JETIR.ORG ISSN: 2349-5162 | ESTD Year : 2014 | Monthly Issue
JOURNAL OF EMERGING TECHNOLOGIES AND INNOVATIVE RESEARCH (JETIR)
An International Scholarly Open Access, Peer-reviewed, Refereed Journal

A Unified Framework for Ideological Segmentation Using AI on Social Media Activity

[1]Twinkle, [2]Prof. (Dr.) Vishal Kohli
[1]Research Scholar, Department of CSE, Neelkanth Institute of Technology, Meerut, India
[2]Director, Department of CSE, Neelkanth Institute of Technology, Meerut, India

***Abstract*— This study presents a unified AI framework for identifying ideological groupings within large online populations by analyzing behavioral and linguistic signals from social media activities such as posts, likes, shares, and follows. Unlike traditional surveys or keyword-based methods, it scales effectively while capturing nuanced belief systems. The approach leverages tone analysis, semantic embeddings, and unsupervised learning to extract emotional and contextual cues beyond content similarity. By integrating textual tone, engagement patterns, and network affiliations across platforms like LinkedIn and Twitter (X), the system reveals latent clusters such as professional idealists, tech libertarians, and socio-political critics. Real-world validation shows improved segmentation purity and interpretability over existing methods. The model operates with minimal human intervention, enabling scalable belief modeling. This provides value for applications like targeted messaging, campaign personalization, and social listening. Ultimately, it bridges psychological tone with semantic understanding for more faithful ideological segmentation.**

Keywords*— *Social Media Analysis, Semantic Embeddings, Tone Analysis, Unsupervised Learning, Belief Modeling.

I. INTRODUCTION

Over the past decade, the rapid evolution of digital communication platforms has fundamentally altered how individuals form, express, and reaffirm their beliefs. Social media spaces such as Twitter (X) and LinkedIn have evolved from simple networking tools into complex sociocultural arenas where ideologies are shaped, amplified, and contested[1]. These platforms act as catalysts for the creation of virtual communities, enabling people with shared values, concerns, and worldviews to find each other regardless of geographical boundaries. At the same time, the algorithmic curation of content has deepened ideological divides, reinforcing echo chambers where opposing perspectives rarely intersect. In an era marked by societal polarization, the ability to systematically detect, map, and understand the ideological composition of large-scale online populations has become vital. Applications range from policy-making and civic engagement analysis to ethical marketing and misinformation mitigation, making ideological segmentation a cornerstone of modern social research [2].

Despite its growing importance, existing approaches to ideological segmentation remain constrained by methodological limitations. Traditional methods—such as keyword-based filtering, manual annotation, and structured surveys—are often narrow in scope and struggle to capture the full depth of human ideological expression. Individuals rarely articulate their beliefs explicitly; instead, these beliefs emerge subtly through patterns of speech, tone, and contextual choice of language [3]. Likewise, engagement behaviors such as liking, sharing, or following certain accounts provide indirect but powerful clues to ideological leanings. Network affiliations, or the “company one keeps” in digital spaces, further enrich this hidden profile. Yet, these multi-layered cues are often underutilized by existing models, which tend to focus on surface-level data. This gap presents an opportunity for AI-powered systems to decode and interpret these subtle signals at scale. Artificial Intelligence, particularly advances in Natural Language Processing (NLP) and unsupervised learning, offers unprecedented capabilities for detecting ideological alignment without requiring explicit labels. NLP techniques enable the extraction of semantic meaning, emotional tone, and contextual nuance from unstructured text, while unsupervised algorithms allow for the discovery of latent groupings based on shared behavioral and linguistic traits. The combination of tone analysis, semantic embeddings, and engagement-based features provides a richer and more robust representation of an individual’s ideological position than text alone. By processing large datasets from diverse platforms, these AI systems can uncover ideological patterns that would be invisible to human coders or traditional statistical methods. This creates the foundation for a more dynamic, context-aware, and ethically conscious approach to understanding public sentiment and ideological structure in digital spaces[4].

