

## پیش‌گفتار

- درباره کتاب ..... ۲۱
- این کتاب چگونه سازمان‌دهی شده؛ نقشه راه ..... ۲۱

## فصل اول: یادگیری ماشین چیست؟

- ۱-۱. آیا برای یادگیری ماشین نیازی به دانش عمیق ریاضی و برنامه‌نویسی دارم؟ ..... ۲۵
- ۱-۱-۱. وقتی به مفهوم فرمول‌ها و کدها بیندیشیم همه چیز واضح‌تر خواهد بود ..... ۲۵
- ۲-۱. یادگیری ماشین دقیقاً چیست؟ ..... ۲۶
- ۱-۲-۱. هوش مصنوعی چیست؟ ..... ۲۶
- ۲-۲-۱. یادگیری ماشین چیست؟ ..... ۲۷
- ۳-۲-۱. یادگیری عمیق چیست؟ ..... ۲۸
- ۳-۱. چگونه می‌توانیم کاری کنیم ماشین با داده‌ها تصمیم‌گیری کند؟ ..... ۲۹
- ۱-۳-۱. انسان‌ها چگونه فکر می‌کنند؟ ..... ۳۰
- ۲-۳-۱. چند اصطلاح یادگیری ماشین؛ مدل و الگوریتم ..... ۳۰
- ۳-۳-۱. چند مثال از مدل‌هایی که انسان‌ها استفاده می‌کنند ..... ۳۱
- ۴-۳-۱. چند مدل نمونه که کامپیوترها استفاده می‌کنند ..... ۳۶
- ۴-۱. خلاصه ..... ۳۷

## فصل دوم: انواع یادگیری ماشین

- ۱-۲. تفاوت داده‌های برچسب‌دار و بدون برچسب چیست؟ ..... ۴۱
- ۱-۱-۲. داده چیست؟ ..... ۴۱
- ۲-۱-۲. ویژگی چیست؟ ..... ۴۲
- ۳-۱-۲. برچسب چیست؟ ..... ۴۲
- ۴-۱-۲. پیش‌بینی‌ها ..... ۴۲
- ۵-۱-۲. داده‌های برچسب‌دار و بدون برچسب ..... ۴۲
- ۲-۲. یادگیری ماشین نظارت‌شده؛ شاخه‌ای از یادگیری ماشین ..... ۴۳
- ۱-۲-۲. مدل‌های رگرسیون، اعداد را پیش‌بینی می‌کنند ..... ۴۵
- ۲-۲-۲. مدل‌های دسته‌بندی، یک وضعیت را پیش‌بینی می‌کنند ..... ۴۶
- ۳-۲. یادگیری بی‌نظارت؛ شاخه‌ای از یادگیری ماشین ..... ۴۷

- ۲-۳-۱. الگوریتم‌های خوشه‌بندی، داده‌ها را براساس شباهت گروه‌بندی می‌کنند. ۴۸
- ۲-۳-۲. کاهش ابعاد، داده‌ها را بدون از دست رفتن حجم زیادی اطلاعات ... ۵۱
- ۲-۳-۳. سایر روش‌های ساده‌سازی داده‌ها؛ فاکتورگیری ماتریس‌ها و تجزیه ... ۵۴
- ۲-۳-۴. یادگیری ماشین مولد ... ۵۵
- ۲-۴. یادگیری تقویتی چیست؟ ... ۵۶
- ۲-۵. خلاصه ... ۵۸

