

کلان داده‌ها در علوم اجتماعی

روش‌ها و ابزارهای علم داده برای پژوهشگران و مجریان

مؤلف:

لان فاستر

مترجمان:

مهدی اوریا

اکبر اصغرزاده طباطبائی



انتشارات دانشگاه فرماندهی و ستاد ارتش جمهوری اسلامی ایران

۱۴۰۳

عنوان و نام پدیدآور	: کلان داده‌ها در علوم اجتماعی (روش‌ها و ابزارهای علم داده برای پژوهشگران و مجریان) / مولف [صحیح: ویراستاران] لان فاستر... [و دیگران] ؛ مترجم مهدی اوریا، اکبر اصغرزاده بناب.
مشخصات نشر	: تهران: ارتش جمهوری اسلامی ایران، دانشگاه فرماندهی و ستاد، انتشارات دافوس، ۱۴۰۳.
مشخصات ظاهری	: ۵۶۶ ص.: جدول (رنگی)، نمودار (رنگی).
شابک	: ۹۷۸-۶۲۲-۸۳۳۰-۷۹-۲
وضعیت فهرست نویسی	: فیپا
یادداشت	: عنوان اصلی: Big data and social science : data science methods and tools for research and practice, 2 nd ed., 2024.
یادداشت	: کتابنامه: ۵۲۲-۵۶۶.
عنوان دیگر	: روش‌ها و ابزارهای علم داده برای پژوهشگران و مجریان.
موضوع	: علوم اجتماعی -- داده‌پردازی Social sciences -- Data processing علوم اجتماعی -- روش‌های آماری Social sciences -- Statistical methods داده‌کاوی Data mining داده‌های کلان Big data
شناسه افزوده	: فاستر، ایان، ۱۹۵۹-م. Foster, Ian, 1959
شناسه افزوده	: اوریا، مهدی، ۱۳۵۳- مترجم
شناسه افزوده	: اصغرزاده بناب، اکبر، ۱۳۶۳- مترجم
شناسه افزوده	: دانشگاه فرماندهی و ستاد آجا. انتشارات دافوس
اطلاعات رکورد کتابشناسی	: فیپا

عنوان: کلان داده‌ها در علوم اجتماعی: روش‌ها و ابزارهای علم داده برای پژوهشگران و مجریان

مترجمان: مهدی اوریا و اکبر اصغرزاده بناب

طراح جلد: علیرضا اکبرپور

صفحه‌آرایی: امیرحسین رضائی

ناشر: دافوس

شمارگان: ۱۰۰۰

تعداد صفحه: ۵۶۶ ص

نوبت چاپ: چاپ اول

تاریخ انتشار: ۱۴۰۳

چاپ و صحافی: مدیریت چاپ، انتشارات و فصلنامه دانشگاه فرماندهی و ستاد آجا

قیمت: ۵۹۰۰۰۰۰ ریال

نشانی: تهران، میدان پاستور، خیابان دانشگاه جنگ، دانشگاه فرماندهی و ستاد، انتشارات دافوس

تلفن: ۰۲۱-۶۶۴۱۴۱۹۱، ۰۲۱-۶۶۴۷۰۴۸۶

مسئولیت صحت مطالب بر عهده مترجمان می‌باشد.

کلیه حقوق برای دافوس آجا محفوظ است. (نقل مطالب با ذکر مأخذ بلامانع است).

فهرست مطالب

عنوان مطلب.....	صفحه
فصل اول: مقدمه.....	۲۱.....
چرا این کتاب؟.....	۲۲.....
تعریف کلان داده و ارزش آن.....	۲۵.....
اهمیت استنتاج.....	۲۷.....
اهمیت درک چگونگی تولید داده ها.....	۳۲.....
ابزارهای جدید برای داده های جدید.....	۳۵.....
مورد استفاده کتاب.....	۳۶.....
ساختار کتاب.....	۴۳.....
بخش اول: جمع آوری و مدیریت داده ها.....	۴۳.....
بخش دوم: مدل سازی و تحلیل.....	۴۶.....
بخش سوم: استنباط و اخلاق.....	۴۸.....
منابع.....	۵۰.....
فصل دوم: کار با داده های وب و API ها.....	۵۳.....
مقدمه.....	۵۴.....
استخراج اطلاعات از وب.....	۵۶.....
به دست آوردن داده ها از وب سایت ها.....	۵۷.....
محدودیت های اسکریپتینگ.....	۶۶.....
رابط های برنامه نویسی کاربردی (APIs).....	۶۷.....
API ها و منابع مرتبط.....	۶۷.....
RESTFUL APIs، داده های بازگشتی، و ریرهای پایتون.....	۶۹.....
استفاده از یک API.....	۷۱.....

