Nicotine and Tobacco Research, 2025, **27**, 988–996
https://doi.org/10.1093/ntr/ntae276
Advance access publication 23 November 2024
Original Investigation

AI for Tobacco Control: Identifying Tobacco-Promoting Social Media Content Using Large Language Models

Hüseyin Küçükali PhD[1,2], **Mehmet Sarper Erdoğan PhD[3]**

[1]Centre for Public Health, Queen's University Belfast, Belfast, UK;
[2]Research Center for Healthcare Systems and Policies, Istanbul Medipol University, Istanbul, Türkiye;
[3]Department of Public Health, Istanbul University—Cerrahpasa, Istanbul, Türkiye

Corresponding Author: Hüseyin Küçükali, PhD, *Queen's University Belfast, Centre for Public Health, BT12 6BA Belfast, UK.*
E-mail: h.kucukali@qub.ac.uk

Abstract

Introduction: Tobacco companies use social media to bypass marketing restrictions. Studies show that exposure to tobacco promotion on social media influences subsequent smoking behavior, yet it is challenging to monitor such content. We developed an artificial intelligence that can automatically identify tobacco-promoting content on social media.

Aims and Methods: In this mixed methods study, 177,684 tobacco-related tweets published on Twitter in Turkish were collected. Through inductive content analysis of a sample of 200 tweets, the main mechanisms by which tobacco is promoted on social media were identified. Then, a sample of 5000 tweets was deductively analyzed and labeled based on those mechanisms. A pre-trained transformer-based Large Language Model was fine-tuned using the labeled dataset. Then, tobacco promotion in all tweets was predicted using this model.

Results: The main mechanisms of tobacco promotion on social media included modeling the behavior, expressing positive attitudes, recommending use, and marketing brands or vendors. The developed model identified tobacco-promoting social media content with 87.8% recall and 81.1% precision. The utility of the model was demonstrated in the analysis of tobacco promotion in tweets for a period of a month.

Conclusions: This tool makes it possible to monitor tobacco promotion in social media and creates new opportunities for tobacco control policy and practice, not only in surveillance and enforcement but also in health promotion.

Implications: Tobacco promotion in social media is a well-known yet hard-to-addressed problem due to the nature of social media. This study leverages a cutting-edge AI approach, Large Language Models, to identify tobacco promotion in social media content automatically and precisely. The developed model offers better prediction performance than previously proposed techniques. The study enables surveillance of tobacco-promoting content both for research purposes and enforcement of tobacco control measures. Furthermore, we suggest a range of health promotion opportunities this tool can help with from developing personal skills to creating supportive environments and strengthening community actions.

Introduction

Prevention of tobacco advertising and promotion constitutes one of the pillars of tobacco control according to the policy package formulated as MPOWER by the World Health Organization (WHO).[1] In this context, laws in many countries ban the marketing of tobacco products through conventional media channels (television, radio, newspapers, etc.). Indicating the ineffectiveness of partial bans, WHO emphasizes that bans should cover the entire media, especially the new media; otherwise, tobacco companies find new ways as enabled by their high marketing budgets.[1]

Social media can turn into the "escape area for illegality" due to its size, uncontrolled nature, and anonymity.[2] International tobacco companies bypass the said bans thanks to social media. In recent years, many news reports have been published that expose the marketing activities undertaken by international tobacco companies all around the world through "influencers."[3–6] In addition, tobacco-promoting content in social media is not only limited to the marketing activities of tobacco companies but can also be produced by ordinary individuals.[7]

The theory of information scanning[8] asserts that individuals may acquire information even if they are not actively seeking it. A prospective study on young adults revealed that passive exposure to social media content promoting tobacco determined smoking behavior 6 months later and this effect was indicative of the dose–response relationship.[7] Again, another larger study demonstrated this effect on media exposure related to electronic cigarettes.[9]

In Türkiye, 66.5% of students watching TV channels, videos, or motion pictures see a tobacco user on these media.[10] However, it is neither known nor monitored to what extent the exposure to social media content promoting tobacco is. While an amendment in the Turkish tobacco control law[11] in 2018 banned the use of images of tobacco products on social media as with other media, there are some challenges in identifying violations in social media due to its characteristics given above. The Tobacco Control Strategy 2018–2023 mentioned that tobacco companies were engaged in advertising and promotional activities through social media due to the advertising ban, and set a strategy for "developing a system to identify campaigns and programs of the tobacco industry which

Received: January 31, 2024. Revised: August 8, 2024. Accepted: November 14 2024.

encourage young people to use tobacco products and facilitate their access to such products."[12] Our freedom of information query to the organization responsible for this strategy showed that no solution has been introduced and the set targets have not been met, yet (Supplementary Material S1).

Considering the production volume and consumption rate for social media content, it is very difficult to identify content promoting tobacco. This difficulty makes it harder not only to monitor tobacco promotion on social media but also to give a proper public health response. Individual investigations by cumbersome bureaucratic organizations upon complaints, penalties to be imposed if the perpetrator can be identified, and access restrictions will not act as a solution to the issue, although these measures are included in the law and the strategy document. Tobacco promotion in social media stands as a major gap in the primary prevention of tobacco addiction.

Artificial Intelligence technologies enable the performance of tasks that normally require human cognition on a faster, larger, and constant basis. Thus, they can contribute to solving this issue. As a result of developments in recent years, artificial intelligence technologies have been very successful in a wide range of tasks, such as comprehension of natural language, text classification, translation, summarization, question answering, etc.[13,14] In this context, identifying any content that promotes tobacco on social media can also be formulated as a text classification problem. If artificial intelligence manages to identify such content successfully, it can serve for surveillance, prevention, and health promotion purposes.[15]

This study aims to develop and demonstrate an artificial intelligence that can automatically identify tobacco-promoting content on social media.[16] Objectives of the study are as follows:

1. Understanding and defining the ways in which tobacco promotion takes place on social media,
2. Using these definitions, create a dataset of a large number of social media content with labels on whether they promote tobacco,
3. Using the labeled dataset, fine-tuning a large language model to be able to classify such content,
4. Using this model, predicting and exploring tobacco promotion on social media within a period of a month.

