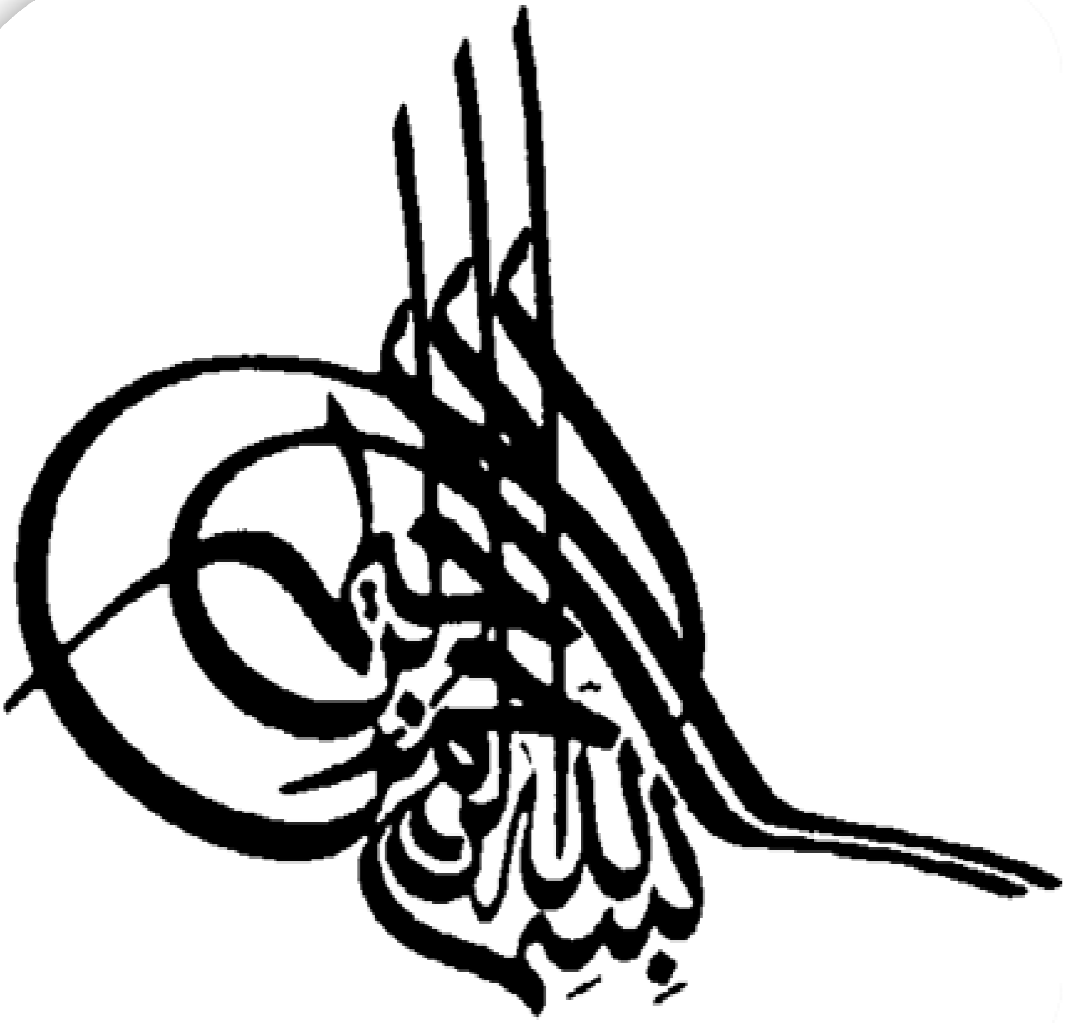


دانشگاه فردوسی مشهد
دانشکده مهندسی - گروه مهندسی کامپیوتر

پایان‌نامه کارشناسی‌ارشد

خوشه‌بندی اسناد متنی مبتنی بر مفاهیم همسایگی و شباهت معنایی

ملیحه دانش

استاد راهنما: جناب آقای دکتر محمود نقیب‌زاده

استاد مشاور: جناب آقای دکتر احد هراتی

تیر ۹۰

تقدیر و تشکر

خداوند را سپاس می‌گویم که به من فرصت داد تا عمر خود را در راه تحصیل علم و دانش سپری کنم و همواره استادانی دلسوز و فرزانه بر سر راهم قرار داد تا در این راه دراز و بی‌پایان علم‌جویی، راهنمای راهم و تسکین آتش سیری‌ناپذیرم باشند. امید آن که به یاد خورشید تابان راهم، شمع کوچکی بر سر راه تشنگان دیگر باشم.

