Department: Applications

From Prediction to Insight: Visual Analytics for Understanding Compound Potency Models

Bahavathy Kathirgamanathan†, *University of Cologne, Cologne, Germany*

Tiago Janela†, *University of Bonn, Bonn, Germany*

Elena Xerxa, *University of Bonn, Bonn, Germany*

Gennady Andrienko, *Fraunhofer IAIS, Sankt Augustin, Germany and City St George's University of London, UK*

Jürgen Bajorath, *University of Bonn, Bonn, Germany*

Natalia Andrienko, *Fraunhofer IAIS, Sankt Augustin, Germany and City St George's University of London, UK*

Abstract—Machine learning (ML) is widely used in medicinal chemistry, but accurate predictions alone are insufficient. Researchers need insight into which molecular features determine compound properties. We present an application-oriented case study that analyzes a trained model for compound potency as a source of domain knowledge. The model is converted into decision rules, and topic-guided visual analytics is used to identify co-occurring feature conditions associated with high predicted potency. These patterns are then mapped back to molecular substructures, yielding chemically interpretable motifs and testable hypotheses about structure–activity relationships. The study demonstrates how combining rule-based representations, topic modeling, and visual exploration can turn potency predictions into mechanistic insight, and outlines a reusable workflow for interpreting ML models of molecular properties.

Machine learning (ML) models are widely used to support scientific discovery, but accuracy alone is insufficient: scientists also want to understand how models make decisions and how this relates to domain knowledge. Explainable AI (XAI) methods aim to increase transparency, yet often emphasize single-feature effects and local explanations, making it hard to uncover multi-feature patterns that matter in practice.

In medicinal chemistry, ML models are often used to estimate compound potency, i.e., how strongly a molecule binds to a biological target given its structural fragments. This task is central for identifying promising drug candidates. We report an application-oriented case study that extracts domain knowledge from a trained potency model for one compound class by uncovering co-occurring structural patterns linked to high activity. The work involved an interdisciplinary team of drug discovery scientists, machine learning researchers, and visual analytics specialists.

The workflow comprised three main steps: (1) transform the trained model into decision rules to capture learned feature interactions; (2) use topic-guided visual analytics to reveal co-occurring conditions and navigate rules predicting high potency; and (3) map identified conditions to molecular substructures to obtain chemical interpretations.

The novelty lies in combining topic-guided exploration of rule-based model representations with feature-to-structure mapping. This addresses unmet needs by revealing multi-feature co-occurrences beyond marginal effects and linking model patterns to specific molecular substructures that domain scientists can interpret. From this case study, we learned that three elements are key: a rule-based model representation, topic-guided filtering, and structural mapping. Together they yielded candidate combinations of molecular components and substructures associated with high predicted potency for the studied class. Limitations include a single compound class, reliance on one model representation, and the lack of com-

Digital Object Identifier 10.1109/XXX.0000.0000000
† These authors contributed equally.

prehensive evaluation. Nonetheless, the case outlines a practical workflow potentially transferrable to other targets and compound classes: train an interpretable model, convert it into rules, and use visual analytics to extract and assess domain hypotheses.

Related Work

Compound potency prediction and model explanation

ML is commonly used to predict compound activity and potency, ranging from linear models to deep learning. Across large benchmarks, complex models often do not outperform simpler ones, which strengthens the case for models that are easier to interpret [1], [2]. To explain predictions, feature attribution methods such as LIME [3] and SHAP [4] are often employed. Attributions can be mapped to molecular substructures to aid interpretation [5]. However, these approaches often focus on estimating single-feature effects and generating local explanations, and may struggle to reveal multi-feature patterns and robust global structures [6]. Our work addresses this gap by analyzing co-occurring conditions in a rule-based model representation and relating them to recognizable substructures.

Visual analytics for rule- and ensemble-based models

Visual analytics (VA) provides approaches to make rule sets and ensembles more interpretable. Representative systems include RuleMatrix for structured rule inspection [7], iForest and Explainable Matrix for ensemble exploration and feature-rule relationships [8], [9], and RuleExplorer for navigating large rule sets [10]. Our work builds on this research but, unlike prior tools that mostly focus on model tuning or local reasoning about specific predictions, aims at spotting recurring multi-feature patterns linked to higher potency and then connects these patterns to molecular structures for domain interpretation.

