

متن کاوی نگرش یادگیری ماشین

مؤلف

چارو سی. آگاروال

مترجم

یوب ترکیان

نیاز دانش

سرشناسه	: آگاروال، چارو سی، ۱۹۷۰ - م Aggarwal, Charu C
عنوان و نام پدیدآور	: متن کاوی نگرش یادگیری ماشین / مولف چارو سی آگاروال؛ مترجم ایوب ترکیان
مشخصات نشر	: تهران: نیاز دانش، ۱۳۹۷
مشخصات ظاهری	: ۶۷۶ ص.: مصور، جدول، نمودار
شابک	: 978-600-8906-33-9
وضعیت فهرست نویسی	: فیفا
یادداشت	: عنوان اصلی: Machine learning for text, 2018.
موضوع	: فراگیری ماشینی Machine learning
موضوع	: متن پردازش (Computer science) Text processing
شناسه افزوده	: ترکیان، ایوب، ۱۳۳۷ - ، مترجم
رده بندی کنگره	: QC۳۲۵/۵/۱۷م۲ ۱۳۹۷
رده بندی دیو	: ۰۰۶/۳۱
شماره کتاب مناسب ملی	: ۵۴۲۹۶۴۳



نام کتاب	: متن کاوی نگش یادگیری ماشین
مؤلف	: چارو سی. آگاروال
مترجم	: ایوب ترکیان
مدیر اجرایی - ناظر بر چاپ	: حمیدرضا محمد شیرازی - محمد مصباحی
ناشر	: نیاز دانش
صفحه آرا	: واحد تولید انتشارات نیاز دانش
نوبت چاپ	: دوم - ۱۳۹۸
شمارگان	: ۱۰۰ نسخه
قیمت	: ۸۵۰۰۰۰ ریال

ISBN:978-600-8906-33-9

شابک: ۹۷۸-۶۰۰-۸۹۰۶-۳۳-۹

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 ماده ۲ قانون حمایت از مؤلفان، مصنفان و هنرمندان تحت پیگرد قانونی قرار می گیرند.

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: ۰۹۱۲-۲۱۰۶۷۰۹

فهرست مطالب

عنوان	شماره صفحه
فصل ۱ / مقدمه	۲۱
۱-۱ مقدمه	۲۱
۱-۱-۱ سازمان فصل	۲۴
۲-۱ اهمیت یادگیری از متن	۲۴
۳-۱ مدل‌های تحلیلی برای متن	۲۶
۱-۳-۱ پیش‌پردازش متن و محاسبه مشابهت	۲۷
۲-۳-۱ کاهش بُعد و فاکتورگیری ماتریس	۲۹
۳-۳-۱ خوشه‌بندی متن	۳۰
۱-۳-۳-۱ روش‌های قطعی و احتمالی فاکتورگیری ماتریس	۳۱
۲-۳-۳-۱ مدل‌های مخلوط احتمالی اسناد	۳۱
۳-۳-۳-۱ الگوریتم‌های مشابهت پایه	۳۲
۴-۳-۳-۱ روش‌های پیشرفته	۳۳
۴-۳-۱ طبقه‌بندی متن و مدل‌سازی رگرسیونی	۳۳
۱-۴-۳-۱ درخت‌های انتخاب	۳۵
۲-۴-۳-۱ طبقه‌گرهای قاعده پایه	۳۵
۳-۴-۳-۱ طبقه‌گر بیز ساده‌نگر	۳۵
۴-۴-۳-۱ طبقه‌گرهای نزدیک‌ترین همسایه	۳۶

۳۶	۵-۴-۳-۱ طبقه‌گرهای خطی
۳۷	۶-۴-۳-۱ موضوعات وسیع‌تر در طبقه‌بندی
۳۷	۵-۳-۱ تحلیل مشترک متن با داده‌های نامتجانس
۳۸	۶-۳-۱ بازیابی اطلاعات و جستجوی وب
۳۹	۷-۳-۱ مدل‌سازی و تعبیه‌گری زبان متوالی
۳۹	۸-۳-۱ خلاصه‌سازی متن
۳۹	۹-۳-۱ استخراج اطلاعات
۴۰	۱۰-۳-۱ نظرکاوی و تحلیل تمایل
۴۰	۱۱-۳-۱ بخش‌بندی متن و ردیابی رویداد
۴۰	۴-۱ خلاصه
۴۱	۵-۱ تمرین

فصل ۲ / آماده‌سازی متن و محاسبه مشابهت

۴۳	۱-۲ مقدمه
۴۴	۱-۱-۲ سازمان فصل
۴۴	۲-۲ استخراج و توکن‌سازی متن خام
۴۸	۱-۲-۲ مسایل خاص وب در استخراج متن
۴۹	۳-۲ استخراج ترم از توکن
۴۹	۱-۳-۲ حذف واژه‌های توقف
۵۰	۲-۳-۲ خط تیره کوتاه
۵۱	۳-۳-۲ تغییر حالت حرف
۵۱	۴-۳-۲ هم‌سان‌سازی کاربردپایه
۵۲	۵-۳-۲ ریشه‌یابی
۵۳	۴-۲ نمایش فضای برداری و نرمال‌سازی
۵۵	۵-۲ محاسبه مشابهت در متن
۵۸	۱-۵-۲ مفیدبودن نرمال‌سازی idf و ریشه‌یابی
۵۹	۶-۲ خلاصه
۶۰	۷-۲ تمرین

