

VIII

Closing

Final Discussion

This chapter closes this thesis. *Section 1.1* provides evidence on how the developed work validates the thesis statement. *Section 1.2* summarizes the contributions and discusses the major contributions of our work.

1.1 Hypothesis Validation

Based on the extensive theoretical and empirical observations collected throughout this thesis, this section aims to test the underlying hypothesis: whether or not focusing the learning on relevant regions improves the performance guarantees of descriptive and associative classification models in high-dimensional data contexts. For this aim, consider the following views on the contributions of our work:

- principles to satisfy the set of formulated requirements. Furthermore, we provided a solution space with integrative and easily parameterizable algorithms guaranteeing that, for a given data context, multiple requirements can be simultaneously satisfied. Tables provided throughout the overview chapters provide an exhaustive list of (non-contradictory) contributions to satisfy the specified requirements to answer the thesis statement;
- empirical evidence showing the superior performance of the proposed algorithms against peer algorithms. We showed the superiority of the proposed biclustering and classification algorithms (*Books III and VI*) against state-of-the-art peers with regards to a wide-variety of performance views, as well as the superiority of learning functions for structured data (*Book IV*) and structured codomains (*Chapter VI-7*) against peer learning functions adapted from related streams of research. We also provide empirical evidence for the adequacy and robustness of the proposed (statistical) assessment methodologies (*Books V and II*);
- application of the proposed learning algorithms over real data from distinct domains was demonstrated to be associated with the discovery of new and meaningful regions. We demonstrated the relevance of: 1) learning of plaid, order-preserving and cascade models from expression and biological network data to retrieve non-trivial yet informative (biologically significant) regulatory modules and responses; 2) discovering arrangements of events from multi-dimensional healthcare (heritage) and commercial (foodmart) databases; and 3) modeling discriminative regions with varying coherency strength, quality and statistical significance across biomedical and social data domains.

In sum, we show that the proposed learning methods satisfy the introduced requirements, outperform state-of-the-art methods and unravel unique putative relations of interest from real data. In this context, we believe that these groups of arguments provide strong evidence in favor of the *approval of the thesis statement*. As a result, we expect an increasing attention given to this field of research and a broader systemic view on the learning conditions and data properties where the thesis statement is verified and (possibly) rejected.

1.2 Thesis Contributions and Implications

Throughout Books II to VI, we addressed different challenges associated with the learning in high-dimensional data contexts. The conducted research gave rise to the following contributions.

Performance Guarantees of Models Learned from High-Dimensional Data

The first part of the thesis established a solid foundation for the assessment of descriptive and classification models in high-dimensional data contexts (contributions listed in Tables II-1 to II-3). First, we proposed robust estimators to bound and compare the performance of classification models, showing how existing estimators can be enhanced to: adequately measure the generalization error, accommodate smoothing factors in the presence of probabilistic outputs, and minimize the problems associated with the inference of performance guarantees from unfeasible error estimates. Second, we extended this assessment towards local descriptive models, showing how error estimates can be collected to infer robust guarantees of performance and parameterized with adequate loss functions. In this context, we overviewed the properties of existing loss functions for descriptors of tabular data (biclustering models) and structured data (triclustering and cascade models), and proposed new loss functions able to provide relevant performance views from synthetic and real data. Finally, to further guarantee non-biased and complete assessments of descriptive and classification models, we proposed data generators for (labeled) tabular and structured data based on the different forms of local and global regularities commonly observed in real-world data.

Implications

These contributions open a new door for robust, unbiased and complete assessments of state-of-the-art descriptive and classification methods. They are also critical to: 1) gain an in-depth understanding of the behavior of these methods in high-dimensional data contexts, 2) unravel their strengths and weakness, and 3) (possibly) guide their improvement. The guarantees of performance inferred under the proposed methodology can be used to weight and validate the high number of implications derived from the analysis of models learned from real data, and to measure the impact of applying procedures for dimensionality reduction prior to learning.

Learning Local Descriptive Models from Tabular Data

The second part of our thesis has contributions (enumerated in Tables III-1 to III-10) to the learning of robust biclustering models with flexible structures, coherency and quality from (high-dimensional) tabular data. Towards this end, we first explored the synergies between biclustering and pattern mining by studying the impact of using distinct patterns (including itemsets, association rules, formal concepts and sequences), dedicated (anti-)monotonic searches, and multiple preprocessing and postprocessing procedures on the properties of biclustering models.

Second, we extended these contributions to learn biclustering models with different forms of coherency commonly observed in biomedical and social data domains, including: additive and multiplicative models (by interpreting shifts and common multiples in data); plaid models (by identifying meaningful interactions between biclusters); order-preserving models (by discovering noise-tolerant orderings of values); and the previous models in the presence of symmetries. In particular, we proposed new algorithms to learn such non-constant models and extended them to deal with non-trivial plaid effects and be able to incorporate meaningful relaxations.

Implications

The learning of flexible biclustering models opens many possibilities for the analysis of biomedical and social data. The discovery of a flexible structures is essential to guarantee an unbiased exploration of the search space. The discovery of additive and multiplicative coherencies is critical to analyze biomedical data with structural differences on the responsiveness associated with gene expression, molecular concentrations or physiological responses. In social applications, these coherencies are useful to model social interactions with non-trivial yet coherent behavior and to group subjects with identical variation of preferences for browsing, (e-)commercing and collaborative filtering. The accommodation of symmetries are used to capture distinct regulatory mechanisms in biological domains and opposed (yet correlated) trading, tweeting and browsing activity.

The possibility to learn plaid models using meaningful relaxations is critical to detect and interpret meaningful interactions between biological processes, clinical responses and social behavior. Similarly, the proposed composition functions also allow the study of interactions between biological processes and social groups from the point of view of their regulatory/behavioral activity, enabling the categorization of the interactions and providing insights to understand complex regulation/behavior.

Furthermore, the possibility to mine (strictly or monotonically increasing) orderings has been shown to be essential to solve different real-world problems, including the analysis of omic, chemical and mutagenesis data; the discovery of (relative) preferences from collaborative filtering data; the support to planning, scheduling and recommendation tasks; among others.

Finally, and for the first time, the proposed contributions also enable the sound application of biclustering over real-world tabular data characterized by a mixture of numeric, ordinal and categoric features.

Third, we revised the previous algorithms to guarantee their robustness to noise. In this context, we extended them with: multi-item assignments in order to guarantee that they are not exposed to discretization problems; principles along the preprocessing, mining and postprocessing steps to guarantee their robustness to varying forms and amount of noise; and multi-value imputations to handle an arbitrary-high number of missings.

Implications

These contributions provide a mark on how to effectively address the item-boundaries problem of discretization procedures by learning from data elements with an arbitrary number of items assigned. Results further show that under adequate searches this option does not visibly hamper the computational complexity of the task. As such, this option seizes the benefits of discretizing data without apparent downsides, thus opening critical considerations for future research.

Learning from data elements with multiple values is also valuable to deal with noise and provide risk-free imputations. Furthermore, its combination with the remaining principles to guarantee robustness appears to be essential to learn from data with an arbitrary-high number of missings (such as often found in biomedical domains) and for the analysis of real-world data with varying forms and amount of noise.

Fourth, we proposed new algorithms to guarantee the efficiency of the biclustering task for dense and high-dimensional data, including new pattern-growth searches for frequent itemset mining (based on annotated frequent pattern trees less prone to memory and time bottlenecks) and sequential pattern mining (based on efficient data projections and item-indexable properties to mine both monotonically and strictly increasing orders). We further guarantee their scalability in the presence of approximate searches and data partitioning principles from pattern mining. Finally, we also guarantee the efficiency of merging, extension and reduction procedures for postprocessing by placing anti-monotonic heuristics and pushing their computation to the mining step.

Implications

By integrating the well-studied principles to guarantee the scalability of pattern mining searches, the computationally complex task of learning flexible biclustering models from large-scale data becomes tractable. In this context, the proposed contributions remain prepared to handle the increasing size and dimensionality of real-world data.

Furthermore, the contributions for the efficient postprocessing of biclustering solutions can be transversally applied to manipulate voluminous outputs in the context of a wide-variety of research streams. As part of these contributions, some of the proposed mappings to solve postprocessing tasks, such as the possibility to merge an arbitrary-high number regions by relying on multi-support pattern mining tasks, open new avenues to guarantee the scalability of these procedures.

Fifth, we extended the proposed algorithms to learn from large-scale network data. For this aim, we revised the underlying data structures and searches, and further investigate the role of pattern-based biclustering to discover modules with non-dense yet meaningful and coherent interactions, as well as to guarantee their robustness to noisy and missing interactions. We showed the applicability towards homogeneous and heterogeneous networks with either quantitative/weighted or qualitative/labeled interactions. Finally, we generalized these principles to learn from sparse data, characterized by the presence of an arbitrary number of uninformative elements and/or missings.

Implications

These contributions provide the unprecedented opportunity to efficiently discover non-trivial (yet meaningful, coherent and significant) modules from network data. The discovery of these modules is essential to characterize, discriminate and/or predict biological functions and social activity. In particular, we expect to entail broader biological analysis to further establish relationships between modules and biological functions, as non-dense models might give clues about the organization of genes, proteins and metabolites, and support the characterization of molecular entities with yet unclear roles. The proposed contributions were also shown to be relevant to identify non-trivial communities from social networks where individual may have different degree of activity/involvement. Furthermore, they can also be applicable for the analysis of network traffic monitors, neural networks, structured financial transactions and coauthorship/citation networks.

The proposed principles to tolerate missing and noisy interactions within the discovered modules can be further used to predict unknown interactions and to test the confidence of the existing interactions. Also, the plaid model can be used to explore the network structure and identify biological/social hubs based on the overlapping interactions between modules.

Finally, we proposed principles for biclustering tabular data in the presence of background knowledge. For this end, we revised the proposed pattern-growth searches in order to explore efficiency gains from succinct, (anti-)monotonic, convertible and prefix-monotone constraints and to accommodate annotations extracted from knowledge bases and literature, and motivated their relevance across biological and social data domains.

Implications

The incorporation of background knowledge within (pattern-based) biclustering through declarative constraints is essential to orient the task according to user expectations and other available sources of knowledges. In particular, these contributions provide the unprecedented possibility to remove uninformative regions and to focus the search on (non-trivial) regions of interest. As such, we expect to see a systematization of constraints with relevance for the analysis of biomedical and social data. Furthermore, the possibility to annotate rows and columns with terms from knowledge repositories, semantic sources and literature (such as biological functions from gene ontologies and well-studied networks) can valuably guide the learning.

We further showed that the described contributions can be soundly integrated and their synergies explored. For this end, we proposed the BicPAMS software with declarative, graphical and programmatic interfaces.

Learning Local Descriptive Models from Structured Data

The third part of this thesis has contributions (listed in Tables IV-1 to IV-4) to the learning of local descriptive models from (high-dimensional) structured data. First, we motivated and formalized the relevant task of learning local descriptive models from three-way time series given by cascades of modules. A new algorithm was proposed to discover flexible modules and their causal relation, and to handle arbitrary-high temporal misalignments between their supporting observations. We further extended the algorithm to guarantee its robustness to noise and to stochastic uncertainties associated with divergent paths, as well as to enhance its scalability based on dissimilarity guarantees, removal of uninformative elements and approximate searches under strict optimality guarantees.

Implications

The learning of cascade models from multivariate time series opens a new door to study regulatory responses to growth, development, drugs and disease progression in biomedical domains, as well as behavioral responses (associated with social interaction, web navigation, commercial activity, financial decisions) to specific events of interest in social domains. In this context, the proposed contributions provide the unprecedented opportunity to handle the inherent stochasticity and complexity of these responses and to model their structural and temporal kinetics.

Second, we proposed a new algorithm to learn local descriptive models given by arrangements of informative events from multi-dimensional databases and collections of records. In this context, we provided a structured view on the relevance of this task, and an adequate data mapping of complex data structures into multi-sets of events for a cohesive tackling of the target task. The proposed algorithm is able to discover sets of temporally-related events derived based on the discovery temporal patterns from multi-sets of events. Multiple principles were proposed to enhance its efficiency, to deal with the structural sparsity of multi-sets of events, to prevent that some arrangements jeopardize the learning, and to tolerate temporal misalignments and noise.

