بسم الله الرحمن الرحيم

**دانـــشگاه یـــزد**

**دانشکده مهندسی برق و کامپیوتر**

**گروه مهندسی کامپیوتر**

**پایان‌نامه**

**برای دریافت درجه کارشناسی ارشد**

**مهندسی کامپیوتر – هوش مصنوعی**

# رتبه‌بندی تصاویر بر مبنای پرس‌وجوی کاربر

استاد راهنما

**دکتر علی‌محمد زارع‌بیدکی**

استاد مشاور

**دکتر مهدی رضائیان**

پژوهش و نگارش

**سید کاظم میردهقان**

شهریورماه ۱۳۹۲

**تقدیم به**

روح پاک **پدرم** که به من آموخت تا چگونه در عرصـه زندگی، ایستادگی را تجربه کنم به **مادرم**، دریای بی‌کران فداکاری و عشـــق که وجودم برایش همه رنج بود و وجودش برایم همه مهر و به **همسرم**، اسطوره زندگی‌ام، پناه خستگی‌ام و امید بودنم.

## تقدیر و تشکر

شکر شایان نثار ایزد منان که توفیق را رفیق راهم ساخت تا این پایان‌نامه را به پایان برسانم. سپس از پدر و مادر که نخستین آموزگارانم بودند تا معلمین، دبیران و استادانم که هر یک به‌زعم خویش بر آموخته‌هایم افزودند، قدردانی می‌نمایم. مراتب سپاس خود را از جناب آقای دکتر علی‌محمد زارع‌بیدکی به‌عنوان استاد راهنما که همواره نگارنده را مورد لطف و محبت خود قرار داده‌اند، ابراز می‌دارم. همچنین از جناب آقای دکتر مهدی رضائیان، استاد مشاور اینجانب که از رهنمودهای ایشان در این پژوهش بهره برده‌ام، کمال تشکر را دارم.

# چکیده

جستجو در اینترنت کنونی به امری پیچیده تبدیل شده است. دلیل این امر رشد روزافزون در تصاویر موجود در وب می‌باشد. برای جستجوی تصاویر دلخواه کاربر، به یک موتور جستجوی[۱] کارامد نیاز است تا در کمترین زمان ممکن، متناسب‌ترین تصاویر با پرس‌وجوی[۲] کاربر را به او ارائه نماید. چالش اصلی موتورهای جستجو، رتبه‌بندی[۳] تصاویر بازیابی شده از پرس‌وجوی کاربر است. برای رتبه‌بندی تصاویر الگوریتم‌های مختلفی وجود دارد که می‌توان آن‌ها را به پنج دسته الگوریتم‌های مبتنی بر محتوای متنی[۴]، الگوریتم‌های مبتنی بر محتوای دیداری[۵]، الگوریتم‌های مبتنی بر اتصال[۶]، الگوریتم‌های مبتنی بر رفتار کاربر[۷] و الگوریتم‌های ترکیبی[۸] تقسیم نمود. برخی از این الگوریتم‌ها از دقت کافی برخوردار نیستند و برخی دیگر دچار درجه‌پیچیدگی بالا در پردازش تصویر می‌باشند.

هدف اصلی این پژوهش، مروری بر الگوریتم‌های رتبه‌بندی تصاویر بازیابی شده از پرس‌وجوی متنی کاربر و بررسی روش عملکرد و نقاط ضعف و قوت هر کدام می‌باشد. در ادامه روش جدیدی برای رتبه‌بندی تصاویر برای استفاده در موتورهای جستجو پیشنهاد می‌گردد که ترکیبی از رتبه‌بندی بر اساس محتوای متنی و به کارگیری محتوای دیداری تصاویر به منظور ادغام اسناد با تصاویر مشابه می‌باشد. این ادغام باعث غنی‌تر شدن محتوای متنی اسناد تصویری شده و دقت جستجو را افزایش می‌دهد. سپس با استفاده از یک الگوریتم یادگیری جفتی[۹] و مجموعه داده محک تولیدی که توسط افراد خبره برچسب‌گذاری شده است، وزن فیلدهای متنی مورد جستجو را تنظیم می‌کنیم. نتایج آزمایش‌ها نشان می‌دهد این روش می‌تواند دقت رتبه‌بندی را تا حد بالایی افزایش دهد.



