

# Human-Based and Automatic Feature Ideation for Time Series Data: A Comparative Study

J. Schmidt<sup>1</sup>, H. Piringer<sup>2</sup>, T. Mühlbacher<sup>2</sup>, and J. Bernard<sup>3,4</sup>

<sup>1</sup>VRVis Zentrum für Virtual Reality und Visualisierung Forschungs-GmbH, Austria

<sup>2</sup>Visplore GmbH, Austria

<sup>3</sup>University of Zurich, Switzerland

<sup>4</sup>Digital Society Initiative, Zurich, Switzerland

## Abstract

Feature ideation is a crucial early step in the feature extraction process, where new features are extracted from raw data. For phenomena existing in time series data, this often includes the ideation of statistical parameters, representations of trends and periodicity, or other geometrical and shape-based characteristics. The strengths of automatic feature ideation methods are their generalizability, applicability, and robustness across cases, whereas human-based feature ideation is most useful in uncharted real-world applications, where incorporating domain knowledge is key. Naturally, both types of methods have proven their right to exist. The motivation for this work is our observation that for time series data, surprisingly few human-based feature ideation approaches exist. In this work, we discuss requirements for human-based feature ideation for VA applications and outline a set of characteristics to assess the goodness of feature sets. Ultimately, we present the results of a comparative study of human-based and automated feature ideation methods, for time series data in a real-world Industry 4.0 setting. One of our results and discussion items is a call to arms for more human-based feature ideation approaches.

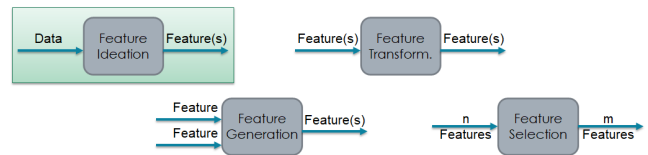
## 1. Introduction

Feature extraction is an important step in many data-driven analysis workflows. While research domains like information retrieval, data mining, and machine learning have extensively studied automatic methods for feature extraction, pioneer visual analytics (VA) approaches have proven the usefulness of human-in-the-loop feature extraction. The feature extraction process can be structured into four steps [CMKK22], which is useful to subdivide challenges, Figure 1 shows a formalization.

*Feature ideation* refers to extracting new features from (raw) input data. These features can be combined in the *feature generation* step, to form more precise representations. *Feature transformation* describes changes made to existing features. Finally, *feature selection* leads to a representative subset of features for further usage. Example VA approaches include FeatureEnVi [CMKK22] for feature generation, transformation, and selection, SomFlow [SKB\*18] for feature transformation, as well as INFUSE [KPB14] and FeatureExplorer [ZKM\*19] for feature selection.

We concentrate on a spot that is comparatively blind in VA research: the study of *feature ideation (FI) for time series data (TS)*. Similarly as for most data types, features for TS are compact and faithful representations of the raw data.

*Automatic methods for FI* widely exist in many research fields and for many data types. The strengths of automatic FI methods are their generalizability, applicability, and robustness across domains, e.g., assessed through empirical evaluations and benchmark tests on descriptors [KK03], or downstream techniques [BLB\*17]



**Figure 1:** Typical inputs and outputs of the four steps in feature extraction [CMKK22] workflows. We focus on the FI for time series.

like clustering and classification. Also, automatic methods can efficiently operate on large datasets, work in a data-driven way, and can be included in data science workflows easily due to the availability of programming libraries. Major downsides of automatic methods are their varying applicability for different data characteristics, i.e., the method choice is a non-trivial problem for data scientists without domain knowledge at hand [Ber15].

*Human-based FI* is most useful when a real-world application requires features that shall include domain knowledge, reflect human mental models, consider special semantics, or cannot be ideated with standard automatic time series descriptors easily. Along the lines of human-based VA principles, domain experts can define individual features in an online process, while directly interacting with the raw data through visual interfaces. Example VA approaches directly supporting the ideation of individual features for text data include Flock [CB15] for the interactive FI from written text and FeatureInsight [BAL\*15] for interactive visual FI for text classification. For TS data, related interactive preprocessing and analysis approaches can help to ideate feature sets through the in-

teractive descriptor definition [BRG\*12], or statistical FI [ZCB11], even if individual features cannot be ideated through direct TS interaction. Interactive TS querying [HS04] and filtering [AS94] approaches enable experts to interact with TS directly, but trigger operations to reduce the number of data objects, rather than to ideate features. Finally, interactive segmentation and labeling approaches enable users to directly engage with TS [BBB\*18], even if FI tasks are not supported. In essence, VA approaches for FI through direct TS interaction are scarce. A possible downside of human-based FI is the risk of overfitting [WWH19] for a particular task and data subset [Ber15], and thus a lack of generalizability.

