فصلنامه علمی-پژوهشی
سیاست علم و فناوری

سال دوم، شماره ۲، تابستان ۱۳۸۸

روشی نوین برای دسته‌بندی هوشمند متون علمی (مطالعه موردی مقالات فناوری نانو متخصصان ایران)

بابک تیمورپور[۱]، محمد مهدی سپهری[۲*]، لیلا پزشک[۳]

۱- دانشجوی دکتری مهندسی صنایع، دانشگاه تربیت مدرس، تهران، ایران

۲- دانشیار، مهندسی صنایع، دانشگاه تربیت مدرس، تهران، ایران

۳- دانشجوی کارشناسی ارشد مدیریت تکنولوژی، دانشگاه علوم و تحقیقات، تهران، ایران

چکیده

مقالات نمایه شده در مؤسسه اطلاعات علمی[۱] (ISI) یکی از معتبرترین شاخص‌های سنجش علم و فناوری می‌باشد که طبقه‌بندی موضوعی آن‌ها یکی از چالش‌های بزرگ مدیریت فناوری است. در این مقاله سعی شده است با استفاده از یک روش نوین متن‌کاوی[۲] به نام SUTC، مقالات متخصصان ایرانی در حوزه فناوری‌نانو که در مجلات ISI نمایه شده‌اند دسته‌بندی شوند. این دسته‌بندی می‌تواند معیاری مناسب برای سیاست‌گذاران در شناخت توانمندی‌های کشور در زمینه‌های مختلف تحقیقاتی فناوری‌نانو قرار گیرد. در این راستا، ابتدا استانداردهای معتبر در فناوری نانو با یکدیگر ادغام شده و طبقه‌بندی جامعی برای نانومواد حاصل گردیده است. سپس، با استفاده از روش‌های بازیابی اطلاعات[۳] و متن‌کاوی، مقالات بدون دانش پیشین از برچسب دسته‌ها به طور هوشمند دسته‌بندی شده‌اند. به منظور ارزیابی روش طراحی شده، دسته‌بندی هوشمند مقالات با دسته‌بندی مقالات به‌وسیله خبرگان حوزه نانو مقایسه شده است. نتایج، حاکی از صحت مناسب روش ارائه شده است.

کلیدواژه‌ها: علم‌سنجی، فناوری نانو، متن‌کاوی، طبقه‌بندی متون، خوشه‌بندی، معیار سیلوئت

پژوهشگاه علوم انسانی و مطالعات فرهنگی
پرتال جامع علوم انسانی

۱- مقدمه

توسعه فناوری‌های نوین و دستیابی به اقتصاد دانایی محور در سال‌های اخیر مورد توجه سیاست‌گذاران و مسئولان کشور بوده و جزو اهداف برنامه چهارم توسعه قرار گرفته است [۱]. یکی از معیارهای ارزیابی فعالیت‌های مراکز علمی و پژوهشی، مقالات و متون علمی منتشر شده از سوی این مراکز است [۲، ۳]. علاوه بر رشد ۱۰۰۰ درصدی تولید مقالات به‌وسیله متخصصان ایرانی که در دوره‌ای ده ساله (۲۰۰۴-۱۹۹۵) اتفاق افتاده [٤] و به تأیید مراجعی نظیر Nature (۲۰۰۵) رسیده است [۵]، آمار منتشره از سوی ستاد ویژه توسعه فناوری‌نانو ـ دفتر همکاری‌های فناوری ریاست جمهوری ایران – حاکی از افزایش تعداد مقالات ISI منتشره به‌وسیله متخصصان ایرانی در این حوزه از ۹ مقاله در سال ۲۰۰۰ به ۷۷۰ مقاله در سال ۲۰۰۸ است [٦]. فناوری نانو، فناوری ساخت و کنترل فرایندها در ابعاد ملکولی است [۷]. رشد و توسعه این فناوری در برنامه‌های سوم و چهارم توسعه در قالب سند راهبرد آینده توسعه فناوری‌نانو در اولویت دولت قرار گرفته است. بررسی حوزه فعالیت و تعیین جهت‌گیری دستاوردهای علمی و فناورانه مراکز علمی و

* نویسنده عهده‌دار مکاتبات: mehdi.sepehri@modares.ac.ir

1. Institution of Scientific Information
2. Text Mining
3. Information Retrieval

پژوهشی کشور در حوزه نانو یکی از برنامه‌های سند راهبرد آینده توسعه فناوری‌نانو است که منجر به ایجاد دانشی عمیق در رابطه با زمینه کاری متخصصان، مراکز پژوهشی و در نتیجه جامعه علمی کشور در این حوزه خواهد شد[۸].

یکی از کاربردی‌ترین روش‌های ارزیابی توسعه علم و فناوری تجزیه و تحلیل متون منتشره در این حوزه، با استفاده از روش‌های علم‌سنجی[1] است. روش‌های متعارف علم‌سنجی شامل تحلیل استنادی[2] و متن‌کاوی می‌باشند [۹،۱۰]. در سال‌های اخیر مطالعات جامعی در حوزه متن‌کاوی برای ایجاد نگاهی کلی به ساختار فنی و زیرساخت‌های ادبیات پژوهشی علم و فناوری نانو انجام شده است که در بخش ۳، مروری بر مطالعات گذشته، به آن می‌پردازیم.

در تحقیق فعلی روش جدیدی ارائه شده که از مفاهیم مطرح در بازیابی اطلاعات و خوشه‌بندی برای دسته‌بندی[3] متون بر اساس درخت نانو استفاده می‌کند. این درخت از تلفیق استانداردهای مطرح موجود جهت دسته‌بندی سلسله‌مراتبی فناوری نانو برای شناخت زمینه‌های مختلف تحقیقاتی فناوری نانو به دست آمده است.

نتایج این روش نویدبخش بوده و افق جدیدی در دسته‌بندی بدون ناظر مطرح می‌کند. همچنین ارزیابی و دسته‌بندی مقالات فناوری نانو به‌وسیله این روش راه‌گشای دسته‌بندی آتی متخصصان و مراکز پژوهشی فعال نانو در ایران خواهد بود.

هدف اصلی این پژوهش دسته‌بندی مقالات بدون دانش پیشین از برچسب دسته‌ها می‌باشد. نتایج دسته‌بندی قابل استفاده برای شناخت حوزه‌های فعال نانو در ایران، حوزه کاری پژوهشگران نانو و حوزه کاری مراکز پژوهشی نانو می‌باشد. این دسته‌بندی در حوزه‌های مختلف فناوری نانو راهکاری قابل اتکا برای ستاد ویژه توسعه فناوری نانو به عنوان متولی این فناوری در کشور در تکمیل سیاست‌های افقی اتخاذ شده قبلی نظیر حمایت تشویقی از محققان با سیاست‌های عمودی جدید نظیر حمایت از مراکز برتر در هر حوزه و در راستای توانمندی‌های کشور است. در حال حاضر این کاربردها در حال تحلیل و انجام می‌باشند.

۲- متن کاوی

متن‌کاوی شاخه‌ای از داده‌کاوی[4] یا همان کشف دانش است. از نظر فییاد[5] و همکارانش (۱۹۹۶)، کشف دانش، فرایند غیربدیهی تشخیص الگوهای معتبر، نو، مفید و در نهایت قابل درک در داده‌هاست [۱۱]. فرق داده‌کاوی با متن‌کاوی این است که در متن‌کاوی الگوها از متن زبان طبیعی استخراج می‌شوند درحالی که داده‌کاوی الگوها را از پایگاه‌های داده ساخت‌یافته به دست می‌آورد. متن‌کاوی، متصل کردن اطلاعات استخراج شده به یکدیگر برای تشکیل حقایق یا فرضیه‌های جدید است تا پس از آن به کمک روش‌های متعارف آزمایش، بررسی بیش‌تری شوند. هدف متن‌کاوی، کشف اطلاعات از قبل ناشناخته می‌باشد که هنوز کسی نمی‌داند و بنابراین مستند نشده است. متن‌کاوی کاربردهای متعددی از جمله بررسی روندهای علمی در علم بیولوژی و هوش تجاری دارد. دوررى[6] و همکارانش براین باورند از آنجا که ۹۰٪ از داده‌های شرکت‌ها و سازمان‌ها با روش‌های متعارف داده‌کاوی قابل کشف دانش نیستند، انگیزه زیادی برای استفاده از متن‌کاوی در صنعت وجود دارد [۱۲]. روش‌های متن‌کاوی اغلب همان روش‌های داده‌کاوی هستند که با تغییراتی برای متون استفاده می‌شوند.

