

بِسْمِ اللَّهِ الرَّحْمَنِ الرَّحِيمِ



دانشگاه علامه طباطبائی

دانشکده روانشناسی و علوم تربیتی

گروه کتابداری و اطلاع رسانی

پایان نامه جهت اخذ کارشناسی ارشد

عنوان:

نگاشت نقشه علمی کتابداری براساس پایان نامه های رشته کتابداری ایران

استاد راهنما

دکتر مهدی علیپور حافظی

استاد مشاور

دکتر عصمت مومنی

پژوهشگر

حمیده جبرئیل زاده

زمستان ۱۳۹۱

تقدیم بہ

پدر و مادر عزیزم

شکر و قدردانی

منت خدای را عزوجل که طاعتش موجب قربت است و به شکر اندرش مزید نعمت! هر نفسی که فرو می رود مدحیات است و چون برمی آید منفرح ذات! پس در هر نفسی دو نعمت موجود بر هر نعمت، شگرتی واجب!

از دست و زبان که برآید کز عمده شکرش به درآید

بدون شک انجام پژوهش و نگارش پایان نامه حاضر بدون بسیاری و کجک استادان محترم میسر نمی گردید. محقق از ابتدای شروع کار پایان نامه، همواره از بهکاری صمیمانه و بی شائبه عزیزانی برخوردار بوده است که وظیفه خود می دانند از آنها شکر و قدردانی کنند.

جناب آقای دکتر مهدی علیپور حافظی

استاد راهنمای محترم که با وجود مشغله کاری زیاد با حوصله و وقت بسیار محقق را قدم به قدم در نگارش پایان نامه همراهی نموده و در تمامی مراحل از کوچکترین مساعدت و گلی مضائقه نداشتند.

سرکار خانم دکتر صحت مومنی

استاد مشاور محترم که افتخار شاگردی ایشان را داشته ام و بارانه نظرات ارزنده، گرانبه و عالی خویش اینجانب را صمیمانه در حفظ انجام این پژوهش یاری نمودند، کمال شکر و امتنان را دارم.

آقای دکتر حمیدرضا جالی مومنی

استاد فرهیخته که زحمت داوری این پایان نامه را بر عهده داشته و بارانه نظرات ارزنده خود موجبات اعتلای پایان نامه را فراهم ساخته اند.

جناب آقای مهندس جلالی منش

استاد محترم که بهواره راهنما و راه‌گشای نخوانده در اتمام و تکمال پایان نامه بوده است.

و همچنین از حمایت‌های معمولی اعضای کمیته پژوهش انجمن کتابداری و اطلاع‌رسانی، اعضای جلسات مشترک کمیته‌های انجمن کتابداری و اطلاع‌رسانی و همه عزیزانی که بنده را در نخواستن این پایان‌نامه یاری نمودند نهایت تشکر را دارم.

چکیده:

هدف پژوهش حاضر نگاشت نقشه علمی کتابداری و اطلاع رسانی براساس پایان نامه های کتابداری و اطلاع رسانی ایران از آغاز دوره کتابداری در ایران تا پایان سال ۹۰ است. رویکرد پژوهش حاضر، در بخش مرور پیشینه ها و ادبیات نظری، مطالعه کتابخانه ای، و در تجزیه و تحلیل یافته ها، روش پیمایش - تحلیلی است و از روش طبقه بندی کننده متن^۱ که از روش های متن کاوی است برای تعیین گرایشات موضوعی پایان نامه ها استفاده شده است. ابتدا اطلاعات پایان نامه های کتابداری جمع آوری شد که در نهایت اطلاعات ۱۴۷۹ پایان نامه گردآوری شد. در این پژوهش برای تفکیک پایان نامه ها از نظر موضوعی، طرح موضوعی LISA انتخاب شد. در پژوهش از روش NB و EM-NB برای طبقه بندی اتوماتیک استفاده شده است. برای تجزیه و تحلیل داده ها از نرم افزار Excel، MATLAB و از نرم افزار NodeXL برای ترسیم نقشه علمی استفاده شده است.

در این پژوهش برای ترسیم نقشه علمی کتابداری فرض بر این است که اگر مدرکی به دو رده تخصیص پیدا کرد، این مدرک ارتباط بین دو موضوع را نشان می دهد. ۶۰۰ پایان نامه از بین همه پایان نامه ها انتخاب شد و در اختیار متخصص موضوعی قرار گرفت تا رده بندی کنند. نقشه علمی برای همین مدارک که بصورت اتوماتیک دسته بندی شده بودند ترسیم شد و با نقشه بدست آمده از رده بندی مدارک توسط متخصص موضوعی مقایسه شد. مقایسه این دو باهم نشان می دهد که تفاوت کمی بین دو نقشه موجود است و اگر دقت الگوریتم در طبقه بندی مدارک افزایش یابد قطعاً نقشه علمی بدست آمده با این روش نیز قابل اتکا خواهد بود.

