ISSN: 2676-4245 Home Page: Ijsp.ut.ac.ir

The University Of Tehran Press

Journal of Social Problems of Iran

The Moral Marketplace of Twitter: Moral Preferences and Political Behavior – A Deep Learning Approach to Big Data from Iran's 2017 Presidential Election

Alireza Haddadi[1]✉

1. Corresponding Author, Assistant Professor, Department of Sociology, Faculty of Social Science, Allameh Tabataba'i University, Tehran, Iran. E-mail: arhaddadi@atu.ac.ir

Article Info

Article type:

Research Article

Article history:

Received: 25 October 2025

Received in revised form: 8 February 2026

Accepted: 7 June 2026

Published online: 23 July 2026

Keywords:

Computational Social Science,

Moral Foundations,

Political Behavior,

Natural Language Processing.

Deep Learning,

Social Network Analysis.

ABSTRACT

Human decision-making, shaped by a confluence of internal dispositions and contextual factors, plays a fundamental role in the formation of social actions and political behaviors. Understanding the moral preferences underlying such decisions is essential for elucidating the mechanisms that govern social order. Adopting an interdisciplinary approach, this study leverages Twitter big data to investigate moral preferences and their relationship with individuals' political orientations. The Moral Foundations framework was operationalized and validated through deep learning models applied to a large-scale dataset of tweets. Our findings identify key constructs—including group cohesion, attention-seeking tendencies, distrust, value orientation, accountability, and cognition—as central to the structuring of political behavior. These constructs exhibit distinct patterns across different political groups (i.e., Reformists, Principlists, Antagonists, and Apolitical users).

Results indicate that user behavior is shaped less by formal factional labels and more by distinctive moral preference profiles. An ethics of responsibility and trust, along with group dynamics favoring conformity, underpinned the stability of political camps. Notably, Reformists and Principlists, despite their discursive opposition, display a similar "moral signature," whereas Antagonists present a distinctive profile characterized by the valorization of power, pronounced distrust, and resistance to stasis. Apolitical users, in contrast, predominantly orient toward close interpersonal relations and micro-identities. Performative solidarity and adversarial discourse among factions, symbolic boundary-making by Principlists, algorithmically driven activism by Antagonists, and the ironic distancing of Apolitical users all reflect adaptive strategies within Twitter's "moral marketplace"—a space where traditional solidarities have fragmented into competing sub-communities.

Cite this article: Haddadi, A. (2026). The Moral Marketplace of Twitter: Moral Preferences and Political Behavior – A Deep Learning Approach to Big Data from Iran's 2017 Presidential Election. *Journal of Social Problems of Iran*, *17*(1), 61-100. DOI: https://doi.org/10.22059/IJSP.2026.405026.671354

 Publisher: The University of Tehran Press.
DOI: https://doi.org/10.22059/IJSP.2026.405026.671354

Introduction

Political behavior is rarely a product of purely rational calculation; rather, it emerges from a complex interplay of internal dispositions, cultural values, and contextual social dynamics. Understanding the moral preferences underlying political decisions is therefore fundamental to explaining social cohesion, ideological cleavages, and collective identity formation. Moral Foundations Theory (MFT) has offered a robust framework for linking intuitive moral judgments to political ideology, yet most empirical work remains confined to survey data and English-language corpora, leaving non-Western digital public spheres largely unexplored.

Social media platforms—particularly Twitter—have transformed political discourse, especially during high-stakes national events. In Iran, Twitter serves as a critical arena for political contestation. However, existing research on Persian Twitter has predominantly focused on network topology, sentiment polarity, or candidate mentions, often neglecting the deeper ethical and psychological drivers of user engagement. This study addresses this gap by adopting an interdisciplinary computational social science approach. Using a large-scale dataset of Persian tweets from the 2017 Iranian presidential election, we pursue three objectives: (1) to operationalize and validate the Eight Bipolar Moral Preferences (EBMP) framework—a culturally grounded moral model—via deep learning; (2) to estimate users' political orientations through social network analysis; and (3) to systematically examine the relationship between expressed moral preferences and political alignment. This paper thus contributes to moral psychology, non-English NLP, and the empirical understanding of political behavior in contemporary Iran.

Methodology

Our mixed-methods design integrates qualitative content analysis, computational network analysis, and supervised deep learning.

Data Collection. We collected 2,308,325 Persian tweets related to the 2017 presidential election over a one-month period surrounding the election date. Data were cleaned by removing irrelevant characters, advertisements, and extra symbols, with orthographic normalization applied to all texts.

Moral Preference Framework and Annotation. The EBMP framework, developed through prior grounded-theory research, comprises eight bipolar dimensions: Group (+/−), Cognition (+/−), Value (+/−), Trust (+/−), Responsibility (+/−), Attention (+/−), Power (+/−), and Inertia (+/−). A random sample of 10,000 tweets was manually annotated by four interdisciplinary experts (sociology, anthropology, political science, and psychology) using a standardized protocol, achieving an inter-coder agreement of 71.68%. To scale annotation, we employed a semi-supervised word-expansion approach. TF-IDF extracted the top 100 distinctive terms per preference; a custom lexicon combined 80% election-period tweets with 20% established psychological dictionaries. Automated labeling occurred when a tweet contained at least two lexicon terms. This expanded the labeled dataset to 120,000 posts, with expert validation on a random 500-sample confirming >89% accuracy.

Deep Learning Models. For each of the eight preferences, we trained separate neural networks using GloVe and FastText 50-dimensional embeddings. Three architectures were tested: LSTM, Transformer, and BERT (with ParsBERT for Persian optimization). Training used the Adam optimizer, categorical cross-entropy loss, and 5-fold cross-validation. The final model achieved a mean accuracy of 68.77% against human validation, which—though slightly below human inter-coder agreement—was deemed sufficient for large-scale classification.

Political Orientation Estimation. To infer users' political leanings, we constructed a multi-layered retweet and follower network. Starting with 4,000 highly active users, we crawled follower/following relationships to build a second-layer graph (32,846 nodes, 3.75M edges) and a third-layer graph (36,224 users, 15.8M edges). Community detection used the Blondel algorithm. Four distinct political clusters were labeled via qualitative content analysis of users' last 100 tweets and biographies: Reformists, Principlists (conservatives), Antagonists (regime opponents), and Apolitical users. Inter-coder reliability for community labeling was 0.77 (Krippendorff's alpha).

Statistical Association. Pearson correlation and regression analyses quantified the relationship between moral preference frequencies (per user) and political group membership, controlling for activity level and network centrality.

Findings

The results reveal a structured moral landscape with clear group-specific patterns.

Overall Moral Preferences. Across the entire dataset, Group+ (23.4%) and Attention+ (21.1%) were most frequent, followed by Value+ (13.2%), Responsibility+ (11.8%), and Trust− (10.5%). Power−, Inertia+, and Attention− were least common. This indicates that users generally prioritize belonging, social recognition, value-commitment, and accountability over radical anti-authoritarianism or passive conservatism.

Reformists and Principlists—A Shared Moral Core. Despite their political rivalry, Reformists and Principlists exhibited remarkably similar "moral signatures." Both groups ranked Group+ and Attention+ highest, followed by Trust−, Value+, and Responsibility+. This overlap suggests that both factions experience politics primarily as a field of collective belonging and recognition within the existing order, with moral sensitivity oriented toward loyalty, care, and interactional fairness. Their differences are largely programmatic and discursive, not foundational.

Antagonists—A Distinctive Profile. Regime opponents displayed a sharply divergent pattern. Power+ was dominant, followed by Attention+, Trust−, and Inertia−. Their discourse centers on authority critique, institutional distrust, and radical change. A higher Group− score also indicates active boundary-making against ruling elites. This aligns with oppositional discourse studies, where moral language organizes around power-resistance and systemic transformation.

Apolitical Users—Micro-Social Morality. Apolitical users centered on Attention+ and Group+, but with significantly lower Trust−, Power+, and Inertia− scores compared to political groups. Their moral concerns are micro-social—interpersonal relationships, emotional expression, and personal identity. A slightly higher Value− score suggests skepticism toward grand ideological narratives, consistent with cultural detachment and a post-ideological orientation.

Beyond Factional Labels. Regression confirmed that moral preferences predicted political orientation more strongly than retweet frequency or network centrality, supporting our core hypothesis that behavior is shaped more by distinct moral preference patterns than official labels alone. Performative solidarity, symbolic boundary-making, algorithmic activism, and ironic distancing all appear as adaptive strategies within Twitter's "moral marketplace," where traditional solidarities have fragmented into competing sub-publics.

Discussion and Conclusion

This study makes three key contributions. **Theoretically**, it validates the EBMP framework as a culturally sensitive extension of MFT. While universal moral foundations (care, fairness, loyalty) remain active, their weighting is significantly modulated by Iran's socio-political context. The moral overlap between Reformists and Principlists challenges assumptions of deep ethical polarization, while the Antagonists' profile highlights how oppositional identities are morally constructed around power-resistance and radical distrust.

Methodologically, we demonstrate that deep learning models, combined with semi-supervised expansion and custom lexicons, achieve acceptable classification accuracy for Persian conversational text—a challenging NLP task. The integration of network analysis and text classification provides a replicable template for non-Western digital contexts.

Empirically, political behavior in Iran's Twittersphere reflects deeper moral motivations, not merely partisan loyalty. The "moral marketplace" metaphor captures how users adapt their moral rhetoric to both audience expectations and platform algorithms, forming stable yet competing moral communities.

Limitations and Future Directions. This study is limited to one election cycle and one platform. Future work should examine temporal variations across multiple elections, incorporate qualitative interviews and surveys for cross-validation, and employ large language models (LLMs) and generative AI for richer contextual understanding. From a policy perspective, we recommend that research funding prioritize the systematic study of moral and value preferences over purely political labels, as these deeper ethical layers offer a more robust analytical basis for understanding polarization, social cleavages, and the potential for constructive dialogue in divided societies.

شاپا الکترونیکی: 2676-4245 سامانه نشریه: Ijsp.ut.ac.ir

بررسی مسائل اجتماعی ایران

انتشارات دانشگاه تهران

بازار اخلاقی توییتر: رابطه ترجیحات اخلاقی و رفتار سیاسی
کاربرد مدل یادگیری عمیق در کلان‌داده انتخابات ریاست‌جمهوری ۱۳۹۶

علیرضا حدادی [۱]✉

۱. نویسنده مسئول، استادیار گروه جامعه‌شناسی، دانشکده علوم اجتماعی، دانشگاه علامه طباطبایی، تهران، ایران. رایانامه: arhaddadi@atu.ac.ir

اطلاعات مقاله

نوع مقاله: مقالهٔ پژوهشی

تاریخ دریافت: ۱۴۰۴/۰۸/۰۳
تاریخ بازنگری: ۱۴۰۴/۱۱/۱۹
تاریخ پذیرش: ۱۴۰۵/۰۳/۱۷
تاریخ انتشار: ۱۴۰۵/۰۵/۰۱

کلیدواژه‌ها:
علوم اجتماعی محاسباتی، بنیادهای اخلاقی، رفتار سیاسی، پردازش زبان طبیعی، یادگیری عمیق، تحلیل شبکه‌های اجتماعی.

چکیده

فرایند تصمیم‌گیری انسانی، که متأثر از عوامل درونی و زمینه‌ای است، نقشی بنیادین در شکل‌گیری کنش‌های اجتماعی و رفتارهای سیاسی ایفا می‌نماید. فهم ترجیحات اخلاقیِ زیربنای این تصمیمات، برای تبیین سازوکارهای حاکم بر امر اجتماعی، ضروری است. مقاله حاضر، با اتخاذ رویکردی میان‌رشته‌ای، از کلان‌داده توییتر بهره می‌گیرد تا به بررسی ترجیحات اخلاقی و نسبت آن با جهت‌گیری سیاسی افراد بپردازد. چارچوب بنیادهای اخلاقی با بهره‌گیری از مدل‌های یادگیری عمیق بر روی مجموعه داده‌ای گسترده از توییت‌ها، تحلیل و تأیید شدند. یافته‌های تحقیق، مفاهیمی چون گروه‌گرایی، تمایل به جلب توجه، بی‌اعتمادی، جهت‌گیری ارزشی، مسئولیت‌پذیری و شناخت را به‌عنوان عناصر کلیدی در شکل‌گیری رفتار سیاسی برجسته می‌سازد. این مفاهیم در میان گروه‌های مختلف (اعم از اصلاح‌طلب، اصول‌گرا، معاند و غیرسیاسی) الگوهای متمایزی از خود بروز می‌دهند.

یافته‌ها نشان می‌دهد که رفتار کاربران بیش از آن‌که صرفاً تابع برچسب‌های رسمیِ جناحی باشد، از الگوهای متمایز ترجیح اخلاقی تأثیر می‌پذیرد. اخلاق مسئولیت‌پذیری و اعتماد، همراه با پویایی‌های گروهیِ متمایل به همرنگی، زمینه‌ساز ثبات اردوگاه‌های سیاسی بوده‌اند. اصلاح‌طلبان و اصول‌گرایان، علی‌رغم تعارض گفتمانی، «امضای اخلاقی» مشابهی دارند، در حالی‌که معاندان پروفایلی متمایز مبتنی بر برجسته‌سازی قدرت، بی‌اعتمادی و مخالفت با سکون نشان می‌دهند و کاربران غیرسیاسی بیش از همه بر روابط نزدیک و هویت‌های خرد تمرکز دارند. همبستگی نمایشی و گفت‌وگوی استدلالی جناح‌ها، مرزگذاری نمادین اصول‌گرایان، کنش‌گری الگوریتمی معاندان و فاصله‌گذاری طنازانه کاربران غیرسیاسی، همگی بازتابی از سازوکارهای تطبیقی با «بازار اخلاقی» توییتر هستند که همبستگی‌های سنتی را به خُرده‌جوامع رقیب تجزیه کرده‌اند.

استناد: حدادی، علیرضا (۱۴۰۵). بازار اخلاقی توییتر: رابطه ترجیحات اخلاقی و رفتار سیاسی کاربرد مدل یادگیری عمیق در کلان‌داده انتخابات ریاست‌جمهوری ۱۳۹۶. *بررسی مسائل اجتماعی ایران*، *۱۷*(۱)، ۶۱-۱۰۰. DOI: https://doi.org/10.22059/IJSP.2026.405026.671354

DOI: https://doi.org/10.22059/IJSP.2026.405026.671354

ناشر: مؤسسهٔ انتشارات دانشگاه تهران.

۱. مقدمه

فهم سازوکارهایی که کنش سیاسی شهروندان را هدایت می‌کنند، یکی از مباحث محوری در ادبیات رفتار رأی‌دهی و تصمیم‌گیری سیاسی است. پژوهش‌های اخیر نشان می‌دهند که تبیین رفتارهای سیاسی تنها با اتکا به گرایش‌های حزبی یا محاسبات عقلانی کافی نیست؛ زیرا این رفتارها در لایه‌های عمیق‌تری از جمله ترجیحات اخلاقی، نظام ارزش‌ها و انگیزه‌های روان‌شناختی ریشه دارند. از این منظر، انتخاب سیاسی فرد بازتابی از الگوهای اخلاقیِ تثبیت‌شده در ذهن و تجربه زیستی اوست و تحلیل رفتار سیاسی بدون توجه به این لایه بنیادین، فهمی ناقص از فرایند تصمیم‌گیری فراهم می‌آورد.

