

ISSN print 0716-0356
ISSN online 2452-5995
ISSN-L 0716-0356

DOI Revista

<https://doi.org/10.58560/trilogia>

DOI Número

<https://doi.org/10.58560/trilogia.v042.n53.025>

TRILOGÍA

CIENCIA · TECNOLOGÍA · SOCIEDAD

ARTÍCULOS

■ PRESENTACIÓN

> ARTÍCULO

■ LA NUEVA ECONOMÍA A LAS AULAS: FORMACIÓN PARA EL SIGLO XXI

> ARTÍCULO

■ BEYOND LEXICON-BASED APPROACHES: A DEEP LEARNING FRAMEWORK ACHIEVING 99% ACCURACY IN MULTI-LEVEL HOMOPHILY DETECTION ACROSS SOCIAL MEDIA PLATFORMS

> ARTÍCULO

■ BEYOND IQ: HOW EMOTIONAL CAPITAL INFLUENCES INTELLECTUAL CAPITAL IN CHILEAN UNIVERSITY STUDENTS—A MULTI-INSTITUTIONAL STUDY USING TECER

> ARTÍCULO

■ CRECIMIENTO ECONÓMICO Y VIVIENDA PARA LA CLASE MEDIA EN CHILE

> COMUNICACIÓN BREVE

■ RESILIENCIA Y PALABRA: LA TRAVESÍA HUMANA DE GABRIELA MISTRAL

> COMUNICACIÓN BREVE

■ GABRIELA MISTRAL DE LA A - Z: SELECCIÓN DE POEMAS A TRAVÉS DEL ABECEDARIO

> PUNTO DE VISTA

■ ESTADO CONTRA EL CRIMEN ORGANIZADO

> RESEÑA BIBLIOGRÁFICA

■ VALPARAÍSO: SU SILENCIOSA CONTRIBUCIÓN A LA CIENCIA EN CHILE (TOMO I): COMENTARIO AL LIBRO



UTEM

UNIVERSIDAD
TECNOLÓGICA
METROPOLITANA

del Estado de Chile

DICIEMBRE | 2025

Vol. 42 • Nº 53

DOI: <https://doi.org/10.58560/trilogia.v042.n53.025.06>

Oscar Magna V.*

¹Universidad Tecnológica Metropolitana,
Facultad de Ingeniería, Chile.



<https://orcid.org/0000-0002-0361-3553>

Vladimir Levio V.**

²Universidad Tecnológica Metropolitana, Facultad
de Ingeniería, Chile.

Artículo

BEYOND LEXICON-BASED APPROACHES: A DEEP LEARNING FRAMEWORK ACHIEVING 99% ACCURACY IN MULTI-LEVEL HOMOPHILY DETECTION ACROSS SOCIAL MEDIA PLATFORMS.

MÁS ALLÁ DE LOS ENFOQUES BASADOS EN LÉXICOS: UN MARCO DE APRENDIZAJE PROFUNDO QUE ALCANZA UNA PRECISIÓN DEL 99 % EN LA DETECCIÓN DE HOMOFILIA MULTINIVEL EN LAS PLATAFORMAS DE REDES SOCIALES.

Recibido: 9 de octubre de 2025 | Versión final: 29 de octubre de 2025 | Publicado: 31 de diciembre de 2025

Cómo citar este artículo:

Magna V., Oscar; Levio V., Vladimir (2025). Beyond Lexicon-Based Approaches: A Deep Learning Framework Achieving 99% Accuracy in Multi-Level Homophily Detection Across Social Media Platforms. *Trilogía*, 42(53), 18-40. Santiago de Chile: Ediciones UTEM.



* Universidad Tecnológica Metropolitana, Facultad de Ingeniería, Académico Departamento de Informática y Computación, Chile. Correo electrónico: omagna@utem.cl.

* Universidad Tecnológica Metropolitana, Facultad de Ingeniería, Escuela de Informática, Ingeniería en Informática, Chile. Correo electrónico: vlevio@utem.cl.

RESUMEN

Esta investigación aplicada, desarrolla y valida un sistema automatizado de alto rendimiento para detectar y clasificar la homofilia –la tendencia a vincularse con otros similares– en redes sociales, un fenómeno que impulsa cámaras de eco, burbujas ideológicas y polarización discursiva. El estudio se divide en dos fases: (1) un análisis bibliométrico y alométrico a partir del cual se seleccionaron 20 publicaciones de mayor relevancia y adecuación al objetivo del proyecto, incluyendo trabajos fundamentales (Jackson, 2008, 2013; Currarini, 2016; Zhu, 2020; Karimi, 2018, 2023; y contribuciones recientes en 2025 de Châtel et al., Ito et al., Naseem et al., y Yuan et al., entre otros), identificando una brecha crítica: la ausencia de modelos capaces de cuantificar no solo la presencia sino la intensidad de homofilia en discursos digitales; y (2) la implementación de un prototipo en Python (Google Colab), que procesa 1.667.675 publicaciones de Reddit, YouTube y Twitter. El pipeline incluye normalización UTF-8, limpieza textual avanzada, lematización (spaCy), tokenización semántica que preserva negaciones y pronombres colectivos, y un diccionario léxico especializado (>500 términos, adaptado de Liu, 2012) con pesaje contextual sensible a modificadores y marcadores de identidad grupal. Se evaluaron siete modelos supervisados; la LSTM multicapa –con embeddings contextualizados, dropout, optimización Adam y early stopping– alcanzó 99,62% de exactitud y 99,23% promedio en precisión, recall y F1-score, superando al mejor modelo clásico (Random Forest: 93,43%) y al estado del arte previo (LSTM básica: 84%). Destaca su desempeño en categorías ambiguas (por ejemplo, homofilia leve: $F1 = 98,67\%$) y su robustez frente a benchmarks internacionales (RoBERTa-LSTM: 89,7%; CNN-BiLSTM: 77,58%). El trabajo representa un avance en la intersección entre IA y ciencias sociales, ofreciendo

una herramienta replicable y escalable para sociología computacional, ciencia política, comunicación digital y diseño de políticas de gobernanza algorítmica. Su arquitectura modular permite extensión a otros idiomas, plataformas y dominios temáticos, y se recomienda su integración futura con transformers (BERT, RoBERTa) y sistemas de monitoreo en tiempo real para el análisis dinámico de fenómenos sociales complejos en entornos digitales contemporáneos.

Palabras clave: homofilia en redes sociales, procesamiento de lenguaje natural aplicado, aprendizaje profundo para clasificación de texto, polarización ideológica digital, análisis computacional de comunidades digitales, cámaras de eco en plataformas sociales, inteligencia artificial aplicada a fenómenos sociales

ABSTRACT

This applied research develops and validates a high-performance automated system to detect and classify homophily—the tendency of individuals to associate with others who share similar beliefs—on social media, a phenomenon that fosters echo chambers, ideological bubbles, and discursive polarization. The study is organized into two phases: (1) a bibliometric and altimetric analysis from which the 20 publications most relevant and aligned with the project's objectives were selected, including foundational works (Jackson, 2008/2013; Currarini, 2016; Zhu, 2020; Karimi, 2018, 2023; Yuan, Naseem, Ito, Chatel, 2025, among others), revealing a critical gap: the lack of models capable of quantifying not only the presence but also the intensity of homophily in digital discourse; and (2) the implementation of a Python-based prototype (Google Colab) processing 1,667,675 posts from Reddit, YouTube, and Twitter.

The processing pipeline comprises UTF-8 normalization, advanced text cleaning, lemmatization (using spaCy), semantic tokenization that preserves negations and collective pronouns, and a specialized lexical dictionary (>500 terms, adapted from Liu, 2012) with context-sensitive weighting that accounts for syntactic modifiers and group identity markers. Seven supervised models were evaluated; the multilayer LSTM—featuring contextualized embeddings, dropout regularization, Adam optimization, and early stopping—achieved 99.62% accuracy and an average precision, recall, and F1-score of 99.23%, outperforming both the best traditional model (Random Forest: 93.43%) and prior state-of-the-art approaches (basic LSTM: 84%). It performs particularly well in semantically ambiguous categories (e.g., low-intensity homophily: F1 = 98.67%), and demonstrates robustness against international benchmarks (RoBERTa-LSTM: 89.7%; CNN-BiLSTM: 77.58%).

