International Research Journal of Engineering and Technology (IRJET) e-ISSN: 2395-0056
Volume: 13 Issue: 10 | Oct 2026 www.irjet.net p-ISSN: 2395-0072

Impact of Language and Social Media Usage on Misinformation Detection Among Individuals in Navi Mumbai: A Comparative Study Across Different Age and Occupational Groups.

Dr.Siji Nambiar[1], Miss.Ankita Yadav[2], Mr. Ashmir Shaikh[3], Miss.Ayesha Shaikh[4]

[1]Assistant Professor, D. Y. Patil Deemed to be University, School of Humanities & Sciences, Belapur, Maharashtra, India.

[234]Students, Department of MCA, D. Y. Patil Deemed to be University, School of Humanities & Sciences, Belapur, Maharashtra, India.

Abstract - *The rapid growth of social media has changed the way people receive, understand, and share information. Along with the easy availability of information, social media has also created opportunities for misinformation to spread quickly among users. In a multilingual environment such as India, the language in which information is presented may also influence how people understand, trust, and evaluate that information. This study examines the relationship between language, social media usage, and misinformation detection among individuals in Navi Mumbai. A survey-based approach was used to understand how people across different age groups and occupational backgrounds interact with information received through social media. The study focuses on users' experiences with different languages, their level of trust in online information, their ability and willingness to identify potentially misleading information, and the methods they use to verify information before accepting or sharing it. The study also considers differences in social media usage patterns among respondents. The collected responses provide insight into how demographic factors, language preferences, and social media habits may influence the way people evaluate information online. The findings help identify common patterns in misinformation awareness and verification behaviour among social media users in Navi Mumbai. The study also provides insight into the difficulties people may face when evaluating information in a multilingual digital environment.*

Keywords - Misinformation, Social Media, Language, Misinformation Detection, Information Verification, Multilingualism, Social Media Users, Navi Mumbai.

INTRODUCTION

Social media has become an important part of daily communication. Platforms such as WhatsApp, Instagram, Facebook and YouTube allow people to receive and share information quickly. However, this easy access has also increased the spread of misinformation, which includes false, inaccurate or misleading information shared intentionally or unintentionally through messages, images, videos and posts.

Language can influence how people understand and judge information, especially in a multilingual country like India. People commonly use English, Hindi, Marathi and Hinglish and may switch between languages while communicating online. Information in a familiar language may feel easier to understand and more trustworthy, but familiarity does not mean that the information is accurate.

Social-media habits and verification practices also play an important role. People differ in how often they receive information, encounter forwarded messages, and check information before believing or sharing it. Similarly, photos, videos, likes and repeated shares may make information appear more believable, even though they do not guarantee that the information is true.

This study focuses on 136 respondents from Navi Mumbai from different age and occupational groups. It examines their social-media and WhatsApp usage, language preferences, verification habits, confidence in identifying misinformation, and the influence of images, videos and number of shares.

The study also uses potentially misleading statements in English, Hindi, Marathi and Hinglish to understand whether language presentation affects people's judgement. Overall, the research looks at misinformation detection through language, social-media exposure, verification behaviour and individual characteristics. The findings can help understand how people in Navi Mumbai deal with misleading information and highlight the importance of digital and media literacy.

LITERATURE REVIEW

Bhosle et al. (2025) explored the use of machine learning for detecting fake news in different Indian languages. Their study showed that automated techniques can be useful for identifying misleading content. However, the study mainly focused on machine-based detection and did not look at how people understand or identify misinformation when it is presented in different languages.

Patil et al. (2024) worked on creating a multilingual fake-news dataset that included Gujarati, Hindi, Marathi, and Telugu. The **dataset** can be useful for developing fake-news detection systems. However, the study was mainly concerned with dataset development and did not study how people judge information or verify whether it is genuine.

Fake News Detection in Indian Languages (2025) examined the use of a CNN-LSTM model for detecting fake news in Hindi. The study showed that deep-learning techniques can be applied to Hindi news content. However, the focus was on the performance of the model rather than on how people themselves identify misleading information.