نتایج نشان می دهد که موضوع کتابخانه ها و مراکز منابع و مواد و منابع موضوعاتی هستند که حجم دانش زیادی را به خود اختصاص داده و هم توانسته اند روابط قوی با موضوعات دیگر برقرار کنند.

واژگان کلیدی: رشته کتابداری و اطلاع رسانی، نقشه علمی، الگوریتم نیوبیز، الگوریتم EM، طبقه بندی متون

¹ Text classification

فهرست مطالب

فصل اول: کلیات تحقیق

عنوان	صفحه
۱-۱ مقدمه	۶
۲-۱ بیان مساله	۷
۳-۱ اهمیت و ضرورت تحقیق	۹
۴-۱ اهداف تحقیق	۹
۵-۱ سوالات پژوهش	۱۰
۶-۱ متغیرهای تحقیق	۱۰
۷-۱ تعاریف مفهومی و عملیاتی	۱۰

فصل دوم: ادبیات تحقیق

۱-۲ مقدمه	۱۴
۲-۲ مبانی نظری	۱۵
۱-۲-۲ مفاهیم اساسی در سنجش علم	۱۵
۱-۱-۲-۲ کتاب‌سنجی	۱۶
۲-۱-۲-۲ علم‌سنجی	۱۷
۳-۱-۲-۲ اطلاع‌سنجی	۱۸
۴-۱-۲-۲ وب‌سنجی	۱۹
۲-۲-۲ نقشه علمی	۲۱
۳-۲-۲ تحلیل استنادی	۲۴
۴-۲-۲ متن کاوی	۲۶
۱-۴-۲-۲ تاریخچه متن کاوی	۲۸
۲-۴-۲-۲ روشهای متن کاوی	۳۰
۳-۴-۲-۲ روشهای مختلف دسته بندی	۳۵
۴-۴-۲-۲ فرایند متن کاوی	۳۹
۵-۴-۲-۲ روشهای رتبه‌بندی و وزن‌دهی واژگان متن	۴۱
۶-۴-۲-۲ بازیابی اطلاعات	۴۳
۸-۴-۲-۲ نرم‌افزار MATLAB	۴۵
۳-۲ پیشینه پژوهش	۴۵
۱-۳-۲ تحقیقات انجام شده در جهان	۴۶