۶۱	فصل ۳ / فاکتورگیری ماتریسی و مدل سازی موضوعی
۶۱	۱-۳ مقدمه
۶۴	۱-۱-۳ سازمان فصل
۶۴	۲-۱-۳ نرمال سازی فاکتورگیری دوطرفه به سه طرفه
۶۶	۲-۳ تجزیه مقدار منفرد
۷۰	۱-۲-۳ مثال SVD
۷۲	۲-۲-۳ روش قدرت پیاده سازی SVD
۷۳	۳-۱-۳ کاربردهای SVD/LSA
۷۳	۴-۱-۳ محاسن و معایب SVD/LSA
۷۴	۳-۳ فاکتورگیری ماتریس غیرمنفی
۷۷	۱-۳-۳ تقریب رتبه ای فاکتورگیری ماتریس غیرمنفی
۷۸	۲-۳-۳ مثال فاکتورگیری ماتریس غیرمنفی
۸۰	۳-۳-۳ کوتاه سازی در اسناد جدید
۸۱	۴-۳-۳ مزایا و معایب فاکتورگیری ماتریس غیرمنفی
۸۱	۴-۳ تحلیل سمانتیک ناپیدا احتمالی
۸۵	۱-۴-۳ ارتباط با فاکتورگیری ماتریس غیرمنفی
۸۶	۲-۴-۳ مقایسه با SVD
۸۶	۳-۴-۳ مثال PLSA
۸۷	۴-۴-۳ محاسن و معایب PLSA
۸۸	۵-۳ بررسی اجمالی تخصیص دیریکله نهفته
۸۸	۱-۵-۳ مدل LDA ساده شده
۹۲	۲-۵-۳ مدل LDA هموار شده
۹۳	۶-۳ تبدیل غیرخطی و مهندسی ویژگی
۹۷	۱-۶-۳ انتخاب تابع مشابهت
۹۷	۱-۱-۶-۳ توابع مشابهت کرنل سنتی
۱۰۰	۲-۱-۶-۳ تعمیم کیف واژه به N-گرام
۱۰۱	۳-۱-۶-۳ کرنل های زیرتوالی رشته
۱۰۵	۴-۱-۶-۳ تسریع زمان عملیات بازگشتی
۱۰۵	۵-۱-۶-۳ کرنل های تابع زبان

۱۰۶	تخمین نیستروم ۲-۶-۳
۱۰۸	موجودیت جزئی ماتریس مشابهت ۳-۶-۳
۱۱۱	خلاصه ۷-۳
۱۱۱	تمرین ۸-۳

فصل ۴ / خوشه‌بندی متن ۱۱۵

۱۱۵	۱-۴ مقدمه
۱۱۷	۱-۴ سازمان فصل
۱۱۷	۱-۴ انتخاب و مهندسی ویژگی
۱۱۸	۱-۲-۴ انتخاب ویژگی
۱۱۸	۱-۲-۴ قدرت ترم
۱۱۹	۲-۲-۴-۱ مدل‌سازی با نظارت برای انتخاب ویژگی بدون نظارت
۱۲۰	۳-۱-۲-۴ لایه‌ی با نظارت با انتخاب ویژگی با نظارت
۱۲۰	۲-۲-۴ مهندسی ویژگی
۱۲۱	۱-۲-۲-۴ روش‌های فاکتورگیری ماتریس
۱۲۲	۲-۲-۲-۴ کاهش بعدیت غیرخطی
۱۲۳	۳-۲-۲-۴ تعبیه‌گری واژه
۱۲۳	۳-۴ مدل‌سازی موضوعی و فاکتورگیری ماتریس
۱۲۳	۱-۳-۴ مدل‌های عضویت آمیخته و خوشه‌های هم‌پوشانی
۱۲۴	۲-۳-۴ خوشه‌های بدون هم‌پوشانی و هم‌خوشه‌گی
۱۲۸	۱-۲-۳-۴ هم‌خوشه‌گی با بخش‌بندی گراف دوجهتی
۱۲۹	۴-۴ مدل‌های آمیخته زایا برای خوشه‌بندی
۱۳۰	۱-۴-۴ مدل برنولی
۱۳۲	۲-۴-۴ مدل مولتی‌نومی
۱۳۴	۳-۴-۴ مقایسه با مدل‌های موضوع با عضویت آمیخته
۱۳۵	۴-۴-۴ ارتباط با مدل بیز ساده‌شده برای طبقه‌بندی
۱۳۶	۵-۴ الگوریتم k-میانگین
۱۳۹	۱-۵-۴ هم‌گرایی و آغازگری
۱۳۹	۲-۵-۴ پیچیدگی محاسباتی

