AOI-Inspired Descriptive Rule Mining: Unlocking Insights from Non-Hierarchical Data

Akshaya Srikrishna[1], Anagha Srikrishna[1], Krishnendu M R[1], and S. Chitrakala[1]

Dept. of Computer Science and Engineering,
College of Engineering Guindy, Anna University, Chennai, India
2022103065@student.annauniv.edu, 2022103066@student.annauniv.edu,
2022103081@student.annauniv.edu, chitras@annauniv.edu

Abstract. Attribute-Oriented Induction (AOI) is a well-established method for pattern discovery in hierarchical data, yet many real-world datasets lack explicit hierarchical structures. This work proposes a systematic methodology (ADRM-NHD - AOI-Inspired Descriptive Rule Mining: Unlocking Insights from Non-Hierarchical Data) that bridges this gap by adapting AOI for effective use on non-hierarchical data. Our approach systematically discretizes continuous and fine-grained variables into lightweight, hierarchy-like proxies to make previously incompatible datasets amenable to generalization-based rule induction. Rule evaluation then integrates two complementary measures: T-weight, to capture a pattern's representativeness within a specific group, and D-weight, to highlight its distinctiveness across different groups. By prioritizing rules with both high T- and D-weight, the methodology produces a concise and interpretable rule set. Applied to the Amazon 2023 Fashion Reviews dataset, our method extracts actionable insights into credibility cues, engagement patterns, and temporal shifts, such as pre- versus post-COVID behaviors. This work demonstrates how to successfully extend AOI principles to flat-feature domains, offering a practical and interpretable approach for knowledge discovery in common behavioral datasets.

Keywords: Non-Hierarchical Data· Amazon Fashion Reviews · Descriptive rules · AOI-inspired rule mining · interpretable rules · T-weight · D-weight · reviewer behavior · structured attributes · targeted marketing · behavioral analytics · temporal trends

1 Introduction

Most real-world data sources, particularly behavioral datasets, are inherently non-hierarchical; they lack the taxonomic structures that classical Attribute-Oriented Induction (AOI) depends on to generalize and abstract patterns. Without such hierarchies, it becomes difficult to roll up fine-grained attributes into higher-level concepts while still retaining interpretability. For instance, on e-commerce platforms like Amazon, review attributes such as length, rating, and helpfulness exist only as flat features, making traditional AOI inapplicable for this common type of data.

E-commerce platforms exemplify this challenge, generating vast volumes of customer reviews that capture consumer opinions across multiple dimensions, including product ratings, review length, helpfulness votes, verification status, and temporal activity. These reviews provide valuable signals of credibility, engagement, and behavioral change. However, their scale and lack of inherent structure make analysis difficult. Unlike structured product taxonomies, these review attributes do not come with the explicit hierarchies that pattern-discovery methods like AOI typically rely on. This absence complicates the task of abstracting fine-grained data into interpretable patterns without introducing significant redundancy or noise.

To address this gap, this study proposes a systematic methodology (ADRM-NHD) tailored to non-hierarchical review attributes. The approach creates "hierarchy-like abstractions" that act as proxies for higher-level concepts by systematically discretizing raw attributes into meaningful, coarse categories. From these generalized forms, rules are induced and then evaluated using two established and complementary measures: T-weight, to capture how representative a rule is within a group, and D-weight, to quantify its distinctiveness across groups. Retaining only rules with high T- and D-weight ensures a concise, human-readable rule set.

