

داده کاوی رسانه‌های اجتماعی

«توییت‌ر، وبلاگ، ویکی‌پدیا»

مؤلفین

گابو زابو

گانگور پولاتکان

اسکار بویکین

آنتونیو چاکیو پولو

مترجم

ایوب تریان

نیاز دانش

www.ketab.ir

عنوان و نام پدیدآور	داده کاوی رسانه‌های اجتماعی: (توییت‌ر، وبلاگ، ویکی‌پدیا) / مولفین گابو زابو... او دیگران؛ مترجم ایوب ترکیان.
مشخصات نشر	تهران: نیاز دانش، ۱۳۹۷.
مشخصات ظاهری	۳۶۰ ص.
شابک	978-600-8906-40-7
وضعیت فهرست‌نویسی	فیمپا
یادداشت	عنوان اصلی: Social Media Data Mining and Analytics, 2019.
یادداشت	مولفین گابو زابو، گانگور پولاتکان، اسکار بویکین، آنتونیو جاکوپولو.
موضوع	برنامه‌ریزی سازمانی -- داده‌پردازی Business planning -- Data processing
موضوع	مصرف‌کنندگان -- رفتار -- آینده‌نگری Consumer behavior-- Forecasting
موضوع	داده کاوی Data mining
موضوع	مصرف‌کننده‌شناسی -- داده‌پردازی Consumer profiling -- Data processing
موضوع	رسانه‌های اجتماعی -- داده‌پردازی Social media -- Data processing
شناسه افزوده	سابو، گابور Szabó, Gábor
شناسه افزوده	ترکیان، ایوب، ۱۳۳۷، مترجم
رده‌بندی کنگره	۱۳۹۷ ۵۲/۲۱۳/۳۰ HD۳۰
رده‌بندی دیویی	۱۱/۳۸۰/۴۰۵۵۸
شماره کتابشناسی ملی	۵۵۳۰۵۱۲



نام کتاب	داده کاوی رسانه‌های اجتماعی: (توییت‌ر، وبلاگ، ویکی‌پدیا)
مؤلفین	گابو زابو - گانگور پولاتکان - اسکار بویکین - آنتونیو جاکوپولو
مترجم	ایوب ترکیان
مدیر اجرایی - ناظر بر چاپ	حمیدرضا محمد شیرازی - محمد شمس
ناشر	نیاز دانش
صفحه‌آرا	واحد تولید انتشارات نیاز دانش
نوبت چاپ	دوم - ۱۴۰۰
شمارگان	۵۰ نسخه
قیمت	۱۲۰۰۰۰۰ ریال

ISBN: 978-600-8906-40-7

شابک: ۹۷۸-۶۰۰-۸۹۰۶-۴۰-۷

هرگونه چاپ و تکثیر (اعم از زیراکس، بازنویسی، ضبط کامپیوتری و تهیه‌ی CD) از محتویات این اثر بدون اجازه‌ی کتبی ناشر ممنوع است، متخلفان به موجب بند ۵ از ماده‌ی ۲ قانون حمایت از مؤلفان، مصنفان و هنرمندان تحت پیگرد قانونی قرار می‌گیرند.

کلیه حقوق این اثر برای ناشر محفوظ است.

آدرس انتشارات: تهران، میدان انقلاب، خیابان ۱۲ فروردین، تقاطع وحید نظری، پلاک ۲۵۵، طبقه ۱، واحد ۲

۰۲۱-۶۶۴۷۸۱۰۶-۶۶۴۷۸۱۰۸-۰۹۱۲۷۰۷۳۹۳۵

www.Niaze-Danesh.com

مشاوره جهت نشر: ۰۹۱۲ - ۲۱۰۶۷۰۹

فهرست مطالب

عنوان شماره صفحه

فصل ۱ / کاربران: چه کسی رسانه‌های اجتماعی ۲۹

۲۹	سنجش تغییرپذیری رفتار کاربر در ویکی‌پدیا
۳۰	تنوع فعالیت‌های کاربر
۴۱	منشأ توزیع فعالیت کاربر
۵۰	تبعات قاعده توان
۵۶	دنباله دراز در فعالیت‌های انسانی
۵۹	دنباله دراز در همه جا: قاعده $۸۰/۲۰$ (قاعده p/q)
۶۴	رفتار برخاط روی توییتر
۶۵	بازیابی توییت‌ها برای کاربر
۶۷	سبببندی لگاریتمی
۶۹	فعالیت‌های کاربر در توییتر
۷۱	خلاصه

