

به نام خدا

۲۱۱۴۱۵۰

شبکه‌های عصبی

و

یادگیری ژرف

تألیف و ترجمه:

دکتر مهدی اسماعیلی

با همکاری: مهندس فائزه اکبری

شبکه‌های عصبی و یادگیری ژرف

تألیف و ترجمه: دکتر مهدی اسماعیلی

با همکاری: مهندس فائزه اکبری

ناشر: انتشارات آتی‌نگر

ناشر همکار: انتشارات وینا

طراحی جلد: صفحه‌آرایی: همتا بیداریان

چاپ اوایل ۱۳۹۹

شمارگان: ۷۵۰ جلد

قیمت: ۶۴۰,۰۰۰ ریال

شابک: ۹۷۸-۶۰۲-۶۱۰۲-۶۶۵۳۳۶-۲

ISBN: 978-622-6102-74-2

حق چاپ برای انتشارات آتی‌نگر محفوظ است.

نشانی دفتر فروش: تهران، خیابان جلالیه، روبروی کوچه رشتچی، پلاک ۱۴۴، واحد ۱

تلفن: ۸-۶۶۵۶۵۳۳۶-۸
بر: ۶۶۵۶۵۳۳۷



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

سرشناسه: اسماعیلی، مهدی، ۱۳۵۰ - [دکتر]

شبکه‌های عصبی و یادگیری ژرف / تألیف و ترجمه: دکتر مهدی اسماعیلی، همکاری: مهندس فائزه اکبری

تهران: آتی‌نگر، وینا ۱۳۹۹

۳۲۸ ص.: مصور، جدول، نمودار.

ISBN: 978-622-6102-74-2

فیبا.

یادداشت: کتاب حاضر برگرفته از کتاب Neural networks and deep learning: a textbook, 2018، اثر چارو سی آگروال است.

موضوع: شبکه‌های عصبی (کامپیوتر) - Neural networks (Computer science)

موضوع: علوم کامپیوتر - Computer science

شناسه افزوده: اکبری، فائزه، ۱۳۶۱ -

شناسه افزوده: آگروال، چارو سی، Aggarwal, Charu C. ۱۹۷۰ - م.

رده‌بندی کنگره

رده‌بندی دیویی

شماره کتابشناسی ملی

QAY۶/۷

۰۰۶/۳۲

۶۱۶۳۳۳

فهرست مطالب

پیشگفتار

۷

فصل اول: مقدمه‌ای بر شبکه‌های عصبی

۹

۱۱	۲-۱ معماری پایه در شبکه‌های عصبی
۱۸	۳-۱ طراحی شبکه عصبی
۲۶	۴-۱ آموزش شبکه عصبی با پس‌انتشار
۳۵	۵-۱ چالش‌هایی در آموزش شبکه‌های عصبی
۴۰	۶-۱ معماری‌های متداوم
۴۳	۷-۱ خلاصه
۴۴	۸-۱ تمرین‌ها

۴۷

فصل دوم: یادگیری ماشین و شبکه‌های عصبی

۴۷	۱-۲ مقدمه
۴۸	۲-۲ معماری‌هایی برای مدل‌های رده‌بندی نودویی
۵۵	۳-۲ معماری‌هایی برای مدل‌های چندرده‌ای (چندکلاس)
۶۰	۴-۲ خودرمزگذارها
۶۸	۵-۲ کاربردی از یک شبکه عصبی ساده؛ word2vec
۷۹	۶-۲ سیستم‌های پیشنهاد دهنده و شبکه‌های عصبی
۸۱	۷-۲ معماری‌های ساده عصبی برای نمایش گراف
۸۳	۸-۲ خلاصه
۸۳	۹-۲ تمرین‌ها

۸۵

فصل سوم: یادگیری شبکه‌های عصبی ژرف

۸۵	۱-۳ مقدمه
۸۵	۲-۳ پس‌انتشار: ساده‌سازی با محاسبات مجزا
۹۰	۳-۳ مقاداردهی اولیه و برپایی شبکه
۹۶	۴-۳ محو و انفجار گرادیان
۱۰۳	۵-۳ استراتژی‌های کاهش گرادیان
۱۱۱	۶-۳ ترندهایی کاربردی برای تسریع و فشرده‌سازی