۳- مروری بر مطالعات گذشته

از زمان آغاز توجه به فناوری نانو در ایران که در انتها و ابتدای سال‌های ۱۳۷۹ و ۱۳۸۰ به وقوع پیوست و همت دولت به تأسیس ستاد ویژه توسعه فناوری نانو (INIC[7]) در تابستان ۱۳۸۲، فعالیت‌های پژوهشی و اجرایی فراوانی به منصه ظهور رسیده‌اند. برای شناخت کلی از این فعالیت‌ها و ورودی به مطالعات گذشته مورد نظر این پژوهش ضرورت دارد به آخرین مقالاتی که به‌تازگی در خصوص وضعیت فناوری نانو در ایران منتشر شده‌اند اشاره کرد. قاضی نوری و حیدری [۱۳] و سرکار و بیت اللهی [۱٤] نویسندگان این مقالات که به ترتیب در سال‌های ۲۰۰۸ و ۲۰۰۹ منتشر شده به موضوع فناوری نانو در کشور پرداخته‌اند. در مقاله اول نویسندگان به موضوع آثار بالقوه توسعه فناوری‌نانو در ایران با نگرشی سیستمی و راهبردی توجه کرده و در مقاله دوم مروری کلی و فشرده بر فعالیت‌های فناوری نانو در ایران با تأکید بر راهبرد آتی این فناوری انجام گرفته است.

حمایت‌های ستاد ویژه توسعه فناوری نانو در سال‌های اخیر توجه پژوهشگران و دانشگاه‌های کشور را به امر پژوهش در حوزه

1. Scientometrics
2. Citation Analysis
3. Classification
4. Data Mining
5. Fayyad
6. Dorre
7. Iranian Nanotechnology Initiative Council

فناوری نانو بیش‌تر معطوف داشته و قرائن نشان می‌دهد که این توجه اثرگذار بوده است. به طور مثال، در یکی از بانک‌های اطلاعاتی پایگاه اینترنتی این ستاد[1] - تا زمان تدوین این مقاله - اطلاعات ۳۱۹۳ پایان‌نامه کارشناسی ارشد و رساله دکتری اعم از پایان یافته یا در شرف انجام فهرست شده‌اند، که ممکن است تعداد واقعی در کشور از این تعداد هم بیش‌تر باشد. گرچه عمده پژوهش‌های انجام شده در کشور به جنبه‌های محسوس و سخت‌افزاری معطوف شده‌اند، با این حال پژوهش‌هایی که به جنبه-های نرم‌افزاری و نامحسوس پرداخته‌اند نیز انجام شده‌اند. در زمینه بررسی و مطالعه متون علمی فناوری نانو که محور اصلی پژوهش مرتبط با این مقاله است پژوهش‌های معدودی در کشور نیز صورت گرفته است که در بند بعدی به آن‌ها می‌پردازیم، با این وصف خاطر نشان می‌شود که پیشرفت‌های به‌وجود آمده در فناوری اطلاعات، امکان جستجو و بررسی مؤثر انتشارات، مندرجات و ارتباط آن‌ها را به منظور ارزیابی مستندات فناوری نانو فراهم آورده است.

همانگونه که بیان شد پیشرفت‌های به‌وجود آمده در فناوری اطلاعات فرصت مناسبی را برای بررسی و پژوهش متون علمی فناوری نانو به‌دست داده است. در این میان می‌توان به تعداد معدودی از پژوهش‌های منتشر و یا اعلان شده در قالب مقالات و یا پایان‌نامه‌های خاتمه یافته در داخل کشور به صورتی فشرده اشاره نمود. قرایلو و همکاران (۲۰۰۶) در مقاله خود بدون استفاده از ابزارهای متن‌کاوی به بررسی آماری ۱۰۹۷ نشریه نمایه شده در ISI پرداخته و در آن ۷۹۰ نشریه را مرتبط با فناوری نانو تشخیص داده‌اند [۱۵]. سپهری و همکاران (۲۰۰۶) در مقاله‌ای مبحث شناسایی روندهای نوظهور - یک حوزه موضوعی که در طول زمان در یک جهت مورد علاقه و مفید رشد می‌کند - در فناوری نانوبیو با رویکرد متن کاوی را مورد پژوهش قرارداده‌اند [۱۶]. در این راستا یگانه فلاح (۱۳۸۶) در پایان نامه کارشناسی ارشد خود با استفاده از زیرساخت نرم افزاری داده‌کاوی متنی TMI[2] و با بهره‌گیری از نظر متخصصین حوزه به تحلیل کتاب شناختی[3] مجموعه‌ای از چکیده مقالات حوزه نانوبیوفناوری که از پایگاه داده INSPEC استخراج نموده به منظور شناسایی روندهای نوظهور در حوزه نانوبیوفناوری می‌پردازد [۱۷]. یکی دیگر از پژوهش‌های انجام شده در حوزه فناوری نانو و متن‌کاوی پایان‌نامه کارشناسی ارشد محمدی با عنوان «ترسیم نقشه علمی کشور در فناوری نانو با رویکرد متن کاوی» است [۱۸]. در این پژوهش وی با رویکرد هم‌رخدادی کلمات، مقالات نانوی ایران را خوشه‌بندی کرده و ارتباط بین حوزه‌های مختلف را بررسی می‌کند. این پژوهشگر در مقاله ای با عنوان «حوزه‌های تشکیل دهنده فناوری و علم نانو در ایران» (۱۳۸۸) ابتدا مقالات مرتبط با فناوری و علم نانو که به‌وسیله پژوهشگران ایرانی از سال ۱۹۷٤ تا ۲۰۰۷ منتشر شده بودند، بر اساس مدل جستجوی مؤسسه فناوری جورجیا در وبگاه علوم[4] بازیابی کرده و سپس به روش متن‌کاوی اصطلاحات عنوان، چکیده و کلید واژه‌ها تمامی مقالات استخراج شده و بر اساس الگوریتم‌های داده کاوی اصطلاحات اصلی فناوری و علم نانو در ایران شناسایی می‌کند [۱۹]. خروجی این مقاله شناسایی زیر حوزه‌های موضوعی فناوری و علم نانو در ایران و ترسیم ساختار سلسله مراتبی بر پایه این شناسایی است. طهماسبی و همکاران به بحث نگارگری و تصویردرآوری نشت و انتشار اطلاعات[5] در حالت پویا پرداخته‌اند [۲۰]. آن‌ها سعی کرده‌اند که با بهره‌گیری از اطلاعات فناوری نانو با استفاده از خوشه‌بندی تجمعی و بررسی جریان اطلاعات بین خوشه‌ها در طول زمان تغییرات در داده‌های با مقیاس بالا را ردیابی کنند. طهماسبی در پایان‌نامه کارشناسی ارشد خود با عنوان «مصورسازی پویای نشت اطلاعات بین حوزه‌های علمی به‌وسیله پیوندکاوی» پژوهش را کامل‌تر می‌کند [۲۱]. در این پایان‌نامه از یک مجموعه داده بزرگ در دنیای واقعی که شامل اطلاعات برگرفته از بیش از ۱۳۰۰۰۰ مقاله منتشر شده در زمینه نانوفناوری در سال‌های ۱۹۶۰ تا ۲۰۰۷ از منبع Scopus هستند استفاده شده است. در این پژوهش با بهره‌گیری از یک روش شناختی منسجم متشکل از چندین فرایند متفاوت ماتریسی و خوشه‌بندی روی داده‌ها، روش‌شناختی جدیدی به نام مپنا[6] یا DyVID[7] معرفی و روی داده‌ها پیاده‌سازی شده است. نتایج نشان می‌دهد که با انجام تحلیل استناد روی نتایج خوشه‌بندی بازه‌های مختلف، نشت اطلاعات بین خوشه‌ها قابل ردیابی است.