اخلاق، به‌مثابه سازوکاری شکل‌دهنده به کنش اجتماعی، حاصل تعامل پیچیده عوامل زیستی، روان‌شناختی، فرهنگی و نهادی است(Graham et al., 2011: 366; Haidt, 2012: 67). ترجیحات اخلاقی هم از الگوهای جهان‌شمول و هم از بافت‌های خاص فرهنگی منشأ می‌گیرند (Haidt & Joseph, 2004:55) و در تلفیق واکنش‌های شهودی و فرایندهای انعکاسی شکل می‌یابند (Greene, 2003:846). بر این اساس، بررسی پیوند میان ترجیحات اخلاقی و کنش سیاسی می‌تواند امکان تبیین دقیق‌تر تأثیر انگیزه‌های ارزش‌محور بر رفتار رأی‌دهی را فراهم کند.

با وجود این اهمیت نظری، روش‌های کلاسیک همچون پیمایش‌ها در آشکارسازی انگیزه‌های عمیق و واکنش‌های آنی که تصمیم‌گیری سیاسی را هدایت می‌کنند، همواره کارآمد نیستند. به‌ویژه در زمینه‌ای که کنشگری سیاسی کاربران به‌شدت دیجیتالی شده و بروز انگیزه‌ها بیش از هر زمان دیگر در بسترهای آنلاین رخ می‌دهد. داده‌های رفتاریِ تولیدشده در رسانه‌های اجتماعی، نظیر توییتر، با فراهم‌ساختن دسترسی به بازنمایی‌های زبانی، عاطفی و تعاملی کاربران، شرایطی کم‌نظیر برای مطالعه الگوهای تصمیم‌گیری فراهم می‌کنند(Tufekci & Wilson, 2012:363; Conover et al., 2011: 89). همراه با این امکانات، پیشرفت‌های حوزه یادگیری ماشین و پردازش زبان طبیعی، ابزارهایی در اختیار پژوهشگران قرار داده‌اند که تحلیل کلان‌داده‌ها و شناسایی الگوهای پنهان در رفتار سیاسی را ممکن می‌سازند(Bollen et al., 2011: 11; Nip & Berthelier, 2024: 1590).

در این چارچوب، بهره‌گیری از داده‌های شبکه‌های اجتماعی امکان ارزیابی تجربی نظریه‌ها و مدل‌های رفتار سیاسی را در مقیاسی فراهم می‌سازد که در مطالعات سنتی کمتر قابل دستیابی بوده است. بر همین اساس، مقاله حاضر به‌جای تمرکز بر یک رویداد خاص، از انتخابات ریاست‌جمهوری ۱۳۹۶ تنها به‌عنوان بستر تجربی برای پیاده‌سازی و آزمون چارچوب نظری بنیادهای اخلاقی

هشت‌گانه دوقطبی (EBMP) که در پژوهش پیشین (Haddadi et al., 2023: 210) معرفی شده است، استفاده می‌کند. هدف اصلی آن است که بررسی شود آیا در محتوای تولیدشده توسط کاربران فارسی‌زبان توییتر، نشانه‌های قابل‌سنجش و قابل‌طبقه‌بندی از مؤلفه‌های این چارچوب نظری وجود دارد یا خیر. در ادامه، مطالعه می‌کوشد نشان دهد که کاربران متعلق به گرایش‌های سیاسی متفاوت تا چه حد و از کدام یک از ترجیحات اخلاقی برای تصمیم‌گیری سیاسی خود بهره می‌گیرند. از این رهگذر، پژوهش حاضر تلاش می‌کند با ترکیب تحلیل کلان‌داده‌های اجتماعی، نظریه بنیادهای اخلاقی و مدلEBMP، گامی در جهت فهم دقیق‌تر نقش انگیزه‌های اخلاقی در شکل‌دهی رفتار رأی‌دهی در زیست‌جهان دیجیتال بردارد. روشی محاسباتی برای تحلیل محتوا و تخمین گرایش سیاسی در شبکه‌های اجتماعی ارائه دهد و با استفاده از آن، ظهور رفتار سیاسی و جناحی را از منظری اخلاقی تبیین نماید.

۲. پیشینهٔ پژوهش

مطالعات روان‌شناختیِ تصمیم‌گیری اغلب بر دو پارادایم توصیفی (مبتنی بر مشاهده تجربی) و هنجاری (متمرکز بر اصول کلی و عقلانی) تقسیم‌بندی می‌شوند (Frankena, 2013: 25). با این حال، نظریه‌های هنجاری معمولاً نسبت به تغییرپذیری داده‌های تجربی کم‌توجه هستند (Roy, 2016: 23–46). در این پژوهش، اخلاق به‌عنوان کیفیتی درونی در انسان در نظر گرفته شده و در چارچوب پارادایم توصیفی قرار می‌گیرد. در این راستا، انتخاب‌های سیاسی بیش از آنکه صرفاً عقلانی باشند، تحت تأثیر احساسات، باورها و سنت‌ها هستند(Smith, 2017: 56).

در حوزه روان‌شناسی اخلاق، مدل‌های مختلفی مانند شهودگرایی اجتماعی و نظریه فرآیند دوگانه به نقش احساسات و تعامل شهود و عقلانیت پرداخته‌اند(Greene, 2003: 845). همچنین، سایمون (1972) با رد عقلانیت کامل، نظریه تصمیم‌گیری رضایت‌بخش را مطرح کرده است. در این میان، نظریه بنیادهای اخلاقی[1] هایت از جایگاه ویژه‌ای برخوردار است که بر پایه شهودهای پیشاعقلانی در شکل‌گیری قضاوت‌های اخلاقی استوار است و تفاوت‌های فردی و فرهنگی را ناشی از تأکید متفاوت بر بنیان‌های تکاملی مشترک می‌داند. این نظریه در ابتدا مبتنی بر پنج بنیان مراقبت، انصاف، وفاداری، اقتدار و تقدس ارائه شد (Haidt & Graham, 2007: 98; Haidt & Joseph, 2004: 150-174).

1 Moral Foundations Theory (MFT)

در بازنگری سال ۲۰۲۳، این نظریه بازصورت‌بندی شده و بنیان انصاف/معاوضه به دو بخش متمایز برابری و تناسب تقسیم شد. همچنین بنیان‌های بالقوه جدیدی مانند آزادی/ضدسلطه، ناموس/شرافت و مالکیت به‌طور جدی مطرح شده‌اند(Atari et al., 2020: 367). بر این اساس، پرسشنامه‌های استانداردی برای سنجش شش بُعد مراقبت، برابری، تناسب، وفاداری، اقتدار و پاکی طراحی شده که در مطالعات میان‌فرهنگی مختلفی به کار رفته است. با این وجود، رویکرد هایت که هنجاری است، با رویکرد استقرایی نظریه‌هایی مانند EBMP که الگوهای اخلاقی را مستقیماً از داده‌ها استخراج می‌کنند، تفاوت دارد.

این دیدگاه‌ها مبنای حساسیت نظری و ساخت نظریه EBMP را شکل داده‌اند، اما چارچوبی تجویزی نداشته‌اند؛ بلکه به‌مثابه زمینه‌ای تفسیری برای تحلیل ابعاد اخلاقی در جامعه ایرانی به کار گرفته شده‌اند.

پژوهش‌های اخیر، بر تعامل میان ارزش‌های اخلاقی، زمینه اجتماعی و ابزارهای ارتباطی در شکل‌دهی به رفتار سیاسی و نتایج انتخاباتی تأکید دارند.

استرلینگ و جاست (۲۰۱۸) با تحلیل حجم بزرگی از توییت‌های آمریکایی نشان دادند که گرایش سیاسی با «امضای اخلاقی» زبان کاربران هم‌بسته است؛ لیبرال‌ها بیشتر از واژگان مرتبط با بنیان‌های مراقبت و انصاف استفاده می‌کنند، در حالی که محافظه‌کاران بر وفاداری، اقتدار و تقدس تأکید بیشتری دارند. همچنین سواد سیاسی کاربران نقش تعدیل‌کننده در الگوی استفاده از زبان اخلاقی دارد. روی، پاچکو و گلدواسر (۲۰۲۱) گام فراتر گذاشته و با استفاده از یادگیری رابطه‌ای، «فریم‌های اخلاقی» را در توییت‌های سیاسی تحلیل کردند. یافته‌ها نشان می‌دهد دموکرات‌ها و جمهوری‌خواهان نه تنها در استفاده از بنیادهای اخلاقی، بلکه در جهت‌گیری هیجانی نسبت به بازیگران مختلف نیز تفاوت سیستماتیک دارند. این مطالعه نشان می‌دهد ترکیب نظریه بنیادهای اخلاقی با مدل‌های پیشرفته یادگیری ماشین می‌تواند برای درک سازوکارهای معنایی و ایدئولوژیک در گفتمان سیاسی شبکه‌های اجتماعی به کار رود. رایترهاس، کوپینیک و لکس (۲۰۲۱) در پژوهش «بررسی تفاوت‌های مبتنی‌بر اخلاق در فریمینگِ توییت‌های سیاسی» به تحلیل فریمینگ اخلاقی در دو پیکره مختلف پرداختند: توییت‌های سیاستمداران آمریکایی و توییت‌های مرتبط با کووید-۱۹ به زبان آلمانی از اتریش. نتایج نشان داد در داده‌های آمریکایی، دموکرات‌ها تمایل بیشتری به فریم‌های مراقبت دارند، در حالی که جمهوری‌خواهان بر وفاداری و بنیادهای گروه‌محور تأکید می‌کنند. در نمونه اتریشی نیز پیروان حزب محافظه‌کار از فریم مراقبت همسو با پیام‌های رسمی کمپین‌هایشان استفاده می‌کردند.

همچنین موج جدیدی از پژوهش‌ها به نقد و بهبود روش‌های سنجش محتوای اخلاقی پرداخته است. استکر و هوپ (۲۰۲۵) در مقاله «عدم هم‌گرایی سنجه‌های خودکار در مانیفست‌های حزبی» به بررسی علت نتایج متناقض در پژوهش‌های مبتنی بر نظریه بنیادهای اخلاقی پرداخته‌اند. چنگ و هیل (۲۰۲۵) نیز در مقاله «سنجش بنیادهای اخلاقی در متن غیرانگلیسی» محدودیت‌های ابزارهای انگلیسی‌محور را در تحلیل متون چینی بررسی کرده و بر ضرورت اعتبارسنجی انسان‌محور و احتیاط در تعمیم خودکار برچسب‌های اخلاقی در بافت‌های فراملی و چندزبانه تأکید کرده‌اند.

توییتر در سال‌های اخیر به بستری کلیدی برای شکل‌دهی به گفتمان سیاسی و تأثیرگذاری بر افکار عمومی، به‌ویژه در کشورهایی چون ایران تبدیل شده است. با وجود رشد پژوهش‌ها درباره توییتر فارسی، اغلب مطالعات ازلحاظ جامعه‌شناختی، رویکردهایی محدود اتخاذ کرده‌اند. اگرچه برخی مطالعات شبکه‌ای به بررسی روابط کاربران و شکل‌گیری اجتماعات مجازی نیز پرداخته‌اند (Saleh & Haddadi, 2022: 245)، اما اغلب ابعاد روان‌شناختی و اخلاقیِ رفتار کاربران را نادیده گرفته‌اند.

در این میان برخی پژوهش‌ها محتوای انتخاباتی را از منظر تحلیل احساسات بررسی کرده‌اند؛ به‌عنوان نمونه، تعیین قطبیت احساسی توییت‌ها در دوره‌های انتخاباتی(Molla Norouzi, 2019: 140-145). خضرایی (۲۰۱۹) با تحلیل ساختاری تأثیرگذاری کاربران در انتخابات ۱۳۹۲، به بررسی شاخص‌هایی همچون تعداد بازتوییت‌ها و تأثیر گروه‌های سیاسی پرداخت؛ با این حال، تمرکز اصلی همچنان بر ذکر اسامی نامزدها و تجسم احساسات باقی ماند، آن‌هم در شرایطی که تنها ۱۵درصد از داده‌ها به زبان فارسی بود و دقت تحلیلی محدودی داشت. عبداللهیان و کرمانی (۱۳۹۹) در پژوهشی مبتنی بر تحلیل محتوا، انتخابات ۱۳۹۶ را با شناسایی خوشه‌های ارزشی در میان گروه‌های اصول‌گرا، اصلاح‌طلب و مهاجر بررسی نمودند. با این حال، خلأ قابل‌توجهی در بررسی ترجیحات اخلاقی و روانی کاربران ایرانی در بستر کلان‌داده‌ها و برقراری نسبت با رفتار و گرایش سیاسی مشهود است. پژوهش حاضر با تمرکز بر همین خلأ، به کاوش در این ابعاد مغفول‌مانده می‌پردازد. این مطالعه با هدف پرکردن خلأهای موجود در پژوهش‌های مرتبط با توییتر فارسی(Khazraee, 2019: 87) ، تقویت قابلیت‌های تحلیل زبان فارسی در فضای دیجیتال(Moniri et al., 2024)، و بومی‌سازی نظریه بنیادهای اخلاقی در بافت فرهنگی ایران (Atari et al., 2020:368) انجام شده است. این پژوهش از طریق تحلیل ترجیحات اخلاقیِ منعکس‌شده در توییت‌ها، چشم‌اندازی نوین درباره‌ی نقش اخلاق در رفتارهای سیاسی و شکل‌گیری افکار عمومی در زیست‌جهان دیجیتال معاصر ارائه می‌دهد.

۱،۲. چارچوب نظری

نویسنده در مقاله پیشین (Haddadi et al., 2023: 210)، با فراتحلیل پژوهش‌های مرتبط و تحلیل محتوای کیفی دادگان توییتر، مدلی نظری را تحت عنوان «هشت ترجیح اخلاقی دوقطبی» (EBMP) ارائه کرده‌است که متناسب با بافت فرهنگی و اجتماعی ایران شکل گرفته است.

برای ساخت چارچوب و کشف بنیادهای اخلاقی ایرانیان، از روش نمونه‌گیری چندگانه نظری (گلوله‌برفی) تا رسیدن به اشباع نظری استفاده شد. ابتدا ۱۵۰ ترجیح اخلاقی شناسایی شده از مطالعات پیشین، با تحلیل محتوای کیفی به ۳۵ ترجیح ادغام شدند. اعتبار صوری مفاهیم از طریق روش دلفی و اجماع هفت متخصص تأمین گردید. در نهایت، هشت ترجیح اخلاقی عام با مؤلفه‌های متعدد به عنوان چارچوب مفهومی اولیه تعیین شد.

در مرحله بعد، داده‌های توئیتر در بازه زمانی ۳۱ فروردین تا ۲۹ اردیبهشت ۱۳۹۶ جمع‌آوری شد، که شامل ۲،۳۰۸،۳۲۵ توییت مرتبط با انتخابات بود. از این مجموعه، نمونه‌ای تصادفی متشکل از ۱۰،۰۰۰ پست انتخاب شد. تحلیل داده‌ها با رویکرد استقرایی و با استفاده از نظریه داده‌بنیاد (Glaser & Strauss, 1967: 134) انجام گرفت. برای شناسایی متغیرهای روان‌شناختی و جامعه‌شناختی، از واژه‌نامه‌های مرتبط با نظریه بنیادهای اخلاقی (Graham et al., 2009: 1028) و ابزار LIWC (Pennebaker et al., 2015) استفاده شد. فرایند کدگذاری شامل کدگذاری باز برای شناسایی دسته‌های اولیه و کدگذاری محوری برای پالایش و ارتباط‌دهی این دسته‌ها با مضامین کلان‌تر اخلاقی بود. در نهایت، با برگزاری جلسات و کدگذاری کیفی مصاحبه‌های کارشناسان، ترجیحات اخلاقی نهایی مقوله‌بندی شدند.