Methods

Study Design

This was a sequential mixed methods study where a qualitative analysis was followed by a quantitative analysis (a.k.a. qual → QUAN[17]). First, the tobacco-related social media content was collected from social media. Second, inductive content analysis[18] was applied to a small sub-sample to define the ways in which tobacco was promoted on social media. Third, a larger sample was labeled by human reviewers by deductive content analysis, using the themes identified in the previous step. Fourth, a large language model (LLM) was fine-tuned by supervised machine learning using the labeled dataset, followed by the evaluation of the model's performance. Fifth, tobacco promotion in all collected social media content was predicted using the fine-tuned model. Lastly, descriptive statistics on tobacco promotion on social media were presented. Figure 1 illustrates the study's workflow.

Data Collection

The data consisted of tweets (n = 177,684) published publicly on Twitter in Turkish on 1st–31st October 2020 and containing at least one tobacco-related keyword. The data were collected by a Python script using Standard Search API v1.1.[19]

Following the recommendations of Kim et al.,[20] a set of keywords was defined using (i) type names of tobacco products, (ii) brand names of conventional tobacco products, and (iii) brand names of electronic tobacco products. The types commonly used in Türkiye were listed with spelling variations. A total of 34 sub-brands of 5 manufacturers were acquired from local retailers for traditional cigarette brands available in the Turkish market. In total, 18 sub-brands of 11 manufacturers from an international market analysis[21] were used for e-cigarettes. Some keywords resulted in a high rate of irrelevant data. For instance, the keyword for a cigarette brand "Maltepe," also yielded tweets referring to a location with the same name. Thus, the keywords yielding relevant results at least half of the time were selected. The final search query is provided in Supplementary Table S1.

The literature lacks an agreed sample size that is sufficient for training/fine-tuning a language model. Based on a review of similar studies[22] 5000 tweets were sampled randomly by weighting the number of tweets per day.

Qualitative Content Analysis: Inductive Phase

Inductive content analysis was used to identify how tobacco was promoted through tweets.

To this end, a small sub-sample consisting of 200 tweets was selected randomly from the 5000-tweet sample. Tweets were coded inductively as to how tobacco promotion occurred, if any. When the codes were grouped, 4 themes emerged as tobacco promotion mechanisms in social media:

1. Behavior: By stating the behavior of using a tobacco product;
2. Attitude: By expressing positive attitudes towards tobacco;
3. Recommendation: By directly recommending the use of a tobacco product;
4. Marketing: By indicating the name of a specific tobacco product or the location where it is sold.

However, the classification of all tweets was not straightforward. Some involved multiple themes, and contradictory statements both promoting and opposing tobacco. Therefore, a 4 × 4 coding frame was developed (Supplementary Table S2) to capture the complicated nature of the content. Each tweet was coded with a category under each theme either (1) against tobacco, (2) favors tobacco, (3) both, or (0) none (0). For example: a tweet that reads "I smoked an e-cigarette for the first time today, it's not meant for me, but anyone curious should give it a try" is coded as behaviour:2, attitude:1, recommendation:2, and marketing:0. Tweets coded favoring tobacco under at least one theme were considered tobacco-promoting.

Qualitative Content Analysis: Deductive Phase

The 5000-tweet sample was deductively labeled by human reviewers using the coding frame.

To reduce possible inter-reviewer bias and support triangulation, a team of 27 volunteers was formed and each tweet

