

یادگیری عمیق

مؤلفین

ایان گودفلو

یاشوا بنجیو

آرون کوروئل

مترجم

ایوب ترکیان

نیاز دانش

Goodfellow, Ian	: گودفلو، ایان	سرشناسه
	: یادگیری عمیق / مولفین: ایان گودفلو، یاشا بنجیو، آرون کوروئل؛ مترجم ایوب ترکیان	عنوان و نام پدیدآور
	: تهران: نیاز دانش، ۱۳۹۷.	مشخصات نشر
	: ۵۵۵۲ ص: نمودار.	مشخصات ظاهری
	: 978-600-8906-41-4	شابک
	: فیبا	وضعیت فهرست‌نویسی
	: Deep learning, [2016].	یادداشت
	: واژه‌نامه	یادداشت
	: یادگیری ژرف	عنوان دیگر
Machine learning	: فراگیری ماشینی	موضوع
Bengio, Yoshua	: بنژیو، یوشوا، ۱۹۶۴ - م	شناسه افزوده
Courville, Aaron	: کوروئل، آرون	شناسه افزوده
	: ترکیان، ایوب، ۱۳۳۷ - ، مترجم.	شناسه افزوده
	: الف ۱۳۹۷ ی ۲۹ گ ۵/۵ Q۳۲۵/۵	رده‌بندی کنگره
	: ۰۰۶/۳۱	رده‌بندی دیویی
	: ۸۸۴۸۳	شماره کتابشناسی مل



: نام کتاب	: یادگیری عمیق
: مؤلفین	: ایان گودفلو - یاشا بنجیو - آرون کوروئل
: مترجم	: ایوب ترکیان
: مدیر اجرایی - ناظر بر چاپ	: حمیدرضا محمد شیرازی - محمد شمس
: ناشر	: نیاز دانش
: صفحه‌آرا	: واحد تولید انتشارات نیاز دانش
: نوبت چاپ	: اول - ۱۳۹۷
: شمارگان	: ۱۰۰ نسخه
: قیمت	: ۶۵۰۰۰۰ ریال

ISBN:978-600-8906-41-4

شابک: ۹۷۸-۶۰۰-۸۹۰۶-۴۱-۴

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 می‌گیرند.

کلیه حقوق این اثر برای ناشر محفوظ است.