فصل ۲ / شبکه‌ها: چگونگی رسانه‌های اجتماعی ۷۳

۷۵	انواع و ویژگی‌های شبکه‌های اجتماعی
۷۵	شبکه‌های صریح
۷۷	گراف‌های جهت‌دار و بدون جهت
۷۸	ویژگی‌های گره و ضلع
۷۹	گراف‌های وزندار
۸۱	شبکه‌های غیرصریح
۸۴	مصورسازی شبکه‌ها
۸۸	درجه شبکه
۹۰	شمارش تعداد ارتباط
۹۲	دنباله دراز در ارتباطات کاربر
۹۶	ورای مدل شبکه ایده‌ها
۹۸	صید هم‌بستگی‌ها
۹۹	مثلث‌های محلی و خوشه‌بندی

۲۸۷	 ساختار Count-Min - مثال Hitter های سنگین
۲۸۸	 ساختار Count-Min - مثال درصد بالا
۲۸۸	 انباشت ساختارهای داده تخمینی
۲۹۰	 خلاصه تخمین ها
۲۹۰	 خلاصه

فصل ۶ / یادگیری، نگاشت، و توصیه ۲۹۳

۲۹۴	 سرویس های رسانه های اجتماعی برخط
۲۹۴	 موتورهای جستجو
۲۹۵	 درگیر شدن با محتوا
۲۹۷	 تعامل با دنیای واقعی
۲۹۷	 تعاملات با افراد
۲۹۹	 فرمولاسیون مسئله
۳۰۱	 یادگیری و نگاشت
۳۰۴	 فاکتورگیری ماتریس
۳۰۶	 یادگیری، آموزش
۳۰۶	 فرو- و فرابرازش
۳۰۹	 تنظیم در فاکتورگیری ماتریس
۳۰۹	 فاکتورگیری ماتریس غیرمنفی و تکمیل
۳۱۰	 نمایش روی درجه بندی فیلم
۳۱۴	 تفسیر کلیشه های یادگیری شده
۳۱۸	 تحلیل کاوشی
۳۲۳	 پیش بینی و توصیه
۳۲۶	 ارزیابی
۳۲۸	 مروری بر متدولوژی ها
۳۲۸	 رویکردهای نزدیکترین همسایه
۳۲۹	 رویکردهای مبتنی بر یادگیری با سرپرست
۳۳۰	 پیش بینی درجه بندی فیلم با رگرسیون لجیستیک
۳۳۹	 سوژه های رایج با ویژگی ها
۳۴۰	 کاربردهای زمینه خاص
۳۴۱	 خلاصه

فصل ۷ / نتیجه گیری ۳۴۳

۳۴۳	 پایداری عجیب الگوی تعامل انسانی
۳۴۶	 میانگین ها، انحراف معیارها، و نمونه برداری
۳۵۴	 حذف داده های پرت

مقدمه

این کتاب در مورد استفاده از داده‌ها برای شناخت نحوه استفاده سرویس‌های رسانه‌های اجتماعی است. از زمان توسعه وب ۲، سایت‌ها و خدماتی که به کاربران خود توان تغییر فعال و سهیم شدن در محتوای خدمات را می‌دهند، بسیار مورد اقبال قرار گرفته است. منشأ رسانه‌های اجتماعی در شبکه‌های اجتماعی و خدمات ارتباطی گروهی اولیه، شامل سامانه‌های بولتن‌بورد (BBS) دهه ۱۹۸۰، گروه‌های خبری Usenet و Geocities در دهه ۹۰ است؛ این گروه‌ها حول گرایش‌های موضوعی سازمان‌دهی می‌شدند و برای کاربران، ارتباطات ایمیلی یا چت فراهم می‌کردند. شبکه ارتباطات اطلاعات جهانی موسوم به اینترنت، سامانه شبکه سطح بالاتری را ایجاد کرد؛ وب ایجاد ارتباط در بین افراد و گروه‌های با گرایش‌های فکری مشابه. اگر چه ایده بنیادی وصل کردن افراد در سطح جهانی از آن موقع تغییر زیادی نکرده، دامنه و تاثیر سرویس‌های رسانه‌های اجتماعی در حد غیر قابل وصفی رشد کرده است. اگر چه طبیعی است که قسمت عمده‌ای از گفتگوها هنوز در «دنیای واقعی» رخ می‌دهد، تغییر جهت به سمت تبادل اطلاعات الکترونیکی در سطح اندرکنش‌های انسانی، قوی‌تر شده است. رشد سریع دستگاه‌های همراه و قابلیت اتصال، «اینترنت را در جیب افراد قرار داده»، و همراه آن، امکان برقراری ارتباط با دوستان، خانواده‌ها، و شرکت‌های تجاری مورد نظر، در هر زمان و در هر مکان، فراهم شده است.

