

شناسایی هیجان در متن فارسی با استفاده از مدل یادگیری ماشینی

Emotion Detection in Persian Text; A Machine Learning Model

Ali Khosravi^{۱*}, Manoochehr Kelarestaghi^۲, Mehdi Purmohammad^۳علی خسروی^{۱*}، منوچهر کلارستاقی^۲، مهدی پورمحمد^۳

1. Ph. D. in Cognitive Modeling, Institute of Cognitive Science Studies (ICSS), Pardis
2. Assistant Professor, Department of Electrical and Computer Engineering, Kharazmi University, Tehran
3. Assistant professor, Department of Cognitive Linguistics Institute for Cognitive Science Studies (ICSS), Pardis

- ۱- دانش آموخته دکتری مدلسازی شناختی، موسسه آموزش عالی علوم شناختی، پردیس
- ۲- استادیار، گروه برق و مهندسی کامپیوتر، دانشگاه خوارزمی، تهران
- ۳- استادیار، گروه موسسه آموزش عالی علوم شناختی، پردیس

Abstract

This study aimed to develop a computational model for recognition of emotion in Persian text as a supervised machine learning problem. We considered Plutchik emotion model as supervised learning criteria and Support Vector Machine (SVM) as baseline classifier. We also used NRC lexicon and contextual features as training data and components of the model. One hundred selected texts including political-social newspaper editorials were used to test the model (real terms). Also in this study, the "support vector machine" algorithm was used as the learning classifier and four indicators of accuracy, accuracy, f-score and recall were used to evaluate the model. The results show that the efficiency of the model in detecting different emotions varies from 79% to 98% and mean precision of the model for all classes was 84%. Using all indexes, the classifier showed more performance in joy category than other 7 types. The results of this study show that using emotion-based approach, supervised learning and minimal contextual features can be useful in automatic identification of emotions. It also showed that a combination of lexical resource and contextual features can be used as learning base for a SVM model.

Keywords:

Emotion Detection, Text Analysis, Machine Learning

چکیده

پژوهش حاضر به منظور ارائه‌ی یک مدل محاسباتی برای بازشناسی هیجان در متن فارسی صورت گرفته است. در مدل پیشنهادی تحقیق، گروه‌بندی هشت‌گانه‌ی پلاچیک در هیجان‌ها به‌عنوان مبنای یادگیری نظارت‌شده، ویژگی‌های باقی متن فارسی و لیست واژگان هیجانی برچسب‌دار به‌عنوان عناصر و اجزای مدل تعیین شده اند. برای آزمون مدل (شرایط واقعی) از صد متن منتخب شامل سرمقاله‌ی روزنامه‌های سیاسی- اجتماعی استفاده شده است. همچنین در این تحقیق از الگوریتم «ماشین بردار پشتیبان» به‌عنوان طبقه‌بند یادگیرنده استفاده گردیده و چهار شاخص دقت، درستی، ضریب F و بازخوانی برای ارزیابی مدل بکار گرفته شده‌اند. نتایج تحقیق نشان می‌دهند که کارایی مدل (دقت) در شناسایی هیجان‌های مختلف از ۷۹ درصد تا ۹۸ درصد متغیر بوده و در ارزیابی کلی مدل، میانگین دقت مدل معادل ۸۴ درصد است. بر اساس سایر شاخص‌ها، می‌توان گفت که بیشترین نرخ درستی مربوط به هیجان خوشی و کمترین آن مربوط به هیجان خشم است. همچنین تطبیق قابل‌توجهی میان شاخص F با شاخص «بازخوانی» وجود دارد و بیشترین مقدار آن در گروه هیجان شادی دیده می‌شود. نتایج این تحقیق نشان می‌دهد استفاده از رهیافت مبتنی بر گروه‌بندی هیجان‌ها، یادگیری نظارت‌شده و نیز ویژگی‌های بافتی حلقه‌ای در متن می‌تواند کارایی مناسبی در شناسایی خودکار هیجان‌ها داشته و می‌توان تلفیقی از گروه‌های اصلی هیجان را به‌منظور افزایش کارایی مدل یادگیری بکار گرفت.

کلیدواژه‌ها:

شناسایی هیجان، تحلیل متن، یادگیری ماشینی

مصنوعی به‌ویژه طراحی رابط کاربری انسان- ماشین مورد توجه قرار گرفته است (رامالینگام، ۲۰۱۸). در این زمینه، روابط متقابل بین متن و هیجان‌ها و به‌ویژه مسائلی مانند چگونگی تأثیر هیجانی متن بر خوانندگان و یا نحوه‌ی بازتاب هیجان نویسنده از موضوعات پژوهشی مهم بوده و حجم قابل‌توجهی از مقالات در این زمینه وجود دارد (مک

مقدمه

هیجان‌ها پدیده‌های پیچیده‌ای هستند که با تغییرات فیزیکی و روان‌شناختی همراه بوده و موجب تأثیر بر افکار و رفتارهای فرد می‌شوند. آن‌ها دربرگیرنده برانگیختگی فیزیولوژیک، رفتارهای معین و تجربیات آگاهانه هستند (میرز، ۲۰۰۴). هیجان‌ها از نخستین مراحل پیدایش روانشناسی علمی تاکنون به‌عنوان یکی از ابعاد مهم رفتار انسانی مورد بررسی بوده‌اند. در تحقیقات اخیر موضوع شناسایی هیجان در متن به نحو فزاینده‌ای در طراحی سیستم‌های هوش