۲-۳-۲ تحقیقات انجام شده در ایران ۴۸

فصل سوم: روش شناسی تحقیق

۱-۳ مقدمه ۵۸

۲-۳ روش پژوهش ۵۸

۳-۳ جامعه پژوهش ۵۹

۴-۳ روش و ابزار گردآوری اطلاعات ۵۹

۵-۳ روش تجزیه و تحلیل اطلاعات ۶۱

فصل چهارم: تجزیه و تحلیل داده ها

۱-۴ مقدمه ۶۸

۲-۴ توصیف داده ها ۶۸

۳-۴ یافته های تحقیق ۶۹

فصل پنجم: نتیجه گیری و پیشنهادات

۱-۵ مقدمه ۱۱۱

۲-۵ بحث درمورد یافته های پرسش های پژوهش ۱۱۱

۳-۵ نتیجه گیری ۱۱۶

۴-۵ محدودیت های پژوهش ۱۱۶

۵-۵ پیشنهادهای پژوهش ۱۱۷

۱-۵-۵ پیشنهادهای اجرایی ۱۱۷

۲-۵-۵ پیشنهاد برای پژوهش های آتی ۱۱۷

فهرست منابع ۱۱۸

پیوست I

فهرست جداول

- جدول ۴-۱. تعداد پایان‌نامه‌ها به تفکیک دانشگاه ۶۸
- جدول ۴-۲. توزیع طبقه بندی مدارک آزمون (رده بندی شده توسط متخصص موضوعی) ۷۰
- جدول ۴-۳. تعداد مدارک تخصیص یافته به هر رده در مرحله اول (تعیین رده مجموعه داده‌های آموزشی) ۷۱
- جدول ۴-۴. توزیع طبقه‌بندی پایان‌نامه‌ها (۱۴۷۹) با استفاده از الگوریتم NB با $T=80\%$ ۷۳
- جدول ۴-۵. توزیع طبقه‌بندی پایان‌نامه‌ها (۱۴۷۹) با استفاده از الگوریتم NB با $T=60\%$ ۷۴
- جدول ۴-۶. توزیع طبقه‌بندی پایان‌نامه‌ها (۱۴۷۹) با استفاده از الگوریتم NB با $T=50\%$ ۷۵
- جدول ۴-۷. دقت الگوریتم NB در طبقه بندی مدارک آزمون با T درصد مختلف ۷۶
- جدول ۴-۸. ارزیابی کارایی هر رده در طبقه‌بندی مدارک آزمون با استفاده از الگوریتم NB ۷۷
- جدول ۴-۹. توزیع طبقه‌بندی پایان‌نامه‌ها (۱۴۷۹) با استفاده از الگوریتم EM-NB با $T=80\%$ ۷۸
- جدول ۴-۱۰. توزیع طبقه‌بندی پایان‌نامه‌ها (۱۴۷۹) با استفاده از الگوریتم EM-NB با $T=60\%$ ۷۹
- جدول ۴-۱۱. توزیع طبقه‌بندی پایان‌نامه‌ها (۱۴۷۹) با استفاده از الگوریتم EM-NB با $T=50\%$ ۸۱
- جدول ۴-۱۲. دقت مدل در طبقه بندی مدارک آزمون با استفاده از الگوریتم EM-NB ۸۲
- جدول ۴-۱۳. مقایسه دقت مدل در طبقه بندی مدارک آزمون با الگوریتم NB و EM-NB با درصد مختلف T ۸۲
- جدول ۴-۱۴. ارزیابی دقت هر رده در طبقه بندی مدارک آزمون با استفاده از الگوریتم EM-NB ۸۳

فهرست اشکال

- شکل ۱-۲. نمودار روابط بین حوزه‌های کتاب‌سنجی، علم‌سنجی، اطلاع‌سنجی، وب‌سنجی و سایبرسنجی (بجورنبرون، ۲۰۰۴) ۲۰
- شکل ۲-۲. یک نمونه از ساخت مدل براساس داده‌های قدیمی ۳۳
- شکل ۳-۲. دسته‌بندی داده‌های وام ۳۴
- شکل ۱-۳. فلوچارت مراحل برچسب زدن داده‌های آموزشی ۶۴
- شکل ۱-۴. نمودار فراوانی پایان‌نامه‌های انجام شده به تفکیک هر پنج سال ۶۹
- شکل ۲-۴. نمودار مقایسه دقت مدل در طبقه بندی مدارک آزمون با الگوریتم NB و EM-NB با درصد مختلف T ۸۳
- شکل ۳-۴. نمودار مقایسه معیار F (میانگین دقت و بازیابی) برای رده‌ها در طبقه بندی متون با الگوریتم NB و EM-NB ۸۵
- شکل ۴-۴. نمودار توزیع طبقه بندی کل پایان‌نامه‌ها در رده های موضوعی لیزا ۸۶
- شکل ۵-۴. نمودار توزیع فراوانی پایان‌نامه‌های انجام شده در بازه زمانی (۱۳۴۶-۱۳۵۰) در همه رده‌ها ۸۸
- شکل ۶-۴. نمودار توزیع فراوانی پایان‌نامه‌های انجام شده در بازه زمانی (۱۳۵۱-۱۳۵۵) در همه رده‌ها ۸۹
- شکل ۷-۴. نمودار توزیع فراوانی پایان‌نامه‌های انجام شده در بازه زمانی (۱۳۵۶-۱۳۶۰) در همه رده‌ها ۹۰
- شکل ۸-۴. نمودار توزیع فراوانی پایان‌نامه‌های انجام شده در بازه زمانی (۱۳۶۰-۱۳۶۵) در همه رده‌ها ۹۱
- شکل ۹-۴. نمودار توزیع فراوانی پایان‌نامه‌های انجام شده در بازه زمانی (۱۳۶۶-۱۳۷۰) در همه رده‌ها ۹۲
- شکل ۱۰-۴. نمودار توزیع فراوانی پایان‌نامه‌های انجام شده در بازه زمانی (۱۳۷۱-۱۳۷۵) در همه رده‌ها ۹۳
- شکل ۱۱-۴. نمودار توزیع فراوانی پایان‌نامه‌های انجام شده در بازه زمانی (۱۳۷۶-۱۳۸۰) در همه رده‌ها ۹۴
- شکل ۱۲-۴. نمودار توزیع فراوانی پایان‌نامه‌های انجام شده در بازه زمانی (۱۳۸۱-۱۳۸۵) در همه رده‌ها ۹۵
- شکل ۱۳-۴. نمودار توزیع فراوانی پایان‌نامه‌های انجام شده در بازه زمانی (۱۳۸۶-۱۳۹۰) در همه رده‌ها ۹۶
- شکل ۱۴-۴. نقشه علمی کتابداری براساس رده‌بندی مدارک با نظر متخصص ۱۰۰
- شکل ۱۵-۴. نقشه علمی رشته کتابداری براساس مدارک رده‌بندی شده بصورت اتوماتیک ۱۰۵