با تشکر از

راهنمای دلسوز و فرزانه

استاد ارجمند

جناب آقای دکتر محمود نقیب‌زاده

مشاور و مشوق راه علم

جناب آقای دکتر احد هراتی

و همه کسانی که تاکنون مرا در راه رسیدن به اهدافم یاری نموده‌اند.

چکیده

خوشه‌بندی، روش داده‌کاوی قدرتمندی است که جهت کشف موضوع از اسناد متنی مورد استفاده قرار می‌گیرد. در این زمینه الگوریتم‌های خانواده k-means به دلیل سادگی و سرعت بالا، در خوشه‌بندی داده‌هایی با ابعاد بالا، کاربرد فراوانی دارند. در این الگوریتم‌ها، معیار شباهت cosine، تنها شباهت میان زوج اسناد را اندازه‌گیری می‌کند که در مواقعی که خوشه‌ها به خوبی تفکیک نشده باشند، عملکرد مناسبی ندارد. درمقابل، مفاهیم همسایگی و اتصال با در نظرگرفتن اطلاعات سراسری در محاسبه میزان نزدیکی دو سند، عملکرد بسیار بهتری دارند. چنانچه میزان شباهت دو سند از حد آستانه‌ای بیشتر باشد آن دو سند همسایه‌اند و تعداد همسایه‌های مشترک میان آنها، مقدار تابع اتصال این دو سند را نشان می‌دهد. بنابراین با توجه به اینکه تنها دو حالت همسایگی و عدم همسایگی داریم که با صفر و یک نمایش داده می‌شوند، مقداری از اطلاعات را در مورد میزان شباهت میان اسناد از دست می‌دهیم که منجر به کاهش دقت خوشه‌بندی حاصل می‌شود. جهت رفع این مشکل، در گام اول لیستی از مقادیر گسسته را برای تعیین بازه‌ای از مقادیر آستانه به جای تنها یک مقدار، در نظر گرفتیم که به دنبال آن درجات متفاوتی از همسایگی، بر اساس میزان شباهت میان اسناد خواهیم داشت. همچنین جهت افزایش هر چه بیشتر دقت نتایج حاصل، از منطق فازی نیز بهره برده و مقدار شباهت میان اسناد را با استفاده از مقادیر عضویت فازی نمایش دادیم. به این ترتیب میزان همبستگی میان اسناد را با استفاده از منطق فازی بهبود داده و گام جدیدی در کاربردهای منطق فازی برداشتیم.

همچنین در این مدل، روابط معنایی میان کلمات نادیده گرفته شده و تنها اسنادی با واژگان مشابه با یکدیگر مرتبط شده‌اند. در این پروژه پایانی از آنتولوژی WordNet جهت ایجاد مدل جدید نمایش اسناد بهره بردیم، بدین صورت که در آن از روابط معنایی به منظور وزن‌گذاری مجدد بسامد کلمات در مدل فضای برداری اسناد استفاده شده است. سپس مفاهیم همسایگی و اتصال را بر روی مدل حاصل اعمال نمودیم. نتایج حاصل از اعمال روش‌های پیشنهادی و ترکیبات آنها بر روی مجموعه داده‌های متن واقعی، حاکی از عملکرد موثر و مناسب‌تر الگوریتم پیشنهادی ما نسبت به روش‌های پیشین می‌باشد و می‌تواند جایگزین خوبی برای الگوریتم‌های پیشین در امر خوشه‌بندی اسناد باشد.