**کلمات کلیدی:** تصویر، رتبه‌بندی، روش ترکیبی، موتور جستجو


---

[1] Search engine
[2] Query
[3] Ranking
[4] Text-based
[5] Content-based
[6] Link-based
[7] User behavior
[8] Combinational
[9] Pair-wise Learning

# فهرست مطالب


۱ مقدمه ..... ۱

۱-۱ معماری موتور جستجوی تصویر ..... ۲

۱-۲ رتبه‌بندی تصاویر در وب ..... ۵

۱-۳ چالش‌های رتبه‌بندی تصاویر در وب ..... ۷

۱-۳-۱ مفهوم مرتبط بودن ..... ۷

۱-۳-۲ توصیفگرهای تصویر ..... ۸

۱-۳-۳ وابستگی به کاربر ..... ۸

۱-۳-۴ شباهت یا تطبیق ..... ۹

۱-۳-۵ نمایه‌ها ..... ۹

۱-۳-۶ نوع پرس‌وجو ..... ۹

۱-۳-۷ بازیابی تصاویر ..... ۱۰

۱-۴ ارزیابی ..... ۱۱

۱-۴-۱ مجموعه داده محک ..... ۱۱

۱-۴-۲ معیارهای ارزیابی ..... ۱۲

۱-۵ ساختار پایان‌نامه ..... ۱۳

۲ مروری بر کارهای گذشته ..... ۱۴

۲-۱ روش‌های مبتنی بر محتوای متنی ..... ۱۴

۲-۱-۱ مدل فضای برداری ..... ۱۵

۲-۱-۲ مدل احتمالی ..... ۱۷

۲-۲ روش‌های مبتنی بر محتوای دیداری تصاویر ..... ۱۹

۲-۲-۱ الگوریتم VisualRank ..... ۲۱
۲-۳ روش‌های مبتنی بر اتصال ..... ۲۳
2-3-1 الگوریتم PageRank ..... ۲۴
۲-۳-۲ الگوریتم HITS ..... ۲۶
۲-۳-۳ الگوریتم PicASHOW ..... ۲۹
۲-۴ روش‌های مبتنی بر رفتار کاربر ..... ۲۹
۲-۵ روش‌های ترکیبی ..... ۳۰
۳ روش پیشنهادی ..... ۳۱
۳-۱ مقدمه ..... ۳۱
۳-۲ الگوریتم پیشنهادی ..... ۳۱
۳-۲-۱ گام اول: شناسایی تصاویر یکسان و مقیاس‌شده ..... ۳۳
۳-۲-۲ گام دوم: ادغام اسناد تصویری یکسان و مقیاس شده و نمایه‌سازی آن‌ها ..... ۳۷
۳-۲-۳ گام سوم: فرآیند جستجوی تصویر ..... ۴۲
۳-۳ پیاده‌سازی روش پیشنهادی و ایجاد مجموعه داده محک ..... ۴۲
۳-۴ بهبود روش پیشنهادی با روش یادگیری جفتی ..... ۴۴
۳-۵ ارزیابی روش‌های پیشنهادی ..... ۴۶
۴ نتیجه‌گیری و کارهای آینده ..... ۴۹
۴-۱ نتیجه‌گیری ..... ۴۹
۴-۲ دستاوردهای پایان‌نامه ..... ۴۹
۴-۳ کارهای آینده ..... ۵۰
پیوست الف. تعریف واریانس و میانگین ..... ۵۲

