

Users' Emotions analysis in urban parks based on data in virtual space; using Model-oriented and Non-model-oriented methods: the case study of Mellat park in Tehran

Maryam Mohammadi ^{1✉}, Fatemeh Ghodousi ²

1. (Corresponding author) *Department of Urban Design, Faculty of Architecture and Urban Planning, Iran University of Art, Tehran, Iran*

Email: m.mohammadi@art.ac.ir

2. *Department of Urban Design, Faculty of Architecture and Urban Planning, Iran University of Art, Tehran, Iran*

Email: fateme.ghodussi@gmail.com

Article Info

Article type:
Research Article

Article History:

Received:

26 November 2024

Received in revised form:

20 February 2025

Accepted:

1 April 2025

Available online:

6 May 2025

Keywords:

Sentiment Analysis,
Park,
Social Network,
Machine Learning,
Lexical Method,
Blending Learning.

ABSTRACT

Among the types of urban spaces, urban green spaces and parks, as city breathing spaces, are lush and relaxing areas that have been selected as the case of this research paper. Therefore, this article aims to analyze the emotions of users of Mellat Park in Tehran in the form of analytical research based on a quantitative method (supervised machine learning approach and lexical-based. After preprocessing and labeling, the data were examined and analyzed using two methods, such as model-oriented and non-model-oriented. Emotions were also examined and analyzed using the Python programming language. The comparison of these two methods revealed that among the machine learning algorithms, XGBoost has the highest accuracy at 87%, while K-nearest neighbors and support vector machines have lower accuracy but are still capable of predicting emotions in green spaces. The lexical method (using the VADER dictionary) has a lower predictive ability compared to machine learning. Finally, the stacking ensemble learning model, which was used to increase the accuracy of the model, has the ability to predict emotions based on the results of the confusion matrix (96%). Therefore, using the method based on virtual space data, it is possible to predict the emotions of users of other urban green spaces with high speed and accuracy in Tehran.

Cite this article: Mohammadi, M., & Ghodousi, F. (2025). Users' Emotions analysis in urban parks based on data in virtual space; using Model-oriented and Non-model-oriented methods: the case study of Mellat park in Tehran. *Geographical Urban Planning Research Quarterly*, 13 (1), 39-57.
<http://doi.org/10.22059/jurbangeo.2025.382327.1988>



Extended Abstract

Introduction

Increasing the quality of urban parks is effective in providing the psychological and emotional comfort of users. Sentiment analysis is an approach to measure the level of psychological comfort and emotional response of users to spaces.

Methodology

The research method in the case study is based on the quantitative method. The following presents the selected social virtual space, the time period of data extraction, and the methods used in both methods for data preparation and analysis. It was found that the Google Map social network is suitable for this study. After the pre-processing, in order to train and test the data, the data is divided into 80% training data and 20% testing data (for evaluating the models). Then, the data is labeled to train machine learning models using the points given by the users to the location. In order to select and allocate the appropriate model or algorithm, according to the subject under investigation and the need to classify emotions and achieve a predictive model, the models under the supervision of machine learning were used. The VADER method was used in the non-model method (classification of feelings by the dictionary method). At the same time, the Blending Ensemble model from the stacking family was also used in this research.

Results and discussion

According to the first method, among the implemented models, the XG Boost algorithm correctly recognized 87% of all messages and was the best algorithm. After that, the k nearest neighbor algorithm correctly recognized about 80% of all messages, support vector machine, 70%, and simple Bayes and linear discriminate analysis about 66% and 57% of all messages. Other models performed less than 50%, and the random forest model performed worse than all algorithms. The results of the second method showed that 22% of the data in the category of negative data were correctly identified, 33% of the neutral data were false, and 44% were identified as false positives. In the category

of neutral data, 7% of false negative data, 20% of true neutral data, and 73% of false positive data have been detected. In positive data, 3% of false negative data, 15% of false neutral data, and 82% of positive data are correctly identified. As a result, in this method, the highest correct percentage associated with positive data is 82% correct detection. As a result, it had the best performance in detecting positive data, but it did not perform well in detecting negative and neutral data. In order to provide a predictive model with higher accuracy, the Blending Ensemble models and the stacking family, which is included in the Ensemble Learning model category, were used. The result of applying the proposed blending model to the test data in the confusion matrix shows that in the category of negative data, 80% of the data are correct, 10% of the data are neutral and false, and 10% have identified false positives from the data. In the category of neutral data, 0% of false negative data, 84% of true neutral data, and 16% of neutral data were detected as false positives. In the positive data section, 0% of false negative data, 3% of false neutral data, and 97% of true positive data have been identified. As a result, in general, in the proposed blending model, the best performance has been in positive data, which has been correctly recognized in 97%, and the most error in negative data has been in general with 20% false neutrals and false positives, which compared to other algorithms have a small percentage of errors and have performed well in all 3 categories of positive, neutral and negative. In relation to the evaluation criteria of this model, a percentage above 95% can be seen in all the evaluation criteria, which indicates the good performance of the model.

Conclusion

The results of the research showed that the model-oriented methods worked better than the word-based method, and the blending method was better than the machine learning algorithms. Therefore, the algorithm trained with the blending method has the ability to predict urban feelings in the park with a high probability. The chosen method has many applications in the

field of urban planning and urban design because it provides the possibility of identifying citizens' feelings about the environment due to its low cost compared to field research and the speed of data analysis.

Funding

There is no funding support.

Authors' Contribution

Authors contributed equally to the conceptualization and writing of the article. All of the authors approved the content of the manuscript and agreed on all aspects of the work declaration of competing interest none.

Conflict of Interest

Authors declared no conflict of interest.

Acknowledgments

We are grateful to all the scientific consultants of this paper.

تحلیل احساسات کاربران پارک‌های شهری مبتنی بر داده‌های فضای مجازی؛ با بهره‌گیری از روش‌های مدل‌گرا و غیرمدل‌گرا مطالعه موردی: پارک ملت تهران

مریم محمدی ✉، فاطمه قدوسی

۱- نویسنده مسئول، گروه طراحی شهری، دانشکده معماری و شهرسازی، دانشگاه هنر ایران، تهران. رایانامه: m.mohammadi@art.ac.ir

۲- گروه طراحی شهری، دانشکده معماری و شهرسازی، دانشگاه هنر ایران، تهران. رایانامه: fateme.ghodussi@gmail.com

چکیده

اطلاعات مقاله

نوع مقاله:

مقاله پژوهشی

تاریخ دریافت:

۱۴۰۳/۰۹/۰۶

تاریخ بازنگری:

۱۴۰۳/۱۲/۰۲

تاریخ پذیرش:

۱۴۰۴/۰۱/۱۲

تاریخ چاپ:

۱۴۰۴/۰۲/۱۶

واژگان کلیدی:

تحلیل احساسات

شهری،

پارک،

شبکه اجتماعی،

یادگیری ماشین،

روش لغوی،

یادگیری گروهی.

از میان انواع فضاهای شهری، فضاهای سبز شهری و پارک‌ها، به‌عنوان تنفس گاه‌های شهر، عرصه‌هایی سرسبز و آرامش‌بخش هستند و، به‌عنوان بستر فضایی پژوهش انتخاب شده‌اند. بنابراین این مقاله، با هدف تحلیل احساسات کاربران از پارک ملت تهران، در قالب پژوهشی تحلیلی و مبتنی بر روش کمی (رویکرد یادگیری ماشین نظارت‌شده و مبتنی بر لغت) است. داده‌ها پس از پیش‌پردازش و برچسب‌گذاری، با دو روش: مدل‌گرا و غیرمدل‌گرا، بررسی و تحلیل احساسات با زبان برنامه‌نویسی پایتون، انجام شده است. مقایسه این دو روش نشان داد که از میان الگوریتم‌های یادگیری ماشین، ایکس.جی.بوست با بیشترین دقت (۸۷٪)، کی-نزدیکترین همسایه و ماشین بردار پشتیبان با دقت کمتر، قابلیت پیش‌بینی احساسات در فضاهای سبز را دارند. روش لغوی (استفاده از فرهنگ لغت ویدر) در مقایسه با یادگیری ماشین، قابلیت پیش‌بینی کمتری دارد. در نهایت مدل یادگیری گروهی از نوع پشته‌سازی که برای بالا بردن دقت مدل استفاده شده که بر اساس نتایج ماتریس درهم‌آمیختگی (۹۶٪) قابلیت پیش‌بینی احساسات را دارد. بنابراین با بهره‌گیری از روش مبتنی بر داده‌های فضای مجازی، می‌توان، به پیش‌بینی احساسات کاربران سایر فضاهای سبز شهری با سرعت و دقت بالا در شهر تهران دست یافت.

استناد: محمدی، مریم و قدوسی، فاطمه. (۱۴۰۳). تحلیل احساسات کاربران پارک‌های شهری مبتنی بر داده‌های فضای مجازی؛ با بهره‌گیری از روش‌های مدل‌گرا و غیرمدل‌گرا مطالعه موردی: پارک ملت تهران. *پژوهش‌های جغرافیای برنامه‌ریزی شهری*، ۱۳ (۱)، ۳۹-۵۷.

<http://doi.org/10.22059/jurbangeo.2025.382327.1988>

مقدمه

موضوع احساس و ادراک کاربران از فضاهای شهری، یکی از پایه‌ای‌ترین موضوعات برای طراحی، بهبود و اصلاح فضاها در شهر است (پای‌کن و پورجعفر، ۱۳۹۸؛ ربانی خوراسگانی و کیانپور، ۱۳۸۸؛ Davidson & Gu et al., 2019; Begley, 2012; Ekman et al., 1969). از آنجاکه از دید متأخرترین نظریه‌پردازان شهری، شهرسازی برای مردم و با آن‌ها است، باید از رویکردهای نخبه‌گرا که طراحی شهری تداعی شونده (راپوپورت، ۱۳۹۲) را موردتوجه قرار نمی‌دهند؛ حذر کرد. از میان تنوع فضاهای شهری در شهرها، فضاهای سبز شهری و پارک‌ها گونه‌ای از فضاها هستند که به دلیل کیفیت‌هایی چون طراوت، سبزی‌گی و تجربه مستقیم طبیعت که در دسترس کاربران قرار می‌دهند (Kovacs-Györi et al., 2018; Lim et al., 2018; Wang et al., 2021)، پناهگاهی برای شهرنشینانی هستند که با محیط‌های شهری شلوغ و پر از نوفه دست‌وپنجه نرم می‌کنند. علی‌رغم انتظاری که از این گونه از فضاها در شهر می‌رود، لازم است تا احساس کاربر در این فضاها را نیز بررسی نمود تا مشخص شود این گونه فضاها تا چه حد از حیث کیفیت‌های پیش‌گفته موفق بوده‌اند.

