تاریخ دریافت مقاله: ۱۴۰۳/۰۳/۰۸ | تاریخ پذیرش نهایی: ۱۴۰۳/۰۷/۰۳ | تاریخ انتشار: ۱۴۰۳/۱۰/۰۱

نشریه علمی نامه معماری و شهرسازی، ۱۷(۴۵)، ۷۹-۱۰۳
DOI: 10.30480/aup.2024.5338.2151
نوع مقاله: پژوهشی

ارائه مدل پیش‌بینی‌کننده تحلیل احساسات داده‌های چندوجهیِ شهری با بهره‌گیری از یادگیری عمیق بررسی موردی: کلان‌شهرهای رسمی ایران*

مریم محمدی
دانشیار گروه طراحی شهری، دانشکده معماری و شهرسازی، دانشگاه هنر ایران، تهران، ایران (نویسنده مسئول مکاتبات)
E-mail : m.mohammadi@art.ac.ir

چکیده

تحلیل احساسات شهری با استفاده از داده‌های فضای مجازی، یکی‌از روش‌های نسبتاً نوین برای تحلیل احساسات در شهر است. هدف مقاله حاضر، ارائه مدلی پیش‌بینی‌کننده برای تحلیل احساسات داده‌های متنی-تصویری (چندوجهی) رسانه اجتماعی توئیتر درباره شهر است. روش پژوهش کمی و رویکرد پژوهش یادگیری عمیق نظارت شده است. تحلیل داده‌ها با زبان برنامه‌نویسی پایتون و در فضای ابری گوگل کولب انجام شده است. نمونه مطالعاتی، داده‌های شهری چندوجهی مربوط به ۸ کلان‌شهر رسمی ایران در بازه‌ای ۱۰ ساله (فرودین ۱۳۹۰ تا اسفند ۱۴۰۰) است. برای آموزش داده‌های متنی از الگوریتم هیبریدی طراحی شده (مبتنی‌بر شبکه عصبی سی.ان.ان.) استفاده شده که دقت آن، ۷۷٪ بوده است. برای آموزش داده‌های تصویری از الگوریتم‌های: شبکه عصبی، شبکه از پیش‌آموزش‌دیده دنس‌نت۱۲۱ و شبکه وی.جی.جی.۱۹ استفاده شد که دقت شبکه سوم با ۸۸٪، از دو الگوریتم دیگر بهتر بوده است. الگوریتم ترکیبی برای تحلیل داده‌های متنی-تصویری، مبتنی‌بر ترکیب مدل‌های ازپیش آموزش‌دیده وی.جی.جی.۱۹ و مدل ترکیبی طراحی شده برای متن بوده که دقت آن ۸۱٪ بوده است. براساس نتایج، می‌توان با ارائه داده‌های متنی-تصویری مربوط به سایر شهرها، احساسات کاربران نسبت به شهر را براساس موقعیت داده‌های ورودی تعیین کرد؛ این امر می‌تواند به برنامه‌ریزان و مدیریت شهری کمک نماید تا پهنه‌های نامناسب و مناسب ازمنظر کاربران را براساس داده‌های ثبت شده آن‌ها در فضای مجازی به دست آورند و نسبت به بهبود فضاهای نامناسب اقدام نمایند.

کلیدواژه‌ها: مدل پیش‌بینی‌کننده، تحلیل احساسات شهری، داده‌های چندوجهی، یادگیری عمیق، کلان‌شهر، ایران

* مقاله حاضر برگرفته از پژوهشی است که در گروه آمارِ دانشکده ریاضیِ دانشگاه واترلو در کانادا در فرصت مطالعاتی یک‌ساله نویسنده انجام شده است.

مقدمه

مسائل برنامه‌ریزی و طراحی شهری، مسائلی پیچیده هستند. از همین‌رو فهم آن‌ها به‌خصوص در مقیاس کلان، صرفاً نمی‌تواند با روش‌های متداول و سنتی شهرسازی، محقق شود. بنابراین استفاده از علم داده‌ها و مدل‌سازی می‌تواند به‌عنوان رویکرد و روشی مؤثر برای فهم آن‌ها باشد. براساس پژوهش‌های انجام شده و گزارش‌های موجود در رابطه با برخی از پروژه‌های واقعی شهری در کشورهای مختلف (EURISY, 2017; Kim *et al*., 2020)، یادگیری ماشین و عمیق می‌توانند به حل مسائل شهری کمک کنند و مزایای دیگری چون، کاهش هزینه‌های اقتصادی ناشی از ممیزی، سرعت و قابلیت پیش‌بینی‌پذیری را به‌همراه داشته باشند. این روش‌ها می‌توانند در مرحله ارزیابی، طراحی و پس‌از اجرا در برنامه‌ریزی و طراحی شهری استفاده شوند (Tebyanian, 2020).

موضوع احساسات استفاده کننده (شامل استرس محیطی، احساس شادمانی، احساس ترس و... (Ekman & Friesen, 1971; Ekman, 1999 به نقل‌از Chapman *et al*., 2018) و یا احساسات منفی و مثبت) به محیط و تأثیر آن بر نوع حضور و فعالیت در شهر (Huang *et al*., 2014)، یکی‌از موضوعات مهم در طراحی فضاها در شهر است و این پژوهش بر آن است تا احساسات کاربران در شهر به‌عنوان موضوع اصلی پژوهش مدّنظر قرار دهد. در رابطه با اهمیت بررسی احساسات در شهر، می‌توان به این موضوع توجه نمود که عمل و اقدام کاربر در فضا، ناشی از احساس و ادراک شخص از فضا است و تمام تلاش شهرسازان، خلق فضاهایی است که کاربران در آن احساس خوبی را تجربه نمایند.

همان‌طور که در ادامه آورده می‌شود، طیف گسترده‌ای از پژوهش‌ها در ارتباط با موضوع احساسات در شهر به‌صورت کمی و کیفی، انجام شده و ارزیابی احساس افراد در فضاها، عموماً در محیط واقعی و یا بعضاً در محیط مجازی انجام پذیرفته است. تحلیل احساسات یا نظرکاوی که پیش‌از ظهور رسانه‌های اجتماعی در عرصه‌های مختلفی چون اقتصاد و بازاریابی، سیاست و جامعه‌شناسی سبقه زیادی داشته؛ امروزه به‌واسطه داده‌های منتشر شده و حجم داده‌های قابل دسترس در فضاهای اجتماعی مجازی، به‌شیوه‌های مدل‌گرا چون یادگیری ماشین تحلیل می‌شود. علی‌رغم آنکه این روش از تحلیل احساسات براساس داده‌های فضای مجازی در حوزه‌های پیش گفته، جایگاه ویژه‌ای یافته، اما این حوزه در مطالعات شهری نوظهور بوده و پژوهش‌های چندان زیادی در این ارتباط انجام نشده است. بنابراین این پژوهش بر آن است تا ضمن تبیین این روش، با بهره‌گیری از روش یادگیری عمیق، مدلی پیش‌بینی‌کننده را ارائه دهد که امکان شناسایی احساسات مثبت و منفی کاربران در رابطه با فضا را فراهم آورد.

همان‌طورکه بیان شد شناسایی و دسته‌بندی فضاها ازحیث نوع پاسخ عاطفی کاربر، مسیری روشن را برای طراحان و مدیران شهری فراهم می‌کند. تصور کنید براساس داده‌های فضای مجازی مکان‌محور و آموزش ماشین بتوان، براساس داده‌های مجازی، بدون هزینه‌های بالای ممیزی و صرف زمان و نیروی انسانی زیاد، فضاهای نامناسب را شناسایی نمود و حتی براساس نوع واژه‌های استفاده شده کاربران و الگوریتم‌های دسته‌بندی یا خوشه‌بندی، فضاها را طبقه‌بندی و تحلیل نمود. بدین‌ترتیب به‌جای استفاده از روش‌های سنتی، صرفاً با بهره‌گیری از داده‌های جدید در یک ماشین آموزش‌دیده می‌توان، ممیزی‌ای کلی نسبت به مناسب یا نامناسب بودن فضا ازنظر کاربران داشت و به‌جای رویکرد گزینشی مبتنی‌بر قدرت برای بهبود یک فضا، فضاهای نیازمند بهسازی را شناسایی و اولویت‌بندی مناسبی برای بهبود آن‌ها داشت. علی‌رغم آنکه می‌توان طیف گسترده‌ای تحلیل‌ها را با استناد به داده‌های در دسترس انجام داد، اما به‌دلایل مختلف و محدودیت‌هایی که در بخش نتیجه‌گیری به آن اشاره می‌شود، در این پژوهش به‌دلیل نوع داده‌های در دسترس، بیشتر بر تبیین نحوه استفاده از این روش و نمونه‌ای از لایه‌های تحلیل احساسات اکتفا شده است.

پیشینه پژوهش

در این بخش نمونه‌های از پژوهش‌های انجام شده درحوزه تحلیل احساسات شهری و کاربرد یادگیری ماشین در شهرسازی و نوع داده‌های انتخاب شده معرفی می‌شوند. هدف این بخش آن است تا ضمن بررسی کاربست تحلیل احساسات در موضوعات شهری، فرایند انجام پژوهش و نوع داده‌های مورد استفاده در هر پروژه نیز مشخص شود. پژوهش‌ها با جست‌وجو در پایگاه داده‌هایی چون ساینس دایرکت، الزویر و یا با جستجو در گوگل اسکولار، ریسرچ گیت و... انتخاب شده‌اند. لیست اولیه مقالات ۵۸ مقاله بوده که با توجه به محتوای بررسی شده، ۲۳ مورد به‌دلیل محتوا و حوزه نتایج، توسط محقق مناسب‌تر تشخیص داده شده است (جدول ۱).

جدول ۱. بررسی پیشینه پژوهش‌های حوزه احساسات شهری

نمونه	مقیاس/ هدف	نوع داده، حجم و نحوه جمع‌آوری آن‌ها/ نحوه تحلیل، ابزار مورد استفاده
نمونه ۱ اندونزی	هدف: ادغام تحلیل احساسات متن، تصویر و مفهوم/ ارائه روش یادگیری ترکیبی و نحوه استخراج مفاهیم از متن و تصویر	داده‌های چندوجهی-متن و تصویر- توئیتر (شامل ۹۹۶۲ توئیت)؛ استفاده از دو حاشیه‌نویس برای برچسب‌زنی داده‌ها و درصورت عدم توافق استفاده از حاشیه‌نویس سوم؛ استفاده از شبکه عصبی پویا؛ استفاده از رویکرد ترکیبی و بهره‌گیری از متاطبقه‌بندی کننده (بگینگ، جنگل تصادفی، ماشین بردار پشتیبان، رگرسیون لجستیک، ایکس جی.بوست و بیز ساده)؛ استخراج مفاهیم از متن با استفاده از سنتیک‌نت ۵ و استخراج مفاهیم از تصویر با استفاده از مدل دیپ سنتی‌بنک.
نمونه ۲	هدف: تلفیق داده‌های متنی، بصری و صوتی با استفاده از سی.اف.اس. و پی.سی.ای./ ادغام ویژگی‌ها با استفاده از روش طبقه‌بندی باینری ۴ طرفه	داده‌های متنی، صوتی و تصویری- تحلیل ۸۰ ویدئو- یوتیوب؛ استخراج ویژگی از داده‌های متنی (با استفاده از سی.ان.ان. و انتقال آن‌ها به اس.وی.ام. و یادگیری چندهسته‌ای)، صوتی (نرم‌افزار منبع باز اپن‌اسمایل) و بصری (استخراج اف.سی.پی. ها)؛ تقسیم ویدئوها به گفته‌ها و برچسب‌زنی داده‌های متنی (استفاده از سنتیک‌نت ۳/۰)؛ استفاده از ۳ حاشیه‌نویس برای برچسب‌زنی داده‌ها؛ تبدیل ویدئوها به فریم تصویر و استخراج ویژگی از آن‌ها (استخراج نقاط مشخصه چهره: استخراج ۶۸ اف.سی.پی. از هر تصویر).
نمونه ۳ کالیفرنیا	هدف: ادغام تحلیل احساسات داده‌های متنی و بصری با استفاده از الگوریتم طبقه‌بندی‌کننده گروهی	داده‌های متنی و بصری- اینستاگرام؛ استفاده از ۱۲۰۰پست اینستاگرامی اولیه و ۱۰۱۰ پست نهایی برای تحلیل؛ استفاده از دو حاشیه‌نویس برای برچسب‌زنی داده‌ها و استفاده از حاشیه‌نویس سوم درصورت تفاوت در حاشیه‌نویسی؛ استفاده از الگوریتم روش نمونه‌گیری اقلیت مصنوعی برای ایجاد داده‌های مصنوعی و متعادل‌سازی داده‌ها؛ استفاده از الگوریتم یادگیری ماشین برای ارزیابی شامل ماشین بردار پشتیبان، جنگل تصادفی و طبقه‌بندی‌کننده بیز ساده؛ اس.ام.اُ.تی.ایی. بکینگ. استفاده از الگوریتم
نمونه۴	هدف: پیشنهاد دو مدل برای تحلیل احساسات داده‌های چندوجهی، شامل مدل چندوجهی همبستگی و یادگیری هایپرگراف	داده‌های متنی و تصویری؛ استفاده از میدان تصادفی مارکوف، در مدل چندوجهی همبستگی؛ استفاده از مدل کوله کلمات برای متن، تصویر و ایموجی در مدل یادگیری چندوجهی هایپرگراف؛ استفاده از فرهنگ لغت احساسات متنی، سنتی‌بنک و فرهنگ لغت احساسات ایموجی در مدل یادگیری چندوجهی هایپرگراف.