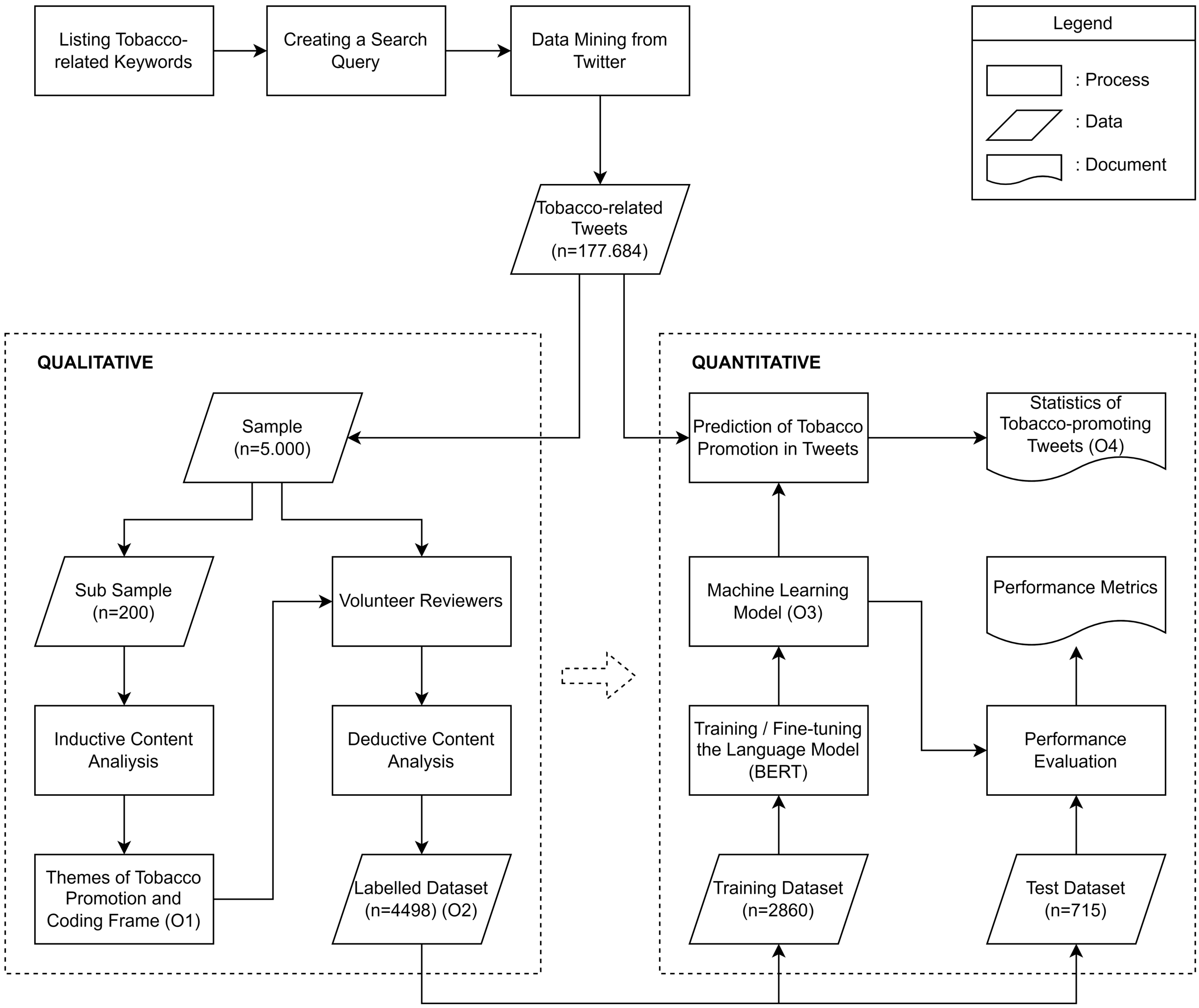

Figure 1. Study processes. O1–O4 indicate study objectives.

was coded by 2 randomly selected reviewers. Reviewers were provided with 4-h training which involved a briefing about the purpose and methods of the project, and instructions on accessing the tweets and usage of the coding framework. The researcher coded a handful of tweets together with volunteers and answered their questions.

A simple chatbot was programmed in Python to randomly assign tweets to reviewers and record their coding. As a reminder, the coding framework (Supplementary Table S2) was provided under each tweet.

Natural Language Processing

Typical in social media data, informal language use with typing mistakes and abbreviations can hinder prediction performance. To avoid that, an academic web service[23] was used to pre-process texts for machine learning by tailoring them into formal written language. Developed specifically for the Turkish language, the service provided seven functions including transformation of letter cases, lexicon lookup, detection of proper nouns, correction of diacritics, restoration of missing vowels, accent normalization, and spelling correction.[23] In addition to that, usernames and URLs were removed using regular expressions while hashtags and emojis were retained.

Machine Learning

A language model that can predict whether a given tweet text is tobacco-promoting was developed and evaluated for its performance.

This study employed the Bidirectional Encoder Representations from Transformers (BERT), an LLM that was developed by Google in 2018 and proven to be more successful in language processing tasks than other methods to date.[24] Typically, an LMM is used in two steps: pre-training and fine-tuning. In the pre-training step, the model is trained to represent the language in general through the use of a very large corpus of texts. At this stage, the model cannot make classification. Then, fine-tuning is carried out by relatively smaller sets of training data, enabling the model that has already “understood” a language in general to learn a specific classification task. This property of BERT allows to use of an advanced model, which would normally require a long time and high computing power, in a specific task through fine-tuning.[24] This study used the open-source BERTurk model with a vocabulary of 128,000 words, which was pre-trained by Schweter in 2020 using 35 GBs of data in Turkish.[25]

The general-purpose BERTurk model was then fine-tuned to the specific objective of the study. Fine-tuning was executed on a server with a Graphics Processing Unit rented from Google. The dataset was divided into 4:1 training and test datasets by random sampling stratified on the categories of the outcome variable. First, the model was fine-tuned by introducing the training dataset to the algorithm. Then, the model was applied to the data that was not encountered by the model during the training, i.e., test dataset, and a value was predicted for each tweet in this dataset. The model performance was evaluated by comparing the predicted values against the values assigned by human reviewers.

The training and testing process was carried out by first including all the reviews (model 1) and then only consistent reviews (model 2) to predict the promotion of tobacco. Tobacco promotion in all tweets collected was predicted using the second model.

Although the primary aim of the study was to identify any promotion of tobacco, additional models were also developed for the prediction of the four themes (behavior, attitude, recommendation, and marketing) individually. These models were developed in six scenarios to better account for the underlying distribution of categories in each theme:

1. Trained on all reviews, predicts four categories (favors tobacco, against tobacco, both, and not applicable),
2. Trained on all reviews, predicts two categories (favors tobacco vs others),
3. Trained on all reviews, predicts two categories (favors tobacco vs against tobacco),
4. Trained on consistent reviews, predicts four categories (favors tobacco, against tobacco, both, and not applicable),
5. Trained on consistent reviews, predicts two categories (favors tobacco vs others),
6. Trained on consistent reviews, predicts two categories (favors tobacco vs against tobacco).

In total, 26 models were developed, 2 for predicting tobacco promotion and 24 for predicting four themes in six scenarios.

Statistical Analysis

Accuracy, sensitivity (a.k.a. recall), specificity, positive predictive value (a.k.a. precision), F1 Score, Pearson's phi (a.k.a. Matthews correlation coefficient), and Cohen's kappa statistics were calculated to evaluate the performance of machine learning models. Descriptive statistics on the distribution of tobacco-promoting tweets among all the collected tweets were computed.

Software

Python scripts were programmed for each step of the study by the researcher. The workflow of these scripts, its connections with cloud services and the database, and the open-source packages used are shared in Supplementary Material S2.

Results

Description of the Dataset

A total of 177,684 tweets containing tobacco-related keywords were collected. On average, 6253.2 ± 1011.1 (min.: 4447, max.: 8252) tweets were posted per day. These tweets were posted by 126,503 individual accounts. A total of 80% of the accounts posted only one tweet and 12.5% posted two. A total of 37.6% of the tweets were original tweets, 37.8% were retweets and quotations of other tweets, and 24.7% were replies. The majority of tweets (57.8%) contained mentions (of other users). In total, 87.8% of tweets were posted through mobile devices. Further characteristics of tobacco-related tweets are provided in Supplementary Figure S1 and Supplementary Table S3.

Qualitative Findings

The main qualitative findings are described in the method section since they informed subsequent processes. Here, a few example quotations are provided for the distinct themes in content analysis. Four themes emerged about tobacco promotion on social media: behavior of using tobacco, attitude toward tobacco, recommendation for tobacco use, and marketing of specific tobacco products and points-of-sale. While the themes in some tweets were simple and easily identifiable (Table 1), the case with most others was more complicated.