روش‌های متعددی برای سنجش احساسات در شهرها موردتوجه شهرسازان است. روش‌های سنتی مبتنی بر پرسش‌نامه یا خود اظهاری جزء اولین روش‌ها بوده و روش‌های متأخرتر مبتنی بر ابزارها یا گجت‌ها است (طاهرطلوع دل و همکاران، ۱۳۹۸). در این میان با ظهور تکنولوژی و وب در عصر فناوری و اطلاعات، مقوله تحلیل احساسات با استفاده از داده‌های منتشرشده در فضای تحت وب (De Oliveira & Painho, 2015; Iaconesi, & Persico, 2013; Hao & Dai, 2016؛ طیرانی نجاران و جلالی، ۱۳۹۷؛ سلطانی‌فر و همکاران، ۱۳۹۲) بیش‌ازپیش موردتوجه قرار گرفت. روش‌های پیش‌گفته در علوم مختلفی کاربرد داشته و در علوم شهری نیز ساختاردهی به داده‌های شبکه‌های اجتماعی و تحلیل آن‌ها برای تحلیل احساسات شهری به‌عنوان روشی نسبتاً نو ظهور مورد تأکید قرار گرفته است. از مزایای این روش می‌توان به انجام آن بدون نیاز به مراجعه مستقیم به کاربران فضا و بدون القاء نظرات مصاحبه‌گران یا آزمون‌گیرندگان، نام برد. درعین‌حال حجم زیاد داده‌ها و نظرات کاربران در شبکه‌های متنوع اجتماعی مجازی، داده‌های کلانی را در دسترس قرار می‌دهد که می‌تواند نوعی برون‌سپاری جمعی مطلوب باشد. با این توضیحات، مسئله پژوهش حاضر این است که آیا در فضاهای شهری تهران نیز می‌توان با بهره‌گیری از داده‌های نشر یافته در فضاهای مجازی، احساسات کاربران در فضاها را بررسی نمود.

مبانی نظری

احساسات و روش‌های تحلیل آن

شهرسازی و طراحی شهری در پی رفع نیازهای جامعه و ایجاد احساس رضایت‌بخش در افراد هستند. با ایجاد رشته روانشناسی محیطی که تعامل مردم و محیط شهری را در نظر می‌گیرد، مطالعات گسترده‌ای آغاز شده‌آغاز شده و امروزه احساسات شهری به حوزه‌ای میان‌رشته‌ای تبدیل شده است (پای‌کن و پورجعفر، ۱۳۹۸). احساسات معانی متفاوتی دارند، واژه ایموشن^۲ در زبان انگلیسی به معنای هیجان موردنظر روانشناسان است که از ریشه لاتین کلمه *ام‌اور*^۴ به معنای حرکت، تحریک، حالت تنش یا تهییج مشتق شده است. واژه دیگر مرتبط به احساسات، کلمه *سنتیمنت*^۶ است که طبق

2. Emotion

4. Emovere

6. Sentiment

تعریف، حالت عاطفی و آمیزه عناصر هیجانی و تخیلی کم‌وبیش پایداری است که در غیاب محرک دوام می‌آورد (ربانی خوراسگانی و کیانپور، ۱۳۸۸: ۳۷). از دید دسمیت^۱ احساس پدیده‌ای است که شامل چهار مؤلفه: ۱) واکنش‌های رفتاری همچون انتخاب شیوه در برخورد با اشخاص یا مسائل؛ ۲) واکنش‌های بروز حالت (برای مثال تغییرات در چهره و صدای شخص)؛ ۳) واکنش‌های فیزیولوژیک (برای مثال تغییرات نرخ ضربان قلب یا نرخ تنفس هنگام ترس)؛ و ۴) تجربه ذهنی و درونی (به‌عنوان نمونه احساس سردرگمی یا ناامیدی) است (محمدرضوی و همکاران، ۱۳۹۲: ۱۰۰). به‌صورت کلی رایج‌ترین انواع احساسات شناخته‌شده شامل خشم، اضطراب، شادی، غم، خنثی، ترس، انزجار، تعصب و خستگی و علاقه هستند (Gu et al., 2019; Davidson & Begley, 2012; Ekman et al., 1969). روش‌های مختلفی برای تحلیل احساسات وجود دارد چون: روش نقشه‌برداری احساسات، دستگاه‌های فیزیولوژیکی، استفاده از واقعیت مجازی (طاهرطلوع دل و همکاران، ۱۳۹۸: ۱۰۸) و استفاده از شبکه‌های اجتماعی (De Oliveira & Painho, 2015; Iaconesi, & Persico, 2013). این پژوهش از داده‌های شبکه‌های اجتماعی بهره برده است. همان‌طور که در پژوهش‌ها بیان شد، امروزه داده‌های بدون ساختار در تحلیل احساسات رایج هستند، چراکه کاربران معمولاً نظرات خود را از طریق کلمات به‌جای اعداد بیان می‌کنند (Hao & Dai, 2016). فناوری‌های تلفن همراه، ابزاری ارزشمند برای جمع‌آوری داده‌های احساسات شهری هستند، زیرا آن‌ها می‌توانند هم‌زمان، مکان فضایی و پست‌های کاربر را که باید حاوی محتوای احساسی یا روحی باشند، جمع‌آوری کنند (De Oliveira & Painho, 2015). با استفاده از الگوریتم‌های کامپیوتری به‌منظور استخراج احساسات از داده‌های موجود در وب، می‌توان به تحلیل احساسات در شبکه‌های اجتماعی دست یافت (طیرانی نجاران و جلالی، ۱۳۹۷: ۶۳).

داده، کلان داده و شبکه اجتماعی مجازی

کلان داده، به مجموعه داده‌های عظیمی اشاره دارد که از منابع مختلف مانند آزمایش‌های علمی، دستگاه‌های حسگر، فعالیت‌های شبکه اجتماعی، داده‌های ارتباطات از راه دور، دوربین‌ها و... به‌دست می‌آید (Mohamed & Al-Jaroodi, 2014). زیکوپولوس و همکاران در سال ۲۰۱۲، در نشریه آی.بی.ام^۲، «کلان داده» را شامل مجموعه‌ای از سه کلمه «وی^۳» توصیف کرده‌اند: حجم^۴، سرعت^۵ و تنوع^۶ (به نقل از O'Leary, 2013; Milne & Watling, 2019). در این میان، شبکه‌های اجتماعی مجازی سایت‌هایی هستند که از یک سایت ساده مانند موتور جستجوگر با اضافه کردن امکاناتی مانند چت و ایمیل و امکانات دیگر، خاصیت اشتراک‌گذاری را به کاربران خود ارائه می‌دهند (یزدخواستی و همکاران، ۱۳۹۲: ۹۲) و عموماً نوع داده متنی در آن‌ها بیشتر است (Zhu, 2020). شهروندان خود را به این شبکه‌های مجازی متصل می‌کنند و از آن به‌عنوان بستری برای اظهار نظر در مورد عرصه‌های مختلف زندگی و همچنین ابراز هویت گروه‌های محروم اجتماعی استفاده می‌کنند (سلطانی‌فر و همکاران، ۱۳۹۲: ۱۷۸)؛ شبکه اجتماعی آنلاین سنتی، مبتنی بر موقعیت جغرافیایی نیستند. این در حالی است که اغلب شبکه‌های اجتماعی پرکاربرد امروزه مبتنی بر موقعیت مکانی هستند. سایت‌های شبکه‌های اجتماعی مبتنی بر مکان، یک چارچوب «۳ + ۱» را شامل می‌شوند. لایه جغرافیایی شامل بررسی‌های تاریخی کاربران است، درحالی‌که لایه اجتماعی حاوی اطلاعات دوستی اجتماعی است. لایه محتوا نیز

1. Desmet

2. IBM

3. V

4. Volume

5. Velocity

6. Variety

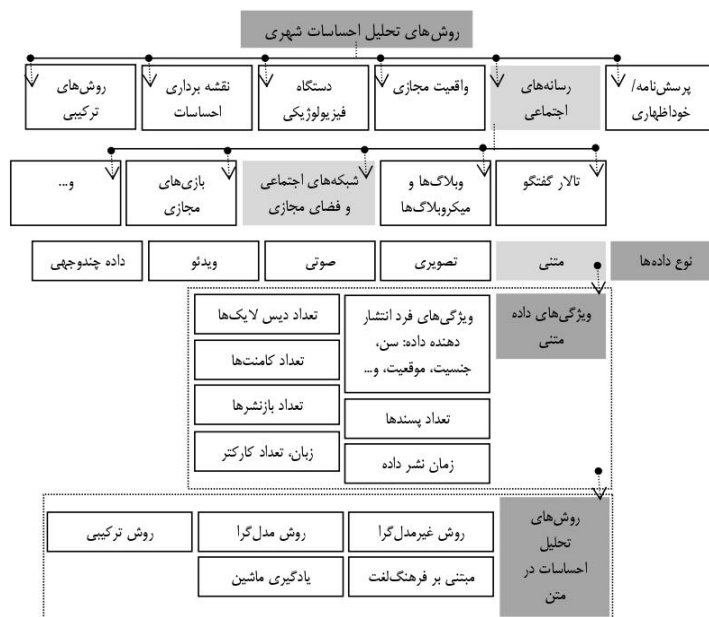
شامل بازخوردهای کاربر یا نکاتی درباره مکان‌های مختلف است. هر سه این لایه‌ها، یک جدول زمانی مشترک دارند و اطلاعات زمانی رفتار «ورود» کاربر را نشان می‌دهند (Gao & Liu, 2014).

تکنیک‌های طبقه‌بندی تحلیل احساسات متنی

داده‌کاوی وب، حوزه‌ای است که با استفاده از تکنیک‌های داده‌کاوی به استخراج اطلاعات از وب می‌پردازد. متن‌کاوی بخشی از محتوا کاوی وب است. متن‌کاوی فرآیند تحلیل متن بدون ساختار، استخراج الگوها و تبدیل آن‌ها به دانش است. تحلیل احساسات، تکنیک طبقه‌بندی متن در داده‌های متنی است که ارزش ذهنی سند متنی را که جهت‌گیری آن، مثبت یا منفی است، تعیین می‌کند (Kumar & Garg, 2020). تکنیک‌های طبقه‌بندی احساسات را می‌توان به سه دسته تقسیم کرد: روش‌های یادگیری ماشین، لغوی و ترکیبی. اولین تکنیک شامل الگوریتم‌های متداول یادگیری ماشین و شامل استفاده از ویژگی‌های زبانی است. مورد دوم تحلیل از طریق مجموعه‌ای از اصطلاحات احساسی است که از قبل در واژه‌نامه احساسات تدوین شده‌اند (Kumar et al., 2019). رویکرد ترکیبی شامل ترکیب رویکردهای یادگیری ماشین و لغوی است (Kumar et al., 2019). تکنیک‌های مبتنی بر یادگیری ماشین^۱ تحت نظارت، نیمه نظارت‌شده یا بدون نظارت هستند (Tyagi et al., 2019).

