GaCoVi: a Correlation Visualization to Support Interpretability-Aware Feature Selection for Regression Models

Diego Rojo, Nyi Nyi Htun and Katrien Verbert

Department of Computer Science, KU Leuven, Belgium

Abstract

The recent growth of interest in explainable artificial intelligence (XAI) has resulted in a large number of research efforts to provide accountable and transparent machine learning systems. Although a large volume of research has focused on algorithm transparency, there are other factors that influence the interpretability of a system, such as end-users' understanding of individual features and the total number of features. Thus, involving end-users in the feature selection process may be key to achieving interpretability. In addition, previous work has suggested that to obtain satisfactory interpretability and predictive performance, the feature selection process should look for a subset of features that are highly correlated with the response variable yet uncorrelated to each other. Taking this into account, in this paper, we present a work-in-progress design study of a novel system for correlation visualization, GaCoVi. GaCoVi is designed to put domain experts in the loop of feature selection for regression models in scenarios where transparency of the machine learning systems is crucial.

CCS Concepts

• ***Human-centered computing*** → ***Visualization systems and tools;*** • ***Computing methodologies*** → ***Feature selection;***

1. Introduction

The extended use of machine learning (ML) on high-stakes decision making in domains such as healthcare, banking, education and employment has led to an increasing concern about the accountability and transparency of ML systems [ACM17]. This matter has influenced even legal regulations, such as the introduction of the European General Data Protection Regulation (GDPR) involving the right for individuals to obtain "meaningful explanations of the logic involved" in automated decision-making systems. Besides, transparency of ML systems is also essential for domain experts to trust them and, thus, use them confidently in their practices [ZC18]. These needs have resulted in the growth of explainable artificial intelligence (XAI), a field that aims at producing more comprehensible systems while maintaining high predictive performance.

A large volume of XAI research is devoted to automatically generating (post-hoc) explanations of black-box models whose logic is not comprehensible for humans [GMR*18]. However, when dealing with structured datasets with naturally meaningful features, some researchers advocate for the use of inherently interpretable models (e.g., linear regression) instead of black-box models, arguing that interpretable models could replace black box models while achieving equal accuracy [Rud19]. Employing an inherently interpretable model leads undoubtedly to more *algorithm transparency*, i.e., a better understanding of the inner working of the system. However, as pointed out by Lipton [Lip18], when analyzing the interpretability of an ML system there are at least two other levels of transparency to consider: *decomposability*, i.e., every part of the model, such as inputs, parameters, and calculations, admits an intuitive explanation, and *simulatability*, i.e., the model is simple enough for a user to grasp it as a whole at once.

The feature selection (FS) process is critical to attaining both decomposability and simulatability. On the one hand, including features that are anonymous, highly engineered, or unknown by the user could prevent from achieving decomposability [Lip18]. On the other hand, to attain simulatability, there should be a cognitively manageable number of features, where this number would depend on whether the user can relate the features to pre-existing mental models or previous knowledge [Rüp06]. Therefore, involving the final users, i.e., the ones that are going to interpret the ML system, in the FS process is crucial to improve the interpretability of ML systems. Nonetheless, to have reliable interpretability and adequate predictive performance, the FS process should also pursue common goals. Namely, the subset selection should include features that are highly correlated with the response variable yet uncorrelated to each other [Hal00, PRKV*17].

In literature, visual analytics systems have been proposed for the task of FS as a preceding step to prediction (classification or regression) tasks. Although most of the literature focuses on scenarios of classification of structured [KPB14] or unstructured [DPL*19] datasets, an example of a system designed for regression scenarios like the ones we consider in this paper is FeatureExplorer [ZKM*19]. FeatureExplorer supports domain experts on FS

