The Impact of Social Media Algorithms on Public Opinion

An integrative research synthesis and proposed evaluation model

Om R Anjikar | Akash Rathod | Vedant Lihtikar | Yash Gabhud

Master of Computer Applications(MCA)

D. Y. Patil Institute of Master of Computer Application and Management, Akurdi, Pune, Maharashtra, India

Abstract

Social media recommendation systems play an important role in deciding what information users see and how often they see it. These systems use ranking algorithms, personalisation, and user engagement patterns to recommend content based on individual interests. Previous research has shown that these mechanisms can influence the type of information people consume, contribute to the spread of misinformation, shape political opinions, affect consumer decisions, and impact users' well-being. However, these effects are often examined separately, and simply showing users certain content does not necessarily mean that their opinions will change over time.

To address this issue, this paper proposes a multi-objective machine learning framework that evaluates social media recommendations using a Civic-Impact Balanced Recommendation Score (CIBRS). The proposed score considers several factors, including content relevance, information quality, diversity of viewpoints, misinformation risk, and potential effects on user well-being. Based on these factors, a constrained re-ranking approach is introduced to recommend useful content while reducing repeated exposure to similar opinions and limiting the promotion of potentially harmful information.

This research does not present new experimental findings from individual participants. Instead, the numerical charts are based on values reported in existing research studies, while the proposed CIBRS framework is presented as a new concept that requires further testing and validation. The overall analysis suggests that the influence of recommendation algorithms is more clearly visible in content exposure and user engagement than in long-term changes in political opinions.

The main contribution of this study is a practical framework for evaluating and improving social media recommendation systems. Rather than focusing only on increasing user engagement, the proposed approach aims to encourage more balanced, diverse, reliable, and responsible information consumption.

Keywords: social media algorithms; machine learning; public opinion; recommendation systems; information diversity; misinformation; CIBRS; responsible AI

1. Introduction

Social media platforms have become an important part of our daily lives, influencing how people access, share, and understand information about the world. Most platforms use recommendation and ranking algorithms to decide what content appears in users' feeds. These algorithms consider different factors, such as previous clicks, likes, shares, watch time, and personal interests, to suggest content that users are likely to engage with. Although this makes it easier to discover information, it can also limit the variety of opinions and perspectives users are exposed to over time. As a result, concerns such as filter bubbles, echo chambers, the spread of misinformation, political polarisation, and the lack of transparency in platform algorithms have received increasing attention.

However, the influence of recommendation algorithms is not the same for every user or situation. A systematic review published in 2026 highlights that these algorithms have a clear influence on the information people see and how frequently certain content appears. However, their impact on long-term political opinions is less consistent and can depend on several factors, including the platform being used, users' social networks, personal identities, the credibility of information sources, and the duration

of exposure. This suggests that changing the information people see does not necessarily mean changing what they think or believe.

To better understand these effects, this paper brings together different research areas, including political communication, consumer decision-making, mental health, and machine learning. It focuses on how recommendation algorithms influence people's information consumption and explores ways to make these systems more responsible. The central idea of this research is that social media recommendation systems should not be judged only by how much engagement or relevance they generate. They should also be evaluated based on the quality of information they recommend, the diversity of viewpoints they provide, and the possible risks they may create for users.

2. Research Problem, Questions, and Objectives

Most social media platforms use engagement-based algorithms to decide what content users see. These algorithms mainly focus on factors such as likes, shares, comments, and watch time to recommend content that attracts attention. However, content that receives high engagement is not always useful or reliable. Sometimes, these systems may repeatedly recommend similar information, promote misleading content, encourage polarised discussions, or expose users to content that can negatively affect their mental well-being.

Another major concern is that existing research generally studies these effects separately rather than examining them together. This makes it difficult to understand how recommendation algorithms create trade-offs between different areas, including public opinion, information quality, economic decisions, and user well-being. Therefore, there is a need for a more balanced approach that evaluates recommendation systems beyond user engagement and considers their wider impact on individuals and society.

Research Questions

RQ1: How do engagement-based ranking algorithms and personalised recommendations influence the diversity and reliability of information users receive?

RQ2: How can machine learning techniques be used together to evaluate content quality, identify misinformation risks, measure viewpoint diversity, and detect potential behavioural risks?

RQ3: Can a multi-objective recommendation approach increase the diversity of information shown to users without significantly affecting the relevance of recommended content?

