Pan, H. Hembrooke, F. Radlinski, and G. Gay. *Evaluating the accuracy of implicit feedback from clicks and query reformulations in web search*. *ACM Transactions on Information Systems (TOIS)*, 25(2), 2007.
- [17] Steve Fox, Kuldeep Karnawat, Mark Mydland, Susan Dumais, and Thomas White, *Evaluating Implicit Measures to Improve Web Search*, *ACM Trans.on Info. Sys.*, 23(2):147-168, 2005.
- [18] HA Feild, J Allan, R Jones , *Predicting Searcher Frustration Proceedings of the 33rd international ACM*, 2010.
- [19] E Di Buccio, M Melucci, D Song, "Combining Interaction and Content for Feedback-Based Ranking", Springer, 2011
- [20] G Velayathan, S Yamada , *Investigating User Browsing Behavior - Web Intelligence and Intelligent* , 2007 - ieeexplore.ieee.org
- [21] E Natsheh, *Personalized Web Documents Filtering by Analyzing User Browsing Behaviors*, *International Journal of Information*, 2013.
- [22] Kelly, D., and Belkin, N. J. (2001). *Reading time, scrolling and interaction: Exploring implicit sources of user preferences for relevance feedback during interactive information retrieval*. In *Proceedings of the 24th Annual International Conference on Research and Development in Information Retrieval (SIGIR '01)*, USA, 408-409.

داده شده روی مرورگر فایرفاکس ثبت گردید. به منظور تخمین رضایت کاربران به هر رویداد مهم وزنی اختصاص داده شد تا سهم مناسبی در تخمین میزان مرتبط بودن نتایج داشته باشد. تحلیل نتایج حاصله نشان داد که مهم‌ترین فعالیت کاربران شامل لغزش و کلیک روی نتایج بوده است. همچنین میزان رضایت کاربران از نتایج پارسی‌جو ۳۰ درصد می‌باشد. در ادامه با افزایش گستره کاربران پارسی‌جو تحلیل‌های بیشتری روی نتایج حاصله از بازخوردهای ضمنی انجام خواهد شد. به عنوان مثال رضایت کاربران از موتور جستجو مبتنی بر موضوعات مختلف یکی از تحلیل‌های ارزشمندی است که به بهبود عملکرد موتور جستجو در حوزه‌های مختلف کمک شایانی خواهد نمود.

مراجع

- [۱] موسوی سبحان، عظیم زاده معصومه، محمودی مریم، یاری علیرضا، "ارائه چارچوبی جامع و کارا برای ارزیابی موتورهای جستجوی فارسی" هجدهمین کنفرانس ملی سالیانه انجمن کامپیوتر، تهران، اسفند ۹۱.
- [۲] عظیم زاده معصومه، سموری شهریار، یاری علیرضا، " بررسی و مقایسه کیفی موتورهای جستجو در حوزه وب فارسی" هجدهمین کنفرانس ملی سالیانه انجمن کامپیوتر، تهران، اسفند ۹۱.
- [۳] محمد صادق زاهدی، معصومه عظیم‌زاده، مریم محمودی، رضا بدیع، "ذاتقه سنجی و تحلیل رفتار کاربران سیار در موتور جستجوی بومی مبتنی بر تحلیل فایل ثبت وقایع"، چهارمین کنفرانس بین‌المللی مدیریت فناوری اطلاعات، ارتباطات و کامپیوتر، تهران، ۹۳.
- [۴] محمد صادق زاهدی، معصومه عظیم‌زاده، نوید فرهادی، علی محمد زارع بدیکی، "تحلیل رفتار کاربران فارسی زبان در موتور جستجوی بومی" نوزدهمین کنفرانس ملی سالانه انجمن کامپیوتر ایران، اسفند ۹۲.
- [5] Reza Badie, Masomeh Azimzadeh, Mohammad Sadegh Zahedi, "Automatic Evaluation of Search Engines", *Seventh International Symposium on Telecommunications (IST2014)*, September 09-11, 2014 .
- [6] Maryam Mahmoudy, Mohammad Sadegh zahedi, Masomeh Azimzadeh, "Evaluating the retrieval effectiveness of search engines using Persian navigational queries", *Seventh International Symposium on Telecommunications (IST2014)*, September 09-11, 2014.
- [7] Markey, K.: *Twenty-five years of end-user searching, Part 1: Research findings*. *American Society for Information Science and Technology* 58(8) 1071-1081, 2007
- [8] Park, S., Ho Lee, J., Jin Bae, H.: *End user searching: A Web log analysis of NAVER, a Korean Web search*