Implications

The proposed contributions open new opportunities to learn informative relations from complex data structures. This direction has been widely termed a hot topic in data mining due to the increasing availability, quality and high-dimensionality of large-scale structured data. Illustrating, corporative and administrative data from public and private sectors typically follow multi-dimensional or relational schema. Also, repositories of heterogeneous health-records, user actions or financial transactions are additional data structures not supported by the majority of existing learning algorithms. In this context, the provided principles to consistently map these databases as multi-sets of events and to adequately learn informative arrangements of events (combining temporal and cross-attribute dependencies) can be seen as a relevant mark in the machine learning community.

Finally, we proposed alternative stochastic methods to model cascades and arrangements of events from structured data. For this aim, we first provided a probabilistic view of itemset sequences by extending hidden Markov models with: new architectural components to model frequent orderings, principles for their effective initialization,

revised learning schema to surpass convergence problems, and traversal procedures to efficiently decode sequential patterns from the learned lattices. As a result, we proposed a stochastic sequential pattern miner tolerant to noise (yet robust to spurious background matches) and able to model dissimilar, lengthy and non-trivial patterns. Finally, we shown that this algorithm can be adequately extended to stochastically learn local descriptive models from three-way time series and multi-sets of events, and confront its benefits and limitations against deterministic peers. For this end, we customized some of the proposed architectures and enhanced them with the unprecedented possibility to disclose time frames associated with state emissions.

Implications

The stochastic modeling of temporal patterns from structured data is relevant to provide compact representations of commonly large outputs; offer a probabilistic and noise-tolerant view of patterns; seize efficiency gains when learning from dense data; support the query-driven decoding of regions with patterns of interest; and focus the search on dissimilar and arbitrary-large regions. Stochastic models are of particular relevance for real-world data with complex stochastic phenomena and without well-defined regions. The proposed contributions enable the stochastic learning of descriptive models from structured data with non-fixed multivariate order and high-dimensionality. In this context, we expect to witness the extension of these contributions (such as new architectural components more conducive to the learning of certain temporal patterns), as well as their generalization for alternative stochastic learning methods, including neural networks, stochastic grammars and dynamic Bayesian networks.

Significance Guarantees of Local Descriptive Models

The fourth part of the thesis has contributions (enumerated in Tables V-1 to V-4) to assess and guarantee the statistical significance of regions modeled from (high-dimensional) data. For this aim, we first proposed statistical tests to robustly assess the statistical significance biclusters and enhanced the proposed biclustering algorithms accordingly. In this context, we revised non-conservative corrections (with principles to efficiently guarantee a minimized risk of accepting non-significant biclusters and rejecting significant biclusters) to guarantee their deviation from expectations. Complementarily, we shown how either local or global expectations on the size of biclusters can be inferred to guide the biclustering task.

Second, we extended the proposed statistical assessment towards noisy biclusters with flexible coherency. For this aim, new statistical views were provided to test additive and multiplicative models (based on allowed shifts and greatest common divisors), plaid models (based on the removal of plaid effects), and order-preserving models (based on the allowed permutations). The properties of search space size for these coherency assumptions were revised, and dynamic programming principles were considered to avoid redundant computations. For the assessment of noisy biclusters, we provided new principles to retrieve the underlying pattern expectations, and further extended the statistical framework to integrate homogeneity and significance views by combining the probability of a given bicluster to deviate from expectations with the probability to have unexpectedly low levels of noise.

Third, we extended the proposed statistical assessment towards real-valued biclusters with continuous factors and (possibly unknown) coherency strength. In this context, we proposed non-biased estimators of their true significance with lower and upper bounds. Additionally, new statistical tests to assess biclusters with continuous shifts and scales were provided based on the integral of the product of slided and scaled density functions. We further guarantee the extensibility of the statistical tests in this book towards tabular data with non-identically distributed features.

Implications

Assessing the statistical significance of biclustering models is critical to filter, validate or weight the increasing number of implications in literature derived from the analysis of local regions from biomedical and social data. Furthermore, it is essential to evaluate and compare state-of-the-art biclustering and pattern mining algorithms with regards to the significance of their solutions, as well as to guide the biclustering tasks, promoting the significance of their outputs and efficiency of the searches.

The provided systemic analysis on how the properties of a given bicluster (support, length, pattern, coherency, quality) and the input data (size, dimensionality, regularities) affect statistical significance, can be further used to shape upcoming biclustering algorithms.

As largely motivated, guaranteeing the significance of biclusters is further relevant to minimize the propensity of a given learning function to over/underfit the observed data, ensuring the adequacy of the capacity term and thus its ability to generalize.

Finally, we provided methods to assess the statistical significance of local descriptive models from structured data. In this context, we provided statistical views to test the deviation of the support of a sequential pattern against expectations derived from null data, as well as alternative tests on probabilistic models (compliant with hidden Markov models) based on the weight unexpectedness and coverage of a sequential pattern. These assessments were extended towards temporal patterns, cascades and arrangements of events, and incorporated within the previously proposed deterministic and stochastic algorithms to guide the learning.

Implications

The proposed contributions are essential to assess and guarantee that regions in structured data contexts are not informative by chance, a prominent problem due to the typical high-dimensionality of structured data given by multivariate time series, multi-sets of events and multi-dimensional data.

Learning Effective Classifiers from Local Descriptive Models

The fifth part of this thesis has contributions (listed in Tables VI-2 to VI-6) to the learning of effective (associative) classifiers from high-dimensional data. Towards this end, we first proposed new associative classifiers with principles to guarantee the adequacy of their discovery, training and testing functions. To guarantee the discovery of relevant regions we proposed: a new weighted notion of support to adequately assess the discriminative power of noisy regions, an adequate exploration of the data space, and the inclusion of regions able to discriminate groups of classes by composing rules with disjunctions of labels. New integrative scores (able to combine the discriminative power, size, coherency and quality of a region) were proposed along with adequate composition criteria to guarantee the adequacy of training functions. Finally, for an effective testing of new observations, we proposed new relaxations on the matching criterion to prevent the scarcity of matched regions and a new calculus of class strength.

Second, to guarantee the adequate inference of decisions in data domains characterized by regions with varying homogeneity and quality, we extended the previous classifiers to guarantee the recovery of with varying coherency strength, coherency assumption and noise. A new penalization schema was proposed for non-constant regions based on their degree of flexibility to prevent that they jeopardize the learning, as well as new matching criteria to test observations against non-constant regions.

Implications

The proposed associative classifiers are essential to guarantee an adequate learning from high-dimensional data domains characterized by the presence of regions of interest (possibly) well-approximated by discriminative biclusters. In this context, they are particularly relevant to learn biological and clinical markers, classify user behavior, evaluate (web) contents, and support trading/administrative/commercial decisions. They are further proposed as viable alternatives to surpass the largely motivated problems associated with feature selection, dimensionality reduction and/or sparse priors.

The proposed contributions provide an unprecedented opportunity to address the major criticisms of existing associative classifiers, including: scarcity of matchings, inability to adequately score noisy regions, inadequate scoring in the presence of imbalanced data, biases towards small (non-significant) regions, inappropriate space exploration, absence of adequate dissimilarity guarantees between regions, and inability to model regions discriminating more than a single class. Furthermore, contrasting with the focus on constant regions given by discriminative patterns of existing associative classifiers, the proposed associative classifiers are able to retrieve the commonly observed non-constant (yet meaningful and coherent) regions underlying biomedical and social data.

These contributions further enable the systematic analysis on how the varying coherency, quality, size and discriminative power of the underlying regions of a high-dimensional data space affect the performance of classifiers. This understanding is essential to revise the behavior of state-of-the-art classifiers and place considerations on the design of new classifiers.

Third, we further extended the previous associative classifiers to guarantee their ability to: learn from sparse data by soundly removing true missings and uninformative elements and adequately interpreting false missings; learn from data with global and local regularities by robustly combining the (probabilistic) outputs of alternative

classifiers; learn from stochastic descriptors of tabular data where membership vectors can be used to promote an adequate space coverage and shape the scoring criteria; and learn from data in the presence of background knowledge by effectively incorporating annotations and constraints with nice properties.

Implications

These contributions provide the unprecedented opportunity to classify sparse data, network data and data with missing elements (associated with monitoring holes, default expectations and errors). Also, they introduce the possibility to remove non-interesting regions to guide the learning and explore efficiency gains. In biological data domains, uninformative elements are typically associated with non-differential regulation of genes or concentration of molecular entities. For other domains, uninformative elements may correspond to: entries with low-counts from text-based data, inconclusive ratings in collaborative filtering data, unprofitable decisions from trading data, or healthy evaluations from medical data.

Finally, the possibility to incorporate a large variety of constraints and knowledge-driven annotations provides the opportunity to flexibly guide the learning of (associative) classifiers.

Fourth, we proposed new associative classifiers able to effectively learn from structured data. For this aim, the previously proposed deterministic algorithms to discover regions from structured data were extended with discriminative criteria. From these regions, rules are inferred, ranked according to a new integrative score and composed within a navigable tree structure. During the testing phase, we proposed new matching criteria sensitive to both structural and temporal misalignments, and revised the class strength calculus accordingly. Principles for selective/decaying memory can be easily incorporated within these classifiers. Complementarily, stochastic classifiers were proposed by testing the likelihood of new observations to be described by class-conditional probabilistic models of a data partition. Finally, we show the applicability of these classifiers to learn from varying data structures, including multivariate time series, itemset sequences, multi-sets of events and multi-dimensional databases.

Implications

The proposed contributions are key for a wide-range of biomedical and social applications, including phenotype discrimination from gene expression time series; disease prediction from repositories of clinical events; classification of user behavior from collections of temporal snapshots of a social network; diagnosis from (multivariate) physiological signals; financial decision support from repositories of trading actions; marketing initiatives from (e-)commerce events, and web content organization from user actions.

Contrasting with existing classifiers, the proposed associative classifiers provide the possibility to learn from multi-sets of events, handling arbitrary levels of sparsity between events and integrating distinct event types (heterogeneous attributes). Similarly, the proposed classifiers are also critical to learn from regions of three-way time series given by discriminative cascades.

Finally, the provided comparison between deterministic and stochastic learners offer key principles for an adequate selection and parameterization of classifiers according to the end goal and the properties of the input data.

Fifth, we proposed a structured view on how to assess and shape the guarantees of statistical significance of classifiers learned from high-dimensional data. For this purpose, we first discussed the benefits from learning from statistically significant regions and extended the previously proposed statistical views of regions to also assess the significance of their discriminative power. This statistical view was used to revise the behavior of associative classifiers, decision trees and random forests, thus minimizing their propensity to underfit high-dimensional data.

We further extended these statistical views to assess the impact that training and testing functions have on the risk towards false positive and negative decisions. To turn these views transparent to the user, we proposed the annotation of rules and classification decisions with an indicative score of their guarantees of statistical significance. Moreover, in order to avoid the blind optimization of the behavior of classifiers according to this criteria, we revised their learning according to both accuracy (average error) and significance (variability of error) views.

Implications

Guaranteeing that classification decisions are inferred from statistically significant models is of heightened importance to learn biological and clinical markers and to support computer-aided decisions associated with medical/trading/marketing/administrative initiatives with either high impact on daily lives or high costs.

These contributions are also vital to guarantee the adequacy of learning from high-dimensional data with a (possibly) limited number of observations. In particular, they are decisive to surpass the: 1) overfitting risk of global classifiers (by guaranteeing that uninformative regions are discarded); 2) underfitting risk of classifiers reliant on procedures for dimensionality reduction (by preventing

the loss of relevant regions and the inclusion of new forms of bias); and 3) underfitting risk of local classifiers (by guaranteeing that decision rules are inferred from significantly informative and discriminative regions).