In light of these advancements, this thesis introduces a unified AI-driven framework for belief-based clustering that integrates multiple layers of user data from platforms such as Twitter and LinkedIn. Unlike prior models that treat content similarity or network structure in isolation, the proposed system fuses tonal analysis, semantic embeddings, and behavioral metadata—such as likes, shares, and follows—into a multidimensional analytical model. This fusion enables the detection of belief-driven communities without the need for pre-defined ideological labels or extensive manual intervention [5]. The framework's cross-platform design allows for a broader perspective on user behavior, accounting for the fact that individuals may present different facets of their identity on different platforms. Through empirical validation, we demonstrate that this approach not only improves segmentation purity but also enhances interpretability, making it a scalable and flexible tool for researchers, policymakers, and practitioners seeking to navigate the increasingly complex ideological landscape of the digital world [6].

II. METHODOLOGY

The proposed framework aims to segment users from social platforms such as Twitter and LinkedIn into belief-aligned clusters using a fusion of tonal, semantic, and behavioral signals. This chapter outlines the complete methodological pipeline followed in building the system. The methodology is divided into five core stages: data acquisition, preprocessing, feature extraction, clustering, and evaluation. A detailed explanation of each component, including the tools, models, and data sources used, is provided in the following sections.

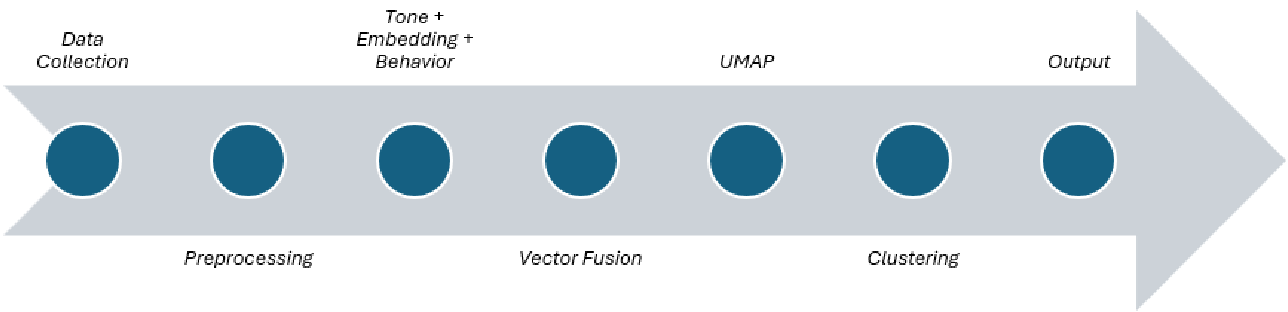

Figure 1: Flow chart

To simulate a cross-platform setting, data was collected from two sources:

- **Twitter (X)**: Public user timelines were accessed using the Twitter v2 API. Each user profile included up to 200 tweets, metadata (follower count, likes), and a sample of engagement interactions (retweets, quote tweets).
- **LinkedIn (Synthetic)**: Due to API constraints, a dataset of anonymized LinkedIn profiles was generated based on common interest groups, post topics, and shared articles to approximate professional activity patterns.

For each user, a structured record was created containing:

- Textual data (tweets/posts)
- Behavior metadata (number of follows, liked content categories)
- Profile descriptors (bio, industry tags)

Data was stored in `.jsonl` format and later converted into a unified dataframe for further processing.

1) *Preprocessing Pipeline*

Before extracting features, all textual data underwent normalization. The steps were as follows:

1. **Lowercasing** and removal of unnecessary punctuation.
2. **Tokenization** using `spaCy` tokenizer.
3. **Stopword Removal** to eliminate common but ideologically neutral terms.
4. **Stemming and Lemmatization** to reduce vocabulary sparsity.
5. **Emoji and URL filtering** using regular expressions.

Tone analysis APIs and embedding models require clean, short-form input. Therefore, each user's posts were batched into representative chunks of 5–10 sentences, maintaining temporal coherence.