در سال‌های اخیر تا کنون روش‌های تحلیل متفاوتی از قبیل تحلیل ثبت اختراع‌ها برای ارائه نتایج تحقیقات تجاری شده [۲۲]، اسناد گرنت برای مطالعه تأثیر تأمین بودجه دولتی بر فناوری نانو [۲۳] و متون علمی برای ارائه تلاش‌های پژوهشی دانشگاهی در

1. http://nano.ir/infobeta/stud_projects.php
2. Textual Data Mining Infrastructure
3. Bibliometrics

4. Web of Science
5. Information Diffusion
٦. مصورسازی پویای نشت اطلاعات بین حوزه های علمی
7. Dynamic Visualization of Information Diffusion

ارزیابی وضعیت توسعه علم و فناوری‌نانو استفاده شده‌اند. به علاوه، در طول ۳۰ سال گذشته تعداد وسیعی از مستندات علمی در حوزه علم و فناوری نانو از سراسر دنیا گردآوری و در پایگاه‌های داده متفاوتی ذخیره شده است [۲].

زیرساخت‌ها (نویسندگان فعال، کلیدی‌ترین مجلات، مؤسسات و کشورها، مستندات با بیش‌ترین ارجاع)، ساختار فنی، کاربردهای بالقوه و ابزارآلات مرتبط با فناوری نانو در متون پژوهشی از جمله مواردی بوده‌اند که در تحقیقی که در این زمینه به‌وسیله کاستاف[1] و همکارانش (۲۰۰۷) انجام شده است به‌دست آمده‌اند [۲٤]. در این تحقیق که با اعمال ۳۰۰ کلیدواژه علم و فناوری نانو [۲۵] بر پایگاه-های داده SCI/SSCI[2] انجام شده از روش‌هایی نظیر زبان‌شناسی محاسباتی[3]، خوشه‌بندی مستندات[4]، طبقه‌بندی سلسله‌مراتبی[5] و الگوریتم خوشه‌سازی فازی[6] برای متن‌کاوی استفاده گردیده است. همین پژوهشگران در مقاله‌ای دیگر (۲۰۰۷) به مرور کلی منابع تحقیق فناوری نانو در سطح جهانی نیز پرداخته‌اند [۲٦]. در این مقاله با استفاده از متن‌کاوی روی پایگاه‌های داده SCI/SSCI به استخراج هوش فنی[7] پرداخته‌اند.

در تحقیق دیگری که به‌وسیله آلنکار[8] و همکارانش (۲۰۰۷) انجام شد، اختراعات مرتبط با فناوری نانو در سه کشور ژاپن، آمریکا و آلمان در سه دسته کلی زنجیره ارزش طبقه‌بندی شدند [۲۷]. ارتباط ۱۹۳۵۱ اختراع در این تحقیق با فناوری‌نانو با اعمال ٤٦ کلیدواژه با استفاده از متن‌کاوی مشخص شده است [۲۸]. در ادامه، این اختراعات در سه دسته کلی زنجیره ارزش نظیر نانومواد خام، نانوواسطها و محصولات نانویی از طریق ارتباط کدهای طبقه‌بندی بین‌المللی پروانه‌های ثبت اختراع[9]، طبقه‌بندی شده‌اند.

در تحقیقی که اخیراً به‌وسیله شین‌لی[10] و همکارانش (۲۰۰۸) انجام شده است، سیستم اینترنتی نگاشت دانش[11] نانو به‌وجود آمده است [۲۹]. این سیستم، اختراعات و گرنت‌های مرتبط با نانو را از سال‌های ۱۹۷٦-۲۰۰٦ از پایگاه‌های داده USPTO[12]، EPO[13] و JPO[14] و NSF[15] تجزیه و تحلیل می‌کند. سیستم مذکور با استفاده از روش‌های متن‌کاوی و تحلیل مفهومی[16]، اقدام به اکتساب و تجزیه داده‌ها و سیستم‌سازی می‌نماید. عملگرهای جستجوی پیشرفته، شبکه ارجاعات و نمودارهای روند به کاربر امکان استفاده برخط از امکانات مختلف سیستم را در سطح کشورها، مؤسسات و زمینه‌های فنی برای دوره‌های زمانی مختلف به‌دست می‌آورد.

٤- روش‌شناسی

مقالات دسته‌بندی شده در این تحقیق، شامل تمام مقالات انتشار یافته در مجلات ISI در سال‌های ۱۳۸۲-۱۳۸۸ می‌باشد که به‌وسیله پژوهشگران مراکز علمی و پژوهشی ایران تالیف شده‌اند. ارتباط این مقالات با فناوری نانو به‌وسیله خبرگان این حوزه در ستاد ویژه توسعه فناوری نانو تعیین شده است. تعداد کل این مقالات ۱۹۹۰ مورد می‌باشد. روش متن‌کاوی به کار گرفته شده با اعمال کلیدواژگان تعیین شده، روی سه قسمت اصلی این مقالات، اعم از عنوان[17]، کلمات کلیدی[18] و چکیده[19] مقالات پیاده شده است. از آنجایی‌که اهمیت کلمات استفاده شده در بخش‌های عنوان و کلید واژه بیش‌تر از بخش کلمات چکیده می‌باشند، ضریب تأثیر برای مشاهده کلیدواژگان در هر یک از بخش‌های فوق طبق نظر خبره به ترتیب ۱ و ۱ و ۰/۳ در نظر گرفته شده‌اند.

مراحل طی شده (نمودار جریان) این تحقیق در شکل ۱ آمده است. این روش «طبقه‌بندی بدون ناظر متون بر پایه معیار سیلوئت»[20] که به اختصار آن را SUTC نامیده‌ایم. روش به کار گرفته شده در این پژوهش، روش عمومی متن‌کاوی با ایجاد تغییرات در مرحله دسته‌بندی هوشمند می‌باشد. در دسته‌بندی متعارف لازم است ابتدا تعدادی از مقالات به‌وسیله خبرگان دسته‌بندی شود و سپس مدلی بر اساس مقالات برچسب خورده ساخته شود. در صورتی که در روش SUTC دسته‌بندی بدون آموزش و طی دو مرحله «تخصیص اولیه به روش بازیابی اطلاعات» و «تخصیص ثانویه بر حسب معیار سیلوئت» انجام می‌شود.

1. Kostoff
2. Science Citation Index/Social Citation Index
3. Computational Linguistics
4. Document Clustering
5. Hierarchical Taxonomy
6. Fuzzy Clustering Algorithm
7. Technical Intelligence
8. Alencar
9. International Patent Classification (IPC)
10. Xin Li
11. Internet Knowledge mapping system
12. United States Patent and Trademark Office
13. European Patent Office
14. Japan Patent Office
15. U.S. National Science Foundation
16. Content analysis
17. Title
18. Keyword
19. Abstract
20. Silhouette based Unsupervised Text Categorization

