



دانشگاه کاشان
University of Kashan

مجله محاسبات نرم

SOFT COMPUTING JOURNAL

تارنمای مجله: scj.kashanu.ac.ir



تشخیص سرقت ادبی در متون علمی مبتنی بر بلوک‌بندی متن و معیار مشابهت کسینوسی[✦]

نگار مجمع[✧]، استادیار، سارا باشتین^۱، دانشجوی کارشناسی ارشد

^۱ گروه مهندسی کامپیوتر، موسسه آموزش عالی نقش جهان، اصفهان، ایران.

اطلاعات مقاله

چکیده

تاریخچه مقاله:

دریافت ۰۶ دی ماه ۱۴۰۰

پذیرش ۰۴ آذر ماه ۱۴۰۱

کلمات کلیدی:

سرقت علمی ادبی

تشخیص اصالت متون علمی

فاصله کسینوسی

بلوک‌بندی متن

در دهه اخیر با گسترش دسترسی به شبکه جهانی اینترنت، سرعت و سهولت در دسترسی به ایده‌ها، مستندات، مقالات، دست نوشته‌ها و داده‌های جمع‌آوری شده توسط دیگران افزایش یافته است. این موضوع باعث شده است که تبادل اطلاعات و افکار بین محققین و تولیدکنندگان علوم آسان‌تر شود، اما در مقابل باعث آسان شدن اعمال رونوشت غیرمجاز، خلاصه نویسی بدون ذکر منبع و در کل سرقت متون ادبی شده است. از آنجایی که دانشگاه‌ها و مراکز آموزشی، منابع علمی و پژوهشی را با سهولت در دسترس اغلب کاربران قرار می‌دهند، تشخیص میزان اصالت متون علمی در این مراکز مهم‌تر و بالطبع آن از حساسیت بیشتری برخوردار است. در این پژوهش روشی ارائه شده تا با استفاده از بلاک‌بندی قطعات اسناد، مقایسه بین قطعات مرتبط انجام شود. در روش پیشنهادی پس از دسته‌بندی اسناد به دو دسته اسناد اصلی و اسناد مشکوک، پیش‌پردازشی با هدف حذف ایست و واژه‌ها و جمله‌بندی جدید صورت پذیرفته است. سپس اسناد قطعه‌بندی شده و با استفاده از شباهت کسینوسی، میزان شباهت متون با یکدیگر تعیین شده است. روش پیشنهادی در آزمون ۵۰ سند موجود در مجموعه داده‌ها، دقت ۹۴ درصدی را کسب کرده که به نسبت به یکی از روش‌های مشابه بهبود ۲ درصدی داشته است.

© ۱۴۰۱ - مجله محاسبات نرم، کلیه حقوق محفوظ است.

۱. مقدمه

محلّی دیگر، موجب شده است تا دانشجویان و دانش‌آموزان به رونوشت غیرمجاز از ایده‌ها و اطلاعات دیگران ترغیب شوند. اغلب موارد رونوشت از آثار ادبی در دانشگاه‌ها است که اسناد به طور معمول به صورت مقاله و گزارش یافت می‌شوند. با این حال، رونوشت از آثار دیگران می‌تواند به صورت نسبی در هر زمینه‌ای شامل مقالات علمی، طراحی‌های هنری و کدهای منبع صورت گیرد [۲]. سرقت علمی می‌تواند در چندین شکل ظاهر شود، استفاده از متن، یک رابطه ریاضی، شکل یا نمودار و غیره بدون ارجاع به صاحب اثر نمونه‌های بارز سرقت هستند. از میان انواع نمونه‌های سرقت ادبی، استفاده از قسمت یا کل یک سند متنی بیشترین میزان رخداد را به خود اختصاص داده است.

سرقت ادبی، تلقی ایده، نظرات و افکار دیگران به نام خود و یا استفاده از آنها بدون ارجاع به منبع اصلی است [۱]. افراد با جابه‌جایی کلمات، جملات و یا عبارات و جایگزینی معانی آنها و همچنین دوباره‌نویسی^۱ متن، سعی در پنهان نمودن موارد رونوشت شده از کار دیگران در کار خود را دارند. امروزه، دسترسی آسان به منابع اطلاعاتی، از طریق وب یا شبکه‌های

✦ نوع مقاله: پژوهشی

* نویسنده مسئول

پست(های) الکترونیک: majma@naghshejahan.ac.ir (مجمع)

sbashtin4@gmail.com (باشتین)

¹ Rewriting