فصل اول

کلیات تحقیق

۱-۱ مقدمه

دهه‌های ۱۹۵۰ و ۱۹۶۰ به دنبال افزایش میزان اطلاعات و گسترش تولیدات علمی و رواج روش‌شناسی پوزیتیویسم^۱، رویکردهای کمی برای سنجیدن میزان تولید اطلاعات علمی در حیطه‌های گوناگون علم وارد شد. در نتیجه، مباحث نوینی پایه‌ریزی شد که از ترکیب واژه‌های "سنجی" و "اسامی حوزه‌های شناخته شده علم مانند روان‌شناسی، زیست‌شناسی و غیره، به صورت روان‌سنجی، زیست‌سنجی و ... در آمدند. بدین ترتیب مباحثی چون کتابخانه‌سنجی، کتاب‌سنجی، علم‌سنجی، اطلاع‌سنجی و وب‌سنجی به علم وارد شدند (عابدی جعفری و همکاران، ۱۳۹۰). علم‌سنجی یکی از رایج‌ترین روش‌های ارزیابی فعالیت‌های علمی است. این روش در شوروی سابق پدید آمده و در کشورهای اروپای شرقی، بویژه مجارستان، برای اندازه‌گیری علوم در سطوح ملی و بین‌المللی استفاده شد. نخستین کسانی که واژه علم‌سنجی را ابداع کردند، دوبروف^۲ و کارنوا^۳ بودند، آنها علم‌سنجی را به عنوان اندازه‌گیری فرایند انفورماتیک تعریف کردند. انفورماتیک از نظر میخائیلوف^۴ عبارت است از: اصول علمی که به بررسی ساختار و ویژگی‌های اطلاعات علمی می‌پردازد و قوانین و فرایندهای این ارتباطات را مورد بحث قرار می‌دهد (سن گوپتا^۵، ۱۹۹۲). به دنبال مطرح شدن این علم، دانشمندان برجسته دیگری از جمله کول^۶ و ایلز^۷ و هولم^۸ نیز از مقالات علمی به عنوان ملاکی برای مقایسه تولید علم کشورهای مختلف استفاده کردند. آنها از این طریق تولیدات علمی کشورهای مختلف را از لحاظ کمی و کیفی با یکدیگر مورد مقایسه قرار داده و وضعیت کشورهای مختلف را در تولید اطلاعات علمی مشخص نمودند (عابدی جعفری و همکاران، ۱۳۹۰).

علم‌سنجی رشته‌ای است که تولیدات و ارتباطات و تعاملات علمی بین موجودیت‌های علمی را مورد مطالعه قرار می‌دهد. علم‌سنجی به دنبال سنجش برون‌داد علمی سازمان‌ها و افرادی است که در حوزه تولید علم فعالیت می‌کنند (لدسدروف^۹ و بسلر^{۱۰}، ۱۹۹۷). علم‌سنجی می‌تواند شکافهای علمی یک حوزه را بگشاید و پیشنهادهای برای رفع آنها ارائه دهد. لذا با استفاده از علم‌سنجی می‌توان به بررسی پژوهش‌ها در رشته‌ای خاص پرداخت و تولید علم آن را مورد ارزیابی قرار داد (شرق و همکاران، ۱۳۹۰).

¹. Positivism

². Dobrov

³. Karrennoi

⁴. Mikhailov

⁵. Sengupta

⁶. Cuil

⁷. eils

⁸. Holm

⁹. Leydesdorff

¹⁰. Besselaar

موضوع نقشه‌های علم، منبعث از دانش علم‌سنجی است. نقشه علمی عبارت است از تجزیه و تحلیل انتشارات یک حوزه علمی از زوایای مختلف و ترسیم نگرشی کلی از آن حوزه که بر پایه‌ی این نقشه و ترسیم سیر تغییر و تحولات، حوزه‌هایی که بیشترین و کمترین نزدیکی را دارند از هم متمایز می‌شوند. بدین ترتیب، هر کاربر، افزون بر ویژگی‌ها و ارتباط بین زیررده‌های هر حوزه از علم می‌تواند تاثیرگذارترین افراد و موسسات تحقیقاتی را نیز در آن حوزه خاص مشخص نماید. هدف از تهیه نقشه علمی، شناسایی نقاطی از دانش است که به اصطلاح، بحث داغ^۱ حوزه‌ی مربوط به خود را پیگیری می‌کنند (نایونز^۲، ۱۹۹۹).