For example, a tweet "Turkish coffee and cigarette, I'll stick with you even to the bitter end" (id: 1322324511711178752) expresses an attitude favoring tobacco, even though cigarette was referred to together with death. Another tweet "Sometimes, it smells so good when someone around me lights a cigarette that I would stick a cigarette up in my ears and nose if I didn't know it would be a waste of time after 8 months of trying" (1317147491914186754) shares

Table 1. Exemplary Quotations for Various Ways of Tobacco Promotion[a]

	Favors tobacco	Against tobacco
Behavior	Look how awesome the night is I'm enjoying a cigarette and coffee along with the rain (1317534881568772097)	I have not been smoking for 3 days ♥ (1321081674495655938)
Attitude	Dude, coffee and cigarette make a very good duo on the balcony when the weather is like this (1314247211987939329)	Those smokers, you're not cool you just look like a walking chimney (1312834039603245059)
Recommendation	Go out on the balcony and light a cigarette. Cheers to your dreams that will not come true even if you die. (1319448024184225792)	You smoke a lot. Quit smoking. (1315241254444294145)
Marketing	I'll just smoke a Marlboro touch blue, it is the best (1317604450052280323)	Let smoking on the sidewalk be banned. (1320074867744690176)

[a]Tweet ids are provided in brackets.

behavior against tobacco use while holding an attitude favoring tobacco, both implicitly. On the contrary, tweets like "I feel so sorry for myself because I'm a cigarette addict" (1320251910314745862) or "Again, I failed to quit goddamn smoking" (1316060965956452357) express an attitude against tobacco, while they contain tobacco use behavior. Tweets, such as "Be like a cigarette in this life. It reads 'kills' on it, yet it is indispensable...!" (1316678772326903808) or "Rain, sad songs and cigarette, these three have a bittersweet taste" (1322127077064232960) or "The urge to smoke and nausea while smoking, these are in every moment of my life now" (1315386587644137473), include both negative and positive attitudes together. The following tweet promotes a specific product while, expressing both the behaviors of smoking and smoking cessation: "I was going to light my last cigarette for old times' sake, but I had two left. Anyway, I'll quit smoking, friends and I'll give away my last single Camel White to one of my followers. (…)" (1317931126154055682). These examples provided insights into the complicated nature of tobacco promotion in social media.

A total of 502 tweets were excluded due to profane language which was captured using an open-sourced list.[26] The remaining 4498 tweets were coded by two reviewers. Tobacco promotion was coded consistently in 3575 (79.5%) tweets. In total, 2005 tweets (56.1%) were promoting tobacco under at least one theme. A high proportion of tweets (35.4%) were containing tobacco use behavior. Raw coding by reviewers and tobacco promotion labels are shared as open data. The distribution of codes by themes and categories is summarized in Supplementary Table S4.

Machine Learning Performance

The performance metrics of models 1 and 2 predicting tobacco promotion in tweets are shown in Table 2. As expected, model 2, which was trained on consistently coded tweets, performed better than model 1, across all metrics. Model 2 could discriminate whether social media content promotes tobacco with 81.7% accuracy and 87.8% precision. The model, a script to run it, and the predictions are shared open access.

The performance metrics of models predicting categories for individual themes are available in Supplementary Table S5. The models trained on consistent reviews were more successful than the ones trained on all reviews. For instance, models 19–22 predicted tobacco use behavior, the attitude toward tobacco, tobacco use recommendation, and tobacco marketing with 81.9%, 89.3%, 98.3%, and 99.2% accuracy, respectively.

Predicted Tobacco-Promotion

A total of 70,803 tweets posted in October 2020 in Turkish were predicted to be tobacco-promoting. This accounted for 39.9% of the tweets containing any tobacco-related keyword and 2360.1 ± 599.4 tweets per day. The percentage of tobacco-promoting tweets per day (min: 28.3% on Oct 11, max: 52.4% on Oct 18) is shown in Supplementary Figure S2. It was similar on different days of the week (min: 37.5% on Saturday, max: 42.2% on Wednesday) or weekdays and weekends (40.0% vs 39.5%).

Figure 2 shows that more tobacco-related tweets were posted at night than in daytime hours (which peaks at 23:00–23:59). Tweets promoting tobacco were more frequent at night (18:00–05:59, 43.7%) than in daytime hours (06:00–17:59, 34.3%), especially after midnight.

Other characteristics of tweets promoting tobacco are provided in Supplementary Table S6. While 58.9% of original tweets promoted tobacco, only 20.5% of retweets promoted tobacco. It was noteworthy that 60.3% of tweets promoting tobacco mentioned another user. The percentage (45.2%) of tobacco promotion was higher among tweets posted via mobile devices than other devices.

The first 50 words that were most frequently used in tweets are shown in the word cloud in Supplementary Figure S3 (stop words[27] were removed). Tobacco-promoting tweets included words about foods (e.g., tea, coffee, etc.) consumed with tobacco as well as environments of tobacco use (e.g., night, balcony, etc.) whereas addictive substances (e.g., alcohol, beer, etc.) and financial words (e.g., TRY, lira, tax, etc.) were frequently mentioned in tweets that do not promote tobacco.

Discussion

Main Outcomes

Different Ways of Tobacco Promotion on Social Media Were Defined in Four Themes

Tobacco was promoted in social media by modeling tobacco use behavior, expressing a positive attitude toward tobacco, recommending the use of tobacco, and marketing a specific brand or point-of-sale. It is worth noting that tweets could contain combinations of themes, which may sometimes contradict each other.

Table 2. Performance Metrics of Machine Learning Models Predicting Tobacco Promotion in Social Media

	Model 1: all codings (n = 8,996)	Model 2: consistent codings (n = 3,575)
Accuracy	79.7%	81.7%
PPV[a] (Precision)	81.9%	81.1%
Sensitivity (Recall)	80.8%	87.8%
Specificity	78.4%	73.9%
F_1 score	81.3%	84.3%
Pearson's phi (MCC[b])	0.59	0.63
Cohen's kappa	0.59	0.62

[a]PPV: Positive predictive value.
[b]MCC: Matthew's correlation coefficient.

