



تحلیل عواطف با استفاده از تکنیک‌های
یادگیری ماشین

مؤلف:

محمد دهقانی



سازمان نشر و دانش
سانجش ودانیش

- سرشناسه
عنوان و نام پدیدآور
مشخصات نشر
مشخصات ظاهری
شابک
وضعیت فهرست نویسی
یادداشت
موضوع
رده بندی کنگره
رده بندی دیویی
شماره کتابشناسی ملی
اطلاعات رکورد کتابشناسی
- دهقانی، محمد. ۱۳۷۴-
تحلیل عواطف با استفاده از تکنیک‌های یادگیری ماشین / مؤلف محمد دهقانی.
تهران: سنجش و دانش، ۱۴۰۰.
ل، ۱۰۳ص.؛ مصور، جدول، نمودار.
۲۵۰۰۰۰ ریال/ 978-622-05-2821-0
فیبا
کتابنامه: ص. ۹۸-۱۰۳.
نویسن
Twitter
شبکه‌های اجتماعی پیوسته -- ایران -- جنبه‌های سیاسی -- نمونه‌پژوهی
Online social networks -- Iran -- Political aspects -- Case studies
HMV۴۲
۰۰۶/۷۵۴
۸۷۴۳۵۱۴
فیبا

www.ketab.ir

عنوان: تحلیل عواطف با استفاده از تکنیکهای یادگیری ماشین		 www.sanjesh.ir sanjeshodanesh@iran.ir
مؤلف: محمد دهقانی		
ناشر: سنجش و دانش		
نوبت چاپ: چاپ اول - ۱۴۰۰		
تیراژ: ۳۰ نسخه		
شابک: ۹۷۸-۶۲۲-۰۵-۲۸۲۱-۰۰		قیمت: ۲۵۰۰۰۰ ریال
نشانی: میدان انقلاب . خیابان انقلاب . خیابان دانشگاه . بلاک ۱۳۶ . تلفن: ۰۲۱۶۱۲۶		
«کلیه حقوق مادی و معنوی این اثر محفوظ می باشد»		

کتاب حاضر، حاصل پایان نامه کارشناسی ارشد نگارنده در رشته فناوری اطلاعات است که در سال ۱۴۰۰ در دانشکده مهندسی صنایع و سیستم‌ها دانشگاه تربیت مدرس به راهنمایی سرکار خانم دکتر آخوندزاده از آن دفاع شده است.

www.ketab.ir

چکیده

تحلیل عواطف فرایندی است که به تحلیل عواطف و نظرات کاربران می‌پردازد و از این طریق، اولویت افراد را شناسایی می‌کند. تاکنون تحقیقات زیادی در مورد تشخیص قطبیت نظرات در نظر کاوی و تجزیه و تحلیل عواطف در زبان‌هایی مثل انگلیسی و چینی انجام شده است اما در زبان فارسی به دلیل کمبود ابزارهایی برای پردازش زبان طبیعی این تحقیقات محدودتر انجام شده است و به دلیل اهمیت روزافزون این بحث در دنیای تجارت، سیاست، معضلات اجتماعی و غیره تمرکز این پژوهش در این حوزه بنا شده است. امروزه و در سطح جهانی کاربران، از شبکه‌های اجتماعی مانند توئیتر به‌صورت مداوم استفاده می‌کنند. اغلب افراد در مواجهه با پدیده‌های مختلف اطرافشان، توئیتر را بستری مناسب برای اظهار نظر انتخاب می‌کنند. در این پژوهش به کمک دو مجموعه داده برچسب دار فارسی، برآیند کلی نظرات افراد مختلف در توئیتر حول عملکرد حاکمیت به کمک تکنیک‌های یادگیری ماشین بررسی می‌شود. همچنین برای جامع‌تر شدن کار، از روش‌های قدیمی‌تر کدگذاری کلمات مثل `bag of words` و از روش‌های جدیدتر مثل مدل زبانی و `word embedding` استفاده می‌شود و نتایج هر دو روش باهم مقایسه می‌شود. در این پژوهش از تکنیک‌های جدید مثل شبکه‌های عصبی و `fastText`، `word2vec` و `bert` استفاده شده که بالاترین صحت مربوط به مدل `bert` است. با توجه به عملکرد مناسبی که مدل‌ها روی دیتاست‌ها داشتند می‌توان از آن‌ها در دنیای واقعی استفاده کرد. به‌صورتی که داده‌های جدید و تصادفی توئیتر فارسی را حول موضوع عملکرد حاکمیت دریافت کرده و به کمک مدلی که قبلاً آموزش داده بودیم بتوانیم برچسب آن‌ها را به‌صورتی پیش‌بینی کنیم. در چنین شرایطی از کنار هم قراردادن برچسب‌های پیش‌بینی شده به یک جمع‌بندی کلی از نظرات افراد مختلف نسبت به عملکرد حاکمیت رسید.