در ساختار علوم، برخی از ارتباطات و پدیده‌ها به صورت انتزاعی برای ذهن قابل درک است، در صورتی که همین روابط به صورت فیزیکی برای چشم ملموس نیستند، پژوهشگران حوزه اطلاع‌رسانی در تلاش‌اند که روابط و پدیده‌های نامرئی موجود در ساختار علم را کشف نموده و بازیابی گرافیکی به صورت چند بعدی در قالب حوزه‌های موضوعی که باهم ارتباط بیشتری دارند در فاصله نزدیک‌تر و حوزه‌هایی که ارتباط کمتری دارند در فاصله بیشتر نمایش داده می‌شوند (نایونز، ۱۹۹۹).

لذا پژوهش حاضر سعی دارد با نگاشت نقشه علمی علوم کتابداری و اطلاع‌رسانی ایران با استفاده از پایان‌نامه‌های انجام شده در این رشته چشم‌اندازی کلی از رشته کتابداری و اطلاع‌رسانی در ایران ارائه دهد.

۱-۲ بیان مساله

نقشه علمی، بازنمونی فضایی از چگونگی پیوند رشته‌ها، حوزه‌ها، متخصصان و مقاله‌های آن‌ها به وجود می‌آورد، این نقشه‌ها را می‌توان به نقشه‌های جغرافیایی تشبیه کرد، که رابطه‌های سیاسی یا جنبه‌های فیزیکی را بر روی زمین نشان می‌دهند (اسمال^۳، ۲۰۰۰، ص. ۸۰۱). شبکه استنادی که در قالب این نقشه‌های فضایی نشان داده می‌شود، نموداری جهت‌دار و پیچیده است که راس‌های آن می‌توانند به ترتیب زمانی مرتب شوند و خط‌های مرزی موجود در این گراف‌ها راس‌های قدیمی را به راس‌های جدید وصل می‌کنند. این شبکه الگوهای ارتباطی و همچنین چگونگی همکاری علمی و روند استندهای ملی و جهانی پژوهشگران را نشان می‌دهند (اسمال، ۲۰۰۰، ص. ۸۰۲).

1. در حوزه‌های داغ، فعالیت بیشتری صورت می‌گیرد.

². Noyons

³. Smal

برای ترسیم ساختار علوم راه‌های گوناگونی وجود دارد که یکی از آن‌ها متن‌کاوی و تجزیه و تحلیل هم‌رخدادی واژگانی است که می‌توان از طریق آن روند شکل‌گیری حوزه‌های علمی یک حوزه را مورد بررسی قرار داد (بوهم^۱ و دیگران، ۲۰۰۲).

فرایند اصلی ترسیم نقشه‌های علم مبتنی بر نظر بورنر و همکارانش، شامل شش مرحله است که عبارتند از:

۱. نخستین گام در هر فرایند نگاشت یا ترسیم نقشه، استخراج اطلاعات مناسب است. در این مرحله استراتژی‌های مختلف جستجو کاربرد دارند.

۲. انتخاب واحدهای تحلیل، بستگی به سوالی دارد که درصدد پاسخ‌گویی به آن هستیم. رایج‌ترین واحدها برای ترسیم نقشه‌ها، نوشته‌ها هستند که عبارتند از مجلات، مدارک، نویسندگان و لغات و اصطلاحات توصیفگر.

۳. واژه‌های تکنیکی بسیاری به عنوان شاخص‌های شناسایی شباهت بین مقالات به کار برده می‌شوند، این واژه‌ها از پیشوندهای *co* و *inter* ساخته شده‌اند.

۴. شباهت‌های بین مدارک (واحدها) معمولاً با روش‌های مختلفی محاسبه می‌شود که رایج‌ترین آن‌ها عبارتند از: ارتباطات ارجاعی، شباهت‌های هم‌رخدادی، مدل بردار فضایی.

۵. روش‌های دسته‌بندی متنوعی با توجه به کاربرد هر یک در ترسیم نقشه‌ها موجود می‌باشد و مهم‌ترین آن‌ها تحلیل خوشه‌ای و مثلث‌بندی هستند.

۶. در آخرین مرحله نوبت به استفاده از فنون نمایش اطلاعات در قالب بصری می‌رسد (بورنر^۲ و همکاران، ۲۰۰۳).