RQ4: How can the effects of recommendation algorithms be regularly monitored and evaluated at both individual user and overall platform levels?

Research Objectives

The main objectives of this research are:

1. To understand how social media recommendation algorithms influence information consumption, public opinion, and other important social outcomes.
2. To identify measurable factors that can help evaluate the quality, diversity, and responsibility of recommended content.
3. To propose a transparent scoring system and re-ranking method that considers multiple factors rather than focusing only on user engagement.
4. To develop a clear and reproducible validation approach for evaluating the performance of the proposed framework.

5. To examine the limitations, ethical concerns, and necessary safeguards involved in developing more responsible recommendation systems.

3. Literature Synthesis

3.1 Political Opinion and Campaigns

Hasibuan et al. (2024) explain how social media algorithms that focus mainly on user engagement can influence the way people understand political information. When users repeatedly receive content that supports their existing beliefs, it can create echo chambers and increase the visibility of misleading information, especially during political campaigns. Similarly, Joselin et al. (2025) discuss issues such as selective exposure, confirmation bias, political microtargeting, and the use of bots on social media platforms. They also highlight the importance of transparency and media literacy in helping users understand the information they consume. Although these studies explain the social and democratic challenges created by recommendation algorithms, they do not provide a single predictive framework that evaluates these effects together.

3.2 Public Decisions and Economic Behaviour

Zafar and Saleem (2025) explore how social media recommendation algorithms influence consumer preferences, purchasing decisions, targeted advertisements, and influencer marketing. Their research also considers how economic uncertainty can affect people's responses to online content. According to their reported platform engagement data, TikTok recorded the highest average likes, shares, and comments among the five platforms studied, while Twitter/X showed comparatively high sharing activity. These findings suggest that different social media platforms influence user engagement in different ways. Therefore, platform features and content types should also be considered when studying the impact of recommendation algorithms on consumer behaviour.

3.3 Mental Health and Behavioural Signals

More and Daund (2026) conducted a survey involving 100 participants from Nashik to understand the relationship between social media usage, user behaviour, and mental well-being. Their study examined factors such as compulsive checking of social media, daily screen time, awareness of filter bubbles, anxiety, and stress after using social media. The researchers reported several interesting relationships, including a correlation of $r = 0.54$ between the urge to check social media and anxiety, $r = 0.41$ between daily usage hours and post-use stress, $r = 0.33$ between filter-bubble awareness and anxiety, and $r = 0.47$ between the urge to check and active engagement.

The study also used Logistic Regression to identify users who may be at higher risk. The reported model achieved 89% accuracy, with a high-risk recall of 0.96 and an F1-score of 0.85. These findings show how machine learning can be used to examine behavioural patterns associated with social media use. However, it is important to note that these statistical results belong to the original study by More and Daund and are not newly generated findings of the present research.

3.4 Systematic Evidence and Mixed Effects

Bajuwaiber (2026) conducted a systematic review covering 36 studies using the PICO and SPIDER research frameworks. The review examined how recommendation algorithms influence information exposure and political opinions across different contexts. It reported findings from cross-national audits showing greater amplification of

mainstream right-leaning content in six out of seven countries examined in one Twitter audit.

However, the review also points out that the relationship between recommendation algorithms and changes in political attitudes is not always consistent. Their effects can vary depending on users' personal choices, social networks, political environments, and research methods. The study therefore highlights that algorithms should not be considered the only reason behind political polarisation. Other social and individual factors also play an important role in shaping people's opinions.

3.5 Public Decision-Making and Accountability

Shekhar and Kumar (2025) discuss several concerns related to social media recommendation systems, including filter bubbles, emotionally engaging content, limited transparency, and the lack of user control over personalised feeds. They suggest that platforms should take greater responsibility for how their algorithms select and recommend information.

Their recommendations include regular algorithmic audits, giving users more control over their personalised feeds, improving digital literacy, and introducing suitable regulatory safeguards. These suggestions highlight the importance of making recommendation systems more transparent and responsible. Providing users with diverse information, allowing greater control over recommendations, and maintaining human oversight can help address some of the challenges associated with algorithm-driven content.

4. Research Gap and Novelty

Research Gap

Existing research shows that the impact of social media recommendation algorithms is spread across several different areas. Some studies focus on political opinions, while others examine consumer behaviour, mental health, misinformation, and user engagement. Since these topics are mostly studied separately, it becomes difficult to understand how recommendation algorithms affect all these areas together.

