

۵۹۷۷۷۲

به نام خدا

یادگیری ژرف

نویسنده: خانم. ایمان گودفلو، یوشوا بنجیو، آرون کورویل

www.ketab.ir

مترجمان:

مهندس سجاد ملکی

مهندس مائده حاجی آقامحسینی

یادگیری ژرف

مترجمان: مهندس سجاد ملکی، مهندس مائده حاجی‌آقامحسینی

ویراستار علمی: مهندس رامین مولاناپور

ناشر: انتشارات آتی‌نگر

ناشر همکار: انتشارات وینا

چاپ اول، ۱۳۹۷

شمارگان: ۵۰۰ نسخه

قیمت: ۵۵۰,۰۰۰ ریال

شابک: ۹۷۸-۶۲۲-۶۱۰۲-۰۰۰-۱

ISBN: 978-622-6102-00-1

حق چاپ برای انتشارات آتی‌نگر محفوظ است.

نشانی دفتر فروش: خیابان ج ۱۱، جنوب، روی کوچه رشتچی، پلاک ۱۴۴ - واحد ۱

تلفن: ۰۲۱-۵۳۳۷۶۶۵۶۵۳۳۶-۸



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

سرشناسه: گودفلو، یان، Goodfellow Ian -

یادگیری ژرف / نویسندگان یان گودفلو، یوشوا بنجیو، آرون کورویل؛ مترجمان سجاد ملکی، مائده حاجی‌آقامحسینی؛

ویراستار علمی: رامین مولاناپور

تهران: آتی‌نگر، وینا ۱۳۹۷

۵۰۴ ص: مصور، جدول، نمودار.

ISBN: 978-622-6102-00-1

فیپا.