### فصل سوم: رسم یک خط نزدیک نقاطمان: رگرسیون خطی

- ۳-۱-۱. مسئله؛ می‌خواهیم قیمت یک خانه را پیش‌بینی کنیم. ۶۴
- ۳-۲-۱. راه‌حل؛ ساخت یک مدل رگرسیون برای قیمت خانه‌ها. ۶۴
- ۳-۲-۲. مرحله یادآوری؛ بررسی قیمت خانه‌های موجود. ۶۵
- ۳-۲-۳. مرحله فرمول‌بندی؛ ساختن قانونی برای تخمین قیمت خانه‌ها. ۶۶
- ۳-۲-۳. مرحله پیش‌بینی؛ وقتی خانه جدیدی در بازار به فروش گذاشته می‌شود. ۶۸
- ۳-۲-۴. رگرسیون خطی چندگانه؛ اگر متغیرهای بیشتری داشته باشیم چه؟ ۶۹
- ۳-۲-۵. بعضی از سوال‌هایی که ممکن است داشته باشید و پاسخ سریع آنها. ۷۰
- ۳-۳-۱. چگونه می‌توانیم از کامپیوتر بخواهیم این خط را برایمان رسم کند: الگوریتم ... ۷۱
- ۳-۳-۱. مقدمه‌ای بر شیب خط و عرض از مبدأ. ۷۴
- ۳-۳-۲. ترفندی ساده برای نزدیک‌تر کردن خط به مجموعه‌ای از نقاط به صورت نقطه به نقطه. ۷۶
- ۳-۳-۳. ترفند مربع؛ روشی هوشمندانه‌تر برای نزدیک کردن خط به یکی از نقاط. ۷۹
- ۳-۳-۴. ترفند قدرمطلق؛ روش مفید دیگری برای انتقال خط به نزدیکی نقاط. ۸۲
- ۳-۳-۵. الگوریتم رگرسیون خطی؛ تکرار ترفند مربع یا قدرمطلق برای نزدیک کردن خط به نقاط ... ۸۳
- ۳-۳-۶. بارگذاری داده‌ها و رسم آنها. ۸۴
- ۳-۳-۷. استفاده از الگوریتم رگرسیون خطی با مجموعه داده‌ها. ۸۵
- ۳-۳-۸. استفاده از مدل برای پیش‌بینی. ۸۶
- ۳-۳-۹. الگوریتم رگرسیون خطی عمومی (مطالعه این قسمت اختیاری است). ۸۶
- ۳-۴-۱. چگونه کیفیت نتایج مدل را می‌سنجیم؟ تابع خطا. ۸۸
- ۳-۴-۱. خطای مطلق؛ سنج‌های که با محاسبه مجموع قدرمطلق فاصله‌ها، کیفیت ... ۸۸
- ۳-۴-۲. خطای مربعات؛ سنج‌های که با محاسبه مجموع مربع فاصله نقاط ... ۸۹
- ۳-۴-۳. خطای میانگین مربعات، ریشه دوم آن و خطای میانگین مطلق در ... ۹۰
- ۳-۴-۴. گرادیان کاهشی؛ روشی برای کمینه کردن تابع خطا با پایین آمدن تدریجی از یک کوه. ۹۱
- ۳-۴-۵. رسم تابع خطا و تشخیص نقطه توقف. ۹۴
- ۳-۴-۶. آیا از یک نقطه برای آموزش استفاده می‌کنیم یا چند نقطه؟ گرادیان کاهشی ... ۹۵
- ۳-۵. کاربردهای دنیای واقعی؛ استفاده از کتابخانه Turi Create برای پیش‌بینی ... ۹۵

|        |                                                                     |     |
|--------|---------------------------------------------------------------------|-----|
| ۳-۶.   | اگر نتوان داده‌ها را با یک خط توصیف کرد چه؟ رگرسیون چند جمله‌ای     | ۹۸  |
| ۳-۶-۱. | نوع خاصی از توابع منحنی؛ چند جمله‌ای‌ها                             | ۹۸  |
| ۳-۶-۲. | داده‌های تان غیرخطی است؟ مشکلی نیست؛ بیایید یک چند جمله‌ای برای ... | ۹۹  |
| ۳-۷.   | پارامترها و هایپرپارامترها                                          | ۱۰۰ |
| ۳-۸.   | کاربردهای رگرسیون                                                   | ۱۰۰ |
| ۳-۸-۱. | سیستم‌های توصیه‌گر                                                  | ۱۰۱ |
| ۳-۸-۲. | توصیه ویدئو و موسیقی                                                | ۱۰۱ |
| ۳-۸-۳. | توصیه محصولات                                                       | ۱۰۱ |
| ۳-۸-۴. | خدمات درمانی                                                        | ۱۰۱ |
| ۳-۹.   | خلاصه                                                               | ۱۰۲ |