جای تعجب ندارد که مجموعه‌ای از خدمات شکل گرفته و نیازهای ارتباط و اشتراک فراهم شده، فضای عمومی و خصوصی را متحول کرده است. از طریق این خدمات، امکان شناخت نحوه تفکر دیگران در مورد سیاست، برند، فراورده، و غیره، فراهم شده است. با به اشتراک گذاشتن شناسنامه‌دار یا بدون نام و نشان، امکان طرح آزاد تفکرات انسان‌ها در مقایسه با رسانه‌های سنتی، بیش از پیش فراهم گردیده است. با توجه به اینکه هر کس در صورت انتخاب، می‌تواند پیام خویش را به دیگران منتقل کند، مسئولیت مجموعه‌های خدمت‌رسانی نیز برای در

دسترس قرار دادن محتوای مربوط و جالب به کاربران، مطرح است. چیزی که در مورد همه این خدمات مشترک است، وابستگی آنها به افراد بوده، و آنها فقط نقش تسهیل‌گری را ایفا می‌کنند. این امر بدین معنی است که تا حدی، نظم ریاضی کشف شده در اثر تحلیل داده‌های کاربری، انعکاسی از رفتار افراد است. بنابراین، انتظار می‌رود در هنگام کار با این مجموعه داده‌ها، نکات و چالش‌های مشابهی مشاهده گردد. هدف این کتاب، نمایان کردن این نظم و تبیین رویکردهای کمک به شناخت نحوه توجه کاربران به این خدمات، از دریچه داده‌های جمع‌آوری شده توسط سامانه‌های این خدمات، است.

سبجش اندرکنش‌های انسانی

نیروی محرکه رسانه‌های اجتماعی، اندرکنش افراد حول محتوایی است که خدمت برخط ارائه می‌دهد. به‌عنوان مثال، شبکه‌سازی اجتماعی، ارتباط افراد با یکدیگر، به اشتراک‌گذاری تصویر و چندرسانه، گزارشات خبری، محتوای وب، و اطلاعات مختلف دیگر، را تسهیل می‌کند. در رایج‌ترین سناریوی کاربرد این خدمات، افراد برای به‌روز شدن در مورد دوستان، خویشاوندان، و برای به اشتراک گذاشتن چیزی در مورد زندگی خود با آنها، از فیسبوک استفاده می‌کنند. برای توییت، چون ارتباط لزومی ندارد دوطرفه باشد، کاربران می‌توانند در مورد نحوه فکر، به اشتراک‌گذاری، یا ارتباط یک فرد با دیگران، اطلاع پیدا کنند. با لینکدین، شبکه اجتماعی متخصصین، هدف، وصل کردن متخصصین با گرایش مشابه یا یکدیگر از طریق شبکه گروه‌های آنها، و وصل کردن شرکت‌های استخدام‌گر با متقاضیان کار است.

سرویس‌های رسانه‌ای دیگری وجود داشته که در آن، جنبه شبکه‌سازی تعاملات اجتماعی، به‌جای ایجاد یا لذت از محتوای اشتراکی (مثلاً، ویکی‌پدیا، یوتیوب، یا اینستاگرام)، بیشتر نقش تسهیل‌گری دارد. اگر چه ارتباط بین کاربران می‌تواند وجود داشته باشد، آنجا هدف، قابل مدیریت کردن کشف محتوا برای کاربران و کارآمد کردن ایجاد محتوا (مثلاً، مقالات ویکی‌پدیا)، است. البته، سایت‌ها و خدمت‌رسانی‌های رسانه‌های اجتماعی زیاد دیگری وجود داشته که گرایش یا زمینه خاصی را هدف‌سازی می‌کنند (هنر، موسیقی، عکس‌برداری، مراکز علمی، اماکن جغرافیایی، مذاهب، سرگرمی‌ها، و غیره). این وضعیت نشانگر این امر است که بیشترین تمایل کاربران برخط برای ارتباط، مبتنی بر همسانی گرایش‌ها یا اشتراکات است.

صرف نظر از زمینه‌های به شدت گسترده تمرکز، یک جنبه این خدمات مشترک است: فلسفه وجودی آنها، وجود کاربر و مخاطب است. این جنبه، چیزی است که آنها را از محل‌های «از قبل ایجادشده» اینترنت استاتیک، نظیر سایت‌های خبری رسانه‌های سنتی، صفحه خانگی شرکت‌ها،