The provided results stress the importance of embracing sound statistical views to assess the propensity of a given classifier to make false positive and false negative decisions. As such, new directions for future work become necessary: the extension of the provided experimental analysis; the exploration of the provided statistical principles to further guide the learning of local classifiers; and the generalization of the proposed contributions towards alternative classification models, such as support vector machines, neural networks and Bayesian classifiers.

Sixth and finally, we tackled the multi-period classification task by proposing new algorithms able to: 1) learn sequences of classes with intricate stochastic dependencies, and 2) embed (single-label) classifiers to adequately learn from tabular and structured data. In this context, we motivated and formalized the task, studied meaningful metrics (sensitive to temporal misalignments and error accumulation on the estimated sequences) to robustly assess multi-period classifiers, and proposed new algorithms reliant on clustering for an adequate reduction and recovery of the space of sequences. Variants of the clustering-based algorithms were proposed based on the segmentation of the sequence of classes (based on their local periodicities) and on the use of sliding-windows to minimize the risk towards false positive and false negative decisions.

Implications

Multi-period classification is essential to answer a wide-set of real-world problems. The prediction of the evolving state of living, geophysical, economic and societal systems (referred as one of the ten most critical data mining challenges for this decade [693]) is necessary to: anticipate epidemics or environmental changes; assess key changes in human health (from clinical, psychophysiological and biological perspectives) and human behavior (based social data); support personalized decisions (prognostics); characterize disease progression; and support administrative decisions in both private industries and public sectors and planning tasks.

The proposed contributions to model the stochastic dependencies between the periods under classification and embed existing learning functions opens up new possibilities for the application of existing classifiers for predictive tasks from varying data structures.

1.3 Future Work

Below, we highlight potentially relevant research for future work according to the five books in this thesis.

Book II *Performance Guarantees of Models Learned from High-Dimensional Data*

- analyze the pros and cons of surveyed/proposed loss functions for local descriptive models by measuring their ability to correctly assess changes in the behavior of biclustering, triclustering and cascade learning algorithms;
- study the variability of performance (over/underfitting components) of state-of-the-art biclustering and triclustering algorithms using the proposed principles to collect error estimates;
- apply the proposed error estimators to better assess the performance of state-of-the-art and new classifiers on real and synthetic data: 1) interpret differences from using dimensionality reduction procedures with varying extent, 2) study generalization capacity (error decomposition), and 3) measure the relevance of smoothing factors on high-dimensional data for a more informed analysis of the error;
- assess the performance of learners using synthetic data with several local and global regularities (heatmaps on their empirical ability to learn from regions with varying structure and homogeneity);

Book III *Learning Local Descriptive Models from Tabular Data*

- systematic comparison of stochastic versus deterministic methods for biclustering, as well as exploration of their synergies for the development of new learning functions (combining benefits) or ensemble models;
- study the relevance and applicability of pattern-based biclustering for heterogenous data analysis (including the role of biclusters with mixture models (Basics III-3.2) to combine sets of features from different sources);
- extend (pattern-based) biclustering for integrative data analysis (such as combined analysis of biological networks and expression data [341, 297, 576]) with a focus on guiding the learning and validating outputs;

- understand to which extent can highly flexible biclustering models (such as the ones proposed in this thesis) be used to exhaustively decompose and characterize the regularities of an input dataset;
- characterize the properties of regions (structure, coherency and quality) observed across data domains;
- study and quantify the relevance of the proposed biclustering models for sparse data analysis, including: 1) structurally sparse data, 2) data with an arbitrary-high number of missings (with an heightened focus on clinical data), and 3) data with a large amount of uninformative elements (e.g. undifferentiated expression, inconclusive ratings, low word-counts, unprofitable decisions, healthy evaluations);
- analyze the role of the proposed algorithms for biclustering network data to characterize the functions of molecular entities with less-studied or with yet unclear roles based on constraint-free searches and searches for coherent modules with succinct constraints (to guarantee their inclusion in solutions);
- apply biclustering to the analysis of non-trivial communities in social networks (where individual may have different degree of activity and involvement);
- apply the plaid model to biological and social networks in order to: 1) explore network structure, 2) identify hubs, and 3) assess the type of connectivity between modules;
- apply additive and multiplicative models to analyze clinical data, with a focus on physiological responses and clinical records (highly affected by individual differences, possibly explained by scales and shifts);
- apply the order-preserving model to group individuals with coherent variation of: 1) preferences during browsing, (e-)commercing and collaborative rating, and 2) behavior for trading, interaction and tweeting;
- systematically identify and categorize the nature of the interactions between biological processes and social groups based on the analysis of their overlapping regulatory and behavioral activity (plaid model), including simplistic relations (such as is-part-of or exchanges-with) and complex interdependencies;
- analyze the relevance of multiple functions and relaxations to describe plaid effects in biological and social domains. Illustrating, scaling functions can be used to model biological processes with catalyzing regulatory effect on another process, while 'in-between' relaxations to model incremental yet non-linear effects;
- develop new pattern mining searches able to discover lengthy and dissimilar approximate patterns using smaller patterns retrieved from an adequate exploration of the different regions in a data space;
- assess of the relevance of noise-free discretization for the different real-world problems;
- overview the classes of computational methods for symbolic data analysis that can be adapted to support multi-item assignments, adapt of these methods accordingly, and quantify the changes in performance;
- study the impact of the proposed postprocessing procedures for alternative tasks (such as indexing);
- study the proposed new task of pattern mining methods with variable support (*Section III-7.2*), and assess its relevance and applicability to answer real-world problems (e.g. maximum bipartite matching);
- enumerate relevant succinct, (anti-)monotonic, convertible and prefix-monotone constraints for biomedical and social data domains, and quantification of the impact of their incorporating;
- assess the benefits of incorporating annotations from knowledge repositories (including those listed in Pointer II-2.2) and literature on the performance of pattern-based algorithms;
- study new imputation methods for missing values by discovering pattern-based biclusters covering these missings (under succinct constraints) and recovering the expected value from the underlying coherency;
- develop and validate new methods to adjust the confidence of molecular associations and predict unknown associations based on the analysis of the sparse and noisy (yet coherent) network modules where nodes of interest appear;

Book IV *Learning Local Descriptive Models from Structured Data*

- assess the relevance of cascade learning to model regulatory responses from expression time series, including the analysis of responses (uniquely) elicited for a given drug, condition or induced stimuli;
- study the temporal kinetics of regulatory cascades by further assessing the variability across observations;

- extend cascade learning to the analysis of temporal biological and social networks given by snapshots of a weighted graph at different time points;
- apply cascade learning to model unique behavioral responses associated with cascades of social interaction, web navigation, commercial activity and financial decisions;
- extract annotations and constraints from stable gene regulatory networks and use them to guide the learning of regulatory cascades from experimental data;
- extend cascade learning to integrative data analysis, and apply it for the combined analysis of expression, metabolic and proteomic time series;
- study the relevance of the proposed algorithms for multi-sets of events in the unsupervised analysis of business databases (from marked facts of interest from multi-dimensional databases or tables of interest from relational databases), as well as loose collections of user actions and financial transactions;
- study the potential of developing new learning algorithms to find alternative types of temporal patterns from structured data (such as calendric rules and chords [467]);
- integrate contributions from constraint-based sequential pattern mining to learn cascades and arrangements of events from structured data in the presence of background knowledge; systematize relevant constraints across domains; and quantify their benefits;
- generalize the proposed principles for the stochastic modeling of itemset sequences to use temporal neural networks and dynamic Bayesian networks (as they can embed Markovian assumptions [478, 74]);
- assess the fit and relevance of the proposed probabilistic methods to model web usage logs, commercial transactions, and financial decisions in the stock market;
- study the potential of the proposed HMMs to dynamically self-learn their architectures from data [603, 233];

Book V *Significance Guarantees of Local Descriptive Models*

- analyze the guarantees of statistical significance provided by (parameter-free) feature selection and dimensionality reduction procedures;
- infer heuristics from the proposed statistical views to guide alternative biclustering algorithms;
- exhaustively compare state-of-the-art biclustering algorithms regarding their: ability to minimize false positives (ability to exclude non-significant biclusters), and minimize false negatives (by testing their ability to recover planted significant biclusters);
- motivate the relevance of the proposed statistical views by collecting a sample of scientific articles (with a focus on the analysis of local associations) and validating whether the inferred scientific statements have a proper ground truth or not;
- systematically analyze the impact of the properties of a cascade/arrangement (support, shape, noise) and of the input data (size, order, time points, sparsity, regularities) on statistical significance;
- extend the proposed statistical views (for symbolic temporal patterns) towards real-valued temporal patterns;
- compare the pros and cons of statistically testing patterns from deterministic vs. probabilistic models;

Book VI *Learning Effective Classifiers from Local Descriptive Models*

- study the extensibility of the proposed contributions towards regression problems. A possible avenue is to explore mappings between classifiers and regression methods able to infer decisions from regions (such as logistic model trees);
- systematize the relevance of learning classifiers from regions with distinct coherency and quality for real-world problems (e.g. impact of order-preserving models for planning, scheduling and recommendation);
- characterize the properties of discriminative regions from labeled biomedical and social data;
- improve biclustering searches by pushing varying discriminative criteria into the mining step;
- study the applicability of branch-and-bound or iterative-deepening searches for (pattern-based) biclustering;
- apply the proposed classifiers over sparse data and networks with labeled nodes;

- study the relevance of the proposed associative classifiers to learn biological and clinical markers;
- validate biological and clinical of putative markers learned from experimental data based on the statistical significance of the learned associations;
- assess the guarantees of statistical significance of existing classifiers by analyzing the properties of the learned models from biomedical and social data domains.
- motivate the relevance of annotating classification decisions with indicators of their statistical significance by applying the proposed principles to annotate medical/trading/marketing/administrative decisions from real data (decisions with high risks and/or costs);
- study the possibility to enhance alternative classification models with the proposed statistical principles, including support vector machines, neural networks and Bayesian classifiers;
- infer heuristics to promote the selection of statistically significant regions to guide the learning of local classifiers (and quantification of their impact and ability to narrow the search space);
- quantify the impact of noise on the significance guarantees of classification rules;
- develop a methodology to transparently assess the propensity of a (local) classifier to make false positive and false negative decisions through the analysis of the learned models;
- study the relevance of modeling regions from heterogeneous data sources in classification;
- apply the proposed classifiers for structured data to a wider-set of problems, including diagnosis from (multivariate) physiological signals, financial decision support from trading actions, marketing initiatives from (e-)commerce events, and content suggestions from user logs;
- extend the proposed associative classifiers to learn from temporal snapshots of a biological or social network;
- analyze the impact of customizing the architectural components of the proposed stochastic methods in accordance with the properties of the input structured data;
- specify relevant constraints across tabular/structured data domains (along with principles for their effective incorporation) to guide the learning of classifiers;
- assess of the impact of annotating observations (extracting an arbitrary number of labels from their values) on the classifiers' behavior;
- extend associative classifiers to adequately learn from ordinal labels;
- apply multi-period classifiers to answer a wider-set of real-world problems, including prognosis of disease, planning of physical resources, and long-term budgeting predictions;
- assess the adequacy of the proposed multi-period classifiers for the prediction of sequences of upcoming events and for the classification of non-uniform and non-convex periods;
- extend multi-period classifiers to deal with lengthy sequences of labels, including principles for memory sampling (to selectively forget estimated periods) and assessing their local-stationarity (to guide segmentation);
- develop updatable multi-period classifiers to anticipate critical events based on continuously incoming data;
- quantify the impact of using similarity functions able to adequately measure temporal mismatches to affect the behavior and assess the true performance of multi-period classifiers;

The listed directions for future research can be essentially divided according to their aim: 1) application, 2) extension, 3) exploration, or 4) validation. First, the developed algorithms for describing and classifying data can be *applied* on alternative real-world data to either perform domain-specific analysis or infer domain-independent principles. Second, the same algorithms can be *extended* in order to: 1) guarantee their further improvement, 2) support new behavior (such as the possibility to interpret new forms of knowledge), and 3) solve new problems (such as the possibility to describe heterogeneous data or to classify ordinal classes). Third, the synergies between the proposed contributions and existing work can be *explored* in order to use the proposed principles to support other tasks (such as dimensionality reduction) or benefit from other tasks (such as structured pattern mining). Finally, the provided statistical and performance views can be use to *validate* available and upcoming learning algorithms and scientific statements inferred from the input data or the learned models.