نمونه	مقیاس/ هدف	نوع داده، حجم و نحوه جمع‌آوری آن‌ها/ نحوه تحلیل، ابزار مورد استفاده
نمونه ۵ شیکاگو و نیویورک آمریکا، ۲۰۱۵-۲۰۱۹	مقیاس: فضای شهری (پارک) هدف: تحلیل احساسات، نوع فعالیت‌ها و رضایت کاربران از حضور در پارک‌ها	داده‌های متنی- توئیتر؛ تعداد توئیت‌های پارک ۶۰۶: ۱۲۹۵۲ توئیت/ تعداد توئیت‌های پارکِ های لاین: ۱۶۵۳۴۷ توئیت؛ استفاده از فرهنگ لغت وی.ای.دی.ایی.آر.[1]
نمونه ۶ بوستون، امریکا، ۲۰۱۸	مقیاس: فضای سبز شهری هدف: دسته‌بندی احساسات در هفت دسته: شادی، غم، عصبانیت، اعتماد، تحلیل، آزمایش و ترس/ مکان‌نگاشت احساسات	داده‌های تصویری- اینستاگرام (۱۰ هزار داده)، که مشتمل بر داده‌هایی چون، تاریخ، محتوای پست، تعداد کامنت‌ها، موقعیت مکانی، عرض‌وطول جغرافیایی هستند؛ استفاده از پیکوداش[2] برای خروجی گرفتن از داده‌های اینستاگرام؛ استفاده از تن آنالیزور[3] برای تحلیل احساسات (Ntshangase, 2018 به نقل‌از Kim *et al*., 2020)؛ ارائه داده در قالب فرمت جئو.جی.سون[4].
نمونه ۷ نیویورک، ۲۰۱۷	مقیاس: فضای سبز شهری هدف: بررسی احساس کاربران هنگامی که در پارک هستند نسبت به زمانی که در مکان‌های دیگر هستند.	داده‌های متنی- توئتیر (توئیت‌هایی با برچسب‌های جغرافیایی مربوط به ۵۴۹ روز از جون ۲۰۱۶ تا ۱۷ دسامبر ۲۰۱۷)؛ داده‌های آموزشی ۱۰ هزار توئیت هستند که به‌صورت تصادفی انتخاب و توسط دو محقق آموزش‌دیده و یک کاربر ترک مکنیکال آمازون برچسب‌گذاری شده‌اند؛ استفاده از الگوریتم رگرسیون لجستیک ساده، ماشین بردار پشتیبان بیز ساده؛ استفاده از فرهنگ لغت احساسی ام.پی.کیو.ای. و لیو.
نمونه ۸ دوبلین، ۲۰۱۰	مقیاس: فضای سبز شهری هدف: شناسایی موضوعات مورد توجه و احساسات کاربران در پارک‌های محبوب	داده‌های متنی- تریپ ادوایزر و فوراسکوئر (مربوط به بازه زمانی مه ۲۰۰۶ تا دسامبر ۲۰۱۰)؛ بررسی دیدگاه‌های ۸ پارک با محبوبیت بالا از میان ۶۰ پارک شناسایی شده در دوبلین؛ استفاده از تخصیص پنهان دیریکله برای شناسایی موضوعات؛ استفاده از روش طبقه‌بندی مبتنی‌بر فرهنگ لغت.
نمونه ۹ لندن	مقیاس: فضای سبز شهری هدف: تحلیل حال‌وهوای افراد هنگام توئیت کردن، استخراج احساسات و تحلیل روند روزانه، هفتگی و فصلی./ خوشه‌بندی پارک‌ها	داده‌های متنی- توئیتر (۷۸ هزار توئیت در ۱۷۰۰ پارک)؛ استفاده از الگوریتم پی.ای.ام. در نرم‌افزار آر. استودیو؛ تحلیل امتیاز احساسات براساس فرهنگ لغت هوو و لیو؛ قطبیت احساس با استفاده از فرهنگ لغت احساسات شورای ملی تحقیقات کانادا، از طریق پکیج سایوژت.
نمونه ۱۰ سانفرانسیسکو	مقیاس: فضای سبز شهری هدف: شناسایی رفتار کاربران حین ورود به پارک و قبل و بعداز آن/ الگوی فراوانی کلمات مورد استفاده	داده متنی- توئیتر (دوره زمانی از ۱۹ ماه می ۲۰۱۶ تا ۲ آگوست ۲۰۱۶ در ۲۲۰ فضای سبز)؛ استفاده از هدونومتر به‌عنوان فرهنگ لغت احساسات با ۱۰۰۲۲ کلمه.

نمونه	مقیاس/ هدف	نوع داده، حجم و نحوه جمع‌آوری آن‌ها/ نحوه تحلیل، ابزار مورد استفاده
نمونه۱۱ سیدنی، استرالیا	مقیاس: شهر هدف: تحلیل احساسات و لحن برای ادراک سطح رضایت کاربران حمل‌ونقل عمومی در رسانه‌های اجتماعی. / شناسایی موضوعات مورد تاکید کاربران/ شناسایی رابطه بین تاخیر حمل‌ونقل و حجم توئیت روزانه	داده‌های متنی- توئیتر (بازه زمانی ژانویه ۲۰۱۸ تا آوریل ۲۰۱۹)؛ شناسایی ۲۶ کلمه کلیدی و طبقه‌بندی براساس آن‌ها؛ تحلیل احساسات با استفاده از ابزار زبان طبیعی ان.ال.نی.کی که مدل منبع باز وی.ای.دی.ایی.آر. است. / تحلیل لحن واتسون آی.بی.ام.؛ مدل‌سازی موضوعی و استفاده از تخصیص پنهان دیریکله (با استفاده از کتابخانه متن باز جنسیم).
نمونه ۱۲ پیجینگ	مقیاس: شهر هدف: تشخیص رویدادهای شهری (۵ رویداد مثبت و ۵ رویداد منفی) با استفاده از داده‌های رسانه‌های اجتماعی دارای برچسب در یک شبکه شطرنجی که هر سلول آن ۱ در ۱ کیلومتر است. / ارزیابی میزان احساسات متن	داده‌های متنی-سیناویبو (۴۲۷۸۶۰۷ میکروبلاگ)؛ استفاده از شبکه عصبی حافظه کوتاه‌مدت؛ استفاده از روش مبتنی‌بر فرهنگ لغت چینی به نام هونت؛
نمونه ۱۳ لندن	مقیاس: شهر هدف: تحلیل رویداد بزرگ ورزشی (المپیک) در سه بازه زمانی حین، قبل و بعداز آن بین ساکنان و بازدیدکنندگان/ تحلیل متنی شامل تحلیل احساسات و مدل‌سازی خودکار/ شناسایی هات اسپات‌ها	داده‌های متنی-توئیتر مربوط به سال ۲۰۱۲ (قبل از رویداد ۲۷ جون تا ۱۳ جولای/ حین رویداد: ۲۷ جون تا ۱۲ آگوست/ پس‌از رویداد: ۲۷ آگوست تا ۱۲ سپتامبر)؛ استفاده از الگوریتم تخصیص پنهان دیریکله؛ استفاده از الگوریتم احساسات براساس کار برین، ۲۰۱۲؛ استفاده از فرهنگ لغت هو لیو.
نمونه ۱۴ چین	مقیاس: ۱۹۵ مقصد گردشگری هدف: تحلیل احساسات از ادراک کیفیت هوا توسط گردشگران	داده‌های متنی- سینا ویبو بازه زمانی از ژانویه ۲۰۱۱ تا دسامبر ۲۰۱۷ (۲۷۵۰۰ دیدگاه)؛ تحلیل محتوای داده‌ها با استفاده از ابزارهای جووسیکر، رست سی.ام. و بوزون ان.ال.پی.؛ تحلیل احساسات با رست.ایی.آی.؛ تعیین ارزش احساسی متون با شبکه عصبی مصنوعی (ای.ان.ان.).
نمونه ۱۵ جونگ ایست در سنگاپور	مقیاس: شهر هدف: ارائه مکانیسم تحلیل احساسات جغرافیایی کالیبره شده توسط کاربر (سی.جی.اس.ای.) و بهبود طبقه‌بندی براساس سرویس کالیبراسیون جمع‌سپاری/ تحلیل احساسات، خوشه‌بندی و تحلیل سری‌های زمانی/ تصویرسازی احساسات اجتماعی به‌صورت دو بعدی و سه بعدی	داده‌های متنی- توئیتر و اینستاگرام؛ استفاده از داده‌های اس.تی.اس.، ساندرز (اس.ای.ان.) و سمی ایوال ۲۰۱۵ (ایی.وی.ای)، به‌عنوان داده‌های آموزشی؛ استفاده از الگوریتم‌های یادگیری ماشین، بیز، جی. ۴۸.

نمونه	مقیاس/ هدف	نوع داده، حجم و نحوه جمع‌آوری آن‌ها/ نحوه تحلیل، ابزار مورد استفاده
نمونه ۱۶ لندن	مقیاس: محله هدف: تحلیل احساسات در محلات مبتنی‌بر الگوریتم حافظه کوتاه مدت طولانی	داده‌های متنی- استفاده از مجموعه داده‌های سنتی هود (شامل ۵۲۱۵ داده)؛ شناسایی لیستی از کلمات کلیدی شامل قیمت، ایمنی، موقعیت حمل‌ونقل و عمومی برای طبقه‌بندی؛ استفاده حاشیه‌نویسی بی.آروای.تی. و انتخاب سه حاشیه‌نویس بدون تخصص؛ استفاده از الگوریتم حافظه کوتاه مدت طولانی.
نمونه ۱۷ ۲۰۱۴	مقیاس: شهر هدف: استخراج احساسات معنایی-مکانی ارائه فرایند تحلیل احساسات	داده‌های متنی- توئیتر (در منطقه نیویورک در طول هفته مد ۶-۱۳ فوریه ۲۰۱۴ و در منطقه بوستون؛ داده‌های رویداد مارتن، ۱۵ آوریل ۲۰۱۳)؛ فراداده توئیت‌ها شامل: مکان، زمان، شناسه کاربر، شناسه پست و...؛ تلفیق احساسات توسط سنسورها، وی.جی.آی.[5] و جی.آی.اس.
نمونه ۱۸ شیکاگو، آمریکا، ۲۰۲۰	مقیاس: فضای باز شهری هدف: بررسی تاثیر تلفیق ویژگی‌های عمیق با ویژگی‌های سان[6] و یولو[7]	داده‌های تصویری-استفاده از شبکه اجتماعی فلیکر؛ نوع داده‌ها: داده‌های توئیتر از دیپ‌سنت[8] (شامل ۱۲۶۹ تصویر توئیتر)/ داده‌های تصاویر بیرونی از اوت‌دُرسنت[9] (شامل ۴۰ هزار و ۵۱۶ تصویر عمومی در دسترس فلیکر در رابطه با شیکاگو)؛ استفاده از یادگیری عمیق.
نمونه ۱۹ آمریکا ۲۰۱۵	مقیاس: شهر (۱۰۰ شهر در دو گونه در حال رشد و رو به زوال) هدف: بررسی تفاوت احساس کاربران شهرهای رو به زوال و در حال رشد از شهر محل سکونت خود/ تحلیل احساسات	داده‌های متنی- توئیتر (۳۰۹ هزار توئیت در ۲ ماه)؛ دانلود توئیت‌ها شامل شناسه کاربر، نام کاربری، متن، طول جغرافیایی، عرض جغرافیایی، زبان و مکان؛ استفاده از فرهنگ لغت ای.اف.آی.ان.ان.[10] که دارای ۲۴۷۷ کلمه است؛ استفاده از رگرسیون لجستیک دوتایی؛ استفاده از الگوریتم بهینه‌سازی پی.اس ـمچ ـ مولتی اس.ای.اس. ماکرو[11].
نمونه ۲۰ ۱۲ شهر در اروپا و ایالات متحده، ۲۰۱۴	مقیاس: شهر هدف: ارائه نقشه‌ای کیفی از شهر برای مسیریابی و تهیه نقشه عرصه‌های شاد/ فضاهای بدبو/ فضاهای پر سر و صدا.	داده‌های تصویری-رسانه‌های اجتماعی.
نمونه ۲۱ چین، ۲۰۱۷	مقیاس: نمای خیابان هدف: ارائه یک مدل کیفیت نمای خیابان براساس یادگیری ماشین	داده‌های بصری- استخراج تصاویر نما از گوگل‌مپ؛ دانلود ۳۶۰۷۹۶ تصویر در فوریه ۲۰۱۶؛ برچسب‌گذاری ۳۵۰۰ تصویر به‌صورت دستی با دو برچسب تصویر ساختمان و خیابان؛ تاکید بر دو جنبه کیفیت نگهداری و تداوم و رتبه‌بندی توسط متخصصین (۸ فارغ‌التحصیل معماری)؛ استفاده از هیستوگرام معمولی اس.آی.اف.تی.، استفاده از دو شبکه پیشرفته پیچشی شامل الکس نت و گوگل نت؛ اعتبارسنجی نتایج مدل پیش‌بینی‌کننده با محیط واقعی.

نمونه	مقیاس/ هدف	نوع داده، حجم و نحوه جمع‌آوری آن‌ها/ نحوه تحلیل، ابزار مورد استفاده
نمونه ۲۲ مکزیک ۲۰۱۹	مقیاس: ۵۰ منطقه شهری هدف: طبقه‌بندی تصاویر ماهواره‌ای برای تشخیص فرم شهری	داده‌های بصری- پلت فرم پلنت (تصاویر ماهواره‌ای مربوط به سال ۲۰۱۵-۲۰۱۶)؛ استفاده از روش طبقه‌بندی حداکثر احتمال پارامتری، به‌عنوان متداول‌ترین روش طبقه‌بندی سنجش از دور؛ استفاده از الگوریتم‌های طبقه‌بندی شامل شبکه‌های عصبی مصنوعی، ماشین بردار پشتیبان، درخت تصمیم و حداکثر احتمال.
نمونه ۲۳ شهر سامرویل در بوستون	مقیاس: شهر هدف: ممیزی ظرفیت پارکینگ پس‌از احداث یک خط جدید مترو با استفاده از یادگیری ماشین نظارت شده	داده‌های تصویری- عکس‌های ماهواره‌ای از کل شهر؛ استفاده از الگوریتم شبکه‌های عصبی پیچشی برای استخراج ویژگی از داده‌های تصویری؛ استفاده از الگوریتم طبقه‌بندی از رگرسیون لجستیک، جنگل تصادفی، آدابوست، ایکس.جی.بوست و شبکه عصبی با استفاده از داده‌های ارزیابی مالیات و ویژگی‌های به دست آمده از برداشت میدانی.