پایان‌نامه و رساله، که گزارش پژوهش در دوره‌های کارشناسی ارشد و دکترای تخصصی هستند (شورای عالی برنامه ریزی، ۱۳۷۲) برای انتقال اطلاعات به مخاطبان خود تدوین می‌شوند و از این رو، نوعی مدرک به شمار می‌روند (باکلند^۳، ۱۹۹۷). پایان‌نامه کارشناسی ارشد و رساله‌های دکترای تخصصی، مدارکی هستند که در فرایند تحصیل در این مقاطع و در نتیجه همکاری میان دانشجویان و استادان و با

¹. Böhm

². Börner

³. Buckland

پشتیبانی موسسه‌های آموزش عالی پدید می‌آیند. داده‌های مجموعه‌ی این مدارک در کنار هم، به خودی خود، و همچنین در مقایسه با داده‌های منابع دیگر می‌توانند اطلاعات جدیدی را فراهم آورند.

رشته کتابداری اساساً ماهیتی میان‌رشته‌ای دارد، باید گفت که مطالعه‌های رشته کتابداری و اطلاع‌رسانی از سه زاویه‌ی موضوع تحقیق، رویکرد آن و روش‌شناسی مورد استفاده، میان رشته‌ای محسوب می‌شود. از سوی دیگر جامعه کتابداری و اطلاع‌رسانی در سال‌های اخیر با مباحث و دیدگاه‌های جدیدی روبرو است و بدیهی است به منظور تحول بنیادین این جامعه تمام اجزاء و لایه‌های آن باید دچار تحول شود. رشته کتابداری و اطلاع‌رسانی باید همچون سایر رشته‌ها از پایه‌های محکمی برخوردار باشد تا بتواند همواره بالنده حرکت کند، برای بالندگی نیاز است تا به مانند دیگر رشته‌ها در دو جنبه به تبیین و توجیه خود پردازد، اول از جنبه نظری و دوم از جنبه عملی و اجرایی.

بنابراین لازم است که جهت بهره‌وری بیشتر از پتانسیل‌های پژوهشی و همچنین مدیریت و برنامه‌ریزی در رشته کتابداری و اطلاع‌رسانی تصویری از ساختار رشته کتابداری و اطلاع‌رسانی در ایران برای پژوهشگران این رشته با هدف بررسی وضعیت موجود و همچنین به عنوان یک راهنما برای برنامه‌ریزی در آینده ترسیم شود. همچنین ضروری است، ساختار این علم به پژوهشگران جوان که در ابتدای راه پژوهش هستند جهت ایجاد فرصت‌های پژوهشی تازه شناسانده شود. با توجه به فرایند شش مرحله‌ای بورنر و همکارانش در این پژوهش سعی شده است تا نقشه علمی رشته کتابداری و اطلاع‌رسانی ایران با استفاده از پایان‌نامه‌های کارشناسی ارشد و دکترای رشته کتابداری انجام شده در ایران ترسیم شود.

۱-۳ اهمیت و ضرورت تحقیق

براساس این پژوهش، پژوهشگران رشته کتابداری و اطلاع‌رسانی خواهند توانست بر پایه پیشینه علمی رشته کتابداری در طی سال‌ها، به تقویت توانمندی‌ها و قلمروهای علمی پرکار و نیز زدودن ضعف‌ها و کاستی‌های موجود پردازند. همچنین، آن‌ها خواهند توانست از قلمروهای پرکار و نویسندگان برجسته به عنوان الگو استفاده نمایند تا دیگر قلمروها و نویسندگان نیز بتوانند به پر بارتر کردن نقشه علم این رشته و ارتقاء ساختار فکری آن کمک کنند.

۱-۴ اهداف تحقیق

هدف اصلی:

هدف از پژوهش حاضر تهیه نگاهی از وضعیت علمی موجود پایان‌نامه‌ها و موضوعات کلیدی رشته کتابداری و اطلاع‌رسانی بر اساس پایان‌نامه‌های رشته کتابداری و اطلاع‌رسانی در ایران است.

