

عنوان مقاله:

ارزیابی میزان تطابق راهکارهای مدل سازی موضوعی بر پایگاه های داده تحت وب گراف محور متن کوتاه پویا

محل انتشار:

هفتمین کنفرانس بین المللی وب پژوهی (سال: 1400)

تعداد صفحات اصل مقاله: 9

نویسندگان:

محدثه طاهرپرور - دانشجوی دکتری، گروه مهندسی کامپیوتر و فناوری اطلاعات، دانشکده فنی و مهندسی، دانشگاه آزاد اسلامی واحد رشت

فاطمه احمدی آبکناری - استادیار، گروه مهندسی کامپیوتر و فناوری اطلاعات، دانشکده فنی و مهندسی، دانشگاه پیام نور مرکز رشت

پیمان بیات - استادیار، گروه مهندسی کامپیوتر و فناوری اطلاعات، دانشکده فنی و مهندسی، دانشگاه آزاد اسلامی واحد رشت

خلاصه مقاله:

مدلسازی موضوعی یک ابزار تحلیلی محبوب برای استخراج موضوع از داده های متنی و خوشه بندی داده های پیکره های متنی است. روشهای زیادی برای مدلسازی موضوعی وجود دارد که انواع روابط و محدودیت ها را در انواع مجموعه داده ها در نظر میگیرند. بسیاری از پژوهشگران به روش مدل سازی تحلیل پنهان دریکله^۱ به دلیل انعطاف پذیری و سازگاری آن علاقه مند هستند. اما انتخاب این روش در خصوص مجموعه داده های پیچیده و خاص با چالش های بسیار همراه است. نظر به گسترش شبکه های اجتماعی و وجود پایگاه های داده پویا و متن کوتاه، بررسی امکان پذیر بودن استفاده از بهترین روش مدل سازی موضوعی بر اساس معیارهای ارزیابی همچون انسجام موضوع^۲، زمان اجرای مدل، انحصارطلبی^۳ و میزان حیرت^۴ مدل هدف پژوهش حاضر است. در این مقاله، رویکردی های مختلفی از روش های مدلسازی موضوعی در خصوص مجموعه داده متنی کوتاه پویا مورد تجزیه و تحلیل قرار گرفته است. مجموعه داده متن کوتاه پویا میتواند کاربردهای متنوعی داشته باشد به عنوان مثال، مجموعه داده های مربوط به موضوع مقالات، مجموعه داده های گردآوری شده از رسانه های اجتماعی، مجموعه داده های نظرات کاربران در خصوص محصول جدید ارائه شده توسط یک شرکت تجاری و موارد دیگر. با توجه به گراف محور بودن پایگاه داده مورد استفاده در این پژوهش، موضوع بدست آمده از خروجی اعمال روش های مدلسازی موضوعی، کمک شایانی در مساله چالش برانگیز تشخیص جوامع در حوزه تحلیل گراف می کند. تشخیص مناسب جوامع میتواند در یافتن گره های تاثیرگذار مناسب در بازاریابی ویروسی موثر باشد. نتایج بدست آمده از بررسی انواع روش های مدل سازی موضوعی بر پایگاه داده DBLP و نوع گره موضوع مقاله و ارزیابی نتایج با معیارهای ارزیابی موضوعی نشان از پایداری و تطابق روش بایترم بر روی این پایگاه داده دارد.

کلمات کلیدی:

تحلیل محتوا، تشخیص جامعه، مدل سازی موضوعی، متن کاوی، یادگیری ماشینی.

لینک ثابت مقاله در پایگاه سیویلیکا:

<https://civilica.com/doc/1236894>

