

بِسْمِ اللّٰهِ الرَّحْمٰنِ الرَّحِیْمِ

دانشنامه سواد فضای مجازی

جلد سوم

دانش اطلاعاتی

صدرا فیروزمند، حسین غفاری، حسین حق‌پناه







## کلان داده

کلیدواژه‌ها: نافراموشی - کلان داده - ردپای دیجیتال - اینترنت اشیا

### — تشکیل کلان داده

در دنیای آنالوگ، ثبت و نگهداری دانش سخت بود. لذا فقط آن دسته از دانش ذخیره می‌شدند که ارزش آن مشخص و نقد بود اصطلاحاً از جنس «اطلاعات» بود؛ مثل شرح احوال بزرگان، نظرات و مقالات دانشمندان. در عوض، داده‌هایی مثل سوابق بیماری افراد معمولی و یا تعداد قدم‌های روزانه آن شهروند اهمیتی نداشت و در عمل دور ریخته می‌شد.

اما با تحول حاصل از انقلاب دیجیتال علاوه بر آسان شدن ثبت اطلاعات، حجم و فضای فیزیکی لازم برای ذخیره داده نیز کم شد. از این رو میزان داده‌ای که در حافظه‌های دیجیتال ذخیره می‌شد به تدریج افزایش یافت. در چنین شرایطی، همه چیز برای ثبت داده‌هایی که تا آن زمان نگهداری آنها دشوار بود و دور ریخته می‌شد، فراهم شد. لذا بسیاری از «اطلاعات» آنالوگ، مثل کتب قدیمی و نسخ خطی یا فیلم‌ها و تصاویر آرشیوی تاریخی دیجیتال شدند.

## — مثال‌هایی از کلان داده

### + کلان داده‌های مربوط به مشخصات فنی دستگاه‌ها

با گسترش فضای سایبر و مبتنی بر خصلت «نافراموشی» این فضا، حتی کوچک‌ترین کنش و رفتار کاربران در سکو (پلتفرم) های مجازی، امکان ثبت و ذخیره‌سازی دارد: مانند شناسه کاربری افراد، مکان اتصال به سکو و یا حتی حرکت موشواره (ماوس) آن‌ها. این موارد تنها برخی از نمونه‌های داده‌هایی است که به آنها «ردپای دیجیتال» می‌گویند. این ردپاها وقتی کنار هم قرار بگیرند، «کلان داده»<sup>۲۲</sup> را تشکیل می‌دهند که با تحلیل آن می‌توان اطلاعاتی مثل اینکه کدام تصویر یا محتوا بیشتر توجهشان را جلب کرده، پراکندگی جغرافیایی بازدیدکنندگان چگونه است، متناسب با دستگاه (دیوایس) و سیستم عامل کاربران چه تغییرات فنی برای بهینه‌سازی سکو اعمال شود و غیره به دست آورد. اما این پایان مسیر داده‌ها نیست.

### + کلان داده‌های مربوط به رفتار کاربران

با گسترش رسانه‌های اجتماعی، کاربران با فعالیت در آنها علاوه بر ردپاهای دیجیتال فنی، ردپاهای دیجیتال رفتاری نیز از خود به جا می‌گذارند. اینکه چه عبارتی جستجو را کرده‌اند، چه کسانی را دنبال می‌کنند، توسط چه کسانی دنبال می‌شوند، چه محتوایی را پسند (لایک) کرده‌اند یا برایشان نظر (کامنت) گذاشته‌اند، چه عکسی برای نمایه (پروفایل) خود گذاشته‌اند و ... کلان داده‌ای است که از آن می‌توان اطلاعاتی درباره عقاید و علایق کاربران به دست آورد.

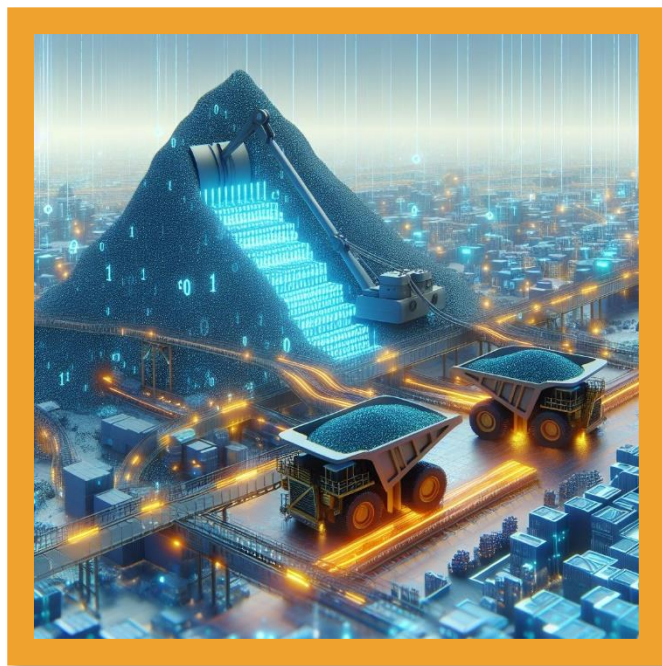
### + کلان داده‌های مربوط به سبک زندگی کاربران

علاوه بر این، فناوری «اینترنت اشیا» (IoT) این بستر را فراهم کرده است که بتوان از ساعت‌های مچی، خودرو، چراغ‌های روشنایی، در و پنجره، جاروبرقی، یخچال و لباسشویی داده‌هایی مثل میزان ضربان قلب و تعداد قدم‌ها، مکان‌هایی که فرد به آنها رفته، ساعات خواب و بیداری، نقشه‌ی خانه و وسایل داخل آن، محتویات یخچال و میزان لباس‌هایش را جمع‌آوری کرد و کلان داده‌ای از سبک زندگی افراد تشکیل داد.

## — کاربرد کلان داده

در نهایت کلان داده حاصل از این ردپاهای دیجیتال، از نوع فنی آن گرفته تا رفتاری و سبک زندگی، به عنوان ماده خام تحلیل در فرایند «داده‌کاوی» و استخراج اطلاعات مورد استفاده در حوزه‌ها و صنایع گوناگونی قرار می‌گیرد.