Fake News Detection on Hindi News Dataset examined fake-news detection using a Hindi news dataset. The study mainly dealt with computational methods for identifying fake news. It did not focus on whether factors such as language familiarity or a person's background affect their ability to recognize misinformation.

Garg et al. (2024) studied fake-news detection in Hindi using multimodal and ensemble-learning techniques. Their research showed that combining different approaches can improve automated detection. However, the study concentrated on machine-based analysis and did not examine people's trust, judgement, or verification habits when they come across questionable information.

Verma (2026) investigated fake-news detection in Indic languages using a fine-tuned MuRIL model. The study focused on improving the ability of an AI model to identify fake news in different Indian languages. However, it did not examine how people themselves decide whether information is true or false across different languages.

Shad et al. studied hate speech and fake narratives in Hinglish social-media content. Their research mainly focused on automatically identifying harmful online content. Although the study highlights the challenges of working with mixed-language content, it did not specifically examine how people understand or verify misinformation written in Hinglish.

Yadav et al. (2025) developed a multi-stage approach for detecting racial hoaxes in code-mixed Hindi-English social-media content. Their work focused on improving machine-learning models for automated detection. It did not study how people from different age groups identify or respond to such misleading information.

Rabbani et al. (2025) compared different approaches, including CharBERT and large language models, for detecting racial hoaxes in code-mixed Hindi-English social-media content. The study mainly evaluated automated detection methods. Human understanding, trust, and verification of information were not the main focus of the research.

Mathur et al. examined offensive content in Hindi-English code-switched tweets. The research focused on automatically identifying offensive language. However, it did not explore how people perceive, trust, or verify information when misinformation is presented using different languages or mixed-language communication.

Upadhyay et al. (2026) studied the detection of harmful content and its targets in code-mixed Hinglish using multilingual and cross-lingual transformer models. The research mainly dealt with language-processing techniques and automated detection. It did not focus on how ordinary social-media users judge or verify potentially misleading information.

OBJECTIVES OF THE STUDY

1. To study the effect of language on understanding and trust in social-media information.
2. To examine misinformation detection across age and occupation groups.
3. To study the relationship between social media usage and misinformation detection.
4. To examine verification habits before believing and sharing information.
5. To compare misinformation perception across English, Hindi, Marathi and Hinglish.
6. To study the influence of language, visuals, shares and mixed-language content.
7. To measure the overall misinformation detection level of respondents.

RESEARCH HYPOTHESES

1. Language

- **Null Hypothesis (H_{01}):** There is no significant association between language and misinformation detection.
- **Alternative Hypothesis (H_{11}):** There is a significant association between language and misinformation detection.

2. Social Media Usage

- **Null Hypothesis (H_{02}):** There is no significant relationship between social media usage and misinformation detection.
- **Alternative Hypothesis (H_{12})**: There is a significant relationship between social media usage and misinformation detection.

3. Age

- **Null Hypothesis (H_{03}):** There is no significant relationship between age and misinformation detection.
- **Alternative Hypothesis (H_{13}):** There is a significant relationship between age and misinformation detection.

4. Verification before Believing

- **Null Hypothesis (H_{04}):** There is no significant relationship between verification before believing and misinformation detection.
- **Alternative Hypothesis (H_{14})**: There is a significant relationship between verification before believing and misinformation detection.

5. Verification before Sharing

- **Null Hypothesis (H_{05}):** There is no significant relationship between verification before sharing and misinformation detection.
- **Alternative Hypothesis (H_{15}):** There is a significant relationship between verification before sharing and misinformation detection.

RESEARCH METHODOLOGY

Research Design

The study used a descriptive survey-based research design to examine how language and social media usage are related to people's ability to identify misinformation.

Study Area

The study was conducted among respondents from Navi Mumbai, which had a population of approximately 16.36 lakh as per the 2025 Census. The respondents represented different occupational groups, with the highest participation from college students, followed by working professionals, self-employed individuals, homemakers, retired persons, and others. This diverse sample helped the study examine differences in social media usage, language preferences, verification behaviour, and misinformation detection among people from different backgrounds.