۱۴۰	ارتباط با مدل‌های احتمالی	۳-۵-۴
۱۴۰	الگوریتم‌های خوشه‌بندی سلسله‌مراتبی	۶-۴
۱۴۴	پیاپی‌سازی کارآمد و پیچیدگی محاسباتی	۱-۶-۴
۱۴۶	ارتباط طبیعی با k-میانگین	۲-۶-۴
۱۴۷	دسته‌جمعی‌های خوشه‌بندی	۷-۴
۱۴۸	انتخاب مؤلفه دسته‌جمعی	۱-۷-۴
۱۴۹	تلفیق نتایج از مؤلفه‌های مختلف	۲-۷-۴
۱۴۹	خوشه‌بندی متن به صورت توالی	۸-۴
۱۵۰	روش‌های کرنلی برای خوشه‌بندی	۱-۸-۴
۱۵۱	k-میانگین کرنلی	۱-۱-۸-۱
۱۵۲	مهندسی ویژگی صریح	۲-۸-۴
۱۵۴	کرنل یا مهندسی ویژگی صریح؟	۱-۸-۴
۱۵۵	کرنل‌های ابعادی خوشه‌بندی طیفی	۲-۸-۴
۱۵۷	تبدیل خوشه‌بندی به دگر با نظارت	۹-۴
۱۵۸	موضوعات عملی	۱-۹-۴
۱۵۹	ارزیابی خوشه‌بندی	۱۰-۴
۱۵۹	چاله‌های سنج‌های صحه‌گذاری داخلی	۱-۱۰-۴
۱۶۰	سنج‌های صحه‌گذاری خارجی	۲-۱۰-۴
۱۶۴	رابطه ارزیابی خوشه‌بندی با نظارت	۱-۲-۱۰-۴
۱۶۵	اشتباهات رایج در ارزیابی	۲-۲-۱۰-۴
۱۶۶	خلاصه	۱۱-۴
۱۶۷	تمرین	۱۲-۴

فصل ۵ / طبقه‌بندی متن: مدل‌های پایه

۱۶۹	مقدمه	۱-۵
۱۷۱	انواع نشانه‌ها و مدل‌سازی رگرسیون	۱-۱-۵
۱۷۲	آموزش و تست	۲-۱-۵
۱۷۳	یادگیرنده‌های استقرایی، شرطی، و استنتاجی	۳-۱-۵
۱۷۴	مدل‌های بنیادی	۴-۱-۵

۱۷۵	۵-۱-۵ چالش‌های خاص متن در طبقه‌گرها
۱۷۵	۱-۵-۱-۵ سازمان فصل
۱۷۶	۲-۵ انتخاب و مهندسی ویژگی
۱۷۶	۱-۲-۵ نمایه Gini
۱۷۷	۲-۲-۵ انتروپی شرطی
۱۷۸	۳-۲-۵ اطلاعات جمعی جفت‌وار
۱۷۸	۴-۲-۵ سنجه‌های با ارتباط نزدیک
۱۸۰	۵-۲-۵ آماره χ^2
۱۸۱	۶-۱-۵ مدل‌های انتخاب ویژگی تعبیه‌گری
۱۸۲	۷-۱-۵ شگردهای مهندسی ویژگی
۱۸۳	۳-۵ مدل ساده‌شده
۱۸۴	۱-۳-۵ مدل برنولی
۱۸۵	۱-۱-۳-۵ فاز پیش‌بینی
۱۸۶	۲-۱-۳-۵ فاز امتحان
۱۸۷	۲-۳-۵ مدل مولتی‌نومی
۱۸۹	۳-۳-۵ مشاهدات عملی
۱۸۹	۴-۳-۵ درجه‌بندی خروجی‌ها با بیز ساده‌شده
۱۹۰	۵-۳-۵ مثال بیز ساده‌شده
۱۹۰	۱-۵-۳-۵ مدل برنولی
۱۹۲	۲-۵-۳-۵ مدل مولتی‌نومی
۱۹۴	۶-۳-۵ بیز ساده‌شده نیمه‌نظارتی
۱۹۷	۴-۵ طبقه‌گر نزدیک‌ترین همسایه
۱۹۸	۱-۴-۵ خصوصیات طبقه‌گر ۱-نزدیک‌ترین همسایه
۲۰۱	۲-۴-۵ طبقه‌بندی Ricchio و نزدیک‌ترین ثقل مرکزی
۲۰۳	۳-۴-۵ نزدیک‌ترین همسایه وزن‌دهی شده
۲۰۴	۱-۳-۴-۵ ۱-نزدیک‌ترین همسایه کیسه‌ای و زیرنمونه
۲۰۶	۴-۴-۵ نزدیک‌ترین همسایه انطباقی: خانواده‌های قدرت‌مند
۲۰۹	۵-۵ درختان تصمیم و جنگل‌های راندومی
۲۱۰	۱-۵-۵ روش پایه برای ساخت درخت تصمیم

۲۱۲	۲-۵-۵ تقسیم گره
۲۱۲	۱-۲-۵-۵ پیش‌بینی
۲۱۳	۳-۵-۵ تقسیم‌های چندمتغیره
۲۱۳	۴-۵-۵ موضوعات مشکل‌آفرین درخت تصمیم
۲۱۵	۵-۵-۵ جنگل‌های تصادفی
۲۱۶	۶-۵-۵ جنگل‌های تصادفی به‌عنوان نزدیک‌ترین همسایه انطباقی
۲۱۷	۶-۵-۶ طبقه‌گرهای قاعده‌پایه
۲۱۸	۱-۶-۵ الگوریتم‌های با پوشش متوالی
۲۱۹	۱-۱-۶-۵ قاعده یک-یادگیری
۲۲۱	۲-۱-۶-۵ هرس قاعده
۲۲۲	۲-۶-۵ زبید قاعده از درخت تصمیم
۲۲۳	۳-۶-۵ طبقه‌بندی انجمنی
۲۲۴	۴-۶-۵ پیش‌بینی
۲۲۴	۷-۵ خلاصه
۲۲۵	۸-۵ تمرین