جدول ۱. خلاصه چارچوب نظری ترجیحات اخلاقی در تصمیم‌گیری سیاسی (حدادی و همکاران، ۱۴۰۲)

ردیف	ترجیح	تعریف عملیاتی
۱	قدرت اجتماعی+	میل به قدرت‌طلبی و پرخاشگری
	قدرت اجتماعی−	میل به اطاعت‌پذیری
۲	اعتماد اجتماعی+	میل به اعتماد و ریسک‌پذیری
	اعتماد اجتماعی−	پرهیز از ارتباط و زیان‌گریز
۳	مسئولیت اجتماعی+	میل به قبول مسئولیت
	مسئولیت اجتماعی−	میل به مسئولیت‌ناپذیری
۴	توجه اجتماعی+	اهمیت‌دادن به ارزیابی دیگران و میل به تأییدپذیری
	توجه اجتماعی−	طردپذیری و بی‌توجه به نظر دیگران
۵	ارزش اجتماعی+	توجیه کنش با توجه به ارزش‌ها
	ارزش اجتماعی−	گریز از ارزش‌های انسانی

ردیف	ترجیح	تعریف عملیاتی
۶	شناخت اجتماعی+	تمایل به استدلال
	شناخت اجتماعی−	تمایل به کنشگری احساسی
۷	گروه اجتماعی+	وفادارای به گروه خودی
	گروه اجتماعی−	تمایل به عام‌گرایی و انصاف و برابری
۸	اینرسی اجتماعی+	میل به حفظ وضع موجود
	اینرسی اجتماعی−	میل به تغییر و پذیرش تجربه‌های جدید

چارچوب EBMP [۱] با تأکید بر پیوند میان ترجیحات اخلاقی و الگوهای بیان سیاسی، بستری نظری برای تحلیل داده‌های شبکه‌های اجتماعی فراهم می‌کند. این چارچوب بر این فرض استوار است که ترجیحات اخلاقی افراد در زبان، روایت‌سازی و الگوهای استدلالی آنان بازتاب می‌یابد. از آنجا که توییتر فضایی برای بروز واکنش‌های اخلاقی و سیاسی است، EBMP توانایی تحلیل این کنش‌های زبانی و پیوند آنها با ترجیحات اخلاقی زیربنایی را دارد.

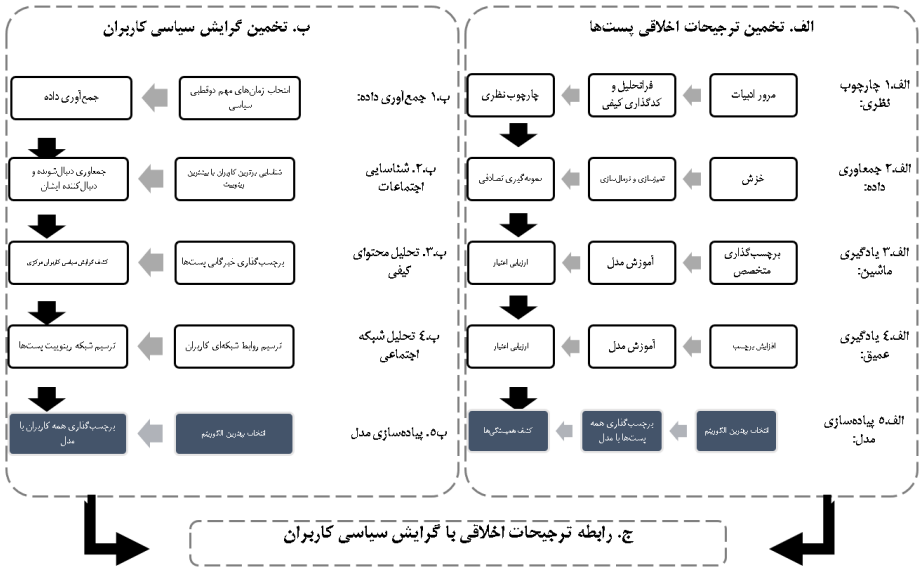
مدل هشت ترجیح اخلاقی دوقطبی (EBMP) در مقایسه با نظریه مبانی اخلاقی هایت (MFT)، ابعاد منحصربه‌فردی مانند اعتماد (با قطب مثبت و منفی) را معرفی می‌کند که روابط بین‌فردی و امنیت اجتماعی را برجسته می‌سازد. همچنین قدرت در EBMP فراتر از احترام/اقتدار در MFTاست و بر سلطه و کنترل تأکید دارد. این افزوده‌ها بازتاب‌دهنده تمرکز EBMP بر زمینه سیاسی و اجتماعی خاص ایران هستند.

EBMP با نظریه سرمایه فرهنگی بوردیو در زمینه هویت گروهی هم‌پوشانی دارد، اما بر ابعاد عاطفی و اخلاقی وفاداری و تعلق تمرکز می‌کند. این مدل با منعطف کردن مدل دوفرایندی، تعادل بین شناخت عقلانی و شهود عاطفی را برقرار می‌سازد. ترجیح شناخت، تصمیمات اخلاقی مبتنی بر استدلال را نشان می‌دهد، در حالی که ترجیح توجه، واکنش‌های شهودی عاطفی را ثبت می‌کند. این ساختار دوقطبی توانایی EBMP در برقراری ارتباط میان چارچوب‌های نظری و ارائه درک دقیق از ترجیحات اخلاقی در توییتر فارسی را نشان می‌دهد.

روش‌شناسی پژوهش

مسیر پژوهشی طی‌شده در مقاله در شکل ۱ نمایش داده شده و در ادامه، هر یک از مراحل به‌تفصیل توضیح داده می‌شود.

۱ Eight Bipolar Moral Preferences که از این به بعد در این مقاله با عنوان EBMP از آن یاد خواهیم کرد.



شکل ۱. مسیر و روش‌های پژوهش

۱،۳. برآورد ترجیحات اخلاقی پست‌ها

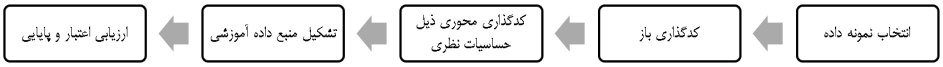
۱،۱،۳. گردآوری پست‌ها

در همکاری با متخصصان فضای مجازی، واژگان کلیدی مرتبط با انتخابات، شناسایی و به‌عنوان کوئری طراحی شد. سپس با استفاده از یک روش خزش[1] سفارشی داده‌های مورد نیاز از توییتر، در بازه زمانی ۱۱ اردیبهشت تا ۸ خرداد ۱۳۹۶، گردآوری شد. برای پاک‌سازی و نرمال‌سازی داده‌ها، حروف به شکل استاندارد تبدیل و محتوای نامرتبط، از جمله تبلیغات، نشانه‌گذاری‌ها و نمادهای اضافی (نظیر @، #، ویرگول و نقطه‌ویرگول) حذف شد. در نهایت، تعداد ۲٬۳۰۸٬۳۲۵ توییت مرتبط با انتخابات در این بازه زمانی گردآوری شد.

۲،۱،۳. یادگیری ماشینی ترجیحات اخلاقی

برای آماده‌سازی داده‌ها جهت آموزش مدل، از همان مجموعه داده کدگذاری‌شده، (دارای نمونه تصادفی شامل ۱۰٬۰۰۰ پست از بازه یک‌ماهه پیش از انتخابات)، استفاده شد.

1 Crawling



شکل ۱. فرایند کدگذاری توییت‌ها برای تهیه مجموعه‌داده آموزشی

سپس برای هر یک از هشت کلاس ترجیح، تعاریف عملیاتی روشن و مثال‌های واقعی توییت برای هر کد (نمونه‌های تیپیک و مرزی) در اختیار کدگذاران قرار گرفت و جلسات آموزش و هماهنگی برگزار شد. برای اطمینان از پایایی کدگذاری، چهار پژوهشگر همکار با تخصص‌های جامعه‌شناسی، انسان‌شناسی، علوم سیاسی و روان‌شناسی، به‌صورت مشارکتی داده‌ها را طبقه‌بندی کردند. این تلاش میان‌رشته‌ای، خطاهای نظام‌مند را به حداقل رسانده و انسجام در کدگذاری را تضمین نمود؛ به‌طوری‌که از طریق جلسات منظم و گفت‌وگوهای مفهومی، درک مشترکی از اصول نظری و اصلاح دسته‌بندی‌ها حاصل شد و ضریب توافق نهایی۷۱٫۶۸٪ بدست آمد.

جدول ۱. ضریب توافق کدگذاران

کارشناس	تعداد تگ مشترک	کاپای کوهن
اول و دوم	۳۷۶	۷۶.۵۹
اول و سوم	۴۷۲	۶۹.۰۶
اول و چهارم	۲۱۶۰	۷۴.۶۲
اول و پنجم	۱۵۲۰	۶۷.۱
میانگین وزنی	۴۵۲۸	۷۱.۶۸

نتیجه برچسب‌گذاری داده‌ها این شد که هر پست در یکی از هشت ترجیح اخلاقی یا در دسته «سایر» (برای محتوای ناسازگار با الگوها) جای گرفت.

مجموعه‌داده استانداردشده حاصل از مرحله کدگذاری باز انسانی، در فاز یادگیری ماشین مورد استفاده قرار گرفت. این داده‌ها به دو بخش تقسیم شدند: ۹۰درصد برای آموزش و ۱۰درصد برای آزمون. در مرحله نخست، از مدل ماشین بردار پشتیبان[1] برای طبقه‌بندی داده‌ها استفاده شد؛ مدلی

1 Support Vector Machine – SVM

که با بیشینه‌سازی حاشیه بین بردارهای پشتیبان، مرز تصمیم‌گیری را بهینـه مـی‌سـازد (Guido et al., 2024).

جدول ۲. روایی مدل‌های کلاسیک یادگیری ماشین (SVM)

قدرت

	precision	recall	f۱-score	support
−۱	۰.۰۰	۰.۰۰	۰.۰۰	۱۶
۰	۰.۹۲	۰.۹۹	۰.۹۵	۵۲۷
۱	۰.۸۶	۰.۳۸	۰.۵۳	۴۷
accuracy			۰,۹۲	۵۹۰
macro avg	۰,۵۹	۰,۴۶	۰,۴۹	۵۹۰
weighted avg	۰,۸۹	۰,۹۲	۰,۸۹	۵۹۰

اعتماد

	precision	recall	f۱-score	support
−۱	۰.۶۳	۰.۲۵	۰.۳۲	۱۶۲
۰	۰.۵۹	۰.۸۶	۰.۷۰	۳۲۰
۱	۰.۶۱	۰.۱۸	۰.۲۷	۱۰۸
accuracy			۰,۵۷	۵۹۰
macro avg	۰,۵۴	۰,۴۳	۰,۴۳	۵۹۰
weighted avg	۰,۵۵	۰,۵۷	۰,۵۲	۵۹۰

مسئولیت

	precision	recall	f۱-score	support
−۱	۰.۴۹	۰.۲۰	۰.۲۹	۱۲۴
۰	۰.۵۲	۰.۴۶	۰.۴۹	۲۳۷
۱	۰.۴۸	۰.۷۰	۰.۵۷	۲۲۹
accuracy			۰,۵۰	۵۹۰
macro avg	۰,۵۰	۰,۴۵	۰,۴۵	۵۹۰
weighted avg	۰,۵۰	۰,۵۰	۰,۴۸	۵۹۰

گروه

	precision	recall	f۱-score	support
−۱	۰.۰۰	۰.۰۰	۰.۰۰	۵۹
۰	۰.۶۶	۰.۴۷	۰.۵۴	۲۰۸
۱	۰.۶۶	۰.۹۰	۰.۷۶	۳۲۳
accuracy			۰,۶۶	۵۹۰
macro avg	۰,۴۴	۰,۴۵	۰,۴۳	۵۹۰
weighted avg	۰,۵۹	۰,۶۶	۰,۶۱	۵۹۰

جلب توجه

	precision	recall	f۱-score	support
−۱	۰.۰۰	۰.۰۰	۰.۰۰	۳۳
۰	۰.۵۷	۰.۴۲	۰.۴۸	۲۵۳
۱	۰.۵۶	۰.۷۵	۰.۶۴	۳۰۴
accuracy			۰,۵۶	۵۹۰
macro avg	۰,۳۸	۰,۳۹	۰,۳۸	۵۹۰
weighted avg	۰,۵۳	۰,۵۶	۰,۵۴	۵۹۰

اینرسی

	precision	recall	f۱-score	support
−۱	۱.۰۰	۰.۳۸	۰.۵۵	۳۲
۰	۰.۹۱	۰.۹۹	۰.۹۵	۵۲۲
۱	۰.۴۰	۰.۰۶	۰.۱۰	۳۶
accuracy			۰,۹۰	۵۹۰
macro avg	۰,۷۷	۰,۴۷	۰,۵۳	۵۹۰
weighted avg	۰,۹۸	۰,۹۷	۰,۹۷	۱۰۰۰

شناخت

	precision	recall	f۱-score	support
−۱	۰.۵۷	۰.۱۱	۰.۱۸	۱۱۹
۰	۰.۵۵	۰.۴۷	۰.۵۱	۲۲۰
۱	۰.۵۲	۰.۷۹	۰.۶۳	۲۵۱
accuracy			۰,۵۳	۵۹۰
macro avg	۰,۵۵	۰,۴۶	۰,۴۴	۵۹۰
weighted avg	۰,۵۴	۰,۵۳	۰,۴۹	۵۹۰

ارزش

	precision	recall	f۱-score	support
−۱	۱.۰۰	۰.۰۵	۰.۰۹	۴۲
۰	۰.۶۸	۰.۹۵	۰.۷۹	۳۹۶
۱	۰.۶۶	۰.۱۲	۰.۱۹	۱۵۲
accuracy			۰,۶۷	۵۹۰
macro avg	۰,۷۱	۰,۳۷	۰,۳۶	۵۹۰
weighted avg	۰,۶۵	۰,۶۷	۰,۵۹	۵۹۰

مدل ماشین بردار پشتیبان (SVM) در مرحله نخست با استفاده از مجموعـه‌داده کیفـی آمـوزش داده شد. این فرایند چندین بار و با افزایش تدریجی حجم داده‌ها تکرار گردید. پس از آمـوزش مـدل SVM، داده‌های به‌روزرسانی‌شده نیز با استفاده از مـدل XGBoost آمـوزش داده شـدند تـا امکـان مقایسه عملکرد فراهم شود. در واقع برای افزایش دقـت پـیش‌بینـی، از تقویـت گرادیـان[1] در قالـب

1 Gradient Boosting

الگوریتم XGBoost استفاده شد؛ روشی پیشرفته در یادگیری تجمعی[1] که به‌طور گسترده در وظایف طبقه‌بندی و رگرسیون به‌کار می‌رود (Chen & Guestrin, 2016:785). با این حال، نتایج مدل XGBoost بهبود معناداری نسبت به SVM نشان نداد.