مبانی نظری

این نوشتار بر آن است تا از رسانه‌های اجتماعی و داده‌های نشر داده شده در آن، برای تحلیل احساسات بهره گیرد. همان‌گونه که در شکل زیر مشهود است روش مبتنی بر تحلیل داده‌های تحت وب می‌تواند مشتمل بر دو دسته غیرمدل‌گرا یا لغوی و مدل‌گرا یا یادگیری ماشین باشد. نوع داده قابل اتکاء از فضای منتخب بسته به نوع فضای اجتماعی مجازی می‌تواند متفاوت باشد. همچنین فارغ از نوع شبکه اجتماعی منتخب داده‌ها به چندین دسته بر اساس محتوای آن‌ها قابل دسته‌بندی هستند. این انواع شامل داده‌های ویدئویی، صوتی، تصویری و متنی هستند. از میان گونه‌ها، عموماً نوع داده‌های منتخب برای تحلیل احساسات، داده‌های متنی هستند، چراکه اغلب با گزاره‌های ذهنی سروکار دارند (Zhu, 2020)؛ هرچند داده‌های تصویری، صوتی و همچنین داده‌های چندوجهی نیز اخیراً موردتوجه قرار گرفته‌اند. در پژوهش حاضر نیز به دلیل نوع داده‌های در دسترس داده‌های متنی انتخاب شده‌اند. در ادامه فرآیند لازم است تا با دقیق‌تر کردن مسئله و انتخاب بستر فضایی پژوهش، فضای مجازی جهت جمع‌آوری داده‌ها استخراج و بعد از پیش‌پردازش آن‌ها، روش‌های تحلیل احساسات روی آن‌ها پیاده‌سازی شوند. در انتها نیز با صحت‌سنجی نتایج و بررسی مقایسه‌ای روش‌ها، روش منتخب، شناسایی و مدل پیش‌بینی کننده پژوهش تدوین شود.



شکل ۱. تحلیل احساسات در رسانه‌های اجتماعی و روش‌های تحلیل آن

روش پژوهش

این پژوهش از نظر هدف، کاربردی است. روش تحقیق در بررسی نمونه مطالعاتی، مبتنی بر روش کمی است. در ادامه فضای مجازی اجتماعی منتخب، بازه زمانی استخراج داده‌ها، و روش‌های مورد استفاده در هر دو روش برای آماده‌سازی داده‌ها و تحلیل آن‌ها ارائه می‌شود.

- **فضای شهری منتخب و فضای مجازی اجتماعی منتخب:** بستر فضایی انتخابی پژوهش، طی فرآیندی تعیین شده است. در ابتدا در رابطه با گونه‌هایی از فضاهای شهری در تهران چون میادین اصلی، مسیرهای اصلی، پارک‌ها و برخی از پروژه‌های بزرگ مقیاس مانند پل طبیعت و...، حجم داده‌های موجود در انواعی از شبکه‌های اینستاگرام، توئیتر، تریپ ادوایزر و گوگل مپ بررسی و پس از استخراج تعداد پست‌ها مشخص شد که شبکه اجتماعی گوگل مپ به دلیل تعداد زیاد داده‌ها، حجم کاراکتر متنی قابل پست کردن و پیام‌های مرتبط به نظرات و احساسات کاربران از سایر مکان‌ها مناسب‌تر است (جدول ۱). با توجه به اینکه تعداد پست‌های منتشر شده در رابطه با پارک ملت تهران، بیشتر از سایر فضاهای بررسی شده بود، این مکان به عنوان نمونه انتخاب شد.

- **بازه زمانی منتخب برای دانلود داده‌ها:** با توجه به محدودیت تعداد داده‌ها، و ضرورت نیاز به حداقلی از داده‌ها برای تحلیل مدل‌گرا و اعتبار نتایج داده‌ها، «بازه زمانی باز» برای جمع‌آوری داده‌ها انتخاب شد.

جدول ۱. نمونه‌ای از تعداد پیام‌های مرتبط با احساسات کاربران در فهرست اولیه مکان‌های منتخب

شبکه اجتماعی	میدان انقلاب	میدان امام خمینی (ره)	میدان ولیعصر (عج)	پارک آب‌و‌آتش	پارک گفتگو	پارک ملت	پیاده راه ۱۵ خرداد	بلوار کشاورز
تعداد پیام مرتبط با احساس و نظرات/کل پیام‌ها								
اینستاگرام	۱۰۰/۰	۱۲/۰	۵۰۰/۰	۱۰۰/۱	۱۰۰/۰	۱۰۰۰/۴	۱۰۰۰/۰	۵۰۰۰/۲۰
توئیتر	۰	۰	۱۰۰ / ۱	۴ / ۱	۱۶/۲	۱۰۰/ ۲	۰	۹۰/۲۰
فورا سکور	۲۵	۰	۶	۰	۰	۰	۰	۰
تریپ ادوایزر	۰	۰	۰	۱۱۱	۰	۱۳۴	۰	۰
گوگل مپ	۴۸۱	۵۸۵	۲۱۷۷	۳۲۵۷	۱۷۹۲	۳۴۶۴	۰	۰

- **حجم داده‌ها:** علی‌رغم آنکه در پژوهش‌های مبتنی بر یادگیری عمیق به داده‌های کلان و حجم حداقل ۳ تا ۴ هزار عددی از داده‌ها نیاز است، اما در پژوهش‌های مبتنی بر یادگیری ماشین، حجم داده‌ها می‌تواند کمتر از مقدار بیان‌شده باشد. درعین‌حال بررسی پژوهش‌های این حوزه نشان داده است که در نمونه‌ای از مطالعات با داده‌های کمتر از ۵۰۰ داده نیز، ماشین به‌خوبی آموزش داده شده و عملکرد مناسبی داشته است. هرچند قطعاً حجم داده‌ها بر یادگیری ماشین مؤثر است، اما محتوای داده انتخابی از آن مهم‌تر است. به این معنی که اگر محتوای داده ورودی دارای اشکال باشد، ماشین علی‌رغم اعتبار بالای به‌دست‌آمده می‌تواند، خروجی نادرستی را ارائه دهد. درعین‌حال با توجه به اینکه در پژوهش حاضر، از رویکرد نظارت‌شده استفاده شده است، ورودی و خروجی (در قالب برچسب) به ماشین داده شده و بنابراین نتیجه تست مدل و بررسی ماتریس درهم‌آمیختگی نشان از اعتبار مدل خواهد داشت. درعین‌حال استفاده از انواع متنوعی از الگوریتم‌ها و نزدیک بودن نتایج تعداد قابل‌توجهی از الگوریتم‌ها به هم، نشان از اعتبار حجم نمونه انتخابی دارد.

- **ابزار و کتابخانه‌های استخراج و داندلود داده‌ها:** استخراج داده‌ها با استفاده از اسکرپر^۱ در زبان برنامه‌نویسی پایتون انجام شد. برای این منظور از کتابخانه‌های پایتون سلیوم^۲ و بیوتیفول سوپ^۳ استفاده شد.

- **روش‌های تحلیل داده‌ها:** ۱- در روش مدل‌گرا از مدل‌های تحت نظارت یادگیری ماشین استفاده شده و از تمام نمونه‌های متداول الگوریتم‌ها که از خانواده و دسته‌های متفاوت که سابقه خوبی در عملکردشان داشتند (Bhavitha et al., 2017)؛ شامل رگرسیون لجستیک، ماشین بردار پشتیبان، جنگل تصادفی، کی نزدیک‌ترین همسایگی، ایکس جی بوست، تحلیل تفکیک خطی^۴، آدابوست، نیو بایس (بیز ساده)، درخت تصمیم‌گیری و مدل ترکیب گروهی، بهره‌برده شده تا عملکرد هر یک به‌منظور گرفتن پاسخ بهتر سنجیده شود. ۲- در روش غیرمدل‌گرا، از روش ویدر^۵ استفاده شد. ۳- درعین‌حال، در این پژوهش از مدل گروه ترکیبی^۶ نیز استفاده شد.

- **اعتبارسنجی نتایج:** برای اعتبارسنجی مدل‌ها از ماتریس درهم‌آمیختگی و معیارهای آن استفاده شده است.

ویژگی داده‌های مستخرج شده

تعداد کل نظرات پارک ملت تهران در گوگل مپ، برابر با ۳۴۶۴ رکورد بوده که پس از استخراج، پیام‌ها در فایل اکسل ذخیره شدند (جدول ۲). داده‌هایی که از پیام‌ها دریافت و استخراج شده‌است استخراج شده‌اند شامل: متن پیام، امتیاز پیام، تاریخ و ساعت ارسال پیام و تعداد امتیازی که بقیه افراد به پیام شخص داده‌اند، بوده است. بعد از پیش‌پردازش، تعداد کل نظرات قابل‌بررسی، معادل ۱۶۹۲ پیام شده است. لازم به ذکر است که به دلیل محدودیت‌های فرهنگ لغت‌های فارسی و همچنین به علت اینکه همه پیام‌ها به زبان فارسی و یا انگلیسی نبودند؛ همه پیام‌ها توسط یک مترجم، به انگلیسی ترجمه و برای انطباق به مترجمی دیگر برای بررسی و صحت ترجمه داده شد. در جدول زیر نمونه‌ای از داده‌های استخراج‌شده، ارائه شده است.

1. Web scraper
2. Python Selenium
3. BeautifulSoup
4. Linear Discriminant Analysis
5. VADER (Valence Aware Dictionary and Sentiment Reasoned)
6. Blending Ensemble

جدول ۲. نمونه‌ای از داده‌های استخراج‌شده از گوگل‌مپ در ارتباط با فضا

Index	Review_text	Review-rating	Review-datetime_utc	Review-likes
0	One of the best park in Tehran with different kind of facilities Its cool place all the year but these pick from autumn shows how amazing is the park in fall Parking is little bit issue in the area There is small public parking in the south but that's so small for such a huge park	5	2020/17/11 18:05:22	30
1	This park is fantastic, fabulous, beautiful, big and always clean, it is located in North of Tehran in high class district of Tehran, lots of adventure and festivals will be played in this park, behind of this park you will see a big and famous cinema, around of this park you will find very good restaurants, and here is a good place for picnic	5	2022/24/04 09:27:29	0
2	It's really beautiful and peaceful place it has lots of beautiful trees and it's really suitable for exercising and spending time with your friends	4	2022/05/20 11:25:20	0
3	Mellat Park is one of the my favorite parks in Iran It's beautiful, clean, well maintained, & it attracts kids, women, & men School trips also come here It's a nice place to spend a time, & explore the nature So, if you are a nature loving, then this peaceful, & beautiful park is waiting for you Entry is free as, it's a public park	5	12/24/2017 04:53:23	5

آماده‌سازی و پیش‌پردازش داده‌ها

به‌منظور پیش‌پردازش داده‌ها، ابتدا داده‌هایی که شامل متن نبوده و فقط امتیاز یا علائم نگارشی در پیام بوده یا پیام فارسی با استفاده از حروف انگلیسی نوشته شده بوده، حذف و سپس مراحل پیش‌پردازش روی داده‌ها انجام شد (جدول ۳).