شکل ۱) نمودار جریان روش طبقه‌بندی بدون ناظر متون بر پایه معیار سیلوئت- SUTC

٤-١ درخت فناوری نانو

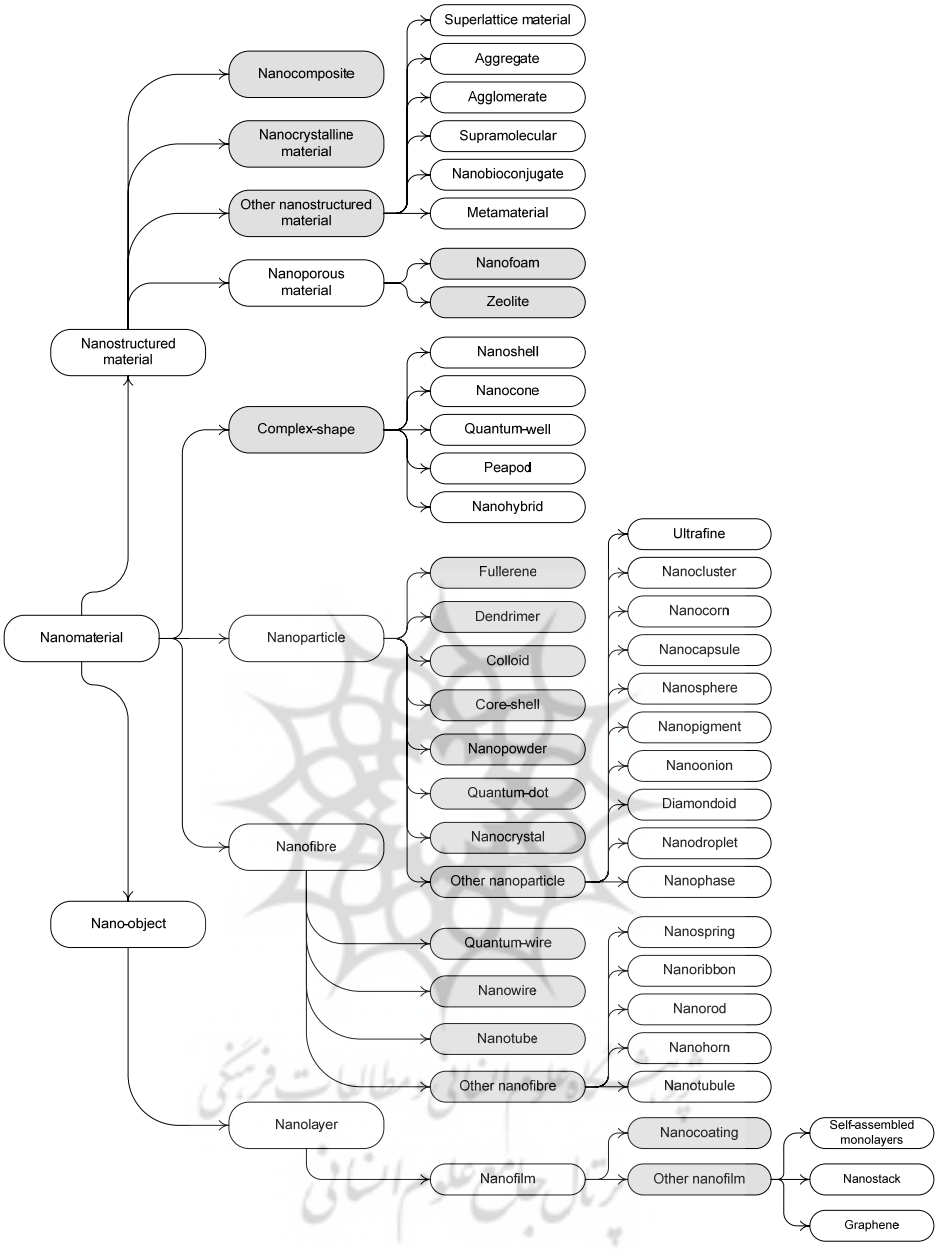
برای دسته‌بندی متون نیاز به یک طبقه‌بندی از پیش تعریف شده داریم. در این تحقیق از دسته‌بندی سلسله‌مراتبی درخت نانو استفاده شده است(شکل ۲). درخت نانوی مورد نظر، زیر نظر ستاد ویژه فناوری نانو و با استفاده از منابع معتبر تهیه شده است. درخت‌های ایجاد شده شامل حوزه‌های مختلفی می‌گردد که عبارتند از: نانو مواد[1]، نانو سیستم‌ها ونانوابزارها[2]، روش‌های ساخت و تولید[3]، روش‌های‌اندازه‌گیری[4]، خواص[5]، علوم نانو[6] و کاربردها[7]. کلیدواژگان مورد اعمال در این دسته‌بندی‌ها از تلفیق استانداردهای معتبری نظیر ISO [۳۳-۳۰] ، BSI [۳۷-۳٤] ، درخت‌های نانو، از انتشارات گروه مطالعاتی آینده‌اندیشی ستاد ویژه توسعه فناوری نانو در سه گروه درخت‌های علم [۳۸]، درخت صنعت [۳۹] و درخت فناوری‌نانو [٤۰] و تحقیقات کاستاف و همکاران [۲٤]، آلنکار و همکاران [۲۷]، و شین لی و همکاران [۲۹] به‌دست آمده‌اند. تلفیق مذکور، پس از پیاده کردن اولیه روش متن‌کاوی مورد نظر روی مقالات با توجه به هر یک از استانداردهای موجود و برای رفع نقص هر یک از استانداردها در معرفی عناصر و واژگان مورد استفاده در فناوری‌نانو صورت گرفته است. باید یادآور شد که کلید واژگان مورد اعمال در این تحقیق متن‌کاوی، تنها در حوزه نانو مواد بوده است. برای دسته‌بندی ابتدا طبق نظر خبرگان ستاد نانو مقطعی از این درخت انتخاب شد. این مقطع شامل دسته‌هایی می‌باشد که اشتراکی با هم نداشته و اجتماع آن‌ها تشکیل دهنده کل حوزه نانومواد است. به عبارت دیگر حوزه نانومواد به ۲۰ دسته که در شکل ۲ با رنگ خاکستری مشخص شده افراز شده است. دسته‌های حاصل از این برش در شکل ۳ در بخش تحلیل ارائه گردیده است.

٤-۲ پیش پردازش متون

پس از تشکیل درخت نانومواد، متن مقالات نانو باید پیش‌پردازش شوند. یکی از مهم‌ترین تفاوت‌های متن‌کاوی با داده‌کاوی معمولی،

1. Nanomaterials
2. Nanosystems and nanodevices
3. Nanomanufacturing processes
4. Nanoscale measurement methods
5. Properties
6. Nanoscience
7. Applications

نحوه پیش‌پردازش است. پیش‌پردازش متون از مراحل زیر تشکیل می‌شود:



شکل ۲) دسته‌بندی سلسله مراتبی نانو مواد برپایه درخت‌های فناوری نانو

۱- **تفکیک به کلمات؛**

۲- **حذف کلمات توقف** شامل افعال عمومی، حروف **اضافه** و ربط؛

۳- **مترادف‌ها** : برای پیش‌گیری از بروز خطا در **محاسبه** تعداد واژگان، لغات هم معنی و عبارات اختصاریافته[1] استخراج می‌گردند. همچنین حالت‌های مختلف نگارش یک عبارت (به عنوان مثال: quantumdot, quanrum-dot, quantum dot) و مشتقات مختلف یک عبارت (به عنوان مثال: nano structure, nano structured) نیز در نظر گرفته شده است.

۴- **ریشه‌یابی**: کلمات با **الگوریتم** پورتر[2] ریشه‌یابی می‌شوند.

۵- **وزن‌دهی به کلمات با مدل فضای برداری.**

1. Abbreviation

2. Porter

همچنین در محاسبه فراوانی هر کلمه، طبق نظر خبره ضرایب اهمیت متفاتی برای عنوان، کلمات کلیدی و چکیده در نظر گرفته شده است.

٤-٣ مدل فضای برداری

در این مدل، متون به شکل بردارهایی در فضای چند بعدی اقلیدسی نمایش داده می‌شود [٤١]. هر محور این فضا متناظر با یک کلمه یا عبارت است.

فراوانی کلمه (TF)[1]: دفعات رخ دادن کلمه t در متن d یعنی n(d,t) است. در این مقاله فراوانی کلمه نرمال نشده است زیرا نرمال سازی کسینوسی طول بردار کلمات که در ادامه توضیح داده می‌شود به طور طبیعی این منظور را نیز برآورده می‌کند.

فراوانی معکوس متن (IDF)[2]: همه محورها در فضای برداری به‌اندازه هم مهم نیستند. IDF به دنبال کاهش اهمیت کلماتی است که در تعدادی زیادی از متون وجود داشته و بنابراین ارزش کمی دارند. اگر N تعداد کل متون و DF_t تعداد متون دارای کلمه t باشد، آنگاه یک شکل متداول وزن‌دهی IDF بر اساس رابطه (۱) است:

$$IDF(t) = -\log_2(DF_t / N) \qquad (۱)$$

اگر $DF_t << N$ آنگاه کلمه t دارای ضریب مقیاس IDF بزرگی خواهد بود و برعکس.

TF و IDF با هم به طوری طبیعی ترکیب می‌شوند تا مدل فضای برداری را تشکیل دهند که در آن مختصات متن d در محور t طبق رابطه (۲) داده می‌شود:

$$w_t = TF(d,t)IDF(t) \qquad (۲)$$

برای‌اندازه‌گیری تشابه بین دو سند متنی $\vec{d}_1$ و $\vec{d}_2$، کسینوس زاویه بین $\vec{d}_1$ و $\vec{d}_2$ طبق رابطه (۳) محاسبه می‌شود.

$$sim(d_1,d_2) = \frac{\vec{d}_1 \bullet \vec{d}_2}{|\vec{d}_1| \times |\vec{d}_2|} = \frac{\sum_{i=1}^{t} w_{i,1} \times w_{i,2}}{\sqrt{\sum_{i=1}^{t} w_{i,1}^2} \times \sqrt{\sum_{i=1}^{t} w_{i,2}^2}} \qquad (۳)$$

این نحوه‌اندازه‌گیری تشابه معادل نرمال کردن کسینوسی طول بردار متون است که در آن وزن کلمات هر متن بر طول بردار آن کلمات آن متن تقسیم می‌شوند.

