

نویسندگان: توماس نیلد

ریاضیات پایه برای علوم داده

کنترل داده‌ها را با مبانی جبر خطی،
آمار و احتمالات در دست بگیرید

مترجمان:

فاضله توسلیان

امین نظری

ریاضیات پایه برای علوم داده

مترجمان: فاضله توسلیان، امین نظری

ویراستار علمی: دکتر رامین مولاناپور

ناشر: انتشارات آتی نگر

مدیر هنری و طراح جلد: همتا بیداریان

چاپ اول، ۱۴۰۳

شمارگان: ۱۰۰ نسخه

قیمت: ۳,۵۰۰,۰۰۰ ریال

شابک: ۹۷۸-۶۲۲-۸۲۴۵-۱۲-۶

ISBN: 978-622-8245-12-6

حق چاپ برای انتشارات آتی نگر محفوظ است.

نشانی دفتر فروش: خیابان جمالزاده جنوبی، روبه روی کوچه رشتچی، پلاک ۱۴۴، واحد ۱

نمابر: ۶۶۵۶۵۳۳۷

تلفن: ۶۶۵۶۵۳۳۶-۸

www.ati-negar.com * info@ati-negar.com



سرشناسه: نیلد، توماس، Nield, Thomas

ریاضیات ضروری برای علوم داده/ نویسندگان توماس نیلد؛ مترجمان فاضله توسلیان، امین نظری / ویراستار علمی: رامین مولاناپور
تهران: آتی نگر ۱۴۰۳.

۳۶۸ ص: مصور، جدول، نمودار.

ISBN: 978-622-8245-12-6

فیپا.

یادداشت: عنوان اصلی کتاب: Essential math for data science : take control of your data with fundamental linear algebra, probability, and statistics

موضوع: داده کاوی - ریاضیات - Mathematics - Data mining - علوم کامپیوتر - ریاضیات - Computer science - Mathematics

موضوع: فراگیری ماشینی - ریاضیات - Mathematics -- Machine learning

موضوع: آمار ریاضی - احتمالات - Probabilities - Mathematical statistics

شناسه افزوده: توسلیان، فاضله، ۱۳۷۱- مترجم

شناسه افزوده: نظری، امین، ۱۳۶۶- مترجم

شناسه افزوده: مولاناپور، رامین، ۱۳۵۲- ویراستار عمکی

شناسه افزوده: بیداریان، همتا، ۱۳۶۱- مدیر هنری

QAY۶/۹

۰۰۶/۳۱۰

۹۴۹۴۳۳۸

رده بندی کنگره

رده بندی دیویی

شماره کتابشناسی ملی

فهرست مطالب

۷.....	مقدمه مترجمان
۹.....	پیشگفتار
۱۳.....	فصل ۱: بررسی ریاضی پایه و حساب دیفرانسیل و انتگرال
۱۴.....	نظریه اعداد
۱۶.....	ترتیب عملیات‌ها
۱۷.....	متغیرها
۱۸.....	توابع
۲۴.....	مجموع
۲۶.....	توان
۲۹.....	لگاریتم
۳۱.....	عدد اویلر و لگاریتم‌های طبیعی
۳۵.....	حد
۳۸.....	مشتق
۴۸.....	انتگرال
۵۴.....	نتیجه‌گیری
۵۴.....	تمرینات
۵۵.....	فصل ۲: احتمال
۵۶.....	درک احتمال
۵۷.....	احتمال در مقابل آمار
۵۸.....	ریاضی احتمالات
۶۵.....	توزیع دو جمله‌ای
۷۵.....	نتیجه‌گیری
۷۶.....	تمرینات
۷۷.....	فصل ۳: آمار توصیفی و استنباطی
۷۷.....	داده چیست؟

۸۰	آمار توصیفی در مقابل آمار استنباطی
۸۰	جمعیت‌ها، نمونه‌ها و بایاس
۸۵	آمار توصیفی
۱۰۶	آمار استنباطی
۱۲۲	توزیع T: کار کردن با نمونه‌های کوچک
۱۲۴	سنجش داده‌های بزرگ و استدلال نادرست
۱۲۶	نتیجه‌گیری
۱۲۶	تمرینات

فصل ۴: جبر خطی ۱۲۹

۱۳۰	بردار چیست؟
۱۴۲	تبدیل‌های خطی
۱۵۰	ضرب ماتریس
۱۵۳	دترمینان
۱۵۷	انواع ماتریس‌های ویژه
۱۵۹	دستگاه‌های معادلاتی و ماتریس‌های وارون
۱۶۴	بردارهای ویژه و مقادیر ویژه
۱۶۷	نتیجه‌گیری
۱۶۷	تمرینات

فصل ۵: رگرسیون خطی ۱۶۹

۱۷۱	رگرسیون خطی پایه
۱۷۹	یافتن بهترین خط مناسب
۱۸۰	معادله شکل بسته
۱۸۲	تکنیک‌های ماتریس وارون
۱۸۴	گرادیان کاهشی
۱۹۲	بیش‌برازش و واریانس
۱۹۴	گرادیان کاهشی تصادفی
۱۹۶	ضریب همبستگی
۲۰۰	اهمیت آماری
۲۰۵	ضریب دترمینان

