

مقدماتی بر سئو

- بیش از ۱/۵ میلیارد وبسایت فعال وجود دارد
- بیش از ۱/۵ میلیون سایت ایرانی وجود دارد
- قانون ۲۰/۸۰ و ویلفرتو پرتو (قانون اصل پرتو)
- اگر می‌خواهید مرده‌ای را دفن کنید، مخفی‌ترین و بی‌رفت آمدترین جا صفحه دوم گوگل است
- موتورهای جستجو لکسیکال یا لغت محور (Lexical)
- موتورهای جستجو هایپرتکست شوال یا لینک محور (Hypertextual)
- موتورهای جستجو آنتی دیورین تد یا سمتیک یا معنایی (Semantic)
- رسالت و ماموریت گوگل تمام اطلاعات جهان را به تمامی افراد دنیا نشان بده و سازماندهی کند
- ۹۲٪ در جهان و ۹۹٪ در ایران منبع بدست آوردن اطلاعات درباره جستجو کاربران در جهان و ایران

Gs.statciunter.com

جلسه اول

موتورهای جستجو قبل از گوگل Lexical یا لغت محور بودن

یعنی به تعداد کلمات در سایت‌ها توجه می‌کردند.

فهرستی که موتورهای جستجو اطلاعات مورد نیاز خود را از صفحات اینترنتی بر می‌دارند را ایندکس موتور جستجو می‌گویند که به شکل یک جدول هستند.

مثال: چگالی کلمه کلیدی

اگر یک متنی ۱۰۰۰ کلمه است و ۵۰۰ تایی آن کلمه کلیدی باشد حتماً تقلب داخل آن است.

قانون Zipf

در متنی‌هایی که توسط انسان‌ها تولید می‌شود تعداد تکرار کلمات با هم دیگر یک رابطه معناداری دارند.

در ساده‌ترین حالت:

رتبه تعداد کلمات $\times$ تعداد تکرارشان

تقریباً یک عدد ثابت خواهد شد.

چگالی کلمه کلیدی:

چگالی کلمه کلیدی یعنی در هر ۱۰۰ کلمه چند بار کلمه کلیدی ما تکرار شده.

۱۰۰۰×۱۰۰/۲۵

اینترنت مثل یک تار عنکبوت است.

لینک‌ها، صفحات مختلف در اینترنت را به یکدیگر وصل می‌کند.

انکر تکست لینک (Anchor text):

برای اطلاعات بیشتر روی این لینک کلیک کنید.

به صفحه‌ای که بعد از کلیک به آن منتقل می‌شویم به آن صفحه مقصد لینک می‌گویند.

اولین الگوریتم گوگل Page rank (برگرفته شده از نام لری پیج موسس گوگل)

Patent یا حق اختراع.

گوگل اولین موتور جستجو بود که علاوه بر لغت محور بودن لینک محور هم بود.

جلسه دوم: شکار دام وب

کتاب فروشگاه همه چیز The everything store (کتابی در مورد تجارت الکترونیک مربوط به سایت آمازون)

کاربران به جای حفظ دامنه شما، اسم شما را گوگل می‌کنند.

اگر توی گوگل پیدا نمی‌شوید، شاید واقعا وجود ندارید.

سئو یک کانال جذب ترافیک پایداره.

جستجوها می‌توانند سرنوشت یک کسب و کار را عوض کنند.

روش‌های بازاریابی

۱. درون‌گرا (Inbound marketing)

۲. برون‌گرا (Outbound marketing)

بازاریابی درون‌گرا: بازاریابی درون‌گرا یعنی خود مخاطبان به سمت ما می‌آیند و مکالمه را آغاز می‌کنند.

بازاریابی برون‌گرا: بازاریابی برون‌گرا یعنی ما هستیم که به سمت مخاطبان می‌رویم و تلاش می‌کنیم تا مکالمه را آغاز کنیم.