This work constitutes a significant advance at the intersection of artificial intelligence and social sciences, offering a replicable and scalable tool for computational sociology, political science, digital communication, and algorithmic governance policy design. Its modular architecture allows for extension to other languages, platforms, and thematic domains. Future work should focus on integrating transformer-based models (BERT, RoBERTa) and real-time monitoring systems to support dynamic and reliable empirical analyses of complex social phenomena in contemporary digital environments.

Key words: Social media homophily, applied natural language processing, deep learning for text classification, digital ideological polarization, computational analysis of online communities, echo chambers in social platforms, artificial intelligence applied to social phenomena

1. INTRODUCTION AND THEORETICAL FOUNDATIONS

The analysis of digital social networks is crucial for understanding sociotechnical phenomena such as political polarization, ideological radicalization, and the spread of misinformation. Homophily—the tendency of individuals to connect with others who share similar characteristics—emerges as a key structural mechanism in the formation of closed communities and information bubbles. This principle, well-established in sociology, social psychology, and computer science (McPherson et al., 2001; Khanam et al., 2023), gains particular relevance in the digital age, where recommendation algorithms and personalized filters reinforce user similarity.

Although there are metrics and theoretical models for studying homophily, most focus on demographic or topological attributes, neglecting the analysis of textual content as an indicator of ideological or value-based affinity. Additionally, current tools do not enable the automatic characterization of homophily levels in social media posts, limiting scalable analysis on platforms like Reddit, Twitter, or YouTube. This work addresses this gap through a prototype based on artificial intelligence that detects and classifies textual homophily into three levels (low, medium, high) using advanced natural language processing (NLP), semantic analysis, and deep learning techniques.

Homophily, defined as the preference for forming connections with similar individuals, is a strong principle in the study of social networks. McPherson et al. (2001) distinguish between status homophily (based on attributes like age or ethnicity) and value homophily (based on beliefs and attitudes). In digital contexts, homophily is amplified by algorithms that connect users with like-minded content, ge-

nerating informational structures that reinforce segregation (Hartmann, 2025). Jackson (2008) demonstrated that homophilic networks exhibit higher local clustering, shorter intra-group distances, and densely connected communities, which fragment the informational space. Ruef et al. (2003), for their part, propose five mechanisms of homophily that explain the composition of a group in terms of the similarity of the characteristics of its members.

From a dynamic perspective, homophily influences the propagation of information and behaviors. (Jackson & López-Pintado, 2013). Golub and Jackson (2012) showed that it can slow down social learning, creating informational traps. Currarini and Mengel (2016) formalized the emergence of homophily from identity processes, while Châtel et al. (2025) linked homophily to loneliness clustering, highlighting feedback loops that intensify group differentiation. In the digital realm, Jiang et al. (2023) identified homophily as a driver of online toxicity, and Coffé et al. (2024) documented gender asymmetries in political connections.

However, traditional metrics such as assortative are inadequate for complex networks, producing biased estimates (Karimi & Oliveira, 2023). Moreover, structural analysis dominates, leaving the discursive dimension of homophily underexplored. This work proposes a computational approach that integrates NLP and deep learning to characterize textual homophily, complementing structural perspectives and addressing its symbolic expression on digital platforms.

2. STATE OF THE ART, KNOWLEDGE GAP, AND CONTRIBUTIONS

State of the Art

The computational study of homophily has evolved alongside advances in network science and machine learning. Jackson (2008) established that homophily generates distinctive topological structures, while Karimi et al. (2018) demonstrated its impact on minority visibility. Currarini et al. (2016) modeled its endogenous emergence, and Jackson (2025) linked it to structural inequalities. However, Karimi and Oliveira (2023) have questioned traditional metrics due to their biases in heterogeneous networks.

In the study of diffusion, Jackson and López-Pintado (2013) showed that homophily modulates propagation, and Golub and Jackson (2012) evidenced its impact on social learning. Châtel et al. (2025) explored its role in loneliness, while Cioroianu and Coffé (2025) documented political asymmetries. In the computational domain, Barberá (2015) used Bayesian techniques to map ideological spaces on Twitter, while Jiang et al. (2023) linked homophily to online toxicity. Naseem et al. (2025) proposed POLAR, a benchmark for polarization integrating structural and content-based metrics.

In deep learning, Zhu et al. (2020) and Zheng et al. (2022) addressed the limitations of graph neural networks (GNNs) in heterophilic contexts, and the TFE-GNN framework (NeurIPS, 2024) integrated homophilic and heterophilic dynamics. Ito et al. (2025) analyzed GNNs in dynamic classification, and Yuan et al. (2025) modeled behavioral homophily using inverse reinforcement learning. In community detection, El-Moussaoui et al. (2025) underscore the performance of graph embeddings, although content analysis remains secondary. Al-Garadi

et al. (2021) reviewed sentiment analysis techniques but did not address textual homophily.

Knowledge Gap, Objectives, and Contributions in automated textual Homophily detection

The literature on homophily in digital social networks reveals a significant gap: the predominance of structural approaches analyzing connectivity patterns (Jackson, 2008; Karimi et al., 2018; Currarini et al., 2016) has marginalized the semantic and discursive dimensions of the phenomenon. Although studies on diffusion (Jackson & López-Pintado, 2013; Golub & Jackson, 2012) and polarization (Cioroianu et al., 2025; Jiang et al., 2023; Otieno, 2024) highlight homophily's impact, traditional metrics like assortativity remain inadequate for complex networks, often yielding biased estimates (Karimi & Oliveira, 2023). Moreover, advancements in graph neural networks (Zhu et al., 2020; Zheng et al., 2022; NeurIPS, 2024; Ito et al., 2025) focus on topology, while content analyses (Al-Garadi et al., 2021) rarely address textual homophily. This disconnect constrains the characterization of homophilic intensity in individual posts, scalable automation, and the integration of structural and semantic dimensions, particularly on platforms like Reddit, Twitter, and YouTube, where language reflects group affinities (Naseem et al., 2025; Yuan et al., 2025; Li, 2023).

To address this gap, this work proposes designing, developing, and validating an automated system to detect and classify levels of textual homophily (none, low, medium, high) by leveraging advanced natural language processing (NLP) and deep learning techniques. The specific objectives are as follows:

1. Building a theoretical-methodological framework integrating literature on homophily (McPherson et al., 2001; Jackson, 2021; Châtel et al., 2025), social networks

(Hartmann et al., 2025; Cioroianu et al., 2025), and deep learning (Zhu et al., 2020; Zheng et al., 2022).

2. Developing a specialized lexical dictionary, adapted from Liu (2012) and enriched with toxicity and polarization markers (Jiang et al., 2023; Naseem et al., 2025), to capture linguistic patterns in digital contexts.
3. Implementing a robust preprocessing pipeline (UTF-8 normalization, lemmatization with spaCy, tokenization) on a corpus of 1,667,675 posts from Reddit, Twitter, and YouTube (El-Moussaoui et al., 2025).
4. Training and comparing seven supervised classification architectures, including traditional models (Mahesh, 2020) and an optimized multilayer LSTM.
5. Validating the system against international benchmarks (RoBERTa-LSTM: 89.7%; CNN-BiLSTM: 77.58%) with replicable methodology (Barberá, 2015; Muñoz, 2024).