نتایج بررسی پیشینه پژوهش نشان داد که بسیاری از پژوهش‌ها بر تحلیل داده‌های متنی تا داده‌های تصویری تمرکز داشته‌اند. در رابطه با مجموعه داده‌های مورد استفاده بیشتر نمونه‌ها متکی بر داده‌های توئیتر بوده‌اند و در پژوهش‌های مربوط به چین از پلتفرم سیناویبو که پلتفرمی مشابه با توئیتر است، استفاده شده است. تحلیل‌ها متکی بر دو شیوه کلی بوده‌اند یا از تحلیل مبتنی‌بر فرهنگ لغت استفاده شده یا از رویکرد مبتنی‌بر یادگیری ماشین بهره برده شده است. در پژوهش‌های شهری که درحوزه‌های مختلف صورت گرفته، بیشتر تأکید بر فضای سبز بوده است. حوزه مورد تأکید در بیشتر پژوهش‌ها، تأثیر حضور در فضاهای سبز بر احساس کاربران بوده است. اما در پژوهش‌های دیگر از ترکیب روش‌های تحلیل احساسات مبتنی‌بر شبکه اجتماعی و تحلیل‌های مبتنی‌بر جی.آی.اس. و یا تحلیل‌های مبتنی‌بر برون‌سپاری استفاده شده است. مزیت استفاده از این روش، کم‌هزینه بودن، سرعت در دست‌یابی به نتایج با بهره‌گیری از کلان‌داده‌ها و گستردگی حجم بررسی است. یافته‌های پژوهش‌ها ضمن بررسی محبوبیت فضاها، مطلوبیت آن‌ها و قطبیت احساسات، به ارائه نقشه یا توصیه‌های مکانی برای کاربران فضاها پرداخته است. درعین‌حال از دیگر مزایای این پژوهش‌ها، درگیری و مشارکت کاربران بوده است. با نگاه به نوع ابزارهای مورد استفاده در تحلیل احساسات می‌توان انواع فرهنگ لغت‌ها و الگوریتم‌های مورد استفاده در پژوهش‌های را به شرح زیر ارائه کرد:

- استفاده از فرهنگ لغت‌های احساسات: ام.پی.کیو.ای.، هوو و لیو، هدونومتر، ای.اف.آی.ان.ان.، وی.ای.دی.ایی.آر.
- استفاده از تحلیل لحن: تحلیل تن آنالیزور و تحلیل لحن واتسون آی.بی.ام.
- الگوریتم‌های مورد استفاده در روش یادگیری ماشین: ماشین بردار پشتیبان، بیز ساده، رگرسیون لجستیک، رگرسیون لجستیک دوتایی، تخصیص پنهان دیریکله، کی میانگین همسایه، زیرو آر.، جی.۴۸، حداکثر احتمال، درخت تصمیم، جنگل تصادفی، آدابوست و ایکس.جی.بوست.
- الگوریتم‌های مورد استفاده در روش یادگیری عمیق: شبکه عصبی مصنوعی، شبکه عصبی حافظه کوتاه مدت، شبکه پیشرفته پیچشی الکس نت، شبکه پیشرفته پیچشی گوگل نت و شبکه عصبی کانولوشنی.

براساس موارد بیان شده ازآنجاکه درحوزه شهر و در رابطه با تحلیل احساسات در شهر با استفاده از روش‌های مدل‌گرا در ایران پژوهش مشخصی انجام نشده، بنابراین انتخاب رویکرد و روش در این پژوهش از نوآوری‌های این پژوهش درحوزه شهرسازی ایران محسوب می‌شود. در ارتباط با نوآوری این پژوهش در مقایسه با

پژوهش‌های انجام شده در سطح جهان، مشاهده می‌شود، علی‌رغم آنکه پژوهش‌های متعددی در ارتباط با موضوع انجام شده، اما از دو حیث این پژوهش در مقایسه با آن‌ها دارای نوآوری است.

- استفاده از رویکرد یادگیری ماشین در تحلیل احساسات داده‌های چندوجهی در مقیاس شهر: علی‌رغم آنکه در پژوهش‌های حوزه علم داده از این روش و از داده‌های چندوجهی استفاده شده، اما درحوزه برنامه‌ریزی و مدیریت شهری در ایران، پژوهش چندانی با این رویکرد انجام نشده است.
- تهیه دیتاستی از داده‌های متنی و تصویری در رابطه با شهر. درواقع بسیاری از پژوهش‌های یادگیری ماشین که مترصد بر توسعه الگوریتم هستند با توجه به مسئله پژوهش، از دیتاست‌های آماده و استاندارد بهره می‌برند. چراکه در موارد بسیاری تولید دیتاست با توجه به محدودیت‌های مالی و همچنین فرایند آماده‌سازی داده و برچسب‌گذاری آن‌ها، عملاً بسیار زمان‌بر و پرهزینه است. از این‌رو دیتاست‌های متعددی در پلت‌فرم‌های متعدد قابل دسترس سایر پژوهشگران است. با این توضیحات ازآنجاکه پژوهش‌های کمی درحوزه مدّنظر پژوهش انجام شده، تهیه دیتاست پژوهش و بارگذاری آن در گیت‌هاب که قابلیت استفاده برای سایر پژوهشگران را فراهم می‌آورد، می‌تواند به‌عنوان نوآوری پژوهش قلمداد شود.

مبانی نظری

تحلیل احساسات کاربران ازطریق شبکه‌های مجازی که حدود دو دهه است مورد توجه قرار گرفته، به حوزه‌ای مهم در علوم مختلف تبدیل شده و مخاطبان آن نه تنها صاحبان شرکت‌ها و سیاست‌مدارن، بلکه کاربران هستند. در این میان این حوزه از متن‌کاوی، در مطالعات شهری هم نفوذ کرده و به‌دلیل روش‌مندی آن مورد استفاده برنامه‌ریزان و طراحان شهری قرار گرفته است؛ چه در قالب پژوهش‌هایی که صرفاً تحلیل احساسات را هدف قرار داده‌اند و چه به‌صورت لایه‌ای تلفیقی در پژوهش‌ها. در ادامه در ابتدا اندکی در رابطه با تحلیل احساسات بحث می‌شود. سپس به بررسی انواع زمینه‌های تحلیل احساسات در شهر و انواع داده‌ها برای تحلیل احساسات پرداخته خواهد شد. سپس انواع روش‌های تحلیل احساسات و جایگاه روش‌های مدل‌گرا و یادگیری عمیق بررسی می‌شود.

تحلیل احساسات و رسانه‌های اجتماعی

در پژوهش‌های مختلف، برای سنجش احساسات از روش‌های مختلفی استفاده شده است. براساس پژوهش‌ها، تحلیل و اندازه‌گیری احساسات از سه طریق انجام می‌شود: ۱) رسانه‌های اجتماعی؛ ۲) اپلیکیشن‌ها؛ و ۳) دستگاه‌های پوشیدنی (Fathullah & S.Willis, 2018, 5). هدف این مقاله، نحوه کاربست رسانه‌های اجتماعی در تحلیل احساسات شهری است. داده‌های رسانه‌های اجتماعی در پژوهش‌های مختلف شهری کاربرد دارند، این حوزه‌ها مشتمل بر حوزه منظر (Chapman *et al*., 2018; Tu *et al*., 2017; Chen *et al*., 2017; Silva *et al*., 2014; Oteros-Rozas *et al*., 2018; Tieskens *et al*., 2018; Wu *et al*., 2016)، کاربری زمین (Chen *et al*., 2016)، فضای سبز، گردشگری (Rossi *et al*., 2018)، حمل‌ونقل (Kenyon & Lyons, 2003; Sdoukopoulos *et al*., 2018; Eboli & Mazzulla, 2011; Mazzulla & Forciniti, 2012)، هویت شهری (Kim *et al*., 2020) و تحلیل احساسات هستند.

براساس تعاریف، تحلیل احساسات، زمینهٔ مطالعاتی مربوط به تحلیل نظرات، نگرش‌ها و احساسات افراد نسبت به موجودیت‌هایی مانند محصولات، خدمات، سازمان، مکان‌ها و رویدادها است (Liu, 2012). در واقع تحلیل احساسات، به معنی فهم احساسات مثبت و منفی کاربر در شبکه‌های اجتماعی است (Sim *et al*.,

2020 به نقل‌از Chapman *et al*., 2018). پژوهش‌هایی نیز برای تعیین نوع و قدرت احساسات، انجام شده است (Liu & Zhang, 2012). در رویکرد احساسات شهری بهتر است از مدل‌های پیچیده‌تر استفاده نمود، زیرا صرفاً قطبیت یک متن، کافی به‌نظر نمی‌رسد (Resch *et al*., 2014). اگر هدف بررسی تحلیل احساسات به‌مثابه روش کمی یا کیفی باشد، می‌توان بیان کرد که این نوع از تحلیل، یک روش تحلیلیِ شبه کیفی است، شکلی از تحلیل محتوا که می‌تواند برای مجموعه‌ی کلان‌داده از جمله مجموعهٔ داده‌های رسانه‌های اجتماعی استفاده شود. خلاف تحلیل محتوای سنتی، تحلیل احساسات در روش نوین، فرآیندی خودکار است که با استفاده از فرهنگ لغت احساسات و برنامه‌ی رایانه‌ای برای تحلیل مجموعه‌ی کلان‌داده‌ها مورد استفاده قرار می‌گیرد (Hollander & Renski, 2015).

انواع داده‌های شبکه‌های اجتماعی برای تحلیل احساسات

به‌صورت‌کلی به‌نظر می‌رسد داده‌های مورد استفاده در تحلیل احساسات در رسانه‌های اجتماعی متنی و غیرمتنی هستند.

داده‌های متنی: تحلیل داده‌های متنی، رواج بیشتری داشته و استخراج متن، که به‌عنوان تحلیل متن نیز شناخته می‌شود، به پژوهشگران اجازه می‌دهد تا نگرش کاربران را در پس داده‌های متنی شناسایی کنند. این حوزه که در تقاطع یازیابی اطلاعات، داده‌کاوی، آمار و زبان‌شناسی قرار دارد، اطلاعات معنی‌داری را از داده‌های متنی بدون ساختار استخراج می‌کند (He *et al*., 2013 به نقل‌از Chapman *et al*., 2018). ظهور رسانه‌های اجتماعی به رشد استفاده از متن‌کاوی کمک کرده است. متن‌کاوی تکنیک‌هایی را از چندین حوزه مانند پردازش زبان طبیعی[۱۲]، یادگیری ماشین[۱۳]، بازیابی اطلاعات[۱۴] و داده‌کاوی[۱۵] وام می‌گیرد (Sim *et al*., 2020 به نقل‌از Chapman *et al*., 2018). در این میان شاخه تحلیل احساسات شهری یکی‌از زیرمجموعه‌های متن‌کاوی است.

داده‌های تصویری: فارغ از بررسی متون برای تحلیل احساسات، امروزه تحلیل احساسات ازطریق تصاویر نیز مورد توجه قرار گرفته است. درواقع اگرچه تحلیل احساسات در محتوای متنی به‌طور قابل توجهی توسعه‌یافته، اما استفاده از آن در رابطه با موضوعات تصویری، موضوعی اساسی و ویژه برای پژوهش است (Campos *et al*., 2017; Chen *et al*., 2014; Dhall *et al*., 2015)؛ چراکه اشتراک‌گذاری تصاویر برای بیان حالات و احساسات بسیار رایج شده است (Parrett, 2016; You & Tuncer, 2016). بنابراین به‌نظر می‌رسد که توسعهٔ تکنیک‌های جدید برای این منظور می‌تواند رویکردهای مبتنی‌بر متن را تکمیل کند (Zhan *et al*., 2019; Zhou & Huang, 2017; Riberio *et al*., 2016; Benevenuto *et al*., 2015). براساس بررسی‌ها به‌نظر می‌رسد حتی اگر از یادگیری عمیق استفاده شود (Baly *et al*., 2016; Huang *et al*., 2014) اما همچنان لازم است تا سرویس‌های جدید در بسترهایی که محتواری مشترک آن‌ها عمدتاً بصری است مانند اسنپ چت، اینستاگرام و فلیکر فعال شوند (Bonasoli de Oliveira *et al*., 2020).

داده‌های چندوجهی: تحلیل داده‌های چندوجهی، شامل داده‌های متنی و تصویری (Setiawan *et al*., 2021; Ji *et al*., 2015; Soleymani *et al*., 2017) و یا داده‌های متنی، تصویری و صوتی (Poria *et al*., 2017) به‌عنوان رویکردی جدید در برخی‌از پژوهش‌ها مورد توجه قرار گرفته است.

ازآنجاکه در این پژوهش داده‌های متنی و تصویری مورد توجه هستند، دلیل استفاده از آن‌ها بیان می‌شود. گاهی به‌دلیل مبهم بودن جمله، پردازش متن به تنهایی کافی نیست و بنابراین استفاده از تصاویر می‌تواند به تحلیل احساسات کمک کند. متن و تصاویر بیشتر از سایر داده‌ها درکنار هم قرار می‌گیرند و بیشتر در پژوهش‌ها استفاده شده‌اند. براساس دیدگاه چن و همکاران تصویر و متن در رسانه‌های اجتماعی همبستگی دارند و مکمل یکدیگر هستند (Chen *et al*., 2015).