فصل ۶ / طبقه‌بندی خطی و گرسیون برای متن

۲۲۹	۱-۶ مقدمه
۲۳۰	۱-۱-۶ تفسیر هندسی مدل‌های خطی
۲۳۲	۲-۱-۶ متغیر سوپه لازم است؟
۲۳۳	۳-۱-۶ تعریف کلی مدل‌های خطی با تنظیم
۲۳۵	۴-۱-۶ تعمیم پیش‌بینی‌های دوتایی به چندطبقه
۲۳۶	۵-۱-۶ خصوصیات مدل‌های خطی برای متن
۲۳۷	۱-۵-۱-۶ نشانه‌گذاری فصل
۲۳۷	۲-۵-۱-۶ سازمان فصل
۲۳۷	۲-۶ رگرسیون کمترین مربعات و طبقه‌بندی
۲۳۸	۱-۲-۶ رگرسیون کمترین مربعات با تنظیم L_1
۲۳۹	۱-۱-۲-۶ پیاده‌سازی کارآمد
۲۳۹	۲-۱-۲-۶ تخمین تقریبی با تجزیه مقدار منفرد

۲۴۱	۳-۱-۲-۶	رابطه با رگرسیون مؤلفه‌های اصلی
۲۴۱	۴-۱-۲-۶	مسیر به رگرسیون کرنلی
۲۴۲	۲-۲-۶	LASSO: رگرسیون کمترین مربعات با تنظیم L_1
۲۴۴	۱-۲-۲-۶	تفسیر LASSO به عنوان انتخاب گر ویژگی
۲۴۵	۳-۲-۶	تمایزگر خطی فشر و طبقه‌بندی کمترین مربعات
۲۴۸	۱-۳-۲-۶	تمایزگر خطی با چند طبقه
۲۴۹	۲-۳-۲-۶	برابری تمایزگر فشر و رگرسیون کمترین مربعات
۲۵۱	۳-۳-۲-۶	طبقه‌بندی کمترین مربعات تنظیم‌شده و LLSF
۲۵۳	۴-۳-۲-۶	پاشنه آشیل طبقه‌بندی کمترین مربعات
۲۵۴	۲-۶	آشپان‌های بردار پشتیبان
۲۵۵	۱-۳-۶	مسیر به بهسازی تنظیم‌شده
۲۵۶	۲-۳-۶	تفسیر حداقل‌تر حاشیه
۲۵۸	۳-۳-۶	Support Vector Machines در اولی
۲۵۹	۱-۳-۳-۶	به‌روزرسانی‌ها، سازگار با تنگی
۲۶۰	۴-۳-۶	فرمولاسیون SVM دو کلاسه
۲۶۳	۵-۳-۶	الگوریتم‌های یادگیری برای SVM کلاسه‌های دوگانه
۲۶۴	۶-۳-۶	تفسیر نزدیکترین همسایه انطباقی برای SVM کلاسه‌های دوگانه
۲۶۶	۴-۶	رگرسیون لجیستیک
۲۶۷	۱-۴-۶	تفسیر بهینه‌سازی تنظیم‌شده
۲۶۹	۲-۴-۶	الگوریتم‌های آموزشی برای رگرسیون لجیستیک
۲۷۰	۳-۴-۶	تفسیر احتمالی رگرسیون لجیستیک
۲۷۱	۱-۳-۴-۶	تفسیر احتمالی گام‌های فرود گرادینتی استوکاستیک
۲۷۲	۲-۳-۴-۶	رابطه بین به‌روزرسانی‌های اولی مدل‌های خطی
۲۷۲	۴-۴-۶	رگرسیون لجیستیک مولتی‌نومی و دیگر تعمیم‌ها
۲۷۴	۵-۴-۶	توضیحات عملکرد رگرسیون لجیستیک
۲۷۵	۵-۶	تعمیم غیرخطی مدل‌های خطی
۲۷۷	۱-۵-۶	SVM‌های کرنلی با تبدیل صریح
۲۷۸	۲-۵-۶	ترویج تفکیک‌پذیری خطی توسط کرنل‌های متعارف
۲۸۱	۳-۵-۶	قوت و ضعف کرنل‌های مختلف

۲۸۲	۱-۳-۵-۶	صید دانش زبان شناسی با کرنل‌ها
۲۸۲	۴-۵-۶	شگرد کرنلی
۲۸۳	۵-۵-۶	کاربرد سیستماتیک شگرد کرنلی
۲۸۹	۶-۶	خلاصه
۲۸۹	۷-۶	تمرین