The 2026 systematic review also highlights an important difference between what users see and what they actually believe. The influence of recommendation algorithms on information exposure is more consistent, but their impact on long-term changes in people's opinions is less clear.

Therefore, the main research gap is not simply the need for another machine learning model to classify or detect content. There is a need for a common evaluation framework that considers multiple factors together, including content relevance, information reliability, diversity of opinions, and potential risks to users. Such a framework should also be transparent enough to explain how different factors influence the final recommendation.

Novel Contribution

To address this research gap, this paper proposes a new framework called the Civic-Impact Balanced Recommendation Score (CIBRS) It is designed as a multi-objective scoring approach that evaluates individual content items as well as complete sets of recommendations shown in a user's feed.

Unlike traditional recommendation systems that mainly focus on user engagement, the proposed CIBRS considers additional factors that are important for responsible content recommendation. It reduces the score of content with poor information quality, repeated presentation of similar viewpoints, and signals associated with potentially harmful user behaviour.

The main idea behind CIBRS is to encourage recommendation systems to provide content that is not only relevant to users but also informative, diverse, and responsible. At this stage, CIBRS is a proposed research framework rather than a fully tested or validated production algorithm. Further experiments and evaluations will be required to measure its effectiveness in real-world situations.

5. Proposed Framework and Machine-Learning Method

5.1 Inputs and Feature Groups

The proposed system uses different types of information to understand users' interests and evaluate the content recommended to them. These inputs are divided into four main groups:

(a) Content Features: These include the actual text of a post, its source, topic, claims, and publication details. They help the system understand what the content is about and assess its quality.

(b) User-Context Features: These include users' stated interests, optional interaction history, and personal preferences for viewing different types of information. These details help the system recommend content that matches users' interests.

(c) Exposure Features:These features examine the type of content users have already seen. They include repeated topics, the distribution of different viewpoints, and the concentration of information from particular sources. This helps identify whether users are continuously receiving similar information.

(d) Safety Signals:These include the estimated likelihood of misinformation and indicators of excessive social media use. They help the system identify potentially harmful recommendation patterns.

The proposed framework does not attempt to identify or store sensitive mental-health diagnoses. Any well-being-related information will be limited to voluntary, non-diagnostic user feedback or general interaction patterns. Such information will only be used with the user's consent.

5.2 Machine-Learning Components

The proposed framework combines multiple machine-learning techniques to evaluate recommended content from different perspectives.

Text Classification: A text classifier or transformer-based model is used to identify the topic and viewpoints expressed in a particular post.

Misinformation Detection: A claim-verification process estimates the possibility of misinformation by comparing claims with reliable sources and supporting evidence. Human review is also included for claims that require additional verification.

Behavioural-Risk Estimation: Models such as Logistic Regression and Random Forest are used as baseline methods to identify general behavioural-risk patterns. These models are selected because their results can be interpreted and compared more easily.

Diversity-Aware Re-ranking: After evaluating the candidate content, a re-ranking process selects a collection of posts that provides a better balance of relevance, information quality, and viewpoint diversity.

The framework begins with relatively simple and interpretable models so that their performance can be properly understood. More advanced models can be introduced later

for comparison. However, human fact-checking remains important, especially when dealing with controversial claims or information that depends heavily on context.

5.3 CIBRS Definition

To evaluate recommended content more fairly, this research introduces the Civic-Impact Balanced Recommendation Score (CIBRS). Instead of considering only how much attention a post receives, the proposed score evaluates five important factors.

For a user u and a candidate content item i, the score is defined as:

$$CIBRS(u,i) = wR \times Rel(u,i) + wQ \times Qual(i) + wD \times Div(u,i) - wM \times Mis(i) - wH \times Risk(u,i)$$

The five components are explained below:

1. Rel (Relevance): Measures how closely the recommended content matches the user's interests, with values ranging from 0 to 1.
2. Qual (Information Quality): Evaluates the reliability of the source and the quality of the claims presented in the content.
3. Div (Diversity): Measures how much a particular post contributes to introducing a different topic or viewpoint.
4. Mis (Misinformation Risk): Estimates the likelihood that the content contains misleading or unsupported information.
5. Risk (Behavioural Risk): Represents a non-diagnostic estimate of possible behavioural risks associated with the content or interaction pattern.