<sup>۲۲</sup> Bigdata



## داده‌کاوی

کلیدواژه‌ها: داده‌کاوی - تبلیغات رفتارمحور - افکارسنجی - رسوایی فیسبوک و کمبریج آنالیتیکا - برگزیت - مسیریاب‌ها

### — چیستی داده‌کاوی

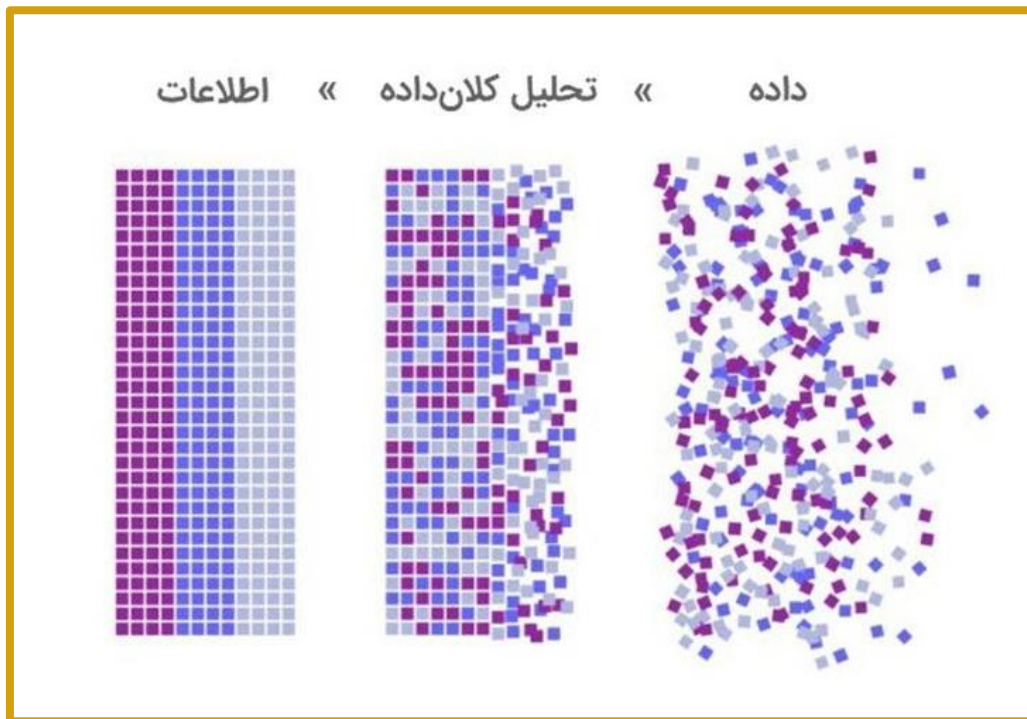
«داده‌کاوی»<sup>۲۳</sup> عبارت است از فرایند پیدا کردن الگوی پنهان میان انبوه داده‌های به ظاهر پراکنده و بی‌معنی و استخراج اطلاعات از آنها. برای مثال اینکه فردی از سوپرمارکت محله‌شان چه تعداد و با چه تنوعی کیک و تنقلات خریداری می‌کند، می‌تواند تعداد اعضای خانواده و فرزندان وی را مشخص کند. یا اینکه با چه فاصله زمانی محصولاتی مثل خمیردندان، مسواک، شامپو، پودر رخت‌شویی و دستمال کاغذی تهیه می‌کند می‌تواند میزان اهمیت دادن به بهداشت و تمیزی در آن خانواده را نشان دهد. خریدن شیر کم‌چرب و پنیر کم‌نمک و مصرف دیر به دیر نمک و روغن، ممکن است به معنی ابتلا به فشار خون یا بیماری قلبی بوده و اهل مراعات سلامتی کردن باشد. و اگر هم آن فرد بدون سابقه قبلی شروع به خرید پوشک بچه کند، می‌توان حدس زد که صاحب فرزند شده و تولدش چه ماهی بوده است. ضمن اینکه ساعت و روزی که معمولاً خریدهایش را انجام می‌دهد یا بانکی که معمولاً با کارت آن مبلغ خرید را حساب می‌کند، می‌تواند نکاتی درباره شغلش داشته باشد...

<sup>۲۳</sup> Data mining

در این مثال صاحب مغازه در واقع داده‌های خامی را که از مشاهداتش درباره خریدهای مشتریان به دست آورده است، تبدیل به اطلاعاتی تقریبی درباره زندگی اهالی محله می‌کند. و از این طریق می‌تواند با تقریب بهتری میزان کالای مورد نیازش را به شرکت‌های پخش‌کننده سفارش دهد. یا با توجه به الگویی که از خرید اهالی محل سراغ دارد، موجود شدن یا تخفیف خوردن کالایی را به آنها یادآوری کند و ... .

این مثال ساده، دقیقاً فرایندی است که فروشگاه‌های بزرگ به منظور مدیریت انبار و زنجیره توزیع خود یا کسب شناخت از ذائقه خریداران و ارائه تخفیف اختصاصی به آنها و در نهایت افزایش فروش و بالا بردن بهره‌وری از طریق داده‌کاوی کلان داده خریدهای مشتریان به کمک هوش مصنوعی انجام می‌دهند.

مثلاً ممکن است فروشگاه‌های از طریق داده‌کاوی این الگو را کشف کند که اغلب مشتریانی که پوشک بچه می‌خرند، بیسکویت مادر هم خریده‌اند. وجه ارتباط این دو کالا با هم مشخص است؛ اما احتمالاً می‌توان حدس زد که مشتریانی هم وجود دارند که پوشک بچه را از قفسه لوازم بهداشتی برداشته‌اند ولی در طول مسیر فراموش کردند به قفسه خوراکی و بیسکویت هم سری بزنند. لذا صاحبان فروشگاه از تحلیل این داده نتیجه می‌گیرند که می‌توان با نزدیک هم قرار دادن قفسه پوشک و بیسکویت، فروش را بالاتر برد.