Appendix

Remaining publications from the candidate during his PhD:

- Rui Henriques and Ana Paiva, 2014, *Seven Principles to Mine Flexible Behavior from Physiological Signals for Effective Emotion Recognition and Description in Affective Interactions*, In *Physiological Computing Systems*, 75-82, ScitePress
- Rui Henriques and Cláudia Antunes, 2014, *Generative Modeling of Itemset Sequences derived from Real Databases*, In *IC on Enterprise Information Systems (ICEIS)*, 264-272, INSTICC, Lisbon, Portugal
- Rui Henriques and Cláudia Antunes, 2014, *Learning Predictive Models from Integrated Healthcare Data: Extending Pattern-based and Generative Models to Capture Temporal and Cross-Attribute Dependencies*, In *Hawai International Conference on System Sciences (HICSS)*, 2562-2569, IEEE CS, Waikoloa, Hawaii
- Rui Henriques and Sara C. Madeira, 2014, *BiP: Effective Discovery of Overlapping Biclusters using Flexible Plaid Models*, In *13th IW of Data Mining in Bioinformatics (BIOKDD'14)*, ACM, New York, US
- Rui Henriques and Ana Paiva, 2014, *Descriptive Models of Emotion: Learning Useful Abstractions from Physiological Responses during Affective Interactions*, In *PhyCS Special Session on Recognition of Affect Signals from Physiological Data for Social Robots (OASIS'14)*, 393-400, SCITEPRESS
- Rui Henriques and Ana Paiva, 2014, *Learning Effective Models of Emotion from Physiological Signals: The 7 Principles*, *Lecture Notes in Computer Science*, Vol. 8908, Chapter 9, Springer
- Rui Henriques, Silvia Pina and Cláudia Antunes, 2013, *Temporal Mining of Integrated Healthcare Data: Methods, Revelings and Implications*, In *SDM DM for Medicine and Healthcare*, 56-64, SIAM, Austin, US
- Rui Henriques, Ana Paiva and Cláudia Antunes, 2013, *Accessing Emotion Patterns from Affective Interactions using Electrodermal Activity*, In *Affective Computing and Intelligent Interaction (ACII'13)*, 43-48, IEEE Computer Society, Geneva, Switzerland
- Rui Henriques, Sara C. Madeira and Cláudia Antunes, 2013, *IndexSpan: Efficient Discovery of Item-Indexable Sequential Patterns*, In *ECML/PKDD IW on New Frontiers to Mine Complex Patterns*, Springer-Verlag, Prague, Czech Republic
- Rui Henriques and Cláudia Antunes, 2012, *On the Need of New Approaches for the Novel Problem of Long-term Prediction over Multi-Dimensional Data*, In *Computer and Information Science 2012*, Volume 429 of *Studies in Computational Intelligence*, 121-138, Springer Berlin Heidelberg, Shangai, China
- Rui Henriques and Cláudia Antunes, 2012, *An Integrated Approach for Healthcare Planning over Multi-Dimensional Data using Long-term Prediction*, In *Health Information Science*, Volume 7231 of *Lecture Notes in Computer Science*, 36-48, Springer Berlin/Heidelberg, Beijing, China
- Rui Henriques, Inês Lynce and Vasco Manquinho, 2012, *On When and How to use SAT to Mine Frequent Itemsets*, IST/INESC-ID, CoRR, abs/1207.6253, Cornell Univ. Library
- Rui Henriques, Ana Paiva and Cláudia Antunes, 2012, *On the Need of New Methods to Mine Electrodermal Activity in Emotion-centered Studies*, In *AAMAS'12, 8th IW on Agents and Data Mining Interaction*, Springer-Verlag LNAI series, Valência, Spain

Bibliography

- [1] *Pattern Extraction for Time Series Classification*, PKDD '01, London, UK, UK, 2001. Springer-Verlag.
- [2] M. Abraham, J.T. Ahlman, A.J. Boudreau, J.L. Connelly, and D.D. Evans. *CPT 2011: Standard Edition*. CPT / Current Procedural Terminology. American Medical Association Press, 2010.
- [3] C. J. Adcock. Sample size determination: a review. *J. of the Royal Statistical Society: Series D (The Statistician)*, 46(2):261–283, 1997.
- [4] Ramesh C Agarwal, Charu C Aggarwal, and VVV Prasad. A tree projection algorithm for generation of frequent item sets. *Journal of parallel and Distributed Computing*, 61(3):350–371, 2001.
- [5] C. Aggarwal and P. Yu. A new framework for itemset generation. In *Symposium on Princ. of database systems*, pages 18–24. ACM, 1998.
- [6] Charu C. Aggarwal, Yan Li, Jianyong Wang, and Jing Wang. Frequent pattern mining with uncertain data. In *Proceedings of the 15th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, KDD '09, pages 29–38, New York, NY, USA, 2009. ACM.
- [7] Rakesh Agrawal, Tomasz Imieliński, and Arun Swami. Mining association rules between sets of items in large databases. *SIGMOD Rec.*, 22(2):207–216, June 1993.
- [8] Rakesh Agrawal and John C Shafer. Parallel mining of association rules. *IEEE Transactions on Knowledge & Data Engineering*, (6):962–969, 1996.
- [9] Rakesh Agrawal and Ramakrishnan Srikant. Fast algorithms for mining association rules in large databases. In *VLDB*, pages 487–499, San Francisco, USA, 1994. Morgan Kaufmann.
- [10] Rakesh Agrawal and Ramakrishnan Srikant. Mining sequential patterns. In *ICDE*, pages 3–14, Washington, DC, USA, 1995. IEEE Computer Society.
- [11] David W. Aha, Dennis Kibler, and Marc K. Albert. Instance-based learning algorithms. *Mach. Learn.*, 6(1):37–66, January 1991.
- [12] HA Ahmed, P Mahanta, DK Bhattacharyya, JK Kalita, and A Ghosh. Intersected coexpressed subcube miner: An effective triclustering algorithm. In *Information and Communication Technologies (WICT), 2011 World Congress on*, pages 846–851. IEEE, 2011.
- [13] H.A. Ahmed, P. Mahanta, D.K. Bhattacharyya, J.K. Kalita, and A. Ghosh. Intersected coexpressed subcube miner: An effective triclustering algorithm. In *Information and Communication Technologies (WICT), 2011 World Congress on*, pages 846–851, dec. 2011.
- [14] Jaegyoon Ahn, Youngmi Yoon, and Sanghyun Park. Noise-robust algorithm for identifying functionally associated biclusters from gene expression data. *Inf. Sci.*, 181(3):435–449, February 2011.
- [15] Tatsuya Akutsu, Satoru Miyano, Satoru Kuhara, et al. Identification of genetic networks from a small number of gene expression patterns under the boolean network model. In *Pacific Symposium on Biocomputing*, volume 4, pages 17–28, 1999.
- [16] Fátima Al-Shahrour, Leonardo Arbiza, Hernán Dopazo, Jaime Huerta-Cepas, Pablo Minguéz, David Montaner, and Joaquín Dopazo. From genes to functional classes in the study of biological systems. *BMC Bioinformatics*, 8, 2007.
- [17] Fadhil Alakwaa, Nahed Solouma, and Yasser Kadah. Construction of gene regulatory networks using biclustering and bayesian networks. *Theoretical Biology and Medical Modelling*, 8(1):39, 2011.
- [18] Ash A. Alizadeh, Michael B. Eisen, R. Eric Davis, Chi Ma, Izidore S. Lossos, Andreas Rosenwald, Jennifer C. Boldrick, Hajeer Sabet, Truc Tran, Xin Yu, John I. Powell, Liming Yang, Gerald E. Marti, Troy Moore, Jr James Hudson, Lisheng Lu, et al. Distinct types of diffuse large B-cell lymphoma identified by gene expression profiling. *Nature*, 403(6769):503–511, 2000.
- [19] Uri Alon. Network motifs: theory and experimental approaches. *Nature Reviews Genetics*, 8(6):450–461, 2007.
- [20] Ronnie Alves, Domingo S. Rodríguez-Baena, and Jesús S. Aguilar-Ruiz. Gene association analysis: a survey of frequent pattern mining from gene expression data. *Briefings in Bioinformatics*, 11(2):210–224, 2010.
- [21] David Amar, Daniel Yekutieli, Adi Maron-Katz, Talma Hendler, and Ron Shamir. A hierarchical bayesian model for flexible module discovery in three-way time-series data. *Bioinformatics*, 31(12):i17–i26, 2015.
- [22] D. Amaratunga, J. Cabrera, and Z. Shkedy. *Exploration and Analysis of DNA Microarray and Other High-Dimensional Data*. Wiley Series in Probability and Statistics. Wiley, 2014.
- [23] IP Androulakis, E Yang, and RR Almon. Analysis of time-series gene expression data: methods, challenges, and opportunities. *Annu. Rev. Biomed. Eng.*, 9:205–228, 2007.
- [24] Cláudia Antunes and Arlindo L Oliveira. Generalization of pattern-growth methods for sequential pattern mining with gap constraints. In *Machine Learning and Data Mining in Pattern Recognition*, pages 239–251. Springer, 2003.
- [25] Claudia Antunes and Arlindo L. Oliveira. Mining patterns using relaxations of user defined constraints. In *Knowledge Discovery in Inductive Databases*, 2004.
- [26] B. Apolloni and C. Gentile. Sample size lower bounds in pac learning by algorithmic complexity theory. *Theoretical Computer Science*, 209(1.2):141 – 162, 1998.
- [27] I. Assent, R. Krieger, E. Muller, and T. Seidl. DUSC: Dimensionality Unbiased Subspace Clustering. In *ICDM*, 2007.
- [28] Ira Assent, Emmanuel Müller, Ralph Krieger, Timm Jansen, and Thomas Seidl. Pleiades: Subspace clustering and evaluation. In Walter Daelemans, Bart Goethals, and Katharina Morik, editors, *ECML/PKDD (2)*, volume 5212 of *Lecture Notes in Computer Science*, pages 666–671. Springer, 2008.
- [29] Gowtham Atluri, Jeremy Bellay, Gaurav Pandey, Chad Myers, and Vipin Kumar. Discovering coherent value bicliques in genetic interaction data. In *BIOKDD*, 2000.
- [30] Gowtham Atluri, Jeremy Bellay, Gaurav Pandey, Chad Myers, and Vipin Kumar. Discovering coherent value bicliques in genetic interaction data. In *IW on Data Mining in Bioinformatics*, 2010.
- [31] Jay Ayres, Jason Flannick, Johannes Gehrke, and Tomi Yiu. Sequential pattern mining using a bitmap representation. In *KDD*, pages 429–435, New York, NY, USA, 2002. ACM.
- [32] Francisco Azuaje. Integrative data analysis for biomarker discovery. *Bioinformatics and Biomarker Discovery: Omic Data Analysis for Personalized Medicine*, pages 137–154, 2011.
- [33] K. Bache and M. Lichman. UCI machine learning repository, 2013.
- [34] Christoph P Bagowski and James E Ferrell Jr. Bistability in the jnk cascade. *Current Biology*, 11(15):1176–1182, 2001.
- [35] R. Rathipriya K. Thangavel J. Bagyamani. Binary particle swarm optimization based biclustering of web usage data. *CoRR*, abs/1108.0748, 2011.