اهداف فرعی:

۱. تعیین میزان کارایی الگوریتم نیویز^۱ و EM-NB در طبقه‌بندی اتوماتیک اطلاعات پایان‌نامه‌های فارسی

۲. تعیین فراوانی پایان‌نامه‌های مورد مطالعه در رده‌های موضوعی لیزا

۳. تعیین ارتباط بین موضوعات در رشته کتابداری و اطلاع‌رسانی

۱-۵ سوالات پژوهش

۱. کارایی الگوریتم نیویز و EM-NB در طبقه‌بندی اتوماتیک اطلاعات پایان‌نامه‌های فارسی به چه میزان است؟

۲. فراوانی پایان‌نامه‌های رشته علوم کتابداری و اطلاع‌رسانی ایران در رده‌های موضوعی لیزا چگونه است؟

۳. توزیع فراوانی پایان‌نامه‌های انجام شده در هر رده در بازه‌های زمانی پنج ساله چگونه است؟

۴. وضعیت نقشه علمی رشته کتابداری و اطلاع‌رسانی در ایران براساس پایان‌نامه‌های اجرا شده چگونه است؟

۱-۶ متغیرهای تحقیق

متغیر مستقل: پایان‌نامه‌های کارشناسی ارشد و دکتری رشته کتابداری و اطلاع‌رسانی در ایران

متغیر وابسته: نقشه علمی حوزه علوم کتابداری و اطلاع‌رسانی

۱-۷ تعاریف مفهومی و عملیاتی

تعاریف متعددی از علم‌سنجی ارائه شده است که در اینجا به چند مورد اشاره می‌کنیم:

^۱ . Naïve Bayes

علم‌سنجی بخشی از جامعه‌شناسی علوم را شکل می‌دهد که اغلب مرتبط با سیاست‌گذاری‌های علمی است. علم‌سنجی درگیر با مطالعات کمی فعالیت‌های علمی و خصوصاً مطالعه‌ی انتشارات است و حوزه‌هایی را در برمی‌گیرد که فراتر از مقوله‌ی کتاب‌سنجی و اطلاع‌سنجی است (دایره‌المعارف بین‌المللی کتابداری و اطلاع‌رسانی، ۲۰۰۳).

علم‌سنجی حوزه‌ای از علم است که با جنبه‌های کمی افراد، گروه‌ها، موضوع‌ها و پدیده‌ها در علم و روابط آن‌ها با یکدیگر سروکار دارد، به طوری که در اصل، به محدوده یک رشته علمی خاص تعلق ندارد. وی هدف از علم‌سنجی را آشکارسازی ویژگی‌های پدیده‌های علم‌سنجی و فرایندهای موجود در پژوهش علمی برای مدیریت موثرتر علم معرفی می‌کند (وینکلر^۱، ۲۰۱۰، ص ۱-۲).

در این پژوهش تعریف عملیاتی علم‌سنجی متناظر با تعریف مفهومی دوم علم‌سنجی است.

نقشه علمی: نقشه علمی، بازنمونی فضایی از چگونگی پیوند رشته‌ها، حوزه‌ها، متخصصان و مقاله‌های آن‌ها به وجود می‌آورد، این نقشه‌ها را می‌توان به نقشه‌های جغرافیایی، که رابطه‌های سیاسی یا جنبه‌های فیزیکی را بر روی زمین نشان می‌دهند، تشبیه کرد (اسمال، ۲۰۰۰، ص ۸۰۱). در این پژوهش منظور از نقشه علمی، به تصویر کشیدن نتایج برآمده از تجزیه و تحلیل پایان‌نامه‌های کتابداری و اطلاع‌رسانی ایران و ترسیم یک نگرش کلی از این رشته است.

متن کاوی: متن کاوی، کاربرد داده کاوی در متون زبان طبیعی است. مهمترین نکته مرتبط این مورد، انتقال فنون عمومی داده کاوی به حوزه بازنمون متون است، که شامل مرحله قبل از پردازش و ویژگی‌های آماری خاص از داده‌های متنی است که زمینه تشکیل الگوریتم‌های خاص داده کاوی را باعث می‌شود (لئوپارد و دیگران، ۲۰۰۴). هدف اولیه متن کاوی، بازیابی اطلاعات از متون ساخت‌نیافته و همچنین ارائه دانش بصورت خالص برای کاربران در یک شکل چکیده است (آنانیودو^۲ و مک‌نوا^۳، ۲۰۰۶).

در این پژوهش منظور از متن کاوی کاوش در اطلاعات پایان‌نامه‌ها به منظور کشف روابط بین آن‌ها است. در این پژوهش از روش طبقه‌بندی متن^۴، که یکی از روش‌های متن کاوی است استفاده شده است.

^۱ vi.

^۲ . Ananiadou

^۳ . McNaught

^۴ . Text Classification

طبقه‌بندی متن: طبقه‌بندی متن به فرآیند تخصیص اتوماتیک یک یا چند سند به یکی از دسته‌های موضوعی است، که قبلاً تعریف شده است. در واقع طبقه‌بندی متن یکی از روش‌های شناسایی الگو برای استفاده موثرتر از اسناد است (توماس و همکاران، ۲۰۱۱).