شکل ۱۱: فرایند تبدیل داده خام به اطلاعات از طریق تحلیل داده‌کاوی

## — کاربردهای داده‌کاوی

برهمن اساس، تحلیل‌های مبتنی بر داده‌کاوی به علت سهولت اجرا و کیفیت نتایج آنها، در حوزه‌ها و صنایع گوناگونی کاربرد دارند:

## — بازاریابی و تبلیغات

سکوی نمایش خانگی (VOD) «نتفلیکس»<sup>۲۴</sup> در سال ۲۰۱۱ به عنوان اولین اقدام خود برای تولید سریال اختصاصی، بیش از ۱۰۰ میلیون دلار روی مجموعه «خانه پوشالی»<sup>۲۵</sup> سرمایه‌گذاری کرد؛ آن‌هم در شرایطی که نخستین فصل سریال دو سال بعد و در سال ۲۰۱۳ قرار بود منتشر شود. در آن زمان بسیاری این سرمایه‌گذاری را نوعی قمار می‌دانستند. اما بعدها مشخص شد که اقدامی آگاهانه و مبتنی بر اطلاعات بوده است:

نتفلیکس پس از بررسی کلان‌داده سوابق تماشای فیلم ۳۳ میلیون کاربری که در آن زمان داشت، دریافت که بخش عمده‌ای از آنها فیلم سینمایی «شبکه اجتماعی» ساخته «دیوید فینچر» را به طور کامل تماشا کرده بودند.

شرکت در ادامه دریافت که اغلب این کاربران فیلم‌هایی که «کیوین اسپیدی» در آنها بازی می‌کند را نیز تماشا کرده‌اند.

شرکت متوجه شد که نسخه انگلیسی سریال «خانه پوشالی» که در سال ۱۹۹۰ ساخته شده بود در انگلستان همچنان پر بیننده است.

بر همین اساس، نتفلیکس تصمیم گرفت روی ساخت نسخه جدیدی از داستان «خانه پوشالی» به کارگردانی «دیوید فینچر» و نقش‌آفرینی «کیوین اسپیدی» به عنوان بازیگر اصلی سرمایه‌گذاری کند. چراکه می‌دانست داستان، کاربران انگلیسی را جذب خواهد کرد و کارگردان و بازیگر، آمریکایی‌ها را!

نتیجه نیز به گونه‌ای بود که طبق برخی تخمین‌های غیررسمی، تعداد کاربران نتفلیکس بعد از فصل اول این سریال در سال ۲۰۱۳، به ۵۷ میلیون مشترک افزایش پیدا کرد!<sup>۲۶</sup> و ۲۷ و ۲۸

<sup>۲۴</sup> «نتفلیکس» (Netflix) یک شرکت آمریکایی در حوزه نمایش فیلم است. این شرکت در سال ۱۹۹۷ تاسیس شد و در آغاز به عنوان یک سرویس فروش و اجاره دی‌وی‌دی فیلم فعالیت می‌کرد. در سال ۲۰۰۷ اما از قابلیت پخش جاری (استریم) فیلم و سریال خود رونمایی کرد. نتفلیکس اکنون به عنوان بزرگترین سکوی نمایش خانگی دنیا با بیش از ۲۴۰ میلیون کاربر فعالیت می‌کند.

<sup>۲۵</sup> House of Cards

<sup>۲۶</sup> <https://builtin.com/big-data/big-data-examples-applications>

<sup>۲۷</sup> <https://b2n.ir/netflix-bigdata>

<sup>۲۸</sup> <https://b2n.ir/netflix-house-of-cards>

البته استفاده از سوابق کاربران در زمینه بازاریابی داده‌کاوی به گونه‌های دیگری هم کاربرد دارد. برای مثال سکوهایی مثل «دیجی‌کالا» یا «طاقچه» بر مبنای داده‌کاوی روی سوابق خرید و جستجوی کاربرانشان، کتاب‌ها یا کالاهای خاصی را تحت عناوینی مثل «پیشنهاد ویژه برای شما» به طور اختصاصی برای آن مخاطب نمایش می‌دهند.

همچنین شرکت‌های تبلیغات اینترنتی مثل «یکتانت» با نصب «کلوچک» (کوکی) روی مرورگر مراجعه‌کنندگان به تارنماهای تحت پوشش شرکتشان، داده ردپای دیجیتال کاربران (تارنماهایی که بازدید می‌کنند، جستجوهای که دارند و ...) را جمع‌آوری و سپس با داده‌کاوی روی آن، سرویس «تبلیغات رفتارمحور»<sup>۲۹</sup> ارائه می‌کنند. بدین صورت که کاربران را بر اساس شاخص‌هایی مثل مرد یا زن بودن، خانه‌دار بودن یا نبودن، کنکوری بودن یا نبودن، داشتن قصد مهاجرت، علاقه‌مندی به کاهش وزن و فیتنس، اهل مد و زیبایی بودن، داشتن فرزند خردسال و حتی مذهبی بودن یا نبودن بخش‌بندی می‌کند و از این طریق به سفارش‌دهندگان تبلیغ و فروشندگان کالا این امکان را می‌دهد که ویژگی‌های مخاطب هدف تبلیغشان را انتخاب کنند و محتوا تبلیغی را به صورت هدفمند و فقط به آن افراد نمایش دهند.

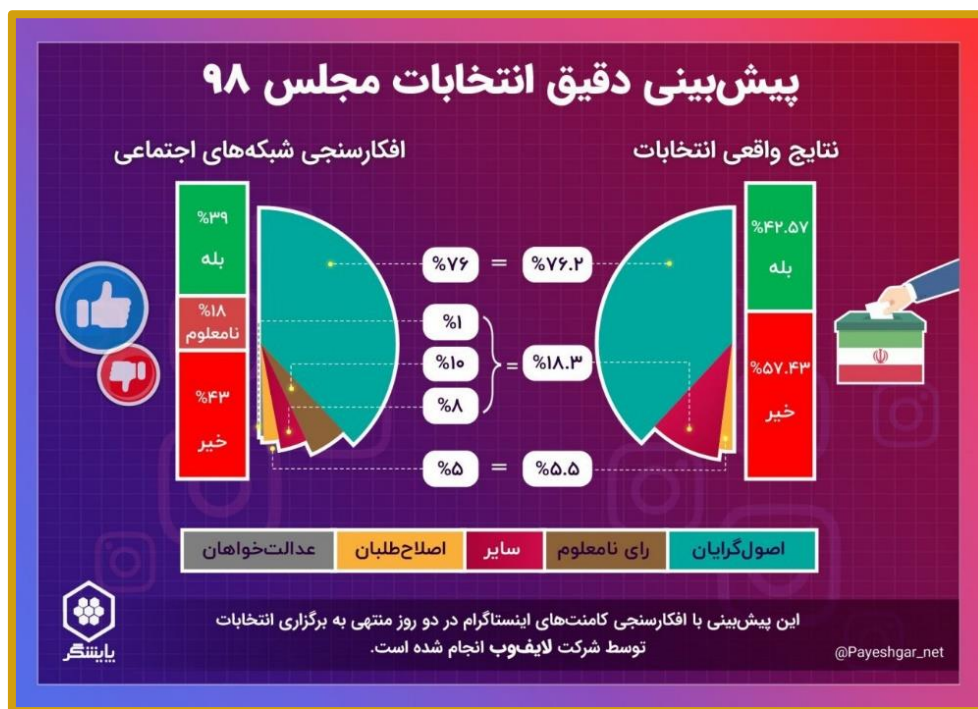
شکل ۱۲: از طریق سرویس تبلیغات رفتارمحور یکتانت، تبلیغ‌ها به صورت اختصاصی هر کاربر و براساس سوابق خرید و جستجوی وی نمایش داده می‌شود.