The weights determine how much importance is given to each factor. They must be non-negative and their total should equal 1.

$$wR + wQ + wD + wM + wH = 1$$

These weights should be decided before the main evaluation and tested under different conditions. They should not be adjusted simply to increase user engagement.

5.4 Set-Level Diversity

Evaluating individual posts is not enough to understand the overall quality of a social media feed. Even if every recommended post is relevant, users may still receive similar information repeatedly.

To address this issue, the proposed framework also evaluates diversity across the complete set of recommended posts.

Let S represent a feed containing k recommended items. Let p_v represent the proportion of items belonging to a particular viewpoint or topic group.

The diversity of the feed can be measured using Shannon Diversity:

$$H(S) = -\Sigma_v \, p_v \ln(p_v)$$

To make the diversity score easier to compare, it can be normalised:

$$H_n(S) = H(S) / \ln(V)$$

Here, V represents the number of different viewpoint or topic groups present in the feed, provided that $V > 1$.

The framework also introduces a repetition penalty to identify situations where one particular viewpoint or topic appears too frequently.

$$Rep(S) = \Sigma_v \max(0, p_v - \tau)^2$$

Here, τ represents a predefined threshold that determines when the concentration of a particular viewpoint becomes excessive.

Using these measurements, the overall feed objective is defined as:

$$J(S) = \Sigma_{i \in S} CIBRS(u,i) + \lambda H_n(S) - \mu Rep(S)$$

The system aims to select content that achieves a good balance between individual relevance and overall diversity. The parameters λ and μ control the importance given to diversity and repetition penalties.

The final selection must also satisfy predefined source-quality and safety requirements. This ensures that increasing diversity does not come at the cost of recommending unreliable or potentially harmful information.

5.5 Model Training and Evaluation

The proposed machine-learning models will be trained using labelled datasets collected with appropriate consent and recorded publication timestamps. To make the evaluation more reliable, the dataset will be divided into training, validation, and testing groups at the user level. This helps prevent information from the same users appearing across different groups and reducing the reliability of the results.

The framework will use stratified cross-validation to maintain a balanced distribution of classes during training. Additional calibration checks and time-based testing will be performed to understand how well the models work on newer data.

To measure the effectiveness of the proposed approach, three recommendation strategies will be compared:

- Approach A – Engagement-Only Ranking :Recommends content mainly according to user engagement signals such as likes, shares, comments, and watch time.
- Approach B – Relevance and Quality Ranking: Considers how relevant the content is to users along with its information quality.
- Approach C – CIBRS-Based Re-ranking: Uses the proposed scoring framework to balance relevance, information quality, viewpoint diversity, misinformation risk, and behavioural risk.

The performance of these approaches will be evaluated using the following measures:

1. Precision@K: Measures how many of the top recommended items are relevant to the user.
2. NDCG@K: Evaluates how effectively relevant content is placed near the top of the recommendation list.
3. Calibration Error: Measures how closely the model's predicted probabilities match actual outcomes.
4. Misinformation Exposure Rate: Measures the proportion of recommended content identified as misleading or unreliable.
5. Viewpoint Diversity: Evaluates the variety of opinions and perspectives included in the recommended feed.
6. Source Concentration: Measures how heavily recommendations depend on a limited number of information sources.
7. User-Reported Usefulness: Collects user feedback about the relevance, reliability, and usefulness of recommended content.

Finally, changes in political opinions should be studied separately through long-term research rather than being assumed from clicks, likes, shares, or other engagement activities. This distinction will help provide a more realistic understanding of how recommendation algorithms influence users over time.

6. Mathematical Formulation

6.1 Engagement-Based Baseline Model

Traditional recommendation systems generally give higher priority to content that receives more user engagement. This engagement can be measured through activities such as clicking, sharing, and watching content.

To represent this approach, the following mathematical model is used:

$$E(u,i) = \alpha_1 \times click(u,i) + \alpha_2 \times share(u,i) + \alpha_3 \times watchtime(u,i)$$

Here:

- E(u,i) represents the engagement score for user u and content item i.
- click(u,i) represents the number or normalised value of clicks.
- share(u,i) represents the sharing activity.
- watchtime(u,i) represents the time spent viewing the content.
- α_1, α_2, and α_3 are the coefficients assigned to each engagement factor.

Before applying the model, all engagement features are normalised so that they can be compared on a similar scale. The coefficients are estimated during model training.