Sample Size

A total of 136 respondents participated in the study. The respondents belonged to different age groups and occupational backgrounds.

Data Collection

Primary data was collected through a structured Google Forms questionnaire. The questionnaire collected information regarding respondents' age, occupation, language preference, social media usage, WhatsApp usage, verification behaviour, trust in information and ability to identify misleading information.

Variables Studied

The study considered different factors, such as the respondents' age, the language they usually use to read or understand news, and how often they receive information through social media and WhatsApp. It also looked at whether they check information before believing or sharing it, how they react to suspicious messages, and how confident they feel when identifying fake news. In addition, the study explored whether people trust information more when it is in their preferred language and whether mixed-language content makes it harder for them to identify misinformation.

Misinformation Detection

Six statements were included to measure the respondents' ability to identify misinformation. The statements covered topics such as charging a phone in airplane mode, WhatsApp forwarded messages, health-related claims, bank-account messages, scholarship messages and the effect of the number of shares.

Each response was compared with the correct status of the statement. A score from 0 to 6 was calculated for each respondent, where a higher score represented better misinformation detection.

Multilingual Comparison

The questionnaire also included similar misleading news statements presented in English, Hindi, Marathi and Hinglish. Respondents were asked which version appeared most likely to be misleading or whether all versions appeared equally misleading.

Data Analysis

The collected data were analysed using frequencies, percentages, mean detection scores and group comparisons. Statistical tests based on the ordinal nature of the questionnaire variables were also used to examine relationships between selected factors and misinformation detection ability.

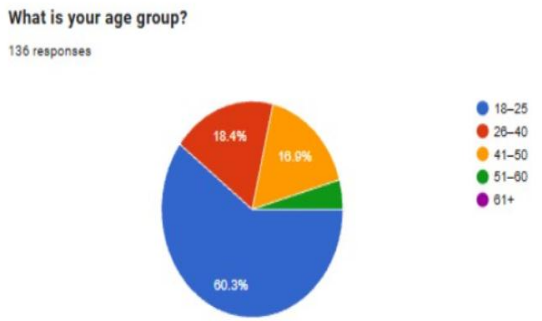
The results were interpreted at a 5% significance level ($\alpha = 0.05$).

4.1 Demographic Profile

A total of 136 respondents participated in the study.

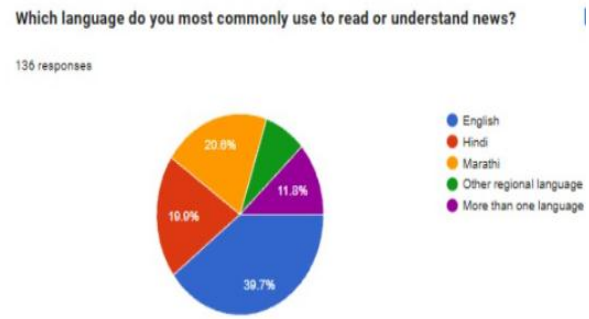
The largest group was 18–25 years, consisting of 82 respondents (60.3%).

In terms of occupation, 65 respondents (47.8%) were students, followed by working respondents (18.4%) and homemakers (15.4%).



4.2 Language Used for Understanding News

English was the most commonly used language for reading or understanding news, reported by 54 respondents (39.7%).



4.3 Social Media as a Source of Information

The responses showed that social media was a regular source of information for most participants.

- Sometimes – 38 (27.9%)
- Rarely – 36 (26.5%)
- Often – 30 (22.1%)
- Always – 28 (20.6%)
- Never – 4 (2.9%)

Only 2.9% reported never receiving news or information through social media.

4.4 WhatsApp and Forwarded Messages



4.5 Verification Before Believing Information

The responses show that verification behaviour differed considerably among participants.



4.6 Verification Before Sharing

Verification before sharing was somewhat more frequent than verification before believing.