<sup>۲۹</sup> <https://www.yektanet.com/blog/11332>

## — سیاست و حکمرانی

مثال جنجالی و پر سروصدای استفاده از داده‌کاوی در سیاست به ماجرای انتقال اطلاعات ۵۰ میلیون کاربران آمریکایی فیسبوک به شرکت تحلیل داده کمبریج آنالیتیکا و سپس ارسال تبلیغات سیاسی هدفمند و متناسب با شخصیت هر کاربر در راستای کمپین انتخاباتی دونالد ترامپ برمی‌گردد که در نهایت باعث پیروزی غیرمنتظره وی در انتخابات ریاست جمهوری سال ۲۰۱۶ این کشور شد!<sup>۳۰</sup> درست همان سال، مشابه همین اتفاق در همه‌پرسی موسوم به «برگزیت» رخ داد و منجر به رای آوردن طرح خروج انگلیس از اتحادیه اروپا شد!<sup>۳۱</sup>

هم‌اکنون در ایران نیز برخی شرکت‌های دانش‌بنیان فعال در حوزه تحلیل داده، پیش از انتخابات‌های مجلس و ریاست‌جمهوری اقدام به جمع‌آوری داده‌های شبکه‌های اجتماعی (مثلاً نظرات و محتوای کاربران له یا علیه یک نامزد یا یک جریان سیاسی) می‌کنند و با تحلیل آنها، میزان مشارکت در انتخابات و حتی نتیجه نهایی آن را پیش‌بینی می‌کنند.<sup>۳۲ و ۳۳</sup>



شکل ۱۳: مقایسه پیش‌بینی سامانه تحلیل داده «پایشگر» از میزان مشارکت و جناح سیاسی منتخب انتخابات مجلس سال ۱۳۹۸ (برپایه تحلیل نظرات اینستاگرام در دو روز منتهی به انتخابات) با نتایج واقعی و نهایی آن<sup>۳۴</sup>

<sup>۳۰</sup> این ماجرا آنقدر در آمریکا بالا گرفت که مارک زاکربرگ، مالک شبکه اجتماعی فیسبوک، مجبور شد در جلسه استماع کنگره آمریکا رسماً بابت نقض حریم خصوصی کاربران - و احتمالاً کمک به رای آوردن ترامپ - عذرخواهی کند!  
<sup>۳۱</sup> فیلم سینمایی «برگزیت» (Brexit) ساخته سال ۲۰۱۹ به چگونگی مداخله در این همه‌پرسی و اثرگذاری احزاب سیاسی بر نظر رای‌دهندگان به کمک تحلیل داده و ارسال تبلیغات اختصاصی در شبکه‌های اجتماعی می‌پردازد.

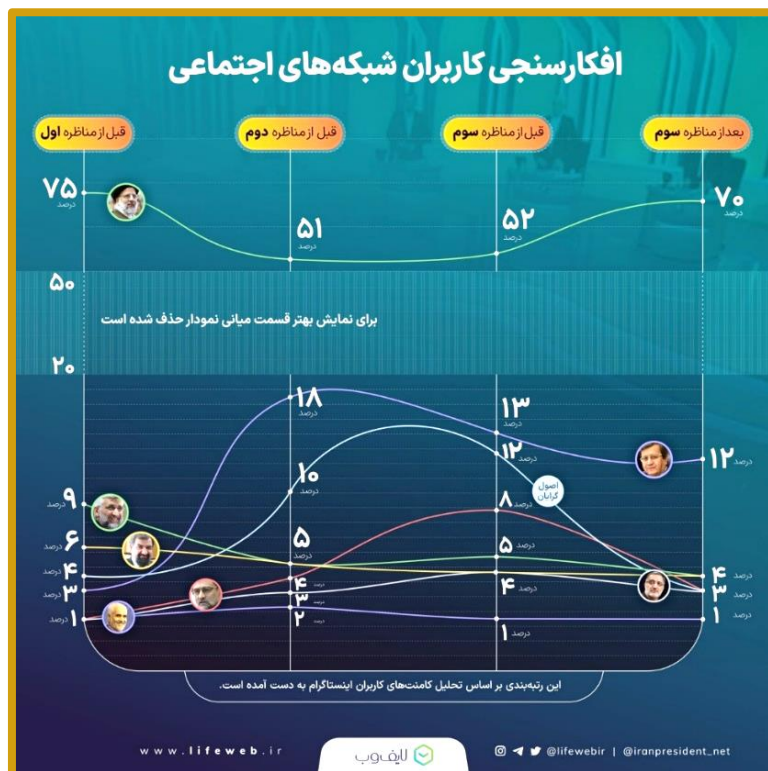
<sup>۳۲</sup> <https://peivast.com/p/103154>

<sup>۳۳</sup> <https://t.me/LifewebIR/115>

<sup>۳۴</sup> <https://t.me/LifewebIR/1828>



شکل ۱۴: نتیجه داده‌کاوی مرکز «بتا» درباره تحلیل احساسات کاربران توییتر بعد از اولین مناظره انتخابات ریاست جمهوری سال ۱۴۰۰ از طریق داده‌کاوی روی ۱/۷ میلیون توییت<sup>۳۵</sup>