جدول ۳. روایی مدل‌های کلاسیک یادگیری ماشین (XGBoost)

قدرت

	precision	recall	f1-score	support
−۱	۰.۰۰	۰.۰۰	۰.۰۰	۱۶
۰	۰.۹۱	۱.۰۰	۰.۹۶	۵۲۷
۱	۱.۰۰	۰.۳۰	۰.۶۶	۴۷
accuracy			۰,۹۲	۵۹۰
macro avg	۰,۶۴	۰,۴۳	۰,۴۷	۵۹۰
weighted avg	۰,۹۰	۰,۹۲	۰,۸۹	۵۹۰

اعتماد

	precision	recall	f1-score	support
−۱	۰.۴۰	۰.۰۵	۰.۰۹	۱۶۲
۰	۰.۵۵	۰.۹۶	۰.۷۰	۳۲۰
۱	۰.۶۷	۰.۰۹	۰.۱۶	۱۰۸
accuracy			۰,۵۵	۵۹۰
macro avg	۰,۵۴	۰,۳۷	۰,۳۲	۵۹۰
weighted avg	۰,۵۳	۰,۵۵	۰,۴۳	۵۹۰

مسئولیت

	precision	recall	f1-score	support
−۱	۰.۴۱	۰.۱۹	۰.۲۶	۱۲۴
۰	۰.۶۶	۰.۷۰	۰.۵۵	۲۳۷
۱	۰.۵۰	۰.۳۸	۰.۴۳	۲۲۹
accuracy			۰,۴۷	۵۹۰
macro avg	۰,۴۶	۰,۴۲	۰,۴۲	۵۹۰
weighted avg	۰,۴۷	۰,۴۷	۰,۴۴	۵۹۰

جلب توجه

	precision	recall	f1-score	support
−۱	۰.۰۰	۰.۰۰	۰.۰۰	۳۳
۰	۰.۵۹	۰.۳۸	۰.۶۶	۲۵۳
۱	۰.۵۶	۰.۷۹	۰.۶۶	۳۰۴
accuracy			۰,۵۷	۵۹۰
macro avg	۰,۳۸	۰,۳۹	۰,۳۷	۵۹۰
weighted avg	۰,۵۴	۰,۵۷	۰,۵۴	۵۹۰

ارزش

	precision	recall	f1-score	support
−۱	۰.۰۰	۰.۰۰	۰.۰۰	۶۲
۰	۰.۶۷	۰.۹۹	۰.۸۰	۳۹۶
۱	۰.۵۷	۰.۰۳	۰.۰۵	۱۵۲
accuracy			۰,۶۷	۵۹۰
macro avg	۰,۴۲	۰,۳۴	۰,۲۸	۵۹۰
weighted avg	۰,۶۰	۰,۶۷	۰,۵۵	۵۹۰

شناخت

	precision	recall	f1-score	support
−۱	۰.۴۵	۰.۲۴	۰.۲۹	۱۱۹
۰	۰.۶۲	۰.۷۵	۰.۵۳	۲۲۰
۱	۰.۶۸	۰.۲۲	۰.۳۰	۲۵۱
accuracy			۰,۴۲	۵۹۰
macro avg	۰,۴۱	۰,۴۰	۰,۳۷	۵۹۰
weighted avg	۰,۴۳	۰,۴۲	۰,۳۸	۵۹۰

اینرسی

	precision	recall	f1-score	support
−۱	۱.۰۰	۰.۳۸	۰.۵۵	۳۲
۰	۰.۹۲	۱.۰۰	۰.۹۶	۵۲۲
۱	۱.۰۰	۰.۳۶	۰.۵۳	۳۶
accuracy			۰,۹۳	۵۹۰
macro avg	۰,۹۷	۰,۵۸	۰,۶۸	۵۹۰
weighted avg	۰,۹۳	۰,۹۳	۰,۹۱	۵۹۰

گروه

	precision	recall	f1-score	support
−۱	۰.۰۰	۰.۰۰	۰.۰۰	۵۹
۰	۰.۵۲	۰.۶۳	۰.۵۷	۲۰۸
۱	۰.۶۹	۰.۷۳	۰.۷۱	۳۳۳
accuracy			۰,۶۲	۵۹۰
macro avg	۰,۴۱	۰,۴۵	۰,۴۳	۵۹۰
weighted avg	۰,۵۶	۰,۶۲	۰,۵۹	۵۹۰

در مجموع، دقت مدل‌ها نسبت به آموزش اولیه بر روی نمونه ۱۰،۰۰۰تایی به‌طور قابل‌توجهی افزایش یافت. در میان مدل‌های یادگیری ماشین، مدل SVM بالاترین دقت را به دست آورد. دسته‌هایی که دقتی بالاتر از ۹۵٪ داشتند، عبارت بودند از: قدرت مثبت، مسئولیت خنثی، مسئولیت

1 ensemble learning

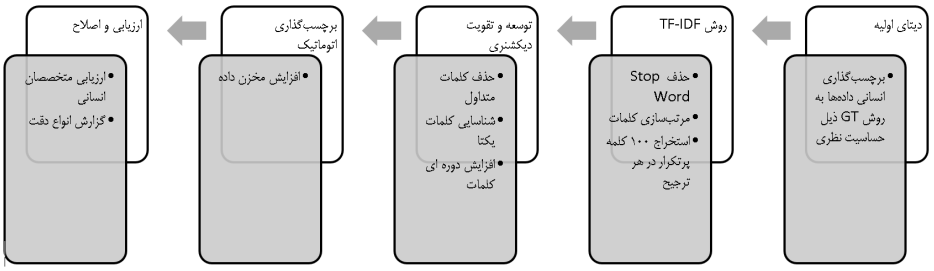
مثبت، گروه‌گرایی منفی، اینرسی منفی، اینرسی خنثی و سایر دسته‌ها دقتی کمتر از ۹۰٪ داشتند. با توجه به این یافته‌ها، ادامه تحلیل با استفاده از مدل‌های یادگیری عمیق[1] به‌منظور ارتقاء دقت و بهبود عملکرد کلی مدل ضروری تشخیص داده شد.

۳،۱،۳. یادگیری عمیق ترجیحات اخلاقی

به‌منظور رفع مشکل دقت پایین مدل، از رویکرد یادگیری نیمه‌نظارتی[2] استفاده شد (Zhu, 2009). برای استخراج واژگان متمایز مرتبط با قطب‌های مختلف مدل EBMP، از روش TF-IDF بهره گرفته شد (Kim & Gil, 2019: 230). در این فرایند، کلمات ممنوعه حذف شدند و ۱۰۰ واژه برتر برای هر دسته استخراج گردید. واژگان مشترک میان ترجیحات حذف شدند و تنها کلمات با قدرت پیش‌بینی بالا حفظ گردیدند.

روش گسترش واژه‌محور نیز برای بهبود دقت نهایی به کار گرفته شد (Tsamardinos et al., 2018: 75). واژه‌نامه نهایی از دو منبع تهیه شد: ۸۰درصد از توییتر فارسی در دوره انتخابات ۱۳۹۶ و ۲۰درصد از واژه‌نامه‌های روان‌شناختی موجود (Pennebaker et al., 2015: 56). در فرایند برچسب‌گذاری خودکار، اگر حداقل دو واژه از واژه‌نامه سفارشی در یک پست وجود داشت، آن پست به یکی از ترجیحات اخلاقی تخصیص داده می‌شد.

بازبینی کارشناسانه موجب گسترش مجموعه‌داده برچسب‌خورده تا ۱۲۰۰۰۰ پست شد. در مرحله اعتبارسنجی، بررسی تصادفی ۵۰۰ پست برچسب‌خورده توسط متخصصان، نرخ دقتی بیش از ۸۹درصد را نشان داد که مؤید اثربخشی روش گسترش واژه‌محور مبتنی بر یادگیری نیمه‌نظارتی بود.



شکل ۲. فرایند برچسب‌گذاری و گسترش داده‌ها با استفاده از یادگیری نیمه‌نظارتی برای مدل EBMP

1 Deep Learning
2 Semi-Supervised Learning

۴،۱،۳. پیاده‌سازی مدل یادگیری عمیق برای استخراج ترجیحات اخلاقی از توییتر فارسی

آموزش مدل: از ۸۰درصد داده‌های برچسب‌خورده برای آموزش هشت شبکه عصبی مجزا (هر کدام برای یک ترجیح اخلاقی) استفاده شد. داده‌های متنی با بردارهای معنایی GloVe (Pennington et al., 2014: 1535) و FastText (Bojanowski et al., 2017: 135) به بردارهای عددی پنجاه بُعدی تبدیل شدند.

معماری مدل: از معماری LSTM برای پردازش توالی داده‌ها و از Transformer و BERT برای بهینه‌سازی با سازوکار خود–توجهی[1] استفاده شد (Vaswani et al., 2017: 34; Devlin et al., 2019: 15). از الگوریتم Adam (Kingma & Ba, 2014: 142) با نرخ یادگیری 10^{-4} برای بهینه‌سازی و تابع هزینه[2] Categorical Cross-Entropy در کاهش خطا استفاده گردید. از لایه تعبیه واژگانی[3] با بُعد ۵۰ استفاده شد که همزمان با سایر لایه‌ها آموزش داده می‌شد.

بهبود کارایی: در این طرح، برای هر ترجیح، کل داده‌ها به پنج بخش هم‌حجم تقسیم شد و در هر تکرار، یکی از بخش‌ها به‌عنوان مجموعه آزمون و چهار بخش دیگر به‌عنوان مجموعه آموزش به مدل ارائه گردید. بدین‌ترتیب، پنج سناریوی متفاوت آموزش و آزمون تولید شد و میانگین و انحراف معیار شاخص‌های عملکردی (از جمله دقت، Recall وF1) بر روی این پنج تکرار گزارش شد. این رویکرد امکان ارزیابی پایداری مدل در برابر تقسیم‌بندی‌های متفاوت داده و سنجش تعمیم‌پذیری آن را فراهم ساخت. الگوهای به‌دست‌آمده از این تحلیل‌ها نشان داد که در اغلب ترجیحات، مدل‌های عمیق نسبت به SVM دقت بالاتری ارائه می‌کنند، هرچند ضعف نسبی در تشخیص کلاس خنثی به‌عنوان یک ویژگی ساختاریِ مسئله باقی می‌ماند و در بخش نتایج و بحث به‌طور تفصیلی تفسیر شده است.

1 self-attention
2 Loss Function
3 Word embedding

جدول ٤. جدول اعتباریابی دقت Ac انواع ترجیحات طبق kfold cross validation

	قدرت	اعتماد	مسئولیت	توجه	ارزش	شناخت	گروه	اینرسی
بازه بالای سطح اطمینان ۹۵	۹۴.۸۸	۹۴.۹۵	۹۴.۸	۹۵.۳۴	۹۸.۶۲	۹۴.۲	۹۳.۵۴	۹۲.۸۳
بالاترین دقت	۹۵	۹۵	۹۵	۹۵	۹۷	۹۴	۹۴	۹۳
پایین ترین دقت	۹۴	۹۱	۹۳	۸۶	۷۴	۹۳	۹۲	۸۷
بازه پایین سطح اطمینان ۹۵	۹۳.۹۱	۹۲.۲۴	۹۲.۴	۸۸.۶۵	۸۱.۷۷	۹۳.۴	۹۲.۰۵	۸۸.۷۶
میانگین	۹۵	۹۴	۹۵	۹۵	۹۰	۹۴	۹۳	۹۲
انحراف معیار	۹۴	۹۴	۹۳	۹۳	۹۵	۹۴	۹۳	۹۱
خطای استاندارد میانگین	۹۴	۹۴	۹۲	۹۱	۹۵	۹۴	۹۲	۹۱

اعتبارسنجی: با استفاده از اعتبارسنجی متقابلk بخشی (k=5) انجام شد. نتایج نشان داد مدل‌های عمیق دقت بالاتری نسبت به SVM دارند، اگرچه در تشخیص کلاس «خنثی» ضعف نسبی نشان می‌دادند (کاهش ۱۰درصد دقت در مجموعه آزمون مستقل). برای حل این مشکل، از لایه تعبیه واژگانی اختصاصی و در برخی موارد از ParsBERT استفاده شد.

در نهایت، مدل موفق به دستیابی به دقت قابل قبولی در طبقه‌بندی ترجیحات اخلاقی در داده‌های توییتری فارسی گردید. این روش امکان طبقه‌بندی دقیق‌تر ترجیحات اخلاقی در داده‌های توییتری فارسی را فراهم کرد و این مدل با استفاده از توییت‌های مربوط به انتخابات ریاست‌جمهوری سال ۱۳۹۶ ایران آموزش دیده و تنظیم شد. نتیجه آن، طبقه‌بندی کل مجموعه‌داده بر اساس ترجیحات اخلاقی بود. سپس این ترجیحات بر اساس فرکانس وقوع اولویت‌بندی شدند تا تصویری دقیق از ابعاد اخلاقی موجود در داده‌ها ارائه شود.

مدل‌های مختلف یادگیری عمیق در چندین مرحله، با استفاده از مجموعه‌های آموزشی متفاوت برای هر ترجیح اخلاقی، آموزش داده شدند.

جدول ۵. ارزیابی عملکرد مدل یادگیری عمیق

ترجیح	داده آموزشی	داده تست	accuracy	f1-score			recall			precision		
				۱	۰	-۱	۱	۰	-۱	۱	۰	-۱
قدرت	۱۰۸۰۵۵۸	۲۷۰۱۴۰	۹۸	۹۸	۹۷	۹۹	۹۹	۹۶	۹۹	۹۸	۹۷	۹۹
اعتماد	۲۱۵۰۳۱۰	۵۳۰۸۲۸	۹۶	۹۸	۹۴	۹۶	۹۸	۹۴	۹۶	۹۸	۹۳	۹۷
مسئولیت	۲۰۶۰۸۶۶	۵۱۰۷۱۷	۹۶	۹۸	۹۳	۹۵	۹۷	۹۶	۹۴	۹۹	۹۱	۹۷
توجه	۲۰۷۰۴۲۴	۵۱۰۸۵۷	۹۷	۹۸	۹۳	۹۷	۹۸	۹۵	۹۵	۹۹	۹۱	۹۸
ارزش	۱۸۱۰۰۴۶	۴۵۰۲۶۲	۹۸	۹۹	۹۴	۹۸	۹۹	۹۳	۹۸	۹۹	۹۴	۹۹
شناخت	۲۱۴۰۳۲۴	۵۳۰۵۸۲	۹۸	۹۹	۹۳	۹۸	۹۸	۹۵	۹۷	۹۹	۹۱	۹۹
گروه	۱۵۸۰۶۱۲	۳۹۰۶۵۳	۹۷	۹۷	۹۲	۹۹	۹۶	۹۴	۹۹	۹۷	۹۱	۹۹
اینرسی	۶۳۰۱۲۸	۱۵۰۷۸۳	۹۶	۹۷	۹۶	۹۷	۹۵	۹۷	۹۷	۹۸	۹۶	۹۷

۵،۱،۳. اعتبارسنجی مدل یادگیری عمیق

برای اعتبارسنجی نهایی مدل یادگیری عمیق، تعداد ۱۶۰۰ پست به‌صورت تصادفی انتخاب و برچسب‌گذاری مجدد انسانی شدند.

جدول ٦. ماتریس اغتشاش[1] تطابق میان برچسب‌گذاری انسانی و خروجی مدل یادگیری عمیق

	برچسب‌گذاری مجدد انسانی (روی کل مجموعه‌داده که طبق آخرین نتایج مدل عمیق برچسب خورده)				
		+	۰	−	Pr
نتایج مدل یادگیری عمیق بر روی مجموعه داده	+	۱۰۵۴	۲۰۵	۱۲۶	۷۶.۱۰
	۰	۴۵۱	۱۴۸۶	۳۶۵	۶۴.۵۵
	−	۱۶۸	۱۸۴	۷۶۱	۶۸.۳۷
	Re	۶۳.۰۰	۷۹.۲۵	۶۰.۷۸	۶۸.۷۷

در جریان اجرای این آزمون نهایی، ضریب توافق میان مدل هوش مصنوعی و پژوهشگر اصلی برابر با ۶۸،۷۷درصد محاسبه شد؛ که کمی کمتر از ضریب توافق میان متخصصان انسانی (۷۱،۶۸درصد) است. با وجود این تفاوت، نتایج تولیدشده توسط مدل هوش مصنوعی سطح قابل‌اعتمادی از دقت را نشان می‌دهند و می‌توان آن‌ها را در زمینه‌ی این پژوهش معتبر و قابل اتکا دانست.