فصل ۷ / عملکرد و ارزیابی طبقه‌گر

۲۹۳	۱-۷	مقدمه
۲۹۴	۱-۷	سازمان فصل
۲۹۴	۲-۷	تفاوت ریه و واریانس
۲۹۶	۱-۲-۷	۱۴۰ رسم
۳۰۱	۲-۲-۷	نشانه‌های گویای ریه و واریانس
۳۰۱	۳-۷	تبعات تفاوت ریه و واریانس عملکرد
۳۰۲	۱-۳-۷	اثر اندازه داده‌ها، آموزش
۳۰۴	۲-۳-۷	اثر بُعدیت داده‌ها
۳۰۵	۳-۳-۷	تبعات انتخاب مدل در متر
۳۰۶	۴-۷	بهبود عملکرد سیستماتیک با هنگردها
۳۰۶	۱-۴-۷	سبب‌سازی و زیرنمونه‌برداری
۳۰۸	۲-۴-۷	تقویت کردن
۳۱۱	۵-۷	ارزیابی طبقه‌گر
۳۱۳	۱-۵-۷	تقسیم به بخش‌های آموزش و تست
۳۱۳	۱-۱-۵-۷	بیرون نگاه داشتن
۳۱۴	۲-۱-۵-۷	صحه‌گذاری تقاطعی
۳۱۴	۲-۵-۷	سنجه صحت مطلق
۳۱۴	۱-۲-۵-۷	صحت طبقه‌بندی
۳۱۶	۲-۲-۵-۷	صحت رگرسیون
۳۱۸	۳-۵-۷	سنجه‌های رتبه‌بندی برای طبقه‌بندی و بازیابی اطلاعات
۳۱۹	۱-۳-۵-۷	خصوصیات عملیاتی دریافت کننده
۳۲۴	۲-۳-۵-۷	قوی‌ترین سنجه‌ها برای لیست‌های درجه‌بندی

۳۳۶	۶-۷ خلاصه
۳۳۶	۷-۷ تمرین

فصل ۸ / متن کاوی مشترک با داده‌های نامتجانس

۳۳۷	۱-۸ مقدمه
۳۳۰	۱-۱-۸ سازمان فصل
۳۳۰	۲-۸ شگرد فاکتورگیری ماتریس اشتراکی
۳۳۱	۱-۸ گراف فاکتورگیری
۳۳۲	۲-۸ کاربرد: فاکتورگیری اشتراکی با متن و لینک وب
۳۳۴	۱-۲-۸ حل مسئله به‌سازی
۳۳۶	۲-۲-۸ تعریف با نظارت
۳۳۷	۳-۲-۸ کاربرد: متن با شبکه‌های اجتماعی بدون جهت
۳۳۸	۱-۳-۲-۸ کاربرد: پیش‌بینی لینک با محتوای متن
۳۳۸	۴-۲-۸ یادگیری انتقالی و تضاد با متن
۳۳۹	۱-۲-۴-۸ یادگیری انتقالی با متن بدون نشانه
۳۴۱	۲-۴-۲-۸ یادگیری انتقالی با متن نشانه‌دار
۳۴۳	۵-۲-۸ کاربرد: سیستم‌های توصیه‌گر با درج بنابر متن
۳۴۵	۶-۲-۸ کاربرد: متن کاوی بین‌زبانی
۳۴۶	۳-۸ ماشین‌های فاکتورگیری
۳۵۱	۴-۸ روش‌های مدل‌سازی احتمالی مشترک
۳۵۲	۱-۴-۸ مدل‌های احتمالی مشترک برای خوشه‌بندی
۳۵۳	۲-۴-۸ طبقه‌گر بیز ساده‌شده
۳۵۳	۵-۸ تبدیل به روش‌های گراف کاوی
۳۵۶	۶-۸ خلاصه
۳۵۷	۷-۸ تمرین

فصل ۹ / بازیابی اطلاعات و موتورهای جستجو

۳۵۹	۱-۹ مقدمه
-----	-----------

۳۶۰	۱-۱-۹ سازمان فصل
۳۶۱	۲-۹ نمایه‌سازی و پردازش پرسمان
۳۶۲	۱-۲-۹ ساختارهای داده فرهنگ‌نامه
۳۶۵	۲-۲-۹ نمایه وارونه
۳۶۶	۳-۲-۹ ساخت نمایه زمان خطی
۳۶۹	۴-۲-۹ پردازش پرسمان
۳۶۹	۱-۴-۲-۹ بازیابی بولی
۳۷۰	۲-۴-۲-۹ بازیابی درجه‌بندی‌شده
۳۷۲	۳-۴-۲-۹ پردازش پرسمان با جمع‌زن ترم-در-هربار
۳۷۴	۴-۴-۲-۹ پردازش پرسمان با جمع‌زن سند-در-هربار
۳۷۵	۵-۲-۹ ترم-در-هربار یا سند-در-هربار؟
۳۷۵	۳-۴-۲-۹ انتخاب نمره رایج
۳۷۵	۶-۴-۲-۹ پرسمان موقیتی
۳۷۷	۸-۴-۲-۹ نمره‌گذاری نام‌های
۳۷۸	۹-۴-۲-۹ یادگیری ماشین از بازیابی اطلاعات
۳۸۱	۱۰-۴-۲-۹ ماشین‌های بردار پشتیبان درجه‌بند
۳۸۳	۵-۲-۹ به‌سازی‌های کارایی
۳۸۳	۱-۵-۲-۹ نشانه‌گرهای گذری
۳۸۴	۲-۵-۲-۹ لیست‌های قهرمان و نمایه‌های نامی
۳۸۵	۳-۵-۲-۹ شگردهای ذخیره موقت
۳۸۶	۴-۵-۲-۹ شگردهای فشرده‌سازی
۳۸۸	۳-۹ نمره‌بندی با مدل‌های بازیابی اطلاعات
۳۸۹	۱-۳-۹ مدل‌های فضای برداری با tf-idf
۳۹۰	۲-۳-۹ مدل استقلال بوتایی
۳۹۳	۳-۳-۹ مدل BM25 با فرکانس‌های ترم
۳۹۵	۴-۳-۹ مدل‌های زبان آماری در بازیابی اطلاعات
۳۹۶	۱-۴-۳-۹ مدل‌های شایدهی پرسمان
۳۹۸	۴-۹ وب‌خیزی و کشف منبع
۳۹۹	۱-۴-۹ یک الگوریتم خزنده پایه