- Always – 45 (33.1%)
- Sometimes – 39 (28.7%)
- Often – 28 (20.6%)
- Rarely – 16 (11.8%)
- Never – 8 (5.9%)

This indicates that respondents were somewhat more likely to verify information before forwarding or sharing it than before accepting it as believable.

4.7 Misinformation Detection Score

Six misinformation statements were included in the questionnaire. Each correct identification of a false statement received one point.

The results were:

- Mean score = 3.20 out of 6
- Average detection percentage = 53.3%
- Minimum score = 0

- Maximum score = 6

The average score of 3.20/6 indicates a moderate level of misinformation detection among the respondents.

4.8 Detection Performance by Age

- 18–25 years: Average score of 3.41 out of 6.
- 26–40 years: Average score of 2.64 out of 6.
- 41–50 years: Average score of 2.96 out of 6.
- 51–60 years: Average score of 3.50 out of 6.

The average scores varied between age groups. However, statistical testing did not show a significant relationship between age group and detection score at the 5% level.

4.9 Detection Performance by Language

Average Misinformation Detection Score by News Language

- English: Average score of 3.44 out of 6.
- More than one language: Average score of 3.44 out of 6.
- Other regional languages: Average score of 3.09 out of 6.
- Marathi: Average score of 2.96 out of 6.
- Hindi: Average score of 2.85 out of 6.

4.10 Effect of Social Media Usage

The average misinformation detection scores increased with greater frequency of receiving information through social media:

Social Media Frequency	Average Score / 6
Never	1.75
Rarely	2.47
Sometimes	3.37
Always	3.39
Often	3.87

A Spearman correlation test showed a positive and statistically significant relationship between the frequency of social media information and misinformation detection score ($\rho = 0.279$, $p = 0.001$).

$$r_s = 1 - \frac{6\sum d^2}{n(n^2-1)}$$

This indicates an association between greater exposure to information through social media and higher detection scores in this sample. However, this result does not establish that social media usage causes better detection ability.

4.11 Verification Before Believing and Detection

The relationship between verification before believing and detection score was positive but did not reach the 5% significance level ($\rho = 0.155$, $p = 0.072$).

Therefore, the data suggests a possible relationship, but the evidence was not statistically significant at $\alpha = 0.05$.

4.12 Verification Before Sharing and Detection

Verification before sharing showed a statistically significant positive relationship with misinformation detection ($\rho = 0.237$, $p = 0.005$).

This suggests that respondents who reported verifying information more frequently before sharing tended to obtain higher misinformation detection scores.

Again, this represents an association and should not be interpreted as proof of causation.

4.13 Multilingual News Statements

For the scholarship-related statement:

- All versions equally misleading – 50 respondents
- English – 28
- Marathi – 21
- Hindi – 19
- Hinglish – 18

For the bank-account statement:

- All versions equally misleading – 49 respondents
- English – 25
- Marathi – 22
- Hindi – 21
- Hinglish – 19

The largest group in both cases selected “All versions are equally misleading.” This suggests that many respondents judged the underlying claim rather than considering one language version inherently more misleading.

However, the remaining responses were distributed across English, Hindi, Marathi and Hinglish, indicating that

language presentation may still influence some respondents' perceptions.

INTERPRETATION

The hypotheses were tested using the questionnaire responses of 136 participants. Since several independent variables were ordinal in nature and misinformation detection was measured using a 0–6 score, **Spearman's rank correlation** was used for ordinal relationships, while a **Kruskal–Wallis test** was used to compare detection scores across language groups. The significance level was set at 5%.

Hypothesis	Test Result	p-value	Decision
Language used for news → Detection score	Kruskal–Wallis H = 3.621	0.460	Not significant
Social media frequency → Detection score	Spearman ρ = 0.279	0.001	Significant
Age group → Detection score	Spearman ρ = −0.144	0.094	Not significant
Verification before believing → Detection score	Spearman ρ = 0.155	0.072	Not significant
Verification before sharing → Detection score	Spearman ρ = 0.237	0.005	Significant

Hypothesis 1: Language and Misinformation Detection

The p-value was 0.460, which is greater than 0.05. Therefore, there was insufficient statistical evidence to reject H_{01}.