1. Term Frequency
2. Inverse document frequency

٤-٤ دسته‌بندی مقالات

در روش‌های متداول دسته‌بندی، تعدادی از مشاهدات از قبل دارای برچسب دسته می‌باشند. این داده‌های دارای برچسب به دو گروه آموزش[3] و آزمون[4] تقسیم می‌شوند. با روش‌های داده‌کاوی، مدلی روی داده‌های آموزش ساخته شده و برای ارزیابی، برچسب داده‌های آزمون با کمک این مدل پیش‌بینی شده و با مقادیر واقعی مقایسه می‌شوند تا دقت مدل تعیین شود. در این تحقیق هر مقاله یک مشاهده و هر گره در سطح مشخص از درخت نانو یک برچسب دسته است.

از آن جا که در این مطالعه، دسته مقالات از قبل ناشناخته می‌باشد، روشی ابتکاری برای دسته‌بندی متون ابداع شد. در این روش ابتدا هر دسته به شکل یک سند متنی در نظر گرفته می‌شود که حاوی عبارات عنوان آن دسته، عبارات زیردرخت آن دسته و مترادفات عبارات می‌شود. برای مثال دسته other nanoparticle شامل عبات ultrafine و بقیه زیردسته‌های مربوط است. سپس عبارات کلیدی هر دسته در هر مقاله شمارش شده و هر مقاله با توجه به فراوانی تجمیعی عبارات موجود در آن و ضریب تأثیر به یکی از دسته‌ها تخصیص می‌یابد. این مرحله *تخصیص اولیه* نامیده می‌شود.

در تخصیص اولیه، بسیاری از مقالات به دلیل نداشتن هیچکدام از کلیدواژگان دسته‌بندی نمی‌شوند. برای دسته‌بندی این مقالات، شباهت هر مقاله دسته‌بندی نشده به مقالات تخصیص داده شده به یک دسته، طبق مدل فضای برداری و شباهت کسینوسی محاسبه می‌گردد. قابل توجه است که این شباهت علاوه بر کلیدواژگان شامل تمام عبارات مهم یک مقاله می‌شود. برای هر مقاله بدون دسته، معیار سیلوئت[5] هر بار با فرض قرار گرفتن آن مقاله در یکی از دسته‌ها محاسبه شده و مقاله به دسته‌ای تخصیص می‌یابد که مقدار سیلوئت بالاتری داشته باشد.

معیار سیلوئت، نسبت متوسط فاصله هر مشاهده به خوشه یا دسته خودش به نزدیکترین خوشه یا دسته دیگر رااندازه می‌گیرد [٤٢]. معیار سیلوئت طبق رابطه (٤) محاسبه می‌شود:

$$s_i = \frac{a_i - b_i}{\max(a_i, b_i)} \qquad (٤)$$

در این رابطه b_i متوسط فاصله مشاهده i ام از خوشه‌ای است که به آن تعلق دارد و a_i متوسط فاصله مشاهده i ام از نزدیک‌ترین خوشه غیر از خوشه خود است. مقدار سیلوئت بین ۱- و ۱+ قرار

3. Train
4. Test
5. Silhouette

دارد. مقدار مثبت و بزرگ سیلوئت یک مشاهده نشان می‌دهد که آن مشاهده به درستی در خوشه مربوط به خود قرار گرفته است. مقدار منفی برای سیلوئت یک مشاهده نشان دهنده تخصیص نامناسب آن مشاهده به خوشه است.

معیار سیلوئت در اصل برای ارزیابی خوشه‌بندی‌های موجود استفاده می‌شود؛ ولی در این جا به شکلی نوین برای دسته‌بندی مقالات تخصیص نیافته استفاده شده است. با توجه به درجه شباهت، هر مقاله به ترتیب اولویت در چند دسته قرار می‌گیرد که نشانگر میان‌رشته‌ای بودن یک مقاله است. لازم به ذکر است که در این تحقیق، نتایج تحلیل تخصیص هر مقاله به دو دسته ارائه می‌شود.

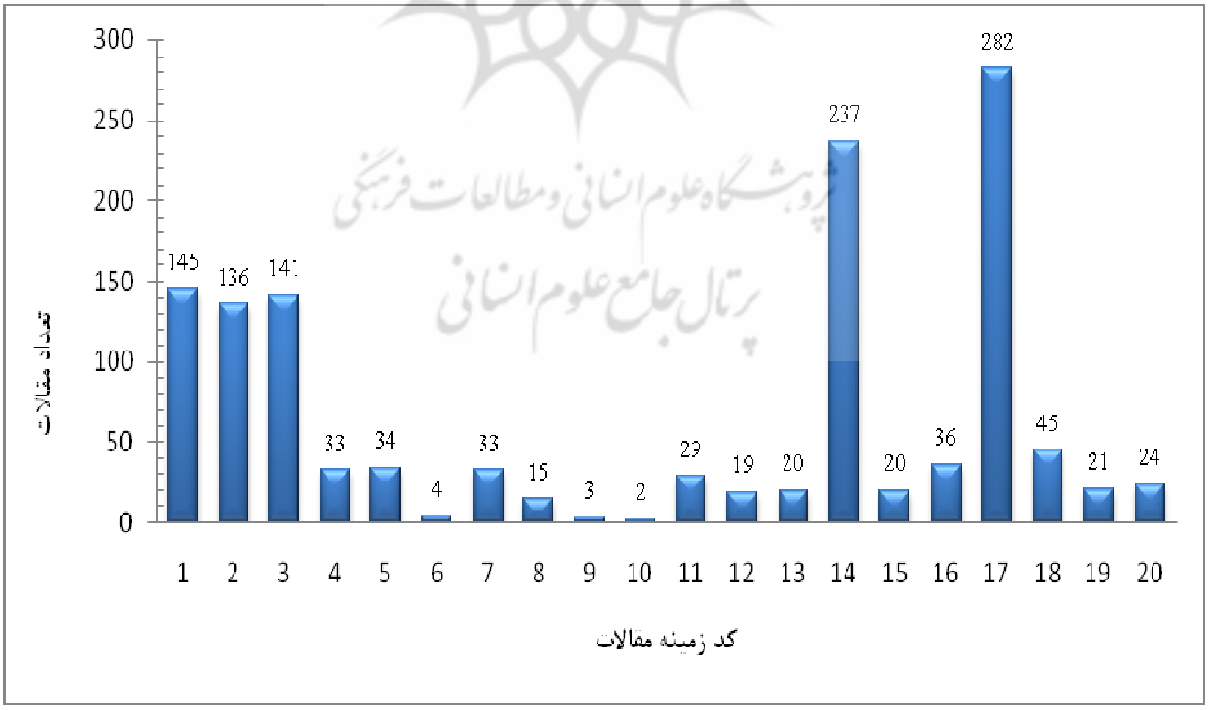
۵- تحلیل

با به‌کارگیری روش‌های شرح داده شده، تعداد ۱۹۹۰ مقاله در ۲۰ طبقه دسته‌بندی شدند. سطح طبقات مورد استفاده در این تحقیق به منظور دسته‌بندی، برشی از درخت شکل ۲ می‌باشد که طبق نظر خبرگان ستاد نانو تعیین شده است. این برش شامل موارد خاکستری بوده و در جدول ۱ فهرست شده است.

تمام پردازش‌ها با زبان برنامه‌نویسی Python انجام شده است. این زبان برای پردازش متون مناسب می‌باشد. برای پردازش زبان طبیعی از ماجول MontyLingua استفاده شده و محاسبات ماتریسی با ماجول Scipy و Numpy انجام شده‌اند.