جایگاه یادگیری ماشین و عمیق در تحلیل احساسات

به‌صورت کلی روش‌های تحلیل احساسات در علوم تحلیل داده مبتنی‌بر روش‌های لغوی[۱۶] (غیرمدل‌گرا) و روش‌های مبتنی‌بر یادگیری ماشین (مدل‌گرا) تقسیم می‌شوند. روش اول که مبتنی‌بر فرهنگ لغت است، با پیش‌پردازش متن آغاز می‌شود. هر پست در ابتدا امتیاز صفر دارد، اگر کلمه‌ای با احساسات مثبت وجود داشته باشد، امتیاز ۱+ به آن اعمال می‌شود و بالعکس. اگر منفی باشد، امتیاز ۱- درنظر گرفته می‌شود (Devika *et al*., 2016 به نقل‌از یوسفی متقاعد و صبوحی، ۱۳۹۸).

همان‌گونه که گفته شد، روش دیگر، روش‌های مدل‌گرا هستند که مشتمل بر یادگیری ماشین و عمیق می‌شوند. یادگیری عمیق، از مباحث جدید درزمینه یادگیری ماشین بوده که اخیراً کاربرد زیادی پیدا کرده است. اصطلاح یادگیری عمیق برای نخستین‌بار در سال ۱۹۸۶ مطرح و در سال ۲۰۰۰ از آن در شبکه‌های عصبی مصنوعی استفاده شد (Hand *et al*., 2001 به نقل‌از زاهدی حقیقی، سخایی و دلیری، ۱۳۹۸؛ سلیمانی و همکاران، ۱۳۹۵ به نقل‌از راستگو و کیانی، ۱۳۹۸). انگیزه به وجود آوردن الگوریتم‌های شبکه عصبی و یادگیری عمیق از ساختار عصبی از مغز انسان الهام گرفته شده که در آن یاخته‌های عصبی با فرستادن پیام به یکدیگر، ادراک را امکان‌پذیر می‌کنند. الگوریتم‌ها در تلاش هستند ویژگی‌های انتزاعی سطح بالا در مجموعه داده‌ها را با استفاده از یک گراف عمیق مدل کنند که دارای چندین لایه پردازشی و متشکل از چندین لایه تبدیلات خطی و غیرخطی هستند. ویژگی‌هایی که از شبکه‌های مبتنی‌بر یادگیری عمیق استخراج می‌شوند، نسبت به ویژگی‌های متداولی مانند لبه، کوله کلمات، هیستوگرام رنگ ازنظر معنا، سطح بالاتری دارند و نزدیک‌تر به درک انسان محسوب می‌شوند، بنابراین استفاده از آن‌ها در دسته‌بندی، منجر به نتایج مطلوب‌تری می‌شود. برای مثال، شبکه عمیق مبتنی‌بر تصاویر چهره ممکن است منجر به استخراج ویژگی‌های معناداری مانند شکل‌های چشم، بینی، فاصله ابروها و... شود (راد، موسوی و وردی، ۱۳۹۹).

با این توضیحات، یادگیری عمیق، روشی جدید در یادگیری ماشین است. در یادگیری عمیق، ویژگی‌های غیرخطی چندین لایه استخراج می‌شود و به یک دسته‌بند اعمال و آن هم این ویژگی‌ها را با هم ترکیب می‌کند تا بتواند پیش‌بینی را انجام دهد (فولادی، فرسی و محمدزاده، ۱۳۹۹). سه ویژگی که سبب برتری روش‌های یادگیری عمیق نسبت به روش‌های سنتی می‌شود عبارت‌اند از: ۱) توانایی این روش‌ها در یادگیری مستقیم از داده‌های خام؛ ۲) داشتن ساختاری سلسله‌مراتبی و عمیق؛ و ۳) عمومی و بهینه بودن روش نسبت به روش‌های سنتی (موسوی، عبادی و کیانی، ۱۳۹۸).

فرایند یادگیری عمیق مشابه فرایند یادگیری ماشین است. با این تفاوت که با توجه به نوع داده‌ها، تفاوت‌هایی نیز در فرایند مشاهده می‌شود. اولین مرحله در یادگیری عمیق، آماده‌سازی داده‌ها و پس‌از آن پیش‌پردازش آن‌ها است. اگر نوع داده‌ها تصویری باشد، باید پیش‌پردازش تصاویر پایگاه داده انجام شود. درواقع در این گام، داده‌هایی که ممکن است شبکه عصبی پیچیده را گمراه کنند، کاهش می‌یابند. برای نمونه، لازم است تا ابعاد ثابت در ورودی مدّنظر قرار گیرد. در گام بعد، تصاویر پیش‌پردازش شده به یک شبکه عصبی عمیق در مرحله بعد وارد می‌شوند و با استفاده از معماری خاص شبکه عصبی عمیق، مراحل آموزش و تست بر روی تصاویر اعمال می‌شود و درنهایت داده‌ها دسته‌بندی می‌شوند (فولادی، فرسی و محمدزاده، ۱۳۹۹). مرحله نهایی در فرایند یادگیری ماشین و یادگیری عمیق، معیار ارزیابی است که براساس ماتریس درهم‌آمیختگی محاسبه می‌شود که دارای معیارهای دقت[۱۷]، صحت[۱۸]، بازخوانی[۱۹] و امتیاز اف.یک.[۲۰] است (راد، موسوی و وردی، ۱۳۹۹).

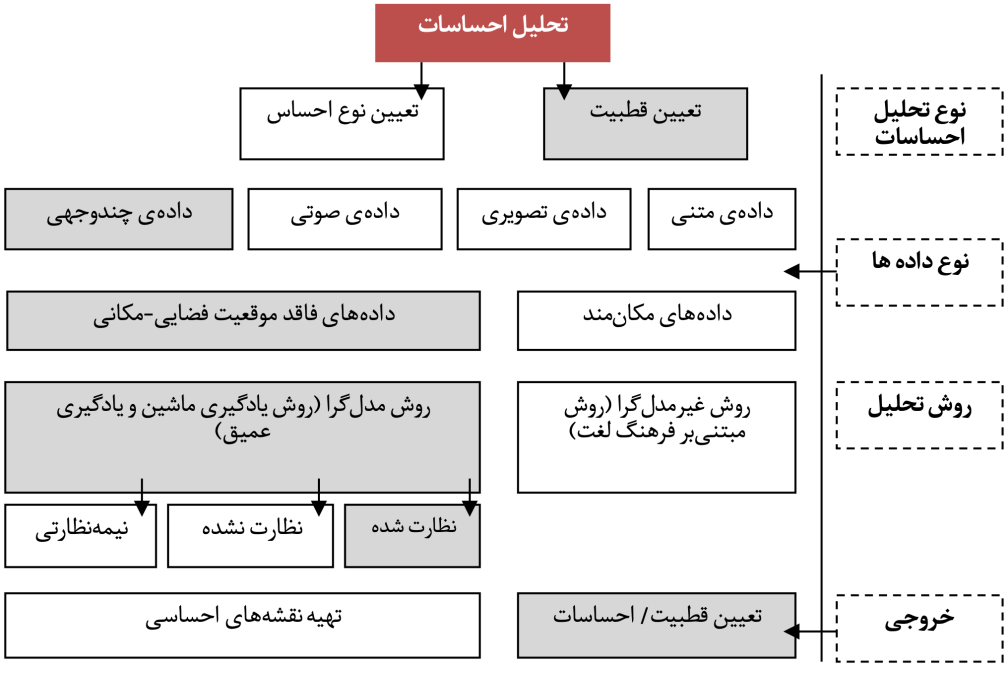
یکی‌از مهم‌ترین مدل‌های یادگیری عمیق، شبکه‌های عصبی پیچشی (سی.ان.ان.) یا همان کانولوشنی هستند (راد، موسوی و وردی، ۱۳۹۹) که در این پژوهش نیز از آن استفاده شده است. این شبکه‌ها از تعدادی

لایه کانولوشن، لایه فعال‌سازی[۲۱]، لایه ادغام[۲۲]، لایه سافت مکس[۲۳] تشکیل شده‌اند (Schmidhuber, 2014 به نقل‌از راستگو و کیانی، ۱۳۹۸). به‌صورت کلی دو نوع روش برای استفاده از الگوریتم‌ها وجود دارد، ۱) محقق می‌تواند شبکه عصبی کانولوشنی خود را با لایه‌های مدنظر خود نوشته و ماشین را با آن آموزش داده و اعتبارسنجی نمایند و ۲) یا از الگوریتم‌های دارای ساختار مشخص و شناخته شده که روی داده‌هایی بسیار بزرگ اعتبارسنجی شده و قابلیت فراخوانی در پلت فرم ژوپیتر یا گوگل کولب دارند، استفاده نماید.

چارچوب مفهومی پژوهش

با توجه به مطالب ارائه شده، در این بخش به‌صورت کلی جایگاه یادگیری ماشین و عمیق در تحلیل احساسات شهری ارائه می‌شود (شکل ۱). همان‌گونه که بیان شد، تحلیل احساسات می‌تواند در قالب تعیین قطبیت یا در قالب تعیین انواع احساسات مانند ۷ دسته اکمن باشد. تحلیل احساسات در عصر اطلاعات و ظهور فضاهای اجتماعی مجازی، مبتنی‌بر داده بوده و انواع داده‌ها می‌توانند به‌صورت متن، تصویر، صوت و چندوجهی باشند. ضمن آنکه نوع داده‌ها می‌تواند مکان‌مند یا غیر از آن باشد.

روش‌های سنجش احساسات در داده‌ها می‌تواند مبتنی‌بر روش‌های مدل‌گرا و غیرمدل‌گرا باشد که در روش دوم مبتنی‌بر رویکرد یادگیری ماشین و عمیق است. به‌منظور تحلیل داده‌های فضاهای مجازی ازطریق روش دوم با توجه به نوع داده ازحیث محتوا یعنی متنی یا غیر آن و همچنین ازحیث برچسب‌گذاری شده یا برچسب‌گذاری نشده از الگوریتم‌های نظارت شده، نظارت نشده و نیمه‌نظارتی بهره برد. در این پژوهش با توجه به نوع داده‌ها نوع تحلیل از نوع تعیین قطبیت، نوع داده‌ها چندوجهی، روش مدل‌گرا و مبتنی‌بر رویکرد یادگیری عمیق نظارت شده و به‌دلیل مکان‌مند نبودن داده‌ها، صرفاً تعیین قطبیت انجام شده است.



شکل ۱. چارچوب پژوهش

روش پژوهش

همان‌گونه که پیش‌از این بیان شد، این پژوهش از نوع کمی و مبتنی‌بر روش‌های مدل‌گرا در تحلیل احساسات است و از الگوریتم‌های یادگیری عمیق بهره برده است. با توجه به این موارد، لازم است تا نوع و حجم داده‌ها برای آموزش ماشین، شبکه اجتماعی منتخب، نحوه برچسب‌زنی داده‌ها، انواع الگوریتم‌ها و حجم داده‌های آموزش و تست مشخص شود.

- شبکه اجتماعی و بستر منتخب: شبکه اجتماعی مورد استفاده در این پژوهش، توئیتر است، اگرچه اینستاگرام نیز، مشتمل بر داده متنی و تصویری است، اما به‌دلیل رایگان نبودن دسترسی به داده‌ها، محدودیت برای دانلود، وجه مثبت بیشتر دیدگاه‌های ثبت شده در آن و شخصی بودن پست‌ها، کنار گذاشته شد. علی‌رغم آنکه نقدهایی چون دسترسی کم کاربران ایرانی به‌دلیل فیلترینگ به توئیتر وجود دارد، اما به‌دلیل نوع داده منتخب در پژوهش و همچنین به‌دلیل قابلیت استفاده از توئیتر استریمینگ جهت دسترسی به داده‌ها، از این پلت‌فرم استفاده شده است.
- نوع داده‌ها: متنی و تصویری یعنی چندوجهی هستند. نوع فراداده استخراج شده با توجه به شبکه اجتماعی منتخب شامل اطلاعات زیر است: عبارت جست‌وجو، لینک تصویر، موقعیت فرد، توئیت آی.دی.، زبان توئیت، تعداد لایک‌ها (پسندها)، تعداد ری‌توئیت‌ها و متن توئیت.
- نوع و محتوای داده‌ها: داده‌هایی جمع‌آوری شده در رابطه با ۸ کلان‌شهر رسمیِ ایران (تهران، مشهد، شیراز، کرج، تبریز، قم، اهواز، اصفهان) هستند. انتخاب کلان‌شهرها به‌عنوان بستری که داده در رابطه با آن جمع‌آوری شده، به‌صورت تصادفی نبوده و به لحاظ زمانی ٤ ماه برای انتخاب داده‌ها و استانداردسازی آن زمان صرف شده است. در ابتدا چند نوع فضا صرفاً در مقیاس شهر تهران شامل محلات، فضاهای گردشگری، پارک‌های شهری و میادین شهری مورد توجه قرار گرفت. در ادامه پس‌از برنامه‌نویسی در فضای گوگل کولب برای استخراج توئیت‌ها، تعداد توئیت‌های موجود در هریک از فضاهای پیش گفته در بازه زمانی یک‌ساله بارگیری و بررسی شد. پس‌از ارائه خروجی‌های اولیه، مشخص شد که در برخی‌از محلات، تعداد توئیت‌ها کم و در تعداد دیگری تعداد توئیت‌ها زیاد است و درباره برخی‌از محلات توئیتی وجود ندارد. لذا امکان تحلیل در مقیاس محلات وجود نداشت. در رابطه با مقاصد گردشگری، میادین و پارک‌های شهری نیز وضعیت مشابه‌ای دیده شد (یکی‌از مشکلاتی که در روند تهیه دیتاست با حجم مطلوب وجود داشت، این بود که تعداد توئیت‌هایی که هم‌زمان شامل متن و تصویر باشد، بسیار محدود بود). به‌همین دلیل در ادامه از عبارت «#شهر_تهران» برای تهیه مجموعه داده‌ها استفاده شد، ازآنجاکه تعداد توئیت‌هایی که هم دربرگیرنده متن و هم تصویر باشد، در یک‌سال به‌عنوان بازه زمانی بسیارکم بود، بازه زمانی ۱۰ ساله (فرودین ماه ۱۳۹۱ تا اسفند ماه ۱٤۰۰) درنظر گرفته شد؛ با این وجود همچنان تعداد توئیت‌های قابل اتکاء کم بود، توئیت‌های مربوط به هفت کلان شهر رسمی دیگر شامل: مشهد، اصفهان، شیراز، تبریز، کرج، قم و اهواز نیز بارگیری شد. پس‌از برچسب‌زنی داده‌ها مشخص شد که بیشتر توئیت‌ها بار احساسی منفی دارند و ممکن است روند یادگیری ماشین را دچار مشکل نمایند، بنابراین به‌منظور افزایش تعداد توئیت‌ها و همچنین افزایش توئیت‌ها با بار احساسی مثبت، لیست نقاط گردشگری شهرها نیز تهیه و توئیت‌های چندوجهی در بازه زمانی تعیین شده در رابطه با آن‌ها در قالب فایل سی.اس.وی. استخراج شد. از مجموع توئیت‌های مربوط به ۸ شهر، که با هشتگ «شهرنام شهر» دانلود شده، حدود٦۰ درصد غیرمرتبط بوده‌اند. همچنین ۸۰ درصد توئیت‌های مربوط به محدوده‌های گردشگری نیز غیرمرتبط تشخیص داده شدند. برخی‌از ملاحظاتی که در نوشتن کد اولیه یا دانلود خروجی درنظر گرفته شد، به شرح زیر است:

- زبان فارسی به‌عنوان عامل محدودکننده لحاظ شد تا توئیت‌هایی که به زبان اردو، عربی و انگلیسی بودند، بارگیری نشوند.
- ری‌توئیت‌ها از روند بارگیری حذف شدند.
- همچنین انواع نگارش‌هایی که برای یک فضا چه به زبان فارسی و چه زبان انگلیسی ممکن بود، در قالب فایل ورودی درنظر گرفته شد.
- توئیت‌هایی که توسط ربات‌ها پست شده بود یا به‌صورت تبلیغاتی بودند، حذف شدند.

- برچسب‌زنی داده‌ها: در این پژوهش تمامی داده‌های بارگیری شده (۳۶۵۵ رکورد) بررسی و پس‌از حذف موارد غیرمرتبط، توئیت‌های باقی‌مانده (۹۲۰ رکورد) برچسب‌گذاری شدند. در ابتدا داده‌های تصویری، براساس قطبیت برچسب‌گذاری شدند. تصاویری که حس مثبت انتقال می‌دادند، برچسب مثبت و با کد ۲، تصاویری که حس منفی القا می‌کردند، منفی و با کد ۰ و تصاویر خنثی، با کد ۱ برچسب‌گذاری شدند. نکته قابل توجه در روند برچسب‌زنی قطبیت داده‌های تصویری آن بود که فارغ از هر نوع قضاوت و نگاه سیاسی، داده‌های تصویری برچسب‌زنی شوند در این بخش تعدادی تصاویر تکراری نیز شناسایی و به‌صورت هایلایت مشخص شدند تا در گام بعد با بررسی متن‌ها، چنانچه هم متن و هم تصویر، تکراری بودند، سطر مورد نظر یا توئیت مربوطه حذف شود. سپس برچسب‌زنی داده‌های متنی مورد توجه قرار گرفت و آن‌ها نیز براساس قطبیت برچسب‌زنی شدند. درحالتی‌که محقق در برچسب‌زنی داده‌ها دچار تردید شده، از نظر دو فرد دیگر برای برچسب‌زنی داده‌ها استفاده شده و نظری که دو یا سه نفر روی آن توافق داشته‌اند، به‌عنوان برچسب نهایی درنظر گرفته شده است.
- الگوریتم‌های مورد استفاده: در این پژوهش از یادگیری عمیق نظارت شده بهره برده شده است. در رابطه با داده‌های متنی، تصویری و چندوجهی، در هر بخش نوع الگوریتم منتخب بیان خواهد شد.
- حجم داده‌های آموزش و تست و روش اعتبارسنجی: در این پژوهش از روش اعتبارسنجی متقابل[۲۴] استفاده شده که روشی برای ارزیابی و بررسی تعمیم‌پذیری مدل است. در این روش، داده‌ها به دو بخش تقسیم می‌شوند و عملکرد مدل روی داده‌های جدید[۲۵] انجام می‌شود. از میان روش‌های چون هولد اوت[۲۶]، کی فولد[۲۷]، لیو-وان-اوت[۲۸]، از روش هولد اوت استفاده شده است. بنابراین در این روش داده‌ها باید به‌صورت تصادفی به دو گروه آموزشی و تست تقسیم شوند. در این پژوهش نسبت داده‌های آموزش به تست، ۷۰ به ۳۰ درنظر گرفته شده و برای مشاهده اعتبارسنجی از ماتریس درهم‌آمیختگی[۲۹] استفاده شده است.

تحلیل داده‌ها و یافته‌ها

در این پژوهش، کدها با استفاده از پایتون و در فضای ابری گوگل کولب نوشته شده است. در ابتدای امر لازم است که کتابخانه‌های مورد نیاز برای یادگیری ماشین شامل پانداز، کِرَس و اسکای.لرن فراخوانی شوند.

پیش‌پردازش داده‌ها

پس‌از فراخوانی داده‌ها، پیش‌پردازش داده‌ها انجام شد. دو نوع داده در فرایند پیش‌پردازش تشخیص داده شد، داده‌های صحیح[۳۰] و غیرصحیح[۳۱] داده‌های رشته‌ای[۳۲] یا شی[۳۳]. برای تعیین داده‌های گم‌شده از میان دو روش معمول، اگر داده، از نوع اول باشد، می‌توان با شاخص‌های آماری چون میانگین، ماکسیمم، مینیمم و... مقداری برای داده‌های گم‌شده درنظر گرفت؛ اما اگر داده از نوع دوم باشد، می‌توان از روش بیشترین تکرار[۳۴] یا مُد استفاده کرد. در ادامه به‌دلیل وجود چند ویژگی رشته‌ای در داده‌ها، ضرورت تبدیل آن‌ها به کد وجود

داشت. برای مثال داده‌های مربوط به موقعیت و عبارت جست‌وجو از این دست بودند که با بهره‌گیری از روش کدگذاری برچسب[۳۵] برای هر رشته، یک کد اختصاص داده شد و داده‌ها جهت آموزش آماده شدند.

نرمال‌سازی داده‌ها

برای نرمال‌سازی داده‌ها از روش‌هایی چون مقیاس‌بندی مینیمم-ماکسیمم[۳۶] و مقیاس‌بندی استاندارد[۳۷] استفاده می‌شود. در این پژوهش از روش اول استفاده شد. هدف این بخش نرمال کردن داده‌ها، بین بازه صفر و یک است. چراکه دامنه اعداد استفاده شده در مجموعه داده‌ها بسیار متغیر است و ممکن است فرایند آموزش ماشین را دچار اختلال کند. در این پژوهش از رابطه زیر استفاده شده است:

رابطه (۱)

$$X\,scaled = \frac{X - Xmin}{Xmin - Xmax}$$

آموزش ماشین

در این بخش با توجه به اینکه داده‌های این پژوهش مشتمل بر متن و تصویر هستند، در ابتدا فرایند آموزش و تست ماشین برای داده‌های متنی و تصویری به‌صورت مجزا انجام و سپس به‌صورت ترکیبی انجام شد. براساس برچسب‌گذاری انجام شده، حدود ۵۰ درصد داده‌ها در دسته قطبیت منفی، حدود ۴۰ درصد در دسته قطبیت مثبت و مابقی آن در دسته قطبیت خنثی هستند. بنابراین حجم داده‌ها در قطبیت خنثی، کم و محدود است. به این منظور به‌جای انتخاب داده‌ها به‌صورت نرمال[۳۸] که در شرایطی که داده‌های هر دسته تقریباً مساوی هستند، مورد استفاده قرار می‌گیرد؛ از انتخاب داده‌ها به‌صورت وزن‌دار[۳۹] استفاده شد.

الف- آموزش داده‌های متنی

تمیز کردن: در ابتدا لازم است داده‌های متنی تمیز[۴۰] شوند. با توجه به اینکه زبان متون، فارسی بوده از کتابخانه هضم[۴۱] استفاده شد. این کتابخانه توسط دانشجویان دانشگاه علم‌وصنعت ایران در سال ۱۳۹۲ به‌صورت متن باز و با استفاده از کتابخانه ان.ال.تی.کی. منتشر شده است (وب‌سایت خانه بیگ دیتای ایران، ۱۳۹۷). برای تمیز کردن داده‌های متنی موارد زیر مورد توجه قرار گرفت:

- حذف تمام اعداد و علائم؛
- حذف لینک‌ها؛
- حذف فاصله‌های اضافی ابتدا و انتها؛
- نرمال‌سازی داده‌ها با دستور normalize Hazm انجام شد.
- توکنایز کردن کلمات؛
- ریشه‌یابی کلمات؛ و
- حذف ایست‌واژه‌ها.

توکنایز کردن: با استفاده از توکنایزر کرس، امکان شماره‌گذاری جملات و تبدیل آن‌ها به بردار وجود دارد. همچنین لازم است که اندازه فرهنگ لغت مشخص شود که در اینجا با تست اعداد مختلف، مقداری عددی اختصاص یافته به MAX_NB_WORDS، ۲۰۰ هزار کلمه درنظر گرفته شد.

همسان‌سازی طول جملات: ازآنجاکه در فرایند برداری کردن جملات، طول جملات یکسان نیستند، لازم است تا با استفاده از دستور لایه‌گذاری صفر[۴۲]، بردارها یکسان شوند.

تعبیه‌سازی کلمات[۴۳]: در ارائه داده‌های متنی به ماشین به‌صورت قابل فهم، می‌توان از روش‌های متعددی استفاده نمود، در این پژوهش به‌جای استفاده کردن از روش برداری کردن یا وکتورایز کردن داده‌ها، از روش تعبیه‌سازی کلمات استفاده شده است. از میان الگوریتم‌هایی چون ورود.وک.۲، گلاو و فست تکست؛ در این

پژوهش از فست تکست[۴۴] استفاده شد. این کتابخانه توسط آزمایشگاه تحقیقات هوش مصنوعی فیس‌بوک ایجاد شده است. این مدل این امکان را به پژوهشگر می‌دهد که یک الگوریتم یادگیری بدون نظارت یا یادگیری نظارت شده برای به دست آوردن نمایش‌های برداری به کلمات را ایجاد کند. طول لایه تعبیه‌سازی[۴۵] در این پژوهش ۳۰۰ درنظر گرفته شد.

الف) آموزش ماشین برای تشخیص قطبیت متن با الگوریتم هیبریدی

در این بخش با بهره‌گیری از شبکه عصبی، دو نوع داده به‌عنوان ورودی به شبکه عصبی درنظر گرفته شده، متن و همچنین سایر ویژگی‌های موجود در مجموعه داده‌ها (به‌عنوان متادیتا شامل موقعیت، تعداد ری‌توئیت‌ها، تعداد پسندها و...). خروجی، برچسب قطبیت که توسط پژوهشگر تعیین شده بود، لحاظ شد (جدول ۲). براساس ماتریس درهم‌آمیختگی مشاهده می‌شود که میزان دقت ماشین با بهره‌گیری از شبکه عصبی در تعیین قطبیت متن ۷۷ درصد است و شبکه هیبریدی طراحی شده مناسب عمل کرده است.

جدول ۲. معماری شبکه عصبی طراحی شده برای تعیین قطبیت متن

```
__________________________________________________________________________________
Layer (type)                 Output Shape              Param #     Connected to
==================================================================================
input_meta (InputLayer)      [(None, 9)]               0           []
txt_input_cnn (InputLayer)   [(None, 100)]             0           []
meta_dense_1 (Dense)         (None, 128)               1280        ['input_meta[0][0]']
Embedd_cnn (Embedding)       (None, 100, 300)          1313400     ['txt_input_cnn[0][0]']
meta_dense_2 (Dense)         (None, 64)                8256        ['meta_dense_1[0][0]']
data_cov_1 (Conv1D)          (None, 100, 256)          153856      ['Embedd_cnn[0][0]']
meta_dense_3 (Dense)         (None, 32)                2080        ['meta_dense_2[0][0]']
flatten_11 (Flatten)         (None, 25600)             0           ['data_cov_1[0][0]']
flatten_10 (Flatten)         (None, 32)                0           ['meta_dense_3[0][0]']
meta_dense_4 (Dense)         (None, 1)                 25601       ['flatten_11[0][0]']
concatenate_3 (Concatenate)  (None, 33)                0           ['flatten_10[0][0]',
                                                                    'meta_dense_4[0][0]']
dropout_15 (Dropout)         (None, 33)                0           ['concatenate_3[0][0]']
dens_1 (Dense)               (None, 512)               17408       ['dropout_15[0][0]']
out_layer (Dense)            (None, 3)                 1539        ['dens_1[0][0]']

==================================================================================
Total params: 1,523,420
Trainable params: 1,523,420
Non-trainable params: 0
__________________________________________________________________________________
```

در معماری شبکه پیشنهادی مشاهده می‌شود، اهمیت متن معادل ۹۰ درصد و اهمیت متادیتا ۱۰ درصد لحاظ شده است. در ادامه با توجه به اینکه این سؤال مطرح بود که آیا صرفاً با درنظر گرفتن متن ممکن است نتیجه بهتری به دست آید، تأثیر متادیتا ۱ درنظر گرفته شد. با اعمال آن و آموزش مدل مشخص شد که لحاظ نکردن متادیتا، تأثیر منفی در دقت مدل داشته و بنابراین اضافه نمودن آن و طراحی شبکه هیبریدی موفق عمل نموده است.