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

۰۲۱-۶۶۴۷۸۱۰۶-۶۶۴۷۸۱۰۸-۰۹۱۲۷۰۷۳۹۳۵

www.Niaze-Danesh.com

مشاوره جهت نشر: ۲۱۰۶۷۰۹ - ۰۹۱۲

فهرست مطالب

عنوان

شماره صفحه

فصل ۱ / مبانی یادگیری ماشین	۹۰
۱-۱ الگوریتم‌های یادگیری	۱۰
۱-۱-۱ فعالیت T	۱۰
۱-۱-۲ معیار عملکرد، P	۱۴
۱-۱-۳ تجربه، E	۱۵
۱-۱-۴ مثال: رگرسیون خطی	۱۸
۲-۱ ظرفیت، فرابرازش و فروبرازش	۲۱
۱-۲-۱ فرضیه نارایگانی	۲۹
۲-۲-۱ تنظیم	۲۹
۳-۱ درایاها و ست‌های صحنه‌گذاری	۳۲
۱-۱-۱ صحنه‌گذاری متقابل	۳۳
۴-۱ تخمین گرهای واریانس	۳۵
۱-۲-۱ تخمین	۳۵
۲-۴-۱ سوبه بایاس	۳۶
۳-۴-۱ واریانس خطا استاندارد	۳۹
۴-۴-۱ مصاحبه سوبه و واریانس در کمیته‌سازی میانگین مجذور خطا	۴۱
۵-۴-۱ سازگاری	۴۱
۵-۱ تخمین حداکثر شایدهی	۴۳
۱-۵-۱ الگاریتم شایدهی مشروط رخی میانگین مجذور	۴۴
۲-۵-۱ خصوصیات حداکثر شایدهی	۴۶
۶-۱ آمار بیزی	۴۷
۱-۶-۱ تخمین حداکثر پسین (MAP)	۵۰
۷-۱ الگوریتم‌های یادگیری با نظارت	۵۱
۱-۷-۱ یادگیری با نظارت احتمالی	۵۲
۲-۷-۱ ماشین‌های بردار پشتیبان	۵۳
۳-۷-۱ دیگر الگوریتم‌های ساده یادگیری با نظارت	۵۵
۸-۱ الگوریتم‌های یادگیری بدون نظارت	۵۸
۱-۸-۱ تحلیل مؤلفه‌های اصلی، PCA	۵۹
۲-۸-۱ خوشه‌بندی کامینز	۶۱
۹-۱ فرود گرادینانی استوکاستیک	۶۳
۱۰-۱ ساخت الگوریتم یادگیری ماشین	۶۵
۱۱-۱ چالش‌های انگیزشی یادگیری عمیق	۶۷
۱-۱۱-۱ نفرین بعد زیاد	۶۷
۲-۱۱-۱ ثبات محلی و تنظیم همواری	۶۸
۳-۱۱-۱ یادگیری مانیفولد	۷۲
فصل ۲ / شبکه‌های عمیق پیش‌خور	۷۷
۱-۲ مثال: یادگیری XOR	۸۰
۲-۲ یادگیری گرادینانی پایه	۸۶
۱-۲-۲ توابع هزینه	۸۷
۱-۱-۲-۲ یادگیری توزیع‌های مشروط با حداکثر شایدهی	۸۸
۲-۱-۲-۲ یادگیری آمار شرطی	۸۹

۹۰	۲-۲-۲ واحدهای خروجی
۹۰	۱-۲-۲-۲ واحدهای خطی برای توزیع‌های خروجی گوسی
۹۱	۲-۲-۲-۲ واحدهای سیگموئید برای توزیع‌های خروجی برنولی
۹۳	۳-۲-۲-۲ سافت‌مکس برای توزیع‌های خروجی مولتی‌نولی
۹۷	۴-۲-۲-۲ انواع خروجی دیگر
۱۰۱	۳-۲ واحدهای پنهان
۱۰۲	۱-۳-۲ واحدهای خطی تک‌سو و تعمیم‌های آنها
۱۰۵	۲-۳-۲ سیگموئید لجیستیک و تانژانت سهموی
۱۰۶	۳-۳-۲ دیگر واحدهای پنهان
۱۰۸	۴-۲ طراحی معماری
۱۰۸	۱-۴-۲ خصوصیات و عمق تخمین فراگیر
۱۱۳	۲-۴-۲ دیگر ملاحظات معماری
۱۱۳	۵-۲ س‌انتشار و دیگر الگوریتم‌های مشتق‌گیری
۱۱۴	۱-۵-۲ گراف‌های محاسباتی
۱۱۶	۵-۱ قاعده زنجیره حسابان
۱۱۷	۳-۵-۲ کاربرد برگشتی زنجیره در به‌دست‌آوردن پس‌انتشار
۱۲۱	۴-۵-۲ مشتق پس‌انتشار در MLP کاملاً متصل
۱۲۲	۵-۵-۲ مشتق‌های نشانده‌به-نشانده
۱۲۴	۶-۵-۲ پس‌انتشار کامل
۱۲۸	۷-۵-۲ مشتق پس‌انتشار برای آموزش MLP
۱۳۱	۸-۵-۲ پیچیدگی‌ها
۱۳۱	۹-۵-۲ مشتق‌گیری خارج از جامعه یادگیری عمیق
۱۳۳	۱۰-۵-۲ مشتق‌های به‌به‌بالا