فهرست مطالب

فهرست جداول

فهرست اشکال

فهرست اختصارات به‌کار رفته در متن

AHC	Agglomerative Hierarchical Clustering
BKM	Bisecting KMeans
DC	Disceret Correlation
DL	Description Logic
FC	Fuzzy Correlation
KDT	Knowledge Discovery in Text
KM	KMeans
LCS	Least Common Subsummer
LSI	Latent Semantic Indexing
NLP	Natural Language Processing
SN	Semantic Neighbor
SVD	Singular Value Decomposition
SVSM	Semantic Vector Space Model
TF-IDF	Term Frequency-Inverse Document Frequency
VSM	Vector Space Model
WLC	Wordnet Lexical Categories

فصل ۱– ورود به مطلب

۱–۱– مقدمه

با ورود به دهه سوم جامعه جهانی وب، انقلاب متنی به عنوان یک تحول شگرف در قابلیت استفاده از اطلاعات موجود در اینترنت شده است. یافتن اطلاعاتی که تقریبا برای هر نیازی که پیش از این خودکار نبوده تنها با فشردن یک کلید و یا یک حرکت ماوس مقدور شده است. در مقایسه با انواع داده‌هایی که در پایگاه داده‌ها ذخیره می‌شوند، متن‌ها، داده‌هایی بدون ساختار، بی نظم و جهت بررسی الگوریتمی مشکل هستند. با این وجود در جامعه متمدن امروزی، متون به عنوان یکی از رایج‌ترین ابزارها جهت تبادل اطلاعات رسمی محسوب می‌شوند و داده‌های متنی بسیار پرکاربرد هستند. در حال حاضر بخش قابل توجهی از اطلاعات قابل دسترس در پایگاه داده‌های متنی (یا پایگاه داده‌های سند[۱]) که شامل مجموعه بزرگی از اسناد منابع مختلف (مثلا مقالات خبری، مقالات، کتاب‌ها، ایمیل‌ها و صفحات وب) ذخیره شده‌اند. پایگاه داده‌های متنی به علت افزایش مقدار اطلاعات موجود به مدل الکترونیکی سریع رشد می‌کنند. امروزه بیشتر اطلاعات در صنعت، کسب و کار و سازمان‌های دیگر به صورت الکترونیکی و در قالب پایگاه داده متنی ذخیره شده‌اند.

به دنبال افزایش داده‌های متنی، مدیریت و تحلیل آنها از اهمیت بسیار زیادی برخوردار است. داده‌های ذخیره شده در بیشتر پایگاه داده‌های متنی، داده‌های نیمه ساخت‌یافته هستند، چون نه به طور کامل غیرساخت یافته و نه به طور کامل ساخت یافته‌اند [Sha 2005]. برای مثال یک سند شامل تعدادی حوزه ساخت یافته مانند عنوان، نویسندگان، تاریخ انتشار، رده[۲] و ... و از طرف دیگر شامل برخی اجزای متنی غیرساخت‌یافته مانند چکیده و محتویات است. روش‌های بازیابی اطلاعات[۳] (IR) مانند (شاخص‌گذاری متن) برای کنترل کردن سندهای غیر ساخت یافته ایجاد شده‌اند. روش‌های بازیابی اطلاعات قدیمی برای مقدار زیادی داده متنی که به طور فزاینده‌ای افزایش می‌یابند، ناکارآمد هستند. بدون دانستن محتویات سندها، فرموله کردن پرس و جوهای مناسب برای تحلیل و استخراج کردن اطلاعات مفید از داده، مشکل است. کاربرها نیاز به ابزارهایی برای مقایسه سندهای مختلف، مرتب کردن سندها بر اساس مرتبط بودن

[1] Document databases
[2] category
[3] Information Retrieval

آن‌ها و یافتن الگوها دارند. بنابراین یکی از جدیدترین زمینه‌های مورد تحقیق در داده‌کاوی، "متن‌کاوی"[1] برای این منظور گسترش یافت.

۱-۲- متن کاوی

متن‌کاوی فناوری ایجاد شده جهت کنترل داده‌های متنی در حال رشد است که در جهت برچینی اطلاعات معنی‌دار از متون زبان طبیعی تلاش می‌کند. متن‌کاوی یعنی جستجوی الگوها در متن غیرساخت یافته و برای کشف خودکار دانش مورد علاقه یا مفید از متن نیمه ساخت‌یافته استفاده می‌شود [Tan 2005].