ب- آموزش داده‌های تصویری

در ابتدا داده‌های تصویری فراخوانی شدند. پیش‌پردازش داده‌های تصاویری به‌دلیل مرتبط بودن داده‌های

تصویری به متنی، امکان‌پذیر نبود. در آموزش ماشین، سه الگوریتم مورد توجه قرار گرفت: الگوریتم شبکه عصبی با چهار بلوک، شبکه از پیش آموزش دیدهٔ دنس‌نت۱۲۱ و شبکه وی.جی.جی.۱۹.

الگوریتم شبکه عصبی

معماری شبکه عصبی طراحی شده برای آموزش ماشین در جدول (۳) ارائه شده است. همان‌طور که مشاهده می‌شود، این شبکه عصبی چهار بلوک را در بر می‌گیرد، اما متأسفانه میزان دقت مدل آموزش دیده، ۶۱ درصد است. بنابراین علی‌رغم آنکه شبکه عصبی در آموزش ماشین در داده‌های متنی به‌خوبی عمل کرده بود، در داده‌های تصویری عملکرد مناسبی نداشت.

جدول ۳. معماری شبکه عصبی پیشنهادی برای آموزش ماشین برای تصویر

Layer (type)	Output Shape	Param #
input_layer (InputLayer)	[(None, 256, 256, 3)]	0
conv1.1 (Conv2D)	(None, 256, 256, 64)	832
conv1.2 (Conv2D)	(None, 256, 256, 64)	16448
block1_pool (MaxPooling2D)	(None, 128, 128, 64)	0
conv2.1 (Conv2D)	(None, 128, 128, 64)	36928
conv2.2 (Conv2D)	(None, 128, 128, 64)	36928
block2_pool (MaxPooling2D)	(None, 64, 64, 64)	0
conv3.1 (Conv2D)	(None, 64, 64, 32)	8224
conv3.2 (Conv2D)	(None, 64, 64, 32)	4128
block3_pool (MaxPooling2D)	(None, 16, 16, 32)	0
conv4.1 (Conv2D)	(None, 16, 16, 16)	12816
conv4.2 (Conv2D)	(None, 16, 16, 16)	6416
block4_pool (MaxPooling2D)	(None, 4, 4, 16)	0
flatten_2 (Flatten)	(None, 256)	0
dropout_1 (Dropout)	(None, 256)	0
dense (Dense)	(None, 256)	65792
dense_1 (Dense)	(None, 3)	771

Total params : 189,283
Trainable params : 189,283
Non-trainable params : 0

الگوریتم شبکه دنس‌نت۱۲۱

این الگوریتم به‌عنوان یک شبکه ازپیش آموزش دیده، مورد استفاده قرار گرفت. این شبکه پیچشی متراکم توسط یکی‌از مقالات سی.وی.پی.آر. ۲۰۱۷ مطرح و با بیش‌از ۲۰۰۰ ارجاع جایزه بهترین مقاله را به خود اختصاص داد. این شبکه حاصل همکاری مشترک دانشگاه کرنول، دانشگاه سینگوا و تیم تحقیقاتی فیس‌بوک ای.آی. است. این شبکه به‌دلیل داشتن اتصالات متراکم، پارامترهای کمتر و دقت بالاتری نسبت به رزنت و پری-اکنیویشنن رزنت[۴۶] دارد.

با اجرای این شبکه و آموزش ماشین براساس آن، دقت ماشین در پیش‌بینی ۵۸ درصد محاسبه شد. بنابراین این شبکه نیز چندان موفق عمل نکرد. بررسی ماتریس درهم‌آمیختگی نشان داد که با توجه به حجم داده‌های

تست، الگوریتم مورد استفاده در طبقه منفی و مثبت، ۷۰ درصد موفق بوده، اما در داده‌های خنثی بسیار ضعیف عمل کرده است.

الگوریتم وی.جی.جی.۱۹

با توجه به اینکه دو الگوریتم پیشین در یادگیری ماشین برای شناسایی دسته‌بندی قطبیت احساسات تصویر چندان مناسب عمل نکردند، از الگوریتم وی.جی.جی.۱۹ استفاده شد. وی.جی.جی.نت یک شبکه عصبی کانولوشنی است که در سال ۲۰۱۴ کارن سایمون و آندره زیسرمن معرفی کردند. این شبکه نسخه‌های مختلفی دارد که این نسخه‌ها شامل ۱۱ تا ۱۹ لایه هستند.

با اجرای الگوریتم بر روی داده‌های تصویری، مشخص شد که این الگوریتم به‌صورت قابل قبولی بسیار موفق‌تر از دو الگوریتم پیشین در اجرای مدل است و دقت آن ۶۸ درصد است. بررسی ماتریس درهم‌آمیختگی نشان می‌دهد که در داده‌های تست، در طبقه منفی، الگوریتم وی.جی.جی.۱۹ در شناسایی تصاویر با بار احساسی منفی، ۸۸ درصد موفق بوده است. در رابطه با داده‌های تصویری طبقه مثبت، ۶۵ درصد موفق بوده است.

ج- آموزش داده‌های چندوجهی (متنی و تصویری)

تا اینجا به‌صورت جداگانه داده‌های متنی با استفاده از الگوریتم سی.ان.ان. و داده‌های تصویری به‌صورت مجزا با سه الگوریتم آموزش داده شدند و یادگیری ماشین انجام شد. همان‌طور که پیش‌از این بیان شد درحوزه داده‌های چندوجهی، داده‌های متنی و تصویری درکنار هم، احساسات را بهتر از داده‌های متنی و تصویری به‌صورت مجزا نشان می‌دهند. براساس نتایج یادگیری ماشین در داده‌های تصویری و متنی به‌صورت مجزا مشخص شد که عموماً متن‌ها گویاتر از تصاویر در انتقال احساسات بوده‌اند. فارغ از این موارد با توجه به هدف پژوهش که آموزش ماشین برای تحلیل داده‌های متنی و تصویری به‌صورت هم‌زمان است، از مدل‌های ترکیبی باید استفاده شود.

مدل ترکیبی یا هیبریدی

در این مدل ترکیبی هم داده‌های متنی، هم داده‌های تصویری و هم داده‌های متا به‌عنوان ورودی به شبکه داده شده و با توجه به اینکه داده‌های متنی، ادراک کاربران از تصویر هستند، قطبیت متن به‌عنوان خروجی درنظر گرفته شد. در این مدل به داده‌های متنی و متا وزن بیشتر و به داده‌های تصویر اهمیت کمتری داده شد. به این معنی که مدل طراحی شده ۱۶ ویژگی از تصویر، ۲۵۶ ویژگی از داده‌های متا و ۵۱۲ ویژگی از داده‌های متنی استفاده شد. به دو دلیل تعداد ویژگی‌های استخراج شده از تصویر، صرفاً ۱۶ ویژگی است: ۱) یادگیری شبکه‌ای که ۳ ورودی دارد، پیچیده است؛ ۲) شبکه طراحی شده برای استخراج ویژگی از تصویر به‌دلیل کم بودن تعداد لایه‌های آن، قابلیت استخراج ویژگی‌های بیشتر را ندارد.

ازآنجاکه درصورت طراحی شبکه با لایه‌های بیشتر، به‌دلیل ۳ ورودی درنظر گرفته شده برای شبکه، امکان مدیریت مدل طراحی شده و یادگیری آن دشوار است؛ به‌نظر می‌رسد راه‌حل، استفاده از الگوریتمی ازپیش طراحی شده است که ماشین به‌واسطه آن به‌خوبی از داده‌های تصویری یاد گرفته است. در ادامه مدل هیبریدی طراحی شده در پایتون ارائه شده است. دقت این مدل، ۶۶ درصد به دست آمده که نشان‌دهنده عملکرد نامناسب آن است.

جدول ۴. معماری مدل هیبریدی برای آموزش ماشین در داده‌های چندوجهی

```
Layer (type)                Output Shape       Param #    Connected to
==================================================================================
 input_image (InputLayer)    [(None, 256, 256, 3  0         []
                             )]
conv1 (Conv2D)               (None, 256, 256, 64  832       ['input_image[0][0]']
                             )
input_meta (InputLayer)      [(None, 9)]        0          []
pool1 (MaxPooling2D)         (None, 128, 128, 64  0         ['conv1[0][0]']
                             )
meta_dense_1 (Dense)         (None, 512)        5120       ['input_meta[0][0]']
txt_input_cnn (InputLayer)   [(None, 100)]      0          []
dropout_46 (Dropout)         (None, 128, 128, 64  0         ['pool1[0][0]']
                             )
dropout_49 (Dropout)         (None, 512)        0          ['meta_dense_1[0][0]']
Embedd_cnn (Embedding)       (None, 100, 300)   1313400    ['txt_input_cnn[0][0]']
conv2 (Conv2D)               (None, 128, 128, 64  16448     ['dropout_46[0][0]']
                             )
meta_dense_2 (Dense)         (None, 256)        131328     ['dropout_49[0][0]']
data_cov_1 (Conv1D)          (None, 100, 256)   153856     ['Embedd_cnn[0][0]']
pool2 (MaxPooling2D)         (None, 25, 25, 64)  0         ['conv2[0][0]']
dropout_50 (Dropout)         (None, 256)        0          ['meta_dense_2[0][0]']
flatten_23 (Flatten)         (None, 25600)      0          ['data_cov_1[0][0]']
 dropout_47 (Dropout)        (None, 25, 25, 64)  0         ['pool2[0][0]']
meta_dense_3 (Dense)         (None, 256)        65792      ['dropout_50[0][0]']
dropout_52 (Dropout)         (None, 25600)      0          ['flatten_23[0][0]']
flatten_21 (Flatten)         (None, 40000)      0          ['dropout_47[0][0]']
flatten_22 (Flatten)         (None, 256)        0          ['meta_dense_3[0][0]']
txt_dense_3 (Dense)          (None, 512)        13107712   ['dropout_52[0][0]']
image_dense_1 (Dense)        (None, 16)         640016     ['flatten_21[0][0]']
dropout_51 (Dropout)         (None, 256)        0          ['flatten_22[0][0]']
dropout_53 (Dropout)         (None, 512)        0          ['txt_dense_3[0][0]']
dropout_48 (Dropout)         (None, 16)         0          ['image_dense_1[0][0]']
concatenate_6 (Concatenate)  (None, 784)        0          ['dropout_51[0][0]',
                                                            'dropout_53[0][0]',
                                                            'dropout_48[0][0]']
dropout_54 (Dropout)         (None, 784)        0          ['concatenate_6[0][0]']
dens_1 (Dense)               (None, 512)        401920     ['dropout_54[0][0]']
out_layer (Dense)            (None, 3)          1539       ['dens_1[0][0]']
==================================================================================
Total params: 15,837,963
Trainable params: 15,837,963
Non-trainable params: 0
```

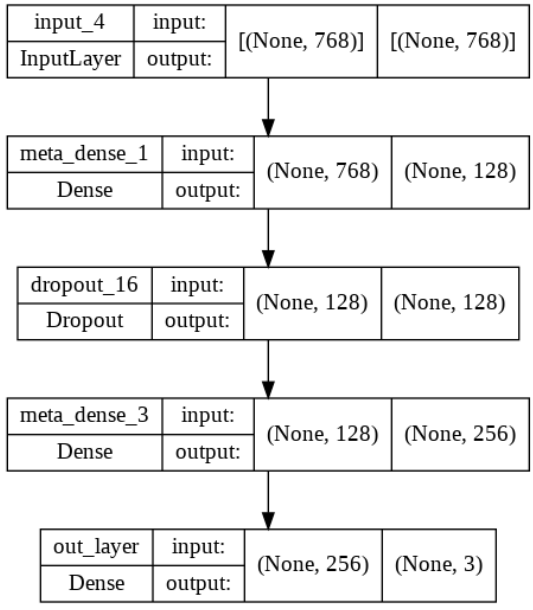
بررسی ماتریس درهم‌آمیختگی در مدل ترکیبی طراحی شده نشان می‌دهد که مدل طراحی شده در شناسایی قطبیت طبقه صفر با احساسات منفی، دقت ۸۵ درصدی داشته و در طبقه ۲ یعنی داده‌ها با احساسات مثبت حدود ۶۵ درصد و در طبقه ۱ یعنی داده‌های خنثی حدودا ۵۰ درصد دقت داشته است.

مدل ترکیبی مبتنی‌بر ترکیب مدل وی.جی.جی.۱۹ مربوط به تصویر و مدل سی.ان.ان. مربوط به متن

در مدل ترکیبی پیشین دو مشکل وجود داشت: ۱) کم بودن تعداد ویژگی‌های تصویر در یادگیری ماشین و ۲) دقت نهایی مدل طراحی شده. بنابراین از روش دوم در تعیین احساسات داده‌های چندوجهی استفاده شد. در این روش ویژگی‌های تصویر و متن و داده‌های متا به‌صورت هم‌زمان به شبکه برای یادگیری داده شد. ازآنجاکه شبکه آموزش دیده شده وی.جی.جی.۱۹ در رابطه با داده‌های تصویری از دقت مناسبی در یادگیری ماشین برخوردار بود از این شبکه استفاده شد و ویژگی‌های تصاویر از این شبکه استخراج شد (شکل ۲). در رابطه با داده‌های متنی و متا هم، شبکه سی.ان.ان. طراحی شده از دقت قابل قبولی برخوردار بود، بنابراین ویژگی‌های مربوط به متن و داده‌های متا نیز از این شبکه استخراج شد. در ادامه برخی از مقادیر تست شده در رابطه با تعداد ویژگی‌های استخراج شده از لایه‌های داده‌های متنی و تصویری که بر دقت مدل تأثیر دارند، مشخص شد حالت بهینه کدام است. همان‌طور که مشاهده می‌شود، درصورت درنظر گرفتن ۶۴ ویژگی از تصویر و ۱۲۸ ویژگی از متن، دقت مدل ۷۸ درصد خواهد بود. یا اگر ۶۴ ویژگی از تصویر استخراج و ۳۲ ویژگی از متن استخراج شود، دقت مدل ۷۹ درصد به دست می‌آید. درنهایت مشخص شد با درنظر گرفتن ۱۲۸ ویژگی از تصویر و ۲۵۶ ویژگی از داده‌های متنی و متا، مدل به دقت ۸۱ درصدی می‌رسد که بهترین دقت ممکن است.