1. Classifier
2. Human-Machine Interface

* نویسنده مسئول: علی خسروی

نشانی: پردیس، موسسه آموزش عالی علوم شناختی، گروه مدلسازی شناختی

تلفن: ۰۲۱-۸۸۹۰۷۶۸۹ (+۹۸)

پست الکترونیک: ali.khosravi.mail@gmail.com

می‌کنند. رهیافت محاسباتی در شناسایی هیجان‌ها در متن عمدتاً بر مبنای یادگیری ماشینی و مدل‌های مربوط به آن است. با وجود اهمیت تهیه‌ی لیست واژگان هیجانی، باید توجه داشت که کارایی سیستم‌هایی که صرفاً مبتنی بر واژگان هستند به کیفیت و جامعیت لیست واژگان بستگی دارد (نویاروسکایا، پردینگر و ایشیزوکا، ۲۰۱۱). در بررسی‌های اخیر پژوهش‌های متعددی به‌منظور تهیه‌ی مجموعه‌ها یا پیکره‌های زبانی حاشیه‌نویسی^۴ شده و همچنین تهیه لیست واژگان برچسب‌گذاری شده در زبان‌های مختلف صورت گرفته است. به‌عنوان نمونه این واژه‌نامه‌ها می‌توان به LIWC (پنباکر و همکاران، ۲۰۰۱)، LEW List (فرانسیسکو و همکاران، ۲۰۱۰) و WordNet Affect (استراپاراوا و والیتوتی، ۲۰۰۴) اشاره کرد. در این واژه‌نامه‌ها، به‌جای معنی هر واژه، نوع هیجان مربوط به آن واژه مشخص شده است. البته هریک از واژه‌نامه‌ها مشخصات گوناگونی را نشان داده و در عین حال محدودیت‌های خاص خود را دارند. به‌عنوان مثال، برخی از آن‌ها صرفاً هیجان‌های مربوط به واژه را نشان داده و امکان تفکیک بین معانی مختلف یک واژه را ایجاد نمی‌کنند.

همان‌گونه که اشاره شد، علاوه بر سطح واژگان، تحلیل هیجانی متن می‌تواند در سطوح گوناگون اعم از سطح جمله یا در سطح کل متن یا سند صورت گیرد (فراه، ۲۰۱۰)؛ بنابراین، می‌توان برچسب‌گذاری و حاشیه‌نویسی را در سطح واژه، جمله و یا کل متن انجام داد و از آن به‌عنوان منبع یادگیری برای مدل استفاده کرد. پیکره‌های زبانی^۵ به‌صورت مجموعه‌ای از متون برچسب‌گذاری شده و حاشیه‌نویسی شده به‌عنوان منبع یادگیری برای الگوریتم‌ها و مدل‌های محاسباتی استفاده می‌شوند. در زبان انگلیسی مجموعه‌ها یا پیکره‌های زبانی متعددی وجود دارند که بر اساس نوع یا ماهیت هیجان حاشیه‌نویسی شده‌اند. متن‌های مربوط به این مجموعه‌ها می‌توانند از جمله، عناوین خبری، جملات یا پاراگراف‌های توثیق شده در شبکه‌های اجتماعی و یا متن‌های طولانی‌تر مانند سرمقالات روزنامه‌ها یا مجلات و همچنین اشعار و متون ادبی انتخاب شوند. همچنین برخی از این مجموعه‌های حاشیه‌نویسی شده در یک گروه موضوعی و تخصصی خاص (مانند خلاصه داستان فیلم‌های سینمایی اکشن) قرار می‌گیرند که می‌توانند به‌عنوان منبع یادگیری برای سیستم‌های اختصاصی استفاده شوند. به‌عنوان نمونه می‌توان به بانک اطلاعات متون احساسی (AffectiveText Dataset) اشاره کرد (استراپاراوا و میهالسی، ۲۰۰۷). این مجموعه دربرگیرنده ۱۲۵۰ عنوان خبری (جمله) است که بر اساس گروه‌های هیجانی و میزان کشش حاشیه‌نویسی شده‌اند. مجموعه‌ی Blogs که توسط آمان و

کیه، ۲۰۱۱). استفاده از کامپیوترها این امکان را فراهم کرده است که حجم زیادی از اطلاعات متنی به شیوه‌ای نظام‌مند تحلیل شود. نرم‌افزارهای متعددی در این زمینه طراحی شده‌اند. شناسایی هیجان‌ها در متن به‌ویژه متونی که در شبکه‌های اجتماعی یا اینترنت قرار می‌گیرند منبع بالارزشی از اطلاعات است چراکه نشان می‌دهد افراد مختلف چگونه نسبت به موضوعات یا شرایط، واکنش نشان می‌دهند (گیند، ۲۰۱۹). به‌عنوان مثال، موتورهای جستجوی امروزی قادرند اطلاعات عاطفی-هیجانی در متن ترانه‌ها را بررسی کرده و مطابق با وضعیت خلقی^۱ شنوندگان، آن‌ها را انتخاب کنند (چو و لی، ۲۰۰۶). بر اساس نظریه‌های موجود، شناسایی هیجان‌ها در متن می‌تواند با استفاده از نشانه‌های متنی انجام شود. به مجموعه‌ی نشانه‌ها و ساختارهای مرتبط با هیجان «زبان هیجان» گفته می‌شود که دارای دو عنصر اصلی نشانه‌های مشهود و نشانه‌های ضمنی است. نشانه‌های مشهود شامل واژه‌ها یا نشانه‌های نوشتاری هستند که مستقیماً به یک هیجان مربوط می‌شوند. می‌توان واژه‌های هیجانی را به دو گروه دیگر شامل واژه‌های مستقیم (گویا) و غیرمستقیم (توصیفی) نیز طبقه‌بندی کرد. علاوه بر واژگان هیجانی، استعاره‌ها، نشانه‌های آوایی و تکواژشناسی نیز پدیده‌های زبانی هستند که هیجان در آن‌ها بازنمایی شده و قابل شناسایی است (ویلس، ۲۰۰۹).

