

نویسنده: کیم فالک

سیستم‌های توصیه‌گر کاربردی

مترجمان:

دکتر نوذر ابراهیمی لامع

دکتر فاطمه ثقفی

(عضو هیئت علمی دانشگاه تهران)

سیستم‌های توصیه‌گر کاربردی

مترجمان: دکتر نوذر ابراهیمی لامع، دکتر فاطمه ثقفی

ویراستار علمی: دکتر رامین مولاناپور

ناشر: انتشارات آتی‌نگر

طراحی جلد و صفحه‌آرایی: همتا بیداریان

چاپ اول، ۱۴۰۲

شمارگان: ۵۰۰ نسخه

قیمت: ۴,۲۰۰,۰۰۰ ریال

شابک: ۹۷۸-۶۲۲-۷۵۷۱-۸۹-۹

ISBN: 978-622-7571-89-9

حق چاپ برای انتشارات آتی‌نگر محفوظ است.

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

نمابر: ۶۶۵۶۵۳۳۷

تلفن: ۸-۶۶۵۶۵۳۳۶



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

سرشناسه: فالک، کیم، FALK، KIM

سیستم‌های توصیه‌گر کاربردی / نویسنده: کیم فالک / مترجمان: نوذر ابراهیمی لامع، فاطمه ثقفی، ویراستار علمی: رامین مولاناپور

تهران: آتی‌نگر ۱۴۰۲

۴۹۸ ص.: مصور، جدول، نمودار.

ISBN: 978-622-7571-89-9

فیبا.

یادداشت: عنوان اصلی کتاب: Practical Recommender Systems, 2019

موضوع: نظام‌های توصیه‌گر

موضوع: Recommender systems (Information filtering)

شناسه‌افزوده: ابراهیمی لامع، نوذر، ۱۳۴۷، - مترجم

شناسه‌افزوده: ثقفی، فاطمه، ۱۳۴۹، - مترجم

شناسه‌افزوده: مولاناپور، رامین، ۱۳۵۲، - ویراستار

شناسه‌افزوده: بیداریان، همتا، ۱۳۶۱، - طراحی کتاب

رده‌بندی کنگره

رده‌بندی دیویی

شماره کتابشناسی ملی

QA۷۶/۹

۰۰۵/۵۶

۹۳۳۵۶۰۲

فهرست مطالب

پیشگفتار.....	۱۳
در مورد این کتاب	۱۵
چه کسی باید این کتاب را بخواند.....	۱۵
نحوه تنظیم این کتاب.....	۱۵
دانلودها.....	۱۷
قراردادهای کد.....	۱۷
انجمن کتاب.....	۱۷
درباره نویسنده.....	۱۸
مقدمه مترجمان.....	۱۹

بخش اول: آشنایی با سیستم‌های توصیه‌گر

فصل اول: سیستم توصیه‌گر چیست؟.....	۲۳
۱-۱ توصیه‌گرهای دنیای واقعی.....	۲۳
۱-۱-۱ سیستم‌های توصیه‌گر در خانه و روی اینترنت.....	۲۵
۱-۱-۲ اصطلاح دنباله بلند.....	۲۶
۱-۱-۳ سیستم توصیه‌گر Netflix.....	۲۶
۱-۱-۴ تعریف سیستم توصیه‌گر.....	۳۴
۱-۲ رده‌بندی سیستم‌های توصیه‌گر.....	۳۷
۱-۲-۱ حوزه.....	۳۷
۱-۲-۲ منظور.....	۳۸
۱-۲-۳ زمینه (Context).....	۳۸
۱-۲-۴ سطح شخصی‌سازی.....	۳۹
۱-۲-۵ نظرات چه کسانی.....	۴۰
۱-۲-۶ حریم خصوصی و اعتماد.....	۴۱
۱-۲-۷ رابط کاربری.....	۴۲
۱-۲-۸ الگوریتم‌ها.....	۴۵
۱-۳ یادگیری ماشین و جایزه Netflix.....	۴۷