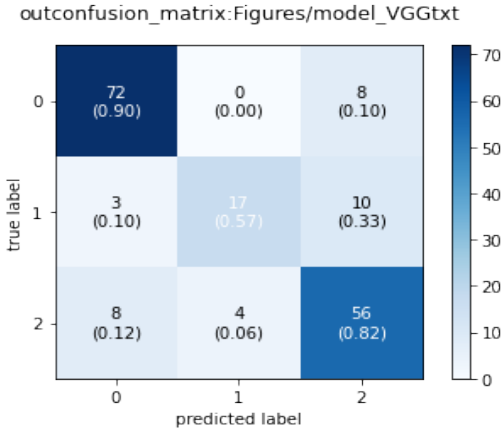
جدول ۵. مقایسه‌ای از نتایج دقت مدل در تعداد ویژگی‌های درنظر گرفته شده در داده‌های مختلف

نوع داده	تعداد ویژگی استخراج شده			
تعداد ویژگی استخراج شده از تصویر	۱۲۸ ویژگی	۶۴ ویژگی	۶۴ ویژگی	۵۱۲ ویژگی
تعداد ویژگی استخراج شده از متا و متن	۲۵۶ ویژگی	۱۲۸ ویژگی	۳۲ ویژگی	۱۰۲۴ ویژگی
دقت مدل به درصد	**٪۸۱**	**٪۷۸**	**٪۷۹**	**٪۷۷**



شکل ۲. معماری شبکه ترکیبی برای آموزش ماشین برای داده‌های چندوجهی

بررسی ماتریس درهم‌آمیختگی این مدل ترکیبی نشان داد (شکل ۳) که دقت این مدل ترکیبی در شناسایی داده‌ها با قطبیت منفی، ۹۰ درصد و در رابطه با داده‌هایی که قطبیت مثبت دارند ۸۲٪ است. هرچند دقت مدل در تعیین داده‌های طبقه خنثی بهتر شده و حدود ۶۰ است، اما همچنان این عدد احتمالاً به‌دلیل حجم کم داده‌های این طبقه بوده است.



شکل ۳. ماتریس درهم‌آمیختگی در ترکیب وی.جی.جی.۱۹ و سی.ان.ان. متن در داده‌های چندوجهی

نتیجه‌گیری

هدف پژوهش حاضر، آموزش ماشین برای پیش‌بینی احساسات کاربران در شهر با استفاده از داده‌های فضای مجازی بوده است. این پژوهش از معدود پژوهش‌های انجام شده درحوزه تحلیل احساسات شهری مبتنی‌بر یادگیری ماشین در ایران است و بنابراین با مرور مبانی نظری و بررسی تجارب جهانی در این حوزه، روش یادگیری ماشین و بهره‌گیری از داده‌های فضای مجازی برای سنجش احساسات در شهر در کلان‌شهرهای ایران مدّنظر قرار گرفته است. همچنین علی‌رغم آنکه پژوهش‌های زیادی در این حوزه در تجارب جهانی انجام شده، اما درحوزه داده‌های چندوجهی که تحلیل احساسات مدّنظر قرار می‌دهند، پژوهش‌های کمتری انجام شده است. براساس یافته‌ها، ماشین آموزش دیده در این پژوهش با الگوریتم هیبریدی استفاده شده برای داده‌های چندوجهی مربوط به کلان‌شهرهای ایران قابلیت تعمیم‌پذیری مطلوبی دارد و می‌توان با ارائه داده‌های جدید مربوط به دیگر شهرها و دیگر فضاها، مثبت با منفی بودن نگرش افراد به فضاها را استخراج نمود و با استناد به آن‌ها، ضمن شناسایی فضاهای نامناسب، با برنامه‌ریزی و مدیریت مناسب شهری آن‌ها را بهبود بخشید. همچنین ازآنجاکه روش مزبور برخلاف روش‌های سنتی برای ممیزی، مشمول زمان و هزینه بالا نیست، می‌تواند با سرعت و دقت قابل قبول از داده‌های نشر داده شده در فضاهای اجتماعی مجازی بهره برد. همان‌گونه که پیش‌از این نیز بیان شد، این روش، از روش‌های مبتنی‌بر فرهنگ لغت که در رابطه با متون استفاده می‌شود، نوین‌تر بوده و ازاین‌رو که داده‌های چندوجهی را مدّنظر داشته است، امکان بهره‌مندی یا مقایسه نتایج این روش با روش غیرمدل‌گرا وجود نداشته است.

فارغ از موارد بیان شده، این پژوهش ازحیث منابع مالی، دسترسی به پایگاه داده و دسترسی کاربران به شبکه مجازی منتخب دارای محدودیت‌هایی بوده است. یکی‌از محدودیت‌های این پژوهش مرتبط با محدودیت‌هایی است که توئیتر در دریافت و دانلود اطلاعات با استفاده از توئیتر استریمینگ ایجاد می‌کند. به‌این‌ترتیب که به‌صورت رندم امکان دسترسی به درصدی از داده‌ها را به پژوهشگر به‌صورت رایگان می‌دهد. همچنین از دیگر محدودیت‌های این پژوهش، استفاده از توئیتر به‌عنوان شبکه اجتماعی منتخب است که به‌علت فیلترینگ در

کشور، عموم کاربران فضای مجازی از آن استفاده نمی‌کنند. در این میان امکان استفاده از سایر شبکه‌های اجتماعی با محدودیت کمتر در زمان انجام پژوهش (چون اینستاگرام) وجود داشت، اما به‌دلیل عدم دسترسی رایگان به داده‌ها و یا برای مثال تأکید بیشتر اپلیکیشن و کاربران بر تصویر تا متن، از انتخاب آن حذر شد. درعین‌حال حجم کم داده‌های چندوجهی قابل دسترس و مرتبط با شهر از دیگر محدودیت‌های پژوهش بوده است. علی‌رغم آنکه پژوهشگر بازه ده ساله و ۸ کلان شهر ایران را مورد بررسی قرار داده است، اما در فرایند بررسی و حذف داده‌های غیرمرتبط، مجموعه داده‌های قابل اتکا، چندان زیاد نبوده است. همچنین نداشتن منابع مالی در رابطه با پژوهش‌هایی که نیازمند حجم داده‌های زیاد جهت آموزش مناسب ماشین یا استفاده از یادگیری عمیق هستند، از دیگر محدودیت‌های پیشِ‌روی این پژوهش بوده است. از دیگر محدودیت‌های این پژوهش، نبود داده‌های کافیِ مکان‌محور بوده که امکان ارائه نقشه‌های تحلیل احساسی را به‌صورت نقطه‌ای یا محدوده‌ای ارائه دهد. چنانچه چنین داده‌هایی در دسترس می‌بود، می‌شد ضمن ارائه نقشه فضاهایی با پاسخ عاطفی مثبت و منفی، میزان احساسات در هر فضا را بررسی نمود. درعین‌حال از دیگر محدودیت‌های پژوهش، مشکلات توسعه‌ای کتابخانه‌های تحلیل احساسات در رابطه با زبان فارسی است که به‌دلیل محدودیت حجم کلمات این کتابخانه‌ها، امکان تحلیل احساسات با مدل اکمن (۱۹۹۹) فراهم نبود. بااین‌وجود علی‌رغم تمامی مسائل موجود، این پژوهش تلاش کرد اهمیت بهره‌مندی از روش منتخب را به‌عنوان روشی نوین در تحلیل احساسات معرفی کند و برخی‌از قابلیت‌های آن را ارائه دهد.

پی‌نوشت‌ها

1. VADER (Valence Aware Dictionary for Sentiment Reasoning) is a model used for text sentiment analysis that is sensitive to both polarity (positive/negative) and intensity (strength) of emotion

در این پژوهش برای تحلیل احساسات، از مدل ویدر استفاده شده که برای داده‌های رسانه‌ای اجتماعی توسعه داده شده است (Hutto & Gibert, 2014). زیرا هم واجد دقت بوده و برای متون کوتاه، در مقایسه با سایر مدل‌های تحلیل احساسات مناسب‌تر است. تحلیل متن، در سطح ابتدایی‌تر، فراوانی کلمات کلیدی منتخب و رابطه‌ی بین آن‌ها را بررسی می‌کند.

۲. Picodash (یک موتور جست‌وجوی اینستاگرام پولی و مبتنی‌بر وب به‌منظور خروجی گرفتن داده‌ها)

۳. Tone Analyzers (تجزیه‌وتحلیل زبانی برای تشخیص لحن احساسی و زبانی در متن نوشتاری)

4. GeoJSON

5. Volunteered geographic information

6. SUN

سان بزرگ‌ترین مجموعه داده شامل ۸۹۹ دسته و ۱۳۰۵۱۹ تصویر است که برای درک صحنه کاربرد دارد. حداقل ۱۰۰ تصویر در هر دسته موجود است (tar.gz.SUN۳۹۷).

۷. YOLO (سیستم تشخیص اشیا است که پیشرفته بوده و مبتنی‌بر زمان واقعی است)

8. DeepSent

9. OutdoorSent

10. AFINN

11. Ps_match_multi-SAS macro

12. NLP

13. ML

14. IR

15. DM

16. Lexical method

17. Accuracy

18. Recall
19. Learning Rate
20. Precision
21. ReLu
22. Pooling
23. Softmax
24. Cross validation
25. Unseen data
26. Hold out
27. K fold
28. Leave-one-out
29. Confusion Matrix
30. Integer
31. Float
32. String
33. Object

در جاوا اسکریپت، اشیاء یا Objects نوعی دیتاتایپ هستند که به برنامه‌نویس اجازه می‌دهند تا گروهی از اطلاعات مشابه را در یکجا نگه دارد.

34. Most frequency
35. Label Encoder
36. MinMaxScaler
37. StandardScaler
38. Normal Sampling
39. Weighted sampling
40. Clean
41. Hazm
42. Zero Padding
43. Embedding
44. Fasttext
45. Embedding
46. ResNet و Pre-Activation ResNet

فهرست منابع

- راد، رؤیا، موسوی، محمد، و وردی، فاطمه (۱۳۹۹). استفاده از یادگیری عمیق در تشخیص خودکار بیماری گیاهان براساس پردازش تصویر برگ. تحقیقات سامانه‌ها و مکانیزاسیون کشاورزی، ۲۱(۷۶)، ۴۹-۶۸.
- راستگو، راضیه، و کیانی، کوروش (۱۳۹۸). شناسایی چهره با استفاده از تنظیم دقیق شبکه‌های کانولوشنی عمیق و رویکرد یادگیری انتقالی. مجله مدل‌سازی در مهندسی، ۱۷ (۵۸)، ۱۰۳-۱۱۱.
- زاهدی حقیقی، سیده سعیده، سخایی، سید محمود، و دلیری، محمدرضا (۱۳۹۸). تشخیص حالت‌های احساسی مبتنی‌بر EEG با استفاده از شبکه یادگیری عمیق. مجله مهندسی پزشکی زیستی، ۱۳ (۲)، ۹۵-۱۰۴.
- فولادی، صابر، فرسی، حسن، و محمدزاده، سجاد (۱۳۹۹). تشخیص و طبقه‌بندی سرطان پوست با استفاده از یادگیری عمیق. مجله علمی دانشگاه علوم پزشکی بیرجند، ۲۶ (۱)، ۴۴-۵۳.
- موسوی، سیدمهدی، عبادی، حمید، و کیانی، عباس (۱۳۹۸). ارائه روشی بهینه مبتنی‌بر یادگیری عمیق به‌منظور طبقه‌بندی طیفی مکانی تصاویر با قدرت تفکیک مکانی بالا در مناطق نیمه‌شهری. نشریه علوم و فنون نقشه‌برداری، ۹ (۲)، ۱۵۱-۱۷۰.
- وب‌سایت خانه بیگ دیتای ایران (۱۳۹۷). پردازش متن با *Jhazm* نسخه جاوا کتابخانه هضم برای پردازش زبان فارسی.

بازیابی شده در ۲۸ بهمن ۱۴۰۱، از https://bigdata-ir.com

- یوسفی متقاعد، مریم، و صبوحی، هادی (۱۳۹۸). مروری بر تحلیل احساسات شبکه‌های اجتماعی درحوزه قطبیت. کنفرانس بین‌المللی وب‌پژوهی، ۳ و ۵ اردیبهشت ماه، تهران، ایران.

- Baly, R., Hobeica, R., Hajj, H., El-Hajj, W., Shaban, K.B., & Al-Sallab, A. (2016). A Meta-Framework for Modeling the Human Reading Process in Sentiment Analysis. *ACM Transactions on Information Systems* (*TOIS*), 35(1), 1-21. https://doi.org/10.1145/2950050.

- Benevenuto, F., Araújo, M., & Ribeiro, F. (2015). Sentiment Analysis Methods for Social Media. *In Proceedings of Brazilian Symposium on Multimedia and the Web* (*WebMedia*). ACM, Manaus, Brazil, 11-11. https://doi.org/10.1145/2820426.2820642.