- [36] P. Baldi and S. Brunak. *Bioinformatics: The Machine Learning Approach*. Adaptive Computation and Machine Learning. MIT Press, 2nd edition, 2001.
- [37] P. Baldi, Y. Chauvin, Y. Hunkapliier, and M. McClure. Hidden Markov models of biological primary sequence information. *Proceedings of the National Academy of Sciences*, 91(3):1059–1063, 1994.
- [38] Arindam Banerjee, Inderjit Dhillon, Joydeep Ghosh, Srujana Merugu, and Dharmendra S Modha. A generalized maximum entropy approach to bregman co-clustering and matrix approximation. In *Proceedings of the tenth ACM SIGKDD international conference on Knowledge discovery and data mining*, pages 509–514. ACM, 2004.
- [39] Mukesh Bansal, Vincenzo Belcastro, Alberto Ambesi-Impiombato, and Diego Di Bernardo. How to infer gene networks from expression profiles. *Molecular systems biology*, 3(1), 2007.
- [40] Ziv Bar-Joseph. Analyzing time series gene expression data. *Bioinformatics*, 20(16):2493–2503, 2004.
- [41] Ziv Bar-Joseph. Analyzing time series gene expression data. *Bioinformatics*, 20(16):2493–2503, 2004.
- [42] Ziv Bar-Joseph, Georg K Gerber, Tong Ihn Lee, Nicola J Rinaldi, Jane Y Yoo, François Robert, D Benjamin Gordon, Ernest Fraenkel, Tommi S Jaakkola, Richard A Young, et al. Computational discovery of gene modules and regulatory networks. *Nature biotechnology*, 21(11):1337–1342, 2003.
- [43] Albert-Laszlo Barabasi and Zoltan N Oltvai. Network biology: understanding the cell’s functional organization. *Nature Reviews Genetics*, 5(2):101–113, 2004.
- [44] Elena Baralis, Luca Cagliero, Tania Cerquitelli, and Paolo Garza. Generalized association rule mining with constraints. *Information Sciences*, 194:68 – 84, 2012. Intelligent Knowledge-Based Models and Methodologies for Complex Information Systems.
- [45] Simon Barkow, Stefan Bleuler, Amela Prelić, Philip Zimmermann, and Eckart Zitzler. Bicat: a biclustering analysis toolbox. *Bioinformatics*, 22(10):1282–1283, May 2006.
- [46] Gustavo E. A. P. A. Batista, Xiaoyue Wang, and Eamonn J. Keogh. A complexity-invariant distance measure for time series. In *SDM’11*, pages 699–710. SIAM, 2011.
- [47] Rohan A. Baxter, Graham J. Williams, and Hongxing He. Feature selection for temporal health records. In *PAKDD*, pages 198–209, London, UK, UK, 2001. Springer-Verlag.
- [48] S. Bay and M. Pazzani. Detecting group differences: Mining contrast sets. *Data Min. Knowl. Discov.*, 5(3):213–246, July 2001.
- [49] Roberto J. Bayardo. Efficiently mining long patterns from databases. *SIGMOD Rec.*, 27(2):85–93, 1998.
- [50] Roberto J. Bayardo, Jr. Efficiently mining long patterns from databases. *SIGMOD Rec.*, 27(2):85–93, June 1998.
- [51] Roberto J Bayardo Jr, Rakesh Agrawal, and Dimitrios Gunopulos. Constraint-based rule mining in large, dense databases. In *Data Engineering, 1999. Proceedings., 15th International Conference on*, pages 188–197. IEEE, 1999.
- [52] Matthew J Beal, Francesco Falciani, Zoubin Ghahramani, Claudia Rangel, and David L Wild. A bayesian approach to reconstructing genetic regulatory networks with hidden factors. *Bioinformatics*, 21(3):349–356, 2005.
- [53] Gürkan Bebek and Jiong Yang. Pathfinder: mining signal transduction pathway segments from protein-protein interaction networks. *BMC Bioinformatics*, 8, 2007.
- [54] Claudia Beleites, Ute Neugebauer, Thomas Bocklitz, Christoph Krafft, and Jurgen Popp. Sample size planning for classification models. *Analytica Chimica Acta*, 760(0):25 – 33, 2013.
- [55] Claudia Beleites, Ute Neugebauer, Thomas Bocklitz, Christoph Krafft, and Jurgen Popp. Sample size planning for classification models. *Analytica Chimica Acta*, 760(0):25 – 33, 2013.
- [56] Jeremy Bellay, Gowtham Atluri, Tina L. Sing, Kiana Toufighi, Michael Costanzo, and et al. Putting genetic interactions in context through a global modular decomposition. *Genome Research*, 21(8):1375–1387, 2011.
- [57] Jeremy Bellay, Gowtham Atluri, Tina L Sing, Kiana Toufighi, Michael Costanzo, Philippe Souza Moraes Ribeiro, Gaurav Pandey, Joshua Baller, Benjamin VanderSluis, Magali Michaut, et al. Putting genetic interactions in context through a global modular decomposition. *Genome research*, 21(8):1375–1387, 2011.
- [58] Riccardo Bellazzi, Fulvia Ferrazzi, and Lucia Sacchi. Predictive data mining in clinical medicine: a focus on selected methods and applications. *Wiley Interdisc. Rev.: Data Mining and Knowledge Discovery*, 1(5):416–430, 2011.
- [59] Amir Ben-Dor, Benny Chor, Richard Karp, and Zohar Yakhini. Discovering local structure in gene expression data: the order-preserving submatrix problem. In *RECOMB*, pages 49–57, New York, NY, USA, 2002. ACM.
- [60] Souhaib Ben Taieb, Gianluca Bontempi, Amir F. Atiya, and Antti Sorjamaa. A review and comparison of strategies for multi-step ahead time series forecasting based on the nn5 forecasting competition. *Expert Syst. Appl.*, 39(8):7067–7083, June 2012.
- [61] Souhaib Ben Taieb, Antti Sorjamaa, and Gianluca Bontempi. Multiple-output modeling for multi-step-ahead time series forecasting. *Neurocomput.*, 73:1950–1957, June 2010.
- [62] Samy Bengio, Françoise Fessant, and Daniel Collobert. Use of modular architectures for time series prediction. *Neural Process. Lett.*, 3:101–106, June 1996.
- [63] Y. Benjamini and Y. Hochberg. Controlling the false discovery rate: A practical and powerful approach to multiple testing. *Royal Statistical Society: Series B*, 57(1):289–300, 1995.
- [64] Yoav Benjamini and Daniel Yekutieli. The control of the false discovery rate in multiple testing under dependency. *Annals of Statistics*, 29:1165–1188, 2001.
- [65] Johannes Berg and Michael Lässig. Local graph alignment and motif search in biological networks. *National Academy of Sciences*, 101(41):14689–14694, 2004.
- [66] Thomas Bernecker, Hans-Peter Kriegel, Matthias Renz, Florian Verhein, and Andreas Zuefle. Probabilistic frequent itemset mining in uncertain databases. In *Proceedings of the 15th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, KDD ’09*, pages 119–128, New York, NY, USA, 2009. ACM.
- [67] G. F. Berriz, O. D. King, B. Bryant, C. Sander, and F. P. Roth. Characterizing gene sets with FuncAssociate. *Bioinformatics*, 19:2502–2504, 2003.
- [68] Jérémy Besson, Céline Robardet, Jean-François Boulicaut, and Sophie Rome. Constraint-based concept mining and its application to microarray data analysis. *Intell. Data Anal.*, 9(1):59–82, January 2005.
- [69] Jérémy Besson, Céline Robardet, Luc De Raedt, and Jean-François Boulicaut. Mining bi-sets in numerical data. In *Knowledge Discovery in Inductive Databases*, volume 4747 of LNCS, pages 11–23. Springer, 2007.
- [70] Claudio Bettini, Xiaoyang Sean Wang, and Sushil Jajodia. Mining temporal relationships with multiple granularities in time sequences. *IEEE Data Eng. Bull.*, 21(1):32–38, 1998.
- [71] Jinbo Bi, Kristin Bennett, Mark Embrechts, Curt Breneman, and Minghu Song. Dimensionality reduction via sparse support vector machines. *The Journal of Machine Learning Research*, 3:1229–1243, 2003.
- [72] Manuele Bicego, Pietro Lovato, Alberto Ferrarini, and Massimo DelleDonne. Biclustering of expression microarray data with topic models. In *IC on Pattern Recognition (ICPR)*, pages 2728–2731. IEEE, 2010.
- [73] Jeff A. Bilmes. What hmms can do. *IEICE journal*, E89-D(3):869–891, 2006.