درون‌گرا مثل: سئو، بازاریابی محتوا، تبلیغات دهن به دهن و ...

برون‌گرا مثل: تلفن مارکتینگ، تبلیغات تلویزیونی، رادیویی، پیامکی و ایمیل و ...

گوگل موتور جستجوی افراد صبور است.

تقلب PBN (Private Blog Network) یا شبکه‌های خصوصی

شخصی که چندین سایت و بلاگ دارد و همه را به یک سایت اصلی لینک می‌دهد.

الگوریتم Cassandra (برای مقابله با PBN درست شد)

تقلب یعنی تمارض کردن به چیزی که واقعا نیستید.

تفاوت سئو کلاه سفید و سیاه

سئو کلاه سفید

به زمان بیشتری نیاز دارد.

پایدار و بلند مدت است.

سقوط‌ها کوچکند و به راحتی قابل بازگشت.

سئو کلاه سیاه

زمان کمتری نیاز دارد.

ناپایدار و کوتاه مدت.

سقوط‌ها ترسناک و گاهی غیرقابل بازگشت هستند.

سئو کلاه خاکستری

سئو کلاه سیاهی است که سیاه بودنش هنوز واسه گوگل قابل تشخیص نیست.

گوگل دنس، رقص گوگل (Google dance)

مربوط به آپدیت‌های الگوریتم‌های گوگل به صورت ماهیانه بود.
گوگل دنس تغییر کرد و به آپدیت همیشگی (Google ever flux) تبدیل شد.

جلسه سوم

درک کیفیت مبدا لینک‌ها:

۱. بررسی ارتباط موضوع صفحه مبدا به لینک
۲. بک لینک‌های صفحه مبدا لینک (سنجش اعتبار صفحه مبدا لینک)

(LSI: Latent Semantic Indexing)

توجه گوگل به محتواهای نزدیک لینک

الگوریتم Hilltop: خرید از یک فرد هندی به نام کریشنا بهارات

۱. خبرگی: Expertise

۲. اعتبار: Authority

در الگوریتم Hilltop اگر ۵ خبره به یک نفر لینک می‌دادند آن لینک دارای اعتبار بود و معتبر بودن بعد از خرید این الگوریتم توسط گوگل عدد ۵ به ۲ کاهش یافت.

مثل سایت‌های Directory یا سایت‌های Edu

لینک نوفالو (Nofollow): یعنی لینک تبلیغاتی است.

لینک فالو: یعنی در رتبه‌بندی تأثیر داشته باشد. (مستقیماً روی رتبه پیج رنک تأثیر می‌گذارد)

جلسه چهارم

مزرعه محتوا (Content Farm)

یک مزرعه محتوا، محتواهایی را از سایت‌های دیگر جمع می‌کند و با یک تغییر کوچک در سایت خود وارد می‌کند و لینک‌سازی انجام می‌دادند.

مثل سایت‌های Demand Media و Answers

آپدیت کافئین (Caffeine Update)

یک آپدیت در ساختار و معماری گوگل است.

اگر آپدیت کافئین اتفاق نمی‌افتاد و امروز اثری از گوگل وجود نداشت.

مشکل زیرساخت گوگل: هنگام سرچ کلمه دو عدد داریم

About ۱۴۱.۰۰۰.۰۰۰ results (۰.۵۴ second)

About تعداد نتایجی که گوگل index کرده است

Results مدت زمانی که طول کشیده تا گوگل جستجو را داخل ایندکس انجام بدهد.

اگر سایتتون تو نتایج گوگل وجود ندارد یا وجود دارد اما اطلاعاتی که از آن نشان می‌دهد آپدیت نیست یعنی ایندکس گوگل نسبت به سایت شما آپدیت نیست.

مشکلات:

۱. ظرفیت ایندکس گوگل نسبت به تعداد محتواهایی که در طول روز منتشر می‌شوند کوچک بوده.
۲. گوگل به خاطر سرعت‌العمل کند از محتواهای جدید یا محتواهای آپدیت شده جا می‌ماند.