۲،۳. برآورد گرایش سیاسی کاربران

تحلیل شبکه ریتوییت در توییتر، امکان ردیابی الگوهای ارتباطی و برآورد گرایش سیاسی کاربران را فراهم می‌سازد. تحلیل شبکه‌های اجتماعی[2] چارچوبی نظام‌مند برای بررسی روابط درون یک جامعه ارائه می‌دهد و به شناسایی ویژگی‌های اجتماعی گره‌های منفرد، مانند سرمایه اجتماعی، و همچنین پویایی‌های کلان‌تر در سطح شبکه کمک می‌کند (Scott, 2000: 53-145). این بینش‌ها امکان کشف الگوهای نفوذ و تعاملاتی را فراهم می‌آورند که می‌توانند گرایش‌های سیاسی کاربران را استنباط کنند.

روش‌های شناسایی اجتماعات[3] نقش مهمی در شناسایی خوشه‌ها یا جناح‌های ایدئولوژیک درون شبکه‌ها ایفا می‌کنند. این روش‌ها دامنه‌ای از رویکردهای گره‌محور[4] و گروه‌محور[5] (Aggarwal,

1. Confusion Matrix
2 SNA
3 Community Detection
4 node-oriented
5 group-oriented

(Alizadeh et al., 2017: 365) [1] را در 2011: 65-78 تا روش‌های مبتنی بر ساختار کلی شبکه بر می‌گیرند. الگوریتم‌هایی مانند الگوریتم پیشنهادی بلوندل[2] (2008) با کارایی بالا به شناسایی خوشه‌ها کمک می‌کنند و امکان پیوندزدن روابط اجتماعی کاربران با گرایش‌های سیاسی احتمالی آن‌ها را فراهم می‌سازند. این مجموعه روش‌ها، ابزاری قدرتمند برای فهم پویش‌های سیاسی در شبکه‌های اجتماعی فراهم می‌آورند.

۱،۲،۳. گردآوری داده‌های کاربران

داده‌ها از طریق خزش توییتر و در بازه زمانی ۱۱ اردیبهشت تا ۸ خرداد ۱۳۹۶ گردآوری شدند و طی آن، ۱۰۷٬۵۳۲ کاربر فارسی‌زبان فعال در این بازه زمانی شناسایی شد.

۲،۲،۳. شناسایی اجتماعات

در این مرحله، دو نوع از شبکه‌های رابطه‌محور در توییتر مورد تحلیل قرار گرفتند. روش ساخت شبکه دنبال‌کنندگان[3] بر اساس الگوی پیشنهادی صالح و حدادی (۲۰۲۲، صص ۲۵۰و۲۵۱) انجام شد. در گام نخست، ۴٬۰۰۰ کاربر فارسی‌زبان که در انتخابات ۱۳۹۶ مشارکت بالایی (بیش از پنج ریتوییت) داشتند، شناسایی شدند. سپس اطلاعات دنبال‌کنندگان و دنبال‌شوندگان این کاربران گردآوری شد.

پس از پاک‌سازی و نرمال‌سازی داده‌ها، گراف لایه دوم با ۳۲٬۸۴۶ گره و ۳٬۷۵۲٬۵۲۱ یال ساخته شد. برای تحلیل شبکه از کتابخانه NetworkX در پایتون استفاده شد (Hagberg, Swart & Sculley, 2008: 15-60). این کتابخانه امکان بازتولید عملکردهای نرم‌افزارهایی همچون Gephi را برای تحلیل شبکه‌های عظیم فراهم می‌سازد (Bastian et al., 2009: 361).

۳،۲،۳. تحلیل محتوای کیفی

برای تعیین گرایش سیاسی کاربران لایه اول (۴٬۰۰۰ نفر)، ۱۰۰ توییت آخر هر کاربر با استفاده از روش تحلیل محتوای کیفی بررسی شد. هر پست بر اساس موضع سیاسی آن برچسب‌گذاری گردید و در نهایت، موضع غالب در میان توییت‌ها به‌عنوان گرایش سیاسی نهایی کاربر در نظر گرفته شد. تحلیل محتوا به‌عنوان روشی نظام‌مند و کیفی برای طبقه‌بندی و تفسیر داده‌های متنی به کار گرفته

1 network-oriented
2 Blondel
3 followers network

شد (Krippendorff, 2019: 136). این روش به‌طور خاص برای شناسایی الگوها و روندهای موجود در محتوای رسانه‌های اجتماعی مؤثر و کاربردی است (Neuendorf, 2017: 89-93). در این مطالعه، دسته‌بندی‌های سیاسی از پیش تعریف‌شده به‌صورت یکسان و منسجم بر روی تمامی پست‌ها اعمال شد تا اعتبار و پایایی نتایج تضمین گردد (Stemler, 2001: 18-37).

۴،۲،۳. تحلیل شبکه اجتماعی

با توجه به محدودیت پردازشی تحلیل محتوای کیفی همه ۱۰۷٬۵۳۲ کاربر فعال در موضوع، از شبکه روابط کاربران لایه اول استفاده شد. بدین منظور ارتباطات میان ۴٬۰۰۰ کاربر اولیه شبکه با بررسی اشتراک‌های موجود در شبکه‌های دنبال‌کننده آن‌ها تعیین شد. این فرایند منجر به شناسایی ۲۸٬۸۴۶ گره جدید در گراف لایه دوم شد (Aggarwal, 2011: 194-204). هرچند ارتباطات میان این گره‌های جدید با کاربران اولیه ثبت شده است، اما روابط مستقیم میان خود گره‌های جدید در این مرحله مشخص نبود (Alizadeh et al., 2017: 380). به‌منظور رفع این محدودیت، داده‌های دنبال‌کنندگان و دنبال‌شوندگان گره‌های لایه دوم نیز گردآوری شد و گراف لایه سوم شکل گرفت. این شبکه نهایی شامل ۳۶٬۲۲۴ کاربر و ۱۵٬۸۵۷٬۴۱۴ یال بود که کلیه ارتباطات میان این کاربران را بازنمایی می‌کرد (Blondel et al., 2008: 43).

در ادامه، خوشه‌های اصلی کاربران توییتر فارسی‌زبان بر اساس الگوی دنبال‌کردن متقابل شناسایی شدند. در مرحله بعد، ویژگی‌هایی که پیونددهنده اعضای هر خوشه بودند مورد تحلیل قرار گرفت. تحلیل عددی شاخص‌های شبکه (مانند شاخص‌های مرکزیت) حساب‌های کاربری اثرگذار را برجسته کرد. با استفاده از نمونه‌گیری کوکران، اجتماعات شناسایی‌شده از طریق بررسی محتوای توییت‌ها و بیوگرافی کاربران برچسب‌گذاری شدند. در نهایت، کاربران در چهار گروه طبقه‌بندی شدند: اصلاح‌طلب، اصول‌گرا، معاند، غیرسیاسی. در پایان، شبکه ریتوییت پویای کاربران در آستانه انتخابات ریاست جمهوری ۱۳۹۶ بررسی شد.

۵،۲،۳. پیاده‌سازی مدل تخمین گرایش سیاسی

شبکه روابط دنبال‌کننده در ذات خود ایستا است و ارتباطاتی را بازنمایی می‌کند که ممکن است با هدف نظارت بر گروه‌های مخالف یا پیش از تغییر گرایش‌های سیاسی کاربران شکل گرفته باشند (Borgatti & Foster, 2003: 68). در مقابل، شبکه ریتوییت پویاتر و در حال تغییر مداوم است و تصویری هم‌زمان از تعاملات کاربران ارائه می‌دهد (Himelboim et al., 2013: 87). در این

پژوهش، کاربرانی که بیش از سه بار یکدیگر را ریتوییت کرده‌اند، دارای برچسب سیاسی یکسان در نظر گرفته شدند؛ به این معنا که نوعی هم‌سویی سیاسی میان آنان وجود دارد (Cohen et al., 2016: 35).

فرایند شناسایی اجتماعات توسط دو متخصص، بر اساس یک پروتکل کدگذاری استانداردشده انجام شد؛ پروتکلی که شامل تعاریف نظری و عملیاتی متغیرها و شاخص‌های کلیدی بود. ضریب پایایی بین کدگذاران برابر با ۰٫۷۷ محاسبه شد که نشانگر توافق قابل‌قبول و پایایی بالای نتایج برچسب‌گذاری است (Krippendorff, 2019: 135).

۳٫۳. ارتباط ترجیحات اخلاقی با گرایش سیاسی

پس از شناسایی ترجیحات اخلاقی بیان‌شده در توییت‌های کاربران، رابطه این ترجیحات با گرایش سیاسی بررسی شد. برای کمّی‌سازی قدرت و جهت این روابط، از روش‌های آماری همچون ضریب همبستگی پیرسون و رگرسیون استفاده شد. نتایج به‌دست‌آمده نشان داد که چگونه ارزش‌های اخلاقی خاص با گرایش‌های سیاسی هم‌راستا می‌شوند و از این طریق، تعامل میان باورهای اخلاقی و دیدگاه‌های سیاسی در جامعه توییتری ایران روشن‌تر گردید.

۳٫۴. ملاحظات روش‌شناختی

تحلیل الگوهای نظم اجتماعی بر پایه ترجیحات اخلاقی و ارزشی، ماهیتی میان‌رشته‌ای دارد و نیازمند تلفیق چارچوب‌هایی از حوزه‌های گوناگون مانند فلسفه، جامعه‌شناسی، روان‌شناسی شناختی، اقتصاد رفتاری و علوم سیاسی است. در این راستا، همکاری با متخصصان این حوزه‌ها (از جمله جامعه‌شناسان و متخصصان هوش مصنوعی) برای دستیابی به فهمی جامع و چندبُعدی ضروری است.

تحلیل کلان‌داده‌های توییتر با چالش‌های منحصربه‌فردی همراه است. زبان غیررسمی کاربران، پردازش متن را دشوار می‌سازد (Paris et al., 2012:539) و محتوای نامربوط، همراه با پیچیدگی زبانی زبان فارسی، فرآیند پالایش و تحلیل داده را با دشواری‌هایی مواجه می‌سازد (Zhang et al., 2020: 32-38). همچنین، محدودیت تعداد کاراکتر در توییتر باعث کاهش غنای محتوایی پست‌ها می‌شود. برای مقابله با این محدودیت‌ها، استفاده ترکیبی از شبکه روابط کاربران و الگوهای بازنشر برای استخراج بینش‌های رفتاری ضروری است. افزون بر این، برچسب‌گذاری دقیق انسانی نقش حیاتی در تولید مجموعه‌داده‌های آموزشی باکیفیت داشته و به افزایش دقت و اعتبار مدل‌های تحلیلی یاری می‌رساند.

یافته‌های پژوهش

۱،۴. ترجیحات اخلاقی در توییتر فارسی

ویژگی‌های مجموعه‌داده مورد استفاده در این پژوهش، در جدول زیر ارائه شده‌اند.

جدول ۷. ویژگی‌های مجموعه‌داده

توییت‌های مرتبط با موضوع	۲،۳۰۸،۳۲۵
تعداد کلمات منحصربفرد کل مجموع داده	۱،۱۷۳،۸۲۳
حجم توییت تصادفی انتخابی برای برچسب انسانی	۱۰،۰۰۰
تعداد کلمات منحصربفرد مجموع داده برچسب‌خورده انسانی	۲۶،۱۵۱
میانگین طول بردار کلمات هر پست مجموع داده برچسب‌خورده انسانی	۱۹,۱۶

در بخش داده‌کاوی، نتایج حاصل از کدگذاری کیفی و تخصصی ۱۰،۰۰۰ پست با بهره‌گیری از روش کیفی ارائه شد. در این بخش، با به‌کارگیری روش یادگیری عمیق بر روی کل مجموعه‌داده، نتایج نهایی داده‌کاوی با سطح دقت توضیح‌داده‌شده در بخش‌های پیشین ارائه می‌گردد.

جدول ۸. فراوانی ترجیحات اخلاقی در توییت‌های انتخابات ۱۳۹۶ (بر اساس مدل یادگیری عمیق)

ترجیح	فراوانی	درصد نسبت به کل پست‌ها
گروه	۴۶۳,۵۹۰	٪ ۴۷.۹۵
اعتماد	۳۷۲,۶۳۲	٪ ۳۸.۵۴
مسئولیت	۳۳۱,۰۴۳	٪ ۳۴.۲۴
توجه	۳۱۴,۷۴۶	٪ ۳۲.۵۵
شناخت	۳۰۳,۱۳۳	٪ ۳۱.۳۵
ارزش	۲۹۲,۹۸۲	٪ ۳۰.۳۰
قدرت	۱۹۴,۸۴۸	٪ ۲۰.۱۵
اینرسی	۱۰۵,۲۳۰	٪ ۱۰.۸۸

جدول ۹. فراوانی ترجیحات اخلاقی دوقطبی (EBMP) در توییت‌های انتخابات ۱۳۹۶ (بر اساس مدل یادگیری عمیق)

ترجیح	فراوانی	درصد نسبت به کل پست‌ها
گروه+	۳۰۷,۱۹۹	٪ ۳۱.۷۷
توجه+	۲۹۸,۶۲۴	٪ ۳۰.۸۹
اعتماد-	۲۷۵,۱۱۴	٪ ۲۸.۴۵
ارزش+	۲۴۱,۴۲۶	٪ ۲۴.۹۷
مسئولیت+	۲۰۸,۷۸۶	٪ ۲۱.۵۹
شناخت+	۱۸۴,۷۷۹	٪ ۱۹.۱۱
گروه-	۱۵۶,۳۹۱	٪ ۱۶.۱۷
قدرت+	۱۳۱,۳۹۱	٪ ۱۳.۵۹
مسئولیت-	۱۲۲,۲۵۷	٪ ۱۲.۶۴
شناخت-	۱۱۸,۳۵۴	٪ ۱۲.۲۴
اعتماد+	۹۷,۵۱۸	٪ ۱۰.۰۸
اینرسی-	۶۶,۶۳۵	٪ ۶.۸۹

۲,۴. گرایش سیاسی در توییتر فارسی

با استفاده از روش تحلیل شبکه اجتماعی مبتنی بر دنبال‌کنندگان، چهار خوشه از کاربران شناسایی شدند.

جدول ۱۰. جامعه آماری جوامع چهارگانه در شبکه دنبال‌کنندگان کاربران توییتر فارسی

	جامعه الف	جامعه ب	جامعه ج	جامعه د
تعداد کاربران	۲۴,۹۵۵	۴,۵۳۹	۵,۳۹۵	۱,۳۳۳
تعداد روابط میان کاربران	۱۱,۰۲۴,۹۵۸	۱,۹۲۹,۲۲۳	۲,۴۸۱,۷۶۸	۴۷۴,۹۶۳
میانگین درجه کلی	۹۵۱	۸۱۵	۱۳۲۸	۱۳۷۴
میانگین درجه درونی	۶۷۵	۴۷۴	۹۵۱	۱۰۲۰
نسبت ارتباطات درون اجتماعی	٪۶۹	٪۷۹	٪۷۸	٪۷۴
چگالی	۰.۰۵۷	۰.۰۱۷	۰.۰۲۵	۰.۰۷۴
میانگین مرکزیت نزدیکی	۰.۵۹	۰.۵۷	۰.۵۵	۰.۵۷
میانگین مرکزیت میانی	۰.۰۰۰۰۱	۰.۰۰۰۰۱	۰.۰۰۰۰۱	۰.۰۰۰۰۱
دسته‌بندی کاربران	اصلاح‌طلب	اصولگرا	معاند	غیرسیاسی

جامعه الف: کاربران در جامعه الف به شدت سیاسی هستند و از اصلاحات سیستماتیک و شخصیت‌هایی مانند روحانی و جهانگیری حمایت می‌کنند. این گروه بزرگترین شبکه را تشکیل

می‌دهد و نمایانگر ایده‌های میانه چپ و راست است که تغییرات را در چارچوب جمهوری اسلامی ترویج می‌دهد.