### فصل چهارم: بهینه‌سازی فرایند آموزش: کم‌برازش، بیش‌برازش، تست و نرمال‌سازی

|        |                                                                          |     |
|--------|--------------------------------------------------------------------------|-----|
| ۴-۱.   | مثالی از کم‌برازش و بیش‌برازش با رگرسیون چند جمله‌ای                     | ۱۰۸ |
| ۴-۲.   | چگونه می‌توانیم کاری کنیم کامپیوتر مدل درست را انتخاب کند؟ تست کردن      | ۱۱۰ |
| ۴-۲-۱. | چگونه مجموعه تست را انتخاب می‌کنیم و اندازه این مجموعه چقدر است؟         | ۱۱۳ |
| ۴-۲-۲. | آیا می‌توانیم از داده‌های تستی برای آموزش مدل استفاده کنیم؟ خیر          | ۱۱۳ |
| ۴-۳.   | کجا قانون طلایی را شکستیم و چگونه می‌توانیم آن را اصلاح کنیم؟ مجموعه ... | ۱۱۴ |
| ۴-۴.   | یک روش عددی برای تصمیم‌گیری درباره پیچیدگی مدل‌ها؛ نمودار پیچیدگی مدل    | ۱۱۴ |
| ۴-۵.   | روشی دیگر برای پیشگیری از بیش‌برازش؛ نرمال‌سازی                          | ۱۱۶ |
| ۴-۵-۱. | مثالی دیگر از بیش‌برازش؛ توصیه فیلم                                      | ۱۱۸ |
| ۴-۵-۲. | اندازه‌گیری پیچیدگی مدل؛ نرم‌های $L1$ و $L2$                             | ۱۱۹ |
| ۴-۵-۳. | تغییر تابع خطا برای حل مشکل مان؛ رگرسیون لسو و ریج                       | ۱۲۱ |
| ۴-۵-۴. | تعدیل عملکرد و پیچیدگی مدل؛ پارامتر نرمال‌سازی                           | ۱۲۲ |
| ۴-۵-۵. | تاثیر نرمال‌سازی $L1$ و $L2$ بر ضرایب مدل                                | ۱۲۲ |
| ۴-۵-۶. | یک روش شهودی برای درک نرمال‌سازی                                         | ۱۲۴ |
| ۴-۶.   | رگرسیون چند جمله‌ای، تست و نرمال‌سازی با کتابخانه Turi Create            | ۱۲۶ |
| ۴-۷.   | خلاصه                                                                    | ۱۳۱ |

### فصل پنجم: استفاده از خط‌ها برای دسته‌بندی نقاط؛ الگوریتم پرسپترون

|        |                                                               |     |
|--------|---------------------------------------------------------------|-----|
| ۵-۱.   | مسئله؛ ما در یک سیاره بیگانه هستیم و زبان آنها را نمی‌دانیم   | ۱۳۷ |
| ۵-۱-۱. | یک سیاره پیچیده‌تر                                            | ۱۴۱ |
| ۵-۱-۲. | آیا دسته‌بندی‌کننده مان باید همیشه درست عمل کند؟              | ۱۴۴ |
| ۵-۱-۳. | یک دسته‌بندی‌کننده کلی‌تر و یک روش متفاوت‌تر برای تعریف خط‌ها | ۱۴۵ |

- ۵-۱-۴. تابع پله‌ای و توابع فعال‌سازی؛ روشی مختصر و مفید برای پیش‌بینی ..... ۱۴۶
- ۵-۱-۵. اگر بیش از دو کلمه داشته باشیم چه اتفاقی می‌افتد؟ تعریف عمومی ..... ۱۴۷
- ۵-۱-۶. بایاس، عرض از مبدأ و خلق یک آدم فضایی ساکت ..... ۱۵۰
- ۵-۲. چگونه می‌فهمیم یک دسته‌بندی‌کننده خوب است یا بد؟ تابع خطا ..... ۱۵۱
- ۵-۲-۱. چگونه می‌توانیم دسته‌بندی‌کننده‌ها را مقایسه کنیم؟ تابع خطا ..... ۱۵۲
- ۵-۳. چگونه می‌توان یک دسته‌بندی‌کننده خوب پیدا کرد؟ الگوریتم پرسپترون ..... ۱۵۹
- ۵-۳-۱. ترفند پرسپترون؛ روشی برای ایجاد یک بهبود کوچک در پرسپترون ..... ۱۶۰
- ۵-۳-۲. تکرار ترفند پرسپترون؛ الگوریتم پرسپترون ..... ۱۶۵
- ۵-۳-۳. گرادیان کاهشی ..... ۱۶۷
- ۵-۳-۴. گرادیان کاهشی تصادفی و گرادیان کاهشی دسته‌ای ..... ۱۶۸
- ۵-۴. پیاده‌سازی الگوریتم پرسپترون ..... ۱۶۹
- ۵-۴-۱. پیاده‌سازی ترفند پرسپترون ..... ۱۷۰
- ۵-۴-۲. پیاده‌سازی الگوریتم پرسپترون ..... ۱۷۲
- ۵-۴-۳. پیاده‌سازی الگوریتم پرسپترون با کتابخانه Turi Create ..... ۱۷۳
- ۵-۵. کاربردهای الگوریتم پرسپترون ..... ۱۷۴
- ۵-۵-۱. جدا کردن ایمیل‌های هم و اسپم ..... ۱۷۴
- ۵-۵-۲. سیستم‌های توصیه‌گر ..... ۱۷۵
- ۵-۵-۳. خدمات درمانی ..... ۱۷۵
- ۵-۵-۴. بینایی ماشین ..... ۱۷۶
- ۵-۶. خلاصه ..... ۱۷۶