The research question that motivates our work is whether methods for the human-based FI for TS data would be useful and worth studying. As such, the study by Wu et al. [WWH19] is related, aiming at assessing the benefit of human-based feature selection for interactive machine learning. We study FI by comparing a human-based approach supported by VA with automatic methods borrowed from data mining, information retrieval, and machine learning. We start with a definition of requirements for human-based FI methods for TS, and the assessment of the applicability of state-of-the-art VA techniques. Next, we discuss a set of qualitative and quantitative characteristics of feature sets for TS data, used for operationalization purposes in our study of FI methods. Ultimately, we study the human-based FI for TS data with a state-of-the-art VA tool applied on a real-world case and compare results from human-based FI with results from automatic methods.

In summary, our contributions are:

- the assessment of the state-of-the-art in VA with respect to different requirements on human-entered FI,
- the characterization of FI methods and resulting feature sets for TS data along the lines of seven criteria,
- the comparative study of two human-based FI with six automatic methods for TS data based on a real-world case, and
- the discussion of and reflection on study results, implications of human-based FI, and a call for action for future VA approaches.

The running example used in this work is representative of industrial cases where machine runs are recorded with sensors, e.g., for predictive maintenance scenarios. Here, experts need to extract features to identify systematic changes in machine runs early, such as steeper progressions, shorter cycles, or varying areas under the curve in high-load situations.

## 2. Requirements for Human-Based Feature Ideation

Automated FI reduce human input in the process to defining input data and parameters. In turn, in human-based FI, experts can leverage their domain knowledge and ideate individual features manually, by directly using the raw data. To support human-based FI on TS data, VA applications need to fulfill certain requirements. We specifically concentrate on VA applications to reach out to domain experts, without the need to use programming tools for analysis.

**Support for Time Series Data** As a main requirement, VA applications need to be able to load and process TS data. Current commercial business analytics tools all support loading CSV files, and most of them also provide support for database connections [BSS\*18]. VA applications need to work with the specifics of TS data (e.g., tabular structure, temporal component).

**Time Series Data Preparation** Data preparation helps to make TS more usable and useful, and differs considerably from approaches for other data types such as tabular data. TS data preparation includes strategies for alignment (Figure 3), normalization [BRG\*12], or segmentation [BBB\*18]. Specific concentration on TS-specific requirements is necessary for VA applications [PTMB09], which is in contrast to many business analytics tools with a rather broad range of data types.

**Visual Representation of Raw Time Series Data** A prerequisite for FI is the ability of tools to represent raw TS data visually, as a basis for the FI. Quality issues of raw data may impede this challenge [Ber15], as data may not have been preprocessed yet, requiring a certain degree of robustness for dirty TS data. One limitation of current commercial business analytics tools is the size of datasets that can be loaded [BSS\*18].

**Generalizable Feature Ideation on Multiples** Tools should be able to highlight patterns in the TS data, i.e., non-singular and non-randomly occurring phenomena. Patterns are particularly worth to be considered for FI, as they are representative of subsets of the data that can be generalized. While many commercial tools do not provide means for FI [BSS\*18], inspiration can be drawn from VA approaches showing aligned and bundled (superimposed) TS [HS04], or utilizing visual clustering to reveal patterns in ts [SKB\*18].

**Interactive Feature Ideation** In contrast to automatic general-purpose methods for FI, most useful human-based features directly result from data inspection and interactive feature definition. Strong motivation comes from interactive TS sketching and brushing techniques that let users directly interact with TS data, including dynamic queries [AS94], timebox widgets [HS04], angular queries [HS04], and free-form sketching [BDF\*15], even if designed for search and filtering operations, rather than for FI.

**Formalization of User Interactions** Building upon the principles of explicit feedback [FWR\*17] and semantic interaction [EFN12], user interactions applied to TS should be interpreted appropriately and be transformed into formal specifications. Specifications of features ensure their re-usability in upcoming workflow steps, including feature engineering.