در این پژوهش منظور از طبقه‌بندی متن اختصاص خودکار پایان‌نامه‌ها به رده‌های از پیش تعریف شده است.

الگوریتم بیز ساده (Naive Bayes): این الگوریتم ارائه کننده مدلی پیش‌بینی کننده در رابطه با احتمال خروجی نتایج خاصی است. الگوریتم NB ارتباط بین داده‌ها را با شمارش تعداد وقوع مشاهده آن‌ها اندازه‌گیری می‌نماید. این الگوریتم سپس مدلی را ایجاد می‌کند که نشان‌دهنده الگوها و ارتباط آن‌ها با یکدیگر است. پس از ایجاد مدل، می‌توان از آن به عنوان الگویی پیش‌بینی کننده استفاده نمود. (غضنفری و همکاران، ۱۳۸۷، ص ۱۸۷).

الگوریتم EM: الگوریتم EM یک روش محاسباتی عمومی برای برآوردهای حداکثر تحت داده‌های ناکامل است. نام این الگوریتم برگرفته از دو گام این الگوریتم است: گام E که محاسبه مقادیر مورد انتظار برای از دست رفته‌است و گام M که محاسبه برآوردهای حداکثر درستی‌نمایی پارامترها با فرض کامل بودن داده‌هاست که براساس دو مرحله پایه‌گذاری شد (دمپستر و همکاران، ۱۹۷۶).

فصل لوم

ادبيات تحقيق

تلاش‌های گارفیلد^۱ در ایجاد و راه اندازی موسسه اطلاعات علمی امریکا در فیلادلفیا سر منشا ظهور حوزه های مختلفی جهت تجزیه و تحلیل دانش گذشتگان در هر حوزه تخصصی از علم است. در نیم قرن اخیر مطالعات و پژوهش های زیادی در حوزه کتاب‌سنجی در دنیا صورت گرفته است. در ایران نیز این مطالعات در سالهای اخیر حجم قابل توجهی را به خود اختصاص داده اند. اکثر مطالعات کتاب‌سنجی و علم سنجی در کشور ما توسط متخصصان کتابداری و اطلاع رسانی صورت گرفته است.

آنچه که سوخت مطالعات سنجشی است، اطلاعاتی است که در منابع چاپی مانند مجلات، پایان نامه ها، کتاب ها و منابع الکترونیک نظیر وب و پایگاههای اطلاعاتی نهفته است. در همین رابطه، امروزه فناوری های اطلاعاتی و ارتباطی نقش مهمی را در تولید، انتقال و استفاده از دانش به عهده دارند. بخشی از مطالعاتی که امروزه در رابطه با اطلاعات و دانش صورت می گیرد در رابطه با دسترسی به اطلاعات می باشد (علیجانی و کرمی، ۱۳۸۷، ص ۱۳).

تلاش گارفیلد در ایجاد نمایه استنادی علوم در دهه ۱۹۵۰ تا حدودی برای رفع این مشکل بود که پژوهشگرانی که می خواهند در یک حوزه تازه شروع به تحقیق کنند باید علاوه بر اینکه از پژوهش های پیشین آگاهی داشته باشند لازم است که بتوانند در یک نظام رتبه‌بندی به ارزیابی بپردازند. حال تحلیل استنادی و مطالعات علم‌سنجی بخشی از این امر را برای پژوهشگران میسر می‌سازد ولی بخش دوم داشتن یک نمای کلی از حوزه تخصصی علم برای پژوهشگران هر حوزه است. از طرفی سیاست‌گذاری علم نیز جزء حوزه‌هایی است که در دنیای جدید با هدف برنامه‌ریزی و سیاست‌گذاری علوم مختلف توسعه پیدا کرده است. در این میان باید بین این دو حوزه از علم که در نقاطی باهم همپوشانی دارند ارتباط شفاف‌تری برقرار شود. یکی از حوزه‌هایی که می‌تواند ارتباط این دو حوزه را به راحتی مشخص نماید، داشتن دید و شمای کلی از یک حوزه خاص از علم در قالب نقشه علمی است. مفهومی که از نقشه به معنی عام در ذهن بسیاری از افراد وجود دارد چنین است که نقشه ارتباط بین کشورها، استانها، شهرها را نشان می‌دهد. اگر کمی به ساختارحوزه های مختلف علم توجه کنیم به راحتی متوجه می شویم که بین حوزه های مختلف ارتباطات و روابطی وجود دارد که می توان این روابط را در قالب نقشه علمی نشان داد.

¹.Garfield