

به نام خدا

مبانی علم داده

نویسندگان:

جان . دی کلهر

برنارد تیرنی

مترجمین:

مهندس امیررضا تجلی

دکتر امیرمحمد رمدانی

دکتر امیرعلی رمدانی



شرکت چاپ و نشر بازرگانی

وابسته به مؤسسه مطالعات و پژوهش های بازرگانی

تهران ۱۳۹۰

سرشناسه	کلپهر، جان دی. ۱۹۷۴- م.
عنوان و نام پدیدآور	مبانی علم داده / نویسندگان جان دی. کلپهر، برندن تیرنی؛ مترجمین امیررضا تجلی، امیرمحمد رمدانی، امیرعلی رمدانی؛
مشخصات نشر	ویراستار علمی خلیل حیدری.
مشخصات ظاهری	تهران: شرکت چاپ و نشر بازرگانی، ۱۴۰۰.
شابک	۹۷۸-۹۶۴-۴۶۸-۹۵۴-۳
وضعیت فهرست نویسی	فیبیا
یادداشت	عنوان اصلی: Data science، ۲۰۱۸.
موضوع	کتابنامه.
	داده‌های کلان
	فراگیری ماشینی
	داده‌کاوی
	تحقیق کمی
شناسه افزوده	تجلی، امیررضا، ۱۳۷۳-، مترجم
شناسه افزوده	رمدانی، امیرمحمد، ۱۳۵۹-، مترجم
شناسه افزوده	رمدانی، امیرعلی، ۱۳۶۵-، مترجم
شناسه افزوده	حیدری، خلیل، ۱۳۴۹-، ویراستار
شناسه افزوده	شرکت چاپ و نشر بازرگانی
زده بندی کنکره	QA۷۷/۹
زده بندی دیویی	۰۰۵/۷
شماره کتابشناسی ملی	۸۵۵۳۹۵۰
اطلاعات رکورده کتابشناسی	فیبیا

Big data
Machine learning
Data mining
Quantitative research



شرکت چاپ و نشر بازرگانی

وابسته به مؤسسه مطالعات و پژوهش‌های کتابخانه‌ای

مبانی علم داده

نویسندگان: جان دی. کلپهر، برندن تیرنی

مترجمین: مهندس امیررضا تجلی، دکتر امیرمحمد رمدانی، دکتر امیرعلی رمدانی

ویراستار علمی: دکتر خلیل حیدری

صفحه‌آرا: سولماز دمندان

طراح جلد: جمال سرور سندان

چاپ اول: زمستان ۱۴۰۰

قیمت: ۷۷،۰۰۰ تومان

تیراژ: ۳۰۰ نسخه

شابک: ۹۷۸-۹۶۴-۴۶۸-۹۵۴-۳

لینوگرافی، چاپ و صحافی: شرکت چاپ و نشر بازرگانی

«کلیه حقوق محفوظ و مخصوص شرکت چاپ و نشر بازرگانی است.»