Applying this methodology to the Amazon 2023 Fashion Reviews dataset uncovers actionable patterns across credibility (verified vs. unverified reviewers), engagement (frequent vs. infrequent contributors), and temporal variation (seasonal and pre-/post-COVID shifts). A key advantage of this approach is its focus on interpretability; unlike black-box classifiers, it produces structured rules that summarize reviewer behaviors in a compact, human-readable form. By concentrating on the dataset's structured features, this approach provides a more direct path to generating these insights, bypassing the computational overhead and ambiguity of complex Natural Language Processing (NLP) on unstructured review text.
The contributions of this work are threefold:

1. Proposing a systematic methodology to adapt AOI principles for non-hierarchical data through attribute discretization;
2. Extending the use of T- and D-weight measures to evaluate rule quality in behavioral mining; and
3. Demonstrating that the resulting rules capture actionable credibility cues, engagement dynamics, and temporal shifts, offering interpretable insights for strategic decision-making.

2 Related work

With the exponential growth of user-generated content on e-commerce platforms such as Amazon, mining reviews has become a central challenge in understanding consumer behavior, trust, and platform credibility. Existing research has largely focused on sentiment classification and credibility analysis. For example, Santhosh et al. [6] used Naïve Bayes with two-phase filtering for classifying Amazon camera reviews, while Aishwarya et al. [1] incorporated verification status and reviewer history for credibility-based filtering. Although such approaches improve classification accuracy and trustworthiness, they remain limited to isolated review-level predictions rather than discovering generalized behavioral patterns across reviewer groups.

Amazon review datasets are inherently semi-structured: they contain numerical, categorical, and textual attributes such as rating, helpfulness votes, review length, and verification status. However, unlike classical structured datasets, they lack natural hierarchies or taxonomies that traditional data mining methods like Attribute-Oriented Induction (AOI) require. This makes directly applying AOI challenging, since its strength lies in abstracting detailed data into higher-level, interpretable rules using concept hierarchies.

AOI was originally introduced by Han et al. [4] as a generalization-based method to derive characteristic and discriminant rules from relational data. Subsequent research has extended AOI in various directions. Noise-free AOI methods were proposed to filter hidden noise and quantify accuracy loss during generalization [10]. To address overgeneralization, entropy-based feature selection and stop criteria were introduced [2], while correlation-based pruning improved rule validity by removing irrelevant attributes [5]. High-level emerging pattern discovery frameworks such as AOI-HEP exploited attribute taxonomies to yield more interpretable and discriminative patterns [9]. Extensions to unstructured text further demonstrated AOI's adaptability, by constructing concept trees from comment data and optimizing induction cost [8]. Simplified pedagogical treatments also exist, providing accessible introductions to AOI's strategic steps [7]. A more general exploration of AOI's capabilities highlights its ability to extract various knowledge rules, including characteristic, discriminant, and multiple-level association rules [3].

Despite these advances, most AOI research operates on the assumption that explicit hierarchies or taxonomies are available, a condition rarely met in semi-structured datasets such as consumer reviews. Our work directly addresses this gap by proposing a systematic methodology for the Amazon 2023 Fashion Reviews dataset. Instead of relying on predefined concept trees, our approach systematically discretizes continuous and fine-grained attributes into "hierarchy-like proxies," enabling the extraction of interpretable characteristic and discriminant rules from data previously considered incompatible with AOI. By applying established measures such as T-weight and D-weight, we derive compact rule sets that highlight representative and contrastive reviewer behaviors, offering actionable insights into credibility, engagement, and temporal dynamics.

3 Proposed Methodology

This work proposes ADRM-NHD, a systematic methodology inspired by Attribute-Oriented Induction (AOI) and specifically adapted to extract interpretable behavioral patterns from non-hierarchical data. Many contemporary datasets, rich in structured features, lack the clear, predefined concept hierarchies (e.g., *City*