DOI: 10.2312/evs.20201060

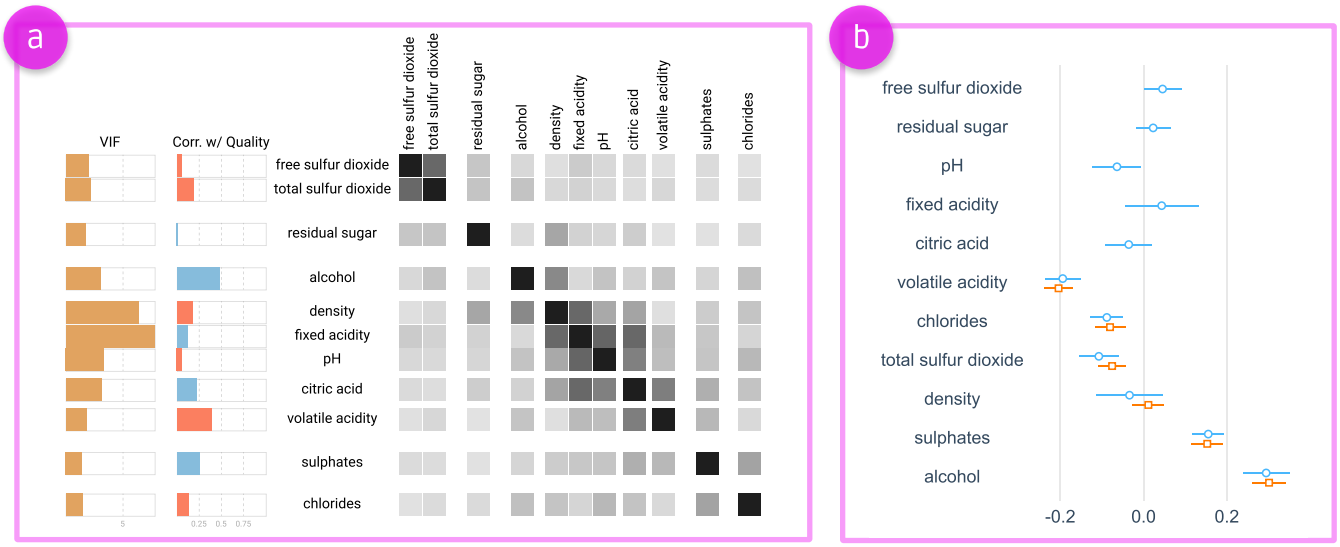

Figure 1: *(a) The GaCovi visualization for the red wine quality dataset [CCA*09] and (b) the 95% confidence intervals of the standardized coefficients of the two multiple linear regression models: without feature selection (blue circles) and with feature selection (orange squares).*

in a plant biomass prediction scenario using, among other components, a corrgram [Fri02] for correlation visualization. However, FeatureExplorer's goals and evaluation are focused on predictive performance and not on interpretability and, thus, the correlation visualization design is not optimized, for example, to avoid the selection of features that are highly correlated with each other, which could affect interpretability. To address this gap and better support domain experts in an interpretability-aware FS process, we developed, using Vega [SRHH16], GaCoVi (Gapped Correlation Visualization). GaCoVi (see Figure 1a) is a novel correlation visualization to include domain experts in the loop of FS for regression models in scenarios where transparency of the ML systems is crucial. This paper presents the problem-oriented design study carried out following the nested model for visualization design proposed by Munzner [Mun09] and shows how GaCoVi can help obtaining more interpretable models in a wine quality prediction scenario.

2. Domain Problem Characterization

Our target users are domain experts who are comfortable doing data analysis, but have little to no ML experience. In particular, users should be familiar with common statistical measures such as Pearson's correlation coefficient (PCC), a measure of the linear association between two variables. This description includes, among others, scientists and academics from most scientific disciplines. The target users would typically be working on some numerical prediction problems using a structured dataset with meaningful features from their domain. We designed our system for datasets where the number of variables is less than 30, and where the response variable, i.e., the target of the prediction problem, and the features, i.e., the explanatory variables, are both numeric.

To support the previously described users to carry out an interpretability-aware FS, we devised a number of challenges that should be addressed by the proposed system:

C1. Identification of irrelevant features. Removing features with little to no predictive power of the response variable is a common feature selection heuristic. Irrelevant features not only will not improve the predictive performance of the models, but can also deteriorate the performance of models sensitive to this superfluous input, such as support vector machines or neural networks [KJ19]. A typical automatic approach is to discard features that had a correlation strength with the response variable lower than a certain threshold. In the FS workflow described in [PRKV*17], this threshold is 0.3, which coincides with Cohen's correlation medium effect size [Coh88]. In an interpretability-aware scenario, also anonymous features and features that have little to no meaning for the user should be considered less relevant.