جامعه ب: این جامعه شامل حامیان اصولگرایی است که با تأکید بر ارزش‌های سنتی انقلاب اسلامی با شخصیت‌هایی مانند رئیسی و قالیباف همسو هستند.

جامعه ج: جامعه ج شامل مخالفان و معترضان آشکار جمهوری اسلامی است که با محتوای به شدت منتقدانه، اغلب برای تغییر رژیم تبلیغ می‌کنند.

جامعه د: این گروه عمدتاً غیرسیاسی است و بیشتر بر موضوعات مربوط به زندگی روزمره مانند ابراز احساسات شخصی، رفتار اجتماعی و سرگرمی‌ها تمرکز دارد. بسیاری از اعضای این جامعه از هویت خود محافظت کرده و از بحث‌های سیاسی اجتناب می‌کنند.

با ترسیم شبکه ریتوییت در بازه اردیبهشت و خرداد ۱۳۹۶، این خوشه‌ها در روابط پویا، متعین شده و ۵۲،۶۲۶ کاربر اصلی در چهار دسته شناسایی شده است.

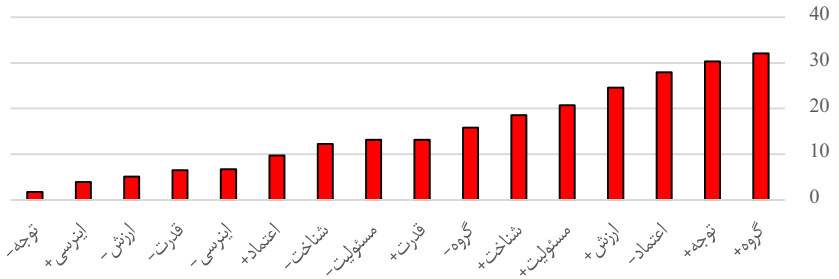
جدول ۱۱. تعداد کاربران فعال اردیبهشت و خرداد ۱۳۹۶ در توییتر بر اساس گرایش سیاسی

گرایش سیاسی	تعداد کاربر	پست	میانگین توئیت تولیدی هر کاربر
اصلاح‌طلب	۳۶,۲۵۵	۵۰۴,۸۹۳	۱۳.۹۲
اصولگرا	۶,۵۹۵	۱۸۹,۶۳۶	۲۸.۷۵
معاند	۷,۸۳۸	۲۱۶,۳۳۰	۲۷.۶۰
غیرسیاسی	۱,۹۳۸	۵۴,۸۲۷	۲۸.۲۹

شکل ۳. نمودار ارتباطی کاربران توییتر در اردیبهشت و خرداد ۱۳۹۶

۳,۴. ارتباط ترجیحات اخلاقی با گرایش سیاسی در توییتر فارسی

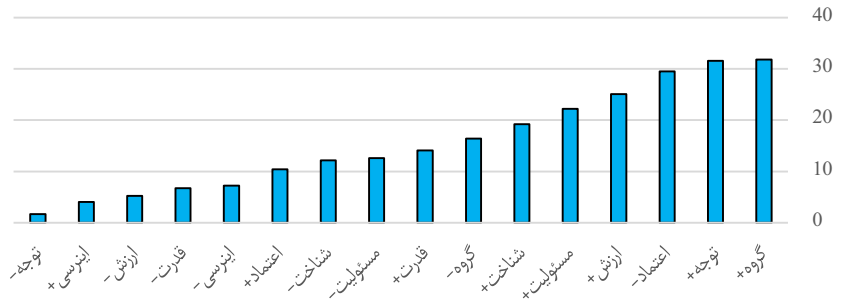
اکنون این امکان فراهم است که پست‌ها برای هر گرایش سیاسی در شکل‌های ۶ تا ۱۰ بررسی شود.



شکل ٤. بررسی EBMP در میان کاربران اصلاح‌طلب

جریان اصلاح‌طلب بر پیوندهای مبتنی بر گروه تأکید می‌کند، ارزش زیادی برای نظرات دیگران قائل است و مسئولیت زیادی در قبال جامعه خود نشان می‌دهد. آن‌ها از آزادی‌های مدنی، برابری جنسیتی و مبارزه با فساد حمایت می‌کنند، در حالی که بی‌اعتمادی و ناامیدی نسبت به وضعیت سیاسی موجود دارند. ویژگی‌های کلیدی مانند گروه‌گرایی، جلب توجه، بی‌اعتمادی، جهت‌گیری ارزشی، مسئولیت و شناخت عقلانی، ترجیحات آن‌ها را تعریف می‌کند.

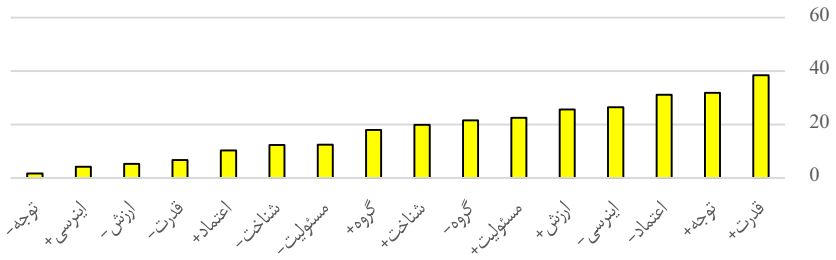
استراتژی سیاسی آن‌ها با نظریه هویت اجتماعی (Tajfel & Turner, 2004: 276) هم‌راستا است که بر همبستگی گروهی و مبارزات جمعی برای ایجاد تغییر تأکید دارد. بی‌اعتمادی نشان‌دهنده شک و تردید آن‌ها نسبت به ساختارهای حکومتی (Hetherington, 1998: 791) است، در حالی که شناخت عقلانی و ارزش‌های غیرمادی (Inglehart, 2003) تمرکز آن‌ها را بر اصلاحات اجتماعی دموکراتیک، با وجود چالش‌های سیستمی، نشان می‌دهد. نظریه ناراحتی شناختی (Festinger, 1957) توضیح می‌دهد که چگونه اصلاح‌طلبان، مشارکت خود را در سیستمی که به آن بی‌اعتمادند، توجیه می‌کنند و آن را به عنوان گامی به سوی تغییر تدریجی و معنادار می‌بینند.



شکل ۵. بررسی EBMP در میان کاربران اصول‌گرا

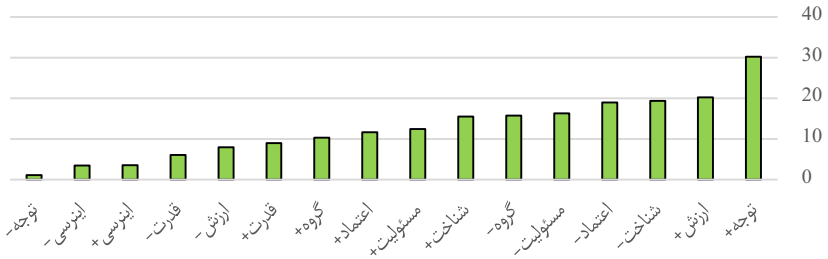
اصول‌گرایان و اصلاح‌طلبان اغلب به عنوان مخالفان سیاسی دیده می‌شوند، اما داده‌ها شباهت‌های چشمگیری در ترجیحات ارزش‌های اصلی آن‌ها نشان می‌دهد. وابستگی به گروه با افزایش توجه و مسئولیت همراه است، در حالی که استقلال و قدرت کاهش می‌یابد. افراد اولویت را به هنجارهای گروهی و همبستگی می‌دهند و رفتار خود را با ارزش‌های جمعی هم‌راستا می‌کنند (Tajfel & Turner, 1979). این امر خودمختاری را کاهش می‌دهد، زیرا وفاداری به گروه منجر به هم‌خوانی می‌شود (Hogg, 2001: 184).

رابطه معکوس گروه‌گرایی با قدرت بر تسلط هنجارهای گروهی بر کنترل فردی تأکید دارد و دیدگاه بندورا (۱۹۸۶) در مورد تأثیر اجتماعی را بازتاب می‌دهد. پیوندهای قوی گروهی ثبات را تقویت کرده و مقاومت در برابر تغییر را به وجود می‌آورد، به طوری که همبستگی را بر ترجیحات فردی اولویت می‌دهند. اصول‌گرایان بر استقلال، عدالت، دیانت و سادگی تأکید دارند که با برخی از ارزش‌های مورد حمایت اصلاح‌طلبان هم‌راستا است. تفاوت اصلی بیشتر در تفسیر و کاربرد این ارزش‌های مشترک است تا در خود ارزش‌ها. این امر ماهیت دقیق ایدئولوژی‌های سیاسی را نشان می‌دهد که در آن ارزش‌های مشترک علی‌رغم اهداف سیاسی متضاد با هم هم‌زیستی دارند (Inglehart, 2003: 120-135; Schwartz, 1992: 71).



شکل ۷. بررسی EBMP در میان کاربران مخالف نظام

مخالفان نظام با مواضع ایدئولوژیک سخت‌گیرانه هم‌راستا هستند و بر بی‌اعتمادی به نظم موجود تأکید می‌کنند. گفتار آن‌ها اغلب از نافرمانی، مشارکت فعال و تغییر سیستماتیک حمایت می‌کند. ترجیحات اصلی که جهت‌گیری آن‌ها را هدایت می‌کنند شامل رفتار جویای قدرت، جلب توجه، بی‌اعتمادی و تمایل قوی به تغییر است. این ویژگی‌ها با نظریه‌های روان‌شناسی اجتماعی دینامیک‌های قدرت هم‌خوانی دارند، جایی که افراد با ویژگی‌های مستبدانه، تعاملات اجتماعی را از منظر رقابت و کنترل می‌بینند (Altemeyer, 1996: 320). از منظر جامعه‌شناسی فرهنگی، مقاومت آن‌ها در برابر وضعیت موجود، با مفهوم هژمونی فرهنگی گرامشی (۱۴۰۳) هم‌راستا است. درخواست‌های آن‌ها برای اصلاح، هم‌زمان با شکایات شخصی و مبارزات جمعی است. این تنش میان میل به استبداد و جستجوی تغییر، برخورد فرهنگی را برجسته می‌کند که در آن نیاز به کنترل با آرزوهای برای تجدید و نوسازی اجتماعی در تضاد است.

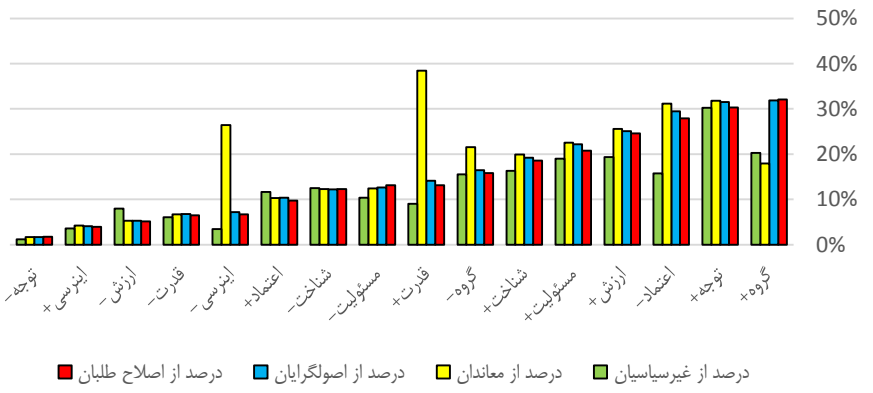


شکل ۸. بررسی EBMP در میان کاربران غیرسیاسی

کاربران غیرسیاسی در شبکه‌های اجتماعی از طریق شوخی، ابراز احساسات و محتوای شخصی فعالیت می‌کنند و از موضوعات سیاسی یا جمعی اجتناب می‌کنند. پست‌های آن‌ها نشان‌دهنده بی‌توجهی به ایدئولوژی‌های سیاسی است، جایی که شوخی و احساسات به عنوان مکانیزم‌های مقابله

یا خودابرازی در محیطی آشفته یا مأیوس‌کننده عمل می‌کنند. این رفتار با مفهوم جداشدگی فرهنگی و نسل «پساایدئولوژیک» هم‌راستا است که با بی‌توجهی بـه سـاختارهای نهـادی شـناخته مـی‌شـود (Couldry, 2012: 148; Fiske, 1994: 12).

این کاربران رفتارهایی برای جلب توجه و شناخت احساسی از خود نشان می‌دهنـد کـه اغلـب بـا بی‌اعتمادی و بی‌مسئولیتی همراه است. جداشدگی آن‌ها از نیازهـای اجتمـاعی مـنعکس‌کننـده یـک دیدگاه جهان‌شناختی پنهان نیهیلیستی است که با بدبینی و انصراف از مشـارکت سیاسـی بـه دلیـل بی‌فایده بودن مشخص می‌شود. این با انتقادات از مقاومت فـردی‌شـده هـم‌خـوانی دارد، جـایی کـه بی‌توجهی نه رد کامل جامعه، بلکه عقب‌نشینی از ناکارآمدی در دستیابی به تغییـرات معنـی‌دار اسـت (Bauman, 2024: 65-68).

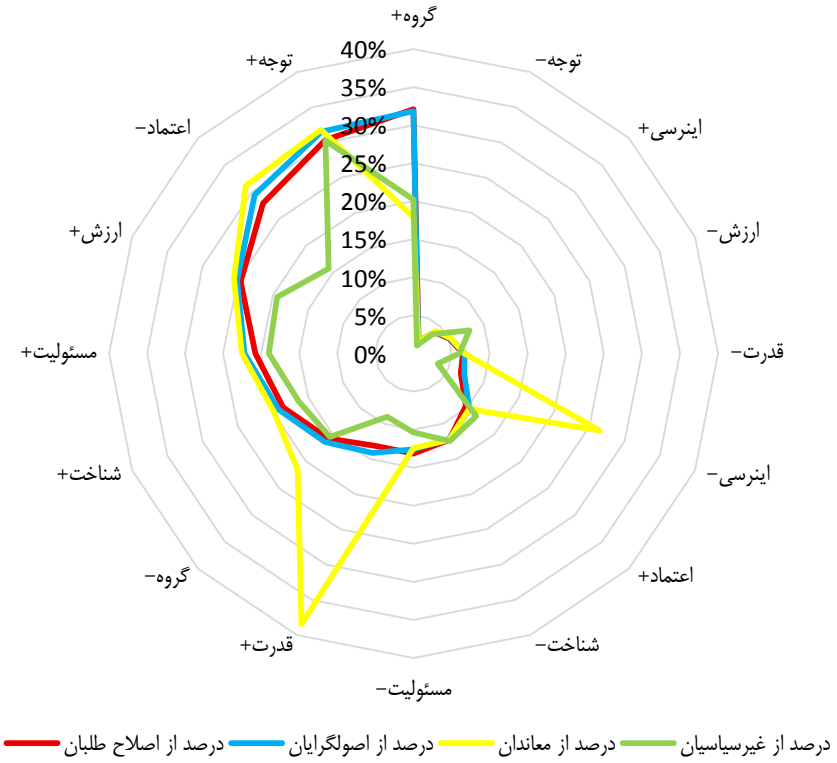


شکل ۹. مقایسه گرایش‌های سیاسی مختلف بر اساس EBMP

در میان وابستگی‌های سیاسی مختلف، افراد اصول اخلاقی تعـاملات میـان‌گروهـی را شناسـایی می‌کنند، همانطور کـه نظریـه بنیـان‌هـای اخلاقـی (Haidt, 2012: 65) پیشـنهاد مـی‌دهـد. ایـن چارچوب‌ها که توسط ایدئولوژی‌های سیاسی شکل مـی‌گیرنـد، راهنمـایی بـرای تفسـیر زمینـه‌هـای اجتماعی ارائه می‌دهند. دورکیم (۱۳۸۳) نیز به‌طور مشابه بـر ایـن نکتـه تأکیـد داشـت کـه چگونـه ارزش‌ها، فراتر از هنجارهای اجتماعی، از طریق ارتباطات پایدار به هویت گروهی تبـدیل مـی‌شـوند. چنین تبادلاتی درک افراد از اهداف مشترک و باورهای اخلاقی در حال تکامل را تصفیه می‌کند.