حل مشکل:

۱. اصلاح معماری خزنده‌ها
۲. افزایش ظرفیت ایندکس

با آپدیت کافئین ظرفیت و سرعت ایندکس گوگل چندین برابر شد.

در ۲۴ تا ۴۸ ساعت، سایت باید ایندکس شود.

آپدیت Attribution Update

اولین محتوا برای اولین بار توسط چه کسی ساخته شده و صاحب اصلی این محتوا کیست.

سرویس Digital Millennium Copyright Act: (DMCA)

تولید کننده اصلی محتوا می‌تواند گزارش کپی شدن محتوای خود به گوگل را به صورت دستی اعلام می‌کند.

الگوریتم پاندا

شخصی به نام **Navneet Panda**:

فاکتورهایی که برای الگوریتم پاندا مهم‌اند:

۱. محتوای تکراری
۲. اعتبار و تخصص
۳. کیفیت نگارش محتوا (غلط املائی نداشته باشد)
۴. کمیت محتوا (کم بودن محتوا و بیش از حد بلند بودن آن بی‌کیفیت است)
۵. ارزشمندی (مثلا دیدگاه می‌گذارند، بوک مارک کردن مقاله، share در شبکه‌های اجتماعی)
۶. جامعیت محتوا (همه جوانب یک موضوع را پوشش می‌دهند)
۷. دسته اول بودن (ارائه ایده‌های جدید و نو، تصاویر اختصاصی، ویدیوهای جدید و جداول)

۸. قدرت برند (با کیفیت بودن یک وبسایت این است که اسم اون شرکت داخل کوئری‌ها هست یا نه)
۹. تبلیغات (پرهیز از تبلیغات زیاد)
۱۰. امنیت (امنیت کاربر در سایت خطر نیفته مثل Https و ...)

الگوریتم پاندا کل محتوای سایت را بررسی می‌کند و بعد به یک نتیجه می‌رسد (نه فقط یک مقاله بلکه کل سایت را بررسی می‌کند)

پاندا جز هسته اصلی گوگل شده و ممکن است روزی چند بار آپدیت شود.
الگوریتم‌های جدید این طور نیست که بلافاصله وارد Core Algorithms شوند.

آپدیت هسته (Core Update - broad core update)

آپدیت‌هایی که هسته گوگل را تشکیل می‌دهند زمانی که آپدیت شوند Core Update است.

آپدیت اصلی Major Update

آپدیت فرعی Minor Update

جلسه پنجم

الگوریتم فرشنس (Freshness Algorithm) الگوریتم تازگی

آقای Amit Singhal: این الگوریتم را ارائه داد.

مدل الگوریتم فرشنس:

QDF: Query Deserves Freshness

مدل ریاضی که تشخیص می‌دهد مردم چه زمانی دنبال اطلاعات جدید هستند و چه زمانی نیستند.

الگوریتم فرشنس بر روی جستجوهای که نیاز به بروزرسانی بیشتر دارد تأثیر بیشتری می‌گذارد.

نشانه‌های از رده خارج شدن محتواهای قدیمی:

۱. آخرین تاریخ بروزرسانی محتوا (تغییر آخرین بروزرسانی محتوا می‌تواند تأثیرگذار باشد که کلاه سیاه است)
۲. نیاز موضوع به بروزرسانی
۳. کاهش نرخ کلیک در بین سایر نتایج کم می‌شود
۴. کاهش زمان حضور کاربران
۵. کم شدن یا ثابت ماندن لینک‌های داخلی (باید به محتوای سایت لینک بدهیم)
۶. کاهش یا ثابت ماندن لینک‌های خارجی

الگوریتم پنگوئن Penguin Algorithm

در الگوریتم پنگوئن لینک‌های Follow فقط مدنظر است.