### فصل ششم: رویکردی پیوسته برای جدا کردن نقاط: دسته‌بندی‌کننده‌های لجستیک

- ۶-۱. دسته‌بندی‌کننده لجستیک؛ نسخه پیوسته پرسپترون ..... ۱۸۱
- ۶-۱-۱. رویکردی احتمالی به دسته‌بندی؛ تابع سیگموئید ..... ۱۸۳
- ۶-۱-۲. مجموعه داده‌ها و پیش‌بینی‌ها ..... ۱۸۵
- ۶-۱-۳. توابع خطا؛ خطای مطلق، مربعات و تابع زیان لگاریتمی ..... ۱۸۶
- ۶-۱-۴. مقایسه دسته‌بندی‌کننده‌ها با زیان لگاریتمی ..... ۱۹۱
- ۶-۲. چگونه می‌توان یک دسته‌بندی‌کننده لجستیک خوب پیدا کرد؟ الگوریتم ..... ۱۹۳
- ۶-۲-۱. ترفند لجستیک؛ روشی برای بهبود جزئی پرسپترون پیوسته ..... ۱۹۳
- ۶-۲-۲. تکرار ترفند لجستیک؛ الگوریتم رگرسیون لجستیک ..... ۱۹۸
- ۶-۲-۳. گرادیان کاهشی تصادفی، گرادیان کاهشی دسته‌ای و گرادیان کاهشی ..... ۱۹۹
- ۶-۳. پیاده‌سازی الگوریتم رگرسیون لجستیک ..... ۲۰۰
- ۶-۳-۱. چگونه خودمان الگوریتم رگرسیون لجستیک را پیاده‌سازی کنیم؟ ..... ۲۰۱

- ۴-۶. کاربردهای دنیای واقعی؛ دسته‌بندی نقد فیلم‌های IMDB با Turi Create ..... ۲۰۴
- ۵-۶. دسته‌بندی به چند کلاس؛ تابع سافت مکس ..... ۲۰۷
- ۶-۶. خلاصه ..... ۲۰۸