الف) متعادل‌سازی داده: به دلیل اینکه داده‌های این پژوهش در طبقه مثبت بیشتر از طبقات منفی و خنثی هست عدم تعادل در طبقه‌بندی و کلاس وجود دارد که می‌تواند دقت مدل را بسیار کاهش دهد. در نتیجه برای حل این مشکل در این پژوهش از تکنیک‌های تغییر معیار عملکرد و تکنیک نمونه‌گیری بیش‌ازحد تصادفی استفاده شده است. تکنیک نمونه‌برداری بیش‌ازحد اقلیت مصنوعی، به افزودن نمونه‌های بیشتر از کلاس اقلیت با داده‌های مصنوعی که تغییرات کوچکی در آن‌ها ایجاد شده است، می‌پردازد و با انتخاب تصادفی یک نقطه از کلاس اقلیت و محاسبه کی نزدیک‌ترین همسایه‌ها برای این نقطه، کار می‌کند. با اجرای تکنیک اسموت^۱ و متعادل‌سازی داده‌ها، تعداد پیام‌ها در هر سه دسته: مثبت، منفی و خنثی ۹۸۴ عدد شده است. تکنیک بعدی استفاده شده تغییر معیار عملکرد است. در عین حال از آنجاکه بیان شده است برای ارزیابی داده‌های نامتعادل، صحت بهترین معیار نیست، از معیارهای دیگری چون دقت، فراخوانی، امتیاز^۲ اف ۱ که بینش بهتری را ارائه می‌دهند، استفاده شد.

ب) تبدیل حروف بزرگ به کوچک و حذف علائم نگارشی: به‌منظور دریافت نتیجه بهتر و بهبود عملکرد فرآیند یادگیری الگوریتم‌ها، داده‌های آموزشی با تبدیل حروف بزرگ در متن به حروف کوچک و همچنین حذف علائم نگارشی آن‌ها چون نقطه (.)، علامت سؤال (?)، علامت تعجب (!)، ویرگول^۳ و... می‌باشند، متن‌ها پاک‌سازی شدند.

ج) حذف کلمات توقف: کلمات توقف چون the, an, a, this, are, am, is و... در متن به‌عنوان نویز در نظر گرفته می‌شوند. در ان. ال. تی. کی^۳ برای حذف کلمات توقف، لیستی از کلمات توقف ایجاد شد تا کلمات موجود فیلتر شوند.

د) جداسازی کلمات: در این پژوهش جمله بر اساس فضاهای سفید تبدیل به کلمات جداگانه می‌شود.

ه) برچسب‌گذاری اجزای واژگانی کلام: هدف اصلی برچسب‌گذاری بخشی از گفتار شناسایی گروه دستوری یک کلمه معین است. در این پژوهش برچسب‌گذاری اجزای کلام با استفاده از ابزار ان. ال. تی. کی- پی. ا. اس- تگ، بعد از توکن‌سازی انجام شده و برای این منظور از تکنیک مبتنی بر قانون برای برچسب‌گذاری استفاده شده است.

1. Smote
2. Comma
3. NLTK

و) نرمال‌سازی متن با استفاده از تکنیک ریشه‌یابی کلمات: نرمال‌سازی واژگان نوع دیگری از نویز را در متن در نظر می‌گیرد و شکل‌های مرتبط با اشتقاق یک کلمه را به یک کلمه ریشه مشترک کاهش می‌دهد که با استفاده از بسته پایتون *in*. *l*. *ti*. *ki* که ابزار زبان طبیعی است، انجام شده است.

ز) نرمال‌سازی متن با استفاده از تکنیک کلمه‌سازی: این تکنیک کلمات را به کلمه اصلی خود که از نظر زبانی لم‌های صحیح واژه است، کاهش می‌دهد. در واقع در این مرحله با استفاده از واژگان و تحلیل مورفولوژیکی، ریشه کلمه تغییر می‌یابد. لم‌تیزشن^۱ یا نرمال‌سازی اطمینان می‌دهد که ریشه کلمه به زبان تعلق دارد. در این پژوهش به این منظور از بسته پایتون *in*. *l*. *ti*. *ki* استفاده شده است.

ط) بردارسازی داده با تکنیک *in*-*g*رم: برای طبقه‌بندی متن، باید متن‌ها را به عدد یا بردار اعداد تبدیل کرد. در این پژوهش از تکنیک *in*-*g*رم در رویکرد کوله کلمات که ساده‌ترین راه برای استخراج ویژگی‌ها از متن هست، استفاده شده است.

ح) بردارسازی داده با بردار شمارش: بردار شمارش^۲ از کتابخانه سایکیت لرن^۳ استفاده شده که برای تبدیل یک متن به بردار بر اساس فراوانی (تعداد) هر کلمه‌ای که در کل متن وجود دارد استفاده می‌کند. بردار شمارش یک ماتریس ایجاد می‌کند که در آن هر کلمه منحصر به فرد با ستونی از ماتریس نشان داده می‌شود و هر نمونه متنی از سند، یک ردیف در ماتریس است. مقدار هر سلول نیز تعداد کلمه در آن نمونه متن خاص است.

	I	like	this	park	here	one	of	the	best	in	Iran	beautiful	place
"I like this park, one of the best in Iran"	1	1	1	1	0	1	1	1	1	1	1	0	0
"I like here . beautiful place"	1	1	0	0	1	0	0	0	0	0	0	1	1

شکل ۳. نمونه برداری سازی داده‌ها

جدول ۳. نمونه‌ای از فرآیند انجام شده در پیش‌پردازش بر روی داده‌ها

مراحل پیش‌پردازش	نمونه داده قبل از پیش‌پردازش	نمونه داده بعد از پیش‌پردازش انجام شده
متعادل‌سازی داده‌ها	تعداد پیام‌های مثبت ۹۸۴ عدد و تعداد پیام خنثی ۷۰ عدد و تعداد پیام منفی ۲۸ عدد بوده است.	تولید داده‌های مصنوعی برای کلاس اقلیت تعداد پیام‌های مثبت ۹۸۴ عدد، تعداد پیام‌های منفی ۹۸۴ عدد، تعداد پیام‌های خنثی ۹۸۴ عدد
تبدیل حروف بزرگ به حروف کوچک	So beautiful and you can ride a bike	so beautiful and you can ride a bike
حذف علائم نگارشی	Bird park, water show, lake, garden, beautiful flowers, gym, and...	Bird park water show lake garden beautiful flowers gym and
حذف کلمات توقف و اعداد	Has beautiful bird and lake park and waterfront	beautiful bird lake park waterfront
توکن‌سازی کلمات	So beautiful and you can ride a bike	"So" "beautiful" "and" "you" "can" "ride" "a" "bike"
برچسب‌گذاری اجزای واژگانی کلام	I went there in the fall.	I(PR ⁴) went(VB ⁵) there(RB ⁶) in(IN ⁷) the(DT ⁸) fall(NN ⁹)
نرمال‌سازی متن با تکنیک ریشه‌یابی	not interesting	not interest
نرمال‌سازی متن با استفاده از تکنیک کلمه‌سازی کلمات	Beautiful park, but needs better maintenance.	Beautiful park, but needs <u>good</u> maintenance.

1. Lemmatization
2. Count Vectorizer
3. Scikit-learn
4. Personal pronouns
5. Verb conjugated
6. Adverb
7. Preposition or Subordinating conjunction
8. Determiner
9. Common noun

I like park "I like" "like park" / (بیوگرام) "park" (تریگرام) "I like park"	I like park. "I like this park, one of the best in Iran" & "I like here. beautiful place"	بردارسازی داده‌ها با تکنیک اِن-گرم بردارسازی داده‌ها با بردار شمارش
---	---	--

تقسیم داده‌های آموزشی و تست: به‌منظور طبقه‌بندی داده‌ها با الگوریتم‌های یادگیری ماشین، ابتدا داده‌ها را به دو قسمت ۸۰ درصد آموزشی (برابر با ۱۳۵۳ پیام) و ۲۰ درصد آزمایشی (۳۳۹ پیام) تقسیم و سپس داده آموزشی مجدد به دو بخش ۱۰۸۲ پیام آموزشی و ۲۷۱ پیام اعتبارسنجی تقسیم شدند.

برچسب‌گذاری داده‌ها: با استفاده از روش اکتشافی، امتیازی که کاربران به مکان داده بوده‌اند، برچسب‌گذاری انجام شده؛ به این صورت که هر کاربری که سه ستاره (امتیاز) به پیام داده باشد، به‌عنوان متن خنثی شناسایی می‌شود و اگر امتیاز چهار و پنج داده باشند، به‌عنوان پیام مثبت شناسایی می‌شود و در نهایت اگر امتیاز یک و دو داده شده باشد، به‌عنوان پیام منفی تلقی می‌شود.

محدوده مورد مطالعه

پارک ملت در شمال تهران و در منطقه ۳ شهرداری تهران قرار دارد. این بوستان از شمال به خیابان جام جم، از جنوب به خیابان نیایش، از شرق به خیابان ولیعصر (عج) و از غرب به خیابان سئول (باشگاه انقلاب) محدود می‌شود (شکل ۲). پارک ملت ابتدا زمین بایری بوده که احداث آن بر اساس موقعیت خاص خود در ۲ فاز متفاوت از سال ۱۳۴۵، آغاز شده است. فاز اول آن شامل قسمتی از پارک که به‌موازات خیابان ولیعصر (عج) است، به بلوار کنار جاده موسوم بوده و بخشی از آن که در قطعه‌ای از پیاده‌رو گسترده شده در سال ۱۳۴۷، بهره‌برداری شد. فاز دوم آن نیز در سال ۱۳۵۳، بهره‌برداری شد که شامل دو بخش تپه‌ای (جنگل‌کاری‌های شمال پارک و چمن‌کاری و گل‌کاری‌های جنوب پارک) است. طراحی بخش دوم پارک، به سبک پارک‌های انگلیسی است و نوع طراحی آن به‌صورت خطوط نامنظم و ارگانیک بوده است. مساحت این بوستان ۳۴ هکتار است که حدود ۲۳ هکتار آن به چمن‌کاری، ۲/۳ هکتار آن به جنگل‌کاری و ۰/۲ هکتار آن به گل‌کاری اختصاص یافته است. در این پارک حدود ۱۲۰ گونه درخت و درختچه و درخت میوه وجود دارد. این پارک دارای پنج زمین‌بازی برای کودکان، چهار سرویس بهداشتی، حدود ۹۰۰ نیمکت، دو بوفه و نمازخانه، دو آب‌نما، دریاچه با قابلیت قایق‌رانی و زمین پینگ‌پنگ است (وب‌سایت سازمان پارک‌ها و فضای سبز شهر تهران، ۱۴ تیر ۱۴۰۱).



