

۲۱۸۲۹۲

مهندسی ویژگی در یادگیری ماشین

مؤلف

پابلو دوبو

مترجم

ایوب ترکیان

www.ketab.ir

نیاز دانش

سرشناسه	: دوبو، پابلو، ۱۹۷۶-م	Duboue, Pablo, 1976
عنوان و نام پدیدآور	: مهندسی ویژگی در یادگیری ماشین / مولف پابلو دوبو؛ مترجم ایوب ترکیان.	
مشخصات نشر	: تهران: نیاز دانش، ۱۳۹۹.	
مشخصات ظاهری	: ص: ۲۴۸، مصور، جدول، نمودار	
شابک	: 978-600-8906-96-4	
وضعیت فهرست نویسی	: فیبا	
یادداشت	: عنوان اصلی: .The art of feature engineering : essentials for machine learning, 2020	
موضوع	: فراگیری ماشین	Machine learning
موضوع	: پایتون (زبان برنامه نویسی کامپیوتر)	Python (Computer program language)
شناسه افزوده	: ترکیان، ایوب، ۱۳۳۷ - مترجم	
رده بندی کنگره	: Q۳۲۵/۵	
رده بندی دیویی	: ۰۰۶/۳۱	
شماره کتابشناسی ملی	: ۷۵۳۰۴۴۹	
وضعیت رکورد	: فیبا	



www.ketab.ir

نام کتاب	: مهندسی ویژگی در یادگیری ماشین
مؤلف	: پابلو دوبو
مترجم	: ایوب ترکیان
مدیر اجرایی - ناظر بر چاپ	: حمیدرضا محمد شیرازی - محمد شمس
ناشر	: نیاز دانش
صفحه آرا	: واحد تولید انتشارات نیاز دانش
نوبت چاپ	: دوم - ۱۴۰۲
شمارگان	: ۳۰۰ نسخه
قیمت	: ۱۵۰۰۰۰۰ ریال

ISBN:978-600-8906-96-4

شابک: ۹۷۸-۶۰۰-۸۹۰۶-۹۶-۴

هرگونه چاپ و تکثیر (اعم از زیراکس، بازنویسی، ضبط کامپیوتری و تهیهی CD) از محتویات این اثر بدون اجازه کتبی ناشر ممنوع است. متخلفان به موجب بند ۵ از ماده ۲ قانون حمایت از مؤلفان، مصنفان و هنرمندان تحت پیگرد قانونی قرار می گیرند.

کلیه حقوق این اثر برای ناشر محفوظ است.

تهران، میدان انقلاب، خیابان ۱۲ فروردین، تقاطع وحید نظری، پلاک ۲۵۵، طبقه ۱، واحد ۲

۰۹۱۲۷۰۷۳۹۳۵ - ۶۶۴۷۸۱۰۸ - ۶۶۴۷۸۱۰۶ - ۰۲۱

www.Niaze-Danesh.com

مشاوره جهت نشر: ۰۹۱۲ - ۲۱۰۶۷۰۹

فهرست

۱۱	فصل اول: مقدمه
۱۴	۱-۱ مهندسی ویژگی
۱۷	۲-۱ ارزیابی
۱۸	۱-۲-۱ سنجه‌ها
۲۰	۲-۲-۱ اعتبارسنجی متقاطع
۲۱	۱-۲-۲-۱ تخمین خارج بخشی
۲۲	۳-۲-۱ فرابرازش
۲۳	۴-۲-۱ نفرین بُعدیت
۲۳	۳-۱ چرخه‌ها
۲۳	۱-۳-۱ چرخه یادگیری ماشین
۲۶	۲-۳-۱ چرخه مهندسی ویژگی
۲۸	۴-۱ تحلیل
۲۸	۱-۴-۱ تحلیل داده اکتشافی
۳۰	۲-۴-۱ تحلیل خطا
۳۱	۵-۱ فرایندهای دیگر
۳۲	۱-۵-۱ مدل‌سازی زمینه‌ای
۳۴	۲-۵-۱ ساخت ویژگی
۳۴	۱-۲-۵-۱ انواع ویژگی
۳۶	۲-۲-۵-۱ شابلون‌های ویژگی
۳۷	۶-۱ بحث
۳۹	فصل دوم: ترکیب ویژگی‌ها
۴۰	۱-۲ نرمال‌سازی ویژگی‌ها
۴۳	۱-۱-۲ استانداردسازی و هم‌بستگی‌زدایی
۴۵	۲-۱-۲ هموارسازی
۴۶	۱-۲-۱-۲ هموارسازی احتمال
۴۷	۳-۱-۲ وزن‌دهی ویژگی