### فصل هفتم: چگونه عملکرد مدل‌های دسته‌بندی را اندازه‌گیری می‌کنیم؟ صحت و مفاهیم مرتبط

- ۱-۷. صحت؛ پیش‌بینی‌های مدل‌مان چقدر درست است؟ ..... ۲۱۳
- ۱-۱-۷. دو مدل نمونه؛ ایمیل‌های اسپم و ویروس کرونا ..... ۲۱۳
- ۲-۱-۷. یک مدل بسیار موثر و به‌شدت به‌دردنخورا ..... ۲۱۴
- ۲-۷. چگونه می‌توان مسئله صحت را حل کرد؟ تعریف انواع مختلف خطا و ..... ۲۱۴
- ۱-۲-۷. مثبت کاذب و منفی کاذب؛ کدام یک بدتر است؟ ..... ۲۱۵
- ۲-۲-۷. نوشتن نقاط درست و اشتباه دسته‌بندی‌شده در یک جدول؛ ماتریس ..... ۲۱۷
- ۳-۲-۷. بازیابی؛ از میان مثال‌های مثبت چند مورد را درست دسته‌بندی کرده‌ایم؟ ..... ۲۱۸
- ۴-۲-۷. دقت؛ از میان مثال‌هایی که مثبت دسته‌بندی کردیم، چه تعداد درست ..... ۲۲۰
- ۵-۲-۷. ترکیب بازیابی و دقت برای بهینه‌سازی هر دوی آنها؛ امتیاز F ..... ۲۲۲
- ۶-۲-۷. بین دقت، بازیابی و امتیازهای F کدام را باید انتخاب کنیم؟ ..... ۲۲۴
- ۳-۷. ابزاری کاربردی برای ارزیابی مدل‌ها؛ منحنی ROC ..... ۲۲۴
- ۱-۳-۷. حساسیت و تشخیص؛ دو معیار جدید برای ارزیابی مدل‌ها ..... ۲۲۵
- ۲-۳-۷. منحنی ROC؛ روشی برای بهینه‌سازی حساسیت و تشخیص مدل ..... ۲۲۷
- ۳-۳-۷. معیاری که کیفیت مدل‌مان را مشخص می‌کند؛ AUC ..... ۲۳۰
- ۴-۳-۷. چگونه به کمک منحنی ROC تصمیم‌گیری کنیم ..... ۲۳۱
- ۵-۳-۷. بازیابی همان حساسیت است، اما دقت و تشخیص تفاوت دارند ..... ۲۳۴
- ۴-۷. خلاصه ..... ۲۳۵

### فصل هشتم: حداکثر استفاده از احتمال‌ها؛ مدل بیز ساده

- ۱-۸. بیمار یا سالم؟ داستانی که قهرمانش قضیه بیز است ..... ۲۴۱
- ۱-۱-۸. مقدمه‌ای بر قضیه بیز؛ احتمال پیشین، رویداد و احتمال پسین ..... ۲۴۵
- ۲-۸. مطالعه موردی؛ مدل تشخیص ایمیل‌های اسپم ..... ۲۴۶
- ۱-۲-۸. پیدا کردن احتمال پیشین؛ احتمال اسپم بودن یک ایمیل چقدر است؟ ..... ۲۴۶
- ۲-۲-۸. پیدا کردن احتمال پسین؛ احتمال اسپم بودن ایمیل به شرط وجود کلمه‌ای خاص ..... ۲۴۷
- ۳-۲-۸. چه اتفاق ریاضی‌ای افتاد؟ تبدیل نسبت‌ها به احتمال‌ها ..... ۲۴۹
- ۴-۲-۸. اگر دو کلمه داشته باشیم چه؟ الگوریتم بیز ساده ..... ۲۵۴
- ۵-۲-۸. اگر بیش از دو کلمه داشته باشیم چه باید کرد؟ ..... ۲۵۸
- ۳-۸. ساخت یک مدل تشخیص ایمیل‌های اسپم با داده‌های واقعی ..... ۲۶۰

- ۸-۳-۱. پردازش داده‌ها ..... ۲۶۰
- ۸-۳-۲. پیدا کردن احتمال‌های پیشین ..... ۲۶۱
- ۸-۳-۳. محاسبه احتمال‌های پسین با قضیه بیز ..... ۲۶۱
- ۸-۳-۴. پیاده‌سازی الگوریتم بیز ساده ..... ۲۶۲
- ۸-۳-۵. کارهای آتی ..... ۲۶۴
- ۸-۴. خلاصه ..... ۲۶۴