شناسایی هیجان در متن با استفاده از تکنیک‌های گوناگون علوم کامپیوتر و زبان‌شناسی محاسباتی^۲ به‌ویژه پردازش زبان طبیعی^۳ به ارائه‌ی سیستم‌های مؤثری در پردازش متون منجر شده و امروزه محیط‌های نرم‌افزاری متعددی برای پردازش ماشینی و خودکار متن در زبان‌های گوناگون ارائه شده است. تفاوت در ابعاد زبان (شامل ابعاد گفتاری، ساختارهای نحوی، سازه‌های معنایی، استعاره‌ها و بسیاری موارد دیگر) باعث شده است تا شیوه‌های پردازشی برای سنجش و شناسایی هیجان در متن برای زبان‌های مختلف تفاوت‌های اساسی داشته باشد. درواقع، هرگونه پژوهش درزمینه‌ی شناسایی ماشینی هیجان در متن، دربرگیرنده‌ی حداقل دو مبنای نظری شامل تئوری‌های هیجان و رهیافت‌های پردازش زبان طبیعی است. پژوهش‌ها نشان می‌دهند که می‌توان ترکیبی از ویژگی‌های واژگانی، معنایی و نحوی را برای یادگیری مدل در تفکیک متن هیجانی و غیرهیجانی بکار گرفت (آمان، ۲۰۰۷). برای این منظور مدل‌های محاسباتی از ویژگی‌های زمینه‌ای به‌عنوان منابع یادگیری و از پیکره‌های زبانی، منابع واژگان هیجانی، بانک‌های واژگان و با توجه به اهداف تحقیق از مدل‌های آماری و یا ریاضی مبتنی بر طبقه‌بندی خودکار استفاده

4. Annotated
5. Linguistic corpus

1. Mood
2. Computational linguistics
3. Natural Language Processing

روزنامه‌های کثیرالانتشار فارسی در طی دوره‌ی زمانی ۹۵ تا ۹۷ بوده است. برای این کار از سرمقاله‌های روزنامه‌های کیهان، ابتکار، همشهری و جوان استفاده شده است. انتخاب متون به صورت تصادفی بوده و در این متن‌ها با توجه به رویه‌ی نوشتاری در این گونه نشریات از ۵۵۰ تا ۱۴۰۰ واژه استفاده شده است. برای ایجاد برچسب‌ها و حاشیه‌نویسی متون (معیار پایه) از پرسشنامه استفاده شده است. این پرسشنامه‌ها توسط ۱۰ نفر با مشخصات زیر حاشیه‌نویسی شده است:

- خواننده‌ی روزنامه‌های کثیرالانتشار (حداقل یک روزنامه به‌طور منظم از طریق اینترنت یا روزنامه چاپی)
- دامنه‌ی سنی بین ۲۰ تا ۴۵
- جنسیت مذکر
- فارسی‌زبان (به‌عنوان زبان مادری)

انتخاب افراد (حاشیه‌نویسان) و تعداد آن‌ها بر اساس روش‌شناسی تحقیقات مشابه با استفاده از داوطلبان انجام شده است (نوویلی، ۲۰۱۸). در پرسشنامه‌ها از مخاطبین خواسته شده است تا احساس خود نسبت به متن خوانده‌شده را در قالب هشت گروه هیجانی مشخص کنند.

برای ارزیابی توافق بین حاشیه‌نویس‌ها از ضریب کاپا استفاده شده است. ضریب کاپا میزان توافق بین دو حاشیه‌نویس (یا هر رتبه‌گذار) را در مقایسه با وضعیت انتخاب تصادفی نشان می‌دهد. از این ضریب برای تعیین قابلیت اعتماد برای هر عملیات گروه‌بندی یا رتبه‌دهی استفاده می‌شود. ضریب کاپا در این تحقیق $0/61$ محاسبه شده که بر اساس رتبه‌بندی به‌عنوان «خوب» یا «مناسب» تلقی می‌شود (دی ملو و گراسر، ۲۰۰۷).

بخش دوم داده‌های ورودی سیستم شامل لیست واژگان دارای برچسب هیجانی است که در این تحقیق از واژه‌نامه استاندارد تهیه‌شده توسط NRC بنام Word-Emotion Association Lexicon استفاده شده است. این واژه‌نامه فارسی به صورت فایل اکسل دربرگیرنده ۱۴۱۸۲ واژه و عبارت است که در قالب گروه‌های هشت‌گانه هیجان (شامل خشم، انتظار، تنفر، ترس، خوشی، غم، تعجب و اعتماد) و نیز بر اساس قطبیت (مثبت و منفی یا خنثی) برچسب‌گذاری شده‌اند. این واژه‌نامه در سال ۲۰۱۶ بر اساس طبقه بندی فوق به صورت دستی و از طریق میانگین‌گیری برچسب‌گذاری شده است.^۲

سومین بخش از اطلاعات مورداستفاده تحقیق در این تحقیق که توسط مدل در فرآیند یادگیری ماشینی استفاده می‌شود، لیست ویژگی

شپاکوویچ (۲۰۰۷) تدوین شده شامل ۵۲۰۵ جمله از ۱۷۳ متن وبلاگ است. این متن‌ها بر اساس گروه‌بندی شش‌گانه اکمن (به انضمام یک گروه با برچسب فاقد هیجان) حاشیه‌نویسی شده‌اند. در پژوهش دیگری، یک مجموعه تحت عنوان EmoBank تهیه شده است که در آن بیش از ده هزار جمله در موضوعات و زمینه‌های مختلف به صورت جداگانه حاشیه‌نویسی شده و طبقه‌بندی شده‌اند (هان و بوچل، ۲۰۱۷).