شاخص ..... ۶۷
مراجع ..... ۷۰

# فهرست شکل‌ها


شکل ۱-۱: معماری یک موتور جستجوی تصویر ..... ۳

شکل ۱-۲: مثالی از نمایه‌ساز معکوس ..... ۴

شکل ۱-۳: فرآیندهای برخط و برون‌خط موتور جستجو ..... ۵

شکل ۱-۴: تصویر به عنوان پرس‌وجو در موتور جستجوی گوگل ..... ۱۰

شکل ۲-۱: دسته‌بندی روش‌های رتبه‌بندی ..... ۱۴

شکل ۲-۲: گراف شباهت در الگوریتم VisualRank ..... ۲۲

شکل ۲-۳: مثالی از الگوریتم PageRank ..... ۲۵

شکل ۲-۴: شمای کلی رتبه‌بندی مبتنی بر رفتار کاربر ..... ۳۰

شکل ۳-۱: ادغام اسناد تصویری یکسان و مقیاس شده ..... ۳۲

شکل ۳-۲: روندنمای الگوریتم تولید شناسه برای هر تصویر ..... ۳۴

شکل ۳-۳: نمونه‌ای از یک تصویر رقمی خاکستری ..... ۳۵

شکل ۳-۴: نمونه‌ای از یک تصویر رقمی رنگی ..... ۳۵

شکل ۳-۵: محاسبه شدت نور پیکسل تصویر خاکستری ..... ۳۶

شکل ۳-۶: تولید کد درهم از تصویر خاکستری ..... ۳۷

شکل ۳-۷: نمونه‌ای از تصویر درون‌هاست ..... ۳۸

شکل ۳-۸: نمونه‌ای از تصویر درون‌دامنه ..... ۳۹

شکل ۳-۹: نمونه‌ای از تصویر خارجی ..... ۳۹

شکل ۳-۱۰: سطوح مختلف محتوای متنی اطراف تصویر ..... ۴۰

شکل ۳-۱۱: نحوه قرار گرفتن تصویر در صفحه وب ..... ۴۱

شکل ۳-۱۲: واسط کاربر سامانه ایجاد مجموعه داده محک ..... ۴۳

شکل ۳-۱۳: ارزیابی روش پیشنهادی پیش و پس از تنظیم وزن‌ها در معیار P@n.....۴۷

شکل ۳-۱۴: ارزیابی روش پیشنهادی پیش و پس از تنظیم وزن‌ها در معیار دقت متوسط۴۸

# فهرست جدول‌ها

جدول ۲-۱: متغیرهای استفاده‌شده در روش BM25[۵]..... ۱۹

# ۱ مقدمه

استفاده از وب به‌عنوان منبع اصلی کسب اطلاعات، با رشد شدیدی همراه بوده است. دلیل این امر، رفع نیاز اطلاعاتی[۱] کاربران می‌باشد. کاربران نیاز دارند تا در کمترین زمان ممکن به اطلاعات مورد نیاز خود دست یابند. از طرف دیگر، اطلاعات موجود روی وب با رشدی نمایی همراه بوده و در قالب‌های مختلف ظاهر شده‌اند. این اطلاعات با سرعت بالایی در حال تغییر می‌باشند. از این رو موتورهای جستجو، ابزارهایی کارامد در این زمینه هستند. کاربران، پرس‌وجوی خود را توسط تعداد محدودی کلیدواژه[۲] مطرح کرده و موتور جستجو مرتبط‌ترین و باکیفیت‌ترین نتایج با پرس‌وجوی کاربر را به او ارائه می‌کند. بر اساس آمار سایت معتبر alexa، موتورهای جستجوی گوگل[۳] و یاهو[۴] به ترتیب در مقام‌های اول و چهارم پر بازدیدترین پایگاه‌های اینترنتی جهان[۱] و مقام‌های اول و دوم پر بازدیدترین پایگاه‌های ایران قرارگرفته‌اند[۲]. بر اساس آمار منتشرشده در همین سایت، از نظر رتبه‌بندی ترافیک منطقه‌ای[۵]، ایران رتبه اول ترافیک جهانی گوگل[۳] و رتبه دوم ترافیک جهانی یاهو[۴] را در اختیار دارد.

یکی از این داده‌های رقمی، تصویر می‌باشد. تصویر بخش بزرگی از اطلاعات وب را شامل می‌شود. در بازیابی تصویر[۶]، هدف یافتن تصاویر مرتبط با پرس‌وجوی کاربر خاص، از میان مجموعه‌ای معمولاً بزرگ از تصاویر است. بازیابی می‌تواند بر روی مجموعه‌های خاص‌منظوره[۷] و یا از بین مجموعه‌ای عمومی از تصاویر انجام پذیرد. به عنوان نمونه بازیابی تصاویر ماهواره‌ای و بازیابی تصاویر پزشکی در دسته تصاویر خاص‌منظوره قرار می‌گیرند. کاربر می‌تواند پرس‌وجوی خود را توسط تعدادی کلیدواژه و یا یک تصویر به موتور جستجو اعلام کند. ولی در اینجا منظور از پرس‌وجو، یک عبارت متنی متشکل از تعدادی کلیدواژه می‌باشد. پس از این که کاربر پرس‌وجوی خود را توسط تعدادی