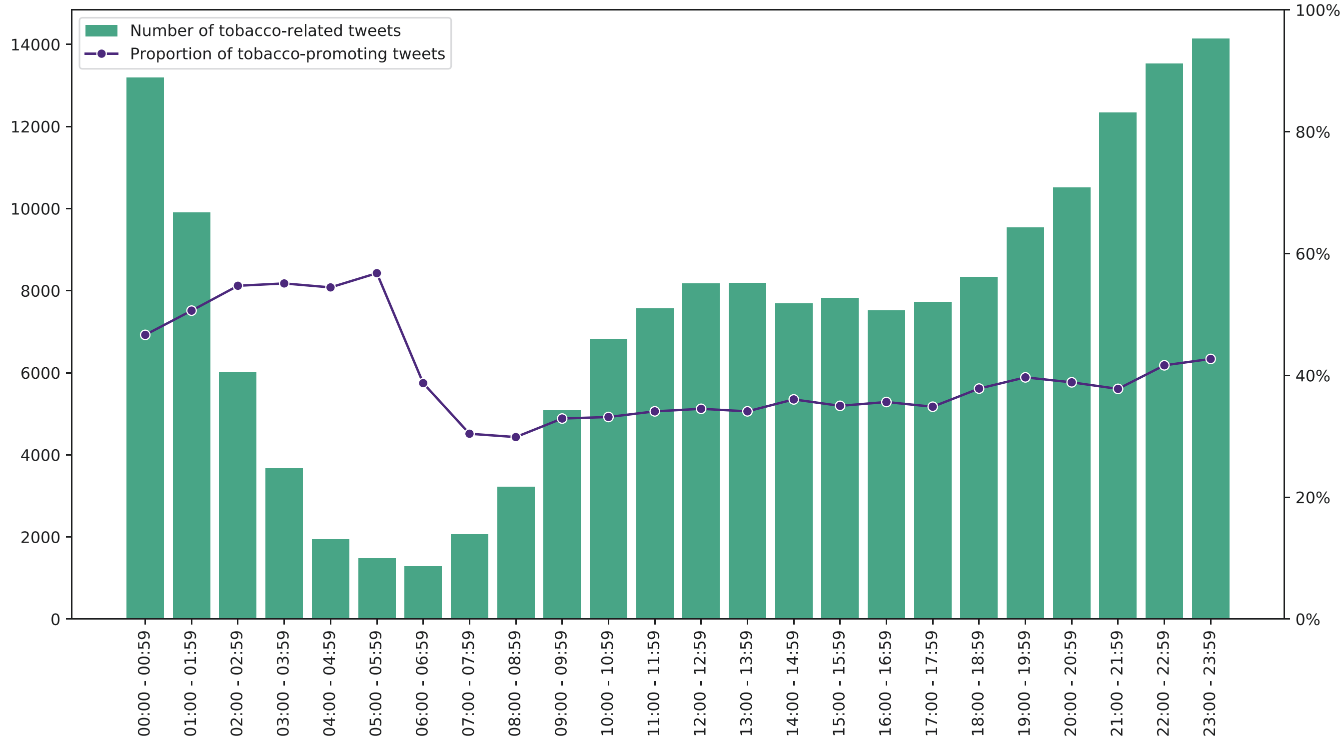

Figure 2. Distribution of tobacco-related and tobacco-promoting tweets by hour of the day.

A Labeled Dataset on Tobacco Promotion in Social Media Was Created

A dataset of 4498 tweets with labels on tobacco promotion and four relevant themes, was published open access. Future studies can utilize the dataset to explore different aspects of the problem or to develop better models.

A Well-Performing Machine Learning Model Was Developed

The model was capable of predicting tobacco-promoting content on social media at high levels of sensitivity (recall) and positive predictive value (precision).

The Utility of the Model Was Demonstrated on Tweets Posted Within a Month

The distribution and some basic characteristics of tobacco-promoting tweets were presented. Tobacco-promoting tweets were published more at night, and with mentions of other users, the foods, and environments related to tobacco use.

Methodological Contributions

Identification of Exposure

Unlike information seeking, "information scanning" is when an individual is exposed to information incidentally and repetitively without direct intent.[8] The study by Bigsby and Hovick[28] indicated that information scanning plays a mediating role in increasing chronic disease risk factors such as smoking. A prospective study by Zhu[7] revealed that young adults exposed to tobacco-promoting social media content use tobacco at higher rates after 6 months and demonstrated its dose–response relationship. Similarly, Liu et al.[9] indicated the effect of information scanning on e-cigarette use among young people. However, exposure data in both studies were self-reported which has been a common limitation in information scanning studies.[8] The current study offers a tool that enables the measurement of exposure to tobacco-promoting social media content more objectively.

Previous social media studies identified some themes in tweets containing a tobacco-related keyword.[29] Themes on tobacco use behavior, attitude, and marketing in some studies[30,31] are consistent with the qualitative findings of this study. However, they failed to address the complicated presentation of themes. Also worth noting, studies analyzed tweets with "tobacco-related" keywords, without a distinction of whether they were "tobacco-promoting." The model developed enables such a distinction.

The thematic analysis in these studies was mostly carried out manually which limited the number of content to be analyzed. On the other hand, a large amount of content can be analyzed by machine learning methods.[29] Liu et al.[32] also stressed the shortfall of manual thematic analysis in the studies investigating conventional or social media content. The mixed methods employed in the current study enabled us to utilize the ability of manual qualitative analysis to capture nuances and the scalability of machine learning.

Machine learning methods were utilized in a few studies, but it was to identify themes such as relevance to tobacco in general,[30] advertisements,[33] smoking cessation,[34] or to detect the sentiment,[35] not to identify tobacco promotion. The current study went beyond identifying particular themes in social media and attempted to define tobacco-promoting content, which might be more useful for tobacco control purposes.

Although the qualitative analysis in this study was based on data in lieu of a theoretical framework, the inductively identified themes had parallels with behavioral theories.[7,9,36] Zhu[7] explained the effect of tobacco-promoting social media content based on the Integrative Model of Behavioral Prediction[37] which suggests that social media content affects the behavior of individuals (tobacco use) through the following mediators: descriptive norms (i.e., how popular tobacco use is), injunctive norms (i.e., what is said about tobacco use), attitude, and self-efficacy.[7,38] The four themes in this study "tobacco use behavior," "attitude toward tobacco," "recommendation for tobacco use," and "tobacco product marketing" in social media content may be linked with the descriptive norm belief, attitude, injunctive norm belief, and self-efficacy, respectively.