۴۸	۱-۴ وبسایت MovieGEEKs
۴۹	۱-۴-۱ طراحی و مشخصات
۵۰	۱-۴-۲ معماری
۵۲	۱-۵ ساخت سیستم توصیه‌گر
۵۴	خلاصه
۵۴	یادداشت پایانی

فصل دوم: رفتار کاربر و نحوه جمع‌آوری آن

۵۸	۲-۱ نحوه جمع‌آوری شواهد هنگام مرور سایت Netflix
۶۰	۲-۱-۱ جمع‌آوری شواهد توسط Netflix
۶۲	۲-۲ یافتن رفتار مفید کاربر
۶۳	۲-۲-۱ ثبت تجربیات بازدیدکنندگان
۶۴	۲-۲-۲ یادگیری از مرورگر یک فروشگاه
۶۹	۲-۲-۳ عمل خرید
۷۰	۲-۲-۴ مصرف محصولات
۷۲	۲-۲-۵ امتیازات بازدیدکنندگان
۷۵	۲-۲-۶ آشنایی با مشتریان خود به روش (قدیمی) Netflix
۷۶	۲-۳ شناسایی کاربران
۷۶	۲-۴ دریافت داده‌های بازدیدکننده از منابع دیگر
۷۷	۲-۵ جمع‌کننده
۷۹	۲-۵-۱ ساخت فایل‌های پروژه
۷۹	۲-۵-۲ مدل داده
۸۰	۲-۵-۳ کیش رفتن: جمع‌کننده شواهد سمت کلاینت
۸۲	۲-۵-۴ ادغام جمع‌کننده در MovieGEEKs
۸۵	۲-۶ کاربران سیستم و نحوه مدل‌سازی آن‌ها
۸۸	خلاصه
۸۸	یادداشت پایانی

فصل سوم: پایش سیستم

۹۰	۳-۱ ایده خوب افزودن داشبورد
۹۱	۳-۱-۱ پاسخ سؤال «حال وبسایت ما چطور است؟»
۹۲	۳-۲ تحلیل‌شناسی
۹۲	۳-۲-۱ تحلیل‌شناسی وب
۹۳	۳-۲-۲ آمار اولیه

۹۴تبدیل‌ها ۳-۲-۳
۹۷تبدیل مسیر منجر به تبدیل ۳-۲-۴
۹۹مسیر تبدیل ۳-۲-۵
۱۰۲کاربرنماها ۳-۳
۱۰۵MovieGEEKs داشبورد ۳-۴
۱۰۵تولید خودکار داده‌ها برای ثبت وقایع ۳-۴-۱
۱۰۶مشخصات طراحی داشبورد تحلیل‌شناسی ۳-۴-۲
۱۰۷چیدمان داشبورد تحلیل‌شناسی ۳-۴-۳
۱۰۸معماری ۳-۴-۴
۱۱۰خلاصه

فصل چهارم: امتیازات و نحوه محاسبه آن‌ها ۱۱۱

۱۱۲ترجیحات ماتریس آیتم - کاربر ۴-۱
۱۱۲تعریف امتیازات ۴-۱-۱
۱۱۳ماتریس کاربر - آیتم ۴-۱-۲
۱۱۵امتیازات صریح یا ضمنی ۴-۲
۱۱۶نحوه استفاده از منابع معتبر برای توصیه‌ها ۴-۲-۱
۱۱۷مرور امتیازات صریح ۴-۳
۱۱۸امتیازات ضمنی چیست؟ ۴-۴
۱۲۰پیشنهاد دادن افراد ۴-۴-۱
۱۲۰ملاحظات محاسبه امتیازات ۴-۴-۲
۱۲۴محاسبه امتیازات ضمنی ۴-۵
۱۲۵نگاهی به داده‌های رفتاری ۴-۵-۱
۱۲۹یک مسئله یادگیری ماشین ۴-۵-۲
۱۳۰نحوه پیاده‌سازی امتیازات ضمنی ۴-۶
۱۳۴افزودن جنبه زمانی ۴-۶-۱
۱۳۷ارزش بیشتر اقلام کم تکرار ۴-۷
۱۴۰خلاصه
۱۴۰یادداشت پایانی