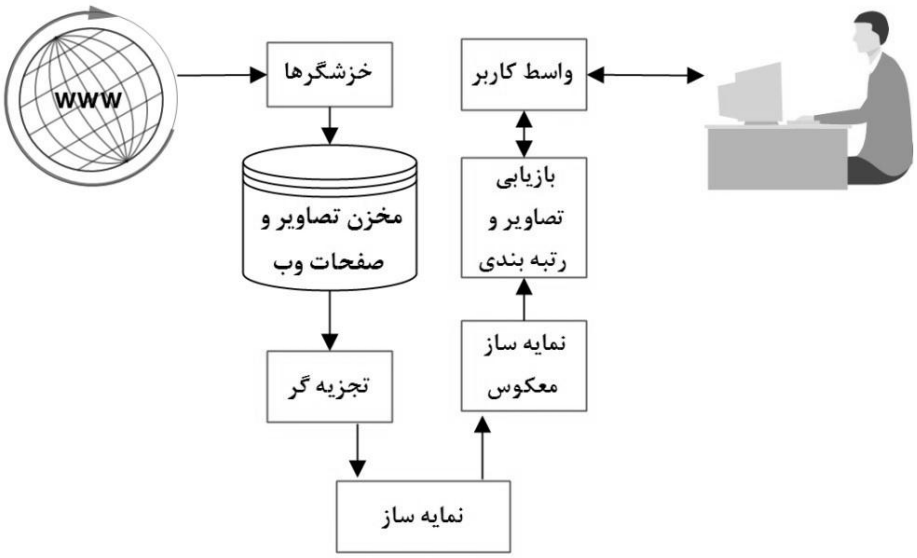
[1] Information need
[2] Keyword
[3] Google
[4] Yahoo
[5] Regional traffic ranks
[6] Image Retrieval
[7] Domain Specific

کلیدواژه به موتور جستجو اعلام می‌کند، موتور جستجو در پاسخ به پرس‌وجوی کاربر ممکن است تعداد زیادی نتیجه را به‌عنوان نتیجه مرتبط کاندید کند. ولی کاربر معمولاً به ۱۰ تا ۲۰ نتیجه اول اکتفا می‌کند[۵] و درصورتی‌که به هدف خود نرسد، پرس‌وجوی خود را تغییر داده و یا از جستجو منصرف می‌شود. در نتیجه رتبه‌بندی نتایج بازیابی شده از اهمیت بالایی برخوردار است تا جایی که می‌تواند میزان محبوبیت[1] یک موتور جستجو را تعیین کند. به همین منظور در این پایان‌نامه، تمرکز ما بیشتر بر روی الگوریتم‌های رتبه‌بندی خواهد بود. ولی برای آشنایی با چگونگی عملکرد یک موتور جستجو، ابتدا ساختار یک موتور جستجوی تصویر را باهم بررسی خواهیم نمود و سپس به شرح الگوریتم خواهیم پرداخت.

## ۱-۱ معماری موتور جستجوی تصویر

در شکل ۱-۱، معماری یک موتور جستجوی تصویر ساده را مشاهده می‌کنید. در مرحله اول، یک یا چند خزش‌گر[2] به‌صورت موازی صفحات وب و تصاویر داخل هر صفحه را خزش نموده و آن‌ها را در مخزن[3] تصاویر و صفحات وب ذخیره می‌کند. فرآیند خزش به صورت کاملاً هوشمند[4] و خودکار[5]، انجام می‌گیرد. خزش‌گر، اسناد وب را بر مبنای پوشش[6] و تازگی[7] جمع‌آوری می‌کند. پوشش به معنای این است که خزش‌گر بتواند هرچه بیشتر از صفحات و تصاویر موجود روی وب را پوشش دهد. از طرف دیگر یک خزش‌گر باید بتواند صفحات و تصاویر را به‌روز نگه دارد. به دلیل حجم بالا، اطلاعات خزش شده به صورت فشرده[8] ذخیره می‌شود.