۴۰۱	 خزنده‌های ترجیحی	۲-۴-۹
۴۰۳	 چند رشته	۳-۴-۹
۴۰۴	 مقابله با تله‌های عنکبوت	۴-۴-۹
۴۰۴	 همپوشی برای آشکارسازی تکرار نزدیک	۵-۴-۹
۴۰۵	 پردازش پرسمان در موتورهای جستجو	۵-۹
۴۰۶	 ساخت نمایه توزیعی	۱-۵-۹
۴۰۷	 به‌روزرسانی پویای نمایه	۲-۵-۹
۴۰۷	 پردازش پرسمان	۳-۵-۹
۴۰۹	 اهمیت اعتبار	۴-۵-۹
۴۱۱	 خوریتیم‌های درجه‌بندی لینک پایه	۶-۹
۴۱۱	 رجه‌بندی	۱۰-۶-۹
۴۱۵	 صفحه حساس به موضوع	۱-۶-۹
۴۱۶	 درجه‌بندی وابستگی	۲-۱-۶-۹
۴۱۸	 HITS	۲-۶-۹
۴۲۰	 خلاصه	۷-۹
۴۲۰	 تمرین	۸-۹

فصل ۱۰ / مدل‌سازی توالی متن و یادگیری عمیق

۴۲۳	 مقدمه	۱-۱۰
۴۲۶	 سازمان فصل	۱-۱-۱۰
۴۲۷	 مدل‌های زبان آماری	۲-۱۰
۴۳۱	 مدل‌های گذرگرام	۱-۲-۱۰
۴۳۲	 رابطه با تعبیه‌گری‌ها	۲-۲-۱۰
۴۳۴	 روش‌های کرنلی	۳-۱۰
۴۳۵	 مدل‌های فاکتورگیری ماتریس واژه-زمینه	۴-۱۰
۴۳۶	 فاکتورگیری ماتریس با شمارش‌ها	۱-۴-۱۰
۴۳۸	 سوژه‌های پس پردازش	۱-۱-۴-۱۰
۴۳۹	 تعبیه‌گری GloVe	۲-۴-۱۰
۴۴۰	 فاکتورگیری ماتریس PPMI	۳-۴-۱۰

۴۴۱	۴-۴-۱۰ فاکتورگیری ماتریس PPMI شیف‌شده
۴۴۱	۵-۴-۱۰ گنجاندن نحو و دیگر ویژگی‌ها
۴۴۲	۵-۱۰ نمایش‌های گرافیکی فاصله واژه‌ها
۴۴۴	۶-۱۰ مدل‌های زبان عصبی
۴۴۴	۱-۶-۱۰ شبکه‌های عصبی: معرفی آرام
۴۴۵	۱-۱-۶-۱۰ لایه محاسباتی تکی: پرسپترون
۴۴۸	۲-۱-۶-۱۰ رابطه با ماشین‌های بردار پشتیبان
۴۵۰	۳-۱-۶-۱۰ انتخاب تابع فعال‌سازی
۴۵۱	۴-۱-۶-۱۰ انتخاب گره‌های خروجی
۴۵۱	۵-۱-۶-۱۰ انتخاب تابع اتلاف
۴۵۳	۶-۱-۶-۱۰ شبکه‌های عصبی چندلایه
۴۵۹	۲-۶-۱۰ تعبیر عصبی با Word2vec
۴۵۹	۱-۲-۶-۱۰ تعبیه‌گری عصبی با کیف‌واژه پیوسته
۴۶۳	۲-۲-۶-۱۰ تعبیه‌گری عصبی با مدل گذرگرام
۴۶۶	۳-۲-۶-۱۰ سوزدهای عمل
۴۶۷	۴-۲-۶-۱۰ گذرگرام با نمونه‌برداری منفی
۴۶۹	۵-۲-۶-۱۰ معماری عصبی SCNS
۴۶۹	۳-۶-۱۰ Word2vec (SGNS) و فاکتورگیری ماتریس لجیستیک
۴۷۲	۱-۳-۶-۱۰ فرود گرادینانی
۴۷۳	۴-۶-۱۰ تعبیه‌گری پاراگراف با Doc2vec
۴۷۵	۷-۱۰ شبکه‌های عصبی برگشتی
۴۷۸	۱-۷-۱۰ سوزدهای عملی
۴۷۸	۲-۷-۱۰ مثال مدل‌سازی زبان RNN
۴۸۰	۱-۲-۷-۱۰ تولید نمونه زبان
۴۸۱	۳-۷-۱۰ کاربرد به توضیح اتوماتیک تصویر
۴۸۳	۴-۷-۱۰ یادگیری توالی به توالی و ترجمه ماشین
۴۸۶	۱-۴-۷-۱۰ سامانه‌های سوال-جواب
۴۸۸	۵-۷-۱۰ کاربرد به طبقه‌بندی سطح جمله
۴۸۹	۶-۷-۱۰ طبقه‌بندی سطح توکن با ویژگی‌های زبان‌شناختی