فصل پنجم: توصیه‌های غیرشخصی ۱۴۱

۱۴۲توصیه غیرشخصی ۵-۱
۱۴۲آگهی بازرگانی ۵-۱-۱
۱۴۴عملکرد یک توصیه ۵-۱-۲

۱۴۴	۵-۲ نحوه ارائه توصیه‌ها بدون داده.....
۱۴۶	۵-۲-۱ لیست ۱۰ آیتم برتر: نمودار آیتم‌ها.....
۱۴۸	۵-۳ پیاده‌سازی نمودار و کارهای مقدماتی برای اجزای سیستم توصیه‌گر.....
۱۴۸	۵-۳-۱ اجزا سیستم توصیه‌گر.....
۱۵۰	۵-۳-۲ کد MovieGEEKs بر گرفته از GitHub.....
۱۵۰	۵-۳-۳ سیستم توصیه‌گر.....
۱۵۰	۵-۳-۴ افزودن نمودار به MovieGEEKs.....
۱۵۲	۵-۳-۵ نمایش جذاب‌تر محتوا.....
۱۵۴	۵-۴ توصیه‌های بذری.....
۱۵۴	۵-۴-۱ آیتم‌های خرید تکراری همانند آیتم مروری شما.....
۱۵۶	۵-۴-۲ قوانین انجمنی.....
۱۶۱	۵-۴-۳ اجرای قوانین انجمنی.....
۱۶۶	۵-۴-۴ ذخیره قوانین انجمنی در پایگاه داده.....
۱۶۸	۵-۴-۵ محاسبه قوانین انجمنی.....
۱۶۹	۵-۴-۶ استفاده از رویدادهای مختلف برای ایجاد قوانین انجمنی.....
۱۷۱	خلاصه.....
۱۷۱	یادداشت پایانی.....

فصل ششم: شروع سرد کاربر (محتوا)..... ۱۷۳

۱۷۳	۶-۱ شروع سرد چیست؟.....
۱۷۵	۶-۱-۱ محصولات سرد.....
۱۷۵	۶-۱-۲ بازدیدکننده سرد.....
۱۷۷	۶-۱-۳ گوسفند خاکستری.....
۱۷۷	۶-۱-۴ نمونه‌هایی از زندگی واقعی.....
۱۷۹	۶-۱-۵ اقدامات لازم در مواجهه با شروع سرد.....
۱۷۹	۶-۲ ردیابی بازدیدکنندگان.....
۱۸۰	۶-۲-۱ کاربران ناشناس پایدار.....
۱۸۰	۶-۳ الگوریتم‌های حل مشکل شروع سرد.....
۱۸۱	۶-۳-۱ استفاده از قوانین انجمنی برای ایجاد توصیه‌ها برای کاربران سرد.....
۱۸۳	۶-۳-۲ استفاده از دانش حوزه و قوانین کسب‌وکار.....
۱۸۳	۶-۳-۳ استفاده از بخش‌ها.....
۱۸۵	۶-۳-۴ استفاده از طبقه‌بندی‌ها برای حل مشکل گوسفند خاکستری و نحوه معرفی محصولات سرد.....
۱۸۷	۶-۴ کسانی که نمی‌پرسند، نمی‌دانند.....