The results did not show a statistically significant association between the language commonly used to understand news and the misinformation detection score.

However, the descriptive results showed differences in average scores between language groups. Therefore, language may influence individual perception or trust without producing a statistically significant difference in overall detection scores in this sample.

Hypothesis 2: Social Media Usage and Misinformation Detection

The Spearman correlation was $\rho = 0.279$ with $p = 0.001$.

Since $p < 0.05$, H_{02} was rejected in favour of H_{12}.

The results showed a statistically significant positive association between frequency of receiving information through social media and misinformation detection score. This should be interpreted as an association rather than a causal relationship.

Hypothesis 3: Age and Misinformation Detection

The Spearman correlation was $\rho = -0.144$ with $p = 0.094$.

Since $p > 0.05$, there was insufficient evidence to reject H_{03}.

Therefore, age group was not significantly associated with misinformation detection score in this sample.

Hypothesis 4: Verification Before Believing and Misinformation Detection

The Spearman correlation was $\rho = 0.155$ with $p = 0.072$.

Since $p > 0.05$, H_{05} was retained.

Although the relationship was positive, it was not statistically significant at the 5% level.

Hypothesis 5: Verification Before Sharing and Misinformation Detection

The Spearman correlation was $\rho = 0.237$ with $p = 0.005$.

Since $p < 0.05$, H_{05} was rejected in favour of H_{15}.

This indicates a statistically significant positive association between verification before sharing and the ability to detect misinformation.

OVERALL INTERPRETATION

The statistical analysis indicates that social media information exposure and verification before sharing were significantly associated with misinformation detection scores. In contrast, the commonly used news language, age group and verification before believing did not show statistically significant relationships with the detection score at the 5% significance level.

The findings therefore suggest that misinformation detection is influenced by several aspects of information-handling behaviour, but the results do not support the conclusion that language or age alone determines a person's ability to identify misinformation.

KEY FINDINGS

1. **Social media is an important source of information:**
   Most of the respondents receive news or other information through social media at least sometimes. This shows that social media has

become a regular part of how people get information.

2. **Young people formed the largest part of the survey:**
   Most respondents were between 18 and 25 years old, and students were the largest occupational group. Therefore, the responses mainly reflect the experience of young social media users.
3. **People use different languages for getting news:**
   English was the most commonly used language, but Hindi and Marathi were also used by many respondents. Some people also use more than one language to understand news.
4. **Verification habits are different among people:**
   Many respondents said that they check information before believing or sharing it. However, some respondents do not verify information regularly, which can increase the chances of believing or spreading misinformation.
5. **Language can affect trust:**
   A good number of respondents said that they are more likely to trust information written in a language they understand well. This suggests that understanding the language can influence how people receive information.
6. **Photos and videos can make information look more believable:**
   Some respondents said that visual content makes information appear more trustworthy. However, not everyone agreed with this, so the effect is different from person to person.
7. **More shares do not always mean more trust:**
   Most respondents did not strongly agree that they would trust information simply because it had been shared many times. This shows that many people understand that popularity does not necessarily mean information is true.
8. **Misinformation detection was moderate:**
   Respondents correctly identified around 3.2 out of 6 misinformation statements on average. This shows that people were able to identify some false information, but identifying misinformation is still difficult in certain cases.
9. **Obvious fake messages were easier to identify:**
   Respondents performed better on claims related to fake scholarship offers and bank-account messages. Some of the more believable claims were harder for them to judge correctly.
10. **Awareness can help reduce misinformation:**
    The responses show that checking the source and verifying information before believing or forwarding it can help people avoid misinformation.

LIMITATIONS OF THE STUDY

1. **Limited Responses:** The study included only respondents who completed the Google Form. A larger sample could have provided broader insights.
2. **Limited Time:** Due to time constraints, the study could not examine all factors that may influence misinformation detection.