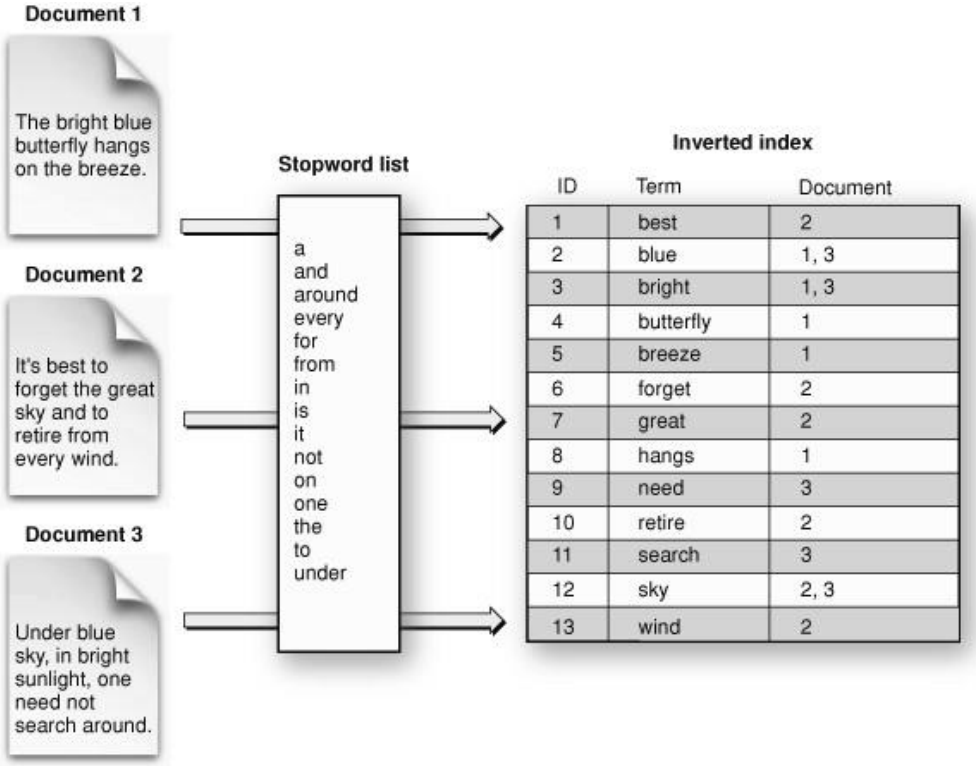
[1] Popularity
[2] Crawler
[3] Repository
[4] Intelligent
[5] Automatic
[6] Coverage
[7] Freshness
[8] Compressed




**شکل ۱-۱: معماری یک موتور جستجوی تصویر**

پس از این مرحله، محتوای صفحات توسط تجزیه‌گر[۱]، استخراج‌شده و پیش‌پردازش[۲]هایی روی آن انجام می‌گیرد. صفحات پیش‌پردازش‌شده به نمایه‌ساز[۳] تحویل داده‌شده و کلمات داخل صفحات استخراج‌شده و ساختمان داده‌ای به نام نمایه[۴] ایجاد می‌شود. نمایه نشان می‌دهد در هر صفحه چه واژگانی وجود دارد. ولی اکثر موتورهای جستجو از نمایه معکوس[۵] استفاده می‌کنند. در نمایه معکوس، مشخص می‌شود که هر واژه در چه اسنادی[۶] آمده است. با استفاده از نمایه معکوس قادریم تا با سرعت بالا، اسنادی که حاوی یک کلمه خاص هستند را بیابیم. نمایه‌ساز معکوس از دو قسمت تشکیل شده است. قسمت واژگان، که شامل تمام واژگان داخل اسناد می‌باشد و لیست رخداد مکانی این واژگان، که شامل تمام اسنادی است که شامل آن واژه هستند. شکل ۱-۲ مثالی از نمایه‌سازی معکوس برای سه سند را نمایش می‌دهد.

[1] Parser
[2] Preprocessing
[3] Indexer
[4] Index
[5] Inverted index
[6] Documents