۱۱۶	۷-۳ خلاصه
۱۱۷	۸-۳ تمرین‌ها

۱۱۹ فصل چهارم: یادگیری در راستای تعمیم بهتر

۱۱۹	۱-۴ مقدمه
۱۲۲	۲-۴ موازنه بایاس و واریانس
۱۲۹	۳-۴ تقسیم داده‌ها به سه بخش آموزشی، آزمایشی و اعتبارسنجی
۱۳۳	۴-۴ تنظیم براساس جریمه (تنظیم L_1 , L_2)
۱۳۵	۵-۴ استفاده از روش‌های تلفیقی
۱۴۰	۶-۴ توقف زودرس
۱۴۱	۷-۴ پیش‌آموزش: روشی برای مقداردهی اولیه اوزان
۱۴۴	۸-۴ ساده‌سازی مسئله
۱۴۵	۹-۴ به اشتراک گذاری با سرمایه
۱۴۶	۱۰-۴ انواع خودرمزگذارها: تعلیم در کاربردهای بیناظر
۱۴۸	۱۱-۴ خلاصه
۱۴۹	۱۲-۴ تمرین‌ها

۱۵۱ فصل پنجم: شبکه‌های تابع پایه شعاعی

۱۵۱	۱-۵ مقدمه
۱۵۱	۲-۵ معماری شبکه‌های RBF
۱۵۳	۳-۵ آموزش یک شبکه RBF
۱۵۹	۴-۵ استفاده از شبکه‌های RBF برای رده‌بندی
۱۶۱	۵-۵ رابطه شبکه‌های RBF با روش‌های کرنل
۱۶۳	۶-۵ خلاصه
۱۶۴	۷-۵ تمرین‌ها

۱۶۵ فصل ششم: ماشین‌های بولتزمن محدود

۱۶۵	۱-۶ مقدمه
۱۶۶	۲-۶ شبکه‌های هاپفیلد
۱۶۹	۳-۶ ماشین بولتزمن
۱۷۳	۴-۶ ماشین‌های بولتزمن محدود
۱۷۷	۵-۶ کاربردهایی از ماشین‌های بولتزمن محدود
۱۸۲	۶-۶ استفاده از RBM فراتر از داده‌های دودویی

۱۸۴	۶-۷ پشتهای از ماشین‌های بولتزمن محدود
۱۸۷	۶-۸ خلاصه
۱۸۷	۶-۹ پرسش‌ها

فصل هفتم: شبکه‌های عصبی برگشتی

۱۸۹	۷-۱ مقدمه
۱۹۰	۷-۲ معماری شبکه‌های عصبی برگشتی
۲۰۲	۷-۳ چالش‌های آموزش شبکه‌های برگشتی
۲۰۴	۷-۴ شبکه‌های حالت اِکو
۲۰۵	۷-۵ حافظه کوتاه مدت طولانی (LSTM)
۲۰۸	۷-۶ واحد عصبی برگشتی گیت دار
۲۱۰	۷-۷ کاربردهایی از شبکه‌های عصبی برگشتی
۲۲۲	۷-۸ خلاصه
۲۲۳	۷-۹ تمرین‌ها

فصل هشتم: شبکه‌های عصبی کانولوشن

۲۲۵	۸-۱ مقدمه
۲۲۶	۸-۲ ساختار اصلی یک شبکه کانولوشن
۲۳۶	۸-۳ آموزش یک شبکه کانولوشن
۲۴۲	۸-۴ چند نمونه از معماری کانولوشن
۲۵۰	۸-۵ برخی از کاربردهای شبکه‌های کانولوشن
۲۵۷	۸-۶ خلاصه
۲۵۷	۸-۷ تمرین‌ها

فصل نهم: یادگیری تقویتی ژرف

۲۵۹	۹-۱ مقدمه
۲۶۱	۹-۲ چالش‌های یادگیری تقویتی
۲۶۲	۹-۳ یادگیری تابع ارزش
۲۶۳	۹-۴ خودراه‌اندازی برای یادگیری تابع ارزش
۲۷۳	۹-۵ روش‌های گرادینان سیاست
۲۸۱	۹-۶ کاربردهای عملی یادگیری تقویتی
۲۸۳	۹-۷ خلاصه
۲۸۴	۹-۸ پرسش‌ها

۲۸۵

فصل دهم: عناوین پیشرفته در یادگیری ژرف

۲۸۵	۱-۱۰ مقدمه
۲۸۶	۲-۱۰ استفاده از مؤلفه توجه
۲۹۳	۳-۱۰ شبکه‌هایی با حافظه خارجی
۲۹۷	۴-۱۰ شبکه‌های تخصصی مولد
۳۰۷	۵-۱۰ یادگیری رقابتی
۳۱۳	۶-۱۰ خلاصه
۳۱۴	۷-۱۰ تمرین‌ها

۳۱۵

ضمیمه الف: شبکه‌های عصبی مصنوعی در عمل

۳۲۱

ضمیمه ب: منابع

برخیز ز خواب تا شرابی بخوریم
زان پیش که از زمانه تابی بخوریم
کاین چرخ ستیزه روی ناگه روزی
چندان ندهد زمان که آبی بخوریم