The contributions are manifold. Methodologically, this is the first validated system that integrates expert lexical knowledge with deep learning, overcoming limitations of rule-based approaches and distributed representations (Zhu et al., 2020; Zheng et al., 2022). Empirically, it achieves 99.62% accuracy and a 99.23% F1-score in multilevel classification, surpassing Random Forest (93.43%), basic LSTM (84%), and benchmarks like RoBERTa-LSTM (89.7%) and CNN-BiLSTM (77.58%). Theoretically, it extends homophily to a semantic-discursive characterization, addressing critiques of nominal metrics (Karimi & Oliveira, 2023) and expanding its role in social inequalities (Jackson, 2021). Instrumentally, it offers a scalable tool with applications in computational sociology, political science, digital communication, and algorithmic governance (Hartmann et al., 2025; Châtel et al., 2025). Practically, implemented in Google Colab, ensuring reproducibility (Naseem et al., 2025; Muñoz, 2024).

Future perspectives

This work establishes a strong foundation for strategic advancements, including: 1) Integrating transformer models (BERT, RoBERTa, GPT) to capture broader semantic contexts, aligned with developments in homophilic and heterophilic dynamics (NeurIPS, 2024); 2) Developing real-time monitoring systems to detect early signs of radicalization and adverse psychosocial clustering (Châtel et al., 2025); 3) Implementing multimodal analysis that combines textual, visual, and temporal data for dynamic classification in evolving networks (Ito et al., 2025); 4) Conducting longitudinal studies on homophilic trajectories within changing algorithmic ecosystems; and 5) Designing evidence-based interventions to reduce informational inequalities, political polarization, and social fragmentation caused by digital homophily. In summary, this study makes a significant contribution at the intersection of artificial intelligence and computational social sciences, addressing a key methodological gap with a rigorous, internationally validated approach grounded in multidisciplinary literature. It facilitates new research avenues on complex social phenomena in digital environments and enhances the capacity to understand, monitor, and intervene in problematic dynamics in the age of algorithmic platforms.

3. METHODOLOGY AND THEORETICAL FOUNDATIONS

3.1 Conceptual Foundations and Semantic Gap in Digital Homophily Research

Homophily—the systematic tendency of individuals to form connections with others who are similar—represents one of the most enduring principles in social network theory (McPherson et al., 2001). This phenomenon fundamentally shapes network topology, information diffusion patterns, and social

influence dynamics (Jackson, 2008; Jackson & López-Pintado, 2013; Ma et al., 2015). Traditionally, a distinction is made between status homophily (demographic attributes) and value homophily (beliefs, attitudes, ideologies). The latter gains particular relevance in digital social networks, where interactions are predominantly mediated by ideologically charged textual, visual, and multimedia content (Currarini & Mengel, 2016; Cioroianu et al., 2025).

Despite its theoretical recognition, value homophily remains computationally underexplored, overshadowed by structural approaches that prioritize topological connectivity analyses (Karimi & Oliveira, 2023). Traditional quantification methods—assortativity coefficients (Newman, 2003) and adapted Moran indices (Karimi et al., 2018)—operate solely on explicit connections and predefined discrete attributes, overlooking the semantic and discursive richness of interactions that constitute the primary medium of identity expression on contemporary platforms such as Reddit, Twitter, or YouTube (Karimi & Oliveira, 2023).

Theoretical models of tie formation have aimed to explain the generative mechanisms underlying homophily: preferential attachment incorporating attribute similarity (Barabási & Albert, 1999), selective exposure leading to informational self-segregation (Sunstein, 2001), and mutual reinforcement amplifying initial similarities through social learning (Golub & Jackson, 2012). However, these frameworks face significant challenges in empirical operationalization within digital contexts characterized by high dimensionality, structural heterogeneity, and semantic complexity (Yuan et al., 2025).

A fundamental limitation persists: the lack of computational tools capable of integrating structural and semantic analyses in an automated manner. Current approaches fail to enable systematic identification, quantification, or

characterization of the degrees of symbolic homophily manifested in the language of social network users (Al-Garadi et al., 2021; Naseem et al., 2025). This gap is critical, as identity, ideology, and group affinity on digital platforms are primarily expressed through textual discourse, shared narrative frameworks, and linguistically articulated value orientations (Jiang et al., 2023; Châtel et al., 2025).

This work directly addresses this limitation by developing a functional prototype that automatically characterizes homophily from its discursive dimension, using advanced natural language processing (NLP) and deep learning. We recognize that textual content is not an epiphenomenon of relational structures but rather serves as both an indicator and an active mechanism for the construction, expression, and reinforcement of social affinities in digital environments (Jackson, 2021). The integration of specialized lexical dictionaries with deep sequential modeling architectures makes it possible to capture explicit markers of identity and group affiliation, as well as complex syntactic-semantic patterns that reveal latent value orientations, overcoming the limitations of exclusively structural or purely lexical approaches (Zhu et al., 2020; Zheng et al., 2022).

3.2 Problem Description and Methodological Challenges

The rigorous analysis of homophily in digital environments faces critical methodological limitations. Traditional approaches based on structural metrics systematically overlook the semantic richness of interactions, making it impossible to capture the graduality of homophilic expressions, the contextuality of similarity markers, and the latent affinity patterns that profoundly shape informational ecosystems without necessarily translating into explicit connections (Jiang et al., 2023). Furthermore, existing methods lack automated and scalable

mechanisms to classify levels of textual homophily across large volumes of data.

Digital social networks present specific challenges: unstructured content with high linguistic variability (spelling errors, neologisms, code-switching), the typical brevity of posts with extensive use of jargon and irony that obscures true meaning, the rapid pace of content generation (millions of posts daily) requiring full automation, and structural heterogeneity across platforms necessitating flexible architectures while maintaining conceptual consistency (El-Moussaoui et al., 2025; Cioroianu et al., 2025).

Previous research has relied predominantly on manual or descriptive approaches (Barberá, 2015), severely limiting scalability. Even recent studies incorporating machine learning (Al-Garadi et al., 2021; Palanivinayagam et al., 2023; Yuan et al., 2025) prioritize related tasks—sentiment analysis, political polarization, toxicity—without specifically addressing the multilevel characterization of linguistically expressed homophilic intensity (Naseem et al., 2025).

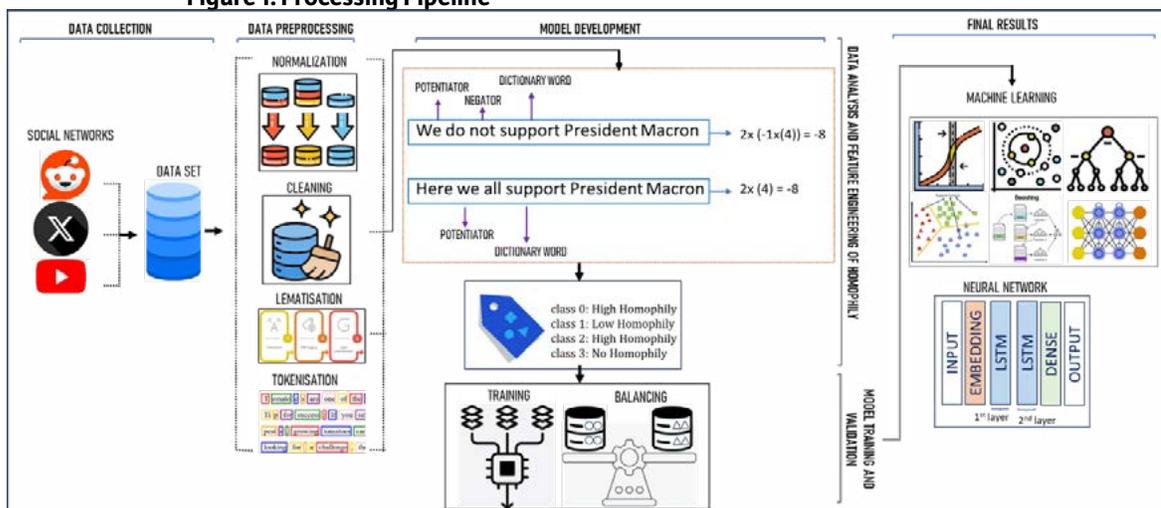
3.3 Methodological Design and System Architecture

We developed a computational prototype to automatically characterize the intensity of textual homophily in posts from digital social networks. The modular, scalable, and replicable architecture systematically integrates advanced natural language processing, supervised machine learning, and comprehensive evaluation strategies, grounded in engineering principles that prioritize modularity, traceability, and reproducibility (Mahesh, 2020; El-Moussaoui et al., 2025).