۴۹۰	شبکه‌های برگشتی چندلایه	۷-۷-۱۰
۴۹۲	حافظه کوتاه مدت طولانی (LSTM)	۱-۷-۷-۱۰
۴۹۵	خلاصه	۸-۱۰
۴۹۶	تمرین	۹-۱۰

فصل ۱۱ / خلاصه‌سازی متن ۴۹۹

۴۹۹	مقدمه	۱-۱۱
۵۰۰	خلاصه‌سازی استخراجی و چکیده‌ای	۱-۱-۱
۵۰۱	گام‌های کلیدی در خلاصه‌سازی استخراجی	۲-۱-۱۱
۵۰۲	فاز پیش‌بندی در خلاصه‌سازی استخراجی	۳-۱-۱۱
۵۰۲	ارماز، مثال	۴-۱-۱۱
۵۰۲	روش‌های موضوع و واژه خلاصه‌سازی استخراجی	۲-۱۱
۵۰۳	احتمالات و اوا	۱-۲-۱۱
۵۰۴	اوزان فرکانس نرمال شده	۲-۲-۱۱
۵۰۶	امضاءهای موضوع	۳-۲-۱۱
۵۰۸	روش‌های انتخاب جمله	۴-۲-۱۱
۵۰۹	روش‌های ناپیدا برای خلاصه‌سازی استخراجی	۳-۱۱
۵۰۹	تحلیل سمانتیک ناپیدا	۱-۳-۱۱
۵۱۱	زنجیره‌های واژگانی	۲-۳-۱۱
۵۱۱	شرح کوتاه WordNet	۱-۲-۳-۱۱
۵۱۲	بهره‌برداری از WordNet برای زنجیره‌های واژگانی	۲-۲-۳-۱۱
۵۱۴	روش‌های گراف پایه	۳-۳-۱۱
۵۱۵	خلاصه‌سازی ثقلی	۴-۳-۱۱
۵۱۶	یادگیری ماشین برای خلاصه‌سازی استخراجی	۴-۱۱
۵۱۷	استخراج ویژگی	۱-۴-۱۱
۵۱۸	طبقه‌گر مورد استفاده	۲-۴-۱۱
۵۱۸	خلاصه‌سازی چندانندی	۵-۱۱
۵۱۸	خلاصه‌سازی ثقل پایه	۱-۵-۱۱
۵۲۰	روش‌های گراف پایه	۲-۵-۱۱

۵۲۱	۶-۱۱ خلاصه‌سازی چکیده‌ای
۵۲۲	۱-۶-۱۱ فشرده‌سازی جمله
۵۲۳	۲-۶-۱۱ فیوژن اطلاعات
۵۲۳	۳-۶-۱۱ مرتب‌سازی اطلاعات
۵۲۴	۷-۱۱ خلاصه
۵۲۴	۸-۱۱ تمرین

فصل ۱۲ / استخراج اطلاعات ۵۲۵

۵۲۵	۱-۱۲ مقدمه
۵۲۸	۱-۱-۱۲ سامر تابیخی
۵۲۹	۲-۱-۱۲ نقش برداشت زبان طبیعی
۵۳۱	۳-۱-۱۲ سازمان فدل
۵۳۱	۲-۱۲ شناسایی مقوله سامدار
۵۳۳	۱-۲-۱۲ روش‌های قاعده پایه
۵۳۵	۱-۱-۲-۱۲ الگوریتم‌های آموزش برای سامانه‌های قاعده پایه
۵۳۶	۲-۱-۲-۱۲ تولید قاعده بالا به پایین
۵۳۷	۳-۱-۲-۱۲ تولید قاعده پایین به بالا
۵۳۸	۲-۲-۱۲ تبدیل به طبقه‌بندی سطح توکن
۵۳۹	۳-۲-۱۲ مدل‌های مارکوف پنهان
۵۴۰	۱-۳-۲-۱۲ مدل‌های مارکوف پنهان و مرئی
۵۴۰	۲-۳-۲-۱۲ سامانه Nymble
۵۴۲	۳-۳-۲-۱۲ آموزش
۵۴۳	۴-۳-۲-۱۲ پیش‌بینی برای بخش تست
۵۴۴	۵-۳-۲-۱۲ گنجاندن ویژگی‌های استخراج‌شده
۵۴۴	۶-۳-۲-۱۲ روایت‌ها و بهبودها
۵۴۵	۴-۲-۱۲ مدل‌های مارکوف انتروپی حداکثر
۵۴۷	۵-۲-۱۲ فیلدهای راندومی شرطی
۵۴۹	۳-۱۲ استخراج رابطه
۵۵۰	۱-۳-۱۲ تبدیل به طبقه‌بندی

۵۵۲	۲-۳-۱۲	پیش‌بینی رابطه با مهندسی ویژگی صریح
۵۵۳	۱-۲-۳-۱۲	استخراج ویژگی از توالی‌های جمله
۵۵۵	۲-۲-۳-۱۲	ساده‌سازی درختان تجزیه با گراف‌های وابستگی
۵۵۶	۳-۳-۱۲	پیش‌بینی رابطه با مهندسی ویژگی: روش‌های کرنلی
۵۵۷	۱-۳-۳-۱۲	کرنل‌ها از گراف‌های وابستگی
۵۵۸	۲-۳-۳-۱۲	کرنل‌های زیرتوالی‌پایه
۵۵۹	۳-۳-۳-۱۲	کرنل‌های درخت‌پایه کانولوشنی
۵۶۲	۴-۱۲	خلاصه
۵۶۲	۱۲	تدریس