Topic modeling in visual analytics

Topic modeling uncovers latent co-occurrence structure in collections of items by treating items as "terms" and collections as "documents" [11]. Visual analytics frameworks integrate interactive steering to improve interpretability and sense-making [12]. We build on these ideas by representing decision rules as "documents" and their conditions as "terms," enabling discovery of frequently co-occurring conditions in large rule sets. To our knowledge, the potential of applying this combination to interpret trained ML models has not been researched so far. Our case study demonstrates how this combination can surface interpretable, chemistry-relevant patterns.

Modeling Compound Potency

Our goal was to understand which molecular structural factors influence potency, i.e., how strongly a compound binds to a biological target. Rather than focusing on predictive accuracy, we sought interpretable insight into the structure–activity relationships learned by an ML model, using the model as a lens on the underlying chemistry. For the domain scientists, the desired outcome was not a "better" model in terms of metrics but a better understanding of how structural features combine to yield active compounds. We therefore designed a workflow that integrates ML, which captures regularities in the data, with visual analytics (VA), which enables experts to explore these regularities interactively and discover chemically meaningful patterns.

We applied this workflow to a class of compounds acting as inhibitors of the MAPK p38 alpha kinase (ChEMBL ID 260), an enzyme that controls inflammatory response. The dataset, derived from the ChEMBL database (version 33) following [13], included 1,495 compounds with experimentally measured activity values. Each molecule was represented by extended connectivity fingerprints (ECFP4) [14], which encode the presence or absence of local atom environments as a high-dimensional binary feature vector. Illustrative examples of how the representation is constructed can be found in the original ECFP paper [14, Figures 1-2].

To keep the model interpretable, we selected a limited set of relevant features using recursive feature elimination based on their importance in a preliminary random forest (RF) model, where multiple decision trees are independently trained on different data subsets. The final feature set contained 30 binary variables representing distinct structural fragments. The RF regression model, trained on 70% of the data and evaluated on the remaining 30%, achieved a mean absolute error adequate for exploratory studies. However, performance was not the main goal; the value of the model lies in how it combines structural conditions to arrive at potency predictions.

To access the internal logic of the model, we converted it into an explicit set of decision rules, each corresponding to a path from root to leaf in a tree. Each rule specifies a combination of feature conditions (e.g., "fragment `266` = present" and "fragment `34` = absent") that yields a predicted potency level. This resulted in 10,670 unique rules comprising 125,841 individual conditions. The rule-based representation captures interactions between structural fragments learned from data and provides a suitable basis for visual and interactive exploration.

Interactive Analysis Workflow

After converting the model into decision rules, we sought to understand what these rules reveal about relationships between structural fragments and compound potency. Tabular rule representations, as in Fig. 1, and systems such as RuleMatrix [7], are commonly used for rule exploration, but our set of over ten thousand multi-condition rules made direct inspection impractical. We therefore adopted an interactive visual workflow, implemented using an extended version of a previously described composition of computational and visual techniques [15], to navigate, summarize, and interpret patterns in the model's logic.

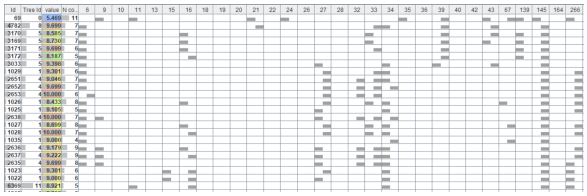

FIGURE 1. Fragment of table-based rule visualization. Each row corresponds to a rule. The first four columns show rule and tree identifiers, predicted potency, and number of conditions; the remaining columns are binary fingerprint features, left-aligned for absent (`0`) and right-aligned for present (`1`).

The workflow followed a progressive exploration strategy, moving from an overview of all rules toward more detailed inspection of promising subsets. This approach combined computational summarization and visual analytics, allowing domain scientists and data analysts to iteratively formulate and test hypotheses about the model's reasoning.