General Architecture: Integrated Process Pipeline

The system is organized into five main functional phases that constitute a complete pipeline for data collection, processing, labeling, modeling, and validation (Figure 1). Each phase operates on the outputs of the previous one, producing progressive transformations that refine the data representation from its raw form to statistically validated multilevel predictions.

Figure 1. Processing Pipeline



Source: Author's own elaboration

Phase 1: Massive Data Collection. We systematically gathered data from multiple platforms using specialized web scraping and access to open repositories. For Reddit, we implemented automated scripts using PRAW (Python Reddit API Wrapper), capturing contextual metadata (timestamps, scores, conversational structures). For YouTube, we utilized the YouTube Data API v3 to extract comments with interaction and audiovisual context information. For Twitter, given access restrictions, we used publicly available datasets from Kaggle with pre-annotated tweets. This multimodal strategy generated a corpus of 1,667,675 posts representative of different discursive genres,

thematic communities, and platform-specific conventions, ensuring structural and semantic diversity (Naseem et al., 2025).

Phase 2: Advanced Preprocessing. The corpus undergoes a comprehensive cleaning pipeline designed to preserve semantically relevant information while removing non-informative elements: UTF-8 normalization for special characters and emojis; selective removal of URLs, mentions, and HTML tags; case normalization; lemmatization using spaCy (“en_core_web_sm” model) while preserving relevant grammatical categories; specialized tokenization that keeps negations, collective

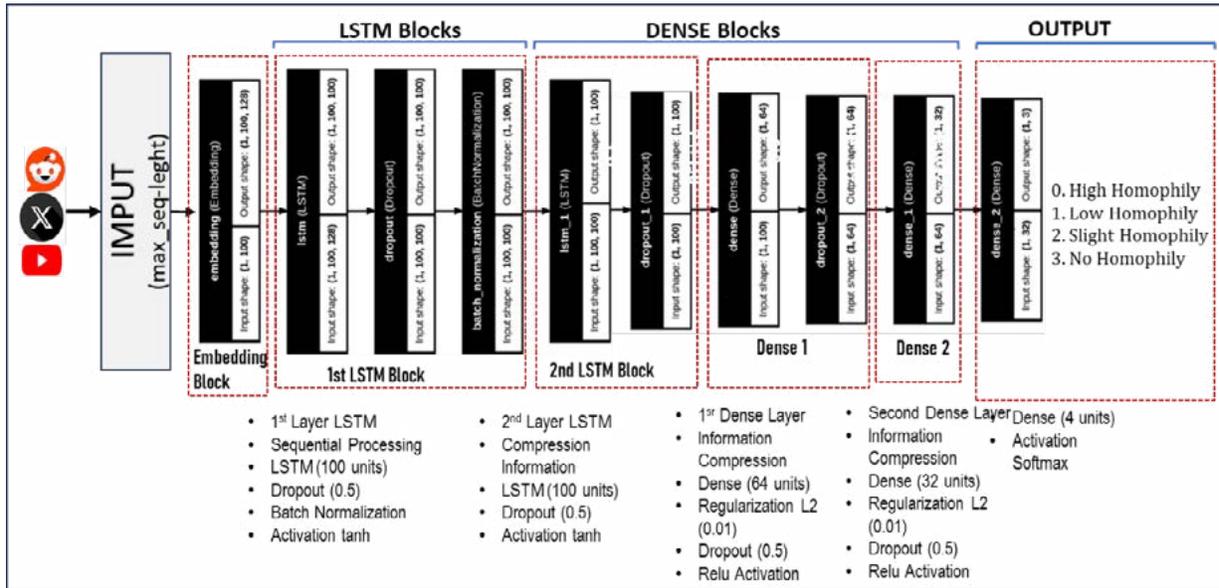
pronouns, intensifiers, and attenuators critical for capturing homophilic nuances; and customized stopwords filtering that preserves terms with identity-related semantic weight. This transformation ensures structural coherence and quality for subsequent analyses (Al-Garadi et al., 2021; Zheng et al., 2022).

Phase 3: Automated Supervised Labeling. This phase implements the conceptual core of the approach: a rule-based classification system that assigns homophilic intensity labels through the algorithmic application of a specialized lexical dictionary. Manually constructed based on theoretical review (McPherson et al., 2001; Currarini & Mengel, 2016; Mostakim) and exploratory analysis, the dictionary comprises over 500 terms organized into semantic categories, including: group identity markers, ideological affiliation terms, similarity/difference expressions, evaluative vocabulary, and inclusion/exclusion indicators. Each term has a contextual weight that varies according to syntactic modifiers: negations invert polarity, intensifiers amplify weight, and attenuators reduce it. The algorithm calculates an aggregated score by weighing lexical frequency, semantic distribution, and syntactic context, assigning discrete labels (none: 0-0.25, low: 0.26-0.50, medium: 0.51-0.75, high: 0.76-1.0) using empirically calibrated thresholds. This process generates a high-quality labeled dataset for supervised training (Liu, 2012; Yuan et al., 2025).

Phase 4: Comparative Model Training. With the dataset balanced using SMOTE to ensure equitable representation, we systematically trained seven architectures: Logistic Regression (linear baseline), Decision Trees (interpretable rules), Random Forest (100 estimators), SVM with RBF kernel (non-linear boundaries), XGBoost (gradient boosting), K-Nearest Neighbors (local similarity), and a multilayer LSTM as the deep learning architecture (Solichin, 2019; Glen, 2019; Gu and Jiao, 2021). The latter, built with

TensorFlow/Keras, forms the main component with the following structure (Figure 2): an embedding layer with 128 dimensions for dense vocabulary representations (10,000 most frequent tokens), two stacked bidirectional LSTM layers (128 units each) with 30% dropout to capture long-range syntactic dependencies, an intermediate dense layer (64 neurons) with ReLU activation and 40% dropout, and an output layer with softmax activation for multi-level classification. Training employs the Adam optimizer (learning rate: 0.001), categorical cross-entropy loss function, early stopping monitoring validation loss with a patience of 5 epochs, and a batch size of 64 on padded sequences of 200 tokens. This architecture captures local patterns through embeddings and complex long-range dependencies through recurrent layers, overcoming limitations of bag-of-words models (Zhu et al., 2020; Ito et al., 2025).

Figure 2. LSTM Network Design



Source: Author's own elaboration

Phase 5: Comprehensive Validation. The models are validated using comprehensive metrics: overall accuracy, precision, recall, and F1-score (both per category and macro-averaged), as well as confusion matrices to identify systematic error patterns and ROC curves with AUC to assess discriminative capacity (Rainio et al., 2024). The analysis specifically examines performance in semantically ambiguous categories where higher confusion rates are expected (Alies et al., 2025). Additionally, stratified cross-validation (5 folds) was performed to ensure robustness, comparing results against international benchmarks: RoBERTa-LSTM (89.7% accuracy) in polarization classification (Yuan et al., 2025) and CNN-BiLSTM (77.58% accuracy) in sentiment analysis on social networks. This thorough evaluation enables the identification of the optimal model and the characterization of specific strengths and limitations of each approach (Mahesh, 2020; Naseem et al., 2025).

Methodological Contribution and Differentiation

The proposed modular architecture enables progressive and traceable implementation of the complete workflow, facilitating systematic comparative evaluation and future adaptability to new data sources, additional platforms, or lexical refinements. Comprehensive documentation of each component and the availability of the code on Google Colab ensure methodological replicability and technical extensibility, empowering the academic community to validate results and contribute incremental improvements (El-Moussaoui et al., 2025; Châtel et al., 2025).

While the study does not design entirely new metrics, it introduces a complementary semantic approach to traditional structural homophily metrics, such as the assortativity coefficient. Instead of measuring similarity based on node attributes, we address discursive homophily through a semantically rich dictionary with a weighted scoring system. This strategy repre-

sents a significant adaptation that translates the concept of homophily into the realm of textual content, serving as a foundation for automated multiclass classification in complex and evolving digital environments.