مسئله‌ی تحقیق حاضر، ارائه‌ی مشخصات مدلی محاسباتی است که در قالب یک الگوریتم ماشینی توانایی شناسایی گروه‌های هیجانی و تفکیک متن بر اساس بار هیجانی را داشته باشد. بر این اساس، مشخصات مدل شامل نوع اطلاعات ورودی، شیوه‌ی پردازش داده‌ها و شیوه‌ی یادگیری مدل می‌بایست در این تحقیق با توجه به ویژگی‌های متن فارسی تعیین شوند. در این مدل، ابعاد شناختی هیجان در تلفیق با ویژگی‌های مختلف ساختاری متن در یک مدل ریاضی تلفیق شده و کارایی سیستم با استفاده از روش‌های ماشینی و غیر ماشینی ارزیابی می‌شود.

حجم محتوای فارسی در محیط اینترنت رشد فزاینده‌ای داشته است و هم‌اکنون برآورد می‌شود که حدود $1/8$ درصد از کل محتوای اینترنت را در بر می‌گیرد.^۱ در این برآورد کتابخانه‌های دیجیتالی خارج از وب، محتواهای خاص شبکه‌های محلی و نیز کانال‌های مربوط به شبکه‌های اجتماعی لحاظ نشده‌اند. با توسعه‌ی محتوای دیجیتالی فارسی، وجود سیستم‌های شناسایی هیجان و احساسات متناسب با ویژگی‌های زبان فارسی در این محیط‌ها به نحو فزاینده‌ای مورد توجه قرار گرفته است. این گونه پژوهش‌ها می‌توانند کاربردهای فراوانی در تحلیل متن، تحلیل و بررسی محتوای شبکه‌های اجتماعی و صفحات اینترنت، نظرسنجی‌های آنلاین، رفتارشناسی، تحلیل واکنش‌ها یا آسیب‌شناسی‌های روانی-اجتماعی و موارد مشابه داشته باشند.

روش

با توجه به هدف پژوهش یعنی ارائه‌ی مدل یادگیری ماشینی برای شناسایی گروه‌های هیجان در متن، مراحل اصلی تحقیق به ترتیب انتخاب مدل یادگیری پایه، طراحی الگوریتم، تعیین اطلاعات ورودی و تعیین پردازش اطلاعات در مدل محاسباتی بوده است. علاوه بر آن می‌بایست آزمون‌های آماری و آزمایش‌های لازم برای اصلاح مدل نیز طراحی شوند.

داده‌ها

داده‌ها و اطلاعات مورداستفاده در این تحقیق شامل سه بخش اصلی است. بخش اول شامل متون مربوط به مقالات اجتماعی

2. <https://github.com/mhbashari/NRC-Persian-Lexicon>

1. <https://w3techs.com>

یادگیری و بخشی دیگر به‌عنوان داده‌ی تست (با فرض داده‌ی واقعی) استفاده می‌شوند. سپس همین فرآیند برای داده‌ها دوباره تکرار شده و بخش دیگر به‌عنوان یادگیری و بقیه‌ی داده‌ها به‌عنوان تست استفاده می‌شوند. معمولاً در این روش‌ها تعداد تکرارها ۱۰ بار است اما می‌توان از تکرارهای دیگری نیز استفاده کرد. درنهایت از دقت‌های به‌دست‌آمده، میانگین گرفته شده و به‌عنوان دقت نهایی بیان می‌شود. طراحی الگوریتم و کد نویسی آن به زبان پایتون صورت گرفته و برای اجرای کدها و تهیه خروجی از محیط بسته‌ی نرم‌افزاری Anaconda navigator 1.8.7 و ماژول Jupyter استفاده شده است.

ارزیابی مدل

در این پژوهش سه آزمون و شاخص آماری مورد استفاده قرار گرفته است که هر کدام در یکی از مراحل اصلی و به‌منظور اطمینان از صحت و کارایی روش‌ها مورد استفاده قرار می‌گیرند.

در ارزیابی مدل از چهار سنجه‌ی اصلی در ارزیابی مدل‌های یادگیری ماشینی یعنی درستی^۴، دقت^۵، بازخوانی^۶ و ضریب f استفاده شده است.

«درستی» درصد کل مشاهداتی است که به‌درستی طبقه‌بندی شده‌اند. «درستی» به‌صورت زیر محاسبه می‌شود:

$$\text{Accuracy} = \frac{TP+TN}{TP+TN+FP+FN} \quad (1)$$

در فرمول (۱)،

TP شامل شناسایی‌های درست، TN شامل حذف‌های درست، FP شناسایی نادرست و FN حذف‌های نادرست است.