- Bonasoli de Oliveira, W., Dorini, L.B., Minetto, R., & Silva, T.H. (2020). OutdooeSent: Setiment analysis of urban outdoor images by using semantic and deep features. *ACM transaction on information systems*, 1(1), 1-29.

- Campos, V., Jou, B., & Giró i Nieto, X. (2017). From pixels to sentiment: Fine-tuning CNNs for visual sentiment prediction. *Image and Vision Computing*, 65, 15 -22. https://doi.org/10.1016/j.imavis.2017.01.011.

- Chapman, L., Resch, B., Sadler, J., Zimmer, S., Roberts, H., & Petutsching, A. (2018). Investigating the emotional responses of individuals to urban green space using Twitter data: a critical comparison of tree different methods of sentiment analysis. *Urban Planning*, 3 (1), 21-33.

- Chen, F., Gao, Y., Cao, D., & Ji, R. (2015). Multimodal hypergraph learning for microblog sentiment prediction. *In IEEE International Conference on Multimedia and Expo* (*ICME*). IEEE, Turin, Italy, 1-6. https://doi.org/10.1109/ICME.2015.7177477.

- Chen, NC., Nagakura, T., & Larson, K. (2016). Social media as Complementary Tool to Evaluate Cities- Data Mining Innovation Districts in Boston, Herneoja, Aulikki; Toni Österlund & Piia Markkanen (eds.), *Complexity & Simplicity - Proceedings of the 34th eCAADe Conference* - Volume 2, University of Oulu, Oulu, Finland, 22-26 August 2016, 447-456.

- Chen, T. (2014). *Deep SentiBank: Visual Sentiment Concept Classification with Deep Convolutional Neural Networks*. CoRR abs/1410.8586, 1-7. arXiv:1410.8586. http://arxiv.org/abs/1410.8586.

- Chen, T., Xu, R., He, Y., & Wang, X. (2017). Improving sentiment analysis via sentence type classification using BiLSTM-CRF and CNN. *Expert Systems with Applications*, 72, 221-230.

- Dhall, A., Goecke, R., & Gedeon, T. (2015). Automatic Group Happiness Intensity Analysis. *IEEE Transactions on Affective Computing*, 6(1), 13-26. https://doi.org/10.1109/TAFFC.2015.2397456.

- Eboli, L., & Mazzulla, G. (2011). A methodology for evaluating transit service quality based on subjective and objective measures from the passenger's point of view. *Transport Policy*, 18(1), 172-181. http://dx.doi.org/10.1016/j.tranpol.2010.07.007.

- EURISY. (2017). *Good City Life: crowdsourcing satellite data and emotions to map our urban landscape*. Retrieved 2022, Feb. 18, from https://goodcitylife.org/

- Fathullah, A., & S.Willis, K. (2018). Engaging the Senses: The Potential of Emotional Data for Participation in Urban Planning. *Urban Science*, 2(4), 1-21. doi:10.3390/urbansci2040098.

- He, W., Zha, S., & Li, L. (2013). Social media competitive analysis and text mining: A case study in the pizza industry. *International Journal of Information Management, 33*, 464-472.

- Hollander, J. B., & Renski, H. (2015). Measuring *urban attitudes using Twitter: an exploratory study*. Lincoln Institute of Land Policy.

- Huang, H., Gartner, G., Klettner, S., & Schmidt, M. (2014). Considering affective responses

towards environments for enhancing location based services. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, 93-96. doi:10.5194/isprsarchives-XL-4-93-2014.

- Hutto, C.J., & Gilbert, E. (2014). VADER: A parsimonious rule-based model foe sentiment analysis of social media text. *Proceedings of the International AAAI Conference on Web and Social Media*, 8(1), 216-225. https://doi.org/10.1609/icwsm.v8i1.14550.
- Ji, R., Cao, D., & Lin, D. (2015). Cross-Modality Sentiment Analysis for Social Multimedia. *IEEE International Conference on Multimedia Big Data*.
- Kenyon, S., & Lyons, G. (2003). The value of integrated multimodal traveler information and its potential contribution to modal change. *Transport Research Part F: Traffic Psychology Behavior*, 6(1), 1-21. http://dx.doi.org/10.1016/s1369-8478(02)00035-9.
- Kim, E., Rosenwasser, D., & Castillo Lopez, J.L.G. (2020). Urban emotion: The interrogation of social media and its implication within urban context. *Proceedings of the 38th eCAADe Conference* - Volume 2, TU Berlin, Berlin, Germany, 16-18 September, 475-482.
- Liu, B., & Zhang, L. (2012). A survey of opinion mining and sentiment analysis. *Mining Text Data*, 415-463.
- Liu, B. (2012). Sentiment Analysis and Opinion Mining. *Synthesis lectures on human language technologies*. 1-167.
- Liu, L., Silva, E.A., Wu, C., & Wang, H. (2017). A machine learning-based method for the large-scale evaluation of qualities of the urban environment. *Computers, environment and urban systems,* 65, 113-125.
- Mazzulla, G., & Forciniti, C. (2012). Spatial association techniques for analyzing trip distribution in an urban area. *European Transport Research Review*, 4(4), 217-233.http://dx.doi.org/10.1007/s12544-012-0082-9.
- Oteros-Rozas, E., Martin-Lpez, B., Fagerholm, N., Bieling, C., & Plieninger, T. (2018). *Using social media photos to explore the relation between cultural ecosystem services and landscape features across five European sites*. Ecol. Indic. 94, 74-86.
- Parrett, G. (2016). *3.5 million photos shared every minute in 2016*. Deloitte. https://goo.gl/uwF81P.
- Poria, S., Peng, H., Hussain, A., Howard, N., & Cambria, E. (2017). Ensemble application of convolutional neural networks and multiple kernel learning for multimodal sentiment analysis. *Neurocomputing*, 261, 217-230.
- Resch, B., Summa, A., Sagl, G., Zeile, P., & Exner, J. (2014). Urban emotions-Geo-semantic emotion extraction from technical sensors, human sensors and crowd sourced data. In G. Gartner & H. Huang (Eds.), *Progress in location- based service, Springer*.
- Ribeiro, F.N., Araujo, M., Goncalves, P., Goncalves, M.A., & Benevenuto, F. (2016). SentiBench -A benchmark comparison of state-of-the-practice sentiment analysis methods. *EPJ Data Science*, 5(1), 1-29. https://doi.org/10.1140/epjds/s13688-016-0085-1.
- Rossi, L., Boscaro, E., & Torsello, A. (2018). Venice through the Lens of Instagram: A Visual Narrative of Tourism in Venice, *18 Companion of the Web Conference*, 11901197-.
- Schwartz, A.J., Dodds, P.S., O'Neil-Dunne, J.P.M., Danforth, C.M., & Ricketts, T. (2019). Visitors to urban greenspace have higher sentiment and lower negativity on Twitter. *People and Nature*, 476-485.
- Sdoukopoulos, A., Nikolaidou, A., Pitsiava-Latinopoulou, M., & Papaioannou, P. (2018). Use of social media for assessing sustainable urban mobility indicators. *International Journal of sustainable*

development and planning, 13(2), 338-348.

- Setiawan, E.I., Juwiantho, H., Santoso, J., Sumpeno, S., Fujisawa, K., & Purnomo, M.H. (2021). Multiview sentiment analysis with Image-Text-Concept features of Indonesian social media posts, *International Journal of Intelligent Engineering & Systems*, 14(2), 521-535.
- Silva, T.H., Vaz de Melo, P.O.S., Almedia, J.M., Salles, J., & Loureiro, A.A.F. (2013). *A comparison of Foursquare and Instagram to the study of city dynamics and urban social behavior, Proc*. ACM SIGKDD Int. Workshop on Urban Computing (UrbComp'13).
- Sim, J., Miller, P., & Swarup, S. (2020). Tweeting the high line life: a social media lens on urban green spaces. *Sustainability*, 12, 1-18.
- Soleymani, M., Garcia, D., Jou, B., Schuller, B., Chang, S., & Pantic, M. (2017). A survey of multimodal sentiment analysis. *Image and Vision Computing*, 65, 3-14.
- Tebyanian, N. (2020). Application of machine learning for urban landscape design: a primer for landscape architect. *Journal of digital landscape architecture*, 5, 217-226.
- Tieskens, K.F., Zantenm B, T.V., Schulp, C.J.E., & Verburg, P.H. (2018). Aesthetic appreciation of the cultural landscape through social media: An analysis of revealed preference in the Dutch river landscape. *Landscape and Urban Planning*, 177, 128-137.
- Tu, W., Cao, J., Yue, Y., Shaw, S., Zhou, M., Wang, Z., Chang, X., Xu, Y., & Li, Q. (2017). *Coupling mobile phone and social media data: A new approach to understanding urban functions and diurnal patterns*. *International Journal of Geographical Information Science*, 31, 2331-2358.
- Wu, J., Lin, Z., & Zha, H. (2016). Multi-view common space learning for emotion recognition in the wild. *In ACM International Conference on Multimodal Interaction*. ACM, Tokyo, Japan, 464-471. https://doi.org/10.1145/2993148.2997631.
- You, L., & Tuncer, B. (2016). Exploring public sentiments for livable places based on a crowd-calibrated sentiment analysis mechanism. *IEEE| ACM international conference on advances in social networks analysis and mining* (*ASONAM*). August 18-21, USA: San Francisco.
- Zhan, X., Wang, Y., Rao, Y., & Li, Q. (2019). Learning from Multi-annotator Data: A Noise-aware Classification Framework. *ACM Transactions on Information Systems* (*TOIS*), 37(2), 1-28. https://doi.org/10.1145/3309543.
- Zhou, G-Y & Huang, J. X. (2017). Modeling and Mining Domain Shared Knowledge for Sentiment Analysis. *ACM Transactions on Information Systems* (*TOIS*), 36(2), 1-36. https://doi.org/10.1145/3091995.

COPYRIGHTS

Copyright for this article is retained by the author(s), with publication rights granted to Journal of Architecture and Urban Planning. This is an open-access article distributed under the terms and conditions of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/).

نحوه ارجاع به این مقاله

محمدی، مریم (۱۴۰۳). ارائه مدل پیش‌بینی‌کننده تحلیل احساسات داده‌های چندوجهیِ شهری با بهره‌گیری از یادگیری عمیق. بررسی موردی: کلان‌شهرهای رسمی ایران. *نشریه علمی نامه معماری و شهرسازی*، ۱۷(۴۵)، ۷۹-۱۰۳.

DOI: 10.30480/aup.2024.5338.2151
URL: https://aup.journal.art.ac.ir/article_1371.html

Vol. 17 • No. 45 • Winter 2025

Received: 28 May, 2024 | Accepted: 24 September, 2024 | Published: 21 December, 2024

Journal of Architecture and Urban Planning, 17(45), 79-103
DOI: 10.30480/aup.2024.5338.2151
Document Type: Research Paper

Presenting a predictive model for sentiment analysis of urban multi-modal data with deep learning
Case study: Iranian formal megacities

Maryam Mohammadi
Associate Professor, Department of Urban Design, Faculty of Architecture and Urban Planning, Iran University of Art, Tehran, Iran (Corresponding Author)

Abstract

Urban planning and design issues are highly complex, making it challenging to achieve a comprehensive understanding of these issues, especially on a large scale, using traditional urban planning methods alone. Therefore, the integration of data science and modeling presents a promising approach for comprehending urban issues. Extensive research and reports on real urban projects implemented across various countries suggest that machine learning and deep learning techniques have the potential to effectively address urban problems and offer additional benefits such as reducing economic costs associated with inefficiencies, improving operational speed, and enhancing predictability. The integration of machine learning and deep learning can be utilized in the evaluation, design, and post-implementation stages of urban planning and design processes. User experience, including environmental stress, happiness, and fear, significantly influences the character and activities within urban spaces. Hence, it is crucial for urban designers to create new spaces or revitalize existing ones in a way that evokes a positive emotional response and a sense of belonging among users. Numerous studies have explored this subject, primarily focusing on evaluating people's emotions within physical environments and occasionally within virtual settings, using AI-based methodologies. Sentiment analysis or opinion mining, previously employed before the rise of social media, is now being investigated and analyzed using computational methods with the aid of vast amounts of data published on social media platforms. This research domain, widely used in economic, marketing, political, and social research, has increasingly permeated the realm of urban studies. However, it is relatively new within the field of urban studies, with only a limited number of research studies conducted thus far. The main objective of this research is to develop a predictive model for understanding users' emotional responses towards urban spaces through an analysis of multi-modal data, including images and textual content, published in virtual spaces. This study incorporates two notable innovations:

- The use of a machine learning approach in sentiment analysis of multi-modal data on an urban scale, which has been explored in data science research but remains insufficiently explored in the realm of urban planning and management in Iran.
- The creation of a dataset comprising textual and visual data relevant to urban environments.

It aims to present a predictive model that effectively analyzes sentiment in textual and visual media data derived from Twitter. The research employs a quantitative methodology and utilizes a supervised deep learning technique. The case study encompasses multi-modal urban data from eight megacities in Iran, collected over a ten-year period until December 2021. An algorithmic hybridization based on the Convolutional Neural Network (CNN) architecture has been developed and employed to train the textual data, resulting in an accuracy rate of 77%. Additionally, the visual data has been trained using three distinct algorithms, with VGG19 demonstrating superior performance and achieving an accuracy rate of 88%. A hybrid algorithm has been devised to train the multi-modal data, yielding an accuracy of 66%. Subsequently, a combination model leveraging pre-trained VGG19 models alongside the developed text combination model has been implemented, achieving an elevated accuracy of 81%. These outcomes demonstrate the machine's capacity to predict sentiment polarity in textual-visual data related to cities.

Keywords: Predictive model, urban sentiment analysis, multi-modal data, deep learning, megacity, Iran