Overview of Feature-Prediction Relationships

As a starting point, we summarized how frequently each structural feature appeared across rules associated with different predicted potency levels. The continuous range of predicted potency values was divided into three intervals: low, medium, and high. The corresponding rule subsets contained 9,376, 1,106, and 188 rules. The selection of these breakpoints was guided by the domain scientists' focus on the highest potency values, while the intermediate interval was introduced to capture potential trends in feature occurrences across different potency levels.

For each interval, we generated an overview display (Fig. 2). Heatmap rows correspond to structural features and columns to binary values 0 for feature absence and 1 for presence. The color intensity encodes how frequently a given feature-value pair appears among the conditions of rules predicting within the respective potency interval, whereas the light blue bar to the left indicates the overall frequency of the feature in that subset of rules.

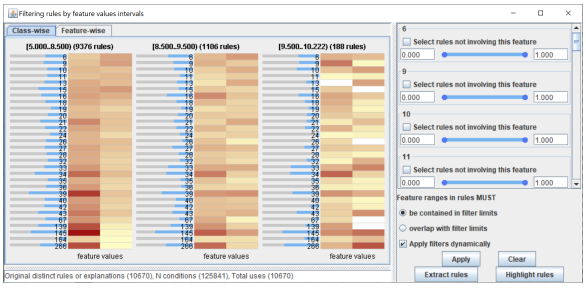

FIGURE 2. Overview of feature-value frequencies across three rule subsets defined by predicted potency intervals: [5, 8.5) - low, [8.5, 9.5) - medium, and [9.5, 10.22] - high.

FIGURE 3. Interactive rule filtering by feature conditions: top, rules with `266 = 1`; bottom, additional filter `145 = 1`.

From this overview, analysts interactively filtered the rule set to focus on specific conditions or combinations (e.g., rules that include or exclude a fragment). Each filtering step dynamically updated the heatmaps, revealing how selections affected the distribution of other feature values across rule subsets and enabling hypothesis-driven exploration of feature interactions and their joint impact on predicted outcomes. This iterative process enabled experts to test hypotheses such as "Does the presence of fragment A consistently

coincide with higher predicted potency?" or "Does fragment B reduce potency when fragment A is present?"

Through such exploration, several fragments emerged as being repeatedly involved in rules predicting higher potency values, while it also became clear that potency could not be attributed to any single fragment in isolation. Rather, specific combinations of fragments tend to co-occur in high-potency rules, hinting at possible structural synergies.

For example, Fig. 3 (top) shows the effect of filtering for rules in which feature `266` equals `1`. This subset contains 241 rules: 52 with low predicted values, 108 with medium values, and 81 with high values. Since most low-prediction rules are excluded by this filter, the presence of the component represented by feature `266` appears to be associated with increased predicted potency. Applying an additional filter for feature `145 = 1`, as shown in Fig. 3 (bottom), further narrows the selection to 61 rules: 5 in the low, 40 in the medium, and 16 in the high prediction intervals. This pattern indicates that the combination of features `266` and `145` may have a synergistic effect, jointly contributing to higher predicted compound potency.

Discovery of Co-Occurring Feature Patterns

While interactive filtering revealed a few promising fragment combinations, exhaustively exploring all possible feature interactions is infeasible due to the combinatorial complexity of the feature space. To make this process more systematic, we applied **topic modeling** using Non-negative Matrix Factorization (NMF) [16]. Each rule was treated as a "document" and each feature–value condition (e.g., `6 = 0`, `6 = 1`) as a "term," yielding a rule–condition matrix indicating which conditions occur in which rules. NMF was then applied to this matrix to identify groups of conditions that tend to co-occur across rules. Each topic corresponds to a set of feature–value conditions with high weights that often appear together, revealing interpretable co-occurrence patterns learned by the model.

Choosing the number of topics is a key modeling decision. We used two complementary strategies.

Embedding-Based Coverage Assessment. Following [17], we ran NMF with 5–25 topics and projected all topics into a common two-dimensional embedding. We selected configurations that provided good coverage of the space, i.e., topics well distributed across the main clusters so that the extracted topics represent diverse patterns without redundancy. By this criterion, models with 12 and 14 topics performed best.