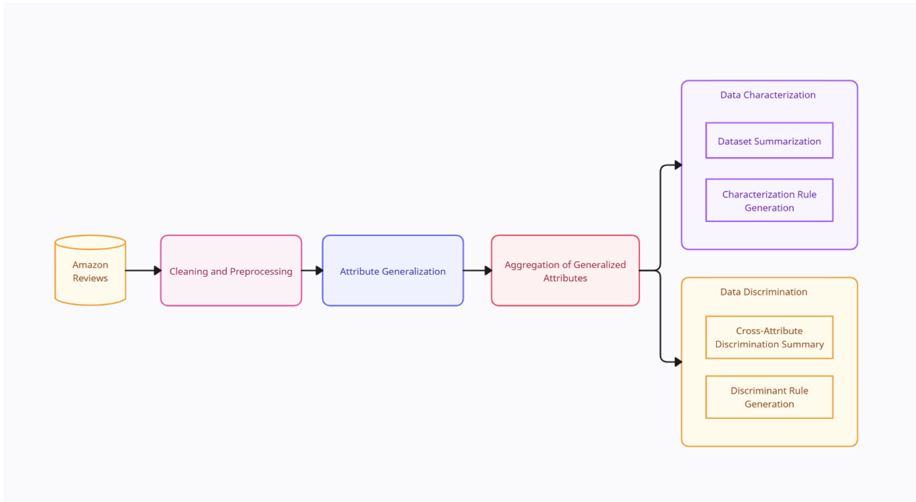

Fig. 1. AOI-Inspired framework for extracting reviewer behavior rules

$\rightarrow$ *State* $\rightarrow$ *Country*) that classical AOI relies upon. Our methodology addresses this gap by creating generalizations on-the-fly, enabling the discovery of insightful rules from this common and challenging data structure.

To demonstrate its efficacy, we apply this methodology to the the Amazon 2023 Fashion Reviews (approximately 2.5M reviews). A key strategic decision is to bypass the ambiguity and computational overhead of complex Natural Language Processing (NLP) on unstructured review text. Instead, we focus on the dataset's structured features (rating, verification status, helpfulness votes, etc.), which provide a less noisy and more direct path to generating compact, human-readable rules about reviewer credibility, participation patterns, and temporal behavior. The overall workflow is illustrated in Figure 1.

Our methodology unfolds in a two-stage pipeline: attribute generalization followed by a unified rule mining process.

The foundational step of our framework is attribute generalization. This stage adapts the core principle of AOI to data lacking explicit hierarchies by discretizing continuous or fine-grained features into meaningful, coarse categories. This process effectively constructs a proxy for a concept hierarchy, reducing data sparsity and noise while preserving essential semantic information. For the Amazon dataset, this transformation is performed as follows:

- Rating: The 1-5 star rating is mapped to coarse sentiment categories: {Negative (1-2 stars), Neutral (3 stars), Positive (4-5 stars)}. This mapping is based on common conventions in sentiment analysis and reflects the typical distribution of review scores.
- Review Length: Measured in tokenized word counts, this continuous feature is grouped into percentile-based categories: {Very Short, Short, Medium, Long}.

Using percentiles ensures the bins are balanced and robust against outliers, such as extremely long or short reviews.

- Helpfulness: The raw count of helpfulness votes, which often follows a power-law distribution, is discretized into levels: {Low, Medium, High} using quantiles. This approach guarantees that each category contains a comparable number of reviews, making subsequent pattern mining more stable.
- Verification: The binary VerificationStatus is already categorical and is retained in its original form: {Verified, Unverified}, serving as a crucial feature for trust and credibility analysis.

This generalization step transforms the raw, noisy data into a clean, conceptualized format, making it amenable to efficient and effective rule induction.

Algorithm 1: Mining Characteristic and Discriminative Rules from Generalized Data

Input: Raw dataset D, Set of target classes $\{C_1, C_2, \ldots, C_n\}$
Output: Generalized dataset D_g, Characteristic rules R_{char}, Discriminative rules R_{disc}

Step 1: Attribute Generalization
Generalize raw dataset attributes: **for** *each review* $r \in D$ **do**
    Map ReviewLength → length categories (Very Short, Short, Medium, Long);
    Map StarRating → sentiment categories (Negative, Neutral, Positive);
    Bin HelpfulVotes → levels (Low, Medium, High);
    Retain VerificationStatus as is;
$D_g \leftarrow$ Generalized dataset;