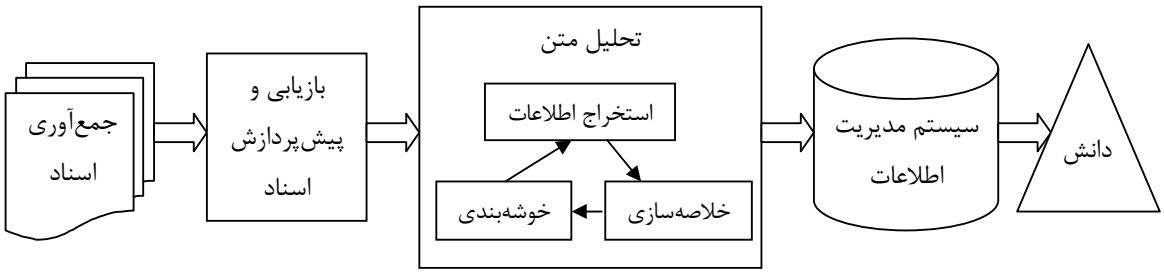
متن‌کاوی تقریبا معادل با تجزیه و تحلیل متون است که وظیفه آن استخراج اطلاعات با کیفیت بالا از متن می‌باشد [Kan 2007]. در موارد معدودی نیز به عنوان فرآیند تحلیل متن جهت استخراج اطلاعاتی که برای اهداف خاصی مفید هستند، تعریف می‌شود. در زمینه کاوش متن معمولا با متونی مواجه هستیم که وظیفه آنها ارتباط اطلاعات حقیقی و یا عقاید می‌باشد و هدف آن استخراج خودکار اطلاعات از چنین متونی می‌باشد، هر چند موفقیت جزئی حاصل گردد [Kan 2007].

به طور کلی متن‌کاوی جهت مشخص کردن سیستمی که بتواند حجم زیادی از متون زبان طبیعی را تحلیل کند و الگوهای مفید زبانی و لغوی را شناسایی کرده و به دنبال آن اطلاعات احتمالا مفید را استخراج کند، استفاده می‌شود [Fan 2005]. شکل ۱-۱ یک مدل کلی از یک کاربرد متن‌کاوی را نشان می‌دهد. این مدل با مجموعه‌ای از اسناد شروع می‌شود، یک ابزار متن کاوی، یک سند خاص را بازیابی و پیش‌پردازش می‌کند. سپس یک مرحله تحلیل متن انجام شده و در مواقعی از شیوه‌های مکرر تا استخراج اطلاعات استفاده می‌شود. سه روش تحلیل متن در این نمونه نشان داده شده اما بسیاری از ترکیبات دیگر نیز بر اساس اهداف سازماندهی می‌توانند استفاده شوند. اطلاعات حاصل می‌تواند در یک سیستم مدیریت اطلاعات قرار داده شود و در نهایت حجم وسیعی از دانش برای کاربر آن سیستم فراهم می‌شود [Fan 2005].

گاهی به جای واژه متن‌کاوی از عبارت "کشف دانش از متن" (KDT[2])، استفاده می‌شود [Sha 2005]. معمولا وظایف متن‌کاوی شامل طبقه‌بندی متن، خوشه‌بندی متن ، استخراج مفهوم، تحلیل معنایی،

[1] Text Mining
[2] Knowledge Discovery in Text

خلاصه‌سازی متن و مدل‌سازی روابط میان نهادها می‌باشد که در ادامه به طور خلاصه این مفاهیم را توضیح می‌دهیم.



شکل۱-۱- نمونه‌ای از متن‌کاوی [Fan 2005]

۱-۲-۱- تعاریف متن‌کاوی

متن‌کاوی یا کشف دانش از متن، از شیوه‌های بازیابی اطلاعات، استخراج اطلاعات و پردازش زبان طبیعی (NLP[1]) استفاده کرده و آنها را به الگوریتم‌ها و روش‌های داده‌کاوی[2]، یادگیری ماشین[3] و آماری[4] مرتبط می‌کند [Mar 2009]. با توجه به ناحیه‌های تحقیق گوناگون، بر هر یک از آنها می‌توان تعاریف مختلفی از متن‌کاوی در نظر گرفت که دو نمونه از آن را در زیر می‌آوریم:

متن‌کاوی = استخراج اطلاعات: در این تعریف، متن‌کاوی متناظر با استخراج اطلاعات در نظر گرفته می‌شود (استخراج واقعیت‌ها[5] از متن) [Wit 2000].