موضوع: عنوان اصلی کتاب: [2016], Deep Learning

موضوع: فراگیری ماشینی — Machine learning

شناسه افزوده: بنجیو، یوشوا Bengio, Yoshua -

شناسه افزوده: کورویل، آرون Courville, Aaron -

شناسه افزوده: ملکی، سجاد، ۱۳۶۸ -

شناسه افزوده: حاجی محسینی، مائده، ۱۳۶۲ -

شناسه افزوده: مولاناپور، رامین، ۱۳۵۲ -، ویراستار

رده‌بندی کنگره

رده‌بندی دیویی

شماره کتابشناسی ملی

۱۳۹۷ ق ۱۳۲۵/۵/۱۹

۰۰۶/۳۱

۵۱۴۶۱۵۳

فهرست مطالب

وبسایت

نمادگذاری

فصل ۱: مقدمه

۱-۱ چه کسانی باید این کتاب را مطالعه کنند؟..... ۲۴

۱-۲ روندهای تاریخی ریاضی ژرف..... ۲۸

بخش اول: مبانی یادگیری ماشین و رد عی کاربردی

فصل ۲: جبر خطی

۲-۱ اسکالرها، بردارها، ماتریس‌ها و تنسورها..... ۴۵

۲-۲ ضرب ماتریس‌ها و بردارها..... ۴۸

۲-۳ ماتریس همانی و ماتریس وارون..... ۵۰

۲-۴ وابستگی خطی و محدوده..... ۵۱

۲-۵ نرم‌ها..... ۵۳

۲-۶ انواع ویژه‌ی ماتریس‌ها و بردارها..... ۵۵

۲-۷ تجزیه‌ی ویژه..... ۵۶

۲-۸ تجزیه‌ی مقدار تکین..... ۵۹

۲-۹ ماتریس شبه‌وارون مور-پنروز..... ۶۰

۲-۱۰ عملگر اثر (trace)..... ۶۱

۲-۱۱ دترمینان..... ۶۲

۲-۱۲ مثال: تحلیل اجزای اصلی..... ۶۲

فصل ۳: نظریه احتمال و اطلاعات

۳-۱ چرا احتمال؟..... ۶۹

۷۲	۳-۲ متغیرهای تصادفی
۷۲	۳-۳ توزیع‌های احتمال
۷۴	۳-۴ احتمال حاشیه‌ای
۷۵	۳-۵ احتمال شرطی
۷۵	۳-۶ قاعده‌ی زنجیره‌ی احتمالات شرطی
۷۶	۳-۷ عدم وابستگی و استقلال شرطی
۷۶	۳-۸ امید ریاضی، واریانس و کواریانس
۷۸	۳-۹ توزیع‌های احتمالاتی متداول
۸۳	۳-۱۰ واریانس و سودمند توابع معمول
۸۶	۳-۱۱ فانی بیز
۸۶	۳-۱۲ جزئیات فنر تغییرهای پیوسته
۸۸	۳-۱۳ نظریه‌ی اطلاعات
۹۲	۳-۱۴ مدل‌های احتمالاتی ساده‌ی یافته

فصل ۳ محاسبات عددی ۹۷

۹۷	۴-۱ سرریز و پاریز
۹۹	۴-۲ شرطی‌سازی ضعیف
۹۹	۴-۳ بهینه‌سازی مبتنی بر گرادینت
۱۰۹	۴-۴ بهینه‌سازی مشروط
۱۱۲	۴-۵ مثال: حداقل مربعات خطی

فصل ۵ اصول یادگیری ماشین ۱۱۵

۱۱۶	۵-۱ الگوریتم‌های یادگیری
۱۲۷	۵-۲ ظرفیت، بیش‌برازش و کم‌برازش
۱۳۷	۵-۳ فرایارامترها و مجموعه‌های اعتبارسنجی
۱۳۹	۵-۴ برآوردگرها، ارببی و واریانس
۱۴۸	۵-۵ برآورد درست‌نمایی پیشینه
۱۵۲	۵-۶ آمار بیزی
۱۵۷	۵-۷ الگوریتم‌های یادگیری با ناظر
۱۶۳	۵-۸ الگوریتم‌های یادگیری بدون ناظر
۱۶۸	۵-۹ گرادینت کاهش تصادفی

۱۷۰..... ۵-۱۰ ساخت یک الگوریتم یادگیری ماشین

۱۷۱..... ۵-۱۱ چالش‌های انگیزه‌بخش یادگیری ژرف

بخش دوم: شبکه‌های عمیق: شبکه‌های نوین

بخش ۶: شبکه‌های پیش‌خور ژرف

۱۸۶..... ۶-۱ مثال: یادگیری بای انحصاری XOR

۱۹۱..... ۶-۲ یادگیری مبتنی بر گرادینت

۲۰۶..... ۶-۳ مدل‌های مخفی

۲۱۲..... ۶-۴ طراحی معماری

۲۱۸..... ۶-۵ پس‌انتشار سایر الگوریتم‌های مشتق‌گیری

۲۳۸..... ۶-۶ یادداشت تاریخچه‌ای

بخش ۷: منظم‌سازی برای یادگیری ژرف

۲۴۵..... ۷-۱ جریمه‌های نرم پارامتر

۲۵۲..... ۷-۲ جریمه‌های نرم به‌عنوان بهینه‌سازی مشروط

۲۵۴..... ۷-۳ منظم‌سازی و مسائل مشروط

۲۵۵..... ۷-۴ تقویت و تکمیل مجموعه داده

۲۵۷..... ۷-۵ پابرجایی نوفه

۲۵۹..... ۷-۶ یادگیری نیمه‌ناظر

۲۶۰..... ۷-۷ یادگیری چندوظیفه‌ای

۲۶۱..... ۷-۸ توقف زودهنگام

۲۶۷..... ۷-۹ اتصال پارامتر و به‌اشتراک‌گذاری پارامتر

۲۶۸..... ۷-۱۰ بازنمایی پراکنده (اسپارس)