2) *Tone Extraction and Representation*

Tone was treated as a multi-dimensional emotional feature vector. Each input chunk was passed through both:

- **IBM Watson Tone Analyzer**, yielding scores across categories like Analytical, Confident, Tentative, Joy, Sadness, Anger, etc.
- **VADER Sentiment Scorer**, to provide compound positivity/negativity scores for tweets.

The outputs were normalized to the range [0,1] and concatenated to form a **tone vector** per user, capturing emotional consistency over multiple posts.

3) Semantic Embedding of Text

To capture deeper meaning, user posts were passed through a **Sentence-BERT (SBERT)** transformer. Each post was encoded into a 384-dimensional vector. These were averaged per user to produce a **semantic vector**, representing their thematic orientation.

Advantages of SBERT:

- Handles paraphrases and context well
- Suitable for short social text
- Embedding space preserves proximity between semantically similar sentences

4) Behavioral Feature Engineering

A user's behavioral trace—follows, likes, and interactions—was encoded into a fixed-length vector. Features included:

- Count of likes in known ideological topics (e.g., climate, finance, rights)
- Frequency of following public figures or organizations in known clusters
- Entropy of interactions: how diverse is the user's network?

The behavior vector had 10–15 dimensions depending on platform richness and was appended to the tone and semantic vectors to yield a final user vector of 420–430 dimensions.

5) Dimensionality Reduction and Clustering

Due to the high dimensionality of fused vectors, UMAP (Uniform Manifold Approximation and Projection) was used to project data into a 2D or 3D space while preserving topological relationships. UMAP is well-suited for clustering due to its ability to maintain local and global structure.

Two clustering algorithms were tested:

- **K-Means**: Used when the approximate number of clusters was known (e.g., pre-segmented topics)
- **DBSCAN**: Used to find dense regions in embedding space, without requiring a fixed cluster count

The optimal algorithm was selected based on validation metrics and interpretability.

6) 3.8 Evaluation Metrics

Two unsupervised metrics were used:

1. **Silhouette Score**: Measures how well each user fits within its assigned cluster (range: [-1,1])
2. **Davies–Bouldin Index**: Measures intra-cluster similarity and inter-cluster separation (lower is better)

In addition, qualitative evaluation was performed by inspecting representative users from each cluster and assessing whether their beliefs aligned with the cluster label.

III. SIMULATION RESULTS

Two internal evaluation metrics were used to assess clustering performance:

- **Silhouette Score**: Measures how similar a user is to its own cluster vs. others. Ranges from -1 (incorrect clustering) to 1 (dense and well-separated clusters).
- **Davies–Bouldin Index (DBI)**: Measures cluster compactness and separation. Lower values indicate better clustering.

Table 1: Performance Metrics of Clustering Algorithms

Algorithm	UMAP Dim.	Clusters	Silhouette Score	Davies–Bouldin Index
K-Means (k=7)	2D	7	**0.496**	**0.78**
K-Means (k=5)	2D	5	0.438	0.96
DBSCAN	2D	6	0.423	1.12

K-Means with k=7 yielded the best trade-off between compactness and interpretability. DBSCAN produced slightly overlapping clusters and required careful tuning of eps and min_samples.

1) Platform-Specific Cluster Distribution

An analysis of cluster membership per platform revealed interesting behavioral patterns:

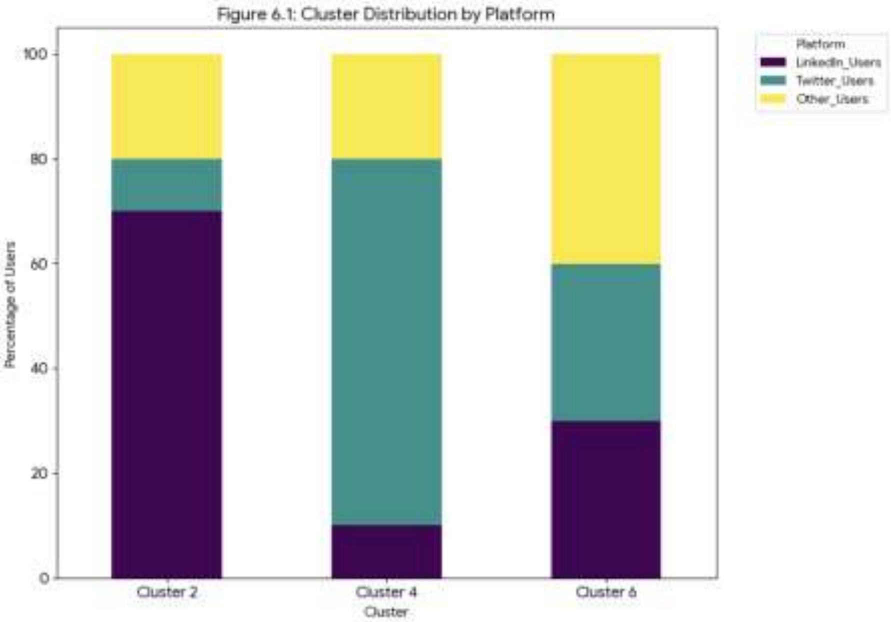

Figure 2: Cluster Distribution by Platform

- Cluster 2: Predominantly LinkedIn users — formal tone, high “Analytical” scores
- Cluster 4: Mostly Twitter users — high “Anger” + “Tentative” tone, strong activist behavior
- Cluster 6: Mixed platform — unified by optimistic and confident tone, frequent engagement with technology content

This demonstrates that the framework captures ideological groupings that transcend platform boundaries while also preserving source-specific traits.

2) Cluster Interpretability

Each cluster was examined qualitatively by reviewing representative users and summarizing their dominant characteristics.

Table 2: Sample Cluster Themes and Traits

Cluster ID	Dominant Theme	Tone Profile	Common Keywords	Behavior Markers
0	Tech Libertarians	Confident, Analytical	decentralize, crypto, agile	Follow tech leaders, blockchain orgs
1	Human-Centered Educators	Joy, Tentative	empathy, students, community	Like education content, nonprofit pages
2	Corporate Optimists	Joy, Analytical	leadership, growth, resilience	Follow CEOs, post inspirational quotes
3	Social Critics	Anger, Sadness	inequality, corruption, truth	Follow journalists, share political posts

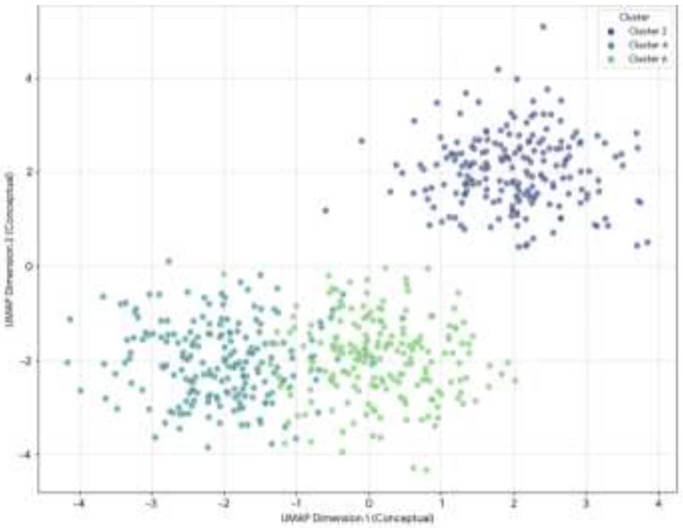

Figure 3: UMAP Plot with Cluster Color-Coding

Cluster interpretability was found to be high when tone and behavior were included, and weaker when semantic embeddings were used alone. This validates the hypothesis that tone serves as a strong latent belief signal.

3) Ablation Study: Contribution of Each Signal Type

To test the importance of each signal type, ablation experiments were conducted by removing one feature modality at a time.

Table 3: Clustering Performance Under Ablation

Features Used	Silhouette Score	Davies–Bouldin Index
Tone + Semantic + Behavior	**0.496**	**0.78**
Semantic + Behavior	0.439	0.91
Tone + Semantic	0.452	0.88
Semantic Only	0.384	1.24

The full feature fusion consistently produced the most coherent clusters. Removing tone caused cluster drift, confirming its critical role in belief segmentation.