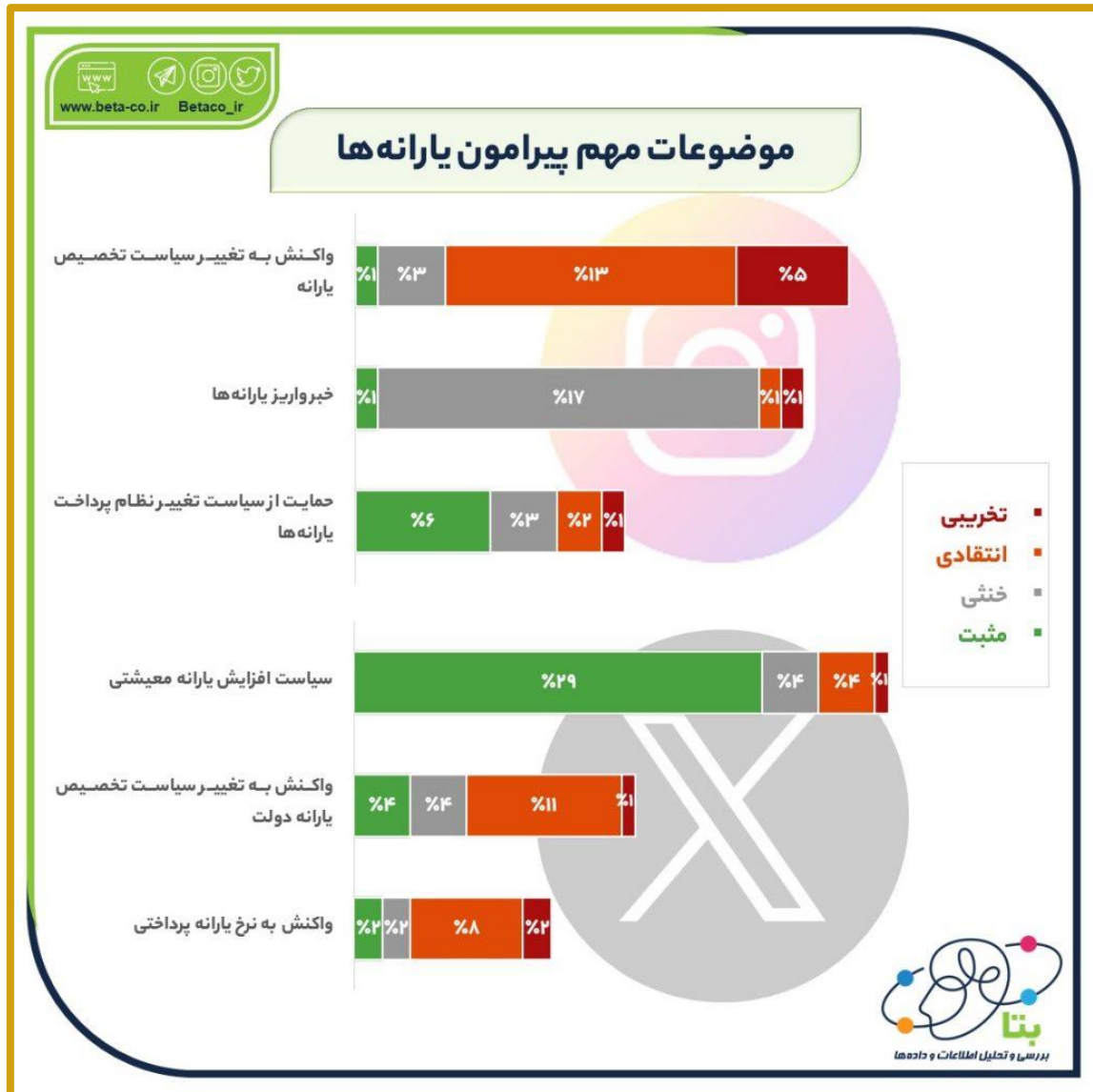


شکل ۱۵: بررسی روند تغییرات محبوبیت نامزدهای انتخابات ریاست جمهوری ۱۴۰۰ قبل و بعد از مناظره‌های تلویزیونی بر اساس تحلیل نظرات کاربران اینستاگرام توسط شرکت «لایف وب»<sup>۳۶</sup>

<sup>۳۵</sup> <https://b2n.ir/beta-co-election1400>

<sup>۳۶</sup> <https://t.me/LifewebIR/2359>

علاوه بر پیش‌بینی نتایج انتخابات، به کمک داده‌کاوی می‌توان اقدام به افکارسنجی از کاربران درباره موضوعات روز جامعه کرد و احیاناً اطلاعات حاصل از آن را در تصمیم‌گیری‌های حکومتی به کار بست: برای مثال بررسی واکنش کاربران به سیاست پرداخت یارانه<sup>۳۷</sup> یا تحلیل عواطف کاربران توئیتر نسبت به مشکلات گاز در زمستان<sup>۳۸</sup> یا نظر کاربران درباره مهاجران افغانستانی<sup>۳۹</sup>