### فصل نهم: تقسیم کردن داده‌ها با پرسیدن سوال: درخت تصمیم‌گیری

- ۹-۱. مسئله: می‌خواهیم براساس دانلودهای کاربر به او برنامه پیشنهاد دهیم ..... ۲۷۵
- ۹-۲. راه حل: ساختن یک سیستم توصیه‌گر برنامه ..... ۲۷۶
- ۹-۲-۱. قدم اول ساخت مدل: پرسیدن بهترین سوال ..... ۲۷۶
- ۹-۲-۲. قدم دوم ساخت مدل: تکرار ..... ۲۸۸
- ۹-۲-۳. قدم آخر: چه زمانی ساخت درخت را متوقف کنیم و سایر هایپرپارامترها ..... ۲۹۰
- ۹-۲-۴. الگوریتم درخت تصمیم‌گیری: چگونه یک درخت تصمیم‌گیری بسازیم و ..... ۲۹۰
- ۹-۳. فراتر از سوال‌های بله / خیر ..... ۲۹۲
- ۹-۳-۱. تقسیم داده‌ها برای ویژگی‌های گسسته که بیش از دو دسته دارند ..... ۲۹۲
- ۹-۳-۲. تقسیم داده‌ها با ویژگی‌های پیوسته مثل سن ..... ۲۹۳
- ۹-۴. مرز بصری درخت‌های تصمیم‌گیری ..... ۲۹۶
- ۹-۴-۱. استفاده از Scikit Learn برای ساخت درخت تصمیم‌گیری ..... ۲۹۸
- ۹-۵. کاربرد دنیای واقعی: مدل‌سازی پذیرش تحصیلی دانشجویان با Scikit-Learn ..... ۲۹۹
- ۹-۵-۱. مقداردهی هایپرپارامترها در Scikit-Learn ..... ۳۰۱
- ۹-۶. درخت‌های تصمیم‌گیری برای رگرسیون ..... ۳۰۳
- ۹-۷. کاربردها ..... ۳۰۷
- ۹-۷-۱. درخت‌های تصمیم‌گیری در خدمات درمانی به صورت گسترده‌ای به کار می‌روند ..... ۳۰۸
- ۹-۷-۲. درخت‌های تصمیم‌گیری در سیستم‌های توصیه‌گر کاربرد گسترده‌ای دارند ..... ۳۰۸
- ۹-۸. خلاصه ..... ۳۰۸

### فصل دهم: شبکه‌های عصبی

- ۱۰-۱. شبکه‌های عصبی با یک مثال: یک سیاره بیگانه پیچیده‌تر ..... ۳۱۴
- ۱۰-۱-۱. راه حل: اگر یک خط کفایت نمی‌کند، از دو خط برای دسته‌بندی داده‌ها استفاده کنید ..... ۳۱۵
- ۱۰-۱-۲. چرا باید از دو خط استفاده کنیم؟ آیا شادی غیرخطی است؟ ..... ۳۱۷
- ۱۰-۱-۳. ترکیب خروجی پرسپترون‌ها به یک پرسپترون دیگر ..... ۳۱۸
- ۱۰-۱-۴. نمایش گرافیکی پرسپترون‌ها ..... ۳۲۰

|     |                                                                    |
|-----|--------------------------------------------------------------------|
| ۳۲۲ | ۵-۱-۱۰. نمایش گرافیکی شبکه‌های عصبی                                |
| ۳۲۳ | ۶-۱-۱۰. مرز تصمیم‌گیری یک شبکه عصبی                                |
| ۳۲۵ | ۷-۱-۱۰. معماری کلی یک شبکه عصبی کاملاً متصل                        |
| ۳۲۷ | ۲-۱۰. آموزش شبکه‌های عصبی                                          |
| ۳۲۸ | ۱-۲-۱۰. تابع خطا؛ روشی برای اندازه‌گیری عملکرد شبکه عصبی           |
| ۳۲۸ | ۲-۲-۱۰. انتشار معکوس؛ مرحله کلیدی آموزش شبکه‌های عصبی              |
| ۳۲۹ | ۳-۲-۱۰. مشکلات بالقوه؛ از بیش‌برازش تا گرادینان محوشونده           |
| ۳۳۰ | ۴-۲-۱۰. تکنیک‌های آموزش شبکه‌های عصبی؛ نرمال‌سازی و حذف تصادفی     |
| ۳۳۲ | ۵-۲-۱۰. توابع فعال‌سازی مختلف؛ tanh و ReLU                         |
| ۳۳۳ | ۶-۲-۱۰. شبکه‌های عصبی با بیش از یک خروجی؛ تابع سافت‌مکس            |
| ۳۳۴ | ۷-۲-۱۰. هایپرپارامترها                                             |
| ۳۳۵ | ۳-۱۰. پیاده‌سازی شبکه‌های عصبی با کتابخانه Keras                   |
| ۳۳۵ | ۱-۳-۱۰. یک مثال گرافیکی دوبعدی                                     |
| ۳۴۰ | ۲-۳-۱۰. آموزش یک شبکه عصبی برای بازشناسی تصویرها                   |
| ۳۴۳ | ۴-۱۰. به کار بستن شبکه‌های عصبی برای رگرسیون                       |
| ۳۴۵ | ۵-۱۰. معماری‌های دیگر برای داده‌های پیچیده‌تر                      |
| ۳۴۵ | ۱-۵-۱۰. شبکه‌های عصبی چگونه می‌بینند؛ شبکه‌های عصبی هم‌گشتی        |
| ۳۴۶ | ۲-۵-۱۰. شبکه‌های عصبی چگونه صحبت می‌کنند؛ RNN‌ها، GRU‌ها و LSTM‌ها |
| ۳۴۷ | ۳-۵-۱۰. شبکه‌های عصبی چطور نقاشی می‌کشند؛ GAN‌ها                   |
| ۳۴۷ | ۶-۱۰. خلاصه                                                        |