۲۰۷	خطای استاندارد تخمین
۲۰۸	فواصل پیش‌بینی
۲۱۲	تقسیم‌بندی داده‌ها به داده‌های آموزشی و آزمایشی
۲۱۹	رگرسیون خطی چندگانه
۲۲۰	نتیجه‌گیری
۲۲۱	تمرینات

فصل ۶: رگرسیون لجستیک و طبقه‌بندی ۲۲۳

۲۲۴	درک رگرسیون لجستیک
۲۲۶	انجام رگرسیون لجستیک
۲۳۱	استفاده از حداکثر درست‌نمایی و گرادیان کاهشی
۲۳۶	رگرسیون لجستیک چندمتغیره
۲۴۰	درک Log-Odds
۲۴۴	مربع R
۲۴۹	P-Values
۲۵۱	تقسیم داده به داده‌های آموزشی / آزمایشی
۲۵۲	ماتریس‌های درهم‌ریختگی
۲۵۶	قضیه و طبقه‌بندی بیز
۲۵۷	مشخصه عملکرد گیرنده یا منحنی ROC
۲۵۹	داده‌های نامتوازن
۲۶۰	نتیجه‌گیری
۲۶۰	تمرینات

فصل ۷: شبکه‌های عصبی ۲۶۳

۲۶۴	چه زمان از شبکه‌های عصبی و یادگیری ژرف استفاده می‌کنیم؟
۲۶۵	یک شبکه عصبی ساده
۲۸۰	پس‌انتشار
۲۸۹	استفاده از scikit-learn
۲۹۱	محدودیت‌های شبکه‌های عصبی و یادگیری ژرف
۲۹۵	نتیجه‌گیری
۲۹۶	تمرینات

فصل ۸: مشاوره شغلی و مسیر روبه جلو ۲۹۷

۲۹۸	بازتعریف علوم داده
۳۰۰	تاریخچه مختصر علوم داده
۳۰۳	پیدا کردن لبه
۳۰۴	مهارت SQL
۳۰۷	مهارت برنامه نویسی
۳۱۱	مصورسازی داده ها
۳۱۳	شناخت صنعت
۳۱۴	یادگیری تمرینش
۳۱۵	عمل ورز در مقابل راهنما
۳۱۸	در شغل علوم داده چه مواردی را باید رعایت کرد
۳۲۸	آیا شغل رویایی شما وجود ندارد؟
۳۲۹	الان به کجا برویم؟
۳۳۰	نتیجه گیری

پیوست الف: عناوین تکمیلی ۳۳۱

۳۳۱	استفاده از رندر LaTeX با SymPy
۳۳۲	توزیع دوجمله ای از ابتدا
۳۳۴	توزیع بتا از ابتدا
۳۳۶	استخراج قضیه بیز
۳۳۷	CDF و CDF وارون از ابتدا
۳۳۸	استفاده از e برای پیش بینی احتمال رویداد در طول زمان
۳۴۰	تپه نوردی و رگرسیون خطی
۳۴۳	تپه نوردی و رگرسیون لجستیک
۳۴۴	مقدمه ای بر برنامه ریزی خطی
۳۵۰	طبقه بند MNIST با استفاده از scikit-learn

پیوست ب: پاسخ تمرینات ۳۵۲

مقدمه مترجمان

کتاب پیش‌رو یکی از بهترین منابع موجود برای کسانی است که می‌خواهند به دنیای وسیع و در حال گسترش هوش مصنوعی، یادگیری ماشین و یا علوم داده قدم بگذارند. در این کتاب ریاضیات پشت برخی از اساسی‌ترین الگوریتم‌های یادگیری ماشین را فرا خواهید گرفت و می‌توانید نحوه پیاده‌سازی این الگوریتم‌ها را در زبان Python بیاموزید.

کتاب اصلی به زبانی ساده و روایت‌گونه به تشریح مطالب مختلف می‌پردازد. مطمئناً اشاره به یک سریال، یک فیلم یا یک مسابقه خاص که در آمریکا برگزار شده است برای خواننده فارسی زبان جذاب نیست. از این‌رو، مترجمان تلاش کردند تا از مثال‌های ملموس‌تری برای توضیحات بیشتر استفاده کنند. به‌علاوه، بخش‌هایی جدیدی به متن اصلی اضافه شده است تا خوانندگان را با مطالب به‌روزتر و مسائل پیچیده‌تر آشنا کند. با این حال هیچ اثری نمی‌تواند خالی از عیب باشد، ما مشتاقانه منتظر دریافت نظرات استادان و دانشجویان گرامی هستیم.

در پایان بر خود واجب می‌دانیم تا از استادان گرامی آقای دکتر محرم منصوری‌زاده و آقای دکتر مهدی عباسی تشکر کنیم.