۱۸۹	۶-۴-۱ هنگامی که بازدیدکننده دیگر جدید نیست.....
۱۸۹	۶-۵ استفاده از قوانین انجمنی برای توصیه سریع
۱۸۹	۶-۵-۱ پیدا کردن آیتمهای جمع‌آوری شده.....
۱۹۰	۶-۵-۲ بازیابی قوانین انجمنی و مرتب‌سازی آن‌ها برحسب اعتماد.....
۱۹۲	۶-۵-۳ نمایش توصیه‌ها (recs).....
۱۹۵	۶-۵-۴ ارزیابی پیاده‌سازی.....
۱۹۵	خلاصه.....
۱۹۶	یادداشت پایانی.....

بخش دوم: الگوریتم‌های سیستم‌های توصیه‌گر

فصل هفتم: یافتن شباهت بین کاربران و بین آیتم‌ها..... ۱۹۹

۲۰۰	۷-۱ چرا شباهت؟.....
۲۰۱	۷-۱-۱ تابع تشابه چیست؟.....
۲۰۱	۷-۲ توابع اصلی شباهت.....
۲۰۳	۷-۲-۱ فاصله ژاکارد.....
۲۰۵	۷-۲-۲ اندازه‌گیری فاصله با هنجارهای Lp.....
۲۰۹	۷-۲-۳ شباهت کسینوسی.....
۲۱۱	۷-۲-۴ یافتن شباهت با ضریب همبستگی پیرسون.....
۲۱۲	۷-۲-۵ محاسبه آزمایشی شباهت پیرسون.....
۲۱۵	۷-۲-۶ تشابه همبستگی پیرسون و همبستگی کسینوسی.....
۲۱۶	۷-۳ خوشه‌بندی K-means.....
۲۱۶	۷-۳-۱ الگوریتم خوشه‌بندی k-means.....
۲۱۹	۷-۳-۲ ترجمه خوشه‌بندی k-means به Python.....
۲۲۴	۷-۴ پیاده‌سازی شباهت‌ها.....
۲۲۷	۷-۴-۱ پیاده‌سازی شباهت در سایت MovieGEEKs.....
۲۳۰	۷-۴-۲ پیاده‌سازی خوشه‌بندی در سایت MovieGEEKs.....
۲۳۴	خلاصه.....
۲۳۵	یادداشت پایانی.....

فصل هشتم: فیلتر اشتراکی / مشارکتی از نوع همسایگی..... ۲۳۷

۲۳۸	۸-۱ فیلترگذاری مشارکتی: یک درس تاریخی.....
۲۳۹	۸-۱-۱ فیلتر اطلاعات به‌روش مشارکتی.....
۲۴۰	۸-۱-۲ کمک به یکدیگر.....

۲۴۲	۸-۱-۳ ماتریس امتیازات
۲۴۳	۸-۱-۴ خط لوله فیلترگذاری اشتراکی
۲۴۴	۸-۱-۵ کدام فیلترگذاری مشارکتی ، کاربر- کاربر یا آیتم-آیتم؟
۲۴۵	۸-۱-۶ الزامات داده‌ها
۲۴۵	۸-۲ محاسبه توصیه‌ها
۲۴۵	۸-۳ محاسبه شباهت‌ها
۲۴۶	۸-۴ الگوریتم آمازون برای پیش محاسبه تشابه آیتم‌ها
۲۵۱	۸-۵ روش‌های انتخاب همسایه‌ها
۲۵۴	۸-۶ یافتن همسایگی مناسب
۲۵۴	۸-۷ روش‌های پیش‌بینی امتیازات
۲۵۷	۸-۸ پیش‌بینی با فیلتر آیتم محور
۲۵۷	۸-۸-۱ محاسبه پیش‌بینی آیتم‌ها
۲۵۸	۸-۹ مشکلات شروع سرد
۲۵۸	۸-۱۰ چند کلمه در مورد اصطلاحات یادگیری ماشین
۲۶۰	۸-۱۱ به کارگیری فیلترگذاری مشارکتی در سایت MovieGEEKs
۲۶۱	۸-۱۱-۱ فیلتر آیتم محور
۲۶۸	۸-۱۲ تفاوت بین توصیه‌های قوانین انجمنی و مشارکتی
۲۶۸	۸-۱۳ تنظیم کردن فیلترگذاری اشتراکی
۲۷۱	۸-۱۴ جوانب مثبت و منفی فیلترهای اشتراکی
۲۷۱	خلاصه
۲۷۲	یادداشت پایانی