لینک‌ها به دو دسته:

۱. ارگانیک یا طبیعی Natural Links
۲. مصنوعی یا غیرطبیعی Unnatural Links

ارگانیک:

لینک‌هایی که کاملاً و بدون دخالت ما ساخته می‌شود.

(لینک‌هایی که از سایت‌های دیگر بدون دخالت ما به سایت ما داده می‌شود)

سایت‌هایی داریم به نام Link Farm (که به صورت Spam وار به سایت‌های مختلف لینک می‌دهند)

لینک ارگانیک یا طبیعی لینکی که طبق تجربه گوگل ساخته شدن آن طبیعی است.

لینک غیر ارگانیک یا مصنوعی طبق تجربه گوگل ساخته شدن آن طبیعی نیست.

ارگانیک:

مشکل: خیلی خیلی سخت بدست میاد.

مصنوعی:

به راحتی در جاهای دیگر لینک می‌سازد.

۱. کامنت گذاشتن در سایت‌های مختلف به قصد لینک گرفتن
۲. خرید رپورتاژ
۳. دریافت لینک از سایت‌های مختلفی که از ابزارها استفاده می‌کنند
۴. دریافت لینک از فوتر سایت‌هایی که از خدمات ما استفاده می‌کنند
۵. خرید بک لینک
۶. تبادل لینک
۷. استفاده از ابزارهای ساخت لینک اتوماتیک

با کیفیت یا بی کیفیت بودن بک لینک در پیج رنک اثر می‌گذارد و لینک‌دهی مصنوعی بی‌کیفیت می‌شود (که یک نوع تقلب است).

مثل:

Overstock.com

پنالتی یا جریمه دستی:

وقتی گوگل متوجه تقلب یک سایت می‌شود به صورت دستی آن سایت جریمه می‌شود و از طریق سرچ کنسول به مدیر آن سایت اطلاع می‌دهد.

مت کاتز: Web spurs bad به الگوریتم پنگوئن تغییر کرد.

سرویس گزارش لینک‌سازی‌های مصنوعی به گوگل.

الگوریتم پنگوئن به موارد زیر توجه می‌کند (برای ساختگی بودن)

۱. تعداد لینک‌های ساخته شده در بازه زمانی مشخص
۲. منبع و محل قرارگیری لینک‌های ساخته شده
۳. حجم لینک‌های ساخته شده با کلمه کلیدی هدف
۴. سابقه لینک‌سازی‌های قبلی
۵. ارتباط سایت‌های مبدا و مقصد لینک

سئو منفی (Negative ESO)

انجام دادن اقدامات کلاه سیاه ولی برای رقیبان

حل مشکل توسط گوگل: Disavow Links To Yours Site

هدف: اثر لینک‌های با کیفیت را از بین نمی‌برد بلکه هدف این سرویس از بین بردن سئو منفی است که توسط رقبا برای ما به وجود آمده است.

الگوریتم پنگوئن به عنوان بخشی از هسته گوگل کار می‌کند.

فرق الگوریتم پنگوئن با پنالتی گوگل

الگوریتم پنگوئن:

۱. می‌تواند باعث افت رتبه سایت شود
۲. گوگل هیچ اطلاعی در مورد کاهش رتبه به صاحب سایت نمی‌دهد
۳. در بیشتر موارد با از بین بردن لینک‌ها به جایگاه قبلی بر نمی‌گردد

۱. می‌تواند باعث افت رتبه سایت شود
۲. گوگل علت افت رتبه را به صاحب سایت اطلاع می‌دهد: از طریق سرچ کنسول
۳. در بیشتر موارد بعد از انجام اصلاحات به جایگاه قبلی که داشتید برمی‌گردید

سرویس Disavow Tools گوگل برای از بین بردن لینک‌های بی‌کیفیت نیست.