Step 2: Characterization Rule Mining
Mine rules that are highly representative for each class: **for** *each target class* $C_j \in \{C_1, \ldots, C_n\}$ **do**
    Generate candidate rules q_a by applying Apriori on D_g;
    for *each candidate rule* q_a **do**
        Compute T-weight(q_a, C_j) (Eq. 1);
    $R_{char}[C_j] \leftarrow$ Select top-k rules from candidate set, ranked by T-weight;

Step 3: Discriminant Rule Mining
Mine rules that best discriminate between classes: Generate candidate rules q_a by applying Apriori on D_g;
for *each candidate rule* q_a **do**
    for *each class* $C_j \in \{C_1, \ldots, C_n\}$ **do**
        Compute T-weight(q_a, C_j) using Eq. 1;
        Compute D-weight(q_a, C_j) using Eq. 2;
$R_{disc} \leftarrow$ Rank and retain top-k rules per class by D-weight;
return D_g, R_{char}, R_{disc};

With the generalized dataset, we proceed to a unified rule mining stage that simultaneously identifies both **characteristic** and **discriminative** behavioral patterns. These two rule types provide a holistic understanding of reviewer behavior.

– **Characteristic Rules** describe the dominant traits *within* a specific class of reviewers. They build a representative profile by answering the question, "What behaviors are most typical for this group?" For example, a characteristic rule might be that 'Positive' reviews are frequently of 'Medium' length.
– **Discriminative Rules** identify the unique patterns that *distinguish* one class from all others. They pinpoint behavioral signatures by answering the question, "What makes this group different?" For instance, a rule stating that 'Unverified' reviewers are uniquely associated with 'Very Short', 'Negative' reviews would be highly discriminative.

To accomplish this, we first employ the Apriori algorithm as an efficient engine to generate a comprehensive set of candidate rules from the generalized data. Apriori excels at finding frequently co-occurring attribute combinations. However, frequency alone does not guarantee insight. Therefore, the generated candidates are subsequently evaluated for both their representativeness within each class (for characteristic rules) and their exclusivity to each class (for discriminative rules). This two-tiered process of generation and evaluation is detailed in Algorithm 1.

By systematically identifying both what is common and what is unique, this methodology produces a rich, dual perspective on reviewer behavior, delivering insights that are far more nuanced than what could be achieved by analyzing frequency or simple statistics alone. The resulting rules are presented in a simple, logical format, making them immediately accessible to analysts and decision-makers.

4 Experimental Results and Analysis

Applying the proposed ADRM-NHD methodolgy to the Amazon 2023 Fashion Reviews dataset yields a set of compact and interpretable behavioral rules. The analysis is presented in two parts: characteristic rules that describe typical behaviors within a group, and discriminant rules that highlight unique patterns distinguishing between groups.

The initial generalization of attributes, which acts as a proxy for hierarchical abstraction, reveals key skews in the data's composition. Ratings were mapped into sentiment categories, showing that positive sentiment dominated with 1,778,595 reviews, while neutral and negative reviews accounted for 245,471 and 476,873, respectively. Review length was discretized into Very Short, Short, and Long, with Short and Very Short reviews being predominant (966,686 and 706,260), compared to only 304,635 Long reviews. Finally, helpfulness votes were aggregated, showing that Helpfulness was rare, with only 205,924 reviews marked Helpful versus over 2 million marked Not Helpful. These foundational distributions provide the context for the subsequent rule induction.

To identify the most typical patterns within a specific group, we mine for characteristic rules, which summarize the most typical patterns within a specific group. To formally measure this representativeness, the framework calculates the T-weight for each potential rule. The T-weight (Eq. 1) of a rule q_a for a given class C_j is defined as the proportion of instances in that class that match the rule, ensuring that the selected rules are truly characteristic of the group they describe.

$$T\text{-weight}(q_a, C_j) = \frac{\text{count}(q_a \in C_j)}{\sum_i \text{count}(q_i \in C_j)} \tag{1}$$

This approach reveals dominant trends, such as verified reviews being strongly associated with short, positive feedback, while unverified reviews more often produced longer narratives. Table 1 presents the top characteristic rules ranked by T-weight, formalizing the insight that verified buyers typically provide concise positive feedback, whereas unverified reviewers tend to submit longer responses.