فصل نهم: ارزیابی و آزمون سیستم‌های توصیه‌گر ۲۷۳

۲۷۴	۹-۱ پیشرفت، فروش متقابل، فروش بالا و مبادله در کسب‌وکار
۲۷۵	۹-۲ چرا ارزیابی مهم است؟
۲۷۶	۹-۳ نحوه تفسیر رفتار کاربر
۲۷۷	۹-۴ چه چیزی باید اندازه‌گیری شود
۲۷۸	۹-۴-۱ درک سلیقه من: به حداقل رساندن خطای پیش‌بینی
۲۷۹	۹-۴-۲ تنوع
۲۷۹	۹-۴-۳ پوشش
۲۸۲	۹-۴-۴ پیشنهاد غیرمترقبه
۲۸۳	۹-۵ قبل از پیاده‌سازی توصیه‌گر
۲۸۳	۹-۵-۱ بررسی الگوریتم
۲۸۴	۹-۵-۲ آزمون برازش

۲۸۵ ۹-۶ انواع ارزیابی
۲۸۶ ۹-۷ ارزیابی آفلاین
۲۸۷ ۹-۷-۱ الگوریتم بدون پیشنهاد
۲۸۷ ۹-۸ آزمایش‌های آفلاین
۲۹۴ ۹-۸-۱ آماده‌سازی داده‌ها برای آزمایش
۳۰۲ ۹-۹ اجرای آزمایشی در MovieGEEKs
۳۰۲ ۹-۹-۱ لیست کارها
۳۰۶ ۹-۱۰ ارزیابی مجموعه آزمون
۳۰۷ ۹-۱۰-۱ شروع با توصیه‌گر پایه
۳۰۹ ۹-۱۰-۲ یافتن پارامترهای مناسب
۳۱۱ ۹-۱۱ ارزیابی آنلاین
۳۱۱ ۹-۱۱-۱ آزمایش‌های کنترل‌شده
۳۱۲ ۹-۱۱-۲ آزمون A/B
۳۱۳ ۹-۱۲ آزمون مداوم با استخراج/کشف کردن
۳۱۴ ۹-۱۲-۱ حلقه‌های بازخورد
۳۱۵ خلاصه
۳۱۶ یادداشت پایانی

فصل دهم: فیلتر محتوا محور ۳۱۹

۳۲۰ ۱۰-۱ مثال توصیفی
۳۲۳ ۱۰-۲ فیلترگذاری محتوا محور
۳۲۴ ۱۰-۳ تحلیل‌کننده محتوا
۳۲۴ ۱۰-۳-۱ استخراج ویژگی برای پروفایل آیتم
۳۲۷ ۱۰-۳-۲ داده‌های طبقه‌بندی با اعداد کوچک
۳۲۸ ۱۰-۳-۳ تبدیل سال به یک ویژگی قابل مقایسه
۳۲۸ ۱۰-۴ استخراج فراداده از توضیحات
۳۳۰ ۱۰-۴-۱ آماده‌سازی توضیحات
۳۳۳ ۱۰-۵ یافتن کلمات مهم با TF-IDF
۳۳۵ ۱۰-۶ مدل‌سازی موضوع با استفاده از LDA
۳۴۲ ۱۰-۶-۱ ابزارهای تنظیم LDA
۳۴۶ ۱۰-۷ یافتن مطالب مشابه
۳۴۶ ۱۰-۸ ایجاد پروفایل کاربر
۳۴۷ ۱۰-۸-۱ ایجاد پروفایل کاربر با LDA
۳۴۷ ۱۰-۸-۲ ایجاد پروفایل کاربر با TF-IDF