Dominant Topic Strength. For each rule, we defined the dominant topic as the one with highest weight.

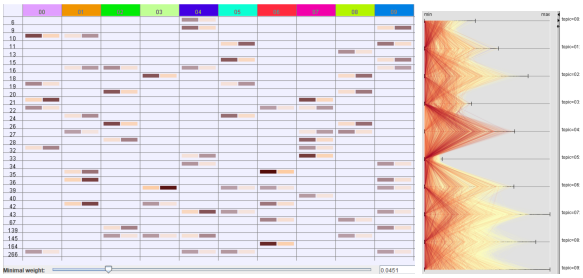

FIGURE 4. Topic modeling views. Left: term–topic matrix from NMF with 10 topics, showing high-weight feature conditions per topic. Right: parallel-coordinates plot of topic weight vectors for rules, colored by predicted potency group.

We preferred models in which most rules had strong associations with at least one topic, indicating that the topic structure captures meaningful regularities rather than distributing similar weights across many topics. By this criterion, 5- and 10-topic models performed best.

As an additional criterion, we evaluated which settings best separated rules predicting high potency from others. Among the candidates (5, 10, 12, and 14 topics), the 10-topic model gave the **clearest separation**, with one topic strongly linked to high-potency rules. We therefore selected the 10-topic model as a good compromise between granularity and interpretability.

Visualization of Topic Modeling Results

To interpret the topic modeling results, we used two complementary visualizations (Fig. 4). The **term–topic matrix** (left) summarizes relationships between feature conditions and topics, allowing us to identify which conditions most strongly characterize each topic. Cells contain segmented horizontal bars where darker left and right halves indicate values `0` and `1`, respectively, and opacity encodes the weight of the corresponding condition. Only conditions exceeding a threshold weight (e.g., 0.045) are shown. For instance, topic 4 is characterized by features `6`, `9`, `16`, `34`, and `266` with value `0` and features `33`, `43`, and `145` with value `1`.

The **parallel-coordinates plot** (PCP, Fig. 4 right) shows how strongly each rule is associated with the topics. Each axis corresponds to one topic, and rules are drawn as polylines connecting their topic weights, colored by predicted potency (low = yellow, medium = orange, high = red). This design supports multi-level comparison: within a rule, the polyline shape reflects relative topic weights; across topics, the density and spread of lines show their overall weight distribution patterns; and across the model, axis lengths reveal the maximal strengths of individual topics.

This dual representation enables joint inspection of

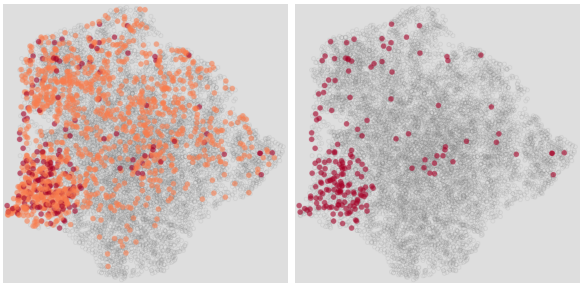

FIGURE 5. 2D projection of rules based on topic-weight vectors, colored by predicted potency (low, medium, high).

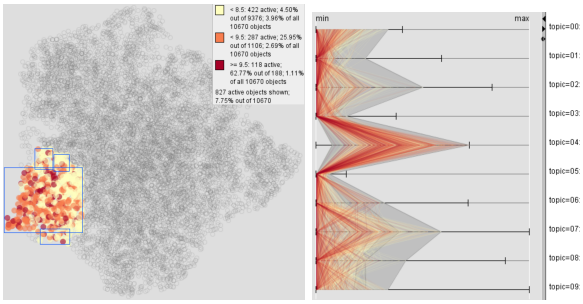

FIGURE 6. Focused exploration of the high-potency cluster. Left: spatial selection in the 2D projection. Right: PCP of selected rules, highlighting strong association with topic 4.

topic composition and topic-rule relationships, helping identify feature combinations that contribute most to model behavior. In our data, many red lines peaked at topic 4, suggesting that this topic captured feature patterns characteristic of highly potent compounds.