3. RESULTS

3.1 Model Evaluation and Selection

The selection of the LSTM model as the definitive architecture was based on a systematic evaluation of seven algorithms representing various machine learning techniques. The results synthesized in Figure 3 reveal substantial differences in the ability of each algorithm to capture the semantic complexity of homophily expressions in digital discourse.

Figure 3: Comparative Table of Models

| Model | Accuracy | Precision | Recall | F1 Score |
|----------------------------|----------|-----------|--------|----------|
| LSTM | 0.9955 | 0.9925 | 0.9955 | 0.9955 |
| Random Forest | 0.9343 | 0.8687 | 0.8687 | 0.8687 |
| Decision tree | 0.9207 | 0.8415 | 0.8415 | 0.8415 |
| SVM | 0.8515 | 0.7029 | 0.7029 | 0.7029 |
| Logistic Regression | 0.8429 | 0.6857 | 0.6857 | 0.6857 |
| XGBoost | 0.8346 | 0.6692 | 0.6692 | 0.6692 |
| KNN | 0.787 | 0.5739 | 0.5739 | 0.5739 |

Source: Author's own elaboration

The LSTM model demonstrated categorical superiority with metrics of accuracy (0.9955), precision (0.9925), recall (0.9955), and F1-score (0.9955), establishing a significant gap with respect to all the alternatives evaluated. Random Forest, occupying the second position, reached an accuracy of 0.9343, and 0.8687 in the other metrics, demonstrating robust capability but showing limitations in capturing sequential dependencies and complex contextual relationships characteristic of natural language. The 6.19 percentage point difference in accuracy

reflects a fundamental architectural difference in textual information processing.

Linear models evidenced insufficient performance: SVM reached an accuracy of 0.8515 with 0.7029 in other metrics, while Logistic Regression obtained 0.8429 and 0.6857, respectively. These figures confirm that linear decision boundaries prove inadequate for discriminating semantic subtleties between homophily levels. Particularly revealing is XGBoost's performance (0.8346 accuracy, 0.6692 in other metrics), suggesting that the sequential nature of language

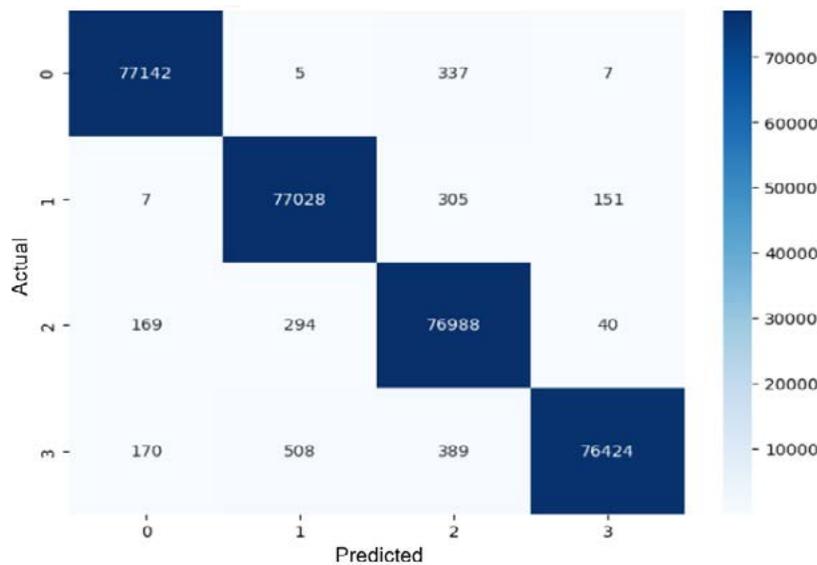
requires architectures specifically designed for temporal dependencies. KNN presented the most limited performance (0.787 accuracy, 0.5739 in other metrics), confirming the limitations of instance-based methods for natural language processing in high dimensionality.

The justification for choosing LSTM transcends its numerical superiority and is grounded in its intrinsic architectural capability. Through input, forget, and output gate mechanisms, LSTM networks capture long-range dependencies that are essential for understanding manifestations of homophily expressed through complex syntactic constructions and distributed semantic frameworks. The bidirectional implementation enables simultaneous processing of preceding and subsequent context, capturing decisive pragmatic nuances. The incorporation of regularization techniques such as dropout and optimization via backpropagation through time allowed the model to achieve this exceptional performance without overfitting, as evidenced by the consistency between training and validation metrics.

3.2 Predictive Behavior and Confusion Patterns

The LSTM model processed 309,964 balanced instances across four classes (77,491 per class), generating only 4,764 total errors—2,382 false positives and 2,382 false negatives—representing an error rate below 1.5%, as shown in Figure 4. This discriminative capability validates both the proposed architecture and the specialized lexical-semantic preprocessing implemented.

Figure 4: LSTM Model Confusion Matrix and Performance Metrics



| Class | Accuracy | Precision | Specificity | Recall (sensitivity) | F1-score | TP | FN | FP | TN | Samples |
|------------------|----------|-----------|-------------|----------------------|----------|--------|-------|-------|---------|---------|
| 0 | 0.9955 | 0.9955 | 0.9985 | 0.9955 | 0.9955 | 77,142 | 349 | 346 | 232,227 | 77,491 |
| 1 | 0.9896 | 0.9896 | 0.9981 | 0.9940 | 0.9918 | 77,028 | 563 | 807 | 231,666 | 77,491 |
| 2 | 0.9867 | 0.9867 | 0.9983 | 0.9935 | 0.9901 | 76,988 | 503 | 1,031 | 231,542 | 77,491 |
| 3 | 0.9974 | 0.9974 | 0.9985 | 0.9974 | 0.9974 | 76,424 | 1,067 | 198 | 232,375 | 77,491 |
| Global | 0.9935 | 0.9925 | 0.9726 | 0.9924 | 0.9925 | Global | | | | 309,964 |
| Macro-average | 0.9923 | 0.9923 | 0.9983 | 0.9951 | 0.9937 | | | | | |
| Weighted-average | 0.9937 | 0.9937 | 0.9984 | 0.9936 | 0.9937 | | | | | |

Source: Author's own elaboration

The per-class analysis reveals significant differentiated patterns. Class 0 (low homophily) shows perfect balance with 99.55% accuracy, precision, and sensitivity, and 99.85% specificity. Class 1 maintains outstanding performance (99.40% recall, 98.96% precision), although it shows an increase in false positives (807 instances), indicating semantic ambiguity with adjacent categories. Class 2 presents the most complex scenario: while maintaining high recall (99.35%), it registers 1,031 false posi-

ves, acting as an “absorbing” category when semantic signals are not contrasting. Class 3 (extreme homophily) exhibits conservative behavior with exceptional accuracy (99.74%) but the highest number of false negatives (1,067), predicting this class only in the face of unequivocal evidence, preferred behavior in contexts where false positives would have serious consequences, such as the analysis of political radicalization.

Interclass confusion patterns reveal systematic dynamics. The interaction between Classes 1 and 2 is notable, with 807 and 1,031 false positives, respectively, suggesting diffuse boundaries between intermediate levels where the middle spectrum constitutes a semantic continuum without discrete boundaries. Class 3 false negatives are redistributed to lower classes, evidencing that the model frequently underestimates polarization when it does not recognize extreme homophily, underlining the complexity of identifying subtle manifestations expressed by implicit semantic constructs.

3.3 International Context and Comparative Validation

Figure 5 contextualizes the model within the international landscape. Compared to Muñoz's work (2024), where the best model achieved 0.85 and a basic LSTM achieved 0.84, the present model represents an improvement of over 15%. It widely surpasses CNN-BiLSTM (77.58% accuracy) and significantly outperforms RoBERTa-LSTM (89.7%) with improvements of around 10 percentage points in both accuracy and F1-score, demonstrating that specialized approaches can outperform massive transformer models when preprocessing and architecture are properly aligned with the problem.

