

کشف قوانین بر روی پرس وجوهای ساخت یافته ی وب معنایی جهت یاری کاربران انسانی در نگارش پرس وجوهای SPARQL

سعید شفیعی حسن آبادی¹، محسن کاهانی²، احسان عسگریان³
آزمایشگاه فناوری وب، گروه مهندسی کامپیوتر، دانشگاه فردوسی مشهد

shafiee.sa@stu-mail.um.ac.ir¹
kahani@um.ac.ir²
ehsan.asgarian@stu-mail.um.ac.ir³

چکیده

یکی از زبان های رایج تولید پرس وجوهای ساخت یافته در وب معنایی SPARQL است. تولید این پرس وجوها، جهت استفاده از داده های وب معنایی با قالب RDF (چهارچوب توصیف منابع)، برای افرادی که با هستان شناس تعریف شده بر بستر مجموعه داده ها آشنایی ندارند، می تواند فرایندی پیچیده باشد. در این بین وجود الگوهای مشخص (از پیش تعریف شده) برای چگونگی ساخت پرس وجوهای SPARQL، کاربران انسانی تازه کار را با مشکل مواجه نموده است. در این پژوهش با استفاده از الگوریتم های یادگیری ماشین، کاوش قوانین ترتیبی و تعریف قالب های ثابتی بر روی الگوهای سه گانه در بدنه ی پرس وجوهای SPARQL ارسال شده از سوی کاربران، می توان به پیش بینی و تخمین اجزای آتی بکارفته در چینش پرس وجوی های آتی پرداخت تا به کاربران آتی وب معنایی در نگارش پرس وجوهای ساخت یافته کمک نمود. دقت های بالای به دست آمده از ارزیابی این قوانین، بر روی مجموعه داده ی معتبر پرس وجوهای پایانه ی DBpedia در سال 2011 نشان از کارایی مناسب این قوانین دارد.

کلمات کلیدی

پرس وجو، SPARQL، قوانین، الگوهای سه گانه، Log، مجموعه داده.

1- مقدمه

شده است. دلیل انتخاب قوانین ترتیبی، فرایند تدریجی پر نمودن الگوهای سه گانه در پرس وجوهای SPARQL است.

با تمرکز بر جنبه ی استفاده ی مستقیم کاربران انسانی از قوانین قابل کشف بر روی زبان SPARQL، سیستم های یاری رسان و مطالعات پیشین انجام گرفته بر روی لاگ پرس وجوهای SPARQL، به عنوان دو جنبه ی کلیدی در این مسیر انتخاب گشت تا بر این اساس، مدلی از قالب های ثابت سودمند (با مطالعه ی لاگ پرس وجوهای SPARQL) معرفی شود تا بتوان بر اساس قوانین استخراجی از پردازش لاگ در پیاده سازی سیستم های آتی (سیستم های یاری رسان) کمک گرفت.

در توصیف بخش های این مقاله، در بخش دوم به بیان پیشینه ای از مطالعات انجام گرفته در زمینه های لاگ پرس وجوهای SPARQL و سیستم های یاری گر کاربر و بررسی قوانین انجمنی بر روی داده های RDF پرداخته می شود. در بخش سوم به تعاریف اولیه مفاهیم مورد نیاز برای درک و بیان راهکار پیشنهادی پرداخته و در بخش چهارم به

در وب معنایی، تولید پرس وجوهای ساخت یافته ی مبتنی بر SPARQL، به دلیل نیاز به آشنایی با مفاهیم تعریف شده در هستان- شناس (تعریف شده بر بستر مجموعه داده) و اطلاع از ساختار زبان SPARQL می تواند فرایندی پیچیده باشد. از سویی در سمت منابع دربرگیرنده ی داده های وب معنایی، تنها داده ها تعریف گشته اند و در این بین الگویی برای چگونگی ساخت پرس وجوهای SPARQL، برای کاربران انسانی تازه کار وجود ندارد. کسب روابط و الگوهای نهان در الگوهای سه گانه موجود در پرس وجوهای SPARQL می تواند به نگارش این پرس وجوها از سوی کاربران جدید انسانی کمک کند. لذا در این مقاله، با استفاده از الگوریتم های یادگیری ماشین، گونه ای از قوانین انجمنی (قوانین ترتیبی)، به بررسی این دسته از قوانین و روابط قابل کشف، در الگوهای سه گانه در پرس وجوهای SPARQL پرداخته