

بازبینی

Application of text mining in the biomedical domain

داریوش حسن پور آده

۹۳۰۸۱۶۴

این مقاله بیشتر به عنوان یک مروری^۱ بر روش‌های استخراج متن^۲ از متون علمی پزشکی می‌باشد. بنابراین نوآوری یا بهبود خاصی در مورد روش‌های موجود برای استخراج متون ارائه نداده است، همچنین راه حلی در رابطه با مشکل خاصی نیز ارائه نداده است؛ بنابراین این بازبینی دارای قسمت‌های الزامی مطرح شده در رابطه با بازبینی‌ها نمی‌باشد. در این نوشتار سعی من بر انتقال خلاصه‌ای از متن مقاله می‌باشد.

در ابتدا مقاله سعی کرده که نشان دهد چرا استخراج متن از متون علمی پزشکی دارای اهمیت زیادی است. سپس یک روال کلی از مراحل مرسوم در استخراج متون آورده است. اولین گام که در استخراج متن باید صورت بگیرد بازیابی منابع متنی^۳ مربوطه با توجه به موضوع مورد علاقه می‌باشد. که این فرایند با استفاده از موتورهای جستجوی پایگاه‌های علمی با استفاده از تعدادی کلمات کلیدی انجام می‌پذیرد. سپس مرحله بعد شامل شناسایی متون موجودیت‌های نامی^۴ در متون بدست آمده از مرحله قبل است. که در این مرحله به دنبال کلمات مرتبط با کلمات کلیدی و ارتباط آنها با یک دیگر هستیم - که شامل کلمات مترادف (معادل) با کلمات کلیدی می‌باشد. مرحله بعد شامل استخراج اطلاعات^۵ می‌باشد. که هدف از این مرحله بدست آوردن ارتباطات بین مفاهیم استخراج شده از متون می‌باشد که در روش برای این مرحله معرفی شده است - Co-occurrence-based methods و NLP-based methods. بعد از مرحله‌ی استخراج اطلاعات به مرحله‌ی کشف دانش^۶ می‌رسیم و که آورده که به سیستم‌های استخراج متن بهتر است با دید استخراج اطلاعات نگاه شود و باید این سیستم بتواند روابط مخفی موجود در متون بین کلمات کلیدی را بر اساس حقایق شناخته شده آشکار کند و در اینجا نیز روشی به نام The ABC principle جهت استخراج اطلاعات آورده است.

در قسمت دیگر مقاله نیز کاربردهایی از استخراج متون برای متون عملی پزشکی آورده است، که ۳ مورد از آنها به طور خلاصه در زیر شرح داده می‌شود.

۱، ۲. Drug repositioning، Drug-target discovery: آورده است که استخراج متن به طور گسترده در جستجو برای اهداف دارویی جدید و دارو استفاده شده است. و فرق این با Drug repositioning در این است که در Drug repositioning اطلاعات در مورد یک داروی شناخته شده و اثرات آن برای جستجوی یک کاربرد جدید مورد استفاده واقع می‌شود.

۳. Adverse events: در حالی که برای کشف دارو و اهداف دارویی^۷ یکی از کاربردهای عمده‌ی استخراج متن در متون پزشکی است، از دیگر کاربردهای مهم دیگر استخراج عوارض جانبی داروها می‌باشد. البته به جز کاربردهای بالا ۳ کاربرد دیگر نیز آورده است که بنده درک درستی از آنها بدست نیاوردم و همچنین در هر کدام از کاربردها، کارهای انجام شده را نیز آورده است که از لیست کردن آنها در این نوشتار اجتناب می‌کنیم.

همانطور که در ابتدای نوشتار آورده‌ام این مقاله بیشتر یک مرور بوده برای همین جای زیادی جهت بازبینی و ارزیابی ندارد.

^۱ Survey

^۲ Text Mining

^۳ Information Retrieval

^۴ Named Entity Recognition

^۵ Information Extracion

^۶ Knowledge discovery

^۷ drug-target