کلیدواژه: تحلیل عواطف، یادگیری ماشین، نظر کاوی، پردازش زبان طبیعی، مدل زبانی.

فہرست مطالب

۱. مقدمات و کلیات ۱

۱-۱. مقدمه ۲

۱-۲. بیان مسئله تحقیق ۲

۱-۳. اهمیت و ضرورت تحقیق ۳

۱-۴. اهداف تحقیق ۴

۱-۵. مفروضات تحقیق ۵

۱-۶. سؤالات تحقیق ۵

۱-۷. بررسی ادبیات موضوع تحقیق ۵

۱-۸. جنبه نوآوری تحقیق ۶

۱-۹. روش انجام تحقیق ۶

۱-۱۰. طرح کلی تحقیق ۶

۱-۱۱. خلاصه فصل ۸

۲. مروری بر ادبیات موضوع ۹

۲-۱. مقدمه ۱۰

۲-۲. تعاریف ۱۰

۲-۲-۱. موجودیت ۱۰

۲-۲-۲. ۱۱

۲-۲-۳. عاطفه و همدلی ۱۱

۲-۲-۴. ۱۱

۲-۲-۵. ۱۲

۲-۲-۶. ۱۲

۲-۳. ۱۲

۲-۳-۱. ۱۲

۲-۳-۲. ۱۲

۲-۳-۳. ۱۲

۲-۳-۴. ۱۲

۲-۳-۵. ۱۲

۲-۳-۶. ۱۲

۲-۳-۷. ۱۲

۲-۳-۸. ۱۲

۲-۳-۹. ۱۲

۲-۳-۱۰. ۱۲

۲-۳-۱۱. ۱۲

۲-۳-۱۲. ۱۲

۲-۳-۱۳. ۱۲

۲-۳-۱۴. ۱۲

۲-۳-۱۵. ۱۲

۲-۳-۱۶. ۱۲

۲-۳-۱۷. ۱۲

۲-۳-۱۸. ۱۲

۲-۳-۱۹. ۱۲

۲-۳-۲۰. ۱۲

۲-۳-۲۱. ۱۲

۲-۳-۲۲. ۱۲

۲-۳-۲۳. ۱۲

۲-۳-۲۴. ۱۲

۲-۳-۲۵. ۱۲

۲-۳-۲۶. ۱۲

۲-۳-۲۷. ۱۲

۲-۳-۲۸. ۱۲

۲-۳-۲۹. ۱۲

۲-۳-۳۰. ۱۲

۲-۳-۳۱. ۱۲

۲-۳-۳۲. ۱۲

۲-۳-۳۳. ۱۲

۲-۳-۳۴. ۱۲

۲-۳-۳۵. ۱۲

۲-۳-۳۶. ۱۲

۲-۳-۳۷. ۱۲

۲-۳-۳۸. ۱۲

۲-۳-۳۹. ۱۲

۲-۳-۴۰. ۱۲

۲-۳-۴۱. ۱۲

۲-۳-۴۲. ۱۲

۲-۳-۴۳. ۱۲

۲-۳-۴۴. ۱۲

۲-۳-۴۵. ۱۲

۲-۳-۴۶. ۱۲

۲-۳-۴۷. ۱۲

۲-۳-۴۸. ۱۲

۲-۳-۴۹. ۱۲

۲-۳-۵۰. ۱۲

۲-۳-۵۱. ۱۲

۲-۳-۵۲. ۱۲

۲-۳-۵۳. ۱۲

۲-۳-۵۴. ۱۲

۲-۳-۵۵. ۱۲

۲-۳-۵۶. ۱۲

۲-۳-۵۷. ۱۲

۲-۳-۵۸. ۱۲

۲-۳-۵۹. ۱۲

۲-۳-۶۰. ۱۲

۲-۳-۶۱. ۱۲

۲-۳-۶۲. ۱۲

۲-۳-۶۳. ۱۲

۲-۳-۶۴. ۱۲

۲-۳-۶۵. ۱۲

۲-۳-۶۶. ۱۲

۲-۳-۶۷. ۱۲

۲-۳-۶۸. ۱۲

۲-۳-۶۹. ۱۲

۲-۳-۷۰. ۱۲

۲-۳-۷۱. ۱۲

۲-۳-۷۲. ۱۲

۲-۳-۷۳. ۱۲

۲-۳-۷۴. ۱۲

۲-۳-۷۵. ۱۲

۲-۳-۷۶. ۱۲

۲-۳-۷۷. ۱۲

۲-۳-۷۸. ۱۲

۲-۳-۷۹. ۱۲

۲-۳-۸۰. ۱۲

۲-۳-۸۱. ۱۲

۲-۳-۸۲. ۱۲

۲-۳-۸۳. ۱۲

۲-۳-۸۴. ۱۲

۲-۳-۸۵. ۱۲