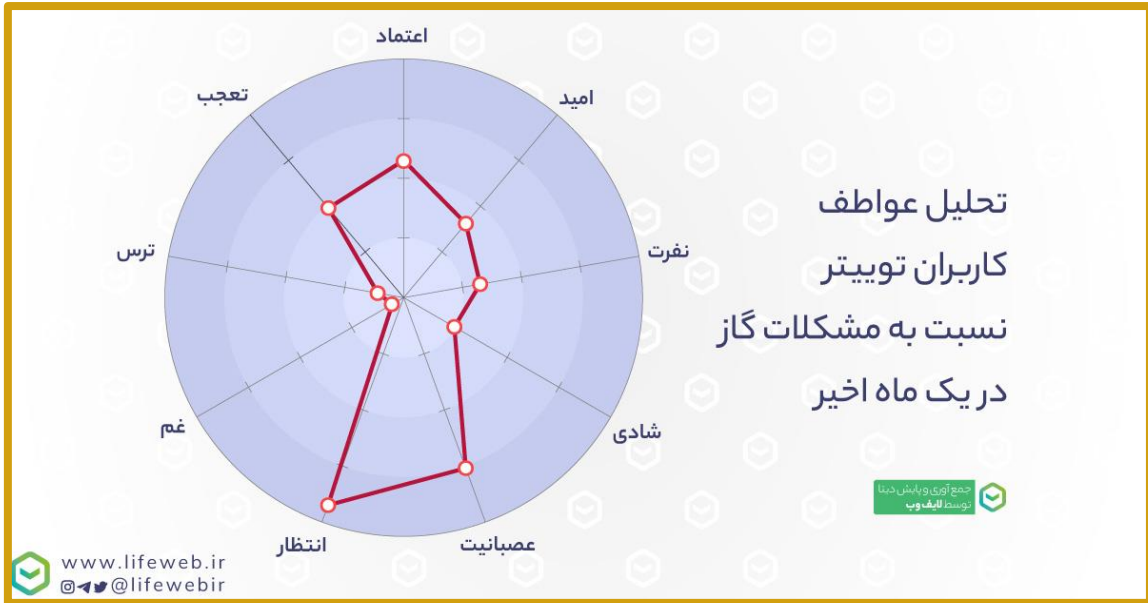


شکل ۱۶: تحلیل بیش از یک میلیون محتوا منتشر شده در توئیتر، اینستاگرام و تلگرام درباره تغییر سیاست‌های پرداخت یارانه در سال ۱۴۰۱

<sup>۳۷</sup> <https://b2n.ir/beta-co-yarane>

<sup>۳۸</sup> <https://lifewebco.com/gas-crisis/>

<sup>۳۹</sup> [https://lifewebco.com/afghan\\_immigrants/](https://lifewebco.com/afghan_immigrants/)



شکل ۱۷: گراف تحلیل توییت‌های کاربران درباره مشکلات گاز در زمستان ۱۴۰۱



شکل ۱۸: نمودارهای تحلیل نظرات کاربران ایکس (توییت‌ها) درباره مهاجران افغان

در کنار افکارسنجی، جمع‌آوری داده و تشکیل کلان‌داده قابلیت‌های ویژه‌ای برای نهادهای نظارتی در زمینه مبارزه با فساد می‌دهد: مثلاً در شیروان فردی که در یک ماه ۱۰ هزار قرص نان را از یک نانوایی خریده و برای مصرف دام و طیور خشکانده بود از طریق بررسی داده کارتهای هوشمند نانوایی‌ها شناسایی و بازداشت شد.<sup>۴۰</sup>

## — بهداشت و درمان

گونه دیگری از داده‌ها که ارزش بسیار بالایی برای داده‌کاوی دارند، داده‌های زیستی - نظیر سوابق بیماری افراد یا داروهای مصرفی یا نتایج آزمایش‌های پزشکی یا حتی ردپاهای دیجیتال زیستی مثل ضربان قلب، میزان اکسیژن خون، تعداد قدم‌ها، ساعات بیداری، کیفیت خواب و غیره که موبایل‌ها و ساعت‌های هوشمند جمع‌آوری می‌کنند - هستند و از آنها می‌توان در حوزه بهداشت و درمان استفاده کرد. برای مثال شرکت‌های بیمه‌ای از طریق تحلیل تاریخچه بیماری‌های فرد، نوع درمان، هزینه‌ها و غیره می‌توانند موارد مشکوک به تقلب<sup>۴۱</sup> را تشخیص دهند.<sup>۴۲</sup>

همچنین از طریق داده‌کاوی می‌توان مشخص کرد که چه افرادی بیشتر به چه بیماری‌هایی مبتلا می‌شوند، چه بیماری‌هایی در آستانه شیوع قرار دارند و برای این منظور چه اقدامات پیشگیرانه‌ای باید اتخاذ گردد. نمونه واقعی چنین کاربردی در همه‌گیری کرونا استفاده شد:

سامانه هوشمند تشخیص کرونا که توسط دانشجویان دانشگاه شریف ساخته شده بود براین اساس کار می‌کرد که تصاویر سی‌تی‌اسکن ریه افراد در آن بارگذاری می‌شد و این سامانه آن را با الگویی که از طریق یادگیری عمیق روی تصاویر سی‌تی‌اسکن ریه تعداد زیادی از مبتلایان کشف کرده بود مقایسه کرده و نتیجه ابتلا را بدون نیاز به پزشک مشخص می‌کرد.<sup>۴۳</sup>

مسیریاب نشان با تحلیل داده موقعیت مکانی کاربرانی که با سرعت پایین حرکت می‌کنند (پس یعنی احتمالاً پیاده هستند)، نواحی پرتردد پیاده را - که احتمال ابتلا به ویروس در آنها بیشتر است - شناسایی کرده و روی نقشه مشخص می‌کرد تا افراد در صورت نیاز به خروج از خانه، از نواحی عبور کنند که ترافیک انسانی کمتری دارد.<sup>۴۴</sup>

پژوهشی که نشان می‌داد از بررسی میزان جستجوی کلمات مرتبط با کرونا در موتور جستجوی گوگل توسط مردم، می‌توان تخمینی از زمان وقوع و میزان ابتلا و تلفات موج بعدی بیماری پیدا کرد.<sup>۴۵</sup>