## فصل یازدهم: یافتن مرزها با سبک: SVM‌ها و روش کرنل

|     |                                                                                    |
|-----|------------------------------------------------------------------------------------|
| ۳۵۴ | ۱-۱۱. استفاده از یک تابع خطای جدید برای ساخت دسته‌بندی‌کننده‌های بهتر              |
| ۳۵۵ | ۱-۱-۱۱. تابع خطای دسته‌بندی؛ تلاش برای دسته‌بندی صحیح نقاط                         |
| ۳۵۷ | ۲-۱-۱۱. تابع خطای فاصله؛ دور کردن دو خط از هم تا حد ممکن                           |
| ۳۵۸ | ۳-۱-۱۱. جمع دو تابع خطا برای محاسبه خطای کل                                        |
| ۳۵۹ | ۴-۱-۱۱. می‌خواهیم تمرکز دسته‌بندی‌کننده بر فاصله دو خط باشد یا دسته‌بندی درست؟ ... |
| ۳۶۱ | ۲-۱۱. پیاده‌سازی SVM‌ها در Scikit-Learn                                            |
| ۳۶۱ | ۱-۲-۱۱. پیاده‌سازی یک SVM ساده                                                     |
| ۳۶۲ | ۲-۲-۱۱. پارامتر C                                                                  |
| ۳۶۳ | ۳-۱۱. آموزش SVM‌هایی با مرزهای غیرخطی؛ روش کرنل                                    |
| ۳۶۴ | ۱-۳-۱۱. استفاده از معادلات چند جمله‌ای؛ کرنل چند جمله‌ای                           |
| ۳۷۰ | ۲-۳-۱۱. استفاده از برآمدگی‌ها در ابعاد بالاتر به نفع ما؛ کرنل RBF                  |

|     |                               |
|-----|-------------------------------|
| ۳۷۵ | ۱۱-۳-۳. آموزش SVM با کرنل RBF |
| ۳۸۱ | ۱۱-۳-۴. پیاده سازی روش کرنل   |
| ۳۸۴ | ۱۱-۴. خلاصه                   |

## فصل دوازدهم: ترکیب مدل ها برای به دست آوردن بهترین نتیجه: یادگیری گروهی

|     |                                                                                      |
|-----|--------------------------------------------------------------------------------------|
| ۳۸۹ | ۱۲-۱. با اندکی کمک از دوستانمان                                                      |
| ۳۹۱ | ۱۲-۲. روش بگینگ: ترکیب تصادفی یادگیرنده های ضعیف برای ساختن یک یادگیرنده قوی         |
| ۳۹۳ | ۱۲-۲-۱. تطبیق یک جنگل تصادفی به صورت دستی                                            |
| ۳۹۵ | ۱۲-۲-۲. آموزش جنگل تصادفی در Scikit-Learn                                            |
| ۳۹۷ | ۱۲-۳. AdaBoost: ساخت یادگیرنده قوی با ترکیب زیرکانه یادگیرنده های ضعیف               |
| ۳۹۸ | ۱۲-۳-۱. تصویر کلی AdaBoost: ساختن یادگیرنده های ضعیف                                 |
| ۴۰۰ | ۱۲-۳-۲. ترکیب یادگیرنده های ضعیف برای ساختن یادگیرنده قوی                            |
| ۴۰۴ | ۱۲-۳-۳. پیاده سازی AdaBoost در Scikit-Learn                                          |
| ۴۰۶ | ۱۲-۴. گرادیان بوستینگ: استفاده از درخت های تصمیم گیری برای ساخت یادگیرنده های قوی تر |
| ۴۱۱ | ۱۲-۵. XGBoost: روشی افراطی برای گرادیان بوستینگ                                      |
| ۴۱۱ | ۱۲-۵-۱. معیار شباهت XGBoost: راهی جدید و موثر برای اندازه گیری شباهت در یک مجموعه    |
| ۴۱۳ | ۱۲-۵-۲. ساخت یادگیرنده های ضعیف                                                      |
| ۴۱۶ | ۱۲-۵-۳. هرس کردن درخت: روشی برای کاهش بیش برآزش با ساده کردن                         |
| ۴۱۷ | ۱۲-۵-۴. ایجاد پیش بینی                                                               |
| ۴۱۹ | ۱۲-۵-۵. آموزش یک مدل XGBoost در پایتون                                               |
| ۴۲۱ | ۱۲-۶. کاربرد مدل های یادگیری گروهی                                                   |
| ۴۲۱ | ۱۲-۷. خلاصه                                                                          |