EMD: Exact Match Domain

تأثیر نام دامنه بر سئو

۱. هدف‌تان را از راه‌اندازی سایت مشخص کنید
۲. حساسیت گوگل بر روی سایت‌های EMD بیشتر می‌شود
۳. تأثیر EMD به مرور زمان کمرنگ‌تر و کمرنگ‌تر خواهد شد
۴. تأثیر یک برند قدرتمند بر روی سئو، صدها برابر بیشتر از وجود کلمات کلیدی در داخل دامنه است

تفاوت سئو ir یا com

دامنه‌های عمومی: net,org,com

دامنه‌های اصلی: ir,uk,de,fr,co,cu,in

دامنه‌های اختصاصی: edu,tv,gov,mil

دامنه‌های کشوری جامعه هدف روی اون کشور تنظیم می‌کند.

نشان دادن سایتتون به کشورهای دیگر کم‌تر است.

استفاده از دامنه‌های جدید: app, tools, club, academy

هرکسی می‌تواند از این نوع دامنه‌ها ثبت کند.

توجه گوگل به این نوع دامنه‌ها باعث ایجاد بازار سیاه و اسپم در دامنه خواهد شد.

معرفی دیرتر این نوع دامنه‌ها، شانس استفاده از آن‌ها را از سایت‌های قدیمی‌تر می‌گیرد.

جلسه هفتم

الگوریتم مرغ مگس خوار (Hummingbird Algorithm)

سال ۲۰۱۲: جستجوها، تک کلمه‌ای یا دو کلمه‌ای هستند.

سال ۲۰۱۳: برای پیدا کردن بهترین نتایج از خود کاربرها کمک گرفت.

بررسی کلیک‌های کاربر بر روی نتایج

تغییراتی که بعد از پروژه مرغ مگس خوار در دنیای سئو ایجاد کرد.

۱. طولانی‌تر شدن جستجوهای کاربران (قبل از این الگوریتم دو یا یک کلمه‌ای بودن که بعد تبدیل ۳ , ۴ ... شدن)

۲. مهم شدن توجه به نیت کاربران در تولید محتوا

نیت جستجو (Search Intent)

گوگل با نگاه به رفتار کاربران متوجه نیت آن‌ها از جستجو می‌شوند.

در تولید محتوا باید به نیت کاربران از جستجویشان فکر کنید.

توجه به نیت جستجو سرنوشت صفحه اول گوگل را رقم می‌زند.

توجه به نیت در انتخاب جنس مناسب محتوا هم تأثیرگذار است.

جلسه هشتم

الگوریتم‌های موبایل گوگل (Google Mobile Algorithm)

دو تا الگوریتم Mobile Friendly, Mobile First Index

موبایل فرندلی: سایت‌هایی که نشان می‌دهند با توجه به دستگاه مختلف صفحه نمایش تغییر می‌کند.

۲۰۱۱: اولین زمانی که رتبه‌بندی متفاوت در موبایل نسبت به دسکتاپ انجام داد.

وقتی سایتی در گوگل ایجاد می‌شود سه مرحله زیر را طی می‌کند:

۱. خزیدن Crawling: هدف آن کامل کردن فهرست index است.

۲. ایندکس کردن Indexing: فهرستی است که اطلاعات مختلف از یک وبسایت را در خود اضافه می‌کند.

۳. رتبه‌بندی Ranking: با توجه به الگوریتم گوگل رتبه می‌گیرد.

Crawl Google Bot: فقط خزنده‌ای است که دسکتاپ را بررسی می‌کند.

همیشه پیشرفت تکنولوژی به معنی تغییرات خیلی زیاد و سریع در رتبه‌بندی‌ها نیست، باید به این فکر کنید که آیا همه عوامل دخیل در تغییر، آمادگی تغییر را دارند.

در سال ۲۰۱۵ الگوریتم موبایل فرندلی:

اسم دیگر موبایل گدون (Mobilegeddon) (فاجعه موبایلی، آخر زمان موبایل)

موبایل فرندلی یکی از فاکتورهای رتبه‌بندی است.

نکته ۱: صفحات سایت مستقل از هم بررسی می‌شوند.