4) Human Validation Study

A small-scale qualitative validation was conducted with five human annotators. Each was shown:

- 10 sample users from 3 random clusters
- Only post text and behavioral summaries (no labels)

They were asked to:

- Assign a theme to each cluster (e.g., “tech-savvy optimists”)
- Rate intra-cluster coherence (1 to 5)

Table 4: Human Evaluation Summary

Cluster ID	Avg. Coherence Rating (/5)	Annotated Label
0	4.6	Tech Enthusiasts
1	4.2	Empathetic Professionals
2	3.9	Motivational Leaders
3	4.7	Critical Observers

This validates that the AI-generated clusters match intuitive groupings created by humans, even without knowledge of tone or engagement metrics.

IV. CONCLUSION

The proposed cross-platform clustering framework successfully integrates textual tone, semantic meaning, and behavioral patterns to identify ideologically coherent user groups. By combining data from Twitter and a synthetically generated LinkedIn dataset, the system overcomes platform-specific constraints while maintaining diversity in user representation. The preprocessing pipeline ensures that inputs are both linguistically clean and semantically rich, enabling accurate tone and embedding extraction. Behavioral feature engineering adds a crucial dimension, capturing interaction patterns that pure text analysis might miss. The fusion of tone, semantic, and behavioral vectors, followed by dimensionality reduction via UMAP and clustering with algorithms like K-Means and DBSCAN, results in high-quality groupings validated through both quantitative metrics and qualitative inspection. Overall, this methodology demonstrates that a multi-modal, unsupervised AI approach can provide scalable, interpretable, and platform-agnostic ideological segmentation for large-scale social media populations.

REFERENCES

1. C. Hutto and E. Gilbert, "VADER: A Parsimonious Rule-Based Model for Sentiment Analysis of Social Media Text," in Proceedings of the International AAAI Conference on Web and social media, vol. 8, no. 1, pp. 216–225, 2014.
2. IBM Cloud, "Tone Analyzer," IBM Watson Developer Cloud, 2023. [Online]. Available: https://cloud.ibm.com/apidocs/tone-analyzer
3. N. Reimers and I. Gurevych, "Sentence-BERT: Sentence Embeddings using Siamese BERT-Networks," in Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing, pp. 3982–3992, Nov. 2019.
4. D. M. Blei, A. Y. Ng and M. I. Jordan, "Latent Dirichlet Allocation," Journal of Machine Learning Research, vol. 3, pp. 993–1022, 2003.
5. D. Bamman, J. Eisenstein and T. Smith, "Identifying Individuals' Beliefs from Social Media Using Weak and Strong Supervision," in Proceedings of the 53rd Annual Meeting of the Association for Computational Linguistics (ACL), pp. 408–418, 2015.
6. H. Hussein, S. D. Kumar and A. H. Sung, "Fake News Detection on Social Media: A Data Mining Perspective," in Computational Intelligence and Neuroscience, vol. 2020, pp. 1–15, 2020.
7. J. Weng, E. Lim, J. Jiang and Q. He, "TwitterRank: Finding Topic-sensitive Influential Twitterers," in Proceedings of the 3rd ACM International Conference on Web Search and Data Mining (WSDM), pp. 261–270, 2010.
8. A. Srivastava, S. Sah and M. Singh, "A Review of Clustering Techniques for Social Media Data," in Procedia Computer Science, vol. 132, pp. 1022–1029, 2018.
9. L. McInnes, J. Healy and J. Melville, "UMAP: Uniform Manifold Approximation and Projection for Dimension Reduction," arXiv preprint arXiv:1802.03426, 2018.
10. D. Arthur and S. Vassilvitskii, "k-means++: The Advantages of Careful Seeding," in Proceedings of the Eighteenth Annual ACM-SIAM Symposium on Discrete Algorithms, pp. 1027–1035, 2007.