شکل ۲. موقعیت جغرافیایی پارک ملت تهران

یافته‌ها

تحلیل قطبیت داده‌ها و شناسایی کلمات مرتبط با آن‌ها

پس از تعیین قطبیت کلی پیام‌ها به تحلیل احساسات در فصول مختلف سال، پرداخته شد. مطابق تحلیل‌های انجام‌شده، مشاهده شد که فصل پاییز ۳۷۹ پیام، فصل بهار ۴۶۸، در فصل تابستان ۴۴۶ و ۳۹۹ پیام در فصل زمستان در ارتباط با پارک ملت در گوگل ثبت‌شده است. در نتیجه همان‌گونه که مشاهده می‌شود، می‌توان نتیجه گرفت که تعداد بازدیدکنندگان از پارک در فصل بهار بیشتر از بقیه فصول سال بوده که می‌تواند نشان‌دهنده دما و هوای مطبوع پارک در فصل بهار، جذابیت بصری بیشتر آن در این فصل و همچنین وجود تعطیلات عید نوروز باشد. پس از فصل بهار با اختلاف بسیار کم، فراوانی پیام در تابستان مشاهده می‌شود. فصل زمستان بعد از بهار و تابستان، بیشترین فراوانی نظرات را دارد، اما تعداد پیام در فصل پاییز کمتر از بقیه فصول سال است. پس از تحلیل احساسات در فصول، تحلیل کلید واژه‌های پرتکرار پیام‌ها انجام شد. برای تعیین این کلمات، ایست واژه‌ها، حروف ربط و کلمات ("park", "Tehran", "place", "people", "mellat", "one", "see", "located", "many", "Park") که برای تعیین احساسات کاربردی ندارند، اما به تعداد زیاد در متن‌ها تکرار شده‌اند، حذف شدند تا کلمات با معنی و پرتکرار مشخص شوند (جدول ۴). همان‌گونه که انتظار می‌رود با توجه به فراوانی پیام‌های مثبت، کلیدواژه‌های مثبت بیشترین تکرار را داشته‌اند که از جمله پرتکرارترین آن‌ها می‌توان به کلمات (زیبا، خوب، بهترین، عالی) که معادل فارسی کلمات انگلیسی بودند اشاره کرد. وجود کلمات «پیاپی‌رو، سینما، دریاچه، باغ‌وحش، پارکینگ و ورزشی» نیز در ستون کلمات پرتکرار مثبت، نشان از احساسات مثبت مردم نسبت به فعالیت‌ها و سرگرمی‌های در پارک دارد. در طبقه‌بندی کلمات منفی پرتکرارترین کلمات، کلمه‌های (بد، متأسفانه، سرد) بوده همچنین وجود کلمات «پارکینگ، اسکیت» در ستون کلمات منفی، نشان‌دهنده وضعیت و دسترسی نامناسب پارک است که احساسات منفی را در کاربران فضا ایجاد کرده است. کلمات «سرد، شب، شلوغ، دسترسی، تنها، سرباز، امنیت، اهانت کردن، مسافر، مهاجر» نشان‌دهنده وضعیت‌ها و موقعیت‌هایی است که به کاربران احساسات منفی داده، مخصوصاً کلمات «مسافر، مهاجر، امنیت، شب، تنها، اهانت کردن» می‌تواند نشان‌دهنده عدم احساس امنیت در پارک باشد. همین‌طور کلمات «کثیف، چیز بی‌تناسب» نیز وجود احساس اغتشاش بصری و دید بصری نامناسب را نشان می‌دهد. کلمات «شهر، دریاچه، مکان، سال، خیابان، شمال، فصل، بهار، تابستان، پاییز، مغازه و شاه، قدیمی» در ستون

جدول ۴. کلیدواژه‌های پرتکرار پیام‌ها در سه دسته نظرات مثبت، منفی و خنثی

ردیف	کلمات مثبت	تعداد	کلمات منفی	تعداد	کلمات خنثی	تعداد
۱	('beautiful'),	۵۱۶	('bad'),	۹۹	('space'),	۹۰
۲	('good'),	۳۵۹	('Unfortunately,')	۴۵	('family'),	۷۲
۳	('best'),	۱۴۹	('cold,')	۲۷	('city'),	۵۴
۴	('great'),	۱۱۳	('parking'),	۲۷	('go'),	۵۴
۵	('Excellent'),	۱۱۱	('space'),	۲۷	('every'),	۵۴
۶	('Valiasr'),	۱۰۸	('lot'),	۲۷	('years'),	۵۴
۷	('Beautiful'),	۱۰۲	–	–	–	–
۸	('nice'),	۹۶	–	–	–	–
۹	('Great'),	۹۰	–	–	–	–
۱۰	('lake'),	۸۸	–	–	–	–



شکل ۴. ابر کلمات

بعد از استخراج و پیش‌پردازش داده‌ها، همان‌گونه که قبلاً بیا شد، به‌منظور یافتن الگوریتم پیش‌بینی کننده بهینه، ابتدا داده‌ها به دو بخش آموزش و تست تقسیم شدند. در جدول ۵، نتایج اجرای الگوریتم‌ها و مدل‌های طبقه‌بندی در مجموعه داده آزمایشی بر اساس معیارهای ارزیابی را می‌توان مشاهده کرد و به مقایسه معیارهای ارزیابی پرداخت. همان‌طور که در جدول نشان داده شده، از بین مدل‌های اجرا شده، الگوریتم ایکس.جی.بوست ۸۷ درصد از کل پیام‌ها را درست تشخیص داده و بهترین الگوریتم بوده است. پس از آن الگوریتم کی-نزدیک‌ترین همسایه، حدود ۸۰ درصد از کل پیام‌ها، ماشین بردار پشتیبان، ۷۰ درصد و بیز ساده و تحلیل تفکیک خطی به ترتیب حدود ۶۶ و ۵۷ درصد از کل پیام‌ها را درست تشخیص داده‌اند. سایر مدل‌ها کمتر از ۵۰ درصد و مدل جنگل تصادفی بدتر از تمام الگوریتم‌ها عمل کرده است.

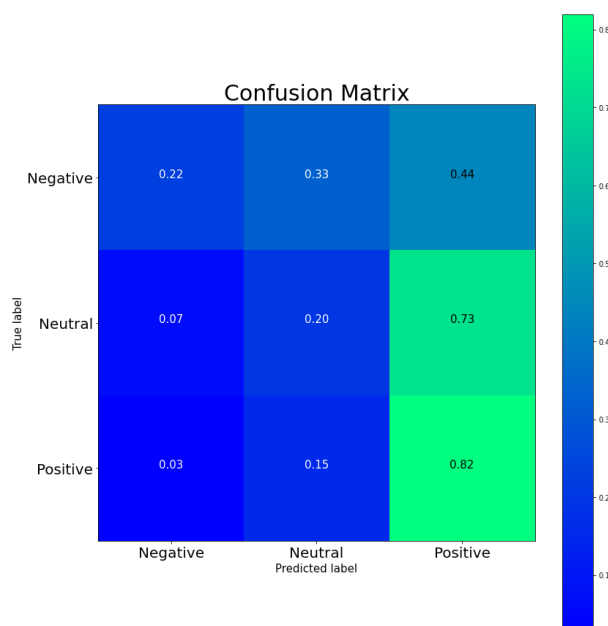
جدول ۵. پارامترهای هر الگوریتم و مقایسه معیارهای ارزیابی در الگوریتم‌های مورد استفاده

مدل و الگوریتم‌ها	معیار امتیاز اف ۱	معیار بازخوانی	معیار دقت ^۱	معیار صحت ^۲	اولویت
روش یادگیری ماشین					
رگرسیون لجستیک	۰.۶۱۰	۰.۴۷۸	۰.۹۰۵	۰.۴۷۸	۶
پارامترها: 'model__C': 0.0001, 'model__penalty': 'l2', 'model__solver': 'newton-cg'					
جنگل تصادفی	۰.۳۴۵	۰.۱۸۳	۰.۹۱۲	۰.۱۸۳	۸
پارامترها: 'model__bootstrap': True, 'model__max_depth': 5, 'model__max_features': 'auto', 'model__min_samples_leaf': 1, 'model__min_samples_split': 2, 'model__n_estimators': 50					
کی نزدیک‌ترین همسایگی	۰.۸۳۲	۰.۸۰۲	۰.۸۶۸	۰.۸۰۲	۲
پارامترها: 1: 'model__algorithm': 'auto', 'model__leaf_size': 10, 'model__n_neighbors': 3, 'model__p': 1					
ایکس جی بوست	۰.۸۷۶	۰.۸۷۰	۰.۸۸۵	۰.۸۷۰	۱
پارامترها: random_state=42, learning_rate: 0.02, max_depth = 4					
درخت تصمیم	۰.۴۶۸	۰.۳۳۶	۰.۹۰۸	۰.۳۳۶	۷
پارامترها: 'criterion': 'entropy', 'max_depth': 7, 'min_samples_leaf': 2, 'min_samples_split': 2, 'splitter': 'best'					
بیز ساده	۰.۷۵۲	۰.۶۵۸	۰.۸۹۱	۰.۶۵۸	۴
آداپوست	۰.۵۱۵	۰.۳۹۲	۰.۸۹۴	۰.۳۹۲	۷
پارامترها: n_estimators=100, random_state=0, 'model__algorithm': 'SAMME.R', 'model__n_estimators': 3					
تحلیل تفکیک خطی	۰.۶۷۹	۰.۵۶۹	۰.۸۸۶	۰.۵۶۹	۵
پارامترها: corpus, id2word = dictionary, num_topics = 10, iterations = 200					
ماشین بردار پشتیبان	۰.۷۸۲	۰.۷۰۸	۰.۸۹۸	۰.۷۰۸	۳
پارامترها: 'model__class_weight': 'balanced', 'model__kernel': 'linear'					

نتیجه اعمال تکنیک مبتنی بر لغت

روش دیگری که بر داده‌ها اعمال شد، روش مبتنی بر واژه بوده با استفاده از ویدر^۳ بود. در این روش احساسات هر پیام بر اساس امتیاز ترکیبی ویدر به‌عنوان مثبت، منفی یا خنثی طبقه‌بندی می‌شود. امتیاز مرکب، ارزش هر کلمه در پیام را در نظر می‌گیرد و با جمع امتیازهای ظرفیت هر کلمه در واژگان محاسبه می‌شود که طبق قوانین تنظیم می‌شود. سپس بین ۱- (بیشترین منفی) و ۱+ (شدیدترین مثبت) نرمال می‌شود. پیامی با امتیاز ترکیبی بزرگ‌تر یا مساوی ۰.۰۵ مثبت، پیامی با امتیاز بین ۰.۰۵- تا ۰.۰۵ خنثی و پیامی با امتیاز کمتر یا مساوی ۰.۰۵- منفی طبقه‌بندی شده که آستانه‌های استاندارد شده بر اساس ادبیات هستند. نتایج نشان داد که ۲۲٪ از داده‌ها در دسته داده‌های منفی درست تشخیص داده شده، ۳۳٪ از داده‌ها خنثی کاذب و ۴۴٪ مثبت کاذب شناسایی شده است. در دسته داده‌های خنثی نیز، ۷٪ از داده‌ها منفی کاذب، ۲۰٪ از داده‌ها خنثی درست و ۷۳٪ از داده‌ها مثبت کاذب تشخیص داده شده است. در داده‌های مثبت ۳٪ از داده‌ها منفی کاذب، ۱۵٪ از داده‌ها خنثی کاذب و ۸۲٪ از داده‌ها مثبت درست شناسایی شده است. در نتیجه در این روش بیشترین درصد درست مرتبط با داده‌های مثبت با ۸۲٪ تشخیص درست است. در نتیجه بهترین عملکرد را در تشخیص داده‌های مثبت داشته است اما عملکرد مطلوبی در تشخیص داده‌های منفی و خنثی نداشته است.