فصل ۳ / تنظیم برای یادگیری عمیق

۱۳۷	۱-۳ جرایم اندازه پارامتر
۱۳۸	۱-۱-۳ تنظیم پارامتر
۱۴۱	۲-۱-۳ تنظیم
۱۴۴	۲-۳ جرایم اندازه به‌عنوان بهینه‌سازی مقید
۱۴۶	۳-۳ تنظیم و مسایل با قید کم
۱۴۷	۴-۳ افزودگی ست داده
۱۴۹	۵-۳ استحکام در مقابل نویز
۱۵۰	۱-۵-۳ تزریق نویز در اهداف خروجی
۱۵۱	۶-۳ یادگیری نیمه‌نظارتی
۱۵۲	۷-۳ یادگیری چندفعالیتی
۱۵۳	۸-۳ توقف زودرس
۱۶۰	۹-۳ پیوند زدن و اشتراک پارامتر
۱۶۲	۱۰-۳ نمایش تنک
۱۶۴	۱۱-۳ انباشتن و دیگر روش‌های آنسبل
۱۶۶	۱۲-۳ دورریزی
۱۷۶	۱۳-۳ آموزش خصمانه
۱۷۸	۱۴-۳ فاصله و انتشار تانژانت، طبقه‌گر تانژانت مانیفولد

فصل ۴ / بهینه‌سازی آموزش مدل‌های یادگیری عمیق

۱۸۴	۱-۴ تفاوت یادگیری و بهینه‌سازی محض
۱۸۵	۱-۱-۴ کمینه‌سازی ریسک تجربی

۱۸۶	۲-۱-۴	توابع اتلاف نیابتی و توقف زودرس
۱۸۶	۳-۱-۴	الگوریتم‌های بیج و مینی بیج
۱۹۲	۲-۴	چالش‌های بهینه‌سازی شبکه عصبی
۱۹۲	۱-۲-۴	بدحالتی
۱۹۳	۲-۲-۴	کمینه محلی
۱۹۵	۳-۲-۴	فلات، زین‌آسی، و دیگر نواحی مسطح
۱۹۸	۴-۲-۴	صخره و گرادیان‌های انفجاری
۱۹۹	۵-۲-۴	وابستگی‌های دراز مدت
۲۰۰	۶-۲-۴	گرادیان‌های غیردقیق
۲۰۰	۷-۲-۴	تناظر ضعیف بین ساختارهای محلی و فراگیر
۲۰۳	۸-۲-۴	محدودیت‌های تئوریک بهینه‌سازی
۲۰۳	۳-۴	الگوریتم‌های بنیادی
۲۰۳	۱-۳-۴	فروود گرادینانی استوکاستیک
۲۰۶	۲-۳-۴	گشتاور
۲۱۰	۳-۴	گشتاور نستروف
۲۱۰	۴-۴	راهبردهای آغازگری پارامتر
۲۱۷	۵-۴	الگوریتم‌های یادگیری تطبیقی
۲۱۷	۵-۴	Adagrad
۲۱۸	۲-۵-۴	Adaptive S
۲۱۹	۳-۵-۴	Adam
۲۲۱	۴-۵-۴	انتخاب الگوریتم به‌سازی مناسب
۲۲۱	۶-۴	روش‌های تخمین درجه ۱
۲۲۱	۱-۶-۴	روش نیوتن
۲۲۳	۲-۶-۴	گرادیان‌های مزدوم
۲۲۷	۳-۶-۴	BFGS
۲۲۸	۷-۴	راهبردهای بهینه‌سازی و مثال‌الگوریتم‌ها
۲۲۸	۱-۷-۴	نرمال‌سازی ناپویسته
۲۳۲	۲-۷-۴	فروود مختصاتی
۲۳۳	۳-۷-۴	میانگین‌گیری Polyak
۲۳۴	۴-۷-۴	پیش‌آموزش با نظارت
۲۳۷	۵-۷-۴	طراحی‌های مدل معین بهینه‌سازی
۲۳۸	۶-۷-۴	روش‌های تداوم و یادگیری دوره‌ای