به نسبت بین تمام مشاهداتی که به‌درستی طبقه‌بندی شده‌اند (TP) به تمام مشاهدات طبقه‌بندی‌شده مثبت (TP+FP)، «دقت» گفته می‌شود (فرمول شماره ۲).

$$\text{Precision} = \frac{TP}{TP+FP} \quad (2)$$

«بازخوانی»، نسبت بین مشاهدات به‌درستی طبقه‌بندی شده به تمام مشاهدات مثبت است. به این سنجه، حساسیت و یا نرخ تأیید درست هم گفته می‌شود (فرمول شماره ۳).

$$\text{Recall} = \frac{TP}{TP+FN} \quad (3)$$

های متنی (زمینه‌ای) است. ویژگی‌های زمینه‌ای را می‌توان در گروه‌های متعددی مانند معنایی، نحوی، واژگانی و غیره طبقه‌بندی کرد. برخی از این ویژگی‌ها به معنا و مفهوم و یا نقش واژه در جمله (مانند صفت، قید و غیره) مربوط می‌باشند و برخی دیگر به روابط بین واژگان در یک جمله (مانند واژه‌های مرتبط یا تکرار واژگان خاص در یک جمله) و همچنین این ویژگی‌ها می‌توانند به تعداد حروف، وجود واژگان خاص (مانند منفی‌سازها) و سایر موارد مرتبط باشند (کوان، ۲۰۱۰).

بررسی ویژگی‌های مزبور در متن فارسی و ارزیابی تأثیر هر گروه از ویژگی‌ها نیازمند تحقیقات متعددی است و در این تحقیق با توجه به اهداف آن مجموعه‌ای از این ویژگی‌های متنی (جدول ۱) در نظر گرفته شده است.

جدول ۱. لیست ویژگی‌های مورد استفاده در مدل یادگیری

ردیف	عنوان	شیوه‌ی محاسبه
۱	تعداد واژگان مثبت	تعداد کل واژه‌های دارای برچسب مثبت
۲	تعداد واژگان منفی	تعداد کل واژه‌های دارای برچسب منفی
۳	نسبت واژگان مثبت	مقدار A تقسیم بر کل واژگان
۴	نسبت واژگان منفی	مقدار B تقسیم بر کل واژگان
۵	علائم	تعداد علامت تعجب
۶	تعدد واژگان هیجانی	تعداد واژگان در هر جمله منفرد تقسیم بر کل تعداد جملات
۷	جایگاه واژه	محل قرار گرفتن واژه در جمله
۸	تشدیدکننده	وجود تشدیدکننده پیش از یک واژه
۹	تکرار	دفعات تکرار هر واژه برچسب دار

ابزار پژوهشی

الگوریتم مورد استفاده در این تحقیق از نوع الگوریتم ماشینی بردار پشتیبان (SVM)، یکی از الگوریتم‌های یادگیری نظارت‌شده در یادگیری ماشینی است که کارایی بسیار خوبی در بسیاری از تکالیف یادگیری از خود نشان می‌دهد (یایونگ، بونچوا و کائینگام، ۲۰۰۹). این الگوریتم در تکالیف پردازش زبان طبیعی مانند برچسب‌گذاری اجزای کلام (POS)، استخراج اطلاعات از متن، استخراج روابط درون‌متنی، برچسب‌گذاری نقش معنایی^۱، ابهام‌زدایی از واژگان مهم^۲، تحلیل روابط و بسیاری موارد دیگر استفاده می‌شود.

در این تحقیق، برای یادگیری مدل از روش اعتبارسنجی متقابل^۳ استفاده شده است. در این روش داده به‌صورت خودکار و توسط الگوریتم تفکیک شده و بخشی از آن‌ها به‌عنوان داده‌ی

4. Accuracy
5. Precision
6. Recall

1. Semantic Role Labeling
2. word sense disambiguation
3. K-fold cross validation

به ضرایب مزبور می‌توان گفت که مدل پیشنهادی برای شناسایی این گروه هیجانی کارایی مناسبی ندارد.

از نظر شناسایی گروه هیجانی انتظار^۱، مدل پیشنهادی دقت و درستی بالایی را نشان می‌دهد (بیش از ۹۰ درصد). این گروه «ضریب f» و نرخ «بازخوانی» اندکی بالاتر از گروه خشم را نشان می‌دهد و به‌طور کلی می‌توان گفت که کارایی مدل در شناسایی این گروه بالاتر از گروه نخست است. در مورد گروه هیجانی «تفر» نیز مدل پیشنهادی از دقت و درستی قابل قبولی برخوردار است و «ضریب f» و «بازخوانی» بالاتری نسبت به دو گروه قبلی دارد. می‌توان گفت که کارایی مدل برای شناسایی و تفکیک این هیجان قابل قبول است.

کارایی مدل در شناسایی و تفکیک هیجان ترس از همه موارد ذکر شده بالاتر است و در رتبه ی دوم از کارایی مدل (پس از خوشی) قرار می‌گیرد. با توجه به بالا بودن سهم نسبی متن‌های دارای برچسب این هیجان از کل متن‌های ورودی به سیستم، می‌توان گفت که افزایش تعداد داده‌های تست در ورودی الگوریتم تأثیر خوبی در بهبود کارایی مدل داشته است.

بیشترین کارایی مدل پیشنهادی در سنجش و شناسایی هیجان «خوشی» دیده می‌شود. نرخ ضریب «بازخوانی» و همچنین "ضریب f" در آن به‌طور نسبی بالا است (به ترتیب ۶۷٪ و ۸۰٪) و درعین حال درستی و دقت آن نیز بالا است. این نتیجه با وجود سهم نسبی پایین این متون در کل متن‌های ورودی نشان‌دهنده کارایی خوب مدل در شناسایی هیجان مربوطه است و در چنین شرایطی دقت و درستی می‌تواند معیارهای خوبی برای کارایی مدل باشند.