## 3. Characterization of Time Series Feature Sets

FI methods typically rely on assumptions about the underlying data, to reveal specific aspects of TS, such as value statistics, temporal aggregations, or shape descriptions. Following the No-Free-Lunch Theorem [WM97], we argue that also no best FI method exists. In this section, we outline important measures that help to reveal main characteristics of feature sets for TS. Beyond traditional metrics used for extensive performance studies of automatic FI methods, the scope of our characterization of FI is broader: Our goal is to also take ideation challenges, the discriminativeness of features, and human-based aspects into account. We will make use of the characterization in our study on human-based and automatic FI methods in Section 4. Overall, we describe seven characteristics that allow a more systematic identification of commonalities and differences across feature sets:

1. **Parameter Sensitivity:** Most FI methods require at least one parameter, and typically parameter values strongly influence the

FI performance and accuracy [BAL\*15]. The *ideation parameterization* characteristic defines the complexity of the parameterization task (i.e., how many parameters) of the FI method.

2. **Computational Scalability:** The computation of the features varies in performance for different methods to be used [OFC21]. The *computational scalability* characteristic defines the time needed to compute ideated features for the underlying TS data.
3. **Redundancy:** Highly correlated features influence the quality of the accuracy of ongoing analysis steps like classification [TL11]. The feature *redundancy* characteristic defines the strength of the correlation between features.
4. **Size:** Depending on the method used, features are defined by vectors or mathematical constructs of different complexity [JSLH22]. The *size* characteristic defines the number of dimensions of the resulting feature vector.
5. **Retrieval Performance:** Features can be used for querying data and content-based retrieval [SCG09]. With the *retrieval performance* characteristic, we take the quality of features for search and retrieval tasks into account.
6. **Classification Performance:** Features strongly influence the outcome of supervised machine learning tasks like classification [Mwa16]. The *classification performance* characteristic assesses the usefulness of feature sets for classification tasks.
7. **Self-Explainability:** Explainability of features is important to users [ZvZCH22] since they need to understand the outcomes of the subsequent analysis processes (e.g., classification). The *self-explainability* characteristic defines the degree to which features are semantically interpretable [BZSA18].

#### 4. Comparative Study of Feature Ideation Methods

We conducted a study to compare differences between features derived from automated FI methods and features derived from human-based ideation supported by VA.

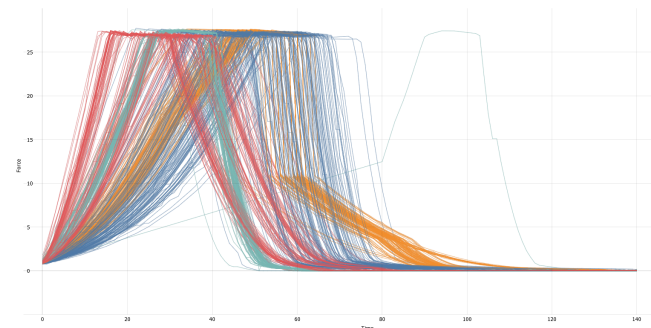
##### 4.1. Study Design

Feature ideation from TS is a common problem in Industry 4.0 applications, where digitization fostered the installation of different sensors in the industrial production process [CSA22]. Analyzing these sensor data is seen as a critical challenge to understand production processes better and detect possible opportunities for improvements. VA techniques are seen as an important key factor for understanding Industry 4.0 data, and we extend this notion for the ideation of features. We, therefore, based our study on a real-world use-case from the domain of Industry 4.0.

**Data** The TS data we use to demonstrate our analysis pipeline stems from a production process of fireproof bricks. Along the production line, bricks must pass through several stations, including composition, pressing, and firing. We recorded 526 pressing processes, producing three different types of stones in four different factories. An illustration of the pressing processes can be seen in Figure 2. A process is characterized by force building up (steep increase at the beginning), a plateau (remaining at maximum pressure for some time), and force dismantling (decrease). Every pressing process takes a maximum of 140 seconds, with a quantization of 1Hz (140 time stamps per TS).

**Expert Users** For the study, we collaborated with process engineers who want to understand their manufacturing processes better. After an initial data analysis step, where we visualized the raw TS

data (Figure 2), the domain experts confirmed that engineers at different factories operate presses in different ways. Further, the domain experts were interested in the retrieval of similar procedures and TS classification to predict the factory of brick production.



**Figure 2:** Pressing processes. 526 time series of an industrial pressing processes are shown, with the x-axis encoding time (from 0 to 140 seconds passed) and the y-axis encoding force. TS are colored according to the four factories they have been executed at.