- [74] C.M. Bishop. *Pattern Recognition and Machine Learning*. Information Science and Statistics. Springer, 2006.
- [75] Sylvain Blachon, Ruggero Pensa, Jeremy Besson, Celine Robardet, Jean-Francois Boulicaut, and Olivier Gandrillon. Clustering formal concepts to discover biologically relevant knowledge from gene expression data. *In Silico Biology*, 7(0033), July 2007.
- [76] Anselm Blumer, A. Ehrenfeucht, David Haussler, and Manfred K. Warmuth. Learnability and the vapnik-chervonenkis dimension. *J. ACM*, 36(4):929–965, October 1989.
- [77] Valeria Bo, Tanya Curtis, Artem Lysenko, Mansoor Saqi, Stephen Swift, and Allan Tucker. Discovering study-specific gene regulatory networks. *PLoS ONE*, 9(9):e106524, 09 2014.
- [78] Dankmar Böhning. Multinomial logistic regression algorithm. *Annals of the Institute of Statistical Mathematics*, 44(1):197–200, 1992.
- [79] Hamid Bolouri and Eric H Davidson. Modeling transcriptional regulatory networks. *BioEssays*, 24(12):1118–1129, 2002.
- [80] Francesco Bonchi, Carlos Castillo, Aristides Gionis, and Alejandro Jaimes. Social network analysis and mining for business applications. *ACM Trans. Intell. Syst. Technol.*, 2(3):22:1–22:37, May 2011.
- [81] Francesco Bonchi, Fosca Giannotti, Alessio Mazzanti, and Dino Pedreschi. Exante: A preprocessing method for frequent-pattern mining. *IEEE Intel. Systems*, 20(3):25–31, May 2005.
- [82] Francesco Bonchi and Bart Goethals. Fp-bonsai: The art of growing and pruning small fp-trees. In Honghua Dai, Ramakrishnan Srikant, and Chengqi Zhang, editors, *Advances in Knowledge Discovery and Data Mining*, volume 3056 of *LNCS*, pages 155–160. Springer Berlin Heidelberg, 2004.
- [83] Francesco Bonchi and Claudio Lucchese. Extending the state-of-the-art of constraint-based pattern discovery. *Data Knowl. Eng.*, 60(2):377–399, 2007.
- [84] P.P. Bonissone, R. Subbu, N. Eklund, and T.R. Kiehl. Evolutionary algorithms + domain knowledge = real-world evolutionary computation. *Evolutionary Computation, IEEE Transactions on*, 10(3):256–280, June 2006.
- [85] G. Bontempi, M. Birattari, and H. Bersini. Lazy learning for iterated time-series prediction. In J. A. K. Suykens and J. Vandewalle, editors, *International Workshop on Advanced Black-Box Techniques for Nonlinear Modeling*, pages 62–68, Leuven, Belgium, 1998. Katholieke Universiteit Leuven.
- [86] Gianluca Bontempi and Souhaib Ben Taieb. Conditionally dependent strategies for multiple-step-ahead prediction in local learning. *Int. Journal of Forecasting*, 27(2004):689–699, 2011.
- [87] Natthaphan Boonyanunta and Panlop Zeephongsekul. Predicting the relationship between the size of training sample and the predictive power of classifiers. In *Knowledge-Based Intelligent Information and Engineering Systems*, volume 3215 of *LNCS*, pages 529–535. Springer Berlin Heidelberg, 2004.
- [88] José Borges and Mark Levene. A dynamic clustering-based markov model for web usage mining. *arXiv preprint cs/0406032*, 2004.
- [89] Jean-François Boulicaut and Jérémy Besson. Actionability and formal concepts: a data mining perspective. In *Proceedings of the 6th international conference on Formal concept analysis, ICFA'08*, pages 14–31, Berlin, Heidelberg, 2008. Springer-Verlag.
- [90] Jean-Francois Boulicaut. Inductive databases and multiple uses of frequent itemsets: The cinq approach. In Rosa Meo, PierLuca Lanzi, and Mika Klemettinen, editors, *Database Support for Data Mining Applications*, volume 2682 of *Lecture Notes in Computer Science*, pages 1–23. Springer Berlin Heidelberg, 2004.
- [91] Doruk Bozdağ, Ashwin S. Kumar, and Umit V. Catalyurek. Comparative analysis of biclustering algorithms. In *BCB*, pages 265–274, New York, NY, USA, 2010. ACM.
- [92] ADRIAN P Bracken and URSULA Bond. Reassembly and protection of small nuclear ribonucleoprotein particles by heat shock proteins in yeast cells. *Rna*, 5(12):1586–1596, 1999.
- [93] P. S. Bradley, C. A. Reina, and U. M. Fayyad. Clustering Very Large Databases Using EM Mixture Models. *Pattern Recognition, International Conference on*, 2:2076+, 2000.
- [94] Sofiane Brahim-Belhouari and Amine Bermak. Gaussian process for nonstationary time series prediction. *Computational Statistics and Data Analysis*, 47(4):705 – 712, 2004.
- [95] Matthew Brand. Structure learning in conditional probability models via an entropic prior and parameter extinction. *Neural Comput.*, 11(5):1155–1182, July 1999.
- [96] Leo Breiman. Random forests. *Machine Learning*, 45(1):5–32, 2001.
- [97] S. Brin, R. Motwani, and C. Silverstein. Beyond market baskets: generalizing association rules to correlations. *SIGMOD Rec.*, 26(2):265–276, June 1997.
- [98] Sergey Brin, Rajeev Motwani, Jeffrey D Ullman, and Shalom Tsur. Dynamic itemset counting and implication rules for market basket data. *ACM SIGMOD Record*, 26(2):255–264, 1997.
- [99] Björn Bringmann, Siegfried Nijssen, and Albrecht Zimmermann. Pattern-based classification: a unifying perspective. *arXiv preprint arXiv:1111.6191*, 2011.
- [100] Bjorn Bringmann and Albrecht Zimmermann. The chosen few: On identifying valuable patterns. In *Proceedings of the 2007 Seventh IEEE International Conference on Data Mining, ICDM '07*, pages 63–72, Washington, DC, USA, 2007. IEEE Computer Society.
- [101] Peter J Brockwell and Richard A Davis. *Time series: theory and methods*. Springer Science & Business Media, 2013.
- [102] George W. Brown. On small-sample estimation. *The Annals of Mathematical Statistics*, 18(4):582–585, 1947.
- [103] Michael Brown, Richard Hughey, Anders Krogh, I. Saira Mian, Kimmen Sjölander, and David Haussler. Using dirichlet mixture priors to derive hidden markov models for protein families. In *1st IC on Int. Sys. for Molecular Bio.*, pages 47–55. AAAI Press, 1993.
- [104] Giulia Bruno and Paolo Garza. Temporal pattern mining for medical applications. In *Data Min.: Found. and Int. Paradigms*, volume 25 of *ISRL*, pages 9–18. Springer Heidelberg, 2012.
- [105] P. Bühlmann and S. van de Geer. *Statistics for High-Dimensional Data: Methods, Theory and Applications*. Springer Series in Statistics. Springer, 2011.
- [106] Douglas Burdick, Manuel Calimlim, and Johannes Gehrke. Mafia: A maximal frequent itemset algorithm for transactional databases. In *ICDE*, pages 443–452. IEEE CS, 2001.
- [107] Christopher J. C. Burges. A tutorial on support vector machines for pattern recognition. *Data Min. Knowl. Discov.*, 2:121–167, June 1998.
- [108] Stanislav Busygin, Nikita Boyko, Panos M Pardalos, Michael Bewernitz, and Georges Ghacibeh. Biclustering eeg data from epileptic patients treated with vagus nerve stimulation. In *Data mining, systems analysis and optimization in biomedicine*, volume 953, pages 220–231. AIP Publishing, 2007.
- [109] Stanislav Busygin, Gerrit Jacobsen, Ewald Krämer, and Contentsoft Ag. Double conjugated clustering applied to leukemia microarray data. In *ICDM IW on clustering high dimensional data*, 2002.
- [110] Richard H. Byrd, Gillian M. Chin, Jorge Nocedal, and Yuchen Wu. Sample size selection in optimization methods for machine learning. *Math. Program.*, 134(1):127–155, August 2012.

- [111] Igor Cadez, David Heckerman, Christopher Meek, Padhraic Smyth, and Steven White. Visualization of navigation patterns on a web site using model-based clustering. In *Proceedings of the Sixth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, KDD '00, pages 280–284, New York, NY, USA, 2000. ACM.
- [112] Tony Cai and Xiaotong Shen. *High-Dimensional Data Analysis (Frontiers of Statistics)*. World Scientific, 2010.
- [113] J. Caldas and S. Kaski. Bayesian biclustering with the plaid model. In *IEEE Workshop on MLSP*, pages 291–296, Oct 2008.
- [114] Toon Calders and Bart Goethals. Mining all non-derivable frequent itemsets. In *Proceedings of the 6th European Conference on Principles of Data Mining and Knowledge Discovery*, PKDD '02, pages 74–85, London, UK, UK, 2002. Springer-Verlag.
- [115] Toon Calders, Bart Goethals, and Szymon Jaroszewicz. Mining rank-correlated sets of numerical attributes. In *KDD*, pages 96–105, New York, NY, USA, 2006. ACM.
- [116] Andrea Califano, Gustavo Stolovitzky, and Yuhai Tu. Analysis of gene expression microarrays for phenotype classification. In *Proc. Int. Conf. Intell. Syst. Mol. Biol*, pages 75–85, 2000.
- [117] Steve E Calvano, Wenzhong Xiao, Daniel R Richards, Ramon M Felciano, Henry V Baker, Raymond J Cho, Richard O Chen, Bernard H Brownstein, J Perren Cobb, S Kevin Tschoeke, et al. A network-based analysis of systemic inflammation in humans. *Nature*, 437(7061):1032–1037, 2005.
- [118] Irene Cantone, Lucia Marucci, Francesco Iorio, Maria Aurelia Ricci, Vincenzo Belcastro, Mukesh Bansal, Stefania Santini, Mario di Bernardo, Diego di Bernardo, and Maria Pia Cosma. A yeast synthetic network for in vivo assessment of reverse-engineering and modeling approaches. *Cell*, 137(1):172–181, 2009.
- [119] Longbing Cao, Yuming Ou, Philip S. Yu, and Gang Wei. Detecting abnormal coupled sequences and sequence changes in group-based manipulative trading behaviors. In *ACM SIGKDD*, pages 85–94, New York, NY, USA, 2010. ACM.
- [120] Pedro Carmona-Saez, Monica Chagoyen, Andres Rodriguez, Oswaldo Trelles, JoseM Carazo, and Alberto Pascual-Montano. Integrated analysis of gene expression by association rules discovery. *BMC Bioinformatics*, 7:1–16, 2006.
- [121] André V Carreiro, Orlando Anunciação, João A Carriço, and Sara C Madeira. Biclustering-based classification of clinical expression time series: A case study in patients with multiple sclerosis. In *5th International Conference on Practical Applications of Computational Biology & Bioinformatics (PACBB 2011)*, pages 229–239. Springer, 2011.
- [122] André Valério Carreiro, Artur J Ferreira, Mário AT Figueiredo, and Sara Cordeiro Madeira. Towards a classification approach using meta-biclustering: Impact of discretization in the analysis of expression time series. *Journal of integrative bioinformatics*, 9(3):207, 2012.
- [123] Malika Charrad and Mohamed Ben Ahmed. Simultaneous clustering: A survey. In *Pattern Recognition and Machine Intelligence*, pages 370–375. Springer, 2011.
- [124] Dehao Chen, Chunrong Lai, Wei Hu, Wenguang Chen, Yimin Zhang, and Weimin Zheng. Tree partition based parallel frequent pattern mining on shared memory systems. In *IPDPS*, page 8, 2006.
- [125] Jingchun Chen and Bo Yuan. Detecting functional modules in the yeast protein-protein interaction network. *Bioinformatics*, 22(18):2283–2290, 2006.
- [126] Ting Chen, Hongyu L He, George M Church, et al. Modeling gene expression with differential equations. In *Pacific symposium on biocomputing*, volume 4, page 4, 1999.
- [127] Ting Chen, Hongyu L He, George M Church, et al. Modeling gene expression with differential equations. In *Pacific symposium on biocomputing*, volume 4, page 4, 1999.
- [128] Xiaodong Chen and Ilias Petrounias. A framework for temporal data mining. In *Database and Expert Systems App.*, DEXA '98, pages 796–805, London, UK, UK, 1998. Springer-Verlag.
- [129] Haibin Cheng, Pang-Ning Tan, Jing Gao, and Jerry Scripps. Multistep-ahead time series prediction. In Ng, Kitsuregawa, Li, and Chang, editors, *Advances in Knowledge Discovery and Data Mining*, volume 3918 of *Lecture Notes in Computer Science*, pages 765–774. Springer Berlin, Heidelberg, 2006.
- [130] Hong Cheng. *Towards Accurate and Efficient Classification: A Discriminative and Frequent Pattern-Based Approach*. ProQuest, 2008.
- [131] Hong Cheng, Xifeng Yan, Jiawei Han, and Philip S Yu. Direct discriminative pattern mining for effective classification. In *Data Engineering, 2008. ICDE 2008. IEEE 24th International Conference on*, pages 169–178. IEEE, 2008.
- [132] Hong Cheng, Philip S. Yu, and Jiawei Han. Approximate frequent itemset mining in the presence of random noise. In Oded Maimon and Lior Rokach, editors, *Soft Computing for Knowl. Disc. and Data Mining*, pages 363–389. Springer, 2008.
- [133] Kin-On Cheng, Ngai-Fong Law, Wan-Chi Siu, and TH Lau. Bivisu: software tool for bicluster detection and visualization. *Bioinformatics*, 23(17):2342–2344, 2007.
- [134] Yizong Cheng and George M. Church. Biclustering of expression data. In *Intelligent Systems for Molecular Biology*, pages 93–103. AAAI Press, 2000.
- [135] J Michael Cherry, Eurie L Hong, Craig Amundsen, Rama Balakrishnan, Gail Binkley, Esther T Chan, Karen R Christie, Maria C Costanzo, Selina S Dwight, Stacia R Engel, et al. *Saccharomyces genome database: the genomics resource of budding yeast*. *Nucleic acids research*, page gkr1029, 2011.
- [136] David W Cheung, Jiawei Han, Vincent T Ng, and CY Wong. Maintenance of discovered association rules in large databases: An incremental updating technique. In *Data Engineering, 1996. Proceedings of the Twelfth International Conference on*, pages 106–114. IEEE, 1996.
- [137] Ding-Ying Chiu, Yi-Hung Wu, and Arbee L. P. Chen. An efficient algorithm for mining frequent sequences by a new strategy without support counting. In *ICDE*, pages 375–, Washington, DC, USA, 2004. IEEE Computer Society.
- [138] Hyungwon Choi, Sinae Kim, Anne-Claude Gingras, and Alexey I Nevizhskii. Analysis of protein complexes through model-based biclustering of label-free quantitative ap-ms data. *Mol Syst Biol*, 6:385, 2010.
- [139] Keunho Choi, Sukhoon Chung, Hyunsill Rhee, and Yongmoo Suh. Classification and sequential pattern analysis for improving managerial efficiency and providing better medical service in public healthcare centers. *Healthc Inform Res.*, 16(2):67–76, 2010.
- [140] Keunho Choi, Sukhoon Chung, Hyunsill Rhee, and Yongmoo Suh. Classification and sequential pattern analysis for improving managerial efficiency and providing better medical service in public healthcare centers. *Healthcare informatics research*, 16(2):67–76, 2010.
- [141] Salim A Chowdhury and Mehmet Koyutürk. Identification of coordinately dysregulated subnetworks in complex phenotypes. In *Pacific Symposium on Biocomputing*, volume 15, pages 133–144. World Scientific, 2010.
- [142] Han-Yu Chuang, Eunjung Lee, Yu-Tsueng Liu, Doheon Lee, and Trey Ideker. Network-based classification of breast cancer metastasis. *Molecular Systems Biology*, 3(1), 2007.
- [143] Darya Chudova and Padhraic Smyth. Pattern discovery in sequences under a markov assumption. In *8th ACM SIGKDD*, KDD '02, pages 153–162, New York, NY, USA, 2002. ACM.
- [144] Recep Colak. *Towards finding the complete modulome: density constrained biclustering*. PhD thesis, Simon Fraser University, 2008.
- [145] Recep Colak, Flavia Moser, Jeffrey Shih-Chieh Chu, Alexander Schönhuth, Nansheng Chen, and Martin Ester. Module discovery by exhaustive search for densely connected, co-expressed regions in biomolecular interaction networks. *PLoS One*, 5(10):e13348, 2010.