C2. Identification of pairs of highly correlated features and selection of the most relevant feature within them. The interpretation of some inherently interpretable models (e.g., multiple linear regression) and some popular methods for explaining black-box models (e.g., permutation feature importance, partial dependence plots, individual conditional expectation plots) become unreliable in the presence of highly correlated features [SBK*08, TL11, HM19]. Besides, the inclusion of highly correlated features can also negatively affect the predictive performance of some models, such as multiple linear regression or Naive Bayes [KJ19, LS94]. Therefore, the user should be able to detect all pairs of highly correlated features and only select one of them, the most relevant one. Note that, in this scenario, the relevance of a feature is determined by a combination of its correlation with the response variable and how interpretable it is for the user.

C3. Reduction to a reasonable subset size. Successfully addressing the previous challenges (C1 and C2) reduces, in general, the size of the selected feature subset. However, the number of remaining features may still be not cognitively manageable for the users, disabling their capacity to grasp the model as a whole. Some inter-

pretability heuristics, such as [WS03], fix the number of features that are cognitively manageable at 7 ± 2 based on Miller's meta-analysis [Mil56], which states that human beings can only deal with that amount of information chunks. However, as pointed out by Rüping [Rüp06], features should not be confused with chunks, since what counts as a chunk depends on the prior knowledge of every person. If a group of features is familiar to the user who has previously grouped and memorized them, they can count as a single information chunk. For example, a doctor could group, thanks to prior knowledge, multiple symptoms into an illness and reason with the latter. Therefore, the user should be able to identify the features in the subset after addressing the previous challenges and continue discarding features, based on their relevance, until the feature subset becomes cognitively manageable.

3. Data Abstraction and Transformations

There are two different approaches to visualize correlations. A first approach is visualizing the raw data and counting on the user to determine the correlation coefficients visually. This is the approach of scatterplot matrices (SPLOM) [Har75] or parallel coordinates [Ins85]. The main problem of these methods is that they need ample screen space as soon as the number of observations or features grows. Friendly [Fri02] suggests SPLOM as a good representation for correlation visualization for datasets of up to 10 variables, which is not enough for the requirements of this scenario, in which we are focusing on datasets of up to 30 features.

The second approach, the one that leads to our data abstraction, consists in directly encoding the correlation coefficients. This is the approach followed by standard correlation matrix based visualizations, such as corrgrams [Fri02] or clustered heatmaps [WF09], that transform this type of raw tabular data into a network data type (which adjacency matrix is the correlation matrix). Precisely, the original data gets transformed into a fully connected undirected network in which each node represents a variable, and the weight of the edge between two nodes corresponds to the correlation coefficient. In the prediction scenario we are analyzing, this data abstraction has an issue: it does not consider the special role the response variable has.

When correlation visualization is used in similar prediction scenarios, this data abstraction sometimes gets modified by adding an extra attribute to the nodes that indicate which is the response variable. However, using this data abstraction, we have only found slightly different visualizations encodings. In particular, in FeatureExplorer [ZKM*19] the authors highlight the name on the response variable in the correlation matrix visualization, and Kozak et al. [KWTH12] suggest giving the first position to the response variable on a correlation table.

As an alternative, we propose to encode the original data into a different multivariate network (MVN) data abstraction in which the response variable is not a node of the network. Instead, its correlation coefficient with each of the features becomes an attribute of each of the features' node. Besides, since we are mostly interested in the strength, i.e., absolute value, of the correlations and not their direction, we decide that the data abstraction has two different edge attributes: the correlation strength that has the role of edge weight, and the correlation direction (positive or negative). The use of this data abstraction is one of the fundamental design decisions of this study as it leads to a different design space on the visual encodings level, the MVN design space. In addition, this encoding makes straight forward the incorporation of other metrics than can aid in the FS process to the data abstraction, such as variance inflation factor (see section 6) or feature importance scores, such as the one used in FeatureExplorer [ZKM*19].