پیشگفتار

شبکه‌های عصبی مصنوعی یکی از تکنیک‌های محبوب و رایج در حوزه یادگیری ماشین محسوب می‌شوند و با پیشرفت‌های اخیر می‌توان گفت که شبکه‌های عصبی امروزه یکی از رایج‌ترین روش‌های است که برای شایع و شناسایی الگوهای موجود در داده‌ها استفاده می‌شود. حجم بالای داده‌های در دسترس و همچنین افزایش قدرت محاسباتی، موفقیت‌هایی را برای شبکه‌های عصبی به همراه داشته و این محبوبیت با توانایی جدیدی برای شبکه‌های عصبی با نام یادگیری ژرف همراه شده است. اگرچه تا نزدیکی عملکرد این مصنوعی به عملکرد انسان راهی طولانی (شاید هم کوتاه) باقیمانده، اما امروزه در برخی از حوزه‌های کاربردی نظیر شناسایی تصاویر، اتومبیل‌های خودران و بازی‌های کامپیوتری، موفقیت‌های چشمگیری برای شبکه‌های عصبی و به خصوص یادگیری ژرف رقم خورده است. توجه داشته باشید که از نظر تئوری، با شرط وجود مجموعه داده‌های کافی؛ شبکه‌های عصبی قادرند هر تابع ریاضی را بیاموزند و برخی از نسخه‌های آن مانند شبکه‌های برگشتی به عنوان تورینگ کامل شناخته می‌شوند.

تمرکز کتاب حاضر بر روی موضوع شبکه‌های عصبی و یادگیری ژرف است و مطالب آن در ده فصل و دو ضمیمه گردآوری و تهیه شده است. در فصل اول با مفاسیم پایه در شبکه‌های عصبی آشنا می‌شویم. شبیه‌سازی مدل‌های متعدد یادگیری ماشین با کمک شبکه‌های عصبی در فصل دوم آمده است. اگرچه در دو فصل اول کتاب مروری اجمالی بر روی روش‌های آموزش شبکه‌های عصبی دیده می‌شود، اما جزییات بیشتر آن و همچنین چالش‌های آموزش شبکه‌های عصبی در فصل‌های سوم و چهارم شرح داده شده است. شبکه‌های تابع پایه شعاعی و ماشین‌های بولتزمن محدود به ترتیب در فصل‌های پنجم و ششم بحث شده‌اند. شاید به جرأت بتوان گفت که شبکه‌های عصبی، موفقیت‌های اخیر خود را مرهون دو معماری برگشتی و کانولوشن است؛ معماری‌هایی که قدرت خود را در کار با متون، صوت و تصویر نشان داده‌اند. شبکه‌های عصبی برگشتی و شبکه‌های عصبی کانولوشن موضوعات فصل‌های هفتم و هشتم کتاب را تشکیل می‌دهند. فصل نهم به نقش یادگیری ژرف در یادگیری

تقویتی تخصیص داده شده است. در فصل آخر نیز به صورت اجمالی به سراغ موضوعات پیشرفته‌ای نظیر استفاده از مؤلفه توجه، مکانیزم‌های تورینگ عصبی و شبکه‌های تخصصی مولد رفته‌ایم.

تمام تلاش خود را انجام داده‌ایم تا نگارش کتاب به گونه‌ای انجام شود که خوانندگان محترم آن بتوانند مفاهیم آن را به راحتی درک کنند. بدون شک ممکن است برای برخی از واژه‌های انگلیسی بتوان معادل‌های بهتری یافت. آنچه مسلم است این است که شاید گاهی تبدیل واژه‌ها آنچنان که باید و شاید انجام نشده است؛ اما در ادای جملات و بیان موضوع تلاش فراوان شده است تا خوانندگان گرامی با متنی مبهم و گیج کننده روبه‌رو نشوند.

در اینجا لازم می‌دانم از همه اساتید و دانشجویان به خاطر راهنمایی‌های ارزشمندشان در حین آماده‌سازی این کتاب سپاسگزاری کنم. بدون شک، همت همکار محترم سرکار خانم فائزه اکبری باعث شد تا این کتاب مهیا شود و بدون لطف ایشان این کار میسر نمی‌شد. همچنین از مدیریت محترم انتشارات آتی‌نگر و دوستان عزیزم جناب رامین مولاناپور نیز به خاطر آماده‌سازی، چاپ و پخش این کتاب تشکر می‌کنم. رهین محبت بی‌دریغ خانواده‌ام هستم که با فراهم‌سازی محیطی مناسب مرا یاری نمودند. اما با وجود همه سعی و تلاشی که در تمام مراحل آماده‌سازی این کتاب انجام گرفته است، یقین دارم که عاری از اشتباه است، چرا که تنها مکتوب بی‌نقص همان معجزه جاوید قرآن کریم است. در آخر ضمن سپاسگزاری از همه کسانی که مرا یاری داده‌اند و با پذیرش مسئولیت هرگونه کاستی احتمالی، امیدوارم که این اندک مفید افتد.