Exploring Topic Associations in Rule Space

To further investigate relationships between topics and rule subsets, we projected the topic-weight vectors into 2D space (Fig. 5) using a dimensionality reduction method from the class of neighborhood-preserving algorithms, which are suited for revealing clusters. Each dot in the projection represents a rule (red = high, orange = medium, gray = low predicted potency). High-potency rules form a distinct cluster in the left region.

To examine this region in more detail, we applied a spatial filter selecting the area containing the high-potency cluster (Fig. 6, left). This selection extracted 118 of 188 high-potency rules, 287 of 1,106 medium, and 422 of 9,376 low-potency rules. The PCP view of the selected subset (Fig. 6, right) shows that most have moderate to high weights for topic 4 (0.0153-0.1444). The combination of projection and filtering thus enables analysts to isolate coherent rule subsets for closer examination.

This analysis identified subsets of rules characterized by co-occurring feature conditions particularly associated with high potency. The subsequent analysis focused on these subsets to derive deeper insights into the structural factors influencing compound activity.

Identifying Candidate Patterns for High Potency

The cluster of rules predicting high potency with similar topic associations suggests that specific structural combinations may increase the likelihood of high potency. However, the corresponding region in projection space also contains medium- and low-potency rules

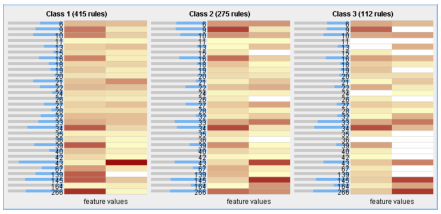

FIGURE 7. Distribution of feature–value pairs across rules of low, medium, and high potency within the selected cluster.

(Fig. 6), motivating a more detailed exploration of which features distinguish high-potency rules from the rest.

For this analysis, we discretized the numeric predictions into three classes (1 = low, 2 = medium, 3 = high potency) and focused on the 827 selected rules (422 + 287 + 118). Mapping numeric outputs to classes can create redundant rules that differ only by additional non-influential conditions, so we applied a redundancy-removal tool [15] that deletes subsumed rules, i.e., those covered by more general rules with the same prediction. After pruning, 802 rules remained (415 + 275 + 112). Figure 7 shows the distribution of feature-value pairs across the three classes. Features `11` and `36` are absent from all class-3 rules, indicating no contribution to high-potency predictions. Apart from this, no single feature uniquely characterizes class 3. The largest frequency differences across classes occur for features `266`, `164`, and `145`.

To reveal combinations distinguishing the classes, we reapplied NMF to the 802 rules, varying the number of topics between 5 and 25 and evaluating configurations by the distribution of dominant topic weights. Models with 5–8 topics showed similar structure: two topics linked to classes 2 and 3, one to class 1, and the rest shared between classes 1 and 2. Because the topic definitions were stable and the 5-topic model

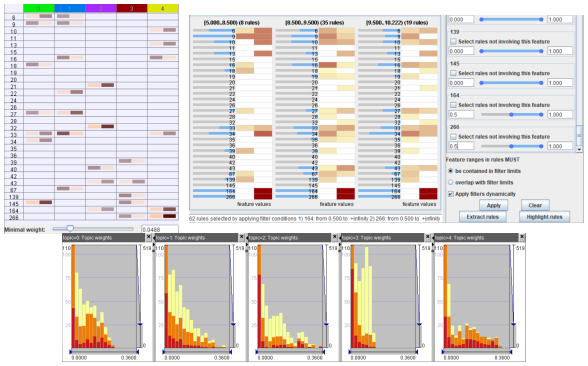

FIGURE 8. Topic-guided analysis of the rule cluster. Top left: term–topic matrix for five topics. Top right: interactive filtering based on topic-defining conditions. Bottom: topic-weight distributions across rules, colored by prediction class.

showed the strongest concentration among class-3 rules, we selected it to identify high-potency patterns. Limited separation between classes 2 and 3 is expected, as both correspond to high potency ranges ([8.5, 9.5) and $\geq$ 9.5).