۲-۳-۸۶. ۱۲

۲-۳-۸۷. ۱۲

۲-۳-۸۸. ۱۲

۲-۳-۸۹. ۱۲

۲-۳-۹۰. ۱۲

۲-۳-۹۱. ۱۲

۲-۳-۹۲. ۱۲

۲-۳-۹۳. ۱۲

۲-۳-۹۴. ۱۲

۲-۳-۹۵. ۱۲

۲-۳-۹۶. ۱۲

۲-۳-۹۷. ۱۲

۲-۳-۹۸. ۱۲

۲-۳-۹۹. ۱۲

۲-۳-۱۰۰. ۱۲

۲-۳-۱۰۱. ۱۲

۲-۳-۱۰۲. ۱۲

۲-۳-۱۰۳. ۱۲

۲-۳-۱۰۴. ۱۲

۲-۳-۱۰۵. ۱۲

۲-۳-۱۰۶. ۱۲

۲-۳-۱۰۷. ۱۲

۲-۳-۱۰۸. ۱۲

۲-۳-۱۰۹. ۱۲

۲-۳-۱۱۰. ۱۲

۲-۳-۱۱۱. ۱۲

۲-۳-۱۱۲. ۱۲

۲-۳-۱۱۳. ۱۲

۲-۳-۱۱۴. ۱۲

۲-۳-۱۱۵. ۱۲

۲-۳-۱۱۶. ۱۲

۲-۳-۱۱۷. ۱۲

۲-۳-۱۱۸. ۱۲

۲-۳-۱۱۹. ۱۲

۲-۳-۱۲۰. ۱۲

۲-۳-۱۲۱. ۱۲

۲-۳-۱۲۲. ۱۲

۲-۳-۱۲۳. ۱۲

۲-۳-۱۲۴. ۱۲

۲-۳-۱۲۵. ۱۲

۲-۳-۱۲۶. ۱۲

۲-۳-۱۲۷. ۱۲

۲-۳-۱۲۸. ۱۲

۲-۳-۱۲۹. ۱۲

۲-۳-۱۳۰. ۱۲

۲-۳-۱۳۱. ۱۲

۲-۳-۱۳۲. ۱۲

۲-۳-۱۳۳. ۱۲

۲-۳-۱۳۴. ۱۲

۲-۳-۱۳۵. ۱۲

۲-۳-۱۳۶. ۱۲

۲-۳-۱۳۷. ۱۲

۲-۳-۱۳۸. ۱۲

۲-۳-۱۳۹. ۱۲

۲-۳-۱۴۰. ۱۲

۲-۳-۱۴۱. ۱۲

۲-۳-۱۴۲. ۱۲

۲-۳-۱۴۳. ۱۲

۲-۳-۱۴۴. ۱۲

۲-۳-۱۴۵. ۱۲

۲-۳-۱۴۶. ۱۲

۲-۳-۱۴۷. ۱۲

۲-۳-۱۴۸. ۱۲

۲-۳-۱۴۹. ۱۲

۲-۳-۱۵۰. ۱۲

۲-۳-۱۵۱. ۱۲

۲-۳-۱۵۲. ۱۲

۲-۳-۱۵۳. ۱۲

۲-۳-۱۵۴. ۱۲

۲-۳-۱۵۵. ۱۲

۲-۳-۱۵۶. ۱۲

۲-۳-۱۵۷. ۱۲

۲-۳-۱۵۸. ۱۲

۲-۳-۱۵۹. ۱۲

۲-۳-۱۶۰. ۱۲

۲-۳-۱۶۱. ۱۲

۲-۳-

۱۶	۴-۲ معرفی الگوریتم‌ها
۱۶	۱-۴-۲ درخت تصمیم
۱۷	۲-۴-۲ ماشین بردار پشتیبان
۱۹	۳-۴-۲ جنگل تصادفی
۱۹	۴-۴-۲ یادگیری عمیق
۲۰	۵-۴-۲ شبکه‌های عصبی
۲۱	۶-۴-۲ پرسپترون
۲۳	۷-۴-۲ تابع فعال‌سازی
۲۵	۸-۴-۲ تابع زیان
۲۶	۹-۴-۲ انتشار به عقب
۲۷	۱۰-۴-۲ الگوریتم بهینه‌سازی Adam
۲۸	۱۱-۴-۲ شبکه‌های عصبی بازگشتی
۳۰	۱۲-۴-۲ حافظه طولانی کوتاه مدت
۳۲	۱۳-۴-۲ حافظه طولانی کوتاه مدت دوسویه
۳۲	۱۴-۴-۲ بیش برارس
۳۴	۵-۲ روش‌های استخراج ویژگی و نمایش متن‌ها
۳۴	۱-۵-۲ one-hot نمایش
۳۴	۶-۲ TF-IDF روش
۳۵	۲-۷-۲ Word2vec
۳۶	۳-۷-۲ fastText
۳۷	۴-۷-۲ bert
۳۷	۸-۲ شبکه‌های ترانسفورم
۳۸	۱۰-۸-۲ فصل