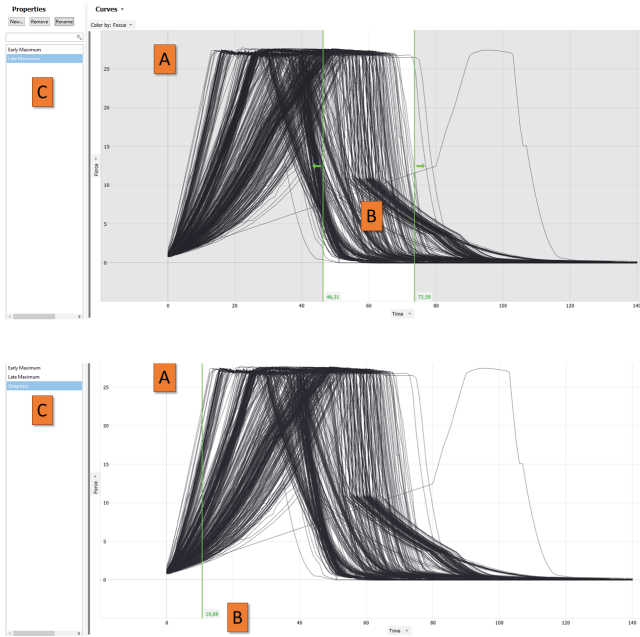
##### 4.2. Independent and Dependent Variables

Overall, we compared eight different FI methods, with a mix of six automated and two human-based methods using a VA application. The independent variables are 1) the *ideator* (automated vs. human-based) and nested, 2) the *ideation method*. To study commonalities and differences between the FI methods, we employ the *characterization of TS feature sets* as dependent variable.

**Automatic Methods** A short list of TS descriptor methods for FI includes Piecewise Aggregate Approximation (PAA) [KCPM01], Perceptual Important Points (PIP) [ZJGK10], Piecewise Linear Approximation (PLA) [KCHP01], Discrete Wavelet Transform (DTW) [Raf99], the Discrete Fourier Transform (DFT) [AFS93], statistical methods [CBNKL18], or deep learning methods that ideate features fully-automatically [NMK\*19] from TS data. We decide for **Raw Data** (140 samples), **PAA**, (25 segments) **PIP** (7 peaks), **DFT**, **Statistics** (mean, area under the curve, auto-correlation), and **Deep**-learned features, for the study. All of them are prominent representatives and available through libraries, while offering a high variety in ideation strategies.

**Human-Based Methods** We provided domain experts with a VA interface where they could create the features by themselves. We asked two domain experts to independently extract important features from the raw data using interactive visual methods. We used the software tool *Visplore* as an exemplary VA application specifically dedicated to TS analysis, has already been used to derive key performance indicators, and meets most requirements discussed in Section 2. It is, e.g., possible to load and visualize raw TS data from files or from a database, and apply data preparation. Different interaction methods (selections, thresholds, angular queries) are provided to ideate features, Figure 3 shows two steps in the FI process. Overall, domain experts ideated three (**HUM3**) and four **HUM4** features from the data, representing two feature sets. The first two features both domain experts selected can be described as the place on the timeline where the TS reach the maximum. Domain experts can use a window interaction (Figure 3, top) to define the interval on the timeline. The third feature in HUM3 was a gradient calculation at the beginning of the procedure (Figure 3, bottom). In





**Figure 3:** Human-based FI, observed in the study. The VA tool Vis-plore was used to visualize the raw data (A). Domain experts can interactively define features for TS (B) using, e.g., area selections (top) or angular queries (bottom). Ideated features are listed on the left (C), also allowing for manipulation and parameterization.

HUM4, the domain expert also added an area selection to capture the different shapes of the TS at the moment of force dismantling.

**Operationalization of Feature Set Characteristics** We evaluated our feature set characteristics (cf. Section 3) in the following way:

1. **Parameter Sensitivity** is calculated as a combination of the number and the sensitivity of the parameters on the outcome (i.e., the necessity to try different parameterizations). For human-based methods, we recorded mouse and keyboard interactions and the need to iterate and refine features.
2. **Computational Scalability** evaluates the runtime of the algorithms. We ran every FI method 1,000 times for the entire dataset, to have a higher comparison precision. Human-based methods were executed using the same procedure.
3. **Redundancy** was measured by computing the arithmetic mean value of all the pairwise correlations between all features.
4. **Size** was defined by taking feature dimensionality into account.
5. **Retrieval Performance** was evaluated by computing the error of k-nearest neighbors, with all TS. We set  $k$  to 5 for all test cases and measured the similarity of the returned TS based on the factory (origin) of the produced brick.
6. **Classification Performance** was measured with the FI score applied to a trained random forest classifier, with ground truth labels representing the four factories (origins) of the TS.
7. **Self-Explainability** was measured based on a coding conducted by the authors, on the agreement on semantic interpretability.