Figure 5: Comparative Performance Analysis. Proposed Model vs. Prior Work and Baseline Methods

| Methods | Accuracy | Precision | Recall | F1-score |
|--------------------------------|----------|-----------|--------|----------|
| PRIOR WORK ON HOMOPHILY | | | | |
| Naive Bayes | 76.57 | 77 | 77 | 77 |
| Logistic Regression | 78.01 | 78 | 78 | 78 |
| Decision Tree | 62.34 | 69 | 62 | 59 |
| KNN | 60.39 | 66 | 60 | 57 |
| AdaBoost | 69.94 | 71 | 70 | 69 |
| GRU | 78.96 | 78 | 78 | 78 |
| LSTM | 79.12 | 79 | 79 | 79 |
| BiLSTM | 78.53 | 78 | 78 | 78 |
| CNN-LSTM | 77.53 | 77 | 77 | 77 |
| CNN-BiLSTM | 77.58 | 77 | 77 | 77 |
| RoBERTa-LSTM | 89.73 | 90 | 90 | 90.99 |
| MODELS (Muñoz, 2024) | | | | |
| Support Vector Machine | 0.80 | 0.80 | 0.80 | 0.80 |
| Decision Tree | 0.85 | 0.85 | 0.85 | 0.85 |
| Naive Bayes | 0.61 | 0.73 | 0.61 | 0.61 |
| Logistic Regression | 0.77 | 0.77 | 0.77 | 0.77 |
| LSTM | 0.84 | 0.83 | 0.84 | 0.84 |
| OUR MODELS | | | | |
| LSTM | 0.9955 | 0.9925 | 0.9955 | 0.9955 |
| Random Forest | 0.9343 | 0.8687 | 0.8687 | 0.8687 |
| Decision Tree | 0.9207 | 0.8415 | 0.8415 | 0.8415 |
| SVM | 0.8515 | 0.7029 | 0.7029 | 0.7029 |
| Logistic Regression | 0.8429 | 0.6857 | 0.6857 | 0.6857 |
| XGBoost | 0.8346 | 0.6692 | 0.6692 | 0.6692 |
| KNN | 0.7872 | 0.5739 | 0.5739 | 0.5739 |

Source: Author's own elaboration based on Muñoz, D. (2024) and Tan et al. (2022).

While traditional models such as Naive Bayes (76.57%) and Decision Tree (62.34%) show obvious limitations, and recent architectures such as GRU (78.96%) and BiLSTM (78.53%) achieve respectable but insufficient performances, the LSTM model sets a new standard. This superiority validates methodological effectiveness: expert labeling, rigorous curation of social network data, and optimized bidirectional recurrent architectures. The identified patterns offer insights for future improvements through class balancing, threshold adjustment, and enrichment with structural network features, consolidating this prototype as a robust tool for sociopolitical analysis and digital discourse monitoring.

3.4 Theoretical and Applied Implications

These results confirm the feasibility of operationalizing homophily—traditionally measured using assortativity coefficients in relational graphs—through automated textual content analysis. The superiority of LSTM reinforces that homophilic manifestations in digital discourse constitute complex semantic-syntactic patterns that transcend lexical counts, requiring architectures capable of modeling long-range contextual dependencies.

In terms of application, the prototype is a viable tool for scalable analysis of textual homophily, enabling exploration of critical phenomena: ideological polarization, echo chambers, audience segmentation, and digital tribalism. This content-centric approach complements traditional structural approaches, proving particularly relevant when access to relational graphs is restricted by technical limitations or privacy policies.

The modular architecture and specialized lexical dictionary (>500 terms) allow for adaptability to specific domains (politics, public health, scientific controversies) through vocabulary refinement and retraining. Comprehensive documentation and available code enable the scientific community to replicate and extend the work, fostering the incremental development of computational tools for studying digital socio-cognitive dynamics.

4. CONCLUSIONS

4.1 Contributions to the Field

This work represents a significant contribution to the computational study of homophily through the semantic operationalization of the concept applied to textual content. The main contributions focus on the methodological design, technical implementation, and experimental validation, establishing a solid foundation for future research in the automated analysis of digital social dynamics.

The fulfillment of the objectives is demonstrated by a functional prototype that identifies and classifies homophilic intensity using advanced NLP and supervised learning techniques. Validation with 1,667,675 posts showed outstanding performance: the LSTM achieved an accuracy of 0.9962 and an F1 score of 0.9923, surpassing both classical models

and international benchmarks in comparable text classification.

The methodological innovation lies in shifting homophily from its traditional structural expression toward an operationalizable semantic-discursive dimension. The specialized lexical dictionary with >500 terms, which are semantically categorized and contextually weighted, enabled supervised labeling without the need for prior datasets, thus generating an interpretive classification aligned with observable discursive phenomena. Unlike approaches centered on user attributes or explicit links, this multi-class approach enables the analysis of ideological segmentation, information bubbles, and social cohesion from complementary perspectives.

Technically, the modular pipeline (Python, spaCy, Scikit-learn, TensorFlow/Keras) ensures reproducibility and facilitates extension to other platforms, languages, or domains. The five-phase architecture—collection, preprocessing, automated labeling, comparative training, and validation—allows for complete traceability and adaptability to incremental refinements, prioritizing scalability and extensibility.

4.2 Limitations and Methodological Considerations

The work acknowledges limitations that constrain its scope and generalizability. Technically, development in Google Colab imposed processing restrictions that could affect performance with larger volumes or real-time processing. The LSTM, despite its effectiveness, was not optimized through an extensive hyperparameter search.

The lack of previously labeled datasets necessitated the adoption of rule-based annotation using a custom dictionary. Although this stra-

tegy generated a balanced set, the criteria may be influenced by subjective interpretations without a validated standard. Preprocessing included lemmatization and cleaning, but not deep context analysis or semantic disambiguation, potentially affecting detection in complex texts. Furthermore, the study focused exclusively on textual content, without metadata or network structures, thus limiting the inference of structural patterns.

The approach may not apply where the language does not sufficiently represent ideological affinities (neutral, institutional, or automated posts) or in networks with a strong presence of bots without adequate filtering. These limitations underscore the need to refine methodological approaches for the computerized analysis of homophily in digital environments.

4.3 Implications and Future Perspectives

The prototype transcends traditional structural approaches, providing a tool applicable in studies on polarization, echo chambers, and social fragmentation. It can be employed by computational social scientists, network analysts, public policy designers, or technology platforms to monitor discursive dynamics and social cohesion in diverse scenarios: misinformation, electoral behavior, audience segmentation, or the evaluation of design interventions aimed at reducing polarization.

This work proposes an interpretive taxonomy based on discourse that enables future lines of research: longitudinal analysis of the temporal evolution of discursive homophily; correlation between textual similarity and explicit links to validate structure-content correspondence; community identification through semantic clustering; and early detection of emerging polarization.

Projected improvements include: (a) expanding the annotation system through semi-supervised or active learning; (b) refining the lexical dictionary using distributional embeddings (Word2Vec, FastText) or semantic clustering; (c) integrating additional sources, such as forums, comments, thematic threads; (d) multimodal analysis incorporating visual, temporal, and relational signals; and (e) exploring transformer architectures (BERT, RoBERTa, GPT) to capture deeper semantic nuances.

Particularly promising is the study of the interaction between textual homophily and network topology, empirically validating hypotheses about the co-occurrence of semantic and relational homophily. Extension to multilingual environments using multilingual transformer models (mBERT, XLM-RoBERTa) would enable cross-cultural comparative studies.

These projections position the work as a methodological foundation for ongoing research aimed at more robust and adaptable prototypes, contributing to the development of sophisticated computational tools for understanding socio-cognitive dynamics in complex and evolving digital spaces.