The most prominent conditions in the five topics are shown in Fig. 8. Topics 0 and 4 are predominantly associated with rules predicting higher potency; topic 4 aligns most strongly with class 3 and is characterized by conditions `266 = 1` and `164 = 1`, whereas topic 0 is driven by `145 = 1`. Other conditions appear with lower weights, implying weaker or context-dependent co-occurrence.

Although the presence of components represented by features `266`, `164`, or `145` correlates with high potency, it is not uniquely predictive. To identify more specific combinations, we used interactive filtering informed by the topic results. As illustrated in Fig. 8 (top right), filtering the complete rule set for `266 = 1` and `164 = 1` leaves only eight low-class rules. Examining these revealed that all lack condition `34 = 0`. Adding this condition excluded the remaining low-class rules, retaining 14 medium- and 9 high-class rules. Adding `27 = 0` further isolated five high-class rules and one near-high rule, all predicting potencies between 9.179 and 9.886. Hence, the combination `266 = 1 & 164 = 1 & 34 = 0 & 27 = 0` consistently yields high-potency predictions.

Repeating this exploration with other topic-defining features showed that setting either `10 = 1` or `13 = 1` alongside `266 = 1` and `164 = 1` required `33 = 0` and `34 = 0` to exclude lower classes, resulting in only two rules predicting the highest potency. Additional explorations starting from single high-weight features

FIGURE 9. Selected molecular fragments corresponding to features `266`, `164`, `145`, `33`, and `34` that frequently occur in topics linked to high-potency predictions; blue and yellow circles mark central and aromatic atoms in the fingerprint-encoded fragments.

(`266`, `164`, or `145`) produced further high-potency combinations. In total, we identified eight feature combinations that drive the model to predict high potency.

Overall, components represented by features `266`, `164`, `145`, `33`, and `34` (Fig. 9) tend to contribute positively to compound potency, although their effect depends on context. Feature `6`, for example, often needs to be absent for others to act effectively. Conversely, features `21`, `22`, `27`, and `32` rarely appear in potent rules. These findings provided candidate structural patterns for domain-level validation and demonstrated the value of combining topic-guided exploration with interactive filtering to identify interpretable mechanisms of high potency.

Domain Interpretation and Validation

Having identified feature combinations associated with high potency, the next step was to validate these patterns and interpret them in chemical terms.

Visualization of Topic Conditions

Figure 9 illustrates several structural components contributing to high potency, represented as node-link diagrams. Features `33` and `34` correspond to a tertiary aliphatic carbon and an ether group, respectively, while larger fragments such as `164` and `266` share a chlorobenzene ring. The consistent occurrence of these fragments in rules predicting high potency suggests that they play a key role in the correct classification of highly active compounds.

Validation of Identified Topic Combinations

To assess the robustness of these findings, we screened the test compounds using the identified feature combinations. Figure 10 shows the distribution of predicted potency values and their associations with the combinations. Six of the eight identified combinations matched at least one test compound, while the remaining two were found only in highly potent

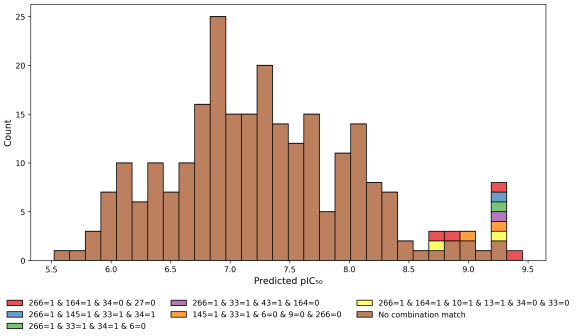

FIGURE 10. Predicted pIC_{50} values for test compounds associated with the identified high-potency feature combinations.

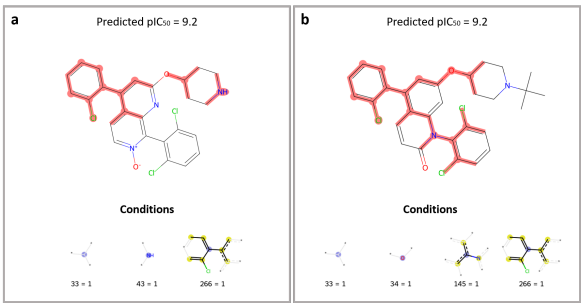

FIGURE 11. Examples of feature-to-structure mapping for two correctly predicted highly potent compounds; atoms belonging to topic-derived features are highlighted in red.

training examples. For matching compounds, predicted values ranged from 8.7 to 9.3 with a mean absolute error of about 0.25, confirming accurate predictions and demonstrating that identified feature combinations reliably capture structural factors linked to high potency.