۳۵۰MovieGEEKs در توصیه‌های محتوای محور در
۳۵۰۱۰-۹-۱ بارگذاری داده‌ها
۳۵۳۱۰-۹-۲ آموزش مدل
۳۵۴۱۰-۹-۳ ایجاد پروفایل آیتم
۳۵۴۱۰-۹-۴ ایجاد پروفایل کاربر
۳۵۷۱۰-۹-۵ نمایش توصیه‌ها
۳۵۸۱۰-۱۰ ارزیابی توصیه‌گر مبتنی بر محتوا
۳۵۹۱۰-۱۱ مزایا و معایب فیلترگذاری محتوای محور
۳۶۰خلاصه
۳۶۰یادداشت پایانی

فصل یازدهم: پیدا کردن ژانرهای پنهان با روش تجزیه ماتریس ۳۶۳

۳۶۴۱۱-۱ کاهش میزان داده‌ها
۳۶۶۱۱-۲ نمونه‌ای از حل مسئله
۳۶۹۱۱-۳ کمی جبر خطی
۳۶۹۱۱-۳-۱ ماتریس
۳۷۲۱۱-۳-۲ تجزیه به عوامل
۳۷۳۱۱-۴ تجزیه با استفاده از SVD
۳۸۰۱۱-۴-۱ افزودن کاربر جدید با روش تاکردن
۳۸۲۱۱-۴-۲ نحوه ارائه توصیه‌ها با SVD
۳۸۳۱۱-۴-۳ پیش‌بینی پایه
۳۸۶۱۱-۴-۴ پویایی زمانی
۳۸۷۱۱-۵ تجزیه به عوامل با استفاده از Funk SVD
۳۸۷۱۱-۵-۱ ریشه میانگین مربعات خطا (RSME)
۳۸۹۱۱-۵-۲ گرادیان کاهش
۳۹۲۱۱-۵-۳ گرادیان کاهش تصادفی
۳۹۳۱۱-۵-۴ تجزیه به عوامل
۳۹۴۱۱-۵-۵ افزودن بایاس‌ها
۳۹۵۱۱-۵-۶ نحوه شروع و زمان توقف
۳۹۹۱۱-۶ ارائه توصیه‌ها با Funk SVD
۴۰۲۱۱-۷ اجرای Funk SVD در MovieGEEKs
۴۰۷۱۱-۷-۱ داده‌های پرت
۴۰۹۱۱-۷-۲ بهنگام نگه‌داشتن مدل

۴۱۰ پیاده‌سازی سریع‌تر
۴۱۰ ۱۱-۸ داده‌های صریح در مقابل ضمنی
۴۱۰ ۱۱-۹ ارزیابی
۴۱۲ ۱۱-۱۰ اهرم‌هایی برای Funk SVD
۴۱۴ خلاصه
۴۱۴ یادداشت پایانی

فصل دوازدهم: استفاده از بهترین الگوریتم‌ها: پیاده‌سازی توصیه‌گرهای ترکیبی ۴۱۷

۴۱۸ ۱۲-۱ دنیای گیج‌کننده ترکیبی‌ها
۴۱۹ ۱۲-۲ ترکیبی یکپارچه
۴۲۰ ۱۲-۲-۱ ترکیب ویژگی‌های محتوامحور با داده‌های رفتاری برای بهبود توصیه‌های فیلترگذاری مشارکتی
۴۲۱ ۱۲-۳ توصیه‌گر ترکیبی مختلط
۴۲۲ ۱۲-۴ ترکیبی گروهی
۴۲۴ ۱۲-۴-۱ توصیه‌گر گروهی تعویضی
۴۲۵ ۱۲-۴-۲ توصیه‌گر گروهی وزن‌دار
۴۲۶ ۱۲-۴-۳ برازش خطی
۴۲۷ ۱۲-۵ پشته‌سازی خطی با ویژگی وزن‌دار (FWLS)
۴۲۸ ۱۲-۵-۱ فرا ویژگی‌ها: وزن‌ها به‌عنوان توابع
۴۳۰ ۱۲-۵-۲ الگوریتم
۴۳۷ ۱۲-۶ پیاده‌سازی
۴۴۸ خلاصه
۴۴۸ یادداشت‌های پایانی