4. Task Abstraction

In this section, we present the domain-independent tasks that the visualization should support. These tasks have been derived from the domain challenges described in section 2 using the data abstraction proposed in section 3. Tasks T1 and T2 are defined in the analysis of tasks for MVN carried out by Pretorius et al. [PPS14].

T1. **Find the nodes with specific attribute values.** This task is referred to as 'nodes (properties)' in [PPS14]. In this scenario, it corresponds with finding the features that are irrelevant based on their attributes, particularly their name and correlation with the response variable.

T2. **Given a node, find the nodes connected by edges with specific attribute values.** This task is referred to as 'links (connected nodes)' in [PPS14]. In our scenario, it corresponds with finding, for each node, all connected nodes whose edge weight attribute, i.e., correlation strength, is high.

5. Visual Encodings and Algorithms

Adding the correlation with the response variable as an attribute necessarily implies the search of different approaches that support its encoding. The MVN design space is thoroughly characterized in the recent survey carried out by Nobre et al. [NMSL19]. In this survey, the authors propose the juxtaposition of node attributes visualizations with the rows or columns of an adjacency matrix as a technique very well suited for dense and even fully connected MVNs where tasks involve analyzing neighborhoods or clusters. This approach is followed in the design of systems such as TaMax [SL19]. The use of adjacency matrices with juxtaposed node attribute seems like a good fit for this scenario since it is similar to common correlation visualizations such as clustered histograms and corrgrams, which makes the visualization more familiar for our users.

In the next subsections, we justify the matrix reordering algorithm and an iconic feature that shapes the name of GaCoVi: the gaps between rows and columns. We then describe the different visual encodings for both nodes and edges.

5.1. Seriation Algorithm and Gaps

Seriation and matrix reordering methods are one of the key factors influencing the efficacy of high-level network tasks such as identifying clusters or highly-connected nodes in adjacency matrix visualizations [NMSL19, BBHR*16]. The adjacency matrix of our problem is $\mathbf{C} = [c_{i,j}]$, where $c_{i,j} = c_{j,i}$ is the correlation strength, i.e., the absolute value of the correlation, between the features corresponding to nodes i and j. In this scenario, we are looking for a seriation algorithm that permutes the order of the rows and columns

of the adjacency matrix in a way that improves the efficacy in the performance of task T2, i.e., finding all nodes that are highly correlated with a given node. If we were able to reorder the rows (resp. columns) of $\mathbf{C}$ so that the correlation strength values monotonically decreases when moving away from the diagonal (see Equation 1), then the closer two rows (resp. columns) are, the higher their correlation strength would be.

$$\begin{cases} c_{x,a} \leq c_{x,b} & \text{for} \quad a < b < x \\ c_{x,a} \geq c_{x,b} & \text{for} \quad x < a < b \end{cases} \tag{1}$$

A symmetric matrix $\mathbf{C}$ that verifies Equation 1 is called a Robinson similarity matrix and a matrix that can be reordered in a way that verifies Equation 1 is called pre-robinsonian matrix [BBHR*16]. Unfortunately, not all correlation strength matrices are pre-robinsonian, so in general, we are not able to reorder the matrix so that all the inequalities of Equation 1 are verified. However, multiple seriation methods approximate the ordering by optimizing different criteria. For task T2, features with similar correlation strength values to a given node should both have the same fate, i.e., either both or none should be selected, so small differences in the inequalities of Equation 1 have less impact on performance. Thus, the selected criterion is the weighted gradient measure, i.e., the difference between agreements and disagreements of the inequalities of Equation 1 weighted by their magnitude. For a small number of variables, it is feasible to use seriation by *branch-and-bound* to find the order that minimizes the weighted gradient measure [BS05]. Otherwise, heuristics methods such as spectral and quadratic assignment problem (QAP) based seriations are suggested, since they perform well on gradient-based criteria [Hah17].