بحث

یافته‌های این تحقیق در راستای تئوری زبان به‌عنوان زمینه‌ی بازنمایی و ادراک هیجان (بارت و همکاران، ۲۰۰۷) نشان‌دهنده‌ی تأثیر ویژگی‌های بافتی متن زبان فارسی و همچنین واژگان در ادراک هیجان و بازنمایی آن است. بر این اساس، تحقیق حاضر نشان می‌دهد که این ویژگی‌ها به‌خوبی می‌توانند در تعیین و تفکیک نوع هیجان و به‌تبع آن نگرش یا گرایش فرد نسبت به یک موضوع بکار گرفته شوند.

برخی از گروه‌های هیجانی قابلیت بیشتری در شناسایی ماشینی از خود نشان داده‌اند که می‌تواند ناشی از تطبیق بهتر آن‌ها با ویژگی‌های زبان نوشتاری فارسی باشد. از سوی دیگر، دلیل این امر را می

ضریب f بر اساس فرمول (۳) محاسبه می‌شود و میانگین هارمونیک بین بازخوانی و دقت است. برای ارزیابی نتایج مدل می‌توان از این ضریب (f-score) استفاده کرد و برای محاسبه‌ی آن از دو پارامتر «دقت» و «بازخوانی» استفاده می‌شود.

$$f\text{-score} = 2 * \frac{\text{precision} \cdot \text{recall}}{\text{precision} + \text{recall}} \quad (۴)$$

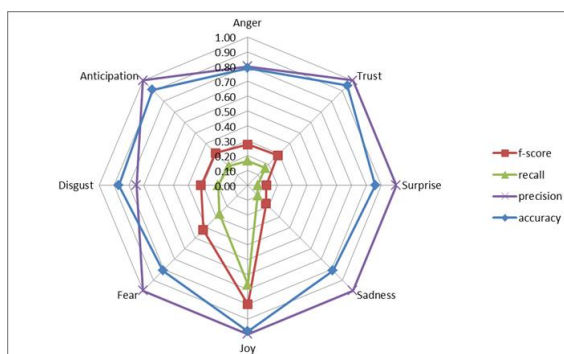
یافته‌ها

نتایج حاصل از ارزیابی الگوریتم استفاده‌شده در جدول ۲ نشان داده شده است. همان‌گونه که مشاهده می‌شود، کارایی مدل (دقت) در شناسایی هیجان‌های مختلف از ۷۹٪ تا ۹۸٪ متغیر است. با توجه به ناهمگنی در تعداد نمونه‌ها (متن‌ها) در هر یک از گروه‌های هیجانی، از میانگین وزنی برای توصیف کلی مدل استفاده شد. بر این اساس، میانگین «دقت» مدل ۸۴٪ محاسبه گردیده است.

جدول ۲. نتایج و شاخص‌های ارزیابی مدل برای گروه‌های هشت‌گانه هیجان

نتایج گروه	TN	FP	FN	TP	درستی	دقت	ضریب f بازخوانی
خشم	۷۵	۱	۲۰	۴	۰/۷۹	۰/۸۰	۰/۲۸
انتظار	۸۹	۰	۹	۲	۰/۹۱	۱/۰۰	۰/۳۱
انزجار	۸۴	۱	۱۲	۳	۰/۸۷	۰/۷۵	۰/۳۲
ترس	۷۴	۰	۱۹	۷	۰/۸۱	۱/۰۰	۰/۴۲
خوشی	۹۴	۰	۲	۴	۰/۹۸	۱/۰۰	۰/۸۰
غم	۷۹	۰	۱۹	۲	۰/۸۱	۱/۰۰	۰/۱۷
تعجب	۸۵	۰	۱۴	۱	۰/۸۶	۱/۰۰	۰/۱۳
اعتماد	۹۴	۰	۵	۱	۰/۹۵	۱/۰۰	۰/۲۹

در نمودار ۱ تمامی گروه‌های هیجانی در قالب یک نمودار نمایش داده شده‌اند. تناسب زیادی در همه‌ی گروه‌ها در مورد شاخص‌های «بازخوانی» و «ضریب f» دیده می‌شود.



شکل ۱. مقایسه‌ی شاخص‌های ارزیابی مدل برای گروه‌های هشت‌گانه‌ی هیجان

گروه هیجان خشم پایین‌ترین میزان درستی را در میان همه‌ی گروه‌ها نشان می‌دهد. همچنین از نظر میزان ضرایب «بازخوانی» و «ضریب f» نیز به ترتیب در رتبه چهارم و پنجم قرار می‌گیرد. با توجه

۱- گروه هیجانی انتظار ترکیبی از رضایت و تهییج است که گاه با گونه‌های اضطراب همراه است اما بر خلاف شگفتی یا تعجب، انتظار وقوع رویداد در فرد وجود دارد. در این پژوهش تحت عنوان چشم انتظاری نیز نام برده شده است.

متون اختصاص یابد. همچنین در این تحقیق، با توجه به پیشینه‌ی پژوهش‌های مرتبط از مدل هشت‌گانه‌ی هیجان‌ها استفاده شده است. استفاده از تئوری‌های مختلف در این زمینه، امکان بیشتری برای تعمیم‌یافته‌های تحقیق فراهم می‌کند.