فصل ۵ / شبکه‌های کانولوشن

۲۴۱	۱-۵	عملیات کانولوشن
۲۴۲	۲-۵	انگیزه
۲۴۵	۳-۵	رای‌گیری
۲۵۰	۴-۵	کانولوشن و رای‌گیری به‌عنوان پیشینی بی‌نهایت قوی
۲۵۷	۵-۵	روایت‌های تابع کانولوشن پایه
۲۶۸	۶-۵	خروجی‌های ساختارمند
۲۶۹	۷-۵	نوع داده
۲۷۱	۸-۵	الگوریتم‌های کانولوشن کارآمد
۲۷۲	۹-۵	ویژگی‌های تصادفی یا بدون نظارت
۲۷۳	۱۰-۵	مبنای علوم اعصاب شبکه‌های کانولوشن
۲۸۱	۱۱-۵	شبکه‌های کانولوشن و سابقه یادگیری عمیق

فصل ۶ / مدل سازی توالی: شبکه های برگشتی و بازگشتی..... ۲۸۳

۲۸۵	۱-۶	وایچی گراف های محاسباتی
۲۸۸	۲-۶	شبکه های عصبی برگشتی
۲۹۲	۱-۲-۶	اجبار معلم و شبکه های دارای برگشت خروجی
۲۹۵	۲-۲-۶	محاسبه گرادیان در شبکه برگشتی
۲۹۷	۳-۲-۶	شبکه RNN به عنوان مدل های گرافیکی جهت دار
۳۰۱	۴-۲-۶	مدل سازی توالی مشروط به متن با RNN
۳۰۴	۳-۶	RNN دوچته
۳۰۶	۴-۶	معماری های توالی به توالی رمزگذار-رمزدا
۳۰۸	۵-۶	شبکه های برگشتی عمیق
۳۰۹	۶-۶	شبکه های عصبی بازگشتی
۳۱۱	۷-۶	پیش و ابستگی های دراز مدت
۳۱۳	۸-۶	شرایط های حالت اکو
۳۱۶	۹-۶	واحدهای نشتی و راهبردهای مقیاس های چند زمانی
۳۱۶	۱-۶	افزودن پیوندهای گذر زمانی
۳۱۷	۲-۹-۶	واحد های نشتی و طیف مقیاس های زمانی مختلف
۳۱۷	۳-۹-۶	حذف اتصالات
۳۱۸	۱۰-۶	حافظه کوتاه مدت و دیگر RNN های درگاهی
۳۱۹	۱-۱۰-۶	STIM
۳۲۱	۲-۱۰-۶	دیگر RNN های درگاهی
۳۲۲	۱۱-۶	بهینه سازی وابستگی های دراز مدت
۳۲۳	۱-۱۱-۶	چین گرادیان
۳۲۵	۲-۱۱-۶	تنظیم برای تشویق جریان اطلاعات
۳۲۶	۱۲-۶	حافظه صریح

فصل ۷ / روش شناسی عملی..... ۳۳۱

۳۳۲	۱-۷	سنجش های عملکرد
۳۳۵	۲-۷	مدل های خط پایه
۳۳۶	۳-۷	تعیین نیاز به داده های بیشتر
۳۳۸	۴-۷	انتخاب فرایارامترها
۳۳۸	۱-۴-۷	تنظیم دستی فرایارامتر
۳۴۲	۲-۴-۷	الگوریتم های بهینه سازی فرایارامتر اتوماتیک
۳۴۳	۳-۴-۷	جستجوی شبکه های
۳۴۵	۴-۴-۷	جستجوی تصادفی
۳۴۶	۵-۴-۷	بهینه سازی فرایارامتر مدل پایه
۳۴۷	۵-۷	راهبردهای خطاگیری
۳۵۱	۶-۷	مثال: شناسایی چند عددی