نکته ۲: روی نتایج دسکتاپ مستقیماً تأثیر گذار نیست.

وظیفه Google bot - mobile این بود که صفحاتی که توسط google bot جمع شده را بررسی کند آیا موبایل فرندلی است یا خیر.

۲۰۱۶: موبایل فرست ایندکس

۱. بیشتر ترافیک از طریق سرچ دسکتاپ: خزنده اصلی: google bot

۲. بیشتر ترافیک از طریق سرچ موبایل: خزنده اصلی: google bot mobile\۲۰۱۸

۲۰۱۸ الگوریتم موبایل فرست ایندکس.

جلسه نهم

روش A/B تست: یک راه حل ساده برای تصمیم‌گیری

روش کار گوگل: به عنوان ماشین

• واکنش مناسب در شرایطی جدید:

روش‌های یادگیری ماشینی می‌توانند حتی در مقابل حالت‌های جدید هم واکنش‌های خوبی نشان دهند.

• کسب تجربه از تلاش‌های قبلی:

ماشین‌ها به کمک روش‌های یادگیری ماشینی می‌توانند تجربه کسب کنند و هر روز بهتر از دیروز شوند.

الگوریتم رنک برین (Rank Brain)

بچه مرغ مگس خوار (مادر این الگوریتم مرغ مگس خوار)

رنک برین جز هسته اصلی گوگل

رنک برین با اندازه‌گیری میزان رضایت از نتایج، جز فاکتورهای مهم رتبه‌بندی است.

بعد از رنک برین تأثیر پارامترهای رتبه‌بندی در حوزه‌ها و جستجوهای مختلف، متفاوت است.

تأثیر الگوریتم رنک برین بر روی جابجایی نتایج گوگل:

هدف تجربه کاربری وارد نتایج گوگل کند.

معیارهای رنک برین برای سنجش رضایت کاربران از نتایج

نرخ کلیک (CTR: Click Through Rate)

CTR : تعداد کلیک/تعداد نمایش

بازگشت به نتایج Pogo-Sticking

بازگشت به صفحه نتایج

نرخ پرش (BR: Bounce Rate)

وارد صفحه دیگر از سایت شده یا دیدگاه گذاشته یا خیر.

مدت زمان ماندگاری (Owel Time)

چقدر در صفحه سایت می‌مانند.

شجره راه وب
تلویزیون
سود ۲×

نگاه کلی گرا:

پنیر

ضرر ۱×

Holistic thinking

(Thinking outside the box)

نگاه جزئی گرا:

پنیر

سود ۱×

Partial thinking

(Thinking inside the box)

تلویزیون

سود ۱×

حداقل کاری که در برابر سوالات اشتباه می‌توان انجام داد خارج شدن از چارچوب اشتباه است.

جلسه دهم

دانش زبان‌های طبیعی (NLP: Natural Language Processing)

وقتی ماشین‌ها بتوانند مثل ما متوجه معانی کلمات شوند...

خلاصه‌سازی متن Text Summarization

چت بات‌ها chat bots

ترجمه‌های ماشینی Machine translation

پاسخ به سوالات Question answering

NTT: موجودیت (چیزهایی که از قبل راجع به آن‌ها اطلاعات داریم) کلماتی که به یک چیز خاص اشاره می‌کند

ENR: Named Entity Recognition

ماشین‌ها هم مثل مغز ما در جملات دنبال کلماتی می‌گردند که از قبل راجع به آن‌ها اطلاعات دارند.

جستجو معنایی Semantic Search

چون، مثل مغز ما هستند Entity Base Search, Entity Oriented Search گفته می‌شود.

۲۰۱۰: سرویس Google Squared Online: کلی اطلاعات در مورد اشخاص، حیوان، مکان و ... اضافه کند. یا اطلاعات را ببیند، این طرح ۲۰۱۱ جمع شد.