Mapping of Topic Conditions

Features from topic combinations were then mapped to compound structures to support domain interpretation. Figure 11 presents two representative high-potency test compounds with features from the corresponding topic conditions highlighted.

In Fig. 11a, the three features (`33`, `43`, `266`) define two coherent substructures: a chlorobenzene ring on the left and a piperidine ring on the right. In Fig. 11b, four features (`33`, `34`, `145`, `266`) form a larger connected substructure combining a dichlorobenzene moiety with a quinolone ring. Here, the piperidine (`33`, `43`) and chlorobenzene (`145`, `266`) groups appear within a series of naphthyridine derivatives [18], a scaffold also present in several other highly potent compounds.

Overall, these mappings indicate that the ring systems highlighted by the topic-guided analysis may not only drive correct predictions but also contribute mechanistically to high compound potency.

CONCLUSION

From this application-oriented case study, we draw several lessons for using visual analytics to move from prediction to insight in machine-learning models of molecular properties. First, representing trained models in a rule-based form and analyzing co-occurrence structure in these rules (e.g., via topic modeling) can reveal multi-fragment patterns that go beyond single-feature attribution and are relevant for structure–property relationships. Second, coupling such topic-guided summaries with interactive visual exploration enables domain experts to navigate large rule spaces, focus on chemically meaningful subsets, and iteratively refine hypotheses about how structural features jointly influence a target property. Third, mapping model-derived feature combinations back to molecular substructures turns these patterns into testable mechanistic hypotheses; in our case, the highlighted ring systems and fragment constellations were consistent with highly potent compounds in training and test data.

Methodologically, the study emphasizes practical prerequisites for this style of analysis: an interpretable or rule-based model representation, a chemically meaningful structural encoding that supports feature–structure mapping, and close collaboration between machine learning, visual analytics, and domain experts during exploration. Under these conditions, the workflow can serve as a reusable template for insight-oriented analysis not only in compound potency modeling but more broadly in data-driven studies of chemical structure–property relationships.

This work is an application-focused case rather than a comprehensive benchmarking or user study. It is limited to considering a single compound class and biological target, relying on one model family, and using predominantly qualitative expert assessment. These limitations point to natural next steps: applying the workflow to additional targets and property types, assessing its impact on decision-making in medicinal chemistry, and systematically comparing and combining it with attribution-based techniques such as SHAP to validate and enrich interpretability analyses.

ACKNOWLEDGMENTS

This work was supported by Federal Ministry of Education and Research of Germany and the state of North-Rhine Westphalia as part of the *Lamarr Institute for Machine Learning and Artificial Intelligence*.

Author Information

Bahavathy Kathirgamanathan is a researcher at University of Cologne, Germany. Contact her at k.bahavathy@gmail.com.

Tiago Janela is a researcher at University of Bonn (Germany). Contact him at janela@bit.uni-bonn.de.

Elena Xerxa is a researcher at University of Bonn (Germany). Contact her at xerxae0@bit.uni-bonn.de.

Gennady and Natalia Andrienko are lead scientists at Fraunhofer IAIS in Sankt Augustin (Germany) and professors at City St George's University of London (UK). Contact them at {gennady|natalia}.andrienko@iais.fraunhofer.de.

Jürgen Bajorath is a professor at University of Bonn (Germany). Contact him at bajorath@bit.uni-bonn.de.