1. Precision
2. Accuracy
3. VADER

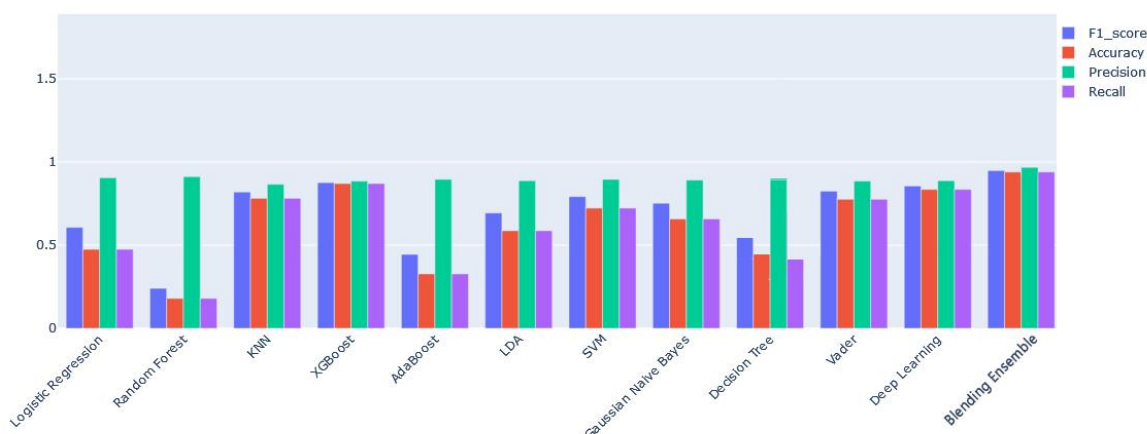


شکل ۵. ماتریس درهم‌آمیختگی روش مبتنی بر واژه

یادگیری گروهی - مدل پیشنهادی

مقایسه دو روش مدل‌گرا و غیرمدل‌گرا نشان داد که روش مبتنی بر واژه نسبت به برخی از الگوریتم‌های یادگیری ماشین نامناسب عمل کرده است. در ادامه به منظور ارائه مدل پیش‌بینی کننده با دقت بالاتر، از مدل‌های ترکیبی^۱ و از خانواده پشته‌سازی^۲ که در دسته مدل‌های یادگیری گروهی^۳ قرار می‌گیرد، استفاده شد. روش ترکیب، یک روش پشته‌سازی است که از اعتبارسنجی متقابل برای تهیه پیش‌بینی‌های خارج از نمونه برای متامدل یا یادگیرنده متا که در این پژوهش یک جنگل تصادفی است، استفاده می‌کند. در واقع، پیش‌بینی‌های هر مدل که از نتیجه اعمال مدل بر روی مجموعه اعتبارسنجی به دست آمده است، به عنوان ویژگی‌هایی برای آموزش و ساختن یک مدل جدید در مدل جنگل تصادفی (یادگیرنده متا) به همراه برچسب‌های صحیح پیام‌ها (به عنوان خروجی) به مدل داده می‌شود تا با استفاده از ورودی و خروجی داده شده، مدل آموزش ببیند. بعد از این مرحله مدل جنگل تصادفی برای پیش‌بینی نهایی روی مجموعه آزمایشی با تعداد ۳۳۹ پیام، ارزیابی می‌شود. نتیجه اعمال مدل ترکیب پیشنهادی به داده‌های آزمایشی در ماتریس درهم‌آمیختگی نشان می‌دهد که در دسته داده‌های منفی، ۸۰٪ از داده‌ها را درست، ۱۰٪ از داده‌ها را خنثی کاذب و ۱۰٪ از داده‌ها را مثبت کاذب شناسایی کرده است. در دسته داده‌های خنثی، ۰٪ از داده‌ها منفی کاذب، ۸۴٪ از داده‌ها خنثی درست و ۱۶٪ از داده‌های خنثی، مثبت کاذب تشخیص داده شده است. در بخش داده‌های مثبت نیز ۰٪ از داده‌ها منفی کاذب، ۳٪ خنثی کاذب و ۹۷٪ مثبت درست شناسایی شده است. در نتیجه به صورت کلی در مدل ترکیب پیشنهادی، بهترین عملکرد در داده‌های مثبت بوده است که ۹۷٪ به درستی تشخیص داده شده است و بیشترین خطا در داده‌های منفی به صورت کلی با ۲۰٪ خنثی کاذب و مثبت کاذب بوده است که در مقایسه با سایر الگوریتم‌ها درصد کمی خطا دارد و در هر سه دسته مثبت، خنثی و منفی مناسب عمل کرده است. در ارتباط با معیارهای ارزیابی این مدل نیز می‌توان درصد بالای ۹۵٪ را در همه معیارهای ارزیابی مشاهده کرد که نشان‌دهنده عملکرد خوب مدل هست.

1. Blending Ensemble
2. Stacking
3. Ensemble Learning



شکل ۵. مقایسه معیارهای ارزیابی در الگوریتم‌های بررسی‌شده

بحث

در این بخش یافته‌های پژوهش حاضر با ساختار دانش موجود مقایسه می‌شوند. همان‌گونه که پیش‌ازاین نیز بیان شد روش‌های متعددی غیر از روش منتخب پژوهش برای تحلیل احساسات در مورد استفاده قرار می‌گیرند. برای نمونه بهره‌گیری از روش‌های مبتنی بر پرسش‌نامه و یا خود اظهاری (Sauter et al., 2010, Nold, 2018, MacDonald, 2014; Pykett et al., 2020)، که می‌توانند در فضای واقعی یا فضای غیرواقعی (مبتنی بر تصویر فضا) انجام شوند. نتیجه تحلیل داده‌های به‌دست‌آمده در این روش‌ها می‌تواند ضمن مشخص کردن انواع احساساتی که در فضا وجود دارد، دلایل بروز احساسات را روشن نمایند.

دسته بعدی، پژوهش‌هایی هستند که مبتنی بر حسگر بوده و بنابراین امکان سنجش تغییرات فیزیولوژیک (Osborne & Jones, 2017) در بدن را به‌واسطه ابزار فراهم می‌آورند و بیشتر در محیط غیرواقعی (مبتنی بر واقعیت مجازی (Valtchanov & Ellard, 2010)) و کمتر در محیط واقعی (با استفاده از ابزارهای قابل حمل (Bergner et al., 2011)) انجام می‌شوند. یافته‌های به‌دست‌آمده با این روش، عموماً به‌شدت به دقت ابزار انتخاب متکی بوده و در موارد بسیاری هزینه انجام آزمون‌ها در این روش، گران است. درعین‌حال در این روش به دلیل کنترل تأثیر یک شاخص بر شاخص دیگر، آزمون‌ها ساده‌سازی شده و یا برخی از معیارها حذف می‌شوند که در نتیجه ممکن است، چندان یافته‌ها قابل اتکاء نباشند.

پژوهش‌های قابل توجهی در کشور با استفاده از هر دو روش در حوزه شهرسازی انجام‌شده که نتایج روش اول، عموماً مرتبط با یک مکان خاص و در روش دوم نتایج عموماً با تأکید بر یک یا چند عنصر خاص بوده است. در هر دو دسته، جمع‌آوری داده‌ها مستلزم صرف زمان و هزینه است. درعین‌حال مقیاس مورد تأکید در هر دو دسته بر خُرد فضاها بوده است. درعین‌حال باید به این موضوع نیز توجه داشت که در برخی از پژوهش‌ها (Fathullah & S.Willis, 2018; Engelniederhammer et al., 2019; Resch et al., 2015) از ترکیب هر دو روش بهره برده شده است.

اما روش منتخب پژوهش حاضر از قابلیت‌های داده‌های نشر داده شده در فضای وب که در ابتدا در حوزه‌های کسب‌وکار، صنعت و تجارت مورد استفاده گرفت و هم‌اکنون نیز در پژوهش‌های شهری بهره گرفته، استفاده نموده است. یکی از ویژگی‌های این روش بهره‌مندی از حجم زیاد داده‌های در دسترس در انواع فضاهای مجازی است که کاربران مختلف فضا به‌صورت داوطلبانه و اختیاری و در زمان‌های مختلف در رابطه با آن ثبت کرده‌اند و بنابراین زمان و هزینه جمع‌آوری حجم زیادی از داده‌ها، در این روش بسیار پایین است. علی‌رغم آنکه تحلیل داده‌ها در بدترین حالت (ناشی از

محتوای داده‌ها) می‌تواند به تعیین قطبیت داده‌ها منجر شود، اما دید کلانی نسبت به فضاهای مختلف پیش‌روی محقق قرار می‌دهد. همچنین این روش، به دلیل قابلیت پیش‌بینی پذیری آن، می‌تواند در زمان کمی قابل‌استفاده برای سایر فضاهای مشابه باشد. با توجه به موارد بیان‌شده، به نظر می‌رسد این روش بتواند در ترکیب با روش‌های پیش‌گفته نیز مؤثر باشد، برای نمونه می‌توان برای سنجش کلی وضعیت از این روش بهره برد و برای اعتبارسنجی صحت نتیجه در درصدی از نمونه‌ها، یکی از دو روش اول استفاده نمود.

نتیجه‌گیری

در این پژوهش تلاش شد تا مقوله احساسات کاربران در فضای سبز شهری بررسی و بر اساس نگاه و روشی نسبتاً نو با بهره‌گیری از داده‌های فضای مجازی، مدلی پیش‌بینی کننده، تدوین شود. نتایج پژوهش نشان داد که روش‌های مدل‌گرا بهتر از روش مبتنی بر واژه عمل کرده و همچنین روش ترکیبی بهتر از الگوریتم‌های یادگیری ماشین است. بنابراین با استفاده از این مدل آموزش‌دیده و تحلیل داده‌های شبکه‌های مجازی اجتماعی در پارک‌های شهری می‌توان تحلیل احساسات را پیش‌بینی و به طبقه‌بندی پارک‌ها از حیث احساسات مثبت و منفی پرداخت و پارک‌هایی که احساسات منفی را متبادر می‌کنند را شناسایی و برای رفع مشکل آن‌ها اقدام نمود.