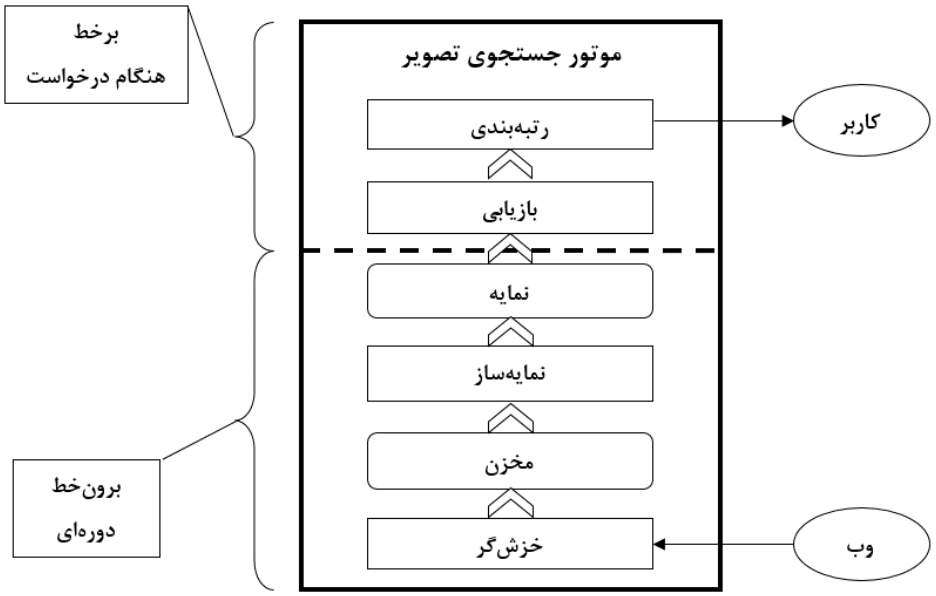

**شکل ۱-۲: مثالی از نمایه‌ساز معکوس**

هدف از نمایه‌سازی استخراج یک نمای منطقی از اسناد می‌باشد. مرسوم‌ترین نمای منطقی در موتورهای جستجو، مدل بسته‌ی کلمات[1] می‌باشد. در این مدل، هر سند از مجموعه‌ای از لیست نامرتب کلمات تشکیل شده است. متن اسناد باید قبل از نمایه‌سازی نرمال شود. هدف از نرمال‌سازی متن، استخراج کلمات کلیدی از سند و حذف حروف اضافی می‌باشد[۶].

در این گام، کاربر پرس‌وجوی خود را از طریق واسط کاربر به موتور جستجو اعلام می‌کند. ماژول بازیابی[2] و رتبه‌بندی، تصاویر مرتبط با آن پرس‌وجو را بازیابی نموده و بر اساس الگوریتم رتبه‌بندی، نتایج جستجو را مرتب نموده و از طریق واسط کاربر، به کاربر ارائه می‌دهد. واسط کاربر به صورت مستقیم با کاربر رابطه دارد. از یک طرف پرس‌وجوی کاربر را دریافت نموده و از طرف دیگر نتایج بازیابی شده پس از رتبه‌بندی را به کاربر نشان می‌دهد.

[1] Bag of words
[2] Retrieval

یک موتور جستجو از دو فرآیند برخط[1] و برون‌خط[2] تشکیل شده است. فرآیند برون‌خط به صورت مداوم و دوره‌ای[3]، تصاویر و صفحات وب را دریافت و ذخیره می‌نماید. سپس نمایه‌ساز این اطلاعات را به نمایه‌های قابل جستجو تبدیل می‌نماید. فرآیند برخط، در هنگام درخواست کاربر اجرا می‌شود. هنگام پرس‌وجوی کاربر، بخش بازیابی تعدادی سند را بازیابی نموده و بخش رتبه‌بندی نتایج بازیابی شده را بر اساس میزان ارتباط با پرس‌وجوی کاربر مرتب نموده و به کاربر ارائه می‌دهد. شکل ۱-۳ این دو فرآیند را در یک موتور جستجو نشان می‌دهد.




**شکل ۱-۳: فرآیندهای برخط و برون‌خط موتور جستجو**

## ۱-۲ رتبه‌بندی تصاویر در وب

رتبه‌بندی یک تصویر به فرآیندی گفته می‌شود که به هر تصویر بازیابی شده با توجه به پرس‌وجوی کاربر، امتیازی نسبت بدهیم که بیان‌کننده میزان مرتبط بودن آن تصویر با پرس‌وجوی کاربر باشد. سپس تصاویر بازیابی شده را بر اساس این امتیاز، مرتب کنیم. با توجه به این‌که به ازای هر پرس‌وجوی کاربر، هزاران تصویر مرتبط وجود دارد، لازم است تا موتور جستجو نتایج را رتبه‌بندی

[1] Online
[2] Offline
[3] Periodically

نموده و تنها چند نتیجه اول را به کاربر نمایش دهد. در حقیقت رتبه‌بندی، میزان مرتبط بودن هر تصویر نتیجه با پرس‌وجوی کاربر را تخمین خواهد زد.