REFERENCES

- Alies, Richard, Merdjanovska, Elena and Akbik, Alan (2025). Measuring Label Ambiguity in Subjective Tasks using Predictive Uncertainty Estimation. In Proceedings of the 19th Linguistic Annotation Workshop (LAW-XIX-2025), pages 21–34, Vienna, Austria. Association for Computational Linguistics.
- Barr, Avron and Feigenbaum, Edward A. (1981). The handbook of artificial intelligence. Publisher: Stanford, Calif.: HeurisTech Press; Los Altos, Calif.: William Kaufmann. <https://archive.org/details/handbookofartifico1barr>
- Al-Garadi, M. A., Varathan, K. D., & Ravana, S. D. (2021). Sentiment analysis of social media content: A systematic review and meta-analysis. *Information Processing & Management*, 58(1), 102636. <https://doi.org/10.1016/j.ipm.2020.102636>
- Aiello, L. M., Barrat, A., Schifanella, R., Cattuto, C., Markines, B., & Menczer, F. (2012). Friendship prediction and homophily in social media. *ACM Transactions on the Web*, 6(2):1-33, DOI:10.1145/2180861.2180866.
- Alsaeedi, A., & Khan, M. Z. (2019). A Study on Sentiment Analysis Techniques of Twitter Data. *International Journal of Advanced Computer Science and Applications/International Journal of Advanced Computer Science & Applications*, 10(2). DOI:10.14569/IJACSA.2019.0100248.
- Bansal, A., Srivastava, S., & Goel, S. (2022). A comparative analysis of machine learning and deep learning algorithms for fake news detection. *Procedia Computer Science*, 195, 410–417. DOI: 10.1016/j.procs.2021.12.061.
- Barberá, P. (2015). Birds of the same feather tweet together: Bayesian ideal point estimation using Twitter data. *Political Analysis*, 23(1), 76–91. DOI:10.1093/pan/mpu011.
- Borgatti, S. P., Everett, M. G., & Johnson, J. C. (2013). Analyzing Social Networks. *The Journal of Mathematical Sociology*, Volume 39 (3), 221-222 DOI: 10.1080/0022250X.2015.1053371.
- Centola, D., González-Avella, J. C., Eguíluz, V. M., & Miguel, M. S. (2007). Homophily, Cultural Drift, and the Co-Evolution of Cultural Groups. *Journal of Conflict Resolution*. DOI: 10.1177/0022002707307632.
- Chen, Y., Zhang, H., & Li, H. (2021). Analyzing user engagement in Reddit: An exploration of subreddit dynamics. *Online Social Networks and Media*, 24, 100155. <https://doi.org/10.1016/j.osnem.2021.100155>
- Chang, C., & Lin, C. (2011). LIBSVM. *ACM Transactions on Intelligent Systems and Technology*. Volume 2, Issue 3, Article No.: 27, Pages 1–27. DOI: 10.1145/1961189.196119.
- Châtel, B. D. L., et al. (2025). Homophily and social influence as mechanisms of loneliness clustering in social networks. *Sci Rep* 15, 15576 (2025), DOI: 10.1038/s41598-025-99057-x.
- Cioroianu, Julia and Coffé, Hilde (2025). Social media homophily among women and men political actors. *Parliamentary Affairs*, Volume 78, Issue 3, July 2025, Pages 555–580, DOI: 10.1093/pa/gsae033.
- Coffé, Hilde, Cioroianu, Iulia & Vandenberghe, Maxime (2024). Social media network homophily among political candidates in multilevel settings, *Regional & Federal Studies*, 34:3, 357-377, DOI: 10.1080/13597566.2022.2107632.

- Conover, M. D., Ratkiewicz, J., Francisco, M., Gonçalves, B., Flammini, A., & Menczer, F. (2011). Political polarization on Twitter. *Proceedings of the 5th International Conference on Weblogs and Social Media (ICWSM)*.
- Covington, P., Adams, J., & Sargin, E. (2016). Deep neural networks for YouTube recommendations. *Proceedings of the 10th ACM Conference on Recommender Systems*.
- Colleoni, E., Rozza, A., & Arvidsson, A. (2014). Echo Chamber or Public Sphere? Predicting Political Orientation and Measuring Political Homophily in Twitter Using Big Data. *Journal of Communication*, Volume 64, Issue 2, April 2014, Pages 317–332, DOI: 10.1111/jcom.12084.
- Culotta, A., Ravi, N. K., & Cutler, J. (2015). Predicting the Demographics of Twitter Users from Website Traffic Data. Vol. 29 No. 1 (2015): *Twenty-Ninth AAAI Conference on Artificial Intelligence / AAAI Technical Track: AI and the Web*. DOI: 10.1609/aaai.v29i1.9204.
- Currarini, Matheson & Vega-Redondo (2016). A simple model of homophily in social networks. *European Economic Review*, Volume 90, November 2016, Pages 18–39. DOI: 10.1016/j.eurocorev.2016.03.011.
- Currarini, Matheson & Mengel (2016). Identity, homophily and in-group bias. *European Economic Review*, Volume 90, November 2016, Pages 40–55. DOI: 10.1016/j.eurocorev.2016.02.015.
- Muñoz, D. (2024). Prototipo funcional para caracterizar la Homofilia en redes sociales. Informe de Proyecto de Titulación, Ingeniería en Informática, UTEM.
- Dandekar, P., Goel, A., & Lee, D. T. (2013). Biased assimilation, homophily, and the dynamics of polarization. *Proceedings of the National Academy of Sciences*, 110(15), 5791–5796. DOI: 10.1073/pnas.1217220110.
- Dang, C. N., García, M. N. M., & De la Prieta, F. (2020). Sentiment Analysis Based on Deep Learning: A Comparative Study. *Electronics*, 9(3), 483. DOI: 10.3390/electronics9030483.
- El-Moussaoui, Mohamen; Mohamed, Hanine, and Kartit, Ali; Garcia Villar, Monica; Garay, Helena and De La Torre Díez, Isabel (2025). A systematic review of deep learning methods for community detection in social networks. *Front. Artif. Intell.*, 21 August 2025, Sec. Machine Learning and Artificial Intelligence, Volume 8–2025. DOI: 10.3389/frai.2025.1572645.
- Hartmann, D., Wang, S.M., Pohlmann, L. et al. A systematic review of echo chamber research: comparative analysis of conceptualizations, operationalizations, and varying outcomes. *J Comput Soc Sc* 8, 52 (2025). <https://doi.org/10.1007/s42001-025-00381-z>.
- Garimella, K., Morales, G. D. F., Gionis, A., & Mathioudakis, M. (2018). Quantifying controversy in social media. *ACM Transactions on Social Computing*, 1(1), 1–27. <https://doi.org/10.1145/3239549>.
- Gers, F. A., Schmidhuber, J., & Cummins, F. (2002). Learning to forget: Continual prediction with LSTM. *Neural Computation*, 12(10), 2451–2471. DOI: 10.1162/089976600300015015.
- Glen, Stephanie (2019). Comparing Classifiers: Decision Trees, K-NN & Naive Bayes, June 19, 2019. TechTarget Blog. <https://www.datasciencecentral.com/comparing-classifiers-decision-trees-knn-naive-bayes/>.