جدول ۱) بیست زمینه طبقه‌بندی مقالات بر پایه نظر خبرگان ستاد نانو

کد	زمینه	کد	زمینه	کد	زمینه	کد	زمینه	کد	زمینه
۱	Nanocomposite	۵	Zeolite	۹	Colloid	۱۳	Nanocrystal	۱۷	Nanotube
۲	Nanocrystalline	٦	Complex shape	۱۰	Core shell	١٤	Nanoparticle	۱۸	Nanofibre
۳	Nanostructured	۷	Fullerene	۱۱	Nanopowder	۱۵	Quantum wire	۱۹	Nanocoating
٤	Nanofoam	۸	Dendrimer	۱۲	Quantum dot	١٦	Nanowire	۲۰	Nanofilm

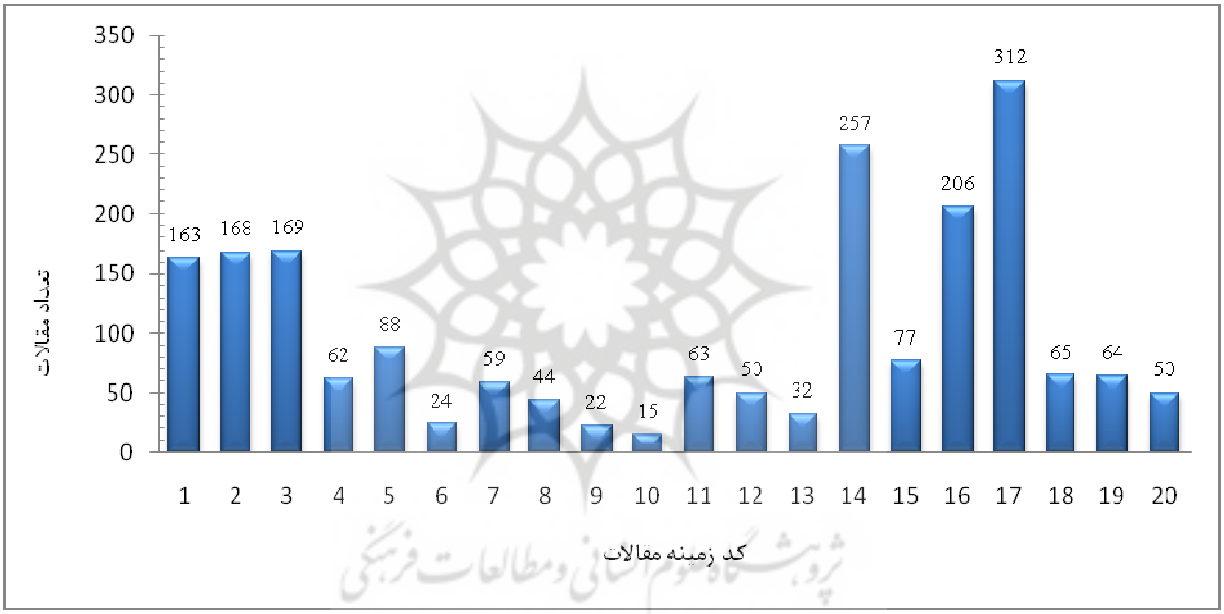


شکل ۳) نمودار توزیع فراوانی مقالات در ۲۰ زمینه در دسته‌بندی اولیه

توزیع تعداد مقالات با دسته‌بندی اولیه، طبق شکل ۳ است. در این مرحله دسته ۱۲۷۹ مقاله مشخص شد از تعداد ۱۹۹۰ مقاله، ۷۱۱ مقاله با دسته‌بندی اولیه بدون دسته ماندند. این مقالات با کمک معیار سیلوئت به خوشه مقالات دسته‌ها تخصیص یافتند. توزیع دسته‌بندی جدید شامل همه ۱۹۹۰ مقاله طبق شکل ٤ است. رشد قابل ملاحظه خوشه‌های کوچک قابل توجه است. این موضوع نشان می‌دهد که تعداد قابل توجهی از مقالات ممکن است فاقد عبارات کلیدی یک دسته بوده؛ ولی با توجه به محتوای متن و دیگر کلمات، متعلق به آن دسته باشند.

۵-۱ ارزیابی دسته‌بندی

در ادامه تعداد ٤۰ مقاله طبق نظر خبرگان دسته‌بندی شدند. نتایج دسته‌بندی خبرگان با دسته‌بندی خودکار مقایسه شده و دقت مدل محاسبه گردید. نتایج دسته‌بندی خودکار در پیوست ۱ فهرست شده است. در ۳ مورد اصولاً مقالات هیچ ارتباطی با نانومواد نداشته و به دسته‌های دیگری از جمله نانو سیستم‌ها و ابزارها، روش‌های تولید و یا دیگر حوزه‌ها تعلق داشته‌اند. از ۳۷ مقاله باقی‌مانده، ٤ مقاله به‌وسیله سیستم خودکار به اشتباه دسته‌بندی شدند. بنابراین دقت مدل به‌کارگرفته شده پس از کنار گذاشتن مقالاتی که متعلق به حوزه نانومواد نیستند ۸۹٪ می‌باشد که این میزان، قابل توجه است.



شکل ٤) نمودار توزیع فراوانی مقالات در ۲۰ زمینه در دسته‌بندی ثانویه با به‌کارگیری معیار سیلوئت

یکی از نکات جالب مدل، تخصیص‌دهی ثانویه است که امکان دسته‌بندی مقالاتی را فراهم می‌کند که حاوی هیچ یک از عبارات کلیدی دسته‌ها نیستند. در تخصیص‌دهی اولیه از تعداد ٤۰ مقاله فوق، ۷ مقاله به هیچ دسته‌ای اختصاص داده نشد. در ۳ مورد اصولا مقالات هیچ ارتباطی با نانومواد نداشتند. از ٤ مقاله دیگر در تخصیص ثانویه طبق معیار سیلوئت به یکی از ۲۰ دسته موجود تخصیص داده شدند. در جدول ۲، تخصیص‌دهی ثانویه این ٤ مقاله طبق معیار سیلوئت با نظر خبره مقایسه شده و مقالاتی که هر دو تخصیص‌دهی در آن‌ها یکسان بوده‌است با T و مقالاتی که یکسان نبوده‌اند با Fنشان داده شده‌اند. دسته‌بندی ۳ مقاله از ٤ مقاله درست می‌باشد. با توجه به این‌که ۷۱۱ مقاله در تخصیص اولیه بدون دسته می‌مانند دسته‌بندی ثانویه اهمیت ویژه‌ای دارد.

جدول ۲) بررسی دقت مدل پس از به‌کارگیری معیار سیلوئت برای دسته‌بندی مقالات بدون دسته در تخصیص اولیه

ردیف	اطلاعات مقالات طبقه‌بندی نشده در تخصیص دهی اولیه	نظر خبره	دسته‌بندی خودکار	نتیجه
1	Title: A flexible design for one-dimensional photonic crystals with controllable photonic bandgap width Keywords: Photonic crystals; Photonic bandgap; Equivalent multilayer structure; Effective optical contrast ratio; Relative bandwidth Abstract: We propose a flexible design for one-dimensional photonic crystals (1D-PCs) with controllable bandgap width which based on the two refractive index mediums. In this structure, each of the low index layers replaced by equivalent three-layers which consist …	Nano-crystal Optical properties	Nanowire Fullerene	F
2	Title: Study of anodization parameters effects on photoconductivity of porous silicon No keywords Abstract: We have prepared porous silicon by etching p-type crystalline silicon in different conditions such as: varying electrolyte concentration, current density, and etching time. The primary objective of this research is to develop a scientifically based...	Photocunductivity Nano-porous	Nanoporous Nanostructured	T
3	Title: Role of surface Cooper pair interactions on critical temperature of ultra thin film superconductors Keywords: Superconductivity; Thin film; BCD theory; Cooper pair interaction; Critical temperature; Fermi surface Abstract: The superconductivity mechanism of Pb thin film on a Si substrate in the weak interaction regime is investigated. A discrete Fermi surface is constructed depend on the film thickness and electron density and crystallographic orientation. We consider two…	Thin film Superconductivity	Quantumwire Nanofilm	T
4	Title: Role of process control agent on mechanical alloying of nano structured TiAl-based alloy Keywords: Titanium aluminides- Mechanical alloying- Process Control Agent Abstract: High energy ball milling was performed on a mixture of titanium and aluminum elemental powders with a composition of Ti-48(at.%)Al. Stearic acid was added to this powder mixture as a process control agent (PCA) to study its effect on the microstructure…	Nanostructured Mechanical alloying	Nanocoating Nanocrystalline	T

٦- نتیجه‌گیری و جمع‌بندی

در این مقاله، مقالات فناوری نانو با توجه به درجه تعلق آن‌ها به طبقات از پیش تعریف شده نانومواد تخصیص می‌یابند. نتایج حاصله بر اساس نظر خبرگان نویدبخش بوده و راه‌گشای دسته‌بندی خودکار مقالات فعلی و آتی است. این روش در شناخت زمینه تحقیقاتی دیگر مستندات منتشره در فناوری نانو و فناورهای دیگر نظیر اختراعات، کتب و پایان‌نامه‌ها می‌تواند استفاده شود.