الگوریتم‌های مختلفی برای رتبه‌بندی تصاویر وجود دارد که به طور کلی می‌توان آن‌ها را در چند دسته زیر تقسیم‌بندی نمود:

- الگوریتم‌های مبتنی بر محتوای متنیت
- الگوریتم‌های مبتنی بر محتوای دیداری تصاویر
- الگوریتم‌های مبتنی بر اتصال
- الگوریتم‌های مبتنی بر رفتار کاربر

الگوریتم‌های مبتنی بر محتوای متنی، به هر تصویر درون یک صفحه، بر اساس میزان مرتبط بودن پرس‌وجوی کاربر با محتوای متنی آن صفحه و ویژگی‌های متنی خود تصویر، امتیازی نسبت می‌دهند. یعنی هرچه پرس‌وجوی کاربر با محتوای متنی یک صفحه که حاوی یک تصویر خاص است مرتبط‌تر باشد، آن تصویر امتیاز بیشتری کسب کرده و در رتبه‌بندی در اولویت بالاتری قرار خواهد گرفت[۷، ۸]. در روش‌های مبتنی بر محتوای متنی از مدل‌های بولی[۱]، احتمالی[۲] و فضای برداری[۳] جهت رتبه‌بندی اسناد بر مبنای محتوای آن‌ها استفاده می‌شود. مهمترین روش در مدل برداری، الگوریتم TF-IDF[۹] و در مدل احتمالی BM25[۱۰] می‌باشد.

الگوریتم‌های مبتنی بر محتوای دیداری تصاویر، بر اساس ویژگی‌های محتوای تصویر، میزان شباهت بین دو به دوی تصاویر بازیابی شده را می‌یابند. سپس گرافی تحت عنوان گراف شباهت تشکیل می‌دهند. حال به هر گره از گره‌های این گراف که همان تصاویر ما می‌باشند بر اساس تعداد یال‌های متصل به آن و وزن هر کدام، امتیازی نسبت داده می‌شود. حال بر اساس این امتیازها، رتبه‌بندی مجددی روی تصاویر بازیابی شده انجام می‌شود[۱۱، ۱۲].

[1] Boolean
[2] Probabilistic
[3] Vector space

الگوریتم‌های مبتنی بر اتصال، امتیازدهی به هر تصویر را بر اساس اعتبار[1] صفحاتی که تصویر درون آن قرارگرفته است، انجام می‌دهند. این اعتبار بر اساس یک رابطه به خود تصویر منتقل می‌شود و به عنوان امتیاز برای آن تلقی می‌شود [۱۳].

الگوریتم‌های مبتنی بر رفتار کاربر، الگوریتم‌هایی هستند که عمل یادگیری[2] را انجام می‌دهند. این الگوریتم‌ها بر اساس بازخورد[3]هایی که از کاربر دریافت می‌کنند، امتیاز یک تصویر را تغییر می‌دهند. این نوع الگوریتم‌ها، معمولاً به‌صورت ترکیبی با الگوریتم‌های دیگر مورد استفاده قرار می‌گیرند [۱۴، ۱۵، ۱۶].

الگوریتم‌های ترکیبی، با ترکیب دو یا چند روش فوق، سعی در بهبود نتایج و سرعت جستجو، نموده‌اند [۱۷].

## ۱-۳ چالش‌های[4] رتبه‌بندی تصاویر در وب

### ۱-۳-۱ مفهوم مرتبط بودن

وقتی کاربر پرس‌وجوی خود را اعلام می‌کند، سامانه[5] باید مجموعه‌ای از تصاویر مرتبط با آن پرس‌وجو را به او ارائه کند. حال منظور از مرتبط بودن پرس‌وجو با یک تصویر چیست؟

در یک دید اولیه، می‌توان گفت اگر محتوای متنی پرس‌وجو با محتوای متن صفحه‌ی در برگیرنده تصویر، اشتراک بیشتری داشته باشد، آن تصویر، تصویر مناسب‌تری برای ارائه به کاربر می‌باشد. مزایای این روش سادگی و سرعت بالای بازیابی نتایج می‌باشد. ولی همواره چنین نیست. صفحاتی که به هر دلیل، از متن و تصویر نامرتبط باهم، استفاده نموده‌اند، باعث می‌شوند نتایج نامرتبط با پرس‌وجو به‌عنوان نتیجه جستجو به کاربر ارائه شود[۱۸]. از دیگر مشکلات الگوریتم‌های مبتنی بر محتوای متنی می‌توان به موجود نبودن این اطلاعات برای هر تصویر و دقت پایین نام برد.

[1] Authority
[2] Learning
[3] Feedback
[4] Challenges
[5] System

پس الگوریتم‌های مبتنی بر محتوای متنی نمی‌توانند به تنهایی در جستجوی یک تصویر در وب، موفق عمل نمایند. در نتیجه می‌توان از ترکیب این الگوریتم با الگوریتم‌های دیگر بهره برد.