- [146] Gao Cong, Kian-Lee Tan, Anthony K. H. Tung, and Xin Xu. Mining top-k covering rule groups for gene expression data. In *Proceedings of the 2005 ACM SIGMOD International Conference on Management of Data*, SIGMOD '05, pages 670–681, New York, NY, USA, 2005. ACM.
- [147] Gao Cong, Kian-Lee Tan, Anthony KH Tung, and Xin Xu. Mining top-k covering rule groups for gene expression data. In *Proceedings of the 2005 ACM SIGMOD international conference on Management of data*, pages 670–681. ACM, 2005.
- [148] Gao Cong, Anthony KH Tung, Xin Xu, Feng Pan, and Jiong Yang. Farmer: Finding interesting rule groups in microarray datasets. In *Proceedings of the 2004 ACM SIGMOD international conference on Management of data*, pages 143–154. ACM, 2004.
- [149] Daniel L Cook, Anthony N Gerber, and Stephen J Tapscott. Modeling stochastic gene expression: implications for haploinsufficiency. *Proceedings of the National Academy of Sciences*, 95(26):15641–15646, 1998.
- [150] Alberto Corigliano and Stefano Mariani. Parameter identification in explicit structural dynamics: performance of the extended kalman filter. *Computer Methods in Applied Mechanics and Engineering*, 193(36):3807–3835, 2004.
- [151] Ivan G. Costa, Alexander Schönhuth, Christoph Hafemeister, and Alexander Schliep. Constrained mixture estimation for analysis and robust classification of clinical time series. *Bioinformatics*, 25(12):i6–i14, June 2009.
- [152] Chad Creighton and Samir Hanash. Mining gene expression databases for association rules. *Bioinformatics*, 19(1):79–86, 2003.
- [153] Nello Cristianini and John Shawe-Taylor. *An introduction to support vector machines and other kernel-based learning methods*. Cambridge university press, 2000.
- [154] Phuong Dao, Recep Colak, Salari, Moser, Davicioni, Schönhuth, and Martin Ester. Inferring cancer subnetwork markers using density-constrained biclustering. *Bioinformatics*, 26(18):i625–i631, 2010.
- [155] Ranajit Das, Sushmita Mitra, Haider Banka, and Subhasis Mukhopadhyay. Evolutionary biclustering with correlation for gene interaction networks. In Ashish Ghosh, RajatK. De, and SankarK. Pal, editors, *Pattern Recognition and Machine Intelligence*, volume 4815 of *Lecture Notes in Computer Science*, pages 416–424. Springer Berlin Heidelberg, 2007.
- [156] Ranajit Das, Sushmita Mitra, and Subhasis Mukhopadhyay. Cross-correlation and evolutionary biclustering: Extracting gene interaction sub-networks. In Santanu Chaudhury, Sushmita Mitra, C.A. Murthy, P.S. Sastry, and SankarK. Pal, editors, *Pattern Recognition and Machine Intelligence*, volume 5909 of *Lecture Notes in Computer Science*, pages 199–204. Springer Berlin Heidelberg, 2009.
- [157] Ranajit Das, Sushmita Mitra, C.A. Murthy, and Subhasis Mukhopadhyay. A least squares fitting-based modeling of gene regulatory sub-networks. In Santanu Chaudhury, Sushmita Mitra, C.A. Murthy, P.S. Sastry, and SankarK. Pal, editors, *Pattern Recognition and Machine Intelligence*, volume 5909 of *Lecture Notes in Computer Science*, pages 165–170. Springer Berlin Heidelberg, 2009.
- [158] Aynur Dayanik, David D. Lewis, David Madigan, Vladimir Menkov, and Alexander Genkin. Constructing informative prior distributions from domain knowledge in text classification. In *Proceedings of the 29th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval*, SIGIR '06, pages 493–500, New York, NY, USA, 2006. ACM.
- [159] P.A.D. de Castro, F.O. de Franga, H.M. Ferreira, and F.J. Von Zuben. Applying biclustering to perform collaborative filtering. In *Intelligent Systems Design and Applications*, pages 421–426, Oct 2007.
- [160] Fabrício Olivetti de França, Guilherme Palermo Coelho, and Fernando J Von Zuben. Predicting missing values with biclustering: A coherence-based approach. *Pattern Recognition*, 46(5):1255–1266, 2013.
- [161] F.O. de Franga and F.J. Von Zuben. Extracting additive and multiplicative coherent biclusters with swarm intelligence. In *Evolutionary Computation (CEC), 2011 IEEE Congress on*, pages 632–638, June 2011.
- [162] Michiel De Hoon, Seiya Imoto, Kazuo Kobayashi, Naotake Ogasawara, and Satoru Miyano. Inferring gene regulatory networks from time-ordered gene expression data of bacillus subtilis using differential equations. In *Biocomputing 2003: Proc. Pacific Symposium*, volume 8, pages 17–28, 2002.
- [163] Michiel De Hoon, Seiya Imoto, Kazuo Kobayashi, Naotake Ogasawara, and Satoru Miyano. Inferring gene regulatory networks from time-ordered gene expression data of bacillus subtilis using differential equations. In *Biocomputing 2003: Proc. Pacific Symposium*, volume 8, pages 17–28, 2003.
- [164] Hidde De Jong. Modeling and simulation of genetic regulatory systems: a literature review. *Journal of computational biology*, 9(1):67–103, 2002.
- [165] M.C.P. de Souto, D.S.A. de Araujo, I.G. Costa, R. Soares, T.B. Ludermir, and A. Schliep. Comparative study on normalization procedures for cluster analysis of gene expression datasets. In *IJCNN*, pages 2792–2798. IEEE, june 2008.
- [166] Janez Demšar. Statistical comparisons of classifiers over multiple data sets. *J. Machine Learning Res.*, 7:1–30, December 2006.
- [167] G. Deng. *Simulation-based optimization*. University of Wisconsin–Madison, 2007.
- [168] Zhaohong Deng, Kup-Sze Choi, Fu-Lai Chung, and Shitong Wang. Enhanced soft subspace clustering integrating within-cluster and between-cluster information. *Pattern Recognition*, 43(3):767–781, 2010.
- [169] Patrik D'haeseleer, Shoudan Liang, and Roland Somogyi. Genetic network inference: from co-expression clustering to reverse engineering. *Bioinformatics*, 16(8):707–726, 2000.
- [170] Patrik D'haeseleer, Xiling Wen, Stefanie Fuhrman, Roland Somogyi, et al. Linear modeling of mrna expression levels during cns development and injury. In *Pacific symposium on biocomputing*, volume 4, pages 41–52, 1999.
- [171] Inderjit S Dhillon, Subramanyam Mallela, and Rahul Kumar. A divisive information theoretic feature clustering algorithm for text classification. *The Journal of Machine Learning Research*, 3:1265–1287, 2003.
- [172] Inderjit S. Dhillon, Subramanyam Mallela, and Dharmendra S. Modha. Information-theoretic co-clustering. In *Proceedings of the ninth ACM SIGKDD international conference on Knowledge discovery and data mining*, KDD '03, pages 89–98, New York, NY, USA, 2003. ACM.
- [173] Tom Dietterich. Overfitting and undercomputing in machine learning. *ACM computing surveys (CSUR)*, 27(3):326–327, 1995.
- [174] C. Ding, Ya Zhang, Tao Li, and S.R. Holbrook. Biclustering protein complex interactions with a biclique finding algorithm. In *ICDM*, pages 178–187, 2006.
- [175] Chris Ding, Ya Zhang, Tao Li, and Stephen R. Holbrook. Biclustering protein complex interactions with a biclique finding algorithm. In *ICDM*, pages 178–187, Washington, DC, USA, 2006. IEEE Computer Society.
- [176] Hui Ding, Goce Trajcevski, Peter Scheuermann, Xiaoyue Wang, and Eamonn J. Keogh. Querying and mining of time series data: experimental comparison of representations and distance measures. *Proceedings of the VLDB Endowment*, 1(2):1542–1552, 2008.
- [177] Marcus T. Ditttrich, Gunnar W. Klau, Andreas Rosenwald, Thomas Dandekar, and Tobias Müller. Identifying functional modules in protein–protein interaction networks: an integrated exact approach. *Bioinformatics*, 24(13):i223–i231, 2008.
- [178] Kevin Dobbin and Richard Simon. Sample size determination in microarray experiments for class comparison and prognostic classification. *Biostatistics*, 6(1):27+.
- [179] Kevin Dobbin and Richard Simon. Sample size determination in microarray experiments for class comparison and prognostic classification. *Biostatistics*, 6(1):27+, 2005.
- [180] Kevin K. Dobbin and Richard M. Simon. Sample size planning for developing classifiers using high-dimensional DNA microarray data. *Biostatistics*, 8(1):101–117, January 2007.
- [181] Sara Dolnicar, Sebastian Kaiser, Katie Lazarevski, and Friedrich Leisch. Biclustering: Overcoming data dimensionality problems in market segmentation. *Journal of Travel Research*, 51(1):41–49, 2012.