### 4.3. Result Analysis and Discussion

We evaluated all FI methods according to the seven characteristics of feature sets (cf. Section 4.1), Table 1 provides a comparative

overview. We decided for a discretization of quantitative results to a qualitative display, to a) have a common scheme for all seven characteristics and b) provide more robust results, given that we applied the study on a single fixed dataset only.

**Comparison of Human-Based and Automatic Methods** The overall results show that human-based methods provide comparable performance with automatic methods for most criteria, by the price of human involvement. We identified possible drawbacks of human-based FI for parameter sensibility. Interestingly, the particular strengths of human-based methods differ from the strengths of automatic methods, opening the space for future studies on the applicability of human-based methods for different settings.

**Findings for Human-Based Methods** A striking advantage of human-based methods was the particularly low feature size (three vs. four), with hardly any redundancy, significantly better than all automatic methods. In addition, human-based FI carry a high degree of self-explainability, as these features have been defined by domain experts externalizing their knowledge. However, these FI can offer a high degree of freedom in terms of parameterization.

**Findings for Automatic Methods** What stood out for the analyzed automated FI methods was their high heterogeneity, with not a single pair of methods showing highly similar results. This echoes the No-Free-Lunch Theorem [WM97], also for the FI in TS. Obviously, choosing a meaningful automatic method for a given dataset is a challenging task for data scientists, especially if only little knowledge about the domain exists. More specifically, setting parameters for automated methods (*Parameter sensitivity*) requires careful tweaking for PAA and PIP, since here the number of segments (PAA) and peaks (PIP) influence the structure of the resulting features. *Computational scalability* reveals very low computation times for the majority of the methods between 0.46s (PAA) and 2.45s (PIP). *Redundancy* is higher for raw data (0.71) and PAA (0.72), and even worse for PIP (0.97). For PAA, PIP, and DFT, the feature *size* depends on the selection of parameters. Here, PIP shows the most compact feature representation. *Retrieval performance* is high for the majority of methods (85.3% on average), with PAA, PIP, and Deep features outperforming this average. *Classification* worked very well when using the raw data (97.5%) and PAA (97.6%), followed by DFT, Statistics, and Deep features. *Self-explainability* was considered to be reasonable for Raw Data and PAA, but very low for DFT and Deep learned features.

### 4.4. Discussion

**Generalization of Human-Based Study Results** Our study was based on a single dataset only. With our decision to present results in a qualitative way, we already took this limitation into account. Future work includes the extension of our study with other datasets and users, to make more generalizable claims, possibly supported with a more quantitative result interpretation.

**Segmentation and Alignment** We have studied a case with almost perfectly segmented and aligned TS. Both experts stated that the segmentation and alignment of TS to reveal patterns relevant for FI often relies on expert knowledge and is not trivial to achieve. In cases where phases in TS curves are often captured or pre-labeled in the data, it is up to humans to find a meaningful specification, as part of a more complex FI process.

	Raw Data automatic	PAA automatic	PIP automatic	DFT automatic	Statistics automatic	Deep automatic	HUM3 human	HUM4 human
1. Parameter Sensitivity	⊘	○	○	+	+	○	-	-
2. Computational Scalability	⊘	+	○	+	+	○	○	○
3. Redundancy	○	○	-	+	○	+	+	+
4. Size	-	○	+	○	○	○	+	+
5. Retrieval Performance	○	+	+	○	○	+	○	+
6. Classification Performance	+	+	○	+	+	+	○	+
7. Self-Explainability	+	+	○	-	○	-	+	+

**Table 1:** Overview of study results as a crosscut between seven feature characteristics (cf. Section 3) and eight FI methods (cf. Section 4.1). We show results in a qualitative strong (+) to average (○) to weak (-) scale. Striking findings are as follows: human-based FI were able to compete with automatic methods for five characteristics. Also, the two human-based FI results are quite homogeneous, but dissimilar to automatic methods, raising an interesting question: when to utilize the benefits of human-based FI? In contrast, the six automatic methods show high result variety: none of the methods is particularly weak, but applicable for at least one characteristic instead.

**Temporal Flexibility** Our case indicated the need for warping support, always when temporal phenomena can be of different lengths, i.e., when bin-to-bin comparisons fail. The situation becomes worse if phenomena consist of several temporal observations that may vary in duration and lengths of breaks. In cases where temporal flexibility is key for effective FI, support through visualization and interaction becomes more difficult.