در این بخش مقایسه حوزه نتیجه پژوهش با ساختار دانش موجود که در بخش پیشینه و مبانی نظری بیان شد، ارائه می‌شود. در پژوهش پای‌کن و پورجعفر (۱۳۹۸)، صرفاً به‌صورت مروری، اهمیت حوزه احساسات در محیط شهری تبیین شده است. در پژوهش‌های ربانی خوراسگانی و کیانپور (۱۳۹۲)، یزدخواستی و همکاران (۱۳۹۲) و نجاران و جلالی (۱۳۹۷) علی‌رغم آنکه تحلیل احساسات مبتنی بر شبکه‌های مجازی تحت وب است، اما بستر پژوهش شهر و فضای شهری نبوده است. پژوهش طاهر طلوع و همکاران (۱۳۹۲) نیز علی‌رغم تأکید بر احساسات شهری کاربران از فضاها، با روش واقعیت مجازی انجام‌شده است. بنابراین در پژوهش‌های داخلی مرتبط با حوزه شهری بسیار کم بر روش‌های تحت وب تأکید شده است. در میان پژوهش‌های خارجی نیز پژوهش‌های قهرمانی و همکاران (۲۰۲۱)، ونگ و همکاران (۲۰۲۱)، کواج-گیوری و همکاران (۲۰۱۸)، شوراتز و همکاران (۲۰۱۹) و لیم و همکاران (۲۰۱۸)، موضوع کیفیت فضاهای سبز شهری را مبتنی بر داده‌های منتشرشده در وب و بر اساس الگوریتم‌های یادگیری ماشین بررسی کرده‌اند و به طبقه‌بندی کاربران فضاهای سبز شهری، و یا طبقه‌بندی پارک‌ها و فضای سبز و یا تفاوت احساس قبل و بعد از حضور در فضاهای سبز پرداخته‌اند. از این‌رو پژوهش حاضر نیز از یکی از روش‌های موردتوجه این پژوهش‌ها یعنی یادگیری ماشین بهره برده است. اما صرفاً بر روش مدل‌گرا تأکید نداشته و از روش‌های غیرمدل‌گرا و لغوی نیز بهره برده است. پژوهش‌های دیگری چون کارتر و همکاران (۲۰۲۰) از اینترنت اشیاء برای تحلیل احساسات بهره برده‌اند و در دسته دیگری از پژوهش‌ها چون راجسورای و شانتیالا (۲۰۱۸) بر توسعه الگوریتم‌ها برای بهبود تشخیص یک احساس خاص چون کنایه و یا در پژوهش رابرتز و همکاران (۲۰۱۸) بر بررسی مقایسه‌ای روش‌های برچسب‌زنی و تأثیر آن بر اعتبار یافته‌ها پرداخته‌اند. در مقایسه با این دسته از پژوهش‌ها از این جهت که پژوهش حاضر در حوزه شهری است، موضوع توسعه الگوریتم یا بررسی روش‌های برچسب‌زنی موردتوجه این نوشتار نبوده است.

بر اساس تمام مطالب بیان‌شده نتایج به‌دست‌آمده در این پژوهش نشان داد که در پژوهش‌های داخلی نیز می‌توان با بهره‌گیری از داده‌های کلان موجود در شبکه‌های مجازی با اطمینان قابل قبولی به تحلیل احساسات پرداخت. درعین حال از حیث انتخاب روش نیز این پژوهش سعی کرده تا بر اساس دو روش اصلی حوزه تحلیل احساسات، داده‌ها را تحلیل نماید و همچنین با بهره‌گیری از تمام الگوریتم‌های متداول در حوزه یادگیری ماشین به تحلیل داده‌ها بپردازد درحالی‌که

در بیشتر پژوهش‌ها، صرفاً بر یک نوع و یا چند الگوریتم تأکید شده بود. درعین‌حال با استفاده از روش یادگیری ترکیبی نیز، این پژوهش سعی در افزایش اعتبار یافته‌های بهره‌مندی از روش مدل‌گرا در تحلیل احساسات داشته است. روش منتخب این پژوهش، کاربردهای فراوانی در حوزه شهرسازی دارد، چراکه به دلیل کم‌هزینه بودن آن در مقابل پژوهش‌های میدانی و سرعت در تحلیل داده‌ها، امکان شناسایی احساسات شهروندان از محیط را فراهم می‌آورد. باین‌حال در رابطه با محدودیت‌های این پژوهش می‌توان بیان کرد که به دلیل وجود محدودیت‌های منابع مالی در جمع‌آوری داده در پژوهش، توانایی انجام و تحلیل احساسات در مقیاس کلان -به‌عنوان مثال تحلیل احساسات پارک‌های متفاوت در شهر تهران یا در یک منطقه- وجود نداشته است. همچنین از دیگر محدودیت‌های پژوهش، مکان محور نبودن داده‌ها، برای تحلیل احساسات در پارک بوده است. در نهایت یکی دیگر از مسائل و محدودیت‌های قابل‌ذکر این نکته است که در زبان فارسی در مقایسه با زبان انگلیسی، فرهنگ لغات احساسات چون مدل ویدر در زبان انگلیسی، که به‌خوبی با مجموعه داده‌های بزرگ آموزش‌دیده باشند، وجود ندارد. همچنین عدم استفاده همه کاربران از شبکه‌های اجتماعی به دلیل محدودیت‌های دسترسی به اینترنت اشیاء را نیز می‌توان بیان نمود.

حامی مالی

این اثر حامی مالی نداشته است.

سهم نویسندگان در پژوهش

نویسندگان در تمام مراحل و بخش‌های انجام پژوهش سهم برابر داشتند.

تضاد منافع

نویسندگان اعلام می‌دارند که هیچ تضاد منافی در رابطه با نویسندگی و یا انتشار این مقاله ندارند.

تقدیر و تشکر

نویسندگان از همه کسانی که در انجام این پژوهش به ما یاری رساندند، به‌ویژه کسانی که کار ارزیابی کیفیت مقالات را انجام دادند، تشکر و قدردانی می‌نمایند.

منابع

- پای‌کن، عصمت و پورجعفر، محمدرضا. (۱۳۹۸). بررسی اهمیت و حوزه‌های کاربرد لایه «احساسات شهری» در طراحی و برنامه‌ریزی شهری. *معماری و شهرسازی/ایران*، ۱۰(۱۸)، ۳۹-۶۰. <https://doi.org/10.30475/isau.2020.103676>
- راپوپورت، ایماس. (۱۳۹۲). *معنی محیط ساخته‌شده: رویکردی در ارتباط غیرکلامی*. (ترجمه: فرح حبیب). تهران: شرکت پردازش و برنامه‌ریزی شهری
- ربانی خوراسگانی، علی و کیانپور، مسعود. (۱۳۸۸). *جامعه‌شناسی احساسات*. *جامعه‌شناسی کاربردی*، ۲۰(۲)، ۳۵-۶۴. [Dor:20.1001.1.20085745.1388.20.2.2.6](https://doi.org/10.22082/cr.2013.23569)
- سلطانی فر، محمد؛ بخشی، شیرزاد و فرامرزینی، سعید. (۱۳۹۲). بررسی تأثیرات شبکه‌های اجتماعی فیس‌بوک و توییتر بر گرایش اعضا به این شبکه‌ها. *پژوهش‌های ارتباطی*، ۲۰(۷۶)، ۱۷۱-۱۹۰. <https://doi.org/10.22082/cr.2013.23569>
- ظاهرطلوع دل، محمدصادق؛ ضرغامی، اسماعیل؛ کمالی تبریزی، سینا و حیدری پور، امید. (۱۳۹۸). بررسی قابلیت ادراک محیط در سیستم واقعیت مجازی بر اساس مؤلفه‌های ادراک بصری. *نشریه علمی اندیشه معماری*، ۳(۵)، ۱۰۶-۱۲۴. [Doi:10.30479/at.2019.10665.1194](https://doi.org/10.30479/at.2019.10665.1194)

محمدرضوی، سیدعلی؛ کیانپور، مسعود و آقابابایی، احسان. (۱۳۹۲). پدیدارشناسی احساسات کاربران در فضای مجازی (مطالعه موردی: جوانان شهر قم). *پژوهش‌های ارتباطی*، ۲۴ (۹۲)، ۷۳-۹۷. <https://doi.org/10.22082/cr.2017.67715.1395>

طیرانی نجاران، نصیر و جلالی، مهرداد (۱۳۹۷). ارائه رویکرد ترکیبی نوین جهت متن‌کاوی تحلیل احساسات در توییتر با استفاده از درخت تصمیم. *مدیریت مهندسی و رایانش نرم*، ۵ (۱)، ۲۱۰-۲۲۷. [Doi:10.22091/jemsc.2018.1272](https://doi.org/10.22091/jemsc.2018.1272)

وبسایت سازمان بوستان‌ها و فضای سبز شهرداری تهران (۱۴۰۱). بازیابی شده در ۱۴ تیر ۱۴۰۱، از [/https://parks.tehran.ir](https://parks.tehran.ir)

یزدخواستی، بهجت؛ عدلی پور، صمد و کیخانی، الهام. (۱۳۹۲). حوزه عمومی و گفتگو در فضای مجازی شبکه‌های اجتماعی (بررسی تأثیر شبکه‌های اجتماعی مجازی بر گفتگوی میان فرهنگ‌ها). *مطالعات فرهنگ-ارتباطات*، ۲۱ (۱۴)، ۸۱-۱۰۱.