پیشنهادها

همان‌گونه که اشاره شد، تحقیق حاضر کارایی کلی رهیافت‌ها و مدل پیشنهادی در شناسایی هیجان‌ها و تحلیل قطبیت متن فارسی را نشان می‌دهد. با این حال ارائه‌ی مدل‌ها و الگوریتم‌های کارآمد و نیز کاربردی سازی تحقیقات در این زمینه، نیازمند انجام پژوهش‌های بیشتری در سه حوزه‌ی زبان‌شناسی محاسباتی، روانشناسی هیجان و علوم کامپیوتر است که می‌توانند چشم‌اندازهای نوینی در تحلیل هیجان به‌ویژه در زبان فارسی را ارائه نمایند. با توجه به نبود پژوهش‌های قابل‌توجه در زبان فارسی در مورد تحلیل هیجان در متن، در حال حاضر امکان ارائه‌ی لیست مناسبی از ویژگی‌های مؤثر (اعم از ویژگی‌های نحوی، معنایی، ساختاری) برای استفاده در مدل محاسباتی وجود ندارد. بدین لحاظ، پژوهش‌هایی برای نشان دادن کارایی هر دسته از ویژگی‌های مزبور با توجه به مشخصات زبان فارسی و همچنین تهیه و تدوین لیست واژگان و نشانه‌های کامل در تحلیل متن (مانند منفی‌سازها، تشدیدکننده‌ها و غیره در زبان فارسی) می‌تواند به‌عنوان موضوع تحقیق در زبان‌شناسی محاسباتی باشند. همچنین با توجه به تأکید پژوهش حاضر بر رهیافت گروه بندی هیجان، می‌بایست مقایسه‌ی دقیقی از کارایی دو رهیافت گروه‌بندی و رهیافت ابعادی (برداری)^۲ و مدل‌های مربوطه در زبان فارسی انجام شود. تفکیک و مقایسه بین کارایی مدل‌ها در استفاده از واژه‌های مستقیم و غیرمستقیم و همچنین ویژگی‌های معنایی و نحوی متن، موضوع مهمی در تحقیقات آینده است. علاوه بر آن به‌منظور افزایش امکانات یادگیری در مدل‌ها، طراحی مجدد الگوریتم در محیط وب با تکمیل و بهینه‌سازی الگوریتم (با یادگیری مداوم از کاربران سیستم به‌عنوان منبع یادگیری) می‌تواند در طراحی نرم‌افزارهای کاربردی برای تحلیل متن در زبان فارسی با کارایی مطلوب به کار گرفته شود.

تشکر و قدردانی

نگارش این مقاله با کمک و راهنمایی آقای دکتر میر محسن پدرام رئیس محترم گروه مدل‌سازی شناختی پژوهشکده علوم شناختی و آقای مهندس احمد گنجی تابش دانشجوی کارشناسی ارشد رشته کامپیوتر صورت گرفته است.

توان به کارایی گروه‌بندی مورد استفاده نسبت داد. این بدان معنی است که برخی مدل‌های هیجان با تعداد گروه کمتر (مانند مدل ۶ تایی اکمن) می‌توانند کارایی بهتری در این الگوریتم‌ها داشته باشند.

این پژوهش یافته‌های تحقیقات گوناگون در زبان‌های دیگر به‌ویژه زبان انگلیسی را در مورد کارایی استفاده از واژگان هیجانی و نشانه‌های متنی در ارائه‌ی مدل شناسایی و تفکیک هیجان‌ها و نیز تحلیل احساسات تأیید می‌کند. نتایج این تحقیق در بررسی تفکیک ماشینی متن‌ها نتایج تحقیق کوان و همکاران (۲۰۱۰) در مورد اهمیت استفاده از واژگان، سایر ویژگی‌ها و نشانه‌های متنی و نیز استفاده از پیکره‌ی زبانی را مورد تأیید قرار می‌دهد. تحقیق مزبور در زبان چینی انجام شده و نشان می‌دهد که حدودی از قابلیت تعمیم نتایج در سایر زبان‌ها نیز وجود دارد. از سوی دیگر، یافته‌های این تحقیق در تأیید پژوهش‌های انجام‌شده توسط محمد (۲۰۱۸) و همچنین یانگ و زو (۲۰۱۴) در مورد کارایی استفاده از واژگان در تحلیل ماشینی متن و نیز شناسایی هیجان‌ها یا احساسات است.

یافته‌های این تحقیق می‌تواند در طراحی الگوریتم‌ها و نرم‌افزارهای شناسایی هیجان در محیط وب، طراحی رابط کاربری نرم‌افزارها، اجرای نظرسنجی‌ها، تحلیل محتوای شبکه‌های اجتماعی و همچنین ارزیابی آسیب‌های روانی و اجتماعی بر اساس گفتار و نوشتار از قبیل تحلیل وضعیت ناهنجاری‌های دانش‌آموزان در مدارس، ارزیابی تأثیرات مثبت یا منفی متون آموزشی، واکنش‌سنجی در مخاطبان رسانه‌ها، افزایش اثربخشی یادگیری محتواهای آموزشی و رسانه‌ای از طریق ارزیابی هیجانی آن‌ها و موارد مشابه مورد استفاده قرار گیرد.