نشانی: تهران، خیابان کارگر شمالی، نبش کوچه همدان، شماره ۱۲۰۴

توزیع و پخش: ۶۶۹۳۹۳۲۹ - ۰۲۱ / فروشگاه مرکزی: ۶۶۴۲۵۱۱۸ - ۰۲۱

واحد نشر مکتوب: ۶۶۴۳۹۲۰۶ - ۰۲۱

واحد نشر الکترونیک: ۶۶۴۳۹۲۰۲ - ۰۲۱ / digital.cppc.ir

WWW.CPPC.ir / WWW.TAKBAB.COM

فروشگاه‌های اینترنتی



فهرست مطالب

پیشگفتار ۹

خلاصه‌ای از رزومه مترجمین ۱۱

پیشگفتاری در مورد مجموعه کتاب‌های دانش ضروری ۱۳

مقدمه ۱۵

سپاس‌گزاری ۱۹

فصل اول: علم داده چیست؟ ۲۱

تاریخچه کوتاهی از علم داده ۲۴

تاریخچه جمع‌آوری داده‌ها ۲۴

تاریخچه تجزیه و تحلیل داده‌ها ۲۸

ظهور و تکامل علم داده ۳۲

علم داده در کجا استفاده می‌شود؟ ۳۷

علم داده در فروش و بازاریابی ۳۸

کاربرد علم داده در دولت‌ها ۳۹

- ۴۰ استفاده از علم داده در ورزش حرفه‌ای
- ۴۱ چرا الان؟
- ۴۴ باورهای غلط و غیرعلمی در مورد علم داده
- ۴۷ فصل دوم: داده‌ها چه هستند و مجموعه داده چیست
- ۵۳ چشم‌انداز داده‌ها
- ۵۷ داده جمع‌آوری می‌شود، خرد چنین نیست!
- ۵۹ فرآیند CRISP-DM
- ۶۷ فصل سوم: اکوسیستم علم داده
- ۷۴ انتقال الگوریتم‌ها به داده‌ها
- ۷۵ پایگاه داده متعارف یا پایگاه داده بدون متعارف
- ۷۹ ساختار کلان داده
- ۸۱ دنیای پایگاه داده ترکیبی
- ۸۳ تهیه و ادغام داده‌ها
- ۸۵ ایجاد جدول تجزیه و تحلیل‌های پایه
- ۸۷ فصل چهارم: یادگیری ماشین
- ۸۸ یادگیری نظارت‌شده در مقابل یادگیری بدون ناظر
- ۹۲ مدل‌های پیش‌بینی در یادگیری
- ۹۳ همبستگی‌ها علیت نیستند اما برخی از آن‌ها مفید هستند
- ۹۹ رگرسیون خطی
- ۱۰۴ شبکه‌های عصبی و یادگیری عمیق
- ۱۱۶ درخت تصمیم
- ۱۲۲ بایاس در مجموعه داده
- ۱۲۳ ارزیابی مدل‌ها: تعمیم نه حفظ

۱۲۶	خلاصه
-----	-------

فصل پنجم: وظایف متعارف علم داده ۱۲۹

۱۳۰	مشتریان ما چه کسانی هستند؟ (خوشه بندی)
۱۳۴	آیا این کلاهبرداری است؟ (تشخیص ناهنجاری)
۱۳۸	آیا همراه آن سیب زمینی سرخ کرده می خواهید؟ (کاوش قاعده همبستگی)
۱۴۴	ریزش یا ماندن سوال این است؟ (طبقه بندی)
۱۴۸	چقدر هزینه دارد؟ (رگرسیون)

فصل ششم: حریم خصوصی و اصول اخلاقی ۱۵۱

۱۵۲	منافع تجاری در برابر حریم خصوصی
۱۵۶	پیامدهای اخلاقی علم داده: نمایه سازی و تبلیغ
۱۶۰	پیامدهای اخلاقی علم داده: ایجاد ساختار سرمایه
۱۶۲	در جستجوی حریم خصوصی از دست رفته
۱۶۳	روش های ریاضی برای حفظ حریم خصوصی
۱۶۶	چارچوب های قانونی استفاده از داده ها و محافظت از حریم خصوصی
۱۶۸	به سوی علم داده اخلاقی

فصل هفتم: روندهای آینده و اصول موفقیت ۱۷۷

۱۷۷	علم داده پزشکی
۱۷۸	شهرهای هوشمند
۱۸۱	اصول پروژه علم داده: علت موفقیت یا شکست پروژه ها
۱۸۲	تمرکز بر موضوع اصلی
۱۸۲	داده ها
۱۸۴	افراد
۱۸۵	مدل ها

پیشگفتار

سپاس خداوند متعال را که این توانمندی را به ما اعطا فرمود تا بتوانیم گام دیگری در راستای ارتقای سطح علمی و دانشی علاقه‌مندان به حوزه‌های نوین برداریم. با توجه به اینکه در دنیای امروز یکی از مولفه‌های اصلی موفقیت دستیابی هرچه سریعتر سازمان‌ها به اطلاعات و تجزیه و تحلیل آن به منظور تصمیم‌گیری‌های اثربخش است، لذا ترجمه این کتاب دو دستور کار ما قرار گرفت و ایده اصلی برای ترجمه شد. علم داده که یکی از دانش‌های میان رشته‌ای و در حال ظهور است، به جمع‌آوری، آماده‌سازی، تحلیل، بصری‌سازی، مدیریت و نگاهداشت اطلاعات در حجم بالا می‌پردازد. این کتاب که با هدف آشنایی هرچه بیشتر علاقه‌مندان به حوزه علم داده ترجمه شده و به موضوعات تخصصی از جمله: ارائه تعاریفی در خصوص علم داده، یادگیری ماشینی، اکوسیستم علم داده، حریم خصوصی و روندهای آتی پرداخته است.

- مخاطبین این کتاب گروه‌های ذیل هستند:
- مدیران و کارشناسان سازمان‌های دولتی
- مدیران و کارشناسان سازمان‌های تولیدی و خدماتی
- مدیران ارشد، مدیران میانی، کارشناسان بنگاه‌های اقتصادی
- صاحبان واحدهای صنعتی، کشاورزی، معدنی، پژوهشگاه‌ها، مراکز تحقیقاتی، مراکز درمانی، دفاتر فنی و مهندسی، دفاتر فروش، شرکت‌های واردکننده و صادرکننده و...