**Feature Hierarchies and Semantic Meaning** Especially in business cases and industrial settings, key performance indicators (KPIs) play an important role, e.g., in predictive maintenance situations, similar to our case. An interesting observation is that the relation between KPIs and features may be 1 : 1 for simplistic or 1 :  $n$  for more complex cases. The latter opens the space for feature hierarchies and higher-level semantics of feature groups or KPIs.

**Pre-Processing-Feature Ideation Ping-Pong** Both experts stated that data quality is most often a critical issue in their data science processes. It is likely and imaginable that, observations made through human-based FI, would lead to more informed decisions with respect to pre-processing and vice versa, triggering an interesting back-and-forth scenario.

**Exploratory Feature Ideation** In the study, we observed that both experts had a very clear information need, and thus were able to formalize their knowledge through interaction effectively. In contrast, an open point of discussion would be a more exploratory setting, where experts require a means to identify patterns, experiment with feature ideation in a back-and-forth manner, to finally arrive at an ideation result. A conceptual workflow would include: Exploratory FI, Informed FI, and Automatable FI.

**Automatic Feature Ideation** The related work, but also the feedback of the two domain experts clearly echo that not every real-world case requires human-based FI. Given the cost of having the human in the loop in this step, an interesting point for future work may include support for deciding on whether the quality of automatic FI is sufficient, or if the real-world case would considerably benefit from a human-based FI component.

## 5. Conclusion

We presented a study on FI with automated and human-based methods. We started with discussing requirements for VA applications

on human-based FI and characterized feature sets based on seven criteria. This characterization also served as the set of dependent variables to compare ideated feature sets. Based on a use-case from Industry 4.0, we extracted 526 TS and employed six automated and two human-based FI. Our study shows that both human-based and automatic FI methods have proven their right to exist. The strengths of automatic methods lies in their generalizability, applicability, and robustness across cases. The strength of human-based methods lies in the creation of features with a low size but with a very high self-explainability, by also leveraging domain knowledge. Human-based FI led to equally good and, for more features, even better results for retrieval and classification tasks for the studied case. In conclusion, we see a clear call to action for conducting research on VA techniques for FI, and for applying human-based FI in the wild. We consider VA as a suitable method to support a close collaboration between humans and automated algorithms. For future work, we would like to work on novel VA tools which support interactive feature ideation and compare the results to automatic methods.

## 6. Acknowledgments

VRVis is funded by BMK, BMAW, Styria, SFG, Tyrol and Vienna Business Agency in the scope of COMET - Competence Centers for Excellent Technologies (879730) which is managed by FFG.

## References

- [AFS93] AGRAWAL R., FALOUTSOS C., SWAMI A. N.: Efficient Similarity Search In Sequence Databases. In *Foundations of Data Organization and Algorithms* (1993), vol. 730, Springer, pp. 69–84. doi:10.1007/3-540-57301-1\_5. 3
- [AS94] AHLBERG C., SHNEIDERMAN B.: Visual Information Seeking: Tight Coupling of Dynamic Query Filters with Starfield Displays. In *SIGCHI Conference on Human Factors in Computing Systems* (1994), ACM, pp. 313—317. doi:10.1145/191666.191775. 2
- [BAL\*15] BROOKS M., AMERSHI S., LEE B., DRUCKER S. M., KAPOOR A., SIMARD P.: FeatureInsight: Visual support for error-driven feature ideation in text classification. In *IEEE Conference on Visual Analytics Science and Technology* (2015), pp. 105–112. doi:10.1109/VAST.2015.7347637. 1, 3
- [BBB\*18] BERNARD J., BORS C., BÖGL M., EICHNER C., GSCHWANDTNER T., MIKSCH S., SCHUMANN H., KOHLHAMMER J.: Combining the Automated Segmentation and Visual Analysis of Multivariate Time Series. In *EuroVis Workshop on Visual Analytics* (2018), Eurographics. doi:10.2312/eurova.20181112. 2