فصل سیزدهم: رتبه‌بندی و یادگیری رتبه‌بندی ۴۴۹

۴۵۰ ۱۳-۱ یادگیری رتبه‌بندی با مثالی از Foursquare
۴۵۵ ۱۳-۲ رتبه‌بندی مجدد
۴۵۵ ۱۳-۳ یادگیری از رتبه‌بندی مجدد
۴۵۶ ۱۳-۳-۱ سه نوع الگوریتم LTR
۴۵۸ ۱۳-۴ رتبه‌بندی شخصی بیزی
۴۶۱ ۱۳-۴-۱ رتبه‌بندی با BPR
۴۶۳ ۱۳-۴-۲ جادوی ریاضی (جادوگری پیشرفته)
۴۶۷ ۱۳-۴-۳ الگوریتم BPR

۴۶۷	۱۳-۴-۴ BPR با تجزیه ماتریس
۴۶۸	۱۳-۵ BPR پیاده‌سازی
۴۷۴	۱۳-۵-۱ توصیه‌ها
۴۷۶	۱۳-۶ ارزیابی
۴۷۸	۱۳-۷ اهرم‌های BPR
۴۸۰	خلاصه
۴۸۰	یادداشت‌های پایانی

فصل چهاردهم: آینده سیستم‌های توصیه‌گر ۴۸۳

۴۸۴	۱۴-۱ این کتاب در چند جمله
۴۸۷	۱۴-۲ موضوعات مورد مطالعه بعدی
۴۸۷	۱۴-۲-۱ مطالعه بیشتر
۴۸۸	۱۴-۲-۲ الگوریتم‌ها
۴۸۹	۱۴-۲-۳ زمینه
۴۸۹	۱۴-۲-۴ تعامل انسان و کامپیوتر
۴۹۰	۱۴-۲-۵ انتخاب یک معماری خوب
۴۹۱	۱۴-۳ آینده سیستم‌های توصیه‌گر
۴۹۶	۱۴-۴ حرف‌های نهایی
۴۹۷	یادداشت‌های پایانی

پیشگفتار

وقتی دانشگاه را در سال ۲۰۰۳ به پایان رساندم، این تهدید وجود داشت که در اروپا به هیچ دانشمند کامپیوتری نیاز نخواهد بود، زیرا همه چیز در کشورهایی که دارای دستمزدهای بسیار پایین تری هستند توسعه می یابد. خدا را شکر که به دلایل زیادی این پیش بینی هرگز محقق نشد. من فکر می کنم یکی از معضلاتی^۱ که شرکتها دست کم می گرفتند عدم شناخت کافی توسعه دهندگان نرم افزارها با فرهنگ جاهایی بود که می خواستند آن ها را پیاده سازی کنند. درخواست های پیاده سازی نرم افزارها انجام شدند، اما عملکرد وظیفه ای^۲ آن ها با آنچه مشتریان انتظار داشتند متفاوت بود.

امروز، خطر مشابهی برای افراد علاقه مند به یادگیری ماشین و علوم داده وجود دارد. اما اکنون تهدید دستمزد پایین نیست، بلکه نرم افزار به منزله خدمت (SaaS) است که در آن داده ها را آپلود می کنید و سپس سیستم کار را برای شما انجام می دهد.

من هم مانند هر کس دیگری نگران این موضوع هستم که ماشین ها حوزه ها و افراد را درک نمی کنند. ماشین ها هنوز آن قدر هوشمند نیستند که بتوانند بدون دخالت انسان ها کار کنند. همه چیز به سرعت پیش می رود، اما من ادعا می کنم هر کسی که این کتاب را بخواند می تواند تا پایان عمر کاری خود با توصیه گرها کار کند.