To increase the efficacy of checking whether two consecutive rows are highly correlated (a necessary step to perform task T2), we introduce a gap between each pair of rows (resp. column). This approach was inspired by gapmaps, i.e., clustered heatmaps with gaps between rows or columns that are not tightly clustered, which have shown promising results in clustering tasks applied to non-symmetrical matrices [EWJP17]. In our case, the gap size is inversely proportional to the correlation strength between the features represented by those rows (resp. columns), taking values from zero to the height (or width) of a cell (see Figure 1a). For mapping the correlation values into the gap, we have opt-in for a decreasing logistic function centered at 0.5, i.e, $f(c) = S/(1+e^{\alpha(c-0.5)})$, where S is the height (or width) of a cell, c is the correlation strength between the corresponding features and α is the steepness of the curve. Although further evaluation is needed, we decided to use a logistic instead of a linear mapping to make it easier to quickly distinguish very weak or very strong correlations. In particular, we select the logistic function $f(c)$ with $\alpha = 14.72$ so that the gaps associated with correlations lower than 0.3 are bigger than 95% of a cell size, and the gaps associated with correlations higher than 0.7 are smaller than 5% of a cell size.

5.2. Edge and Node Attributes Encoding

Our multivariate network has two edge attributes, Pearson's correlation coefficient (PCC) direction (positive or negative) and its strength (numerical attribute). We encode the PCC strength into the color luminance of the adjacency matrix cells, using the Vega [SRHH16] sequential 'Greys' color scheme (see Figure 1a). Other alternatives are encoding it into the size of glyphs such as circles or bars [Fri02], but this can obscure the matrix grid structure [ABHR*13] and, in our case, also affect the gap perception. Since it is not essential for any of the tasks, the PCC direction is only shown on demand, when hovering over a cell, by coloring the borders of the adjacency matrix cell in either red (when negative) or blue (when positive). The nodes' attributes are encoded in a tabular layout, enabling the comparisons between nodes on the same scale. The PCC strength with the response variable is encoded as the length of a bar, which color (red or blue) also encodes the PCC direction. A bar mark is also used in Figure 1a for the variance inflation factor (VIF), a measure of the amount of multicollinearity in a set of multivariate regression features.

6. Use Cases: Wine Quality Prediction

In this section, we show anecdotal evidence of how GaCoVi can lead to more interpretable results. The dataset used is a red wine quality dataset devised by Cortez et al. [CCA*09], which is available at UCI ML repository [DG17]. The 11 input features can be seen in Figure 1, and the response variable is a quality measure scored from 0 to 10. The dataset contains 1599 instances of red wine. Since all features are naturally meaningful, we suppose the user has a similar knowledge of all of them.

Using GaCoVi (see Figure 1a), the user starts by discarding the features that seem not relevant based on shallow correlation values with the response variable. In this case, a user would probably discard free sulfur dioxide, residual sugar, and pH. Note that after removing these three values, all VIF values are below 5, so the multicollinearity issues seem fixed. Next, the user discards fixed acidity since it is highly correlated with density, but it is less relevant. Finally, a user would probably discard citric acid since it is not very relevant and is quite highly correlated with volatile acidity.

To evaluate the feature selection, we trained a multiple linear regression model with all the features ($R^2 = 0.3606$) and with the selected features ($R^2 = 0.3516$). Although the predictive performance of the model with all the features is slightly better, its standardized coefficients exhibit a higher variance, as can be seen in Figure 1b, where its 95% confidence intervals (in blue) are wider. The higher variance of the coefficients makes their interpretation as the strength of association with the quality of wine less reliable.

7. Conclusions and Future Work

In this paper, we introduce a novel correlation visualization to include domain experts in the loop of interpretability-aware feature selection for regression models. Although we have carried out some validations, such as the justification of the visual encoding or the proposed algorithms, future work should include downstream validations involving the target users. Besides, we consider working on exposing uncertainty in the visualization and researching techniques to aid in the use of GaCoVi with high-dimensional data.

Acknowledgements

This project was supported by the BIGDATAGRAPES project (EU H2020 RIA, grant agreement 780751).