[Dor:20.1001.1.20088760.1392.14.21.4.9](https://doi.org/10.22091/jemsc.2018.1272)

References

- Bergner, B. S., Zeile, P., Papastefanou, G., Rech, W., & Streich, B. (2011). Emotional Barrier-GIS – A new Approach to Integrate Barrier-Free Planning in Urban Planning Processes. (M. Schernk, V. Popovich & P. Zeile Eds.) *Real Corp, changing for stability: Lifecycles of Cities and Regions*, 247-257. <https://api.semanticscholar.org/CorpusID:25935756>
- Bhavitha, B. K., Rodrigues, A. P., & Chiplunkar, N. N. (2017). Comparative Study of Machine Learning Techniques in Sentimental Analysis. In *2017 International Conference on Inventive Communication and Computational Technologies (ICICCT)* (pp. 216-221). IEEE. [doi: 10.1109/ICICCT.2017.7975191](https://doi.org/10.1109/ICICCT.2017.7975191)
- Carter, E., Adam, P., Tsakis, D., Shaw, S., Watson, R. & Ryan, P. (2020). Enhancing Pedestrian Mobility in Smart Cities Using Big Data. *Journal of Management Analytics*, 7(2), 173-188. [DOI: 10.1080/23270012.2020.1741039](https://doi.org/10.1080/23270012.2020.1741039)
- Davidson, R. J., & Begley, S. (2012). The Science of Emotion: Exploring the Basics of Emotional Psychology. *Journal of Social and Clinical Psychology*, 31(4), 394-402.
- De Oliveira, T. H. M., & Painho, M. (2015). Emotion & Stress Mapping: Assembling an Ambient Geographic Information-based Methodology in Order to Understand Smart Cities. In *2015 10th Iberian Conference on Information Systems and Technologies (CISTI)* (pp. 1-4). IEEE. [DOI: 10.1109/CISTI.2015.7170469](https://doi.org/10.1109/CISTI.2015.7170469)
- Ekman, P., Sorenson, E. R., & Friesen, W. V. (1969). Pan-cultural Elements in Facial Displays of Emotion. *Science*, 164(3875), 86-88. <https://doi.org/10.1126/science.164.3875.86>
- Engelniederhammer, A., Papastefanou, G., & Xiang, L. (2019). Crowding density in urban environment and its effects on emotional responding of pedestrians: Using wearable device technology with sensors capturing proximity and psychophysiological emotion responses while walking in the street. *Journal of Human Behavior in the Social Environment*, 29(5), 630-646. <https://doi.org/10.1080/10911359.2019.1579149>
- Fathullah, A., & S. Willis, K. (2018). Engaging the Senses: The Potential of Emotional Data for Participation in Urban Planning. *Urban Science*, 2(98), 1-21. [doi:10.3390/urbansci2040098](https://doi.org/10.3390/urbansci2040098)
- Gao, H., & Liu, H. (2014). Data Analysis on Location-based Social Networks. *Mobile Social Networking*, 165-194.
- Ghahramani, M., Galle, N. J., Ratti, C., & Pilla, F. (2021). Tales of a City: Sentiment Analysis of Urban Green Space in Dublin. *Cities*, 119, 103395. <https://doi.org/10.1016/j.cities.2021.103395>
- Gu, S., Wang, F., Patel, N. P., Bourgeois, J. A., Huang, J. H. (2019). A Model for Basic Emotions Using Observations of Behavior in *Drosophila*. *Frontiers in Psychology*, 10(781). [DOI:10.3389/fpsyg.2019.00781](https://doi.org/10.3389/fpsyg.2019.00781)
- Gupta, B., Negi, M., Vishwakarma, K., Rawat, G., Badhani, P., & Tech, B. (2017). Study of Twitter Sentiment Analysis Using Machine Learning Algorithms on Python. *International Journal of Computer Applications*, 165(9), 29-34. [Doi:10.5120/ijca2017914022](https://doi.org/10.5120/ijca2017914022)
- Hao, J., & Dai, H. (2016). Social Media Content and Sentiment Analysis on Consumer Security Breaches. *Journal of Financial Crime*, 23(4), 855-869. [Doi:10.1108/JFC-01-2016-0001](https://doi.org/10.1108/JFC-01-2016-0001)
- Houtkamp, J. M. (2012). *Affective appraisal of virtual environments*. Unpublished PhD Thesis. Leiden University.

- Iaconesi, S., & Persico, O. (2013). An Emotional Compass Harvesting Geo-located Emotional States from User Generated Content on Social Networks and Using them to Create a Novel Experience of Cities. ESSEM Conference, Italy: Turin.
- Kovacs-Györi, A., Ristea, A., Kolcsar, R., Resch, B., Crivellari, A., & Blaschke, T. (2018). Beyond Spatial Proximity-Classifying Parks and their Visitors in London Based on Spatiotemporal and Sentiment Analysis of Twitter Data. *ISPRS International Journal of Geo-Information*, 7(9), 378. <https://doi.org/10.3390/ijgi7090378>
- Kumar, A., & Garg, G. (2020). Systematic Literature Review on Context-based Sentiment Analysis in Social Multimedia. *Multimedia Tools and Applications*, 79(21), 15349-15380. <https://doi.org/10.1007/s11042-019-7346-5>
- Kumar, D. P., Amgoth, T., & Annavarapu, C. S. R. (2019). Machine Learning Algorithms for Wireless Sensor Networks: A survey. *Information Fusion*, 49, 1-25. <https://doi.org/10.1016/j.inffus.2018.09.013>
- Lim, K. H., Lee, K. E., Kendal, D., Rashidi, L., Naghizade, E., Winter, S., & Vasardani, M. (2018). The Grass is Greener on the Other Side: Understanding the Effects of Green Spaces on Twitter User Sentiments. In *Companion Proceedings of the The Web Conference 2018* (pp. 275-282). DOI:10.1145/3184558.3186337
- MacDonald, G. (2014). Bodies Moving and Being Moved: Mapping Affect in Christian Nold's Bio Mapping. *Somatechnics*, 4(1), 108-132. doi: 10.3366/ soma.2014.0115
- Milne, D., & Watling, D. (2019). Big Data and Understanding Change in the Context of Planning Transport Systems. *Journal of Transport Geography*, 76, 235-244. https://jemsc.qom.ac.ir/article_1272.html?lang=en
- Mohamed, N., & Al-Jaroodi, J. (2014,). Real-time Big Data Analytics: Applications and Challenges. In *2014 International Conference on High Performance Computing & Simulation (HPCS)* (pp. 305-310). IEEE. Doi:10.1109/HPCSim.2014.6903700
- Mohammad Razavi, S. A., Kianpour, M., & Aghababaei, E. (2018) . Phenomenology of Users' Emotions in Cyberspace (Case study: Youth of Qom). *Communication Research*, 24(92), 73-97. https://cr.tribresearch.ir/?_action=article&kw=105233&_kw=%D8%A7%D8%AD%D8%B3%D8%A7%D8%B3%D8%A7%D8%AA+%D9%85%D8%AC%D8%A7%D8%B2%DB%8C&lang=en [in Persian].
- O'Leary, D. E. (2013). Big Data, the 'Internet of Things'and the 'Internet of Signs. *Intelligent Systems in Accounting, Finance and Management*, 20(1), 53-65. <https://doi.org/10.1002/isaf.1336>
- Osborne, T. & Jones, P.I. (2017). Biosensing and geography: a mixed methods approach. *Applied Geography*, 87, 160-169. <https://doi.org/10.1016/j.apgeog.2017.08.006>
- Paikan, E., & Pourjafar, M.R. (2020). The Importance and Application of "Urban Emotions" in Urban Design and Planning. *Journal of Iranian Architecture & Urbanism*, 10(2), 39-60. https://www.isau.ir/article_103684.html?lang=en [in Persian].
- Pykett, J., Chrisinger, B. W., Kyriakou, K., Osborne, T., Resch, B., Stathi, A. & Whittaker, A. C. (2020). Urban Emotion Sensing Beyond 'Affective Capture': Advancing Critical Interdisciplinary Methods. *International Journal of Environmental Research and Public Health*, 17(23), 1-22. doi:10.3390/ijerph17239003
- Rabbani Khorasghani, A., & Kianpour, M. (2009).The Sociology of Emotions. *Journal of Applied Sociology*, 20 (2), 35-64. https://jas.ui.ac.ir/article_18178.html?lang=en [in Persian].
- Rahman, A., & Hossen, M. S. (2019). Sentiment Analysis on Movie Review Data Using Machine Learning Approach. *2019 International Conference on Bangla Speech and Language Processing (ICBSLP)*, Bangladesh: Sylhet. doi: 10.1109/ICBSLP47725.2019.201470.
- Rajeshwari, K., & ShanthiBala, P. (2018). Recognition of Sarcastic Emotions of Individuals on Social Network. *International Journal of Pure and Applied Mathematics*, 18(7), 253-259. <http://www.ijpam.eu>
- Rapaport, A. (1990). *The Meaning of the Built Environment: A Nonverbal Communication Approach*. (Translated by Farah Habib). US: University of Arizona Press. [in Persian].
- Resch, B., Sudmanns, M., Sagl, G., Summa, A., Zeile, P., & Exner, J.-P. (2015). Crowdsourcing Physiological Conditions and Subjective Emotions by Coupling Technical and Human Mobile Sensors. *Geographic Information Science*, 1, 514-524. doi:10.1553/giscience2015s514

- Roberts, H., Sadler, J., & Chapman, L. (2018). The Value of Twitter Data for Determining the Emotional Responses of People to Urban Green Spaces. *Urban Studies*, 1-18. <https://doi.org/10.1177/0042098017748544>
- Sauter, D., Hogertz, C., Tight, M., Thomas, R., & Zaidel, D. (2010). Emotions of the urban pedestrian: sensory mapping. Pedestrians' Quality Needs, Cheltenham: WALK21. 31-52.
- Schwartz, A. J., Dodds, P. S., O'Neil-Dunne, J. P., Danforth, C. M., & Ricketts, T. H. (2019). Visitors to Urban Green Space Have Higher Sentiment and Lower Negativity on Twitter. *People and Nature*, 1(4), 476-485. DOI: 10.1002/pan3.100
- Soltanifar, M., Bakhshi, S., & Faramarziyani, S. (2013). Studying the Impacts of Facebook and Twitter on Members' Propensity towards Such Networks. *Communication Research*, 20(76), 171-190. [in Persian].
- Taher Toloudel, M.S., Zarghami, E., Kamali Tabrizi, S., & Heydaripour, O. (2019). The Analysis on Potential of Environmental Perception in System of Virtual Reality Based on Elements of Visual Perception. *Journal of Architectural Thought*, 3(5), 106-124. [in Persian].
- Tayarani Najaran, N., & Jalali, M. (2019). Presenting a Novel Hybrid Approach of Text Mining Sentiment Analysis in Twitter Using CART Decision Tree. *Engineering Management & Soft Computing*, 5(1), 210-227. <https://doi.org/10.22091/jemsc.2018.1272> [in Persian].
- Tehran Municipality Parks and Green Space Organization website (2022). Retrieved on 14 July 2022, <https://parks.tehran.ir> [in Persian].
- Tyagi, P., Chakraborty, S., Tripathi, R. C., & Choudhury, T. (2019). Literature Review of Sentiment Analysis Techniques for Micro blogging Site. In *International Conference on Advances in Engineering Science Management & Technology (ICAESMT)-2019, Uttaranchal University, India: Dehradun*. <http://dx.doi.org/10.2139/ssrn.3403968>
- Valtchanov, D., & Ellard, C. (2010). Physiological and affective responses To immersion in Virtual reality: effects of nature and urban settings. *Journal of Cyber Therapy & Rehabilitation*, 3(4), 359-373.
- Wang, Z., Zhu, Z., Xu, M., & Qureshi, S. (2021). Fine-grained Assessment of Green Space Satisfaction at Regional Scale Using Content Analysis of Social Media and Machine Learning. *Science of the Total Environment*, 776, 145908
- YazdKhasti, B., Adlipour, S., & Keikhaei, E. (2013). Public Sphere and Dialogue in Social Networks (The Study of the Effect of Virtual Social Networks on Cross-cultural Dialogues). *Journal of Culture-Communication Studies*, 14(21), 81-101. [in Persian].
- Zhu, Y. (2020). The prediction Model of Personality in Social Networks by Using Data Mining Deep Learning Algorithm and Random Walk Model. *The International Journal of Electrical Engineering & Education*, 0020720920936839. <https://doi.org/10.1177/0020720920936839>