FUTURE SCOPE OF STUDY

Future research can be conducted using a larger, more diverse sample to better understand misinformation among people in Navi Mumbai and other regions. The study can compare English, Hindi, Marathi, and Hinglish to examine the role of language in identifying misinformation. It can also study different social media platforms such as WhatsApp, Instagram, YouTube, and Facebook separately. Further research can focus on digital awareness and fact-checking habits, including whether simple fact-checking education can improve people's ability to identify fake news.

CONCLUSION

This study explored how people in Navi Mumbai receive and judge information on social media, with a focus on how language influences their understanding and trust. The findings show that social media and WhatsApp are common sources of information, and people may feel more comfortable trusting content shared in a language they understand well.

However, knowing the language does not always mean that a person can identify whether information is true or false. The source of the information, its content, and the habit of checking facts also play an important role. The responses showed that while participants could identify some misleading information, they found certain claims difficult to judge.

Overall, the study shows why it is important to think carefully before trusting or sharing information online. Checking the source and verifying facts before forwarding a message can help people make better decisions and prevent misinformation from spreading further.

REFERENCES

[1] A. Bhosle, K. Patil, S. Thite, S. Dash, A. Sengar, and Y. Lahoti, “Multilingual Fake News Detection: A Machine Learning Approach for Indian Languages,” Advances in Intelligent Systems Research, vol. 200, pp. 384–397, 2026, doi: 10.2991/978-94-6463-948-3_28.

[2] K. Patil, P. Gandhi, A. Chauhan, V. Patil, and A. Pawar, "Multilingual Fake News Detection Dataset: Gujarati, Hindi, Marathi, and Telugu," Zenodo, 2024.

[3] "Fake News Detection in Indian Languages: A Case Study with Hindi Using CNN-LSTM," Procedia Computer Science, vol. 259, pp. 150–160, 2025, doi: 10.1016/j.procs.2025.03.316.

[4] "Fake News Detection on Hindi News Dataset," Global Transitions Proceedings, vol. 3, no. 1, pp. 289–297, 2022, doi: 10.1016/j.gltp.2022.03.014.

[5] S. Garg et al., "Fake News Detection in the Hindi Language Using Multi-Modality via Transfer and Ensemble Learning," Internet Technology Letters, 2025, doi: 10.1002/itl2.523.

[6] A. Verma, "Multilingual Fake News Detection in Indic Languages Using Fine-Tuned MuRIL," SSRN, 2026.

[7] M. Shad et al., "Detecting Hate Speech and Fake Narratives in Code-Mixed Hinglish Social Media Text," in Proc. ICON-FauxHate, 2024.

[8] A. S. Yadav, D. Sharma, A. Singh, and V. K. Singh, "Hope_for_best@LT-EDI 2025: Detecting Racial Hoaxes in Code-Mixed Hindi-English Social Media Data Using a Multi-Phase Fine-Tuning Strategy," in Proc. LT-EDI, 2025.

[9] A. H. Rabbani, D. D. Droba, M. A. Labib, S. Rahman, and H. Murad, "Hoax Terminators@LT-EDI 2025: CharBERT's Dominance over LLM Models in the Detection of Racial Hoaxes in Code-Mixed Hindi-English Social Media Data," in Proc. LT-EDI, 2025.

[10] P. Mathur, R. Shah, R. Sawhney, and D. Mahata, "Detecting Offensive Tweets in Hindi-English Code-Switched Language," in Proc. Sixth International Workshop on NLP for Social Media, pp. 18–26, 2018, doi: 10.18653/v1/W18-3504.

[11] A. Upadhyay, E. Khatri, H. P. Singh, S. K. Mishra, and V. K. Rai, "Span-Level Target and Harmfulness Detection in Code-Mixed Hinglish Using Cross-Lingual and Multilingual Transformers," Expert Systems with Applications, vol. 308, 2026, Art. no. 131033, doi: 10.1016/j.eswa.2025.131033.

[12] "Explainable Multilingual and Multimodal Fake-News Detection: Toward Robust and Trustworthy AI for Combating Misinformation," Frontiers in Artificial Intelligence, 2025, doi: 10.3389/frai.2025.1690616.