This equation provides a simplified representation of engagement-based recommendation. It does not represent the actual or proprietary algorithm used by any particular social media platform.

6.2 Relevance and Misinformation Prediction Models

Machine learning can help estimate whether recommended content is relevant to a particular user. For this purpose, the proposed framework uses Logistic Regression as a baseline classification model.

The probability of content being relevant is calculated as:

$$P(rel = 1 \mid u,i) = \sigma(\beta_0 + \beta^T x)$$

Here, x represents the input features, β represents the model coefficients, and β_0 is the intercept.

The sigmoid function is defined as:

$$\sigma(z) = 1 / (1 + e^{-z})$$

This function converts the model output into a probability between 0 and 1, making it easier to estimate the likelihood that a particular content item is relevant to the user.

A similar approach can be used to estimate the possibility of misinformation:

$$P(mis = 1 \mid i) = \sigma(\gamma_0 + \gamma^T z)$$

In this equation, z represents the content-related features, while γ_0 and γ represent the model parameters.

The predicted probability indicates the estimated likelihood of misinformation based on the available training data. A Random Forest model can also be used as a comparison to examine whether a non-linear approach provides better prediction performance.

6.3 Correlation Analysis

Correlation analysis is used to understand whether two variables are related and how strongly they move together. In this research, it can help examine relationships between factors such as social media usage, user engagement, anxiety, and other behavioural indicators.

The Pearson correlation coefficient is calculated using the following formula:

$$r_{x\gamma} = \Sigma(x_j - \bar{x})(y_j - \bar{y}) / \sqrt{[\Sigma(x_j - \bar{x})^2 \, \Sigma(y_j - \bar{y})^2]}$$

Here:

- x and y represent the two variables being studied.
- $\bar{x}$ and $\bar{y}$ represent their respective mean values.
- $r_{x\gamma}$ represents the correlation coefficient.

The correlation value ranges from -1 to +1. A positive value indicates that both variables tend to increase together, while a negative value indicates that one tends to decrease when the other increases. A value close to zero suggests a weak linear relationship.

It is important to understand that correlation only identifies an association between variables. It does not prove that one variable directly causes changes in another.

6.4 Classification Performance Metrics

To understand how accurately the machine-learning models identify different categories, several classification metrics are used.

The main metrics include:

1. Precision

Precision measures how many of the items predicted as positive were actually positive.
Precision = TP / (TP + FP)
2. Recall
Recall measures how many of the actual positive cases were correctly identified by the model.
Recall = TP / (TP + FN)
3. F1-Score
The F1-score combines precision and recall into a single measurement. It is useful when both types of errors need to be considered.
F1 = 2 × Precision × Recall / (Precision + Recall)
4. Accuracy
Accuracy measures the overall proportion of correct predictions made by the model.
Accuracy = (TP + TN) / (TP + TN + FP + FN)
Where:
- TP = True Positive
- TN = True Negative
- FP = False Positive
- FN = False Negative
For identifying potentially risky content or behaviour, recall and probability calibration are particularly important alongside precision. This helps the system identify more genuine high-risk cases while avoiding unnecessary false alarms.

6.5 Weight Configuration for CIBRS

The proposed CIBRS framework assigns different importance levels to relevance, information quality, viewpoint diversity, misinformation risk, and behavioural risk.
To maintain a balanced scoring system, each weight must be non-negative, and the sum of all weights must equal 1.
$w_j \geq 0$
$\Sigma_j w_j = 1$
As an initial configuration for future experiments, the following weights are proposed:
CIBRS Component = Weight :-
Relevance (wR) =0.25, Information Quality (wQ) = 0.25,Viewpoint Diversity (wD) = 0.20 , Misinformation Risk (wM) = 0.15, Behavioural Risk (wH) = 0.15, Total = 1.00
Under this configuration, relevance and information quality receive the highest individual weights, followed by viewpoint diversity. Misinformation and behavioural risks are also included to ensure that potentially harmful content does not receive a high recommendation score simply because it attracts attention.
These values are only proposed starting weights for future experiments. They have not been obtained through empirical estimation or real-world testing. Their suitability must be examined through stakeholder feedback, experimental validation, and sensitivity analysis before they can be considered reliable.

