اصول و روش‌های یادگیری

علم داده

تألیف:
Sinan Ozdemir

مترجمان:
دکترحامد تابش
الهام نظری، پرنیان عسگری، مرضیه افکن پور
مهران آغه میری، محبوبه اسلامی، شیوا قادری

انتشارات پندار پارس

سرشناسه : ازدمیر، سینان
Ozdemir, Sinan
عنوان و نام پدیدآور : اصول و روشهای یادگیری علم داده/ تالیف سینان ازدمیر؛ مترجمان حامد تابش...[و دیگران].
مشخصات نشر : تهران : پندار پارس، ۱۳۹۸.
مشخصات ظاهری : ۴۱۰ ص.: مصور، جدول، نمودار.
شابک : 978-600-8201-72-4
وضعیت فهرست نویسی : فیپا
یادداشت : عنوان اصلی : Principles of Data Science: Learn the techniques and math you need to start making sense of your data ,2016.
یادداشت : مترجمان حامد تابش، الهام نظری، پرنیان عسگری، مرضیه افکن‌پور، مهران آغه‌میری، محبوبه اسلامی، شیوا قادری.
موضوع : دادمکاوی
موضوع : Data mining
موضوع : کسب و کار -- داده‌پردازی
موضوع : Business -- Data processing
شناسه افزوده : تابش، حامد، ۱۳۵۶-، مترجم
رده بندی کنگره : QA۹/۶
رده بندی دیویی : ۳۰۰/۰۰۶
شماره کتابشناسی ملی : ۱۳۰

انتشارات پندارپارس

دفتر فروش: انقلاب، ابتدای کارگرجنوبی، کوی رشتچی، پلاک ۴، واحد ۱۶ www.pendarepars.com

تلفن: ۶۶۵۷۲۳۳۵ - تلفکس: ۶۶۹۲۶۵۷۸همراه: ۰۹۱۲۲۴۵۲۳۴۸ info@pendarepars.com

نام کتاب : اصول و روش‌های یادگیری علم داده

ناشر : انتشارات پندار پارس

تالیف : سینان ازدمیر

ترجمه : حامد تابش، الهام نظری، پرنیان عسگری، مرضیه افکن‌پور، مهران آغه‌میری، محبوبه اسلامی، شیوا قادری

چاپ نخست : شهریور ۹۸

شمارگان : ۵۰۰ نسخه

طرح جلد : رامین شکرالهی

چاپ، صحافی : فرشیوه

قیمت : ۸۰۰۰۰ تومان شابک : ۴- ۷۲-۸۲۰۱-۶۰۰-۹۷۸

فهرست

مقدمه

موضوع این کتاب علم داده است که درزمینه تحقیق و کاربرد علم داده توضیحات بسیار ارزشمندی را عنوان می‌کند. علم داده در چند دهه گذشته به‌سرعت در حال رشد و توسعه بوده است. به‌عنوان یک زمینه رو به رشد، توجه زیادی در رسانه‌ها و همچنین در بازار کار به دست آورده است. توجه به علم داده پس از ظهور شرکت‌های فناوری مدل‌سازی از اهمیت ویژه‌ای برخوردار گردید و اخیرا شروع به استخدام تیم‌های داده کاو متخصص کرده‌اند.

این کتاب تلاش خواهد کرد تا شکاف بین تخصص ریاضی / برنامه‌نویسی / داده‌کاوی را متوقف کند. امروزه اکثر مردم حداقل یک (یا شاید دو) تخصص دارند، اما علم داده‌کاوی به بیش از سه تخصص نیاز دارد، و به موضوعاتی از هر سه تخصص وارد شدیم و مشکلات پیچیده را قابل‌حل ساخته‌ایم.

برای ارزیابی نتایج علمی دقیق، داده‌ها را تمیز، کشف و تحلیل خواهیم کرد. یادگیری ماشین و روش‌های یادگیری عمیق که برای حل وظایف پیچیده داده مورداستفاده قرار می‌گیرد را بیان می‌کنیم.

این کتاب چه مواردی را پوشش می‌دهد؟!!