Prediction Performance

The models developed by Cole-Lewis et al.,[39] focusing only on e-cigarettes, predicted behavior, attitude, and marketing with accuracies of 84%, 76%, and 91%, respectively. Models developed in this study predicted behavior, attitude, and marketing with accuracies of 94%, 82.6%, and 99.2%, respectively. Utilizing larger resources, Liu et al.[32] predicted the tobacco use behavior in tweets with 71% sensitivity and 52% positive predictive value. A model in this study (model 19) predicted tobacco use behavior with 81.8% sensitivity and 81.8% positive predictive value. Superiority in prediction performance may be due to the utilization of the BERT algorithm, in contrast to relatively older and limited algorithms used in previous studies.[22] In addition, using a coding framework (Supplementary Table S2) that better represents the complicated nature of tobacco promotion on social media and training on data consistently coded by two reviewers might also improve the performance.

Implications on Tobacco Control Policy and Practice

Infoveillance

Global tobacco surveys showed an increasing trend in exposure to traditional media in Türkiye,[10] yet the exposure to tobacco-promoting social media content is unknown. This study predicted that on average 2360 tweets per day promoting tobacco were posted on Turkish Twitter. Developed models can be used to monitor the trends of exposure, objectively. This may enable the impact assessment of national interventions by before–after comparison.

Enforcement

Although some actions have been listed in the Tobacco Control Strategy in Türkiye[12] regarding tobacco promotion in social media, stipulated means fall short of the ambition. For instance, to prevent online advertising (action 3.1), the means was "accelerating the process of blocking access to websites."[12] To strengthen the monitoring system (actions 3.4 and 3.6), the strategy includes "establishing new departments" that will identify harmful content.[12] Such actions might be effective against ordinary websites selling tobacco products. Yet, it is neither feasible nor effective to perform the monitoring and identification by manual assessment, considering that the half-life of this content is as short as 6–18 min for Twitter and 78–90 min for Facebook.[40] Indeed, our freedom of information query (Supplementary Material S1) revealed that the goals for these actions have not been achieved. The model developed in this study can be integrated with the above-mentioned systems to help identify potential influencer marketing campaigns.

Health Promotion

While the content on traditional media is produced only by institutional agents, individuals also become content creators on social media. The study acknowledges that tobacco promotion on social media is not only by the tobacco industry accounts but also by individual users who may or may not affiliated with the industry. This study deliberately focuses on social media *content* instead of *users* and offers a tool to detect harmful exposure. By doing so, it creates opportunities not only for surveillance and enforcement but also for health promotion. Following Ottawa Charter domains,[41] we suggest several use cases:

1. *Developing Personal Skills:* This tool can help to implement more targeted health education for those who are exposed to tobacco promotion.[29] Also, social media platforms can show automatic warnings when a user attempts to post tobacco-promoting content, informing about the potential impact of such content, and developing responsible publishing skills and motivation.
2. *Creating Supportive Environments:* Individuals who do not want to see tobacco-promoting content (e.g., children or people who quit smoking recently) can use a plugin application that integrates our models to filter that content out while they are browsing social media. Furthermore, social media platforms themselves can integrate similar models into their systems (i) to offer an option for not showing tobacco-promoting content to users, (ii) to display an automatic warning message along with that content, (iii) to exclude that content from recommendation algorithms, and (iv) from monetization mechanisms.
3. *Strengthening Community Action:* On the flip side, anti-tobacco content can be supported or rewarded through retweets or likes by public health organizations and volunteers. In addition, this tool can provide quantitative evidence for advocacy groups to urge tobacco companies to stop marketing on social media and for social media companies to restrict tobacco advertising.[5]

Limitations and Future Directions

This study focused only on the text of the tweets since previous studies have already developed well-performing machine-learning models to identify tobacco products in images.[42] Future studies may achieve better predictive performance by adopting multimodal (text + image) approaches.[29] An in-depth examination of the false positives could provide valuable insights to further refine the proposed approach. In addition, the use of stratified random sampling in this study may not fully account for the time series nature of Twitter data. Future research should consider advanced sampling methods to ensure that the model generalizes better to the evolving nature of social media discussions. Future research could also benefit from examining the prevalence of tobacco-promoting material disseminated by bots versus genuine user accounts.

The main purpose of this study was to demonstrate the application of AI technologies in a public health issue—tobacco control. Hence, we adopted the state-of-the-art method available during the study period. Due to the lack of benchmarks for this specific task, study findings are compared with the most closely related studies available. Given the rapidly evolving AI field, new and better-performing LLM models have already become available. Future studies should test and compare different model architectures. Our work, with its detailed methodology and open data sharing, can serve as a benchmark for such studies. Whether such sophisticated models are necessary for this task is another open research question, particularly considering the environmental implications of LLMs.

We acknowledge the recent changes in Twitter's data access policy after its rebranding as X made conducting similar studies very difficult and expensive.[43] This further highlights the value of the data made available through our study, particularly for model development. Even so, implementing these models may require more innovative approaches. Efforts like those by the European Commission[44] to restore research access could alleviate these challenges in the future. Nevertheless, social media platforms themselves have the capacity and, arguably, the responsibility to implement such models internally, as they did in the face of COVID-19 misinformation.[45] Future research should also explore the adaptability and transferability of prediction models to other platforms with more accessible data policies.

Conclusion

We developed an artificial intelligence tool that can automatically identify tobacco-promoting content on social media. This tool enables future studies on tobacco control, information scanning, and behavioral change to assess the exposure to tobacco-promoting social media content more objectively and at scale. The developed machine learning model outperforms similar models in the literature. Recommendations are presented to facilitate the implementation of this tool in tobacco control policies and practices, specifically in the infoveillence of tobacco-promoting content, enforcement of existing bans, and health promotion activities. The data and the model in the study are published open access.[46,47]

Supplementary Material

Supplementary material is available at *Nicotine and Tobacco Research* online.

Funding

This study is supported by the Turkish Green Crescent Society (No: 2019/7). Volunteers who worked in data labeling were recruited from their university network. Except that, it did not have any role in the design of the study, collection, and analysis of data, or interpretation of findings.