۲۷۰..... ۷-۱۱ بگینگ و سایر روش‌های گردآوری و مجموعه‌سازی

۲۷۲..... ۷-۱۲ حذف تصادفی

۲۸۲..... ۷-۱۳ آموزش هم‌آورد

۲۸۴..... ۷-۱۴ فاصله تانژانت، انتشار تانژانت و دسته‌بندی‌کننده تانژانت خمینه

بخش ۸: بهینه‌سازی برای آموزش مدل‌های ژرف

۲۹۰..... ۸-۱ تفاوت یادگیری با بهینه‌سازی محض در چیست؟

۲۹۷	۸-۲ چالش‌های موجود در بهینه‌سازی شبکه‌ی عصبی
۳۰۸	۸-۳ الگوریتم‌های پایه
۳۱۶	۸-۴ استراتژی‌های مقداردهی اولیه‌ی پارامترها
۳۲۱	۸-۵ الگوریتم‌هایی با نرخ یادگیری انطباقی
۳۲۶	۸-۶ متدهای مرتبه‌دوم تقریبی
۳۳۴	۸-۷ استراتژی‌های بهینه‌سازی و الگوریتم‌های متا

فصل ۹: شبکه‌های کانولوشنی ۳۴۷

۳۴۸	۹-۱ مدل ریاضی همگشت
۳۵۱	۹-۲ انگ
۳۵۶	۹-۳ دغام
۳۶۲	۹-۴ همگشت و ادغام: عنوان یک دانش پیشین بی‌نهایت قدرتمند
۳۶۳	۹-۵ انواع توابع همگشتی
۳۷۳	۹-۶ خروجی‌های شبکه‌ی پایانه
۳۷۵	۹-۷ انواع داده
۳۷۷	۹-۸ الگوریتم‌های کارآمد همگشتی
۳۷۷	۹-۹ ویژگی‌های تصادفی یا بدون ناظر
۳۷۹	۹-۱۰ مبنای عصب‌شناسی شبکه‌های همگشتی
۳۸۶	۹-۱۱ شبکه‌های همگشتی و تاریخچه‌ی یادگیری ژرف

فصل ۱۰: مدل‌سازی دنباله: شبکه‌های بازگشتی و تراجعی ۳۸۹

۳۹۰	۱۰-۱ تاگشایی گراف‌های محاسباتی
۳۹۴	۱۰-۲ شبکه‌های عصبی تراجعی
۴۰۹	۱۰-۳ RNN‌های دوسویه
۴۱۱	۱۰-۴ معماری‌های دنباله-به-دنباله‌ی رمزگذار-رمزگشا
۴۱۳	۱۰-۵ شبکه‌های تراجعی ژرف
۴۱۴	۱۰-۶ شبکه‌های عصبی تراجعی
۴۱۶	۱۰-۷ چالش وابستگی‌های بلندمدت
۴۱۸	۱۰-۸ شبکه‌های حالت اکو
۴۲۱	۱۰-۹ واحدهای نشی و استراتژی‌های دیگر برای مقیاس زمانی چندگانه
۴۲۳	۱۰-۱۰ حافظه‌ی کوتاه‌مدت ماندگار و سایر RNN‌های گیت‌دار

۱۰-۱۱ بهینه‌سازی برای وابستگی‌های بلندمدت..... ۴۲۷

۱۰-۱۲ حافظه‌ی آشکار..... ۴۳۱

۱۱-۱ معیارهای عملکردی..... ۴۳۶

۱۱-۲ مدل‌های پایه پیش‌فرض..... ۴۳۹

۱۱-۳ تعیین اینکه داده‌های بیشتر گردآوری شوند یا نه..... ۴۴۰

۱۱-۴ انتخاب فرآیندها..... ۴۴۲

۱۱-۵ استرانی‌های اشکال‌زدایی..... ۴۵۰

۱۱-۶ مثال: تشخیص عدد پندرقمی..... ۴۵۵

۱۲-۱ یادگیری ژرف در مقیاس بزرگ..... ۴۵۷

۱۲-۲ بینایی کامپیوتری..... ۴۶۶

۱۲-۳ تشخیص گفتار..... ۴۷۲

۱۲-۴ پردازش زبان طبیعی..... ۴۷۵

۱۲-۵ سایر کاربردها..... ۴۹۲