فصل ۸ / کاربردها..... ۳۵۵

۳۵۵	۱-۸	یادگیری عمیق مقیاس بزرگ
۳۵۶	۱-۱-۸	پیاده سازی CPU سریع
۳۵۶	۲-۱-۸	پیاده سازی GPU
۳۵۹	۳-۱-۸	پیاده سازی توزیعی مقیاس بزرگ
۳۶۰	۴-۱-۸	فشرده سازی مدل
۳۶۱	۵-۱-۸	ساختار دینامیکی

۳۶۳	۶-۱-۸ پیاده‌سازی سخت‌افزار خاص شبکه عمیق
۳۶۵	۲-۸ بینایی ماشین
۳۶۵	۱-۲-۸ پیش‌پردازش
۳۶۶	۱-۱-۲-۸ نرمال‌سازی کانتراست
۳۷۰	۲-۱-۲-۸ تقویت مجموعه داده
۳۷۱	۳-۸ شناسایی گفتار
۳۷۴	۴-۸ پردازش زبان طبیعی
۳۷۴	۱-۴-۸ n-گرام
۳۷۷	۲-۴-۸ مدل‌های زبان عصبی
۳۷۸	۳-۴-۸ خروجی‌های با ابعاد زیاد
۳۷۹	۱-۳-۴-۸ استفاده از لیست کوتاه
۳۸۰	۲-۳-۴-۸ سافت مکس سلسله‌مراتبی
۳۸۳	۳-۳-۴-۸ نمونه‌برداری اهمیتی
۳۸۵	۴-۳-۴-۸ تخمین نویز کانتراستی و اتلاف درجه‌بندی
۳۸۵	۴-۲۰ تلفیق مدل‌های زبان عصبی با n-گرام
۳۸۶	۵-۴-۸ ترجمه ماشین عصبی
۳۸۸	۵-۵-۸ استفاده مکانیسم توجه و تراز قطعات داده
۳۹۰	۶-۴-۸ چشم‌انداز تاریخی
۳۹۱	۵-۸ کاربرد یادگیری
۳۹۱	۱-۵-۸ مدل‌های صیه‌گر
۳۹۴	۱-۱-۵-۸ کاوش در مقابل انتفاع
۳۹۶	۲-۵-۸ نمایش دانش، استدلال و پاسخ‌دادن سوال

فصل ۹ / خودرمزگذارها

۴۰۰	۱-۹ خودرمزگذارهای فرو کامل
۴۰۱	۲-۹ خودرمزگذارهای تنظیم‌شده
۴۰۲	۱-۲-۹ خودرمزگذارهای تنک
۴۰۴	۲-۲-۹ خودرمزگذارهای نویززا
۴۰۵	۳-۲-۹ تنظیم با جریمه کردن مشق‌ها
۴۰۶	۳-۹ قدرت نمایشی، اندازه لایه و عمق
۴۰۶	۴-۹ رمزگذارها و رمززدهای استوکاستیک
۴۰۸	۵-۹ خودرمزگذارهای نویززا
۴۰۹	۱-۵-۹ تخمین نمره
۴۱۲	۱-۱-۵-۹ چشم‌انداز تاریخی
۴۱۲	۶-۹ مانیفولدهای یادگیری با خودرمزگذارها
۴۱۷	۷-۹ خودرمزگذارهای انقباضی
۴۲۰	۸-۹ تجزیه تنک پیش‌بینانه
۴۲۰	۹-۹ کاربردهای خودرمزگذار

فصل ۱۰ / یادگیری نمایشی

۴۲۵	۱-۱۰ پیش‌آموزش بدون نظارت لایه‌وار مقتصدانه
۴۲۷	۱-۱-۱۰ چه موقع و چرا پیش‌آموزش بدون نظارت کار می‌کند؟
۴۳۳	۲-۱۰ یادگیری انتقالی و انطباق دامنه
۴۳۷	۳-۱۰ واپیچی نیمه‌نظارتی عوامل علی
۴۴۳	۴-۱۰ نمایش توزیعی
۴۵۰	۵-۱۰ نفع‌نمایی از عمق