- [BDF\*15] BERNARD J., DABERKOW D., FELLNER D., FISCHER K., KOEPLER O., KOHLHAMMER J., RUNNWERTH M., RUPPERT T., SCHRECK T., SENS I.: VisInfo: a digital library system for time series research data based on exploratory search - a user-centered design approach. *International Journal on Digital Libraries* 16, 1 (2015), 37–59. doi:10.1007/s00799-014-0134-y. 2
- [Ber15] BERNARD J.: Exploratory search in time-oriented primary data. *Dissertation, PhD Thesis (Dissertation), Technische Universität Darmstadt* (2015). URL: <http://tuprints.ulb.tu-darmstadt.de/5173/>. 1, 2
- [BLB\*17] BAGNALL A., LINES J., BOSTROM A., LARGE J., KEOGH E.: The great time series classification bake off: a review and experimental evaluation of recent algorithmic advances. *Data Mining and Knowledge Discovery* 31 (2017), 606–660. doi:10.1007/s10618-016-0483-9. 1
- [BRG\*12] BERNARD J., RUPPERT T., GOROLL O., MAY T., KOHLHAMMER J.: Visual-Interactive Preprocessing of Time Series Data. In *SIGRAD, Swedish Chapter of Eurographics* (2012), vol. 81, Linköping University Electronic Press, pp. 39–48. doi:10.1111/cgf.13698. 2
- [BSS\*18] BEHRISCH M., STREEB D., STOFFEL F., SEEBACHER D., MATEJEK B., WEBER S. H., MITTELSTAEDT S., PFISTER H., KEIM D. A.: Commercial Visual Analytics Systems-Advances in the Big Data Analytics Field. *IEEE Transactions on Visualization and Computer Graphics* 25, 10 (2018), 3011–3031. doi:10.1109/TVCG.2018.2859973. 2
- [BZSA18] BERNARD J., ZEPPELZAUER M., SEDLMAIR M., AIGNER W.: VIAL: a unified process for visual interactive labeling. *The Visual Computer (TVCJ)* 34, 9 (2018), 1189–1207. doi:10.1007/s00371-018-1500-3. 3
- [CB15] CHENG J., BERNSTEIN M. S.: Flock: Hybrid Crowd-Machine Learning Classifiers. In *ACM Conference on Computer Supported Cooperative Work & Social Computing* (2015), pp. 600–611. doi:10.1145/2675133.2675214. 1
- [CBNKL18] CHRIST M., BRAUN N., NEUFFER J., KEMPA-LIEHR A. W.: Time Series Feature Extraction on basis of Scalable Hypothesis tests (tsfresh – A Python package). *Neurocomputing* 307 (2018), 72–77. doi:10.1016/j.neucom.2018.03.067. 3
- [CMKK22] CHATZIMPAMPAS A., MARTINS R. M., KUCHER K., KERREN A.: FeatureEnVi: Visual Analytics for Feature Engineering Using Stepwise Selection and Semi-Automatic Extraction Approaches. *IEEE Transactions on Visualization and Computer Graphics* 28, 4 (2022), 1773–1791. doi:10.1109/TVCG.2022.3141040. 1
- [CSA22] CIBULSKI L., SCHMIDT J., AIGNER W.: Reflections on Visualization Research Projects in the Manufacturing Industry. *IEEE Computer Graphics and Applications* 42, 2 (2022), 21–32. doi:10.1109/MCG.2022.3156846. 3
- [EFN12] ENDERT A., FIAUX P., NORTH C.: Semantic Interaction for Sensemaking: Inferring Analytical Reasoning for Model Steering. *IEEE Transactions on Visualization and Computer Graphics* 18, 12 (2012), 2879–2888. doi:10.1109/TVCG.2012.260. 2
- [FWR\*17] FEDERICO P., WAGNER M., RIND A., AMOR-AMOROS A., MIKSCH S., AIGNER W.: The Role of Explicit Knowledge: A Conceptual Model of Knowledge-Assisted Visual Analytics. In *IEEE Conference on Visual Analytics Science and Technology* (2017), pp. 92–103. doi:10.1109/VAST.2017.8585498. 2
- [HS04] HOCHHEISER H., SHNEIDERMAN B.: Dynamic query tools for time series data sets: timebox widgets for interactive exploration. *Information Visualization* 3, 1 (2004), 1–18. doi:10.1057/palgrave.ivs.9500061. 2
- [JSLH22] JIA W., SUN M., LIAN J., HOU S.: Feature dimensionality reduction: a review. *Complex & Intelligent Systems* 8 (2022), 2663–2693. doi:10.1007/s40747-021-00637-x. 3
- [KCHP01] KEOGH E. J., CHU S., HART D. M., PAZZANI M. J.: An Online Algorithm for Segmenting Time Series. In *IEEE International Conference on Data Mining* (2001), IEEE, pp. 289–296. doi:10.1109/ICDM.2001.989531. 3