فصل ۱۳ / نظر‌کاری و تحلیل گرایش

۵۶۵	۱-۱۳	مقدمه
۵۶۸	۱-۱-۱۳	واژگان نذر
۵۶۹	۱-۱-۱-۱۳	رویکرد ای و نام‌پایه
۵۶۹	۲-۱-۱-۱۳	رویکردهای یک‌پایه
۵۷۰	۲-۱-۱۳	نظر‌کاوی به‌عنوان سلول کرنی و استخراج اطلاعات
۵۷۲	۳-۱-۱۳	سازمان فصل
۵۷۲	۲-۱۳	طبقه‌بندی گرایش سطح سند
۵۷۴	۱-۲-۱۳	رویکردهای بدون نظارت به طبقه‌بندی
۵۷۶	۳-۱۳	طبقه‌بندی گرایش سطح جمله و سطح عبارت
۵۷۷	۱-۳-۱۳	کاربردهای تحلیل سطح جمله و سطح عبارت
۵۷۸	۲-۳-۱۳	تبدیل طبقه‌بندی ذهنیت به مسئله قطع کمینه
۵۷۸	۳-۳-۱۳	زمینه در تحلیل قطبیت سطح جمله و عبارت
۵۷۹	۴-۱۳	نظر‌کاوی جنبه‌پایه به‌عنوان استخراج اطلاعات
۵۸۰	۱-۴-۱۳	رویکرد بدون نظارت Hu & Liu
۵۸۲	۲-۴-۱۳	OPINE: رویکرد بدون نظارت
۵۸۳	۳-۴-۱۳	استخراج نظر با نظارت به‌عنوان طبقه‌بندی سطح توکن
۵۸۵	۵-۱۳	هرز نظر
۵۸۶	۱-۵-۱۳	روش‌های با نظارت برای آشکارسازی هرز‌نظر

۵۸۷	۱-۱-۵-۱۳	نشانه‌گذاری هرزنظر فریب‌کار
۵۸۸	۲-۱-۵-۱۳	استخراج ویژگی
۵۸۹	۲-۵-۱۳	روش‌های بدون نظارت برای آشکارسازی هرزه‌گر
۵۹۰	۶-۱۳	خلاصه‌سازی نظر
۵۹۰	۱-۶-۱۳	خلاصه درجه‌بندی
۵۹۱	۲-۶-۱۳	خلاصه گرایش
۵۹۱	۳-۶-۱۳	خلاصه گرایش با عبارات و جملات
۵۹۱	۴-۶-۱۳	خلاصه‌های استخراجی و چکیده‌ای
۵۹۲	۷-۱۳	خلاصه
۵۹۲	۸-۱۳	تمرین

فصل ۱۴ / بخش بندی در متن و آشکارسازی رویداد ۵۹۳

۵۹۳	۱-۱۴	مقدمه
۵۹۴	۱-۱-۱۴	رابطه با آشکارسازی و دیایی موضوع
۵۹۴	۲-۱-۱۴	سازمان فصل
۵۹۵	۲-۱۴	بخش‌بندی متن
۵۹۶	۱-۲-۱۴	قطعه‌بندی متن
۵۹۷	۲-۲-۱۴	رویکرد C99
۵۹۹	۳-۲-۱۴	بخش‌بندی با نظارت با طبقه‌گرهای آماده
۶۰۱	۴-۲-۱۴	بخش‌بندی با نظارت با مدل‌های مارکوفی
۶۰۴	۳-۱۴	کاوش جریان‌های متن
۶۰۴	۱-۳-۱۴	خوشه‌بندی متن جاری
۶۰۵	۲-۳-۱۴	کاربرد به آشکارسازی اولین داستان
۶۰۷	۴-۱۴	آشکارسازی رویداد
۶۰۸	۱-۴-۱۴	آشکارسازی رویداد بدون نظارت
۶۰۸	۱-۱-۴-۱۴	روش نزدیک‌ترین همسایه پنجره‌پایه
۶۰۹	۲-۱-۴-۱۴	بهره‌برداری از مدل‌های زایشی
۶۱۰	۳-۱-۴-۱۴	آشکارسازی رویداد در جریان‌های اجتماعی
۶۱۰	۲-۴-۱۴	آشکارسازی رویداد و بخش‌بندی با نظارت

- ۶۱۱ ۳-۴-۱۴ آشکارسازی رویداد و استخراج اطلاعات
- ۶۱۲ ۱-۳-۴-۱۴ تبدیل به طبقه‌بندی سطح توکن
- ۶۱۳ ۲-۳-۴-۱۴ استخراج رویداد زمینه باز
- ۶۱۵ ۵-۱۴ خلاصه
- ۶۱۵ ۶-۱۴ تمرین

- ۶۱۷ پیوست «الف» - آموزش عملی متن کاوی
- ۶۶۹ پیوست «ب» - فرهنگ‌نامه اصطلاحات

www.ketab.ir