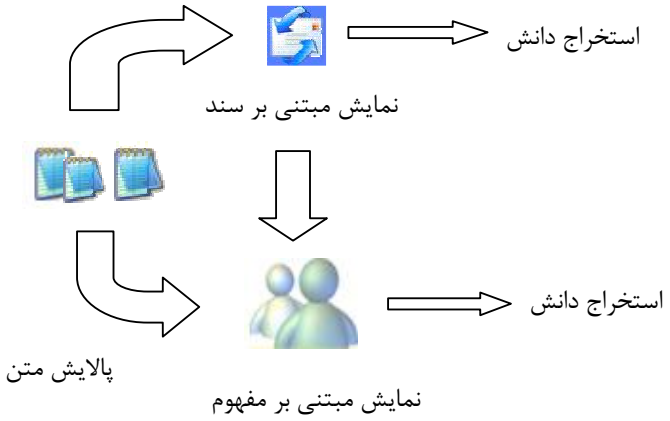
متن‌کاوی = کشف داده متنی: متن‌کاوی را می‌توان به عنوان روش‌ها و الگوریتم‌هایی از حوزه‌های یادگیری ماشین و آماری برای متن‌ها با هدف پیدا کردن الگوهای مفید در نظر گرفت. برای این هدف پیش پردازش کردن متون ضروری است. در بسیاری از روش‌ها، روش‌های استخراج اطلاعات، پردازش کردن زبان طبیعی یا برخی پیش پردازش‌های ساده برای استخراج داده از متون استفاده می‌شود. سپس می‌توان الگوریتم‌های داده‌کاوی را بر روی داده های استخراج شده اعمال کرد [Seb 2002].

[1] Natural language processing
[2] Data Mining
[3] Machine Learning
[4] Statistic
[5] Facts

در این پروژه پایانی بیشتر متن‌کاوی را به عنوان کشف داده متنی در نظر می‌گیریم و بیشتر بر روی روش‌های استخراج الگوهای مفید از متن برای دسته‌بندی مجموعه‌های متنی و یا استخراج اطلاعات مفید متمرکز می‌شویم.

۱-۳- مراحل اصلی فرآیند متن‌کاوی

متن‌کاوی فرآیندی است که شامل حوزه‌های تکنولوژیکی فراوانی است. بازیابی اطلاعات، داده‌کاوی، هوش مصنوعی و زبان‌شناسی محاسباتی، همگی حوزه‌هایی هستند که در این زمینه، نقش دارند. اما به طور کلی دو مرحله اصلی را در فرآیند متن‌کاوی می‌توان در نظر گرفت که در شکل ۱-۲ نشان داده شده است [Zhu 2004].



شکل ۱-۲- فرآیند متن کاوی [Zhu 2004]

اولین مرحله پیش پردازش مستندات است. خروجی این مرحله می تواند دو شکل مختلف داشته باشد:۱) مبتنی بر سند[1] ۲) مبتنی بر مفهوم[2]. در قالب نمایش مبتنی بر سند، آنچه که مهم است، نحوه‌ی نمایش بهتر برای مستندات است. مثلا تبدیل اسناد به یک قالب میانی و نیمه ساخت‌یافته[3]، یا به کار بردن یک شاخص بر روی آن‌ها یا هر نوع نمایش دیگری که کار کردن با اسناد را کاراتر می کند. هر موجودیت[4] در این نمایش در نهایت باز هم یک سند خواهد بود. در نوع دوم نمایش اسناد بهبود بخشیده می‌شود، مفاهیم و معانی موجود در سند و نیز ارتباط میان آنها و هر نوع اطلاعات مفهومی دیگری که قابل

[1] Document based
[2] Concept based
[3] Semi-Structured
[4] Entity

استخراج است، از متن استخراج می‌شود. در این نوع نمایش دیگر با مستندات به عنوان یک موجودیت مواجه نیستیم بلکه با مفاهیمی که از این مستندات استخراج شده‌اند، روبرو هستیم.