۴۷	۱-۳-۱-۲- وزن دهی فرکانس سند معکوس
۴۸	۲-۳-۱-۲- کالیراسیون دوربین
۴۸	۲-۲- گسسته سازی و سبب بندی
۴۹	۱-۲-۲- گسسته سازی بدون نظارت
۴۹	۱-۱-۲-۲- سبب بندی
۵۱	۲-۱-۲-۲- خوشه بندی
۵۲	۲-۲-۲- گسسته سازی با نظارت
۵۳	۱-۲-۲-۲- ChiMerge
۵۳	۲-۲-۲-۲- MDLP
۵۴	۳-۲-۲-۲- CAIM
۵۵	۳-۲- ویژگی های توصیفی
۵۵	۱-۳-۲- هیستوگرام ها
۵۷	۲-۳-۲- دیگر ویژگی های توصیفی
۵۸	۴-۲- برخورد با داده های پرت
۵۹	۱-۴-۲- آشکار سازی داده های پرت
۶۰	۵-۲- تکنیک های پیشرفته
۶۳	فصل سوم: بسط ویژگی ها
۶۴	۱-۳- ویژگی های قابل محاسبه
۷۱	۲-۳- جانهی
۷۴	۳-۳- تجزیه توابع پیچیده
۷۷	۴-۳- بسط ویژگی القا شده توسط کرنل
۸۳	فصل چهارم: کاهش ویژگی ها
۸۵	۱-۱-۴- سنجه ها
۸۵	۱-۱-۴-۱- سنجه های استفاده ویژگی
۹۰	۲-۱-۴-۱- سنجه های ویژگی های چندگانه
۹۲	۳-۱-۴-۱- طبقه گرهای ویژگی تکی
۹۲	۴-۱-۴-۱- زرورق
۹۳	۲-۱-۴- سرهم کردن مجموعه ویژگی: جستجو و فیلتر
۹۴	۱-۲-۴-۱- مقتصدانه
۹۵	۲-۲-۴-۱- معیار توقف

۹۶	۳-۱-۴- تکنیک‌های پیشرفته
۹۷	۱-۳-۱-۴- پایداری
۹۸	۲-۳-۱-۴- لیست سیاه ویژگی‌ها
۹۸	۲-۴- تنظیم و انتخاب ویژگی نهفته
۱۰۰	۱-۲-۴- تنظیم L_2 : رگرسیون صخره
۱۰۰	۲-۲-۴- تنظیم L_1 : LASSO
۱۰۱	۱-۲-۲-۴- ElasticNet
۱۰۲	۳-۲-۴- الگوریتم‌های دیگر
۱۰۳	۳-۴- کاهش بُعدیت
۱۰۴	۱-۳-۴- هش کردن ویژگی‌ها
۱۰۵	۲-۳-۴- افکنش راندومی
۱۰۶	۳-۳-۴- SVD
۱۰۷	۴-۳-۴- تخصیص دیریکله نهفته
۱۰۸	۵-۳-۴- خوشه‌بندی
۱۰۸	۶-۳-۴- دیگر تکنیک‌های کاهش بُعدیت
۱۱۰	۷-۳-۴- نهفته‌سازی
۱۱۲	۱-۷-۳-۴- نهفته‌سازی فراگیر
۱۱۴	۲-۷-۳-۴- نهفته‌سازی محلی: Word2Vec
۱۱۴	۳-۷-۳-۴- نهفته‌سازی‌های دیگر: GloVe و ULMFiT
۱۱۵	فصل پنجم: موضوعات پیشرفته
۱۱۶	۱-۵- بردارهای با طول متغیر
۱۱۶	۱-۱-۵- مجموعه‌ها
۱۱۷	۲-۱-۵- لیست‌ها
۱۱۹	۳-۱-۵- درخت
۱۲۲	۴-۱-۵- گراف
۱۲۲	۵-۱-۵- سری زمانی
۱۲۳	۱-۵-۱-۵- خودهم‌بستگی
۱۲۴	۲-۵-۱-۵- روند، چرخه، و مؤلفه‌های فصلی
۱۲۵	۳-۵-۱-۵- مانایی
۱۲۶	۴-۵-۱-۵- مدل‌های سری زمانی: ARIMA