محدودیت‌های تحقیق

این پژوهش دارای محدودیت‌هایی است که می‌تواند بر کیفیت نتایج آن و یا قابلیت تعمیم نتایج تأثیر داشته باشد. یکی از مهم‌ترین این موارد، کمبود منابع و الگوریتم‌های شناسایی هیجان در متن است که باعث می‌شود امکان استفاده از لیست واژگان مناسب و طبقه‌بندی‌های متنوع هیجان‌ها فراهم نباشد. همچنین مراجع مناسبی برای تهیه‌ی منابع کامل برای واژگان تشدیدکننده، منفی‌سازها و غیره در دسترس نبوده است. استفاده از این‌گونه مراجع باعث تقویت مدل شده و درعین‌حال امکان بیشتری برای ارزیابی مدل‌ها در زبان فارسی را فراهم می‌کند. از سوی دیگر، نبود منابع معتبر برای متون برجسب گذاری شده در زبان فارسی باعث می‌شود اولاً امکان چندانی برای فرایند یادگیری مدل در پژوهش‌های مشابه وجود نداشته باشد و دوم این‌که بخشی از روش‌شناسی و فعالیت‌های پژوهشی به تهیه‌ی این

References

- Aman S. & Szpakowicz S. (2007) Identifying Expressions of Emotion in Text. In: Matoušek V., Mautner P. (eds) Text, Speech and Dialogue. TSD 2007. Lecture Notes in Computer Science, vol 4629. Springer, Berlin, Heidelberg.
- Aman, S. (2007). Recognizing Emotions in Text. Ottawa-Carleton Institute for Computer Science School of Information Technology and Engineering. University of Ottawa.
- Barrett, L. F., Lindquist, K. A., & Gendron, M. (2007). Language as context for the perception of emotion. *Trends in cognitive sciences*, 11(8), 327-32. [DOI:10.1016/j.tics.2007.06.003] [PMid:17625952 PMCID:PMC2225544]
- Cho, Y. H., & Lee, K. J. (2006). Automatic affect recognition using natural language processing techniques and manually built affect lexicon. *IEICE Transactions on Information and Systems*, 89(12), 2964-2971. [DOI: 10.1093/ietisy/e89-d.12.2964]
- D'Mello, S. K., Picard, R., & Graesser, A. C. (2007). Towards an Affect Sensitive Auto-Tutor, *IEEE Intelligent Systems*, 22(4), 53-61. [DOI:10.1109/MIS.2007.79]
- Farra, N., Challita, E., Abou A. R. & Hajj, H. (2010). Sentence-Level and Document-Level Sentiment Mining for Arabic Texts. 1114-1119. [DOI:10.1109/ICDMW.2010.95]
- Francisco V, Gervás P. & Peinado F. (2010). Ontological reasoning for improving the treatment of emotions in text. *Knowl Inf Syst* 25(3):421-443. [DOI:10.1007/s10115-010-0320-1]
- Gaind, Bharat & Syal, Varun & Padgalwar, Sneha. (2019). Emotion Detection and Analysis on Social Media.
- Hahn, U. & Buechel, S. (2017). EmoBank: Studying the Impact of Annotation Perspective and Representation Format on Dimensional Emotion Analysis. *EACL*. [DOI:10.18653/v1/E17-2092]
- Mac Kim, S. (2011). Recognising Emotions and Sentiments in Text. Ph.D. thesis, University of Sydney.
- Mohammad, S. (2018). Word Affect Intensities. *CoRR*, abs/1704.08798.
- Myers, D. G. (2004). *Theories of Emotion*. Psychology: Seventh Edition, New York, NY: Worth Publishers.
- Neviarouskaya, A., Prendinger, H., & Ishizuka, M. (2011b). SentiFul: A Lexicon for sentiment analysis. *IEEE Transactions on Affective Computing*, 2(1), 22-36. [DOI:10.1109/T-AFFC.2011.1]
- Novielli, N., Calefato, F. & Lanubile, F. (2018). A gold standard for emotion annotation in stack overflow. 14-17. [DOI:10.1145/3196398.3196453]
- Pennebaker, J. & E. Francis, M & J. Booth, R. (2001). *Linguistic Inquiry and Word Count (LIWC): LIWC2001*. 71.
- Quan, C. & Ren, F. (2010). Sentence Emotion Analysis and Recognition Based on Emotion Words Using Ren-CECps. *International Journal of Advanced Intelligence*, 2, 105-117.
- Ramalingam, V V & Pandian, A & Jaiswal, Abhijeet & Bhatia, Nikhar. (2018). Emotion detection from text. *Journal of Physics: Conference Series*. 1000. 012027. [DOI:10.1088/1742-6596/1000/1/012027]
- Strapparava, C., & Mihalcea, R. (2007). Semeval-2007 task 14: Affective text. In *Proceedings of the 4th International Workshop on Semantic Evaluations* (pp. 70-74). Prague. [DOI:10.3115/1621474.1621487]
- Strapparava, C., & Valitutti, A. (2004). *WordNet Affect: an Affective Extension of WordNet*.
- Wilce, J. (2009). *Language and Emotion (Studies in the Social and Cultural Foundations of Language)*. Cambridge: Cambridge University Press. [DOI:10.1017/CBO9780511626692]
- Yang, Min & Zhu, D. & Chow, K.P. (2014). A Topic Model for Building Fine-grained Domain-specific Emotion Lexicon. 421-426. [DOI:10.3115/v1/P14-2069]
- Yaoyong Li, Bontcheva K. & Cunningham H. (2009). Adapting SVM for Natural Language Learning: A Case Study Involving Information Extraction. *Natural Language Engineering*, 15(2), 241-271. Cambridge University Press New York, NY, USA. [DOI:10.1017/S1351324908004968]