۴۲۹	 منابع سوگیری
۴۳۰	 سوگیری نمونه
۴۳۱	 سوگیری برچسب (نتیجه)
۴۳۲	 سوگیری در فرآیند یادگیری ماشین
۴۳۵	 سوگیری در کاربرد
۴۳۶	 در نظر گرفتن سوگیری هنگام پیاده‌سازی مدل
۴۳۶	 مقابله با سوگیری
۴۳۶	 تعریف سوگیری
۴۳۷	 تعاریف
۴۴۰	 انتخاب معیارهای سوگیری
۴۴۲	 مثال تشبیهی
۴۴۳	 تعداد مثبت‌های کاذب
۴۴۴	 مثبت‌های کاذب بر اساس با اندازه گروه
۴۴۵	 نرخ کشف نادرست (FDR)
۴۴۵	 نرخ مثبت کاذب (FPR)
۴۴۷	 موازنه‌ها در انتخاب معیار
۴۴۹	 مثال حمایتی
۴۴۹	 تعداد منفی‌های کاذب
۴۵۰	 منفی‌های کاذب تعدیل‌شده بر اساس اندازه گروه
۴۵۱	 نرخ حذف کاذب (FOR)
۴۵۲	 نرخ منفی کاذب (FNR)
۴۵۳	 مورد خاص: برنامه‌های با منابع محدود
۴۵۵	 کاهش سوگیری
۴۵۵	 ممیزی نتایج مدل

۴۵۶	انتخاب مدل
۴۵۸	گزینه‌های دیگر برای کاهش سوگیری
۴۶۰	ملاحظات بیشتر
۴۶۰	مقایسه با چه چیزی؟
۴۶۱	هزینه‌های هر دو نوع خطا
۴۶۲	جمعیت مرتبط کدام است؟
۴۶۲	نتایج پیوسته
۴۶۳	ملاحظات برای اندازه‌گیری مستمر
۴۶۴	عدالت در عمل
۴۶۵	نام‌های دیگری که ممکن است ببینید
۴۶۸	مطالعات موردی
۴۶۸	پیش‌بینی عود جرم با استفاده از COMPAS
۴۷۰	تشخیص چهره
۴۷۴	هدف‌گیری تبلیغات در فیسبوک
۴۷۶	AEQUITAS - یک ابزار برای بررسی سوگیری و عدالت در مدل‌های یادگیری ماشین
۴۸۱	فصل دوازدهم: حریم خصوصی و حرمانگی
۴۸۲	مقدمه
۴۹۱	چرا دسترسی به داده‌ها مهم است؟
۴۹۴	فراهم کردن دسترسی
۵۰۰	داده‌های غیرجدولی
۵۰۱	چالش‌های جدید
۵۰۴	چارچوب قانونی و اخلاقی
۵۰۷	خلاصه
۵۰۹	منابع

۵۱۱	فصل سیزدهم: کتاب‌های کار
۵۱۲	مقدمه
۵۱۴	دفترچه‌ها
۵۱۴	پایگاه داده‌ها
۵۱۴	بررسی و مصورسازی داده‌ها
۵۱۵	API ها
۵۱۵	پیوند رکورد
۵۱۶	تحلیل متن
۵۱۶	شبکه‌ها
۵۱۷	یادگیری ماشین - ایجاد برجسب‌ها
۵۱۷	یادگیری ماشین - ایجاد ویژگی‌ها
۵۱۸	یادگیری ماشین - آموزش و ارزیابی مدل
۵۱۸	سوگیری و عدالت
۵۱۹	خطاها و استنتاج
۵۱۹	دفترچه‌های اضافی
۵۲۱	منابع و مراجع