<sup>۴۰</sup> <https://akharinkhabar.ir/social/9820340>

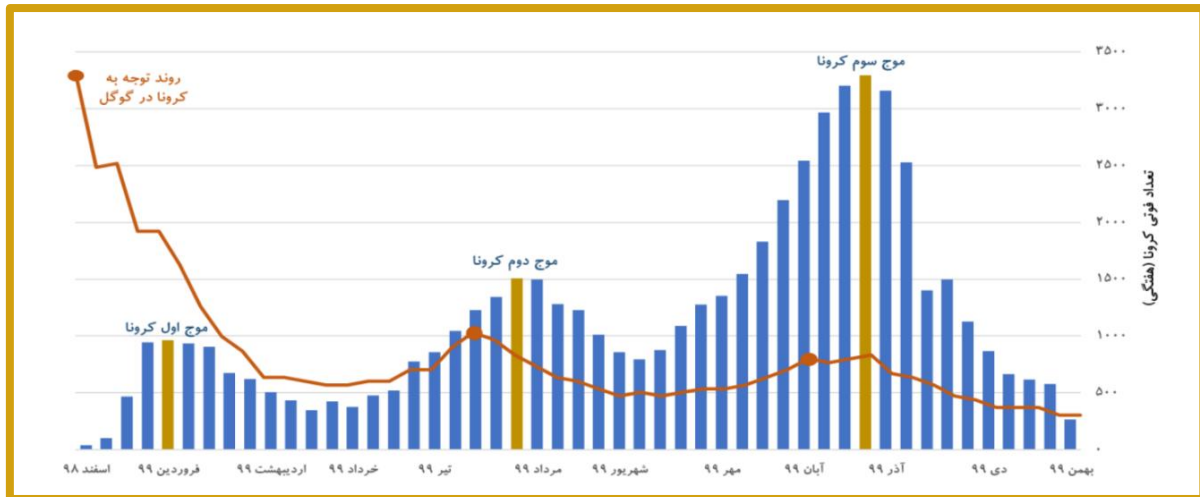
<sup>۴۱</sup> Fraud Detection

<sup>۴۲</sup> <https://b2n.ir/dataak-datamining-health>

<sup>۴۳</sup> <https://mehrnews.com/xRBrM>

<sup>۴۴</sup> <https://vrgl.ir/fb3nd>

<sup>۴۵</sup> <http://payeshjournal.ir/article-1-1528-fa.html>



شکل ۱۹: مقایسه میزان جستجوی مرتبط با کرونا در گوگل توسط مردم و تعداد فوتی‌های کرونا نشان می‌دهد تقریباً ۱۰ الی ۲۰ روز بعد از اوج جستجوها (ابتلای مردم بیماری و جستجو درباره درمان آن) شاهد نقطه اوج فوتی‌ها هستیم.

گونه دیگری از داده‌های زیستی که بسیار مورد توجه پلتفرم‌هاست، تصاویر چهره است. چراکه از این تصاویر می‌توان به عنوان داده ورودی یادگیری ماشین برای ساخت سامانه‌های تشخیص چهره و حتی هوش مصنوعی و ساخت تصاویر و ویدیوهای «جعل عمیق» (Deep Fake) استفاده کرد؛ هوش مصنوعی با تلفیق زدن از تصاویر اولیه‌ای که دارد، تصاویر جدید تولید می‌کند... لذا هر چه تعداد این تصاویر بالاتر باشد تصاویر بهتر و بدیع‌تری خلق خواهد کرد.

از این رو برخی شرکت‌ها تصاویری که کاربران در شبکه‌های اجتماعی به اشتراک می‌گذارند را جمع‌آوری کرده و کلان‌داده آن را برای چنین استفاده‌هایی به فروش می‌گذارند. کما اینکه شرکت متا (فیس‌بوک سابق) در صفحه معرفی پروژه تشخیص چهره‌ای که در سال ۲۰۱۴/۱۳۹۳ با نام «دیپ‌فیس»<sup>۴۶</sup> راه‌اندازی کرده بود، عنوان می‌کند که الگوریتم آن از داده‌کاوی روی چهار میلیون تصویر چهره ایجاد شده است!<sup>۴۷</sup>

شرکت‌های کوچک‌تر اما از روش‌های خلاقانه دیگری برای جمع‌آوری آسان و گسترده تصاویر چهره‌ها استفاده می‌کنند: برای مثال برنامهک (اپلیکیشن) «فیس‌آپ» با ظاهر سرگرم‌کننده خود که در آن کاربران تصاویر چهره‌شان را بارگذاری کرده و نسخه پیر یا جوان شده آن را دریافت می‌کردند توانسته بود تنها دو سال بعد از انتشار، بیش از ۱۵۰ میلیون عکس چهره جمع‌آوری کند.<sup>۴۸</sup> این موضوع زمانی در جهان - خصوصاً آمریکا - جنجالی شد که مشخص شد مالکیت این برنامهک در اختیار شرکت «وایرلس لَب»<sup>۴۹</sup>

<sup>۴۶</sup> شرکت متا ادعا می‌کند الگوریتمی که در پروژه «DeepFace» توسعه داده است، می‌تواند ویژگی‌هایی نظیر جنسیت، سن، احساس و نژاد چهره‌ها را با دقت ۷/۹۳٪ شناسایی کند. [منبع: <https://towardsdatascience.com/using-deepface-for-face-recognition-5f8d1e43f2a6>]

<sup>۴۷</sup> <https://b2n.ir/deepface-facebook>

<sup>۴۸</sup> <https://b2n.ir/faceapp-forbes>

<sup>۴۹</sup> Wireless Lab

روسیه است. از اینرو نگرانی از احتمال استفاده جانبی و بدون نظارت - آمریکایی‌ها - از تصاویر ذخیره شده روی سرورهای روس، واکنش کارشناسان آمریکایی را برانگیخت<sup>۵۰</sup> و<sup>۵۱</sup> «چاک شومر» - سناتور و رهبر اکثریت دموکرات سنای آمریکا - از «اداره تحقیقات فدرال» (FBI) خواست درباره این برنامه به عنوان یک تهدید سایبری برای ایالات متحده تحقیق کند<sup>۵۲</sup> و افبی‌آی نیز در پاسخ به وی، برنامه فیس‌آپ را «تهدید بالقوه جاسوسی» دانست.<sup>۵۳</sup>



شکل ۲۰: نمونه‌ای از خروجی مازول پیر کردن عکس توسط اپلیکیشن فیس‌آپ.



شکل ۲۱: توییت آقای چاک شومر درباره درخواست از نهادهای امنیتی آمریکا برای تحقیق در زمینه احتمال سوءاستفاده اپلیکیشن روسی فیس‌آپ از تصاویر کاربران آمریکایی

<sup>۵۰</sup> <https://b2n.ir/faceapp-technologyreview>

<sup>۵۱</sup> <https://b2n.ir/faceapp-onezeromedium>

<sup>۵۲</sup> <https://b2n.ir/chuck-schumer-tweet>

<sup>۵۳</sup> <https://www.reuters.com/article/idUSKBN1Y62D0>

از این مثال می‌توان ارزش، کاربرد و اهمیت سایر داده‌های زیستی مثل اثر انگشت، ضربان قلب، صدا، عنبیه و قرنیه که افشای جمع‌آوری هر کدام از آنها توسط شرکت‌هایی مثل «تیک‌تاک»<sup>۵۴</sup> و «فیس‌بوک»<sup>۵۶</sup>، «المارت»<sup>۵۷</sup>، موضوع شکایت‌ها و صدور جریمه‌های مختلفی برای آنها در آمریکا بوده است را متوجه شد.