نوآوری اول این مقاله در نحوه نگرش به درخت نانو به شکل عبارات جستجو در بازیابی اطلاعات است. این نگرش امکان دسته‌بندی مقالات را در یک طبقه‌بندی با توجه به برچسب متنی طبقات فراهم می‌کند. مزیت این کار در عدم لزوم آموزش طبق یک مدل دسته‌بندی است و در نتیجه لازم نیست تعداد زیادی از مقالات از پیش به‌وسیله خبرگان برچسب‌گذاری شوند.

نوآوری دوم در استفاده از معیار سیلوئت برای دسته‌بندی بدون ناظر مقالات باقی مانده می‌باشد. همان طور که قبلاً اشاره شد معیار سیلوئت در اصل برای ارزیابی خوشه‌بندی‌های موجود استفاده می‌شود و طبق دانش نویسندگان، تاکنون به این شکل برای دسته‌بندی متون استفاده نشده است. این روش در عین ساده بودن کارایی مناسبی داشته و راه را برای دسته‌بندی بدون ناظر در حوزه‌های مختلف علم و فناوری هموار می‌کند.

برای تحقیقات آتی، استفاده از روش‌های نرمال‌سازی پیشرفته‌تر همچون نرمال‌سازی محوری برای افزایش دقت مدل و نیز استخراج خودکار ضرایب تأثیر هر بخش از متن به تفکیک عنوان، کلیدواژگان و چکیده از روی اطلاعات برچسب‌های تعیین شده به‌وسیله خبرگان تحت بررسی است.

۷- تشکر و قدردانی

نویسندگان مراتب سپاسگزاری خود را نسبت به آقای مهندس مهدی حبیب‌نژاد کورایم و ستاد ویژه توسعه فناوری نانو که در مراحل مختلف تحقیق و نگارش این مقاله همکاری مؤثری داشته‌اند، ابراز می‌دارند. همچنین نویسندگان مراتب سپاسگزاری خود را نسبت به آقای دکتر سید سپهر قاضی نوری، دانشیار دانشگاه تربیت مدرس، به‌دلیل راهنمایی‌های با ارزش ایشان اعلان داشته و نیز از داوران محترم ناشناس که با پیشنهادهای ارزشمند خود بر غنای مطلب افزوده‌اند، سپاسگزاری می‌نمایند.

این مقاله بخشی از پژوهشی است که از حمایت مالی مرکز تحقیقات مخابرات ایران (ITRC) به شماره قرارداد ۵۰۰/۲۰۱۲۷/ت برخوردار بوده است.

منابع References

[۱] قاضی نوری، س.س.، حیدری، ا.، ۱۳۸۷، «بیم‌ها و امیدها درباره آثار اجتماعی فناوری نانو (با نگاهی به ایران)»؛ مرکز تحقیقات سیاست‌های علمی کشور، تهران.

[2] Kostoff, R., Murday, J., Lau, C., Tolles, W., 2006, "The seminal literature of nanotechnology research", J Nanopart Res; Vol. 8, No. 2, pp. 193–213, Doi: 10.1007/s11051-005-9034-9

[3] Schummer, J., 2004, "Multidisciplinarity, interdisciplinarity, and patterns of research collaboration in nanoscience and nanotechnology", Scientometrics, Vol. 59, No. 3, pp. 425–465, Doi: 10.1023/B:SCIE. 0000018542.71314.38.

[4] Mansuri, R., 2005, "Iran's science and technology policies and indicators", Presented at Seminar on UK & Iran Collaboration in Higher Education and Research: Achivements and Challenges; Tehran.

[5] Nature, Editorial, 2005, Iran's long march, pp: 435,247-248.

[٦] پایگاه اینترنتی ستاد ویژه توسعه فناوری نانو (www.nano.ir).

[۷] حبیب نژاد کورایم، م.، ۱۳۸٦، «نانو مواد: تعاریف، روش‌های تولید، خواص و کاربردها»؛ انتشارات عقیق، تهران.

[۸] سند راهبرد ۱۰ ساله آینده توسعه فناوری نانو در جمهوری اسلامی ایران، ستاد ویژه توسعه فناوری نانو، تهران، ۱۳۸۵.

[9] Janssens, F., Glanzel, W., De Moor, B., 2008, "A hybrid mapping of information science", Scientometrics, Vol. 75, No. 3, pp: 607-631.

[10] Glenisson, P., Glanzel, W., Janssens, F., De Moor, B., 2005, "Combining full-text and bibliometric information in mapping scientific disciplines", Information Processing and Management, Special Issue on Informetrics, Vol. 41, No. 6, pp: 1548-1572.

[11] Fayyad, U., Piatetsky-Shapiro, G., Smyth, P., Uthurusamy, R., 1996, "Advances in Knowledge Discovery and Data Mining", MIT Press.

[12] Dörre, J., Gerstl, P., Seiffert, R., 1999, "Text Mining: Finding Nuggets in Mountains of Textual Data", KDD-99.

[13] Ghazinoory,S., Heydari, E., 2008, "Potential impacts of nanotechnology development in Iran", Technology and Society Magazine, Vol. 27, No. 4, pp: 37, 44.

[14] Sarkar, S., Beitollahi, A., 2009, "An overview on nanotechnology activities in Iran", Iranian J Publ Health, Vol. 38, No. 1, pp: 65-68.

[15] Gharailou, D., Maghrebi, M., Abbasi, A., 2006, "Comprehensive evaluation of nanotechnology related ISI journals", First International Congress on Nanoscience and Nanotechnology, Faculty of Engineering, University of Tehran, 18th to 20th, December.

[16] Sepehri, M.M., Jahanshahi, M., Yeganeh Fallah, A., 2006, "Identification of emerging trends in Nano-Bio-Technology: A Text Mining Approach", First International Congress on Nanoscience and Nanotechnology, Faculty of Engineering , University of Tehran, 18th to 20th, December.

[۱۷]یگانه فلاح، آ.، ۱۳۸٦، «شناسایی روند نوظهور در دانش نانوبیوتکنولوژی»؛ *پایان‌نامه کارشناسی ارشد مهندسی سیستم‌های اقتصادی- اجتماعی*، دانشگاه تربیت مدرس، دانشکده فنی و مهندسی، تهران.

[۱۸] محمدی، ا.، ۱۳۸۷، «ترسیم نقشه علمی نانو فناوری در ایران»؛ *پایان‌نامه کارشناسی ارشد علوم کتابداری و اطلاع‌رسانی*، دانشگاه آزاد

اسلامی، واحد علوم و تحقیقات، دانشکده علوم انسانی، تهران.

[۱۹] محمدی، ا.، ۱۳۸۸، «حوزه‌های تشکیل دهنده فناوری و علم نانو در ایران»؛ *پنجمین همایش دانشجویی فناوری نانو، ۹-۱۱ خردادماه*، دانشکده فناوری‌های نوین، دانشگاه علوم پزشکی تهران، تهران.

[۲۰] طهماسبی، ن.، سپهری، م.م. و تیمورپور، ب.، ۱۳۸٦، «مصورسازی پویای نشت اطلاعات بین حوزه‌های علمی به‌وسیله پیوندکاوی»؛ کنفرانس داده‌کاوی ایران، IDMC'07، ۲۹-۳۰ آباتماه، دانشگاه صنعتی امیرکبیر، تهران.

[۲۱] طهماسبی، ن.، «مصورسازی پویای نشت اطلاعات بین حوزه‌های علمی به‌وسیله پیوندکاوی»؛ *پایان‌نامه کارشناسی ارشد مهندسی فناوری اطلاعات، دانشکده فنی و مهندسی*، دانشگاه تربیت مدرس ، تهران.

[22] Huang, Z., Chen, HC., Chen, ZK., Roco, MC., 2003, "International Nanotechnology Development in: country, institution, and technology field analysis based on USPTO patent database", J Nanopart Res, 2004, Vol. 6, No. 4, pp: 325–354.