7. Evidence-Based Findings

This paper presents an evidence-based analysis and a proposed framework. It does not pool raw participant-level datasets, and it does not claim to have trained or tested a new model. The following figures present selected quantitative findings reported in the cited studies.
Figure 1 summarises four correlations reported by More and Daund (2026), n=100. The largest listed association is between urge to check and anxiety (r=0.54). These associations are not evidence that algorithm exposure caused anxiety.
Figure 2 reproduces the platform interaction counts in Zafar and Saleem (2025). TikTok records the highest listed averages for likes (2,300), shares (480), and comments (310). Such counts describe

engagement in that study's analysed content; they should not be interpreted as a direct measure of persuasion or opinion change.

Additional reported model metrics from More and Daund (2026): Logistic Regression accuracy 89%, high-risk recall 0.96, and F1-score 0.85. These figures are reproduced as published and cannot be independently verified without the original dataset, split strategy, class balance, and code.

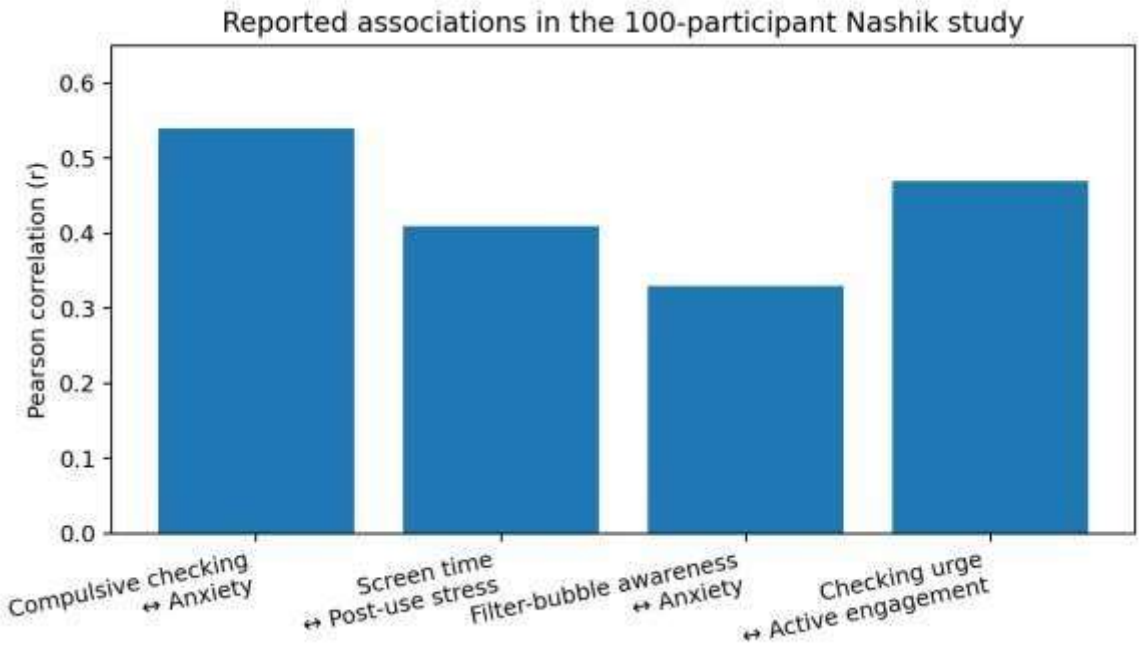


Figure 1. Correlations reported by More and Daund (2026); values are source-reported, not newly calculated.

