

استخراج دانش از پایگاه های داده تحت وب عمیق با استفاده از

Correlation Mining

شهره آجودانیان¹؛ محمد داورپناه جزی²؛ محمد حسین سرایی³

چکیده

اخیرا شاهد رشد سریع پایگاه‌های داده تحت وب هستیم. به منظور مدل سازی و یکپارچه سازی پایگاه‌های داده تحت وب اولین مسئله که باید به آن پرداخته شود این است که واسط پرس و جو چگونه عمل می‌کند یا یک منبع چه قابلیت های پرس و جویی فراهم می‌کند. رویکردی که استفاده می شود بر پایه این اصل است که چارچوب کلی واسط‌های پرس و جو شبیه یکدیگر هستند. پس از استخراج دانش از پایگاه های داده تحت وب با استفاده از روش پارسینگ واسط پرس و جو، از الگوریتم‌های داده کاوی به منظور تطبیق تعداد زیادی شمای دریک زمان در پایگاه‌های داده استفاده می‌شود که استفاده از این الگوریتم‌ها سرعت تطبیق اطلاعات را بالا می برد. در مورد تطبیق معانی و دانش بدست آمده از پایگاه‌های داده، از رویکرد correlation mining و الگوهای co-occurrence استفاده شده است. روش‌های تطبیق شمای موجود معمولا تطبیق یک به یک انجام می‌دهند و صفات مطابق را پیدا می‌کنند. با استفاده از correlation mining قادریم تطبیق شمای m:n انجام دهیم.

در این مقاله یک سیستم استخراج و تطبیق دانش در پایگاه های داده تحت وب عمیق ارائه شده است. این سیستم به طور اتوماتیک وظایف خود را در دو گام انجام می دهد. کاربرد به عنوان ورودی، پرس و جوی خود را که در دامنه خاصی می‌باشد به سیستم می‌دهد. در گام اول دانش موجود در واسط های پرس و جو با استفاده از روش پارسینگ صفحه وب در قالب HTML مورد جستجو، استخراج می گردد. اطلاعات بدست آمده، به عنوان ورودی به گام دوم که وظیفه اش تطبیق صفات است، فرستاده می‌شود و در این گام با استفاده از الگوریتم داده کاوی تهیه شده که بر مبنای رویکرد correlation mining می‌باشد تطبیق‌های پیچیده بین صفات را انجام می‌دهد.

کلمات کلیدی

استخراج دانش، پایگاه‌های داده تحت وب عمیق، پارسینگ واسط پرس و جو، تطبیق معانی با استفاده از الگوریتم‌های داده کاوی

-
1. شهره آجودانیان؛ دانشجوی کارشناسی ارشد دانشگاه آزاد اسلامی واحد نجف آباد؛ shajoudanian@gmail.com
 2. دکتر محمد داورپناه جزی؛ عضو هیئت علمی دانشکده برق و کامپیوتر دانشگاه صنعتی اصفهان؛ استاد مدعو دانشگاه آزاد اسلامی واحد نجف آباد؛ mdjazi@cc.iut.ac.ir
 3. دکتر محمد حسین سرایی؛ عضو هیئت علمی دانشکده برق و کامپیوتر دانشگاه صنعتی اصفهان؛ استاد مدعو دانشگاه آزاد اسلامی واحد نجف آباد؛ sarabee@cc.iut.ac.ir