۲۰۴	۱-۵-۸- واژه‌های توقف
۲۰۴	۲-۵-۸- توکن‌سازی: ریشه‌یابی
۲۰۵	۳-۵-۸- سومین ویژگی‌سازی
۲۰۶	۱-۳-۵-۸- تحلیل خطا
۲۰۶	۶-۸- ویژگی‌ها در بافت
۲۰۶	۱-۶-۸- بایگرام‌ها
۲۰۷	۲-۶-۸- چهارمین ویژگی‌سازی
۲۰۷	۷-۸- بایگرام‌های گذر و هش کردن ویژگی
۲۰۸	۱-۷-۸- بایگرام گذر
۲۰۸	۲-۷-۸- پنجمین ویژگی‌سازی
۲۰۹	۸-۸- کاهش بُعدیت و نهفته‌سازی
۲۰۹	۱-۸-۸- نهفته‌سازی‌ها
۲۱۰	۲-۸-۸- وزن‌دهی ویژگی TF-IDF
۲۱۱	۳-۸-۸- ششمین ویژگی‌سازی
۲۱۱	۹-۸- اشارات پایانی
۲۱۳	۱-۹-۸- بسط محتوا
۲۱۴	۲-۹-۸- ساختار در متن
۲۱۵	فصل نهم: داده‌های تصویری
۲۱۶	مرور فصل
۲۱۷	۱-۹- شهرهای ویکی: تصاویر ماهواره‌ای
۲۱۸	۲-۹- تحلیل داده اکتشافی
۲۱۹	۳-۹- پیکسل به عنوان ویژگی
۲۲۰	۱-۳-۹- اولین ویژگی‌سازی
۲۲۰	۱-۱-۳-۹- تحلیل خطا
۲۲۲	۲-۳-۹- ویژگی‌های قابل محاسبه: تیرگی گوسی
۲۲۳	۳-۳-۹- سفیدسازی
۲۲۵	۴-۳-۹- تحلیل خطا روی تغییرپذیری‌ها
۲۲۶	۴-۹- بسط مجموعه داده اتوماتیک
۲۲۶	۱-۴-۹- تبدیل‌های آفین
۲۲۶	۲-۴-۹- دومین ویژگی‌سازی

۲۲۷	۱-۲-۴-۹- تحلیل خطا
۲۲۷	۵-۹- ویژگی‌های توصیفی: هیستوگرام‌ها
۲۲۷	۱-۵-۹- سومین ویژگی‌سازی
۲۲۸	۱-۱-۵-۹- تحلیل خطا
۲۲۸	۶-۹- آشکارسازهای ویژگی محلی: گوشه‌ها
۲۲۹	۱-۶-۹- آشکارساز گوشه هریس
۲۲۹	۲-۶-۹- چهارمین ویژگی‌سازی
۲۳۰	۱-۲-۶-۹- تحلیل خطا
۲۳۱	۷-۹- کاهش بُعدیت: HOG-ها
۲۳۲	۱-۷-۹- پنجمین ویژگی‌سازی
۲۳۲	۸-۹- اشارات پایانی
۲۳۳	۱-۸-۹- موضوعات دیگر در بینایی کامپیوتر
۲۳۵	فصل دهم: زمینه‌های دیگر
۲۳۶	۱-۱۰- ویدئو
۲۳۷	۱-۱-۱۰- داده‌ها: نمایش صفحه
۲۳۷	۲-۱-۱۰- آشکارسازی قاب کلیدی
۲۳۹	۳-۱-۱۰- ردیابی لکه‌ها: شیفت میانگین
۲۳۹	۱-۳-۱-۱۰- پس‌افکنش هیستوگرام
۲۳۹	۲-۳-۱-۱۰- شیفت میانگین
۲۴۰	۴-۱-۱۰- یادگیری بیشتر
۲۴۱	۲-۱۰- ویژگی‌های جغرافیایی
۲۴۱	۱-۲-۱۰- داده‌ها: مهاجرات حیوان
۲۴۲	۱-۱-۲-۱۰- فاصله شعاعی به علامات راهنما
۲۴۳	۲-۱-۲-۱۰- یادگیری بیشتر
۲۴۳	۳-۱۰- ترجیحات
۲۴۴	۱-۳-۱۰- داده‌ها: کرنل تعهدهای لینوکس
۲۴۵	۲-۳-۱۰- جانمایی ترجیحات
۲۴۶	۳-۳-۱۰- یادگیری بیشتر