اخلاق مسئولیت از همدلی و انتظارات متقابل نشأت می‌گیرد و رفتارهایی را شکل می‌دهد که بـا ارزش‌های اخلاقی جمعی هماهنگ است(Davis, 1996: 138). ارزش‌های مشترک گروهی مرجعی

برای ارزیابی اقدامات فراهم می‌آورد که به افراد این امکان را می‌دهد تا رفتارها را پیش‌بینی کنند و تأثیرات گسترده‌تری که این اقدامات بر انسجام گروهی و اخلاق دارند، شناسایی کنند.



شکل ۱۰. مقایسه راداری گرایش‌های سیاسی مختلف بر اساس EBMP

الگوی به‌دست‌آمده از جدول ترجیحات اخلاقی چهار گرایش سیاسی در ایران را می‌توان در چارچوب نظریهٔ بنیادهای اخلاقی هایت تفسیر کرد. در سطح کلی، بیشترین فراوانی در همهٔ گروه‌ها به ترجیحات «گروه+»، «توجه+» و سپس «ارزش+» و «مسئولیت+» اختصاص دارد؛ امری که نشان می‌دهد صرف‌نظر از گرایش سیاسی، حساسیت به تعلق و همبستگی گروهی، دیده‌شدن و به‌رسمیت‌شناخته‌شدن، و نیز پای‌بندی به ارزش‌ها و مسئولیت‌های اخلاقی، در مرکز قضاوت‌های اخلاقی مشارکت‌کنندگان قرار دارد. در مقابل، ترجیحات مرتبط با «قدرت−»، «اینرسی+» و «توجه−» در همهٔ گروه‌ها کم‌فراوانی‌ترین گزینه‌ها هستند؛ بنابراین، نه اخلاقِ ضدقدرتِ رادیکال و نه دفاع از

سکون و حفظ وضع موجود، برای اکثریت پاسخ‌گویان محوریت ندارد و اخلاق روزمره بیشتر حول روابط اجتماعی و هنجارهای تعاملی تعریف می‌شود.

مقایسهٔ دو جریان رسمی نظام یعنی اصلاح‌طلبان و اصول‌گرایان نشان می‌دهد که «امضای اخلاقی» این دو گروه بیش از آن‌که متعارض باشد، هم‌پوشان است. در هر دو طیف، «گروه+» و «توجه+» بالاترین درصدها را دارند و پس از آن، «اعتماد−»، «ارزش+» و «مسئولیت+» قرار می‌گیرند. این الگو دلالت بر آن دارد که هر دو جریان، سیاست را اساساً میدان تعلق به یک «ما»ی جمعی و جست‌وجوی به رسمیت‌شناخته‌شدن در چارچوب نظم موجود تجربه می‌کنند و حساسیت اخلاقی‌شان معطوف به وفاداری گروهی، مراقبت از اعضا و عدالت تعاملی است. نزد هر دو گروه، ترجیحات مرتبط با «قدرت+»، «اینرسی−» و «گروه−» نسبتا پایین است؛ ازاین‌رو، اختلاف آن‌ها بیشتر در سطح برنامه‌ها و منازعات درون‌سیستمی قابل‌فهم است تا در سطح بنیادهای اخلاقی متمایز. این یافته با نتایج پژوهش‌هایی هم‌سو است که نشان می‌دهند جناح‌های رقیب درون نظام‌های سیاسی، علی‌رغم تفاوت گفتمانی، اغلب بر مجموعه‌ای مشترک از بنیادهای اخلاقی تکیه می‌کنند.

در مقابل، پروفایل اخلاقی معاندان به‌طور معناداری از سه گروه دیگر متمایز است. در این گروه، ترجیح «قدرت+» با فاصله‌ای زیاد بالاترین مقدار را دارد و به‌دنبالش «توجه+»، «اعتماد−» و «اینرسی−» قرار می‌گیرند. تمرکز هم‌زمان بر برجسته‌سازی قدرت، بی‌اعتمادی و مخالفت با سکون، نشان‌دهندهٔ آن است که قضاوت‌های اخلاقی معاندان عمدتاً حول نقد ساختارهای اقتدار، افشای فساد و ظلم ادراک‌شده و مطالبه‌ی تغییر بنیادین در نظم سیاسی شکل می‌گیرد؛ امری که می‌توان آن را فعال شدن شدید بنیاد «اقتدار/زیردستی» در قالب مخالفت با اقتدار موجود دانست. همچنین سهم بالاتر «گروه−» در میان معاندان حاکی از آن است که مرزبندی و طرد گروه‌های حاکم، بخشی از چارچوب اخلاقی آنان را تشکیل می‌دهد. این الگو شباهت زیادی به الگوهای گزارش‌شده در مطالعات تحلیل گفتمان مخالفان سیاسی در شبکه‌های اجتماعی دارد که در آن‌ها، زبان اخلاقی حول محور قدرت، بی‌اعتمادی و تغییر رادیکال سازمان می‌یابد.

پروفایل غیرسیاسی‌ها تصویری نسبتاً متفاوت ارائه می‌کند. در این گروه، گرچه «توجه+» و «گروه+» همچنان در صدر قرار دارند، اما وزن ترجیحات مرتبط با «اعتماد−»، «قدرت+» و «اینرسی−» به‌طور محسوس کمتر از گروه‌های سیاسی است. این امر نشان می‌دهد که برای افراد غیرسیاسی، سیاست میدان اصلی بروز دغدغه‌های اخلاقی نیست و اخلاق آنان بیشتر در سطح روابط نزدیک، مراقبت متقابل و هویت‌های خرد تعریف می‌شود. جالب آن‌که سهم «ارزش−» نزد

غیرسیاسی‌ها کمی بالاتر از سه گروه دیگر است؛ می‌توان این یافته را نشانه‌ای از بدبینی نسبی آنان نسبت به گفتمان‌های کلان ارزشی و ایدئولوژیک دانست که از منظر ایشان بیشتر به عرصهٔ نزاع‌های سیاسی تعلق دارد تا زندگی روزمره.

در مجموع، نتایج این مطالعه نشان می‌دهد که اگرچه در سطح کلان، بنیادهای مراقبت، انصاف و وفاداری در میان همهٔ پاسخ‌گویان فعال‌اند، اما جهت‌گیری نسبت به اقتدار، اعتماد و تغییر، «امضای اخلاقی» متمایزی برای هر یک از گرایش‌های سیاسی می‌سازد. شباهت چشم‌گیر پروفایل اصلاح‌طلبان و اصول‌گرایان و در عین حال، فاصله‌ی معنادار معاندان از هر دو، یادآور یافته‌های استرلینگ و جاست (۲۰۱۸)، روی و همکاران (۲۰۲۱) و رایترهاس و همکاران (۲۰۲۱) است که نشان می‌دهند اختلافات ایدئولوژیک، بیش از آن‌که صرفاً در سطح مواضع موضوعی بروز کند، در قالب الگوهای متفاوت فریمینگ اخلاقی و فعال‌سازی بنیادهای اخلاقی نیز بازتاب می‌یابد. به‌این‌ترتیب، تحلیل حاضر علاوه بر فراهم کردن شواهدی از کاربرد نظریهٔ بنیادهای اخلاقی در بافت ایران، پیوند معناداری با ادبیات بین‌المللی دربارهٔ گفتمان سیاسی و شبکه‌های اجتماعی برقرار می‌کند.

بحث

مقاله حاضر ترجیحات اخلاقی بیان‌شده در توییتر در جریان انتخابات ریاست‌جمهوری ایران در سال ۱۳۹۶ را بررسی کرده و روش‌های کیفی و محاسباتی را مقایسه کرده و اجتماعات کلیدی و تمایلات اخلاقی آن‌ها را تحلیل کرده است. نخست آنکه، ساختار توییتر فارسی مملو از میکرو اینفلوئنسرها و گروه‌های مرجع اجتماعی است، که در چشم‌انداز سیاسی ایران با وجود محدودیت‌های پلتفرم، شبکه‌های حیاتی را تشکیل می‌دهند. این گروه‌ها نقش مهمی در تصمیم‌سازی‌های سیاسی ایفا می‌کنند.

دوم، یافته‌ها نشان می‌دهند که الگوهای رفتاری کاربران بیشتر تحت تأثیر ترجیحات اخلاقی قرار دارند تا گرایش سیاسی آن‌ها. به‌ویژه، اصول‌گرایان و اصلاح‌طلبان، دو جناح متعارض سنتی، الگوهای مشابهی از ترجیحات اخلاقی را نشان می‌دهند.

سوم، هشت ترجیح اخلاقی کلیدی (گروه، شناخت، ارزش، اعتماد، مسئولیت، توجه، قدرت و اینرسی) شناسایی‌شده در تصمیم‌گیری ایرانیان تثبیت شدند و مبنای تأیید نظریه هشت ترجیح اخلاقی دوقطبی(EBMP) را تشکیل می‌دهند.

چهارم، نتایج ما نشان می‌دهند که رفتار سیاسی در توییتر به‌طور عمیق ریشه در مکانیسم‌های جامعه‌شناختی دارد. با استفاده از نظریه مبانی اخلاقی (Haidt, 2012: 67-94; Interian, 2024:

(8، مشاهده می‌کنیم که گروه‌های سیاسی متعارض، ارزش‌های اخلاقی مختلفی را اولویت‌بندی می‌کنند که از طریق تعاملات میان‌گروهی که هویت‌های جمعی را شکل می‌دهند، انتقال و تقویت می‌شود(Durkheim, 2004: 38-57). دینامیک‌های گروهی که به همگرایی و کاهش خودمختاری می‌انجامند(Tajfel & Turner, 1979: 33-47)، پایداری و مقاومت مشاهده‌شده در درون جناح‌های سیاسی را پشتیبانی می‌کند. همچنین دینامیک‌های قدرت و اینرسی، تنش میان نیروهای تغییر و مدافعان وضعیت موجود را توضیح می‌دهند (Foucault,2020: 195-229) و (Kohn, 1989: 83-115). در این چارچوب، جناح‌های سیاسی متمایزی که در توییتر شناسایی کرده‌ایم (اصلاح‌طلبان، اصول‌گرایان، مخالفان نظام و کاربران غیرسیاسی) این فرایندهای جامعه‌شناختی را به شیوه‌های متفاوتی به نمایش می‌گذارند که تحت تأثیر مشوق‌های الگوریتمی توییتر است(Gillespie, 2018: 17-45). همبستگی اصلاح‌طلبان در همبستگی نمایشی و مشارکت عقلانی، کارکرد مرزگذاری نمادین اصول‌گرایان، تحریک الگوریتمی معاندان و بی‌توجهی کنایه‌آمیز کاربران غیرسیاسی همه تطبیق‌هایشان را با بازار اخلاقی پلتفرم نشان می‌دهد که همبستگی‌های سنتی را به خُرده‌عموم‌های رقیب تبدیل می‌کند(Bauman, 2024: 25-48) و (Papacharissi, 2015: 9-34; Cheney-Lippold, 2017: 15-18).

در نهایت، ترجیحاتی مانند اعتماد، گروه، مسئولیت، توجه، شناخت و ارزش بر گفتمان سیاسی تسلط دارند. قطب‌های مثبت ترجیحات گروه و توجه بیش از ۳۰درصد از پست‌ها را تشکیل می‌دهند، در حالی که قطب منفی اعتماد و قطب‌های مثبت ارزش و مسئولیت بیش از ۲۰درصد از پست‌ها را به خود اختصاص می‌دهند. مدل EBMP که از رویکرد استقرایی و مبتنی بر داده‌ها استخراج شده بود (Haddadi et al, 2023: 225)، در یک دسته‌بندی سیستماتیک از ترجیحات اخلاقی پیچیده تأیید شد. این چارچوب، تلاش می‌کند فاصله‌ها را با نظریه‌های اخلاقی موجود در جامعه‌شناسی و روان‌شناسی پر کرده و چشم‌انداز جدیدی برای درک رفتار اخلاقی در زمینه‌های اجتماعی و سیاسی ارائه می‌دهد.

نتیجه‌گیری

مشارکت‌های روش‌شناختی این مقاله به سه جنبه تقسیم می‌شوند. اولاً، روش محاسباتی به ضریب توافقی دست یافته است که قابل مقایسه با توافق انسانی است و به‌طور مؤثر متون مکالماتی فارسی را به متغیرهای روان‌شناسی اجتماعی دسته‌بندی کرده است. دوم، واژه‌نامه ترکیبی متون و زبان‌شناسی دقت مجموعه داده‌های آموزشی را به‌طور چشمگیری بهبود بخشیده است. سوم،

تکنیک‌های یادگیری عمیق برای استخراج ترجیحات از متن دقت بیشتری نسبت به روش‌های سنتی یادگیری ماشین نشان دادند.

پیشنهادها

تحقیقات آینده می‌تواند تغییرات در ترجیحات اخلاقی در انتخابات‌های بعدی که دوقطبی شکل گرفته باشد، را بررسی کند و مصاحبه‌های کیفی، نظرسنجی‌های آنلاین و جلسات خبرگانی را برای تأیید و گسترش این یافته‌ها به کار گیرد. دیدگاه‌های اضافی شامل استفاده از روش‌های خوشه‌بندی بدون داده‌های آموزشی و مقایسه این نتایج با مدل‌های موجود است. گسترش داده‌های آموزشی و استفاده از مدل‌های زبانی بزرگ[1] و تکنولوژی‌های هوش مصنوعی مولد می‌تواند تجزیه‌وتحلیل ترجیحات اخلاقی را ارتقا دهد و بینش‌های عمیق‌تری در تصمیم‌گیری‌های اخلاقی و سیاسی ارائه دهد.

مدل هشت ترجیح اخلاقی دوقطبی(EBMP) یک چارچوب استنتاجی برای درک ترجیحات اخلاقی، به‌ویژه در جامعه ایرانی، ارائه می‌دهد و پتانسیل قابل توجهی برای تحقیقات آینده در زمینه‌های جامعه‌شناسی و فرهنگی دارد. در حالی که این مدل با نظریه‌هایی مانند نظریه مبانی اخلاقی و سرمایه فرهنگی بوردیو هم‌راستا است، EBMP بر تقاطع عوامل فرهنگی، عاطفی و شناختی تأکید می‌کند که آن را از چارچوب‌های موجود متمایز می‌سازد. این مدل از شکاف‌های موجود در چشم‌انداز نظریه سیاست آغاز کرده و بینش‌های جدیدی در مورد تعامل میان ترجیحات اخلاقی و زمینه‌های سیاسی‌اجتماعی فراهم می‌آورد.