فصل ۱، چگونه به‌عنوان یک متخصص داده در نظر برسیم: مقدمه‌ای بر اصطلاحات اساسی استفاده شده توسط متخصصان داده و نگاهی به انواع مشکلات که در سراسر این کتاب حل خواهد شد.

فصل ۲، انواع داده‌ها: به سطوح مختلف و نحوه دستکاری انواع داده نگاه اساسی می‌کند.

این فصل شروعی برای بررسی ریاضیات موردنیاز برای علم داده است.

فصل ۳، پنج مرحله علوم داده: پنج مرحله اساسی در انجام علوم داده ازجمله دستکاری و تمیز کردن داده‌ها را باز می‌کند، و نمونه‌هایی از هر مرحله را بیان می‌کند.

فصل ۴، ریاضیات پایه: به ما کمک می‌کند تا اصول پایه ریاضی را کشف کنیم که عملکردهای متخصصان داده را با دیدن و حل نمونه‌هایی در حساب، جبر خطی و غیره هدایت می‌کنند.

فصل ۵، غیرممکن یا غیرقابل پیش‌بینی-مقدمه‌ای ساده درباره احتمال: یک نگاه مبتدی به نظریه احتمال و نحوه استفاده از آن برای به دست آوردن درک جهان تصادفی ما است.

فصل ۶، احتمال پیشرفته: از اصول فصل قبلی استفاده می‌کند و قضیه‌ها و نظریه‌هایی نظیر قضیه بایاس را به کار می‌گیرد و امید به کشف معنای پنهان در دنیای داده‌ها دارد.

فصل ۷، آمار پایه: در پی بررسی انواع مشکلی است که استنتاج آماری تلاش می‌کند با استفاده از اصول آزمایشی، نرمال‌سازی و نمونه‌گیری تصادفی توضیح دهد.

فصل ۸، آمار پیشرفته: از آزمون فرضیه‌ها و فاصله اطمینان استفاده می‌کند تا بینشی را از آزمایش‌های ما به دست آورد. داشتن توانایی انتخاب تست مناسب و نحوه تفسیر مقادیر p و سایر نتایج نیز بسیار مهم است.

فصل ۹، ارتباطات داده‌ها: توضیح می‌دهد که چگونه ارتباطات و علیت بر تفسیر ما از داده تأثیر می‌گذارد. همچنین با استفاده از تجسم و بصری‌سازی به دنبال به اشتراک گذاشتن نتایج خود با جهان پیرامونمان است.

فصل ۱۰، مربوط به ملزومات یادگیری ماشین است: بر تعریف یادگیری ماشین متمرکز است و به مثال‌های واقعی در مورد نحوه و زمان استفاده از یادگیری ماشینی می‌پردازد.

فصل ۱۱، پیش‌بینی‌ها و درخت تصمیم: به مدل‌های پیچیده‌تر یادگیری ماشین، مانند درخت تصمیم‌گیری و پیش‌بینی‌های خطی نیز می‌پردازیم، به‌منظور حل وظایف مربوط به تحلیل پیچیده‌تر داده‌ها نگاه می‌کند.

فصل ۱۲، فراتر از ملزومات: برخی از فروع‌های پرموز را که علوم داده‌ها را هدایت می‌کنند معرفی می‌کند، ازجمله بایاس و واریانس. در این فصل شبکه‌های عصبی به‌عنوان روش یادگیری عمیق مدرن معرفی می‌شوند.

فصل ۱۳، مطالعات موردی: از مجموعه‌ای از مطالعات موردی استفاده می‌کند تا ایده‌های علوم داده را تقویت کند. همچنین نمونه‌ها و مثال‌های مختلف، ازجمله پیش‌بینی قیمت سهام و تشخیص دستخط، بررسی خواهد شد.

آنچه شما در این کتاب نیاز دارید:

این کتاب از پایتون برای تمام نمونه‌های کد و حل مثال‌ها استفاده می‌کند. یک رایانه کامپیوتر با سیستم‌عامل لینوکس / مک / ویندوز با دسترسی به ترمینال یونیکس با پایتون ۲.۷ نصب‌شده موردنیاز است.

نصب و راه‌اندازی توزیع Anaconda نیز توصیه می‌شود، بسیاری از بسته‌های مورداستفاده نمونه‌ها در کتاب از آن استفاده کرده است.