Declaration of Interests

None declared.

Acknowledgments

We thank volunteers from Istanbul Medipol University Young Green Crescent Club for their contributions to data labeling, Gülşen Eryiğit and the ITU-NLP team for granting us access to their tools for text normalization, and Sabanur Çavdar for proof-reading the final draft.

Author Contributions

Hüseyin Küçükali (Conceptualization, Data curation, Formal analysis, Funding acquisition, Investigation, Methodology, Project administration, Resources, Software, Visualization, Writing—original draft, Writing—review & editing [lead]) and Mehmet Erdoğan (Supervision [lead], Writing—review & editing [supporting])

Data Availability

The data underlying this article are available in The Open Science Framework, at doi:10.17605/OSF.IO/ZCQA6. The language model trained in this study is available in Hugging Face, at doi:10.57967/hf/3607.

Ethics Approval

This study is approved by Istanbul University—Cerrahpasa Non-Invasive Studies Ethical Committee (6 May 2020, No:2020-06).

References

1. World Health Organization. *MPOWER: A Policy Package to Reverse the Tobacco Epidemic*. World Health Organization; 2008. https://iris.who.int/bitstream/handle/10665/43888/9789241596633_eng.pdf. Accessed November 29, 2024.
2. Arklan Ü, Rençber H. Social media as the escape area for illegality. In: Doğan E, Geçğin E, eds. *Current Debates in Public Relations Cultural & Media Studies*. London, UK: IJOPEC Publication Limited; 2018:31–52.
3. Kirkham C. Exclusive: Philip Morris suspends social media campaign after Reuters exposes young 'Influencers'. Reuters Media. Published November 5, 2019. https://www.reuters.com/article/us-philipmorris-ecigs-instagram-exclusiv/exclusive-philip-morris-suspends-social-media-campaign-after-reuters-exposes-young-influencers-idUSKCN1SH02K. Accessed April 16, 2021.
4. Kaplan S. Big tobacco's global reach on social media. The New York Times. Published August 24, 2018. https://www.nytimes.com/2018/08/24/health/tobacco-social-media-smoking.html. Accessed April 17, 2021.
5. Rowell A. Big tobacco wants social media influencers to promote its products—Can the platforms stop it? The Conversation. Published January 23, 2020. https://theconversation.com/big-tobacco-wants-social-media-influencers-to-promote-its-products-can-the-platforms-stop-it-129957. Accessed April 17, 2021.
6. Davies R, Chapman M. Tobacco giant bets £1bn on influencers to boost 'More Lung-Friendly' Sales. The Guardian. Published February 20, 2021. https://www.theguardian.com/business/2021/feb/20/tobacco-giant-bets-1bn-on-social-media-influencers-to-boost-lung-friendlier-sales. Accessed April 17, 2021.
7. Zhu Y. Pro-smoking information scanning using social media predicts young adults' smoking behavior. *Comput Hum Behav.* 2017;77:19–24. doi:10.1016/j.chb.2017.08.004
8. Lewis N. Information seeking and scanning. In: Rössler P, ed. *The International Encyclopedia of Media Effects*. American Cancer Society. Wiley-Blackwell; 2017. doi:10.1002/9781118783764.wbieme0156
9. Liu J, Lochbuehler K, Yang Q, Gibson LA, Hornik RC. Breadth of media scanning leads to vaping among youth and young adults: evidence of direct and indirect pathways from a national longitudinal survey. *J Health Commun.* 2020;25(2):91–104. doi:10.1080/10810730.2019.1709925
10. Centers for Disease Control and Prevention. Global Tobacco Surveillance System Data. Published December 2, 2019. https://www.cdc.gov/tobacco/global/gtss/gtssdata/index.html. Accessed April 15, 2021.
11. Law No. 4207. Tütün Ürünlerinin Zararlarının Önlenmesi ve Kontrolü Hakkında Kanun [Law on Prevention and Control of Harmful Effects of Tobacco Products]. *Republic of Türkiye*. 1996.