در سطح سیاست‌گذاری پژوهشی، در مطالعات میان‌رشته‌ای که داده‌های بزرگ، روش‌های محاسباتی و مفاهیم نظری علوم اجتماعی را تلفیق می‌کنند، ضروری است تمامی ابعاد اعتبار و پایایی (از جمله روایی نظری، روایی درونی و بیرونی، پایایی سنجش و پایایی الگوریتمی) به‌صورت شفاف، صریح و مستند گزارش شوند تا امکان داوری علمی و بازتولید نتایج فراهم شود. از منظر جامعه‌شناختی نیز توصیه می‌شود سیاست‌های پژوهشی و حمایتی، فراتر از تمرکز صرف بر «گرایش‌های سیاسی»، به مطالعه نظام‌مند «ترجیحات اخلاقی و ارزشی» کنش‌گران در فضاهای آنلاین و آفلاین اولویت دهند؛ چرا که ترجیحات اخلاقی، لایه‌ای عمیق‌تر از سلیقه سیاسی را منعکس می‌کند و می‌تواند در فهم شکاف‌ها، قطبی‌شدن‌ها و امکان‌های گفت‌وگوی سازنده در جامعه، مبنای تحلیلی دقیق‌تر و غنی‌تری فراهم آورد.

1 LLM

تعارض منافع

مقاله فاقد هرگونه تعارض منافعی می باشد.

References

Abdollahian, H., & Kermani, H. (2020). Networked Publics in Persian Twitter; Analysis of Network Clusters on Twitter. Journal of New Media Studies, 6(22), 70–100. https://doi.org/10.22054/nms.2021.44855.784 (In Persian)

Aggarwal, C. C. (2011). Social network data analytics. Springer Science & Business Media. https://doi.org/10.1007/978-1-4419-8462-3

Alizadeh, M., Cioffi-Revilla, C., & Crooks, A. (2017). Generating and analyzing spatial social networks. Computational and Mathematical Organization Theory, 23(4), 362–390. https://doi.org/10.1007/s10588-016-9232-2

Altemeyer, B. (1996). The authoritarian specter. Harvard University Press.

Arabi, A., Haddadi, A., & Karami, F. (2023). Political Participation, Ethnicity, and Media Space: A Look at Online and Offline Participation in Fars Province. Journal of Social Development. https://doi.org/10.22055/qjsd.2023.18276 (In Persian)

Asadi, M. (2018). Improving the efficiency of online event detection in Twitter textual data streams considering retweet features [Master's thesis, University of Science and Culture]. (In Persian)

Atari, M., Graham, J., & Dehghani, M. (2020). Foundations of morality in Iran. Evolution and Human Behavior, 41(5), 367–384. https://doi.org/10.1016/j.evolhumbehav.2020.07.014

Bastian, M., Heymann, S., & Jacomy, M. (2009). Gephi: An open-source software for exploring and manipulating networks. Proceedings of the International AAAI Conference on Web and Social Media, 3(1), 361–362. https://doi.org/10.1609/icwsm.v3i1.13937

Bauman, Z. (2024). Liquid Modernity (H. Mohamadzadeh, Trans.). Naqde Farhang. (In Persian)

Blondel, V. D., Guillaume, J. L., Lambiotte, R., & Lefebvre, E. (2008). Fast unfolding of communities in large networks. Journal of Statistical Mechanics: Theory and Experiment, 2008(10), P10008. https://doi.org/10.1088/1742-5468/2008/10/P10008

Bojanowski, P., Grave, E., Mikolov, T., Puhrsch, C., & Joulin, A. (2017). Enriching word vectors with subword information. Transactions of the Association for Computational Linguistics, 5, 135–146. https://doi.org/10.1162/tacl_a_00051

Bollen, J., Mao, H., & Zeng, X. (2011). Twitter mood predicts the stock market. Journal of Computational Science, 2(1), 1–8. https://doi.org/10.1016/j.jocs.2010.12.007

Borgatti, S. P., & Foster, P. C. (2003). The network paradigm in organizational research: A review and typology. Journal of Management, 29(6), 991–1013. https://doi.org/10.1016/S0149-2063(03)00087-4

Bourdieu, P. (2018). Distinction: A Social Critique of the Judgement of Taste (H. Chavoshian, Trans.). Sales. (In Persian)

Cano-Marin, E., Mora-Cantallops, M., & Sánchez-Alonso, S. (2023). Twitter as a predictive system: A systematic literature review. Journal of Business Research, 157, 113561. https://doi.org/10.1016/j.jbusres.2022.113561

Chen, T., & Guestrin, C. (2016). XGBoost: A scalable tree boosting system. In Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (pp. 785–794). ACM. https://doi.org/10.1145/2939672.2939785

Cheney-Lippold, J. (2017). We are data: Algorithms and the making of our digital selves. New York University Press.

Cheng, C. Y., & Hale, S. A. (2025). Beyond English: Evaluating Automated Measurement of Moral Foundations in Non-English Discourse with a Chinese Case Study. arXiv preprint arXiv:2502.02451. https://doi.org/10.48550/arXiv.2502.02451

Cohen, R., Tsang, A., Vaidyanathan, K., & Zhang, H. (2016). Analyzing opinion dynamics in online social networks. Big Data & Information Analytics, 1(4), 279–298. https://doi.org/10.3934/bdia.2016011

Conover, M., Ratkiewicz, J., Francisco, M., Goncalves, B., Menczer, F., & Flammini, A. (2011). Political polarization on Twitter. Proceedings of the International AAAI Conference on Web and Social Media, 5(1), 89–96. https://doi.org/10.1609/icwsm.v5i1.14126

Couldry, N. (2012). Media, society, world: Social theory and digital media practice. Polity Press.

Davis, M. H. (1996). Empathy: A social psychological approach. Westview Press.

Devlin, J., Chang, M. W., Lee, K., & Toutanova, K. (2019). BERT: Pre-training of deep bidirectional transformers for language understanding. In Proceedings of NAACL-HLT 2019. https://doi.org/10.18653/v1/N19-1423

Durkheim, E. (2004). The Elementary Forms of Religious Life (B. Parham, Trans.). Nashr-e Markaz. (In Persian)

Festinger, L. (1957). A theory of cognitive dissonance. Stanford University Press.

Fiske, J. (1994). Media matters: Race and gender in U.S. politics. University of Minnesota Press.

Foucault, M. (2020). Discipline and Punish: The Birth of the Prison (A. Jahandideh & N. Sorkhosh, Trans.). Nashr-e Ney. (In Persian)

Frankena, W. K. (2013). Ethics (H. Sadeghi, Trans.). Taha. (In Persian)

Gillespie, T. (2018). Custodians of the internet: Platforms, content moderation, and the hidden decisions that shape social media. Yale University Press. https://doi.org/10.12987/9780300235029

Glaser, B. G., & Strauss, A. L. (1967). The discovery of grounded theory: Strategies for qualitative research. Aldine Publishing.

Graham, J., Haidt, J., & Nosek, B. A. (2009). Liberals and conservatives rely on different sets of moral foundations. Journal of Personality and Social Psychology, 96(5), 1029–1046. https://doi.org/10.1037/a0015141

Graham, J., Nosek, B. A., Haidt, J., Iyer, R., Koleva, S., & Ditto, P. H. (2011). Mapping the moral domain. Journal of Personality and Social Psychology, 101(2), 366–385. https://doi.org/10.1037/a0021847

Gramsci, A. (2024). Prison Notebooks (A. Masoumbeigi, Trans.). Nashr-e Charkh. (In Persian)

Greene, J. D. (2003). From neural “is” to moral “ought”: What are the moral implications of neuroscientific moral psychology? Nature Reviews Neuroscience, 4(10), 846–850. https://doi.org/10.1038/nrn1224

Guido, R., Ferrisi, S., Lofaro, D., & Conforti, D. (2024). An overview on the advancements of support vector machine models in healthcare applications: A review. Information, 15(4), 235. https://doi.org/10.3390/info15040235

Haddadi, A., Etemadi, S. M., & Hatami, J. (2023). Discovering Moral Preferences on Twitter Social Media (Qualitative Study of the 2017 Election). Cultural Studies & Communication, 19(70), 203–240. https://doi.org/10.22034/jcsc.2022.552920.2565 (In Persian)

Hagberg, A. A., Swart, P. J., & Sculley, D. (2008). NetworkX: Network analysis with Python. Los Alamos National Laboratory.

Haidt, J. (2012). The righteous mind: Why good people are divided by politics and religion. Pantheon Books.

Haidt, J., & Graham, J. (2007). When morality opposes justice: Conservatives have moral intuitions that liberals may not recognize. Social Justice Research, 20(1), 98–116. https://doi.org/10.1007/s11211-007-0034-z

Haidt, J., & Joseph, C. (2004). Intuitive ethics: How innately prepared intuitions generate culturally variable virtues. Daedalus, 133(4), 55–66. https://doi.org/10.1162/0011526042365555

Hetherington, M. J. (1998). The political relevance of political trust. American Political Science Review, 92(4), 791–808. https://doi.org/10.2307/2586304

Himelboim, I., McCreery, S., & Smith, M. (2013). Birds of a feather tweet together: Integrating network and content analyses to examine cross-ideology exposure on Twitter. Journal of Computer-Mediated Communication, 18(2), 154–174. https://doi.org/10.1111/jcc4.12001

Hirschman, A. O. (2018). The Passions and the Interests: Political Arguments for Capitalism before its Triumph (M. Maljo, Trans.). Shiraze-ye Ketab-e Ma. (In Persian)

Hogg, M. A. (2001). A social identity theory of leadership. Personality and Social Psychology Review, 5(3), 184–200. https://doi.org/10.1207/S15327957PSPR0503_1

Inglehart, R. (2003). Modernization and postmodernization: Cultural, economic, and political change in 43 societies. Princeton University Press. https://doi.org/10.2307/j.ctv10vm2ns

Interian, R. (2024). A political radicalization framework based on Moral Foundations Theory. arXiv. https://doi.org/10.48550/arXiv.2406.19878

Khazraee, E. (2019). Mapping the political landscape of Persian Twitter: The case of 2013 presidential election. Big Data & Society, 6(1), 1–15. https://doi.org/10.1177/2053951719835232

Kim, S. W., & Gil, J. M. (2019). Keyword extraction for sentiment analysis using tf-idf. International Journal of Advanced Computer Science and Applications, 10(6), 1–6.

Kingma, D. P., & Ba, J. (2014). Adam: A method for stochastic optimization. International Conference on Learning Representations (ICLR). https://doi.org/10.48550/arXiv.1412.6980

Kohn, M. L. (1989). Class and conformity: A study in values. University of Chicago Press.

Krippendorff, K. (2019). Content analysis: An introduction to its methodology (4th ed.). Sage Publications. https://doi.org/10.4135/9781071878781

Mead, G. H. (2021). Mind, Self, and Society (M. Saffar, Trans.). SAMT. (In Persian)

Molla Norouzi, M. (2019). A Method for Sentiment Analysis on Twitter as a Social Network at Big Data Scale [Master's thesis, University of Science and Culture]. (In Persian)

Moniri, S., Schlosser, T., & Kowerko, D. (2024). Investigating the challenges and opportunities in Persian language information retrieval through standardized data collections and deep learning. Computers, 13(8), 212. https://doi.org/10.3390/computers13080212

Neuendorf, K. A. (2017). The content analysis guidebook (2nd ed.). Sage Publications. https://doi.org/10.4135/9781071802878

Nip, J. Y. M., & Berthelier, B. (2024). Social media sentiment analysis. Encyclopedia, 4(4), 1590–1598. https://doi.org/10.3390/encyclopedia4040104

Papacharissi, Z. (2015). Affective publics: Sentiment, technology, and politics. Oxford University Press. https://doi.org/10.1093/acprof:oso/9780199999736.001.0001

Paris, C., Thomas, P., & Wan, S. (2012). Differences in language and style between two social media communities. Proceedings of the International AAAI Conference on Web and Social Media, 6(1), 539–542. https://doi.org/10.1609/icwsm.v6i1.14307

Pennebaker, J. W., Booth, R. J., & Francis, M. E. (2015). Linguistic Inquiry and Word Count: LIWC2015. Pennebaker Conglomerates.

Pennington, J., Socher, R., & Manning, C. D. (2014). GloVe: Global vectors for word representation. Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP), 1532–1543. https://doi.org/10.3115/v1/D14-1162

Reiter-Haas, M., Kopeinik, S., & Lex, E. (2021). Studying moral-based differences in the framing of political tweets. In Proceedings of the Fifteenth International AAAI Conference on Web and Social Media (ICWSM 2021) (pp. 1085–1089). AAAI Press. https://doi.org/10.1609/icwsm.v15i1.18135

Roy, S., Pacheco, M. L., & Goldwasser, D. (2021). Identifying morality frames in political tweets using relational learning. In Proceedings of the 2021 Conference on Empirical Methods in Natural Language Processing (pp. 9939–9958). Association for Computational Linguistics. https://doi.org/10.18653/v1/2021.emnlp-main.783

Roy, Sisir. (2016). Decision Making and Modelling in Cognitive Science. Springer. https://doi.org/10.1007/978-81-322-3622-1

Saleh, S., & Haddadi, A. (2022). Examining the Conditions for the Formation of the Public Sphere on Twitter Social Media. Journal of New Media Studies, 8(29), 239–276. https://doi.org/10.22054/nms.2022.51531.958 (In Persian)

Schwartz, S. H. (1992). Universals in the content and structure of values: Theoretical advances and empirical tests in 20 countries. In M. P. Zanna (Ed.), Advances in experimental social psychology (Vol. 25, pp. 1–65). Academic Press. https://doi.org/10.1016/S0065-2601(08)60281-6

Scott, J. (2000). Social network analysis: A handbook (2nd ed.). Sage Publications.

Simon, H. A. (1972). Theories of bounded rationality. In C. B. McGuire & R. Radner (Eds.), Decision and organization (pp. 161–176). North-Holland.

Smith, A. (2017). The theory of moral sentiments. CreateSpace Independent Publishing Platform.

Stecker, M., & Hopp, F. R. (2026). Moral Foundation Measurements Fail to Converge on Multilingual Party Manifestos. Political Analysis, 34(2), 166–187. https://doi.org/10.1017/pan.2025.10011

Stemler, S. (2001). An overview of content analysis. Practical Assessment, Research, and Evaluation, 7(17), 1–10. https://doi.org/10.7275/z6fm-2e34

Sterling, J., & Jost, J. T. (2018). Moral discourse in the Twitterverse: Effects of ideology and political sophistication on language use among U.S. citizens and members of Congress. Journal of Language and Politics, 17(2), 195–221. https://doi.org/10.1075/jlp.17034.ste

Tajfel, H., & Turner, J. C. (1979). An integrative theory of intergroup conflict. In W. G. Austin & S. Worchel (Eds.), The social psychology of intergroup relations (pp. 33–47). Brooks/Cole.

Tajfel, H., & Turner, J. C. (2004). The social identity theory of intergroup behavior. In J. T. Jost & J. Sidanius (Eds.), Political psychology: Key readings (pp. 276–293).

Tsamardinos, I., Greasidou, E., & Borboudakis, G. (2018). Bootstrapping the out-of-sample predictions for efficient and accurate cross-validation. Machine Learning, 107, 1895–1922. https://doi.org/10.1007/s10994-018-5714-4

Tufekci, Z., & Wilson, C. (2012). Social media and the decision to participate in political protest: Observations from Tahrir Square. Journal of Communication, 62(2), 363–379. https://doi.org/10.1111/j.1460-2466.2012.01629.x

Turner, J. H. (2018). The Problem of Emotions in Societies (M. R. Hassani, Trans.). Elmi va Farhangi. (In Persian)

Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. A., Kaiser, Ł., & Polosukhin, I. (2017). Attention is all you need. Proceedings of NeurIPS 2017. https://doi.org/10.48550/arXiv.1706.03762

Zhang, Z., Hou, R., & Yang, J. (2020). Detection of social network spam based on improved extreme learning machine. IEEE Access, 8, 112003–112014. https://doi.org/10.1109/ACCESS.2020.3002940

Zhu, X. (2009). Semi-supervised learning literature survey. Computer Sciences Technical Report 1530, University of Wisconsin–Madison.