- Golub, B., & Jackson, M. O. (2012). How homophily affects the speed of learning and best-response dynamics. *The Quarterly Journal of Economics*, 127(3), 1287–1338. <https://web.stanford.edu/~jacksonm/homophily.pdf>.
- Gu, Huirong and Jiao, Jiyuan (2021). Comparison of classifiers for different data in application of classification. August 2021, *Journal of Physics Conference Series* 1994(1):012015. DOI:10.1088/1742-6596/1994/1/012015.
- Greff, K., Srivastava, R. K., Koutník, J., Steunebrink, B. R., & Schmidhuber, J. (2017). LSTM: A search space odyssey. *IEEE Transactions on Neural Networks and Learning Systems*, 28(10), 2222–2232.
- Haenlein, M., & Kaplan, A. M. (2021). A brief history of artificial intelligence: On the past, present, and future of artificial intelligence. *California Management Review*, 61(4), 5–14.
- Harzing, A.W. (2007) Publish or Perish, available from <https://harzing.com/resources/publish-or-perish>.
- Hung, M., Lauren, E., Hon, E. S., Birmingham, W. C., Xu, J., Su, S., Hon, S. D., Park, J., Dang, P., & Lipsky, M. S. (2020). Social Network Analysis of COVID-19 Sentiments: Application of Artificial Intelligence. *JMIR*, Vol. 22, No 8 (2020): August. DOI: 10.2196/22590.
- Ito, M., Koutra, D., & Wiens, J. (2025). Understanding GNNs and Homophily in Dynamic Node Classification. *arXiv / AISTATS 2025*, Proceedings of the 28th International Conference on Artificial Intelligence and Statistics (AISTATS), Thailand. PMLR: Volume 258. DOI: 10.48550/arXiv.2504.20421.
- Jackson, Mathew O. (2025). Inequality's economic and social roots: the role of social networks and homophily. January 2021, *SSRN Electronic Journal*, DOI:10.48550/arXiv.2506.13016.
- Jackson, Mathew O. (2008). Average distance, diameter, and clustering in social networks with homophily. *Lecture Notes in Computer Science* 5385, Conference: Proceedings of the 4th International Workshop on Internet and Network Economics, DOI:10.1007/978-3-540-92185-1_3.
- Jackson, Mathew O. and López-Pintado (2013). Diffusion and contagion in networks with heterogeneous agents and homophily. *Arxiv.org*, [physics.soc-ph] 31 Oct 2011. <https://arxiv.org/pdf/1111.0073>.
- Jackson, Mathew O. (2021). Inequality's economic and social roots: the role of social networks and homophily. January 2021 *SSRN, Electronic Journal*, DOI: arXiv:2506.13016v1 [econ.GN] 16 Jun 2025.
- Ibarra, H. (1992). Homophily and differential returns: Sex differences in network structure and access in an advertising firm. *Administrative Science Quarterly*, 37(3), 422–447. DOI: 10.2307/2393451.
- Ito, M., Koutra, D. & Wiens, J. (2025). Understanding GNNs and Homophily in Dynamic Node Classification. Proceedings of the 28th International Conference on Artificial Intelligence and Statistics, in *Proceedings of Machine Learning Research* 258:2764-2772. <https://proceedings.mlr.press/v258/ito25b.html>.
- Jiang, J., Luceri, L., Walther, J., & Ferrara, E. (2023). Social Approval and Network Homophily as Motivators of Online Toxicity. DOI: 10.48550/arXiv.2310.07779.

- Khanam, K.Z., Srivastava, G. & Mago, V. (2023). The homophily principle in social network analysis: A survey. *Multimed Tools Appl* 82, 8811–8854 (2023). DOI: 10.1007/s11042-021-11857-1
- Karimi, F., Génois, M., Wagner, C., Singer, P., & Strohmaier, M. (2018). Homophily influences ranking of minorities in social networks. *Scientific Reports*, 8, 11077.
- Karimi, F., & Oliveira, M. (2023). On the inadequacy of nominal assortativity for assessing homophily in networks. *Nature Communications*, 14, 1430.
- Legorreta, D. (2015). Comparación entre Máquina de Soporte Vectorial, Naive Bayes, Árboles de Decisión y Métodos Lineales. 9 julio, 2015. <https://dlegorreta.wordpress.com/2015/07/09/comparacion-entre-svm-naive-bayes-arboles-de-decision-y-metodos-lineales/>.
- Liu, B. (2012). *Sentiment Analysis and Opinion Mining*. Morgan & Claypool Publishers. Part of the book series: Synthesis Lectures on Human Language Technologies. DOI: <https://link.springer.com/book/10.1007/978-3-031-02145-9>.
- Li, Y. (2023). Replication Data for Examining Homophily, Language Coordination, and Analytical Thinking in Online Conversations about Vaccines on Reddit: A Study Using Deep Neural Network Language Models and Computer-assisted Conversational Analyses [Conjunto de datos]. En Harvard Dataverse. <https://doi.org/10.7910/dvn/1rns7y>
- Ma, L., Krishnan, R., & Montgomery, A. L. (2015). Latent Homophily or Social Influence? An Empirical Analysis of Purchase Within a Social Network. *Management Science*, 61(2), 454-473. DOI: 10.1287/mnsc.2014.1928.
- Magna, Oscar (2021). Marco de Trabajo para Proyectos de Titulación en modalidad de Investigación v2. Depto. de Informática y Computación, Facultad de Ingeniería, Universidad Tecnológica Metropolitana (UTEM, Chile).
- Mahesh, B. (2020). Machine learning algorithms – a review. *International Journal of Science and Research (IJSR)*, 9(1), 381–386. DOI: 10.21275/ART20203995.
- McPherson, M., Smith-Lovin, L., & Cook, J. M. (2001). Birds of a feather: Homophily in social networks. *Annual Review of Sociology*, 27, 415–444. <https://doi.org/10.1146/annurev.soc.27.1.415>.
- Mostakim, S. A., Ehsan, F., Hasan, S. M., Islam, S., & Shatabda, S. (2018). Bangla content categorization using text based supervised learning methods. <http://103.109.52.4:8080/bitstream/handle/52243/470/inalthesisreport.pdf?sequence=1&isAllowed=y>.
- Nandwani, P., & Verma, R. (2021). A review on sentiment analysis and emotion detection from text. *Social Network Analysis and Mining*, 11(1). DOI: 10.1007/s13278-021-00776-6.
- Naseem, U., Ren, J., Anwar, S., et al. (2025). POLAR: A Benchmark for Multilingual, Multicultural, and Multi-Event Online Polarization. DOI: 10.48550/arXiv.2505.20624.
- Otieno, Polycarp (2024). The Impact of Social Media on Political Polarization. *Journal of Communication* 4(1):56-68, DOI:10.47941/jcomm.1686. <https://carijournals.org/journals/index.php/JCOMM/article/view/1686>.

Palanivinayagam, A.; El-Bayeh, C.Z.; Damaševičius, R. (2023). Twenty Years of Machine Learning-Based Text Classification: A Systematic Review. *Algorithms* 2023, 16, 236. DOI: 10.3390/a16050236. <https://www.mdpi.com/1999-4893/16/5/236>.

Rainio, O., Teuvo, J. & Klén, R. Evaluation metrics and statistical tests for machine learning. *Sci Rep* 14, 6086 (2024). <https://doi.org/10.1038/s41598-024-56706-x>.

Ruef, M., Howard, E. A., Carter, N. M. (2003). The Structure of Founding Teams: Homophily, Strong Ties, and Isolation among U.S. Entrepreneurs. *American Sociological Review*, Vol. 68, No. 2 (Apr., 2003), pp. 195-222. DOI: 10.2307/1519766.

Yuan, Lanqin; Schneider, Philipp J. and Rizoioiu, Marian-Andrei (2025). Behavioral Homophily in Social Media via Inverse Reinforcement Learning. arXiv:2502.02943, DOI: 10.1145/3696410.3714618.

Zheng, X., Wang, Y., Liu, Y., Li, M., Zhang, M., Jin, D., & Yu, P.S. (2022). Graph neural networks for graphs with heterophily: A survey. *IEEE Transactions on Knowledge and Data Engineering*, 35(1), 1–24. DOI: 10.48550/arXiv.2202.07082.

Zhu, J., Xu, H., Yu, Y., & Gong, Z. (2020). Beyond homophily in graph neural networks: Current limitations and effective designs. *Advances in Neural Information Processing Systems*, 33, 7793–7804. DOI: 10.48550/arXiv.2006.11468.



Esta obra está bajo una licencia Creative Commons
Atribución-Compartir Igual 4.0 Internacional.
Atribución: debe otorgar el crédito apropiado
a la Universidad Tecnológica Metropolitana
como editora y citar al autor original. Compartir
igual: si reorganiza, transforma o desarrolla el
material, debe distribuir bajo la misma licencia
que el original.



UNIVERSIDAD
TECNOLÓGICA
METROPOLITANA
del Estado de Chile

TRILOGÍA

CIENCIA · TECNOLOGÍA · SOCIEDAD



EDICIONES UNIVERSIDAD
TECNOLÓGICA METROPOLITANA