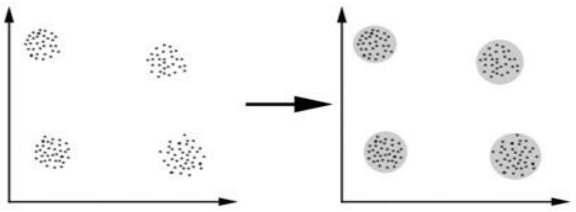
قدم بعدی استخراج دانش از این اشکال میانی نمایش اسناد است. بر اساس نحوه‌ی نمایش یک سند، روش استخراج دانش از یک سند متفاوت است. نمایش مبتنی بر سند، برای خوشه‌بندی، طبقه‌بندی، تجسم‌سازی و نظایر این‌ها استفاده می‌شود، در حالی که نمایش مبتنی بر مفهوم برای یافتن روابط میان مفاهیم، ساختن خودکار لغت‌نامه و آنتولوژی و نظایر آن بکار می‌رود [Zhu 2004].

۱-۴- کاربردهای متن‌کاوی

از جمله متداول‌ترین کاربردهای متن‌کاوی می‌توان موتورهای جستجو را نام برد که در آن کاربر یک عبارت یا کلمه (که ممکن است غلط املایی هم داشته باشد) را تایپ می‌کند و موتورهای جستجو توسط انباره بزرگی که از اسناد دارند، مرتبط‌ترین اسناد را پیدا می‌کنند [Zam 1997]. از جمله دیگر کاربردهای متن‌کاوی شامل طبقه‌بندی متن [Fab 2002]، خوشه‌بندی متن [Sal 1989]، استخراج مفهوم، تحلیل معنایی و خلاصه‌سازی متن [Raj 1997] را می‌توان نام برد. در این مستند، به کاربرد متن‌کاوی در خوشه‌بندی سند می‌پردازیم.

۱-۵- خوشه بندی

ما در جهانی پر از داده زندگی می‌کنیم. هر روزه انسان‌ها با حجم وسیعی از اطلاعات روبه‌رو هستند که باید آنها را ذخیره‌سازی کنند و یا نمایش دهند. یکی از روش‌های حیاتی کنترل و مدیریت داده‌ها، کلاس‌بندی یا گروه‌بندی داده‌های با خواص مشابه، درون مجموعه‌ای از دسته‌ها یا خوشه‌ها می‌باشد. خوشه‌بندی را می‌توان به عنوان مهمترین مسئله در یادگیری بدون نظارت در نظرگرفت. خوشه‌بندی با یافتن یک ساختار درون یک مجموعه از داده‌های بدون برچسب درگیر است. خوشه به مجموعه‌ای از داده‌ها گفته می‌شود که به هم شباهت داشته باشند. در خوشه‌بندی سعی می‌شود تا داده‌ها به خوشه‌هایی تقسیم شوند که شباهت بین داده‌های درون هر خوشه حداکثر و شباهت میان داده‌های درون خوشه‌های متفاوت حداقل شود [Ros 2008]. نمونه‌ای از آن در شکل ۱-۳ آمده است.



شکل۱-۳- نمونه‌ای از اعمال خوشه‌بندی روی یک مجموعه از داده‌ها که از معیار فاصله به عنوان عدم شباهت بین داده‌ها استفاده شده است[Ros 2008]

۱-۶- هدف از خوشه‌بندی

دسته‌بندی، جزئی از طبیعت انسان است. چیزهای محدودی وجود دارند که برای بقای زندگی بسیار مهم و ضروری هستند. تاریخ بشر به تاریخچه سرعت رشد دانش و پیاده‌سازی روش‌های تقسیم اشیاء به دسته‌های معلوم و قابل درک نیز برمی‌گردد. امروزه جهت بهبود توانمندی در دسته بندی، بشر بسیاری از روش‌های هوشمند را توسعه داده است. دو نمونه از موثرترین آنها شامل زبان‌های نوشتاری و کامپیوتر می‌باشد. با استفاده از هر دوی این امکانات ساختارهایی را توسعه داده و رده‌بندی‌های جدیدی را ایجاد کردیم.

