

**Academic Year 2025-26**

**IV B. Tech I Semester Project work**

**LITERATURE REVIEW PLAN**

**Batch Code:** 22CSEA06 **Supervisor:** P. Krishna Moorthy

**Project Title:** A Heuristic-Based Scoring Algorithm for Suspicious Follower Detection in Simulated Instagram Networks.

**1. REVIEW OBJECTIVES**

* Understand the current landscape of fake follower/suspicious user detection in social media in Instagram.
* Analyse rule-based, heuristic, and graph-based techniques used in bot or fake follower detection.
* Compare datasets, methodologies, and evaluation strategies from various studies.
* Identify gaps or overlooked aspects in existing research (e.g., privacy issues, use of real datasets).
* Build a strong foundation for our heuristic scoring-based model with simulated data.

**2. SOURCES FOR ARTICLES**

* Google Scholar
* IEEE Xplore
* Springer Link
* ResearchGate
* ACM Digital Library
* arXiv (Open Access Preprints)
* Semantic Scholar
* CORE (Open Access Repository Aggregator)
* Directory of Open Access Journals (DOAJ)

**3. SEARCH KEYWORD / PROMPT**

* Fake follower identification social media
* Rule-based fake follower detection Instagram
* Bot detection in Instagram social media
* Fake account detection heuristic algorithm
* Instagram influencer analysis suspicious followers
* Heuristic scoring social media user detection
* Rule-based spam profile detection in networks
* Suspicious follower detection in Instagram
* Heuristic-based bot detection
* Rule-based user classification in social networks
* Instagram bot detection using graph analysis
* Influencer fraud detection
* Fake profile detection using heuristic scoring
* Instagram fake engagement detection
* Social media graph pattern analysis

**4. INCLUSION CRITERIA**

* Relevance to Suspicious Follower detection in Instagram
* Articles from 2015 onwards
* Articles from Transactions, Journal or Peer Reviewed Journals
* Scopus Indexed Conference Papers
* Articles from Q1, Q2 and Q3 Scopus Indexed
* Articles Published in English

**5. EXCLUS ION CRITERIA**

* Social media other than Instagram
* Articles prior to 2015
* Reports, policy Literature, Working Papers, Newsletters, Government Documents, Speeches
* Articles from Books, Blogs other web sources
* Articles with a Journal Impact Factor > 1.0
* Not accessible
* Predatory or Non-Indexed Journals

**6. QUALITY ASSESSMENT CHECKLIST**

|  |  |  |
| --- | --- | --- |
| Sl. No | Parameter | Criteria |
| 1 | Relevance | Does the paper directly address fake/suspicious follower or bot/user detection on Instagram or similar social media platforms? |
| 2 | Method | Does the paper propose or utilize heuristic-based scoring, rules, or graph theory (e.g., network analysis) to detect suspicious users? |
| 3 | Dataset | Does the paper use simulated datasets or public real-world datasets (e.g., Instagram data, Botometer, Twitter datasets)? Is the dataset available for reuse? |
| 4 | Metrics | Does the paper provide evaluation using metrics like accuracy, precision, recall, F1-score, AUC, etc.? |
| 5 | Evaluation | Does the study include sound experimental validation (e.g., cross-validation, baseline comparison, ablation studies)? |
| 6 | Reproduction | Are the methods clearly described with enough detail for replication? Are code or datasets shared? |
| 7 | Publication | Is the paper published in a peer-reviewed journal, conference, or trusted repository (e.g., IEEE, ACM, Elsevier, Springer, arXiv)? |
| 8 | Privacy | Does the paper discuss user privacy, data ethics, or anonymization in the detection process? |
| 9 | Real Time | Does the approach emphasize low computational complexity, scalability, or suitability for real-time application? |
| 10 | Citation | How well-cited is the paper? |

**7. DATA EXTRACTION**

|  |  |
| --- | --- |
| **Field** | **Description** |
| Title | Full title of the research paper |
| Author(s) | Lead author and co-authors |
| Year | Year of publication |
| Source | Database/source of the paper (Google Scholar, IEEE, Springer, etc.) |
| Challenges Addressed | What specific technical or societal problem is the paper trying to solve |
| Proposed Method | The approach or technique used (e.g., heuristic, rule-based, graph analysis) |
| Dataset Used | Dataset used in the study (Instagram, simulated, Twitter, or custom dataset) |
| Evaluation Metrics | Metrics used to validate the method (Accuracy, Precision, Recall, F1, AUC, etc.) |
| Unique Findings | Notable or standout observations from their research |
| Advantages | Benefits and strengths of their approach |
| Limitations & Future Scope | Any shortcomings identified and future work suggested by the authors |
| Privacy Consideration | Whether the paper respects data privacy or addresses consent (Yes/No) |
| Notable Gaps | Gaps identified by the team (e.g., lacks privacy, no heuristic scoring, etc.) |

**8. OUTCOME OF LITERATURE SURVEY**

* Literature Gap Identification
* Highlight review findings
* Review Article

Signature of the Supervisor