

عنوان مقاله:

ارائه الگوریتمی جدید برای یادگیری قواعد استخراج اطلاعات از متن

محل انتشار:

دومین کنفرانس ملی علوم مدیریت نوین و برنامه ریزی پایدار ایران (سال: 1395)

تعداد صفحات اصل مقاله: 14

نویسندگان:

فاطمه حیدری کیا - دانشجوی کارشناسی ارشد نرم افزار دانشگاه آزاد اسلامی قزوین

مریم حیدری کیا - دکتری مدیریت دولتی منابع انسانی

خلاصه مقاله:

با رشد سریع حجم متون قابل دسترس به خصوص در شبکه ی جهانی اینترنت، کار استخراج اطلاعات از متن اهمیتتروز افزون یافته است . عملیات اصلی سیستم های استخراج اطلاعات از دو مرحله ساخت پایگاه دانش مورد نیاز برای استخراجاطلاعات و استفاده از پایگاه دانش برای استخراج اطلاعات از متون ورودی تشکیل شده است . در این مقاله برای تولید خودکارپایگاه دانش مورد نیاز الگوریتمی ارائه شده است که می تواند قواعد لازم برای استخراج اطلاعات از متون نیمه ساختارمند (مانندصفحات) وب طول یک فرآیند یادگیری بانظارت تولید کند . در این الگوریتم که یک الگوریتم استقرای قاعده است، از یکقاعده عمومی شروع می شود و بر اساس ورودی و خروجی هایی که درمثال های آموزشی وجود دارد و نیز معیارهای ارزیابیقواعد، آن قاعده و ATP اختصاصی تر و دقیق تر می شود . تعریف معیارهای جدید برای مقایسه و ارزیابی قواعد استخراجاطلاعات، تعریف AATP پس پردازش نهایی برای بهتر کردن قاعده و استفاده از یادگیری به روش تعاملی با کاربر برایبیدست آوردن نتیجه بهتر از دیگر موارد طراحی شده در این الگوریتم است . هر قاعده استخراج اطلاعات به صورت الگویی استکه باید در متن جستجو شود و در صورت تطبیق با قسمتی از متن، اطلاعات مورد نیاز بر اساس پارامترهای تعریف شده در الگواستخراج یابد . برای ارزیابی این الگوریتم، نتایج به دست آمده بر یکی از سیستم های مثال های آزمایشی با نتایج سیستم WHISK (یکی از سیستم های موفق در استخراج اطلاعات) مقایسه شده است که بهبود قابل توجهی را در هر دو معیارRecall و Precision نشان می دهد.

کلمات کلیدی:

استخراج اطلاعات از متن، متن کاوی ، وب کاوی ، یادگیری قواعد، استقرای قاعده، سیستم های استخراج اطلاعات از متن

لینک ثابت مقاله در پایگاه سیویلیکا:

<https://civilica.com/doc/511395>