۴۵۱-۶-۱۰ تأمین سرینخ برای کشف علل نهشته

فصل ۱۱ / مدل‌های زایشی عمیق

۴۵۵	ماشین‌های بولتزمن	۱-۱۱
۴۵۷	ماشین‌های بولتزمن محدود	۲-۱۱
۴۵۹	توزیع‌های شرطی	۱-۲-۱۱
۴۶۱	آموزش ماشین‌های بولتزمن محدود	۲-۲-۱۱
۴۶۱	شبکه‌های باور عمیق	۳-۱۱
۴۶۴	ماشین‌های بولتزمن عمیق	۴-۱۱
۴۶۷	خصوصیات جالب	۱-۴-۱۱
۴۶۸	استنباط میانگین میدان DBM	۲-۴-۱۱
۴۷۰	یادگیری پارامتر DBM	۳-۴-۱۱
۴۷۱	پیش‌آموزش لایه‌وار	۴-۴-۱۱
۴۷۴	آموزش مشترک ماشین‌های بولتزمن عمیق	۵-۴-۱۱
۴۷۷	ماشین بولتزمن برای داده‌های با مقدار حقیقی	۵-۱۱
۴۷۸	RF، ای گوسی-برنولی	۵-۱
۴۸۰	مدل‌های بدون جهت کوواریانس شرطی	۲-۱۱
۴۸۵	ماشین‌های بولتزمن با لایه‌های	۶-۱۱
۴۸۷	ماشین بولتزمن برای خروجی ساختارمند یا متوالی	۷-۱۱
۴۸۸	دیگر ماشین‌های بولتزمن	۸-۱۱
۴۸۹	پس‌انتشار از داخل عملیات رانده	۹-۱۱
۴۹۱	پس‌انتشار از داخل عملیات ناپیوسته استوکاستیک	۱۰-۹-۱۱
۴۹۴	شبکه‌های زایشی جهت‌دار	۱۰-۱۱
۴۹۵	شبکه‌های باور سیستمی	۱-۱۰-۱۱
۴۹۶	شبکه‌های ژنراتور مسطح	۲-۱۰-۱۱
۴۹۹	خودرمزگذارهای تغییرپذیر	۳-۱۰-۱۱
۵۰۲	شبکه‌های خصمانه زایشی	۴-۱۰-۱۱
۵۰۶	شبکه‌های انطباق مومان زایشی	۵-۱۰-۱۱
۵۰۷	شبکه‌های زایشی کانولوشن	۶-۱۰-۱۱
۵۰۸	شبکه‌های اتورگرسیو	۷-۱۰-۱۱
۵۰۹	شبکه‌های اتورگرسیو خطی	۸-۱۰-۱۱
۵۰۹	شبکه‌های اتورگرسیو عصبی	۹-۱۰-۱۱
۵۱۱	NADE	۱۰-۱۰-۱۱
۵۱۴	برداشت نمونه از خودرمزگذارها	۱۱-۱۱
۵۱۴	زنجیره مارکوف مرتبط با هر خودرمزگذار نویززا	۱-۱۱-۱۱
۵۱۵	نمونه‌برداری گیرهای و شرطی	۲-۱۱-۱۱
۵۱۶	روش آموزش پس‌رو	۳-۱۱-۱۱
۵۱۶	شبکه‌های استوکاستیک زایشی	۱۲-۱۱
۵۱۷	GSN‌های تمایزی	۱-۱۲-۱۱
۵۱۸	دیگر شماهای تولید	۱۳-۱۱
۵۱۹	ارزیابی مدل‌های زایشی	۱۴-۱۱
۵۲۲	نتیجه‌گیری	۱۵-۱۱

پیوست الف / فرهنگ‌نامه اصطلاحات

۵۲۳	فارسی به انگلیسی
۵۳۹	انگلیسی به فارسی