Table 1. Top Characteristic Rules by T-Weight

Rule	**T-Weight**
IF Sentiment = Positive AND Verification = Verified THEN Review Length = Short	38.99
IF Sentiment = Positive AND Verification = Unverified THEN Review Length = Long	33.63

While characterization captures typicality, discriminant rules highlight the unique behavioral traits that distinguish one group from another. To identify strongly discriminant patterns, we use a two-part evaluation process for each candidate rule. First, we measure the strength of the association using Lift (Eq. 2). Lift quantifies how much more likely the conclusion of a rule is when the condition is met, compared to its general likelihood. A Lift value greater than 1 indicates a positive correlation, ensuring the pattern is more significant than random chance.

$$\text{Lift}(A \to B) = \frac{\text{Support}(A \cup B)}{\text{Support}(A) \times \text{Support}(B)} = \frac{P(B|A)}{P(B)} \tag{2}$$

Second, to specifically measure a rule's exclusivity to a certain group, we compute the D-weight (Eq. 3). This metric measures the degree to which a rule q_a is concentrated in a particular class C_j compared to all other classes, effectively isolating behavioral signatures.

$$D\text{-weight}(q_a, C_j) = \frac{\text{count}(q_a \in C_j)}{\sum_k \text{count}(q_a \in C_k)} \tag{3}$$

A rule is considered strongly discriminant only if it has both a Lift > 1 and a high D-weight. The following analysis highlights the top discriminant rules found across various dimensions.

Applying these metrics across credibility, engagement, and temporal dimensions provides powerful distinguishing patterns. When comparing frequent versus infrequent reviewers, the framework finds that frequent reviewers are distinguished by a tendency to provide short, unverified positive reviews, reflecting habitual but potentially less credible activity. Infrequent reviewers, in contrast, are uniquely characterized by leaving longer, verified reviews that are more often marked as helpful, as shown in Table 2. Crucially, the D-weight of 100.00% indicates the pattern is entirely exclusive to this group in our dataset.

Table 2. Top Rules and Metrics for Reviewer Types

Rule	**Metrics**
IF Reviewer_Type = Infrequent THEN Helpfulness = More Helpful	Lift: 1.172 T Weight: 43.063 D Weight: 100.000
IF Reviewer_Type = Frequent THEN Verified = Not Verified AND Sentiment = Positive	Lift: 1.184 T Weight: 9.520 D Weight: 100.000

Temporal analysis further uncovers a significant shift in consumer behavior. Pre-COVID reviews were largely positive, verified, and high-rated, while post-COVID reviews are discriminately shorter, more negative, and less helpful (Table 3).

Table 3. Top Rules for Pre- vs. Post-COVID

Rule	**Metrics**
IF lockdown_period = Pre-COVID THEN season = Winter AND Sentiment = Positive AND Verified = Verified AND Rating = High	Lift: 1.073 T Weight: 23.89% D Weight: 100.00%
IF lockdown_period = Post-COVID THEN Sentiment = Negative AND Review_Length = Short	Lift: 1.196 T Weight: 6.38% D Weight: 100.00%

This trend is also reflected in seasonal contrasts, where summer reviews are uniquely associated with the post-COVID period and verified purchases, whereas winter reviews align with pre-COVID conditions, a positive tone, and greater helpfulness, as summarized in Table 4. Again, the 100.00% D-weight for these top rules suggests a near-complete shift in review patterns between the two periods.