## — حمل و نقل

نرم‌افزارهای مسیریاب مثل «نشان»، «بلد» یا «گوگل مپس» در مواردی ممکن است برای نمایش ترافیک زنده خیابان‌ها از اطلاعات محلی مثل دوربین‌های ترافیکی پلیس یا محدودیت‌های ترافیکی اعلام شده توسط نهادهای شهری استفاده کنند؛ اما مسلماً کافی نیست، چون در همه معابر چنین اطلاعاتی وجود ندارد. علاوه‌براین آنها از گزارش کاربران نرم‌افزار مبنی بر وجود ترافیک یا رخ دادن تصادف در مسیر نیز استفاده می‌کنند؛ که باز هم به دلایل مختلف نمی‌توان به آن اطمینان کامل داشت. اما روش سوم که رایج‌تر و کاربردی‌تر نیز می‌باشد، داده‌کاوی روی کلان‌داده موقعیت مکانی کاربران است.<sup>۵۸</sup> بدین صورت که داده GPS تمام گوشی‌های موبایل را جمع‌آوری کرده و با تحلیل تعداد، موقعیت و سرعت جابه‌جایی گوشی‌های حاضر در یک خیابان تشخیص می‌دهند که آیا آنجا ترافیک هست یا خیر.<sup>۵۹</sup> و ۶۰ و ۶۱

علاوه‌براین، کلان‌داده موقعیت مکانی کاربران می‌تواند کاربردهای دیگری هم داشته باشد. مثلاً:

- مسیریاب «نشان» در اولین گزارش سالانه خود براساس تحلیل داده‌های ترافیکی و ساعت حرکت کاربران، سحرخیزترین و شب‌زنده‌دارترین شهرهای ایران را به ترتیب کرج و قم معرفی کرده بود.<sup>۶۲</sup>
- همچنین از طریق مقایسه آرشیو داده‌های ترافیکی روزهای مختلف می‌توان تشخیص داد تصمیمات و حوادث (از قبیل اجرای طرح زوج و فرد یا تعطیلی مدارس یا بارش برف) چه اثری بر عبور و مرور خودروها در معابر شهر داشته است.

<sup>۵۴</sup> <https://b2n.ir/tiktok-findbiometrics>

<sup>۵۵</sup> <https://tcrn.ch/2Rn4Egk>

<sup>۵۶</sup> <https://www.nevis.net/en/blog/are-selfie-apps-and-filter-a-threat-to-online-security>

<sup>۵۷</sup> <https://b2n.ir/walmart-classaction>

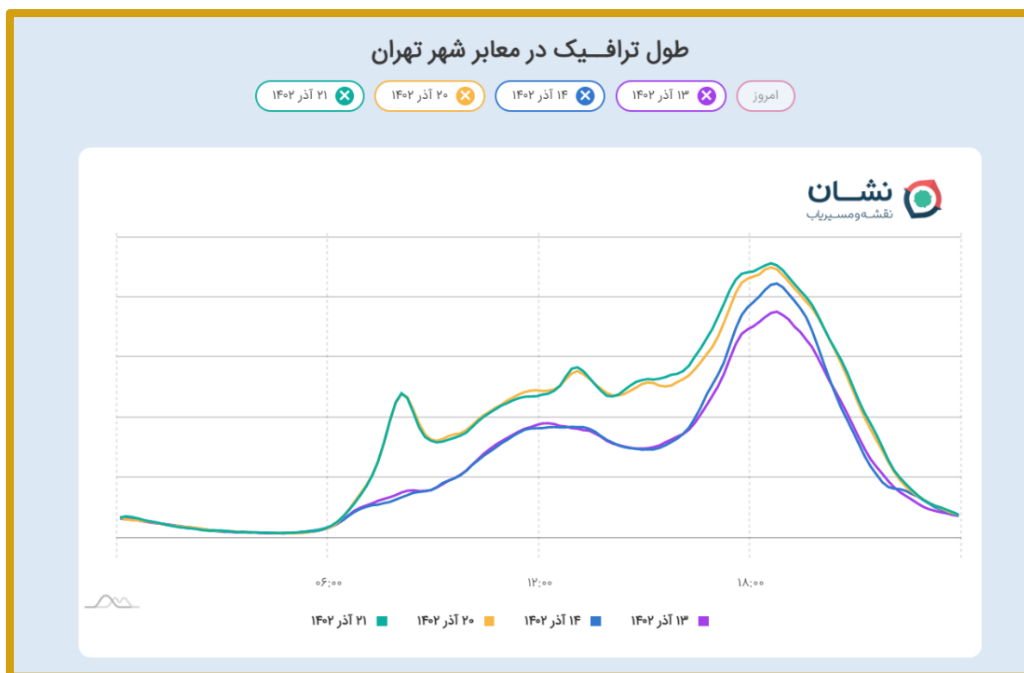
<sup>۵۸</sup> <https://article.tebyan.net/483566>

<sup>۵۹</sup> برای بررسی صحت این موضوع، یک کاربر رند آلمانی آزمایشی طراحی کرد و ۹۹ گوشی موبایل را در یک چرخ‌دستی قرار داد و در سطح شهر حرکت کرد! نتیجه کار را در این ویدیو ببینید: [apararat.com/v/zscl8](https://www.aparat.com/v/zscl8)

<sup>۶۰</sup> اگر با گوشی اندروید خود در حساب گوگل خود وارد شده باشید، می‌توانید آرشیو تمام موقعیت‌ها و مکان‌هایی که در آنها بوده‌اید در این نشانی ببینید: [maps.google.com/locationhistory](https://maps.google.com/locationhistory)

<sup>۶۱</sup> <https://dgto.ir/1f10>

<sup>۶۲</sup> <https://vrgl.ir/dnym8>



شکل ۲۲: مقایسه میانگین ترافیک معابر شهر تهران در روزهای دوشنبه و سه‌شنبه ۱۳ و ۱۴ آذر (تعطیلی مدارس به علت آلودگی هوا) و ۲۰ و ۲۱ آذر (وضعیت عادی) به وضوح نشان می‌دهد تعطیلی شدن مدرسه‌ها علاوه بر حذف اوج شلوغی صبح تهران، بار ترافیکی شهر در طول روز را نیز به طور محسوسی کاهش می‌دهد.<sup>۶۳</sup>

<sup>۶۳</sup> <https://traffic.neshan.org>