https://www.mevzuat.gov.tr/mevzuat?MevzuatNo=4207&MevzuatTur=1&MevzuatTertip=5. Accessed November 22, 2024.
12. Ministry of Health. Tütün Kontrolü Strateji Belgesi ve Eylem Planı 2018-2023 *[Tobacco Control Strategy Document and Action Plan 2018-2023]*. 2018. https://hsgm.saglik.gov.tr/depo/birimler/tutun-mucadele-bagimlilik-db/haberler/tutun_eylem_plani/Tutun_Kontrolu_Strateji_Belgesi_ve_Eylem_Plani.pdf. Accessed April 14, 2021.
13. Wolf T, Debut L, Sanh V, *et al*. Transformers: state-of-the-art natural language processing. In: Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing: System Demonstrations. Association for Computational Linguistics; Online. 2020:38–45. doi:10.18653/v1/2020.emnlp-demos.6
14. Kahn JAI. Breakthroughs in Natural-Language Processing are Big for Business. https://fortune.com/2020/01/20/natural-language-processing-business. Accessed April 24, 2021.
15. Küçükali H. Halk Sağlığında Yapay Zeka. *SD (Sağlık Düşüncesi ve Tıp Kültürü) Dergisi*. 2021;58:92–95. https://sdplatform.com/halk-sagliginda-yapay-zeka. Accessed November 22, 2024.
16. Küçükali H. *Artificial Intelligence in Public Health: In Case of Tobacco Addiction* [*PhD*]. Istanbul University—Cerrahpasa; 2021. https://tez.yok.gov.tr/UlusalTezMerkezi/TezGoster?key=v7BkNnnepTnbhn8rNR77LT3cEzbT8m6XL7nKUsfts4iPLhWZA2sgG_tqKk2AqwEa
17. Schoonenboom J, Johnson RB. How to construct a mixed methods research design. *Kolner Z Soz Sozpsychol*. 2017;69(Suppl 2):107–131. doi:10.1007/s11577-017-0454-1
18. Azungah T. Qualitative research: deductive and inductive approaches to data analysis. *Qual Res J*. 2018;18(4):383–400. doi:10.1108/QRJ-D-18-00035
19. Twitter. Standard API v1.1. Published 2021. https://developer.twitter.com/en/docs/twitter-api/v1. Accessed April 3, 2021.
20. Kim Y, Huang J, Emery S. Garbage in, garbage out: data collection, quality assessment and reporting standards for social media data use in health research, infodemiology and digital disease detection. *J Med Internet Res*. 2016;18(2):e41. doi:10.2196/jmir.4738
21. Grand View Research. E-Cigarette and Vape Market Size & Share Report. https://www.grandviewresearch.com/industry-analysis/e-cigarette-vaping-market. Accessed October 11, 2020.
22. Küçükali H, Küçükali İ, Rakkad N, Rakkad A. Using artificial intelligence to identify tobacco-related contents in social media: a systematic scoping review of the literature. 4th International 22nd National Public Health Congress; December 18, 2020; Online, Türkiye. https://uhsk.org/2020/kongre-kitabi.pdf
23. Eryiğit G, Torunoğlu Selamet D. Social media text normalization for Turkish. *Nat Lang Eng*. 2017;23(6):835–875. doi:10.1017/S1351324917000134
24. Devlin J, Chang M-W, Lee K, Toutanova K. *BERT: Pre-Training of Deep Bidirectional Transformers for Language Understanding*. 2018. http://arxiv.org/pdf/1810.04805v2
25. Schweter S. *BERTurk - BERT Models for Turkish*. Zenodo. 2020. doi:10.5281/zenodo.3770924
26. Oğuz Ö. Türkçe Küfür Karalistesi. Published April 3, 2021. https://github.com/ooguz/turkce-kufur-karaliste. Accessed April 3, 2021.
27. Aksoy A. Türkçe Dolgu Sözcükleri. https://github.com/ahmetax/trstop. Accessed April 24, 2021.
28. Bigsby E, Hovick SR. Understanding associations between information seeking and scanning and health risk behaviors: an early test of the structural influence model. *Health Commun*. 2018;33(3):315–325. doi:10.1080/10410236.2016.1266575
29. Lienemann BA, Unger JB, Cruz TB, Chu K-H. Methods for coding tobacco-related twitter data: a systematic review. *J Med Internet Res*. 2017;19(3):e91. doi:10.2196/jmir.7022
30. Myslín M, Zhu S-H, Chapman W, Conway M. Using twitter to examine smoking behavior and perceptions of emerging tobacco products. *J Med Internet Res*. 2013;15(8):e174. doi:10.2196/jmir.2534
31. van der Tempel J, Noormohamed A, Schwartz R, *et al*. Vape, quit, tweet? Electronic cigarettes and smoking cessation on Twitter. *Int J Public Health*. 2016;61(2):249–256. doi:10.1007/s00038-016-0791-2
32. Liu J, Siegel L, Gibson LA, *et al*. Toward an aggregate, implicit, and dynamic model of norm formation: capturing large-scale media representations of dynamic descriptive norms through automated and crowdsourced content analysis. *J Commun*. 2019;69(6):563–588. doi:10.1093/joc/jqz033
33. Clark EM, Jones CA, Williams JR, *et al*. Vaporous marketing: uncovering pervasive electronic cigarette advertisements on Twitter. *PLoS One*. 2016;11(7):e0157304. doi:10.1371/journal.pone.0157304
34. Aphinyanaphongs Y, Lulejian A, Brown DP, Bonneau R, Krebs P. Text classification for automatic detection of e-cigarette use and use for smoking cessation from twitter: a feasibility pilot. *Pac Symp Biocomput*. 2016;21:480–491. doi:10.1142/9789814749411_0044
35. Allem J-P, Ramanujam J, Lerman K, *et al*. Identifying sentiment of hookah-related posts on Twitter. *JMIR Public Health Surveill*. 2017;3(4):e74. doi:10.2196/publichealth.8133
36. Yoo W, Yang J, Cho E. How social media influence college students' smoking attitudes and intentions. *Comput Human Behav*. 2016;64:173–182. doi:10.1016/j.chb.2016.06.061
37. Fishbein M, Yzer MC. Using theory to design effective health behavior interventions. *Commun Theory*. 2003;13(2):164–183. doi:10.1111/j.1468-2885.2003.tb00287.x
38. Michie S, West R, Campbell R, Brown J, Gainforth H. *ABC of Behaviour Change Theories*. Great Britain: Silverback Publishing; 2014.
39. Cole-Lewis H, Varghese A, Sanders A, *et al*. Assessing electronic cigarette-related tweets for sentiment and content using supervised machine learning. *J Med Internet Res*. 2015;17(8):e208. doi:10.2196/jmir.4392
40. SocialTalent. The Short Half-Life of Your Social Recruiting Posts. Published September 26, 2017. https://www.socialtalent.com/blog/recruitment/the-short-half-life-of-your-social-media-posts. Accessed April 20, 2021.
41. World Health Organization. Ottawa Charter for Health Promotion. https://www.euro.who.int/__data/assets/pdf_file/0004/129532/Ottawa_Charter.pdf. Accessed April 20, 2021.
42. Kharade A, Abhishek K, Dwibedi D, *et al*. Image analytics to detect cigarette in an image using deep learning. In: Merchant SN, Warhade K, Adhikari D, eds. *Advances in Signal and Data Processing: Select Proceedings of ICSDP 2019*. Singapore: Springer; 2021:659–678.
43. Kupferschmidt K. Twitter's plan to cut off free data access evokes 'Fair Amount of Panic' among scientists. *ScienceInsider*. Published February 8, 2024. doi:10.1126/science.adh0813
44. European Commission. *Commission Sends Preliminary Findings to X for Breach of the Digital Services Act*. 12.07.2024. https://ec.europa.eu/commission/presscorner/detail/en/ip_24_3761. Accessed November 22, 2024.
45. Roth Y, Pickles N. Updating our approach to misleading information. Published October 23, 2022. https://blog.x.com/en_us/topics/product/2020/updating-our-approach-to-misleading-information. Accessed July 22, 2024.
46. Küçükali H. *Berturk-tobacco-control (Revision 9880aa5)*. Hugging Face; 2024. doi:10.57967/hf/3607
47. Küçükali H. *AI for Tobacco Control*. OSF; 2024. doi:10.17605/OSF.IO/ZCQA6