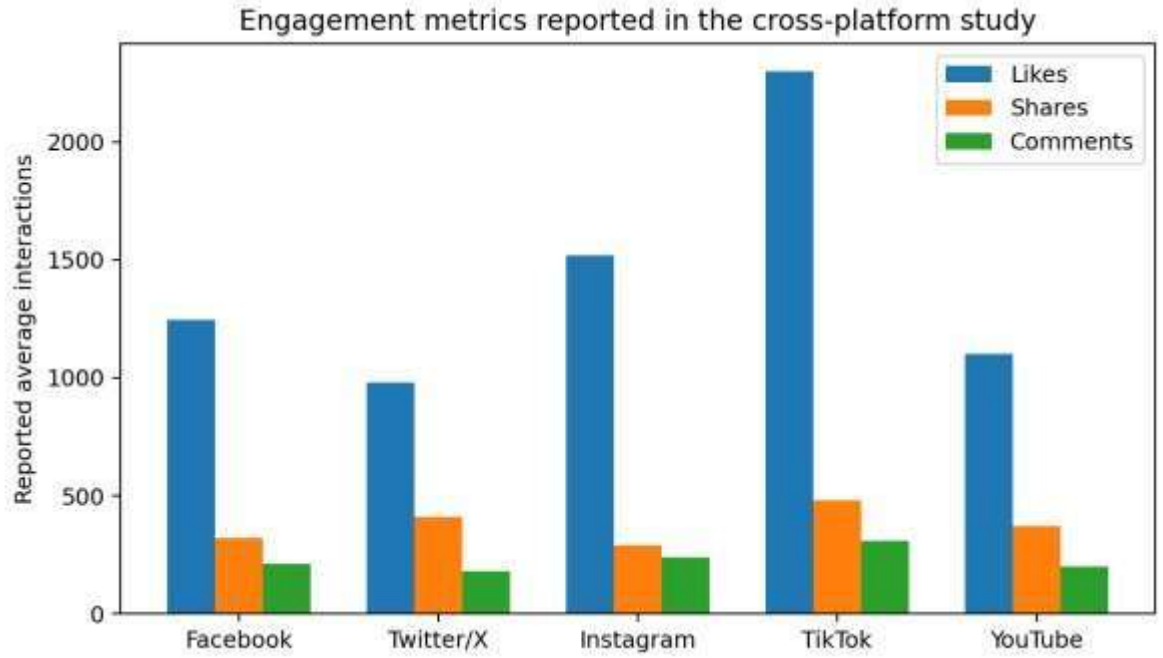


Figure 2. Average interaction counts reported by Zafar and Saleem (2025).

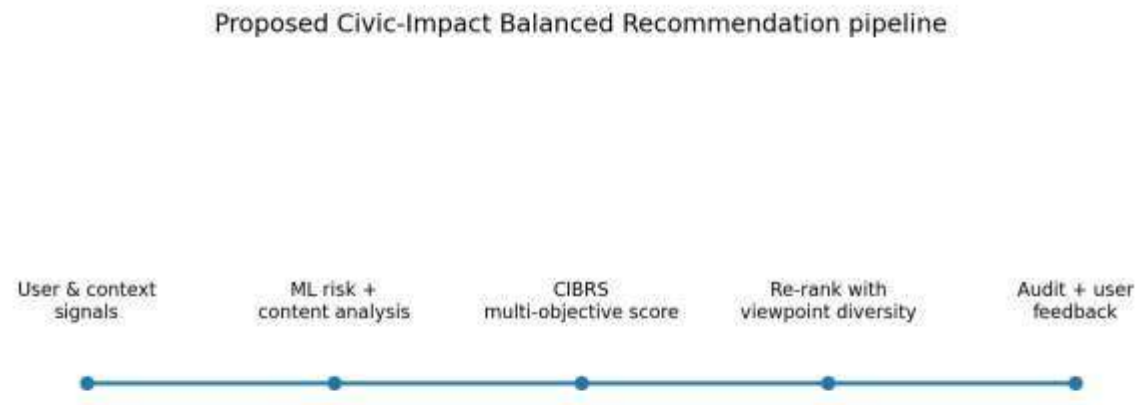


Figure 3. Proposed CIBRS pipeline (conceptual design developed in this paper).

8. Discussion: What the Combined Evidence Means

The reviewed studies highlight that social media algorithms have a stronger and more measurable influence on the information users see and engage with than on their long-term opinions. Although these algorithms can increase the visibility of political content, this does not necessarily mean they directly change people's beliefs or voting preferences.

Another important finding is that user engagement does not always indicate content quality. A post receiving more likes, shares, or comments is not necessarily accurate or informative. Similarly, personalisation can help users discover relevant content, but excessive repetition and low-quality information can limit their exposure to different viewpoints.

The proposed CIBRS framework addresses these concerns by considering relevance, information quality, viewpoint diversity, and potential user risks together. Its goal is to maintain useful recommendations while reducing repetitive and potentially harmful content.

The framework also supports greater user control through options such as chronological feeds, adjustable diversity settings, and explanations for recommended posts. Privacy remains essential, and behavioural-risk predictions should never be treated as medical diagnoses or used for political profiling.

Overall, the proposed approach aims to make social media recommendations more balanced, transparent, and responsible rather than focusing only on user engagement.