References

[ABHR*13] ALPER B., BACH B., HENRY RICHE N., ISENBERG T., FEKETE J. D.: Weighted graph comparison techniques for brain connectivity analysis. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (New York, NY, USA, 2013), CHI '13, Association for Computing Machinery, p. 483–492. doi:10.1145/2470654.2470724. 4

[ACM17] ACM U.S. PUBLIC POLICY COUNCIL AND ACM EUROPE POLICY COMMITTEE: Statement on algorithmic transparency and accountability, 2017. 1

[BBHR*16] BEHRISCH M., BACH B., HENRY RICHE N., SCHRECK T., FEKETE J. D.: Matrix reordering methods for table and network visualization. *Computer Graphics Forum 35*, 3 (2016), 693–716. doi:10.1111/cgf.12935. 3, 4

[BS05] BRUSCO M. J., STAHL S.: *Branch-and-bound applications in combinatorial data analysis*. Springer-Verlag New York, 2005. 4

[CCA*09] CORTEZ P., CERDEIRA A., ALMEIDA F., MATOS T., REIS J.: Modeling wine preferences by data mining from physicochemical properties. *Decision Support Systems 47*, 4 (2009), 547–553. 2, 4

[Coh88] COHEN J.: *Statistical power analysis for the social sciences*. Routledge, New York, 1988. doi:10.4324/9780203771587. 2

[DG17] DUA D., GRAFF C.: UCI machine learning repository, 2017. URL: http://archive.ics.uci.edu/ml. 4

[DPL*19] DENNIG F. L., POLK T., LIN Z., SCHRECK T., PFISTER H., BEHRISCH M.: Fdive: Learning relevance models using pattern-based similarity measures. In *2019 IEEE Conference on Visual Analytics Science and Technology (VAST)* (Oct 2019), pp. 69–80. doi:10.1109/VAST47406.2019.8986940. 1

[EWJP17] ENGLE S., WHALEN S., JOSHI A., POLLARD K. S.: Unboxing cluster heatmaps. *BMC Bioinformatics 18*, 2 (2017), 63. doi:10.1186/s12859-016-1442-6. 4

[Fri02] FRIENDLY M.: Corrgrams. *The American Statistician 56*, 4 (2002), 316–324. doi:10.1198/000313002533. 2, 3, 4

[GMR*18] GUIDOTTI R., MONREALE A., RUGGIERI S., TURINI F., GIANNOTTI F., PEDRESCHI D.: A survey of methods for explaining black box models. *ACM Comput. Surv. 51*, 5 (Aug. 2018). doi:10.1145/3236009. 1

[Hah17] HAHSLER M.: An experimental comparison of seriation methods for one-mode two-way data. *European Journal of Operational Research 257*, 1 (2017), 133 – 143. doi:https://doi.org/10.1016/j.ejor.2016.08.066. 4

[Hal00] HALL M. A.: Correlation-based feature selection for discrete and numeric class machine learning. In *Proceedings of the Seventeenth International Conference on Machine Learning* (San Francisco, CA, USA, 2000), ICML '00, Morgan Kaufmann Publishers Inc., p. 359–366. 1

[Har75] HARTIGAN J.: Printer graphics for clustering. *Journal of Statistical Computation and Simulation 4*, 3 (1975), 187–213. doi:10.1080/00949657508810123. 3

[HM19] HOOKER G., MENTCH L.: Please stop permuting features: An explanation and alternatives, 2019. arXiv:1905.03151. 2

[Ins85] INSELBERG A.: The plane with parallel coordinates. *The Visual Computer 1*, 2 (1985), 69–91. doi:10.1007/BF01898350. 3

[KJ19] KUHN M., JOHNSON K.: *Feature engineering and selection: A practical approach for predictive models*. Chapman and Hall/CRC, 2019. 2

[KPB14] KRAUSE J., PERER A., BERTINI E.: Infuse: Interactive feature selection for predictive modeling of high dimensional data. *IEEE Transactions on Visualization and Computer Graphics 20*, 12 (Dec 2014), 1614–1623. doi:10.1109/TVCG.2014.2346482. 1

[KWTH12] KOZAK M., WNUK A., TARTANUS M., HARTLEY J.: Key variables come first! how best to design a correlation table when there is one key variable. *Journal of Information Science 38*, 4 (2012), 399–403. doi:10.1177/0165551512445159. 3

[Lip18] LIPTON Z. C.: The mythos of model interpretability. *Commun. ACM 61*, 10 (Sept. 2018), 36–43. doi:10.1145/3233231. 1