در عصر حاضر، کلان داده‌ها^۱ به یکی از ارکان اساسی در تحول و پیشرفت علوم مختلف تبدیل شده‌اند. این پدیده، به ویژه در حوزه علوم اجتماعی، فرصتی بی‌نظیر برای پژوهشگران و مجریان فراهم آورده است تا با بهره‌گیری از داده‌های وسیع و متنوع، به تحلیل رفتارها، تعاملات و ساختارهای اجتماعی بپردازند. کلان داده‌ها نه تنها به ما این امکان را می‌دهند که الگوها و روندهای جدیدی را شناسایی کنیم، بلکه به ما کمک می‌کنند تا درک عمیق‌تری از پیچیدگی‌های جامعه و چالش‌های آن به دست آوریم. با پیشرفت فناوری‌های اطلاعات و ارتباطات، حجم و تنوع داده‌ها به طرز چشمگیری افزایش یافته است. این داده‌ها از منابع مختلفی مانند شبکه‌های اجتماعی، نظرسنجی‌ها، انواع پایگاه‌های داده و داده‌های حسگرهای اینترنت اشیاء^۲ جمع‌آوری می‌شوند. این تنوع و حجم بالای داده‌ها، پژوهشگران را به چالش می‌کشد تا از روش‌ها و ابزارهای مناسب علم داده برای تحلیل و تفسیر این داده‌ها استفاده کنند.^۳ این کتاب به بررسی روش‌ها و ابزارهای علم داده پرداخته است و به پژوهشگران علوم اجتماعی کمک می‌کند تا کلان داده‌ها بهره‌برداری کنند. از تحلیل داده‌های آماری و یادگیری ماشین^۴ گرفته تا تحلیل متن^۵ و تحلیل شبکه‌های اجتماعی^۶، هر یک از این ابزارها به ما این امکان را می‌دهند که به درک بهتری از رفتارها و الگوهای اجتماعی دست یابیم. همچنین، تصویربرداری داده‌ها^۶ به ما کمک می‌کند تا نتایج خود را به صورت بصری و قابل فهم ارائه دهیم.

هدف کتاب، ارائه یک راهنمای جامع برای پژوهشگران و مجریان علوم اجتماعی است. کلان داده‌ها و علم داده، امکانات بی‌نظیری برای پژوهشگران و مجریان علوم اجتماعی فراهم کرده است تا پژوهشگران بتوانند به درک عمیق‌تری از رفتارها و الگوهای اجتماعی دست یابند و تصمیمات بهتری

^۱ -Big Data

^۲ -IoT: Internet Of Things

^۳ -Machine Learning

^۴ -Text Analysis

^۵ -Social Network Analysis

^۶ -Data Visualization

در زمینه سیاست‌گذاری و برنامه‌ریزی اجتماعی اتخاذ کنند و کیفیت پژوهش‌های خود را بهبود ببخشند. امیدواریم این اثر، به عنوان منبعی ارزشمند برای علاقه‌مندان به این حوزه، به ارتقای دانش و مهارت‌های آن‌ها کمک کند و در نهایت، به بهبود کیفیت زندگی و ارتقای جامعه منجر شود.

با آرزوی موفقیت برای همه پژوهشگران و مجریان در این مسیر هیجان‌انگیز و چالش‌برانگیز در انتها لازم می‌دانیم از مدیریت انتشارات دانشگاه فرماندهی و ستاد آجا که پشتیبان تولید این اثر بوده‌اند، تشکر نموده و از کلیه استادان گرانقدر، همکاران محترم خالصانه قدردانی نماییم. بدیهی است علی‌رغم دقت فراوان برای تقدیم کتاب بی‌عیب و نقص، ترجمه این اثر خالی از اشکال نیست، از استادان محترم دانشگاه‌ها، دانشجویان عزیز و خوانندگان گرامی خواهشمندیم با ارائه نقطه نظرات اصلاحی و تکمیلی مترجمان را مورد لطف خود قرار دهند.

مهدی اوریا^۱ و اکبر اصغرزاده بناب^۲

www.ketab.ir

^۱ -Mahdi.ouria@chmail.ir

^۲ -Akbar.asgharzadeh@casu.ac.ir