Table 4. Top Rules for Summer vs Winter Season

Rule	**Metrics**
IF season = Summer THEN Verified = Verified AND lockdown_period = Post-COVID	Lift: 1.685 T Weight: 15.66% D Weight: 100.00%
IF season = Winter THEN Sentiment = Positive AND lockdown_period = Pre-COVID AND Helpfulness = More AND Review_Length = Short	Lift: 1.162 T Weight: 15.58% D Weight: 100.00%

5 Conclusion and Future Scope

This study proposes ADRM-NHD, a systematic methodology for mining structured patterns from non-hierarchical data, demonstrated using the Amazon 2023 Fashion Reviews dataset. By discretizing semi-structured attributes such as review length, rating, helpfulness, and verification, the framework generated compact, human-readable characteristic (T-weight) and discriminant (D-weight) rules. The extracted rules revealed meaningful behavioral patterns: verified reviewers tend to give concise, positive feedback; infrequent reviewers often provide more credible and helpful insights; and temporal shifts, such as pre- vs. post-COVID trends, shape review tone and length. These findings demonstrate how AOI principles can be adapted to domains where natural hierarchies are absent, offering interpretable and actionable insights that extend beyond surface-level sentiment analysis.

Future work can extend this framework by incorporating aspect-level sentiment, multimodal cues (e.g., images or metadata), and domain-specific taxonomies to enrich interpretability. Further exploration of advanced AOI variants and comparisons with alternative rule-mining approaches will help strengthen the generalizability, scalability, and practical utility of AOI-inspired review analysis.

References

1. Aishwarya, N., S, B.L., N, K.: Summarization and prioritization of amazon reviews based on multi-level credibility attributes. In: 2021 International Conference on Recent Trends on Electronics, Information, Communication Technology (RTEICT). pp. 5–9 (2021). https://doi.org/10.1109/RTEICT52294.2021.9573744
2. Almamory, S., Talib, S., Hammid, K.: Enhancing attribute oriented induction of data mining p. (04 2019)
3. Han, J., Fu, Y.: Exploration of the power of attribute-oriented induction in data mining (02 1999)
4. Jiawei Han, Micheline Kamber, J.P.: Data Mining: Concepts and Techniques. Morgan Kauffmann (2012)
5. Qu, Y., Li, X., Wang, H.: Improvement of attribute-oriented induction method based on attribute correlation with target attribute. In: 2014 International Conference on Machine Learning and Cybernetics. vol. 2, pp. 670–674 (2014). https://doi.org/10.1109/ICMLC.2014.7009689

6. Santhosh, R., Vignesh, S.V., Rithish, E., Mahendhiran, P.D.: Sentimental analysis on amazon camera reviews using naive bayes algorithm. In: 2023 International Conference on Sustainable Computing and Smart Systems (ICSCSS). pp. 1–8 (2023). https://doi.org/10.1109/ICSCSS57650.2023.10169302
7. Spits Warnars, H.L.H., Wijaya, M.I., Tjung, H., Xaverius, D., Hauten, D., Sasmoko, S.: Easy understanding of attribute oriented induction (aoi) characteristic rule algorithm. International Journal of Applied Engineering Research **11**, 5369–5375 (05 2016)
8. Tu, Y.N., Kuo, C.Y.: Unstructured text data rule generalization in attribute-oriented induction. In: 2022 IEEE 5th International Conference on Knowledge Innovation and Invention (ICKII). pp. 210–214 (2022). https://doi.org/10.1109/ICKII55100.2022.9983603
9. Warnars, S.: Mining frequent pattern with attribute oriented induction high level emerging pattern (aoi-hep). In: 2014 2nd International Conference on Information and Communication Technology (ICoICT). pp. 149–154 (2014). https://doi.org/10.1109/ICoICT.2014.6914056
10. Wu, C., Chen, Y.C., Yu, M.: Mining specific and representative information by the attribute-oriented induction method. Expert Systems **38** (2020). https://doi.org/10.1111/exsy.12643