[LS94] LANGLEY P., SAGE S.: Induction of selective bayesian classifiers. In *Uncertainty Proceedings 1994*, de Mantaras R. L., Poole D., (Eds.). Morgan Kaufmann, San Francisco (CA), 1994, pp. 399 – 406. doi:10.1016/B978-1-55860-332-5.50055-9. 2

[Mil56] MILLER G. A.: The magical number seven, plus or minus two: some limits on our capacity for processing information. *Psychological Review 63*, 2 (1956), 81–97. doi:10.1037/h0043158. 3

[Mun09] MUNZNER T.: A nested model for visualization design and validation. *IEEE Transactions on Visualization and Computer Graphics 15*, 6 (Nov. 2009), 921–928. doi:10.1109/TVCG.2009.111. 2

[NMSL19] NOBRE C., MEYER M., STREIT M., LEX A.: The state of the art in visualizing multivariate networks. *Computer Graphics Forum 38*, 3 (2019), 807–832. doi:10.1111/cgf.13728. 3

[PPS14] PRETORIUS J., PURCHASE H. C., STASKO J. T.: *Tasks for Multivariate Network Analysis*. Springer International Publishing, Cham, 2014, pp. 77–95. doi:10.1007/978-3-319-06793-3_5. 3

[PRKV*17] PEREZ-RIVEROL Y., KUHN M., VIZCAÍNO J. A., HITZ M.-P., AUDAIN E.: Accurate and fast feature selection workflow for high-dimensional omics data. *PLOS ONE 12*, 12 (12 2017), 1–14. doi:10.1371/journal.pone.0189875. 1, 2

[Rud19] RUDIN C.: Stop explaining black box machine learning models for high stakes decisions and use interpretable models instead. *Nature Machine Intelligence 1*, 5 (2019), 206–215. doi:10.1038/s42256-019-0048-x. 1

[Rüp06] RÜPING S.: *Learning interpretable models*. PhD thesis, Univ. Dortmund, 2006. doi:10.17877/DE290R-8863. 1, 3

[SBK*08] STROBL C., BOULESTEIX A.-L., KNEIB T., AUGUSTIN T., ZEILEIS A.: Conditional variable importance for random forests. *BMC Bioinformatics 9*, 1 (2008), 307. doi:10.1186/1471-2105-9-307. 2

[SL19] SAFARLI I., LEX A.: Tamax: Visualizing dense multivariate networks with adjacency matrices. In *Proceedings of the IEEE Information Visualization Conference – Posters (InfoVis '19)* (2019). 3

[SRHH16] SATYANARAYAN A., RUSSELL R., HOFFSWELL J., HEER J.: Reactive vega: A streaming dataflow architecture for declarative interactive visualization. *IEEE Transactions on Visualization and Computer Graphics 22*, 1 (Jan 2016), 659–668. doi:10.1109/TVCG.2015.2467091. 2, 4

[TL11] TOLOŞI L., LENGAUER T.: Classification with correlated features: unreliability of feature ranking and solutions. *Bioinformatics 27*, 14 (05 2011), 1986–1994. doi:10.1093/bioinformatics/btr300. 2

[WF09] WILKINSON L., FRIENDLY M.: The history of the cluster heat map. *The American Statistician 63*, 2 (2009), 179–184. doi:10.1198/tas.2009.0033. 3

[WS03] WEIHS C., SONDHAUSS U. M.: Combining mental fit and data fit for classification rule selection. In *Exploratory Data Analysis in Empirical Research* (Berlin, Heidelberg, 2003), Schwaiger M., Opitz O., (Eds.), Springer Berlin Heidelberg, pp. 188–203. 3

[ZC18] ZHOU J., CHEN F.: *2D Transparency Space—Bring Domain Users and Machine Learning Experts Together*. Springer International Publishing, Cham, 2018, pp. 3–19. doi:10.1007/978-3-319-90403-0_1. 1

[ZKM*19] ZHAO J., KARIMZADEH M., MASJEDI A., WANG T., ZHANG X., CRAWFORD M. M., EBERT D. S.: Featureexplorer: Interactive feature selection and exploration of regression models for hyperspectral images. In *2019 IEEE Visualization Conference (VIS)* (Oct 2019), pp. 161–165. doi:10.1109/VISUAL.2019.8933619. 1, 3