[23] Roco, MC., 2005, "International perspective on government nanotechnology funding in 2005", J Nanopart Res; Vol. 7, No. 6, pp: 707–712, Doi: 10.1007/s11051-005-3141-5.

[24] Kostoff, R.N., Koytcheff, R.G., Lau, C.G.Y., 2007, "Global nanotechnology research literature overview", Technological Forecasting & Social change, Vol. 74, pp: 1733-1747.

[25] Kostoff, R.N., Koytcheff, R.G., Lau, C.G.Y., 2007, "Structure of the global nanoscience and nanotechnology research literature", DTIC Technical Report Number ADA461930; Defense Technical Information Center; Fort Belvoir; VA.

[26] Kostoff, R.N., Koytcheff, R.G., Lau, C.G.Y., 2007, "Global nanotechnology research literature overview", Current Science, 92:11, June, pp: 1492-1498.

[27] Alencar, M.S.M., Porter, A.L., Antunes, A.M.S., 2007, "Nanopatenting patterns in relation product life cycle", Technological Forecasting & Social change; Vol. 74, pp: 1660-1680.

[28] Antunes, A., Canongia, C., Alencar, M.S.M., Mangueira, A.C., Tibau, F., Arnor, M., 2004, "Prospecção desenvolvimento científico de países_chave e países competidores [Nanotechnology Foresight: key and competitors countries]", (Study Commissioned by the CGEE/MCT), Rio de Janeiro: Information Industry, Escola de Química, Universidade Federal do Rio de Janeiro, Brazil.

[29] Li, x., Hu, D., Dang, Y., Chen, H., Roco, M.C., Larson, C.A., Chan, J., 2008, "Nano Mapper: an Internet knowledge mapping system for nanotechnology development", *Journal of Nanoparticles Research*, DOI 10.1007/s11051-008-9491-z.

[30] ISO/TS 27687, Nanotechnologies, terminology & definitions for nanoparticles.

[31] ISO/WD TS 11751, Nanotechnologies, terminology & definitions for carbon nanomaterials.

[32] ISO/WD TS 12921, Nanotechnologies, terminology & definitions for nanostructured materials.

[33] ISO/JWG 1- Ver.1 TR, Nanotechnologies, terminology, Framework for core concepts.

[34] BSI, PAS 133, Terminology for nanoscale measurement & instrumentation.

[35] BSI, PAS 135, Terminology for nanofabrication.

[36] BSI, PAS 71, Terminology for nanomaterials.

[37] BSI, PAS 71, Vocabulary / nanoparticles.

[38] Science Tree, Available from <http://www.nano.ir/sub_forsight.php?page= foresigh_Sci_Tree> (Cited at 2009/6/28).

[39] Industry Tree; Available from <http://www.nano.ir/sub_forsight.php?page= foresight_Ind_Tree> (Cited at 2009/6/28).

[40] Technology Tree; Available from <http://www.nano.ir/sub_forsight.php?page= foresight_Tech_Tree> (Cited at 2009/6/28).

[41] Salton, G., Wong, A., Yang, C.S., 1975, "A vector space model for automatic indexing", Communications of the ACM, Vol. 18, Issue 11, pp: 613-620.

[42] Rousseeuw, P.J., 1987, "Silhouettes: A graphical aid to the interpretation and validation of cluster analysis", *Journal of Computational and Applied Mathematics*, Vol. 20, pp: 53-65.

پیوست ۱) دسته‌بندی خودکار ٤٠ مقاله برای مقایسه با نظر خبرگان

عنوان مقاله	طبقه
A novel method for porosity measurement of various surface layers of nanofibers mat using image analysis for tissue engineering applications	nanofibre
Optimum self-ordered nanopore arrays with 130–270 nm interpore distances formed by hard anodization in sulfuric/oxalic acid mixtures	Nanoporous
Investigation of Ni Nanocrystallization and The Effect of Al2O3 Addition by High-Energy Ball Milling	Nanocrystalline Nanocomposite
Szeged Index of HAC5C6C7[k, p] Nanotube	Nanotube
Influences of cellulose nanofiber on the epoxy network formation	Nanofibre
A flexible design for one-dimensional photonic crystals with controllable photonic bandgap width	nanowire fullerene
Study of anodization parameters effects on photoconductivity of porous silicon	Nanoporous Nanostructured
Synthesis of ZnO nano powder by a gel combustion method	nanopowder
Mechanochemical synthesis of fluoridated hydroxyapatite nanopowder	nanopowder
Magnetic properties of nanostructured MnZn ferrite	Nanocrystalline Nanostructured
Plasmon hybridization in metallic nanotubes	nanotube
Direct ultrasonic-assisted synthesis of sphere-like nanocrystals of spinel Co3O4 and Mn3O4	Nanocrystal Nanoparticle
Nanoparticles of a new mercury(II) coordination polymer: Synthesis, characterization, thermal and structural studies	nanoparticle
Effect of milling time on mechanochemical synthesis of TiO2 nanoparticles	nanoparticle
Differential pulse voltammetric determination of N-acetylcysteine by the electrocatalytic oxidation at the surface of carbon nanotube-paste electrode modified with cobalt salophen complexes	nanotube
Taguchi optimization for combustion synthesis of aluminum oxide nano-particles	Nanostructured Nanopowder
Structural, compositional and optical characterization of water soluble CdS nanoparticles synthesized by ultrasonic irradiation	Nanostructured Nanoparticle
Crystalline order and mechanical properties of as-electrospun and post-treated bundles of uniaxially aligned polyacrylonitrile nanofiber	nanofibre
Carburizing of low-melting-point metals by pulsed nanocrystalline plasma electrolytic carburizing	nanocrystalline
Preparation and characterization of nanocomposite membranes made of poly(2,6-dimethyl-1,4phenylene oxide) and montmorillonite for direct methanol fuel cells	nanocomposite
Effective coordination number model for the size dependency of physical properties of nanocrystals	nanocrystal
Mechanochemical synthesis and characterization of nanostructure forsterite Bioceramics	Nanocrystalline Nanostructured
Study of Er-Doped ZnS Quantum Dots Synthesized by Chemical Capping Method	Quantumdot
Simulation of STM technique for electron transport through boron-nitride nanotubes	nanotube
Investigation of superconductivity in the single-walled carbon nanotubes	nanotube
Computational modeling of an ion-driven nanomotor	Nanocrystal Nanoporous
Role of surface Cooper pair interactions on critical temperature of ultra thin film superconductors	Quantumwire Nanofilm
Modeling the effect of external electric field and current on the stochastic dynamics of ATPase nano-biomolecular motors	Nanoporous Nanocrystal
Role of process control agent on mechanical alloying of nano structured TiAl-based alloy	Nanocoating Nanocrystalline
Electrospun Poly (e-Caprolactone) Nanofiber Mat as Extracellular Matrix	nanofibre
Hydrothermal Synthesis Of nanostructure NaA Zeolite: The effect of Synthesis Parameters on Zeolite Seed Size And Crystallinity	Zeolite Nanostructured
Electrodeposition of nickel oxide nanoparticles on glassy carbon surfaces: application	nanoparticle

to the direct electron transfer of tyrosinase	
Nanocrystalline ZnO for Knoevenagel Condensation and Reduction of the Carbon,Carbon Double Bond in Conjugated Alkenes	nanocrystalline
Electrochemical Properties and electrocatalytic activity of FAD immobilized onto cobalt oxide nanoparticles: Application to nitrite detection	nanoparticle
Host (Nanodimensional Pores of Mesoporous Material) ،Guest (Semiconductor Nanoparticles) Nanocomposite Materials	Nanoparticle Nanocomposite
Electrosorption of Os(lll)-Compelex at Single-well carbon nanotubes immobilized on glassy carbon electrode: Application to nanomolar detection of bromate, periodate and iodate	nanotube
Effect of Nano and Micro Emulsion Silicone Softeners on Properties of Polyester Fibers	Nanocoating Nanowire
Effect of gold nanoparticle as a novel nanocatalyst on luminol-hydrazine chemiluminescence system and its analytical application	nanoparticle
Synthesis of nano-structured mullite by sol-gel method with intermediate mechanical activation	nanostructured
Heat treatment effects on structure and properties of synthesized nanocrystalline NiTi intermetallic by mechanical alloying	nanocrystalline

پژوهشگاه علوم انسانی و مطالعات فرهنگی
پرتال جامع علوم انسانی