کجا وارد کار شدم؟ من به عنوان یک مهندس نرم افزار در ایتالیا کار می کردم و به انگلستان نقل مکان کردم و به شغلی نیاز داشتم که بیشتر از انجام عملیات CRUD^۳ در پایگاه داده نیاز به تفکر داشت. خوشبختانه، یک استخدام کننده بزرگ از RedRock Consulting Ltd با من تماس گرفت. آن ها من را به یک ارائه دهنده سیستم توصیه گر معرفی کردند و من روی موتور سیستم کار می کردم. من در یادگیری ماشین گم شده بودم («گم شده» به معنای واقعاً علاقه مند و درگیر). علاوه بر کار روی سیستم های توصیه گر، من به جست و جوی دانش در اینترنت پرداختم و کتاب های بی شماری در این زمینه و موضوعات مرتبط مطالعه کردم.

امروز نمی توانید یک مطلب عنوان کنید بدون اینکه حداقل ۱۰ نفر سعی کنند چیزی در مورد یادگیری ماشین به شما یاد دهند. وقتی در اینترنت آموزش های یک صفحه ای یا یک ساعته ای را می بینم که ادعا می کنند

1- Issues

2- Functionality

۳- CRUD به چهار عملیات اساسی اشاره دارد که یک نرم افزار باید قادر به انجام آن باشد- ایجاد، خواندن، بهنگام رسانی و حذف.

همه چیز را در مورد یادگیری ماشین به شما آموزش می‌دهند، برایم حیرت‌آور است. من هم می‌توانم یک آموزش مؤثر مشابه در مورد چگونگی خلبان شدن برای یک جنگنده ایجاد کنم: شما با استفاده از یک اهرم بلند می‌شوید و پرواز می‌کنید. اگر نیاز به تیراندازی دارید، یک دکمه را فشار دهید. سپس، قبل از تمام شدن بنزین فرود می‌آیید.

یک آموزش خلبانی جنگنده مانند این احتمالاً برای شروع عالی خواهد بود- من هم از اینجا شروع کردم. اما خودتان را گول نزنید: درگ یادگیری ماشین پیچیده است. عامل انسانی را نیز به آن اضافه کنید که همیشه همه چیز را کمی لرزان‌تر می‌کند.

برگردیم به داستان من، با توصیه‌گرها کار کردم و از این بابت خوشحال بودم و بعد شغلم را تغییر دادم. در شغل جدیدم، قرار بود به کار روی سیستم‌های توصیه‌گر ادامه دهم، اما آن پروژه به تعویق افتاد. در آن زمان از این‌که دیگر با توصیه‌گرها کار نخواهم کرد، عصبی بودم اما در همان زمان انتشارات Manning به من این فرصت را داد که کتابی درباره سیستم‌های توصیه‌گر بنویسم. چه کاری می‌توانستم بکنم، غیر از شرع کار؟ بلافاصله پس از امضای قرارداد، پروژه توصیه‌گر شروع شد. نوشتن این کتاب یک تجربه یادگیری عالی بود و امیدوارم شما هم از آن بهره و لذت ببرید.

آرمان^۱ این کتاب این است که شما را با سیستم‌های توصیه‌گر آشنا کند- نه تنها الگوریتم‌ها، بلکه اکوسیستم سیستم توصیه‌گر. الگوریتم‌ها خیلی پیچیده نیستند، اما برای درک و اجرای آن‌ها نیاز به شناخت کاربرانی است که باید توصیه‌ها را دریافت کنند. محتوای کتاب در طول نگارش تکامل یافته است، چون سعی کردم مطالب بیشتری در آن ارائه دهم. امیدوارم خواندن این کتاب همه مواردی را که برای شروع به کار با توصیه‌گرها نیاز دارید فراهم کند و پایه محکمی برای ساختن سیستم‌های توصیه‌گر به شما بدهد.