در حال حاضر از کامپیوترها جهت توسعه و سازماندهی اطلاعات زیادی که یک نمونه از آنها به شکل متون هستند، استفاده می‌کنیم. ما انسان‌ها متون را به روش‌های مختلفی دسته‌بندی می‌کنیم. در اغلب موارد یک فرد ممکن است از یک مجموعه متون یکسان افرازهای بسیار متمایز و درعین حال مجاز و ارزشمندی را تولید کند. این امر ضرورتا چیز بدی نیست. هر افراز جدید از یک مجموعه از متون در صورتی که با استدلال درستی ایجاد شده باشد، ممکن است به ما دیدگاه و بصیرت جدیدی بدهد. افرازهای متون ممکن است در مواقعی منسوخ و بی‌ربط باشند. همچنین متون جدید ممکن است در یک ساختار قدیمی جای نگیرند یا اینکه ممکن است مجبور به ایجاد تغییراتی در یک ساختار شوند. ساخت یک افراز جدید به طور دستی بسیار گران و زمانبر است. ابزارهای خودکاری که متون را افراز می‌کنند، یا گروه‌های منطقی و معقولی را از متون استخراج کنند، می‌توانند بسیار ارزشمند باشند. حتی اگر افرازها و یا اجزای تولید شده از ابزارهای خودکار، نتایج بدتری از آنچه که به طور دستی و توسط بشر جمع‌آوری شده، داشته باشند باز هم ارزشمند هستند چرا که در اغلب موارد کسی تاکنون یک افراز به طور دستی ایجاد نکرده است. از طرفی، اطلاعات مورد جستجو در محیط‌های الکترونیکی چنانچه توسط کاربرانی هدایت و بازیابی شوند، می‌توانند بهتر بازیابی شوند.

از یک دیدگاه، آنچه که به عنوان یک افراز خوب از یک مجموعه از متون محسوب می‌شود وابسته به استدلالی است که در ایجاد آن افراز استفاده شده که آیا صحیح است یا خیر و اینکه آیا به روش سازگاری استفاده شده یا خیر. در زمینه سازگاری، کامپیوتر بسیار ممتاز است اما استدلال باید به طریقی توسط انسان تامین شود. خوشه‌بندی با استفاده از روشی موثر در نمایش و تجسم اسناد و گروه‌بندی اسناد مشابه و مرتبط با هم و تشکیل مجموعه‌ای از اسناد، در این راستا گام موثری برداشته است. خوشه‌بندی روشی مفید جهت سازماندهی حجم وسیعی از اسناد متنی نامنظم در داخل تعداد محدودی از خوشه‌های معنی‌دار و منسجم می‌باشد که به موجب آن مکانیزم‌هایی جهت مرور و هدایت آگاهانه اطلاعات را تامین می‌کند.

۱-۷- کاربردهای خوشه‌بندی

خوشه‌بندی در زمینه‌های بسیاری کاربرد دارد،م در زیر لیستی از مهم‌ترین آنها را یادآور می‌شویم:

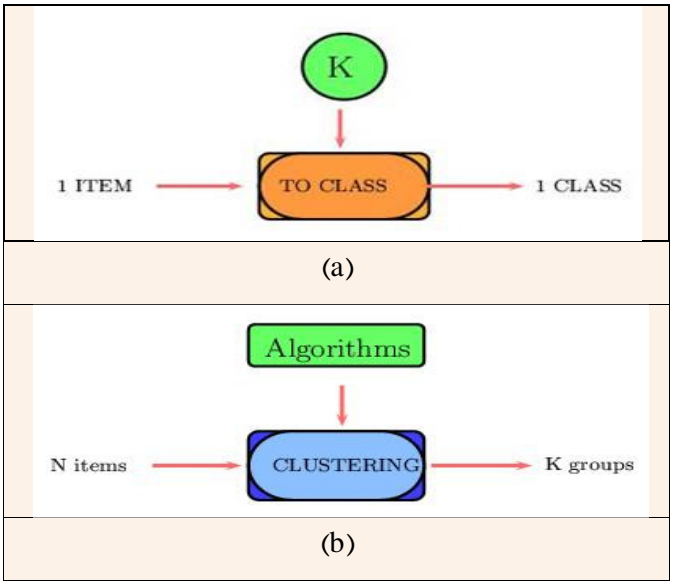
- در زمینه مهندسی (یادگیری ماشین، هوش مصنوعی، تشخیص الگو، مهندسی مکانیک و الکترونیک)
- علوم کامپیوتر (کاوش وب، تحلیل پایگاه داده فضایی، جمع آوری مستندات متنی، تقسیم بندی تصویر)
- علوم پزشکی (ژنتیک، زیست شناسی، میکروب شناسی، فسیل شناسی، روان شناسی، بالین، آسیب شناسی)
- علوم زمین‌شناسی (جغرافیا، زمین‌شناسی، نقشه برداری از زمین)
- علوم اجتماعی (جامعه شناسی، روان‌شناسی، تاریخ، آموزش وپرورش)
- اقتصاد (بازاریابی، تجارت)

همچنین خوشه‌بندی ممکن است با نام‌های دیگری از قبیل علم رده‌بندی عددی [Gre 1996] و یادگیری بدون معلم (یا یادگیری بدون نظارت[1]) [Wil 1988] بکار برده شود. امروزه، خوشه‌بندی نقش حیاتی در سازماندهی مجموعه‌های بزرگ مستندات متنی مانند وب دارد که در ادامه به آن می‌پردازیم.

[1] Unsupervised Learning

۱-۸- خوشه‌بندی در مقابل طبقه‌بندی

با طبقه‌بندی خودکار، به ماشین این توانایی را می‌دهیم که تصمیم بگیرد یک متن، به کدام یک از مجموعه دسته‌های از پیش تعریف شده تعلق دارد. در خوشه‌بندی، ماشین تصمیم می‌گیرد که چطور یک مجموعه متن داده شده را باید تقسیم‌بندی کند. طبقه‌بندی زمانی مناسب است که شخص بخواهد متون جدیدی را بر طبق یک دسته‌بندی معلوم و مشخصی دسته‌بندی کند. خوشه‌بندی زمانی است که شخص بخواهد ساختارهای جدیدی را که پیش از این شناخته نشده بود، کشف کند. هر دو روش ممکن است که نتایج جالبی را بر روی داده‌های متنی نامعلوم ارائه بدهند ، طبقه‌بندی آنها را مطابق با یک ساختار معلوم مرتب می‌کند و خوشه‌بندی ساختار این مجموعه معین را نمایش می‌دهد. در طبقه‌بندی هر داده به یک طبقه (کلاس) از پیشین مشخص شده تخصیص می‌یابد ولی در خوشه‌بندی هیچ اطلاعی از کلاس‌های موجود درون داده‌ها وجود ندارد و به عبارتی خود خوشه‌ها نیز از داده‌ها استخراج می‌شوند [Ros 2008].



شکل۱-۴- a) در طبقه‌بندی با استفاده یک سری اطلاعات اولیه داده‌ها به دسته‌های معلومی نسبت داده می‌شوند. b) در خوشه‌بندی داده‌ها با توجه به الگوریتم انتخاب شده به خوشه‌هایی نسبت داده می‌شوند [Ros 2008].

بنابراین برخلاف طبقه‌بندی، در خوشه‌بندی گروه‌ها از قبل مشخص نمی‌باشند و همچنین معلوم نیست که برحسب کدام خصوصیات گروه‌بندی صورت می‌گیرد. در نتیجه پس از انجام خوشه‌بندی باید یک فرد خبره خوشه‌های ایجاد شده را تفسیر کند و در بعضی مواقع، لازم است که پس از بررسی خوشه‌ها بعضی از پارامترهایی که در خوشه‌بندی در نظر گرفته شده‌اند ولی بی‌ربط بوده یا اهمیت چندانی ندارند حذف