۲۹	۲-۸-۲. دقت
۳۹	۳-۸-۲. بادآوری
۳۹	۴-۸-۲. امتیاز F1
۳۹	۵-۸-۲. روش اعتبارسنجی متقابل
۴۰	۹-۲. توینر
۴۱	۱۰-۹-۲. بررسی پژوهش‌های مربوط به توینر
۴۵	۱۰-۲. روش‌ها
۴۶	۱-۱۰-۲. یادگیری ماشین
۴۷	۲-۱۰-۲. مبتنی بر واژه
۴۹	۳-۱۰-۲. مبتنی بر گراف
۵۰	۴-۱۰-۲. ترکیبی
۵۱	۱۱-۲. بررسی سیستم پژوهش برای مون آنکلیسی
۵۲	۱۲-۲. بررسی سیستم پژوهش برای مون فارسی
۵۵	۱۳-۲. جمع‌بندی مرور ادبیات
۵۸	۱۴-۲. خلاصه فصل
۵۹	۳. روش‌شناسی تحقیق
۶۰	۱-۳. مقدمه
۶۰	۲-۳. ابزار جمع‌آوری داده
۶۱	۳-۳. ساختار کلی روش‌شناسی پژوهش
۶۲	۴-۳. ابزارهای پردازش زبانی
۶۲	۱-۴-۳. تحلیل سازه
۶۳	۲-۴-۳. اعتبار پایایی
۶۴	۳-۴-۳. نوکتاب‌بر
۶۴	۴-۴-۳. بسط‌پذیری موجودیت نامدار

۶۴	۵-۴-۳ ریشه‌یاب
۶۴	۶-۴-۳ قطعه بند
۶۵	۷-۴-۳ بار سرب با تجزیه گر جملات
۶۵	۸-۴-۳ تصحیح خطای املایی
۶۵	۹-۴-۳ برجسب گذار اجزای واژگانی کلام
۶۵	۵ ۳ بیش بردارس منون
۶۵	۶-۳ استخراج ویژگی
۶۶	۷ ۳ انتخاب مدل
۶۶	۸-۳ ابزارها
۶۶	۹-۳ خلاصه فصل
۶۸	۴. نتایج انجام پژوهش
۶۹	۱. ۴ مقدمه
۶۹	۲. ۴ معرفی داده‌ها
۷۳	۳-۴. مأمور سازی ویژگی های استخراج شده
۷۵	۴. ۴ بیش برداری
۷۵	۱-۴-۴. شبکه‌ها
۷۶	۲-۴-۴. آدرس های اینترنتی
۷۷	۳-۴-۴. حذف فضاهای خالی اضافی
۷۷	۴-۴-۴. موجودیت های HTML
۷۸	۵-۴-۴. دنباله ای از کاراکترهای تکرار شده
۷۸	۲. توسعه مدل متن محلی و ارزیابی سطح
۷۸	۱-۵-۴. فاز اول توسعه مدل
۷۹	۲. ۵-۴. فاز دوم توسعه مدل
۸۱	۶. ۴ تحلیل نتایج

۸۳	۷-۴. خلاصه فصل
۸۴	۵. نتیجه‌گیری و پیشنهادها
۸۵	۱-۵. مقدمه
۸۵	۲-۵. تحلیل نتایج تحقیق
۸۵	۳-۵. بررسی نتایج فاز اول
۸۵	۳-۵-۱. استخراج ویژگی
۸۵	۳-۵-۲. انتخاب مدل
۸۷	۴-۵. بررسی نتایج فاز دوم
۸۷	۴-۵-۱. استخراج ویژگی
۸۷	۴-۵-۲. انتخاب مدل
۸۸	۵-۵. پاسخ به سؤالات تحقیق
۸۹	۶-۵. ارائه پیشنهادهایی برای تحقیقات آتی
۹۰	۷-۵. منابع و محدودیت‌های تحقیق
۹۰	۸-۵. خلاصه فصل
۹۱	پیوست
۹۸	۷. منابع و مراجع