## فصل سیزدهم: جمع بندی: یک مثال دنیای واقعی از مهندسی داده ها و یادگیری ماشین

|     |                                                                        |
|-----|------------------------------------------------------------------------|
| ۴۲۵ | ۱۳-۱. مجموعه داده های تایپانیک                                         |
| ۴۲۵ | ۱۳-۱-۱. ویژگی های مجموعه داده ها                                       |
| ۴۲۶ | ۱۳-۱-۲. استفاده از Pandas برای بارگذاری داده ها                        |
| ۴۲۸ | ۱۳-۱-۳. استفاده از Pandas برای مطالعه و بررسی مجموعه داده ها           |
| ۴۲۹ | ۱۳-۲. پاک سازی داده ها: داده های گم شده و چگونگی مدیریت آنها           |
| ۴۳۰ | ۱۳-۲-۱. حذف ستون هایی که داده های گم شده دارند                         |
| ۴۳۱ | ۱۳-۲-۲. چگونه کل ستون را از دست ندهیم: مقداردی به داده های گم شده      |
| ۴۳۲ | ۱۳-۳. مهندسی ویژگی ها: تبدیل ویژگی های مجموعه داده ها قبل از آموزش مدل |
| ۴۳۳ | ۱۳-۳-۱. تبدیل داده های دسته ای به عددی: کدگذاری وان هات                |

|     |                                                                       |
|-----|-----------------------------------------------------------------------|
| ۴۳۶ | تبدیل داده‌های عددی به دسته‌ای (و چرایی آن): ظرف‌بندی                 |
| ۴۳۷ | انتخاب ویژگی‌ها: خلاص شدن از ویژگی‌های اضافی                          |
| ۴۳۸ | آموزش مدل‌ها                                                          |
| ۴۳۹ | تقسیم داده‌ها به ویژگی‌ها، برچسب‌ها و داده‌های آموزشی و اعتبارسنجی    |
| ۴۴۰ | آموزش چند مدل با مجموعه داده‌ها                                       |
| ۴۴۱ | کدام مدل بهتر است؟ ارزیابی مدل‌ها                                     |
| ۴۴۲ | تست مدل                                                               |
| ۴۴۳ | جست‌وجوی شبکه‌ای: تنظیم هاپرپارامترها برای پیدا کردن بهترین مدل       |
| ۴۴۷ | کاربرد اعتبارسنجی متقابل k-fold برای استفاده مجدد از داده‌ها برای ... |
| ۴۴۸ | خلاصه                                                                 |

## پیوست اول: پاسخ تمرین‌ها ..... ۴۵۱

## پیوست دوم: ریاضیات گرادیان کاهشی: پایین آمدن از کوه با استفاده از مشتقات و شیب‌ها ..... ۴۸۷

|     |                                                                           |
|-----|---------------------------------------------------------------------------|
| ۴۸۹ | استفاده از گرادیان کاهشی برای کمینه کردن توابع                            |
| ۴۹۱ | استفاده از گرادیان کاهشی برای آموزش مدل‌ها                                |
| ۴۹۳ | استفاده از گرادیان کاهشی برای آموزش مدل‌های رگرسیون خطی                   |
| ۴۹۶ | استفاده از گرادیان کاهشی برای آموزش مدل‌های دسته‌بندی                     |
| ۵۰۱ | استفاده از گرادیان کاهشی برای آموزش شبکه‌های عصبی                         |
| ۵۰۴ | استفاده از گرادیان کاهشی برای نرمال‌سازی                                  |
| ۵۰۵ | گیر کردن در یک کمینه محلی؛ چطور اتفاق می‌افتد و چگونه این مشکل را حل کنیم |