- [KCPM01] KEOGH E. J., CHAKRABARTI K., PAZZANI M. J., MEHROTRA S.: Dimensionality Reduction for Fast Similarity Search in Large Time Series Databases. *Knowledge and Information Systems* 3, 3 (2001), 263–286. doi:10.1007/PL00011669. 3
- [KK03] KEOGH E. J., KASETTY S.: On the Need for Time Series Data Mining Benchmarks: A Survey and Empirical Demonstration. *Data Mining and Knowledge Discovery* 7, 4 (2003), 349–371. doi:10.1023/A:1024988512476. 1
- [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 (2014), 1614–1623. doi:10.1109/TVCG.2014.2346482. 1
- [Mwa16] MWADULO M. W.: A Review on Feature Selection Methods For Classification Tasks. *International Journal of Computer Applications Technology and Research* 5, 6 (2016), 395–402. doi:10.7753/IJCATR0506.1013. 3
- [NMK\*19] NTAKARIS A., MIRONE G., KANNIAINEN J., GABBOUJ M., IOSIFIDIS A.: Feature Engineering for Mid-Price Prediction With Deep Learning. *IEEE Access* 7 (2019), 82390–82412. doi:10.1109/ACCESS.2019.2924353. 3
- [OFC21] OLIVEIRA A. J., FERREIRA B. M., CRUZ N. A.: A Performance Analysis of Feature Extraction Algorithms for Acoustic Image-Based Underwater Navigation. *Journal of Marine Science and Engineering* 9, 4 (2021), 361. doi:10.3390/jmse9040361. 3
- [PTMB09] PIRINGER H., TOMINSKI C., MUIGG P., BERGER W.: A Multi-Threading Architecture to Support Interactive Visual Exploration. *IEEE Transactions on Visualization and Computer Graphics* 15, 6 (2009), 1113–1120. doi:10.1109/TVCG.2009.110. 2
- [Raf99] RAFIEI D.: On Similarity-Based Queries for Time Series Data. In *Conference on Data Engineering* (1999), IEEE, pp. 410–417. doi:10.1109/ICDE.1999.754957. 3
- [SCG09] SCHETTINI R., CIOCCA G., GAGLIARDI I.: *Feature Extraction for Content-Based Image Retrieval*. Springer US, 2009, pp. 1115–1119. doi:10.1007/978-0-387-39940-9\_162. 3
- [SKB\*18] SACHA D., KRAUS M., BERNARD J., BEHRISCH M., SCHRECK T., ASANO Y., KEIM D. A.: SOMFlow: Guided Exploratory Cluster Analysis with Self-Organizing Maps and Analytic Provenance. *IEEE Trans. on Visualization and Computer Graphics* 24, 1 (2018), 120–130. doi:10.1109/TVCG.2017.2744805. 1, 2
- [TL11] TOLOŞI L., LENGAUER T.: Classification with correlated features: unreliability of feature ranking and solutions. *Bioinformatics* 27, 14 (2011), 1986–1994. doi:10.1093/bioinformatics/btr300. 3
- [WM97] WOLPERT D. H., MACREADY W. G.: No free lunch theorems for optimization. *IEEE Transactions on Evolutionary Computation* 1, 1 (1997), 67–82. doi:10.1109/4235.585893. 2, 4
- [WWH19] WU T., WELD D. S., HEER J.: Local Decision Pitfalls in Interactive Machine Learning: An Investigation into Feature Selection in Sentiment Analysis. *ACM Transactions on Computer-Human Interaction* 26, 4 (2019), 24. doi:10.1145/3319616. 2
- [ZCB11] ZHAO J., CHEVALIER F., BALAKRISHNAN R.: KronoMiner: using multi-foci navigation for the visual exploration of time-series data. In *Conference on Human Factors in Computing Systems* (2011), ACM, pp. 1737–1746. doi:10.1145/1978942.1979195. 2
- [ZJGK10] ZIEGLER H., JENNY M., GRUSE T., KEIM D. A.: Visual market sector analysis for financial time series data. In *IEEE Conference on Visual Analytics Science and Technology* (2010), pp. 83–90. doi:10.1109/VAST.2010.5652530. 3
- [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 *IEEE Visualization Conference (VIS), Short Papers* (2019), pp. 161–165. doi:10.1109/VISUAL.2019.8933619. 1
- [ZvZCH22] ZACHARIAS J., VON ZAHN M., CHEN J., HINZ O.: Designing a feature selection method based on explainable artificial intelligence. *Electronic Markets* 32, 4 (2022), 2159–2184. doi:10.1007/s12525-022-00608-1. 3