9. Proposed Experimental Design and Hypotheses

To evaluate the proposed CIBRS framework, a future study can involve a diverse group of adult participants divided into three groups. Each group will experience a different recommendation approach: engagement-only ranking, relevance-and-quality ranking, and CIBRS ranking.

User interactions and exposure patterns will be recorded to measure content diversity, source reliability, and misinformation exposure. Questionnaires before and after the study will help evaluate information quality, user trust, mental workload, and control over recommendations. A follow-up after a few weeks can examine whether these effects continue over time. The required sample size will be determined through statistical power analysis.

Research Hypotheses

H1: CIBRS will improve viewpoint and topic diversity compared with engagement-only ranking.

H2: CIBRS will reduce exposure to low-credibility content while maintaining content relevance.
H3: Recommendation explanations and user controls will improve users' sense of control.
H4: Changes in political opinions will be less consistent than changes in information exposure.
H5: Results will vary depending on platform, topic, user interests, and media literacy.

Evaluation Metrics

Primary Metrics:Normalised viewpoint diversity (H_n), Exposure to low-credibility content, NDCG@K (recommendation relevance).

Secondary Metrics:Perceived content relevance, User trust and autonomy, Cognitive overload.

The study should also report statistical confidence intervals, effect sizes, missing data, and possible negative effects to ensure reliable evaluation.

10. Limitations and Ethical Considerations

This research has some limitations because the reviewed studies use different research methods, participants, platforms, and evaluation measures. Some studies are based on surveys, while others use literature reviews or qualitative analysis. Therefore, their numerical findings cannot be directly combined. For example, the mental-health results come from a survey of 100 participants, whereas platform engagement statistics come from a separate study.

The proposed CIBRS framework is still a conceptual model and has not yet been experimentally validated. Its scoring weights, features, and prediction accuracy require further testing. Additionally, identifying misinformation and classifying viewpoints can be challenging because information and opinions may change depending on context. Diversity measurements may also fail to capture the actual quality of different perspectives.

To address these challenges, the framework should include independent audits, human review, privacy protection, and regular model updates.

From an ethical perspective, the system must ensure informed consent, secure data storage, minimum data collection, and user control over participation. The proposed framework should never be used for political targeting, voter manipulation, employment screening, or medical diagnosis. Transparency, fairness, and user privacy must remain priorities throughout its development.

11. Conclusion

The research literature shows that social-media algorithms shape information visibility, engagement, and the conditions under which public opinion and decisions form. They also show that impacts vary across platforms, people, and contexts, and that exposure effects should not be mistaken for automatic attitude change. The unique contribution of this paper is the proposed CIBRS framework: a transparent, multi-objective approach that combines relevance and information quality with diversity and risk penalties, then evaluates feed sets rather than clicks alone.

The next research step is empirical validation using a preregistered, controlled comparison of engagement-only, quality-aware, and CIBRS ranking. Until such testing is completed, CIBRS should be treated as a testable design hypothesis—not as a proven solution.

References

[1] E. J. Hasibuan, A. D. R. Putra, and A. S. Dirgantari, "The Role of Social Media Algorithms in Shaping Public Opinion During Political Campaigns," IJSH, vol. 1, no. 2, Dec. 2024.

[2] S. More and R. V. Daund, "Impact of Social Media Algorithms on Public Opinion and Mental Health," International Journal of Advanced Research in Science, Communication and Technology, vol. 6, no. 14, Apr. 2026, doi: 10.48175/IJARSCT-34090.

[3] J. Joselin, A. Vaishya, D. U, and P. V, "A Study on the Influence of Social Media Algorithms on Public and Political Opinion," International Journal of Advanced Research in Science, Communication and Technology, vol. 5, no. 1, Aug. 2025, doi: 10.48175/568.

[4] M. Zafar and S. Saleem, "Algorithmic Influence of Social Media Platforms on Public Opinion and Economic Decisions in the Digital Age," Journal of International Relations and Social Dynamics, vol. 4, no. 1, pp. 48–59, May 2025.

[5] N. Bajuwaiber, "Exploring the impact of social media algorithms on public discourse: a critical review of their role in shaping political opinions and media consumption," Telematics and Informatics Reports, vol. 24, 100384, 2026, doi: 10.1016/j.teler.2026.100384.

[6] R. Shekhar and V. Kumar, "Impact of Social Media Algorithms on Public Decision," International Journal of Scientific Research in Engineering and Management, vol. 9, no. 5, May 2025, doi: 10.55041/IJSREM48984.