REFERENCES

1. W. P. Walters and R. Barzilay, "Applications of Deep Learning in Molecule Generation and Molecular Property Prediction," *Accounts of Chemical Research*, vol. 54, no. 2, pp. 263–270, 2021.
2. T. Janela and J. Bajorath, "Large-Scale Predictions of Compound Potency with Original and Modified Activity Classes Reveal General Prediction Characteristics and Intrinsic Limitations of Conventional Benchmarking Calculations," *Pharmaceuticals*, vol. 16, no. 4, p. 530, 2023.
3. M. Ribeiro, S. Singh, and C. Guestrin, ""Why Should I Trust You?": Explaining the Predictions of Any Classifier," in *ACM SIGKDD*, 2016, pp. 1135–1144.
4. S. M. Lundberg and S.-I. Lee, "A unified approach to interpreting model predictions," in *Advances in Neural Information Processing Systems*, vol. 30, 2017.
5. R. Rodríguez-Pérez and J. Bajorath, "Interpretation of machine learning models using shapley values: Application to compound potency and multi-target activity predictions," *Journal of Computer-Aided Molecular Design*, vol. 34, no. 10, pp. 1013–1026, 2020.
6. J. P. Roth and J. Bajorath, "Machine learning models with distinct Shapley value explanations decouple feature attribution and interpretation for chemical compound predictions," *Cell Reports Physical Science*, vol. 5, no. 8, p. 102110, 2024.
7. Y. Ming, H. Qu, and E. Bertini, "RuleMatrix: Visualizing and understanding classifiers with rules," *IEEE transactions on visualization and computer graphics*, vol. 25, no. 1, pp. 342–352, 2018.
8. X. Zhao, Y. Wu, D. L. Lee, and W. Cui, "iForest: Interpreting random forests via visual analytics," *IEEE Transactions on Visualization and Computer Graphics*, vol. 25, no. 1, pp. 407–416, 2019.
9. M. P. Neto and F. V. Paulovich, "Explainable matrix - visualization for global and local interpretability of random forest classification ensembles," *IEEE Transactions on Visualization and Computer Graphics*, vol. 27, no. 2, pp. 1427–1437, 2021.
10. Z. Li, W. Yang, J. Yuan, J. Wu, C. Chen, Y. Ming, F. Yang, H. Zhang, and S. Liu, "RuleExplorer: A scalable matrix visualization for understanding tree ensemble classifiers," 2024. [Online]. Available: https://arxiv.org/abs/2409.03164
11. I. Vayansky and S. A. Kumar, "A review of topic modeling methods," *Information Systems*, vol. 94, p. 101582, 2020.
12. M. El-Assady, F. Sperrle, O. Deussen, D. Keim, and C. Collins, "Visual analytics for topic model optimization based on user-steerable speculative execution," *IEEE Transactions on Visualization and Computer Graphics*, vol. 25, no. 1, pp. 374–384, 2019.
13. T. Janela and J. Bajorath, "Anatomy of potency predictions focusing on structural analogues with increasing potency differences including activity cliffs," *Journal of Chemical Information and Modeling*, vol. 63, no. 22, pp. 7032–7044, 2023.
14. D. Rogers and M. Hahn, "Extended-connectivity fingerprints," *Journal of Chemical Information and Modeling*, vol. 50, no. 5, pp. 742–754, 2010.
15. L. Adilova, M. Kamp, G. Andrienko, and N. Andrienko, "Re-interpreting rules interpretability," *International Journal of Data Science and Analytics*, vol. 20, no. 1, pp. 25–45, 2025.
16. J. Choo, C. Lee, C. K. Reddy, and H. Park, "Utopian: User-driven topic modeling based on interactive nonnegative matrix factorization," *IEEE Transactions on Visualization and Computer Graphics*, vol. 19, no. 12, pp. 1992–2001, 2013.
17. G. Andrienko, N. Andrienko, and D. Hecker, "Topic modelling for spatial insights: Uncovering space use from movement data," *Computers & Graphics*, vol. 122, p. 103989, 2024.
18. W. Lumeras, L. Vidal, B. Vidal, C. Balagué, A. Orellana, M. Maldonado, M. Domínguez, V. Segarra, and F. Caturla, "1,7-Naphthyridine 1-Oxides as Novel Potent and Selective Inhibitors of p38 Mitogen Activated Protein Kinase," *Journal of Medicinal Chemistry*, vol. 54, no. 22, pp. 7899–7910.