سرویس Google Knowledge Graph: اطلاعاتی که در سرویس قبل جمع شده بود را به گراف تبدیل می‌کرد و اطلاعات مرتبط را به هم وصل می‌کند.

تکنیک word ۲ Vec: تبدیل کلمات به بردارهایی ریاضی.

تا سال ۲۰۱۷ هنوز گوگل یک موتور جستجو لغت محور بود ولی کمی بهتر.

مکانیزم attention هیچ حرف یا کلمه‌ای از جمله حذف نمی‌شود و سعی می‌کند متوجه شود که هر کلمه به کدام کلمه و جمله مربوط می‌شود.

مشکل: نمی‌شد با شبکه‌های عصبی مختلف انجامش داد.

مدل BERT; Bidirectional Encoder Representations From Transformers

فهم فوق‌العاده از کلمات و زبان‌ها پیدا کنند و آن‌ها را متوجه شوند.

همه حرف‌های اضافه را متوجه می‌شود.

از BERT در گوگل ترسنتیت هم استفاده می‌شود.

BERT یک الگوریتم open source است

۲۰۱۹: مدل BERT به یک اطلاعات بزرگ مثل ویکی پدیا آموزش داد.

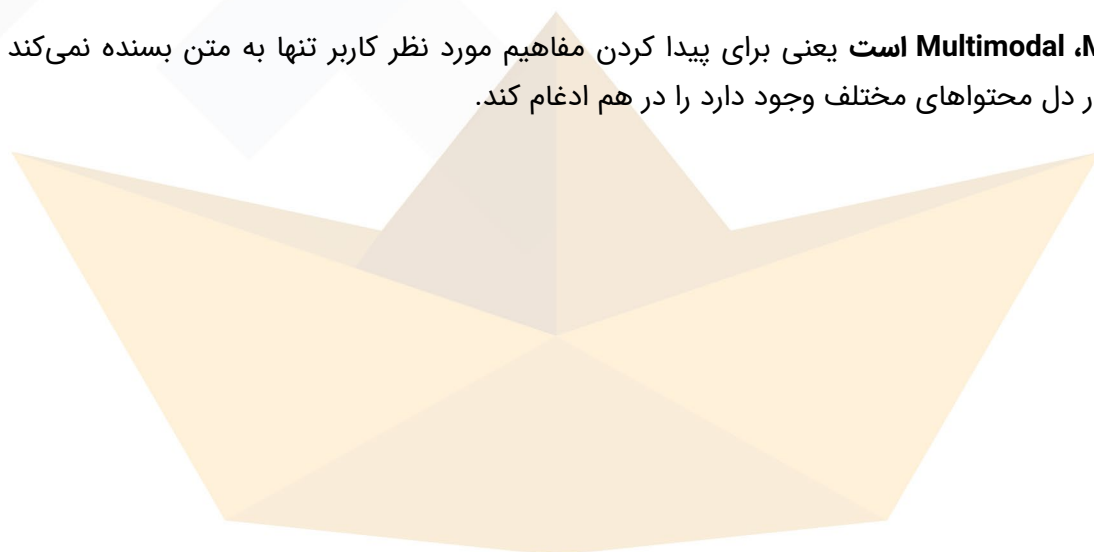
بعد از الگوریتم برت گوگل می‌تواند مفاهیم مختلف را در یک زبان یاد بگیرد و در زبان‌های دیگر از آن استفاده کند.

مدل Google MUM

Multitask Unified Model

الگوریتم MUM هم همانند BERT از معماری Transformer ها و تکنیک Attention برای درک بهتر محتواها استفاده می‌کند اما تا ۱۰۰۰ برابر سریع‌تر و قدرتمندتر است.

الگوریتم MUM، Multimodal است یعنی برای پیدا کردن مفاهیم مورد نظر کاربر تنها به متن بسنده نمی‌کند و می‌تواند مفاهیمی که در دل محتواهای مختلف وجود دارد را در هم ادغام کند.



شاهراه وب