### ۱-۳-۲ توصیف‌گرهای تصویر[1]

در الگوریتم‌های مبتنی بر محتوای دیداری، به ویژگی‌هایی از تصویر نیاز داریم که بتواند به نوعی بازگوکننده محتوای دیداری تصویر باشد. چنین ویژگی‌هایی با نام توصیفگر تصویر، شناخته می‌شوند. توصیفگرهای تصویر، به طور دقیق‌تری، یک تصویر را توصیف می‌کنند ولی استخراج آن‌ها از تصویر با محدودیت در سرعت همراه است[۱۸]. زیرا در هر تصویر با میلیون‌ها پیکسل به عنوان داده مورد پردازش روبرو هستیم و الگوریتم‌های تعیین میزان شباهت بین تصاویر از درجه پیچیدگی بالایی برخوردارند. مشکل دیگر در استفاده از توصیفگرهای تصویر این است که استنتاج مفاهیم معنایی[2] از چنین داده‌هایی، کار آسانی نیست. زیرا این توصیفگرها، تصویر را در سطح پایین[3]، توصیف می‌کنند. همچنین یافتن شباهت بین دو به دوی تصاویر کاری بس پیچیده است. به همین منظور به راهکارهایی نیاز است تا این فرآیند، با سرعت بالاتری انجام گیرد.

### ۱-۳-۳ وابستگی به کاربر

وابستگی به بازخورد کاربر در الگوریتم‌هایی که بر اساس رفتار کاربر عمل می‌کنند، خود یک نوع چالش محسوب می‌شود. تجربه نشان داده است که کاربران علاقه‌ای به ارائه بازخورد مستقیم نخواهند داشت. به همین دلیل چنین الگوریتم‌هایی در موتورهای جستجوی تصویر، کاربرد کمی دارند.

[1] Image Descriptor
[2] Semantic Concepts
[3] Low Level

### ۱-۳-۴ شباهت[1] یا تطبیق[2]

در جستجوی تصاویر توسط الگوریتم رتبه‌بندی محتوای تصویر، می‌توان با انتخاب کاربر، تصاویر مشابه و یا تصاویر دقیقاً منطبق بر هم را به کاربر ارائه داد[۱۸]. یعنی پس از جستجو توسط کاربر، بتوانیم تصاویر مشابه و یا تصاویر منطبق با یک تصویر نتیجه را به کاربر ارائه کنیم.

### ۱-۳-۵ نمایه‌ها[3]

یکی از مسائل پیش روی موتورهای جستجو، حجم نمایه‌های حاصل از ویژگی‌های متنی صفحات، ویژگی‌های متنی اختصاص‌یافته[4] به تصویر و توصیفگرهای تصاویر می‌باشد. بسته به الگوریتم رتبه‌بندی و راهکارهای آن، حجم نمایه‌ها می‌تواند متفاوت باشد. هرچه قابلیت‌های جستجو بالاتر باشد، حجم نمایه‌ها نیز زیادتر می‌شود. به عنوان نمونه اگر بخواهیم مکان واژه‌های پرس‌وجوی کاربر را در نتیجه نشان دهیم و یا این‌که بخواهیم جستجو را بر روی رنگ و اندازه تصاویر انجام دهیم، نیاز به حجم بالاتری برای ذخیره نمایه‌ها داریم. در نتیجه در طراحی الگوریتم رتبه‌بندی باید محدودیت‌های مربوط به فضای مورد نیاز برای نمایه‌ها را هم باید مدنظر قرار دهیم. در صورتی که حجم نمایه‌ها خیلی زیاد باشد، باید اسناد وب را به چندین نمایه در جاهای مختلف تقسیم نمود و عمل جستجو را به صورت موازی انجام داد. برای این منظور می‌بایست از روش‌های توزیع‌شده[5] بهره برد[۱۸].

### ۱-۳-۶ نوع پرس‌وجو

نوع پرس‌وجوی کاربر می‌تواند متفاوت باشد. پرس‌وجو توسط متن، پرس‌وجو توسط یک تصویر به‌عنوان ورودی و یا ترکیبی از این دو حالت را به عنوان ورودی داشته‌باشیم. برای هر یک از این حالات، الگوریتم‌های متفاوتی برای رتبه‌بندی به‌کار می‌روند. در این حالت، کاربر پرس‌وجوی خود را

[1] Similarity
[2] Matching
[3] Indexes
[4] Associated Text
[5] Distributed