- [182] Pedro Domingos. A unified bias-variance decomposition and its applications. In *IC on Machine Learning*, pages 231–238. Morgan Kaufmann, 2000.
- [183] A. Donders, G. van der Heijden, T. Stijnen, and K. Moons. Review: a gentle introduction to imputation of missing values. *Clinical epidemiology*, 59(10):1087–91, 2006.
- [184] Guozhu Dong and Jinyan Li. Efficient mining of emerging patterns: Discovering trends and differences. In *KDD*, pages 43–52. ACM, 1999.
- [185] Marie-Therese Doolin, Anthony L Johnson, Leland H Johnston, and Geraldine Butler. Overlapping and distinct roles of the duplicated yeast transcription factors ace2p and swi5p. *Molecular microbiology*, 40(2):422–432, 2001.
- [186] E. R. Dougherty, C. Sima, J. Hua, B. Hanczar, and U. M. Braga-Neto. Performance of error estimators for classification. *Current Bioinformatics*, 5(1):53–67, 2010.
- [187] L. Duan, W. N. Street, and E. Xu. Healthcare information systems: data mining methods in the creation of a clinical recommender system. *Enterprise Information Systems*, 5(2):169–181, 2011.
- [188] Florence Duchêne, Catherine Garbay, and Vincent Rialle. Learning recurrent behaviors from heterogeneous multivariate time-series. *Artificial intelligence in medicine*, 39(1):25–47, 2007.
- [189] W. DuMouchel and D. Pregibon. Empirical bayes screening for multi-item associations. In *KDD*, pages 67–76. ACM, 2001.
- [190] William DuMouchel. Bayesian Data Mining in Large Frequency Tables, with an Application to the FDA Spontaneous Reporting System. *The American Statistician*, 53(3), 1999.
- [191] M.H. Dunham. *Data Mining: Introductory And Advanced Topics*. Pearson Education, 2006.
- [192] Pierre Dupont, Jérôme Callut, Grégoire Doods, Jean-Noël Monette, Yves Deville, and BP Sainte. Relevant subgraph extraction from random walks in a graph. *Universite catholique de Louvain, UCL/INGI, Number RR, 7*, 2006.
- [193] Pierre Dupont, François Denis, and Yann Esposito. Links between probabilistic automata and hidden markov models: probability distributions, learning models and induction algorithms. *Pattern recognition*, 38(9):1349–1371, 2005.
- [194] Eran Eden, Roy Navon, Israel Steinfeld, Doron Lipson, and Zohar Yakhini. Gorilla: a tool for discovery and visualization of enriched go terms in ranked gene lists. *BMC Bioinformatics*, 10(1):48, 2009.
- [195] Michael Eichler. Granger causality and path diagrams for multivariate time series. *Journal of Econometrics*, 137(2):334–353, 2007.
- [196] Michael Eichler. Graphical modelling of multivariate time series. *Probability Theory and Related Fields*, 153(1-2):233–268, 2012.
- [197] Magdalini Eirinaki and Michalis Vazirgiannis. Web mining for web personalization. *ACM Transactions on Internet Technology (TOIT)*, 3(1):1–27, 2003.
- [198] M. Eisen, P. Spellman, P. Brown, and D. Botstein. Cluster analysis and display of genome-wide expression patterns. *National Academy of Sciences*, 95(25):14863–14868, 1998.
- [199] M.A. El-Gamal, H.L. Abdel-Malek, and M.A. Sorour. Automatic circuit tuning using unsupervised learning procedures. In *Circuits and Systems, 2003 IEEE 46th Midwest Symposium on*, volume 1, pages 125–128 Vol. 1, dec. 2003.
- [200] T. S. El-Sheikh and A. G. Wacker. Effect of dimensionality and estimation on the performance of gaussian classifiers. *Pattern Recognition*, 12(3):115–126, 1980.
- [201] E. Elhamifar and R. Vidal. Sparse subspace clustering. In *Computer Vision and Pattern Recognition (CVPR)*, pages 2790–2797, June 2009.
- [202] Kemal Eren, Mehmet Deveci, Onur Küçüktunç, and Ümit V Çatalyürek. A comparative analysis of biclustering algorithms for gene expression data. *Briefings in bioinformatics*, 14(3):279–292, 2013.
- [203] G.J. Escobar, J.D. Greene, P. Scheirer, M.N. Gardner, D. Draper, and P. Kipnis. Risk-adjusting hospital inpatient mortality using automated inpatient, outpatient, and laboratory databases. *Medical Care*, 46(3):232–239, 2008.
- [204] Themis P. Exarchos, Markos G. Tsipouras, Costas Papaloukas, and Dimitrios I. Fotiadis. A two-stage methodology for sequence classification based on sequential pattern mining and optimization. *Data Knowl. Eng.*, 66(3):467–487, September 2008.
- [205] Neng Fan, Nikita Boyko, and Panos M Pardalos. Recent advances of data biclustering with application in computational neuroscience. In *Computational Neuroscience*, pages 85–112. Springer, 2010.
- [206] Wei Fan, Kun Zhang, Hong Cheng, Jing Gao, Xifeng Yan, Jiawei Han, Philip Yu, and Olivier Verscheure. Direct mining of discriminative and essential frequent patterns via model-based search tree. In *KDD*, pages 230–238. ACM, 2008.
- [207] Gang Fang, Majda Haznadar, Wen Wang, Haoyu Yu, Michael Steinbach, Timothy R. Church, William S. Oetting, Brian Van Ness, and Vipin Kumar. High-Order SNP Combinations Associated with Complex Diseases: Efficient Discovery, Statistical Power and Functional Interactions. *Plos One*, 7, 2012.
- [208] Gang Fang, Rui Kuang, Gaurav Pandey, Michael Steinbach, Chad L. Myers, and Vipin Kumar. Subspace differential coexpression analysis: Problem definition and a general approach. In *Pacific Symposium on Biocomputing*, pages 145–156. World Scientific Publishing, 2010.
- [209] Qiong Fang, Wilfred Ng, Jianlin Feng, and Yuliang Li. Mining order-preserving submatrices from probabilistic matrices. *ACM Trans. Database Syst.*, 39(1):6:1–6:43, January 2014.
- [210] F. Farhadi, E. Hoseini, S. Hashemi, and A. Hamzeh. Teamfinder: A co-clustering based framework for finding an effective team of experts in social networks. In *Data Mining Workshops (ICDMW), 2012 IEEE 12th International Conference on*, pages 107–114, Dec 2012.
- [211] Paolo Favaro, René Vidal, and Avinash Ravichandran. A closed form solution to robust subspace estimation and clustering. In *Computer Vision and Pattern Recognition (CVPR)*, pages 1801–1807. IEEE, 2011.
- [212] Usama M. Fayyad and Keki B. Irani. Multi-interval discretization of continuous-valued attributes for classification learning. In *IJCAI*, pages 1022–1029, 1993.
- [213] Mario Figueiredo. Feature selection for supervised and unsupervised learning, 2004.
- [214] Mário AT Figueiredo. Lecture notes on bayesian estimation and classification, 2004.
- [215] Mário AT Figueiredo and Anil K Jain. Bayesian learning of sparse classifiers. In *Computer Vision and Pattern Recognition, 2001. CVPR 2001. Proceedings of the 2001 IEEE Computer Society Conference on*, volume 1, pages 1–35. IEEE, 2001.
- [216] Mário AT Figueiredo, Anil K Jain, and Martin H Law. A feature selection wrapper for mixtures. In *Pattern Recognition and Image Analysis*, pages 229–237. Springer, 2003.
- [217] Mário AT Figueiredo and Robert D Nowak. Wavelet-based image estimation: an empirical bayes approach using jeffrey’s noninformative prior. *Image Processing, IEEE Transactions on*, 10(9):1322–1331, 2001.
- [218] Rosa L. Figueroa, Qing Zeng-Treitler, Sasikiran Kandula, and Long H. Ngo. Predicting sample size required for classification performance. *BMC Med. Inf. & Decision Making*, 12:8, 2012.
- [219] Rosa L. Figueroa, Qing Zeng-Treitler, Sasikiran Kandula, and Long H. Ngo. Predicting sample size required for classification performance. *BMC Med. Inf. & Decision Making*, 12:8, 2012.
- [220] Philippe Flajolet, Wojciech Szpankowski, and Brigitte Vallée. Hidden word statistics. *Journal of the ACM (JACM)*, 53(1):147–183, 2006.
- [221] Joseph L. Fleiss. *Statistical Methods for Rates and Proportions*. Wiley P. in Applied Statistics. Wiley, 1981.

- [222] Philippe Fournier-Viger, Antonio Gomariz, Manuel Campos, and Rincy Thomas. Fast vertical mining of sequential patterns using co-occurrence information. In Vincent S. Tseng, TuBao Ho, Zhi-Hua Zhou, Arbee L. P. Chen, and Hung-Yu Kao, editors, *Advances in Knowledge Discovery and Data Mining*, volume 8443 of *Lecture Notes in Computer Science*, pages 40–52. Springer International Publishing, 2014.
- [223] Philippe Fournier-Viger, Antonio Gomariz, Ted Gueniche, Espérance Mwamikazi, and Rincy Thomas. Tks: Efficient mining of top-k sequential patterns. In *Advanced Data Mining and Applications*, pages 109–120. Springer, 2013.
- [224] Philippe Fournier-Viger, Antonio Gomariz, Ted Gueniche, Espérance Mwamikazi, and Rincy Thomas. Tks: Efficient mining of top-k sequential patterns. In *Advanced Data Mining and Applications*, pages 109–120. Springer, 2013.
- [225] Philippe Fournier-Viger, Ted Gueniche, Souleymane Zida, and Vincent S Tseng. Erminer: Sequential rule mining using equivalence classes. In *Advances in Intelligent Data Analysis XIII*, pages 108–119. Springer, 2014.
- [226] Philippe Fournier-Viger, Roger Nkambou, and Vincent Shin-Mu Tseng. Rulegrowth: mining sequential rules common to several sequences by pattern-growth. In *Proceedings of the 2011 ACM symposium on applied computing*, pages 956–961. ACM, 2011.
- [227] Philippe Fournier-Viger and Vincent S. Tseng. Tns: Mining top-k non-redundant sequential rules. In *Proceedings of the 28th Annual ACM Symposium on Applied Computing*, SAC '13, pages 164–166, New York, NY, USA, 2013. ACM.
- [228] Philippe Fournier-Viger and Vincent S. Tseng. Mining top-k non-redundant association rules. In Li Chen, Alexander Felfernig, Jiming Liu, and Zbigniew W. Ras, editors, *Foundations of Intelligent Systems*, volume 7661 of *Lecture Notes in Computer Science*, pages 31–40. Springer Berlin Heidelberg, 2012.
- [229] Eibe Frank and Ian H. Witten. Generating accurate rule sets without global optimization. In J. Shavlik, editor, *Machine Learning*, pages 144–151. Morgan Kaufmann, 1998.
- [230] Adelaide Freitas, Wassim Ayadi, Mourad Elloumi, José Luis, and Jin-Kao Hao Oliveira. Survey on biclustering of gene expression data. *Biological Knowledge Discovery Handbook: Preprocessing, Mining, and Postprocessing of Biological Data*, pages 591–608, 2012.
- [231] Nir Friedman. Inferring cellular networks using probabilistic graphical models. *Science*, 303(5659):799–805, 2004.
- [232] Wenjiang J Fu. Penalized regressions: the bridge versus the lasso. *Journal of computational and graphical statistics*, 7(3):397–416, 1998.
- [233] Yukiko Fujiwara, Minoru Asogawa, and Akihiko Konagaya. Stochastic motif extraction using hidden markov model. In *ISMB*, pages 121–129. AAAI, 1994.
- [234] Guojun Gan and Jianhong Wu. A convergence theorem for the fuzzy subspace clustering (fsc) algorithm. *Pattern Recognition*, 41(6):1939–1947, 2008.
- [235] Xiangchao Gan, Alan Liew, and Hong Yan. Discovering biclusters in gene expression data based on high-dimensional linear geometries. *BMC Bioinformatics*, 9(1):209, 2008.
- [236] Salvador Garcia and Francisco Herrera. An extension on "statistical comparisons of classifiers over multiple data sets" for all pairwise comparisons. *Journal of Machine Learning Research*, 9:2677–2694, 2009.
- [237] Minos N Garofalakis, Rajeev Rastogi, and Kyuseok Shim. Spirit: Sequential pattern mining with regular expression constraints. In *VLDB*, volume 99, pages 7–10, 1999.
- [238] Audrey P Gasch, Paul T Spellman, Camilla M Kao, Orna Carmel-Harel, Michael B Eisen, Gisela Storz, David Botstein, and Patrick O Brown. Genomic expression programs in the response of yeast cells to environmental changes. *Molecular biology of the cell*, 11(12):4241–4257, 2000.
- [239] Audrey P. Gasch, Paul T. Spellman, Camilla M. Kao, Orna Carmel-Harel, Michael B. Eisen, Gisela Storz, David Botstein, and Patrick O. Brown. Genomic expression programs in the response of yeast cells to environmental changes. *Molecular Biology of the Cell*, 11(12):4241–4257, 2000.
- [240] Audrey P Gasch and Margaret Werner-Washburne. The genomics of yeast responses to environmental stress and starvation. *Functional & integrative genomics*, 2(4-5):181–192, 2002.
- [241] Xianping Ge and Padhraic Smyth. Deformable markov model templates for time-series pattern matching. In *ACM SIGKDD*, pages 81–90, New York, NY, USA, 2000. ACM.
- [242] Floris Geerts, Bart Goethals, and Jan Van Den Bussche. Tight upper bounds on the number of candidate patterns. *ACM Transactions on Database Systems (TODS)*, 30(2):333–363, 2005.
- [243] Elisabeth Georgii, Sabine Dietmann, Takeaki Uno, Philipp Pagel, and Koji Tsuda. Enumeration of condition-dependent dense modules in protein interaction networks. *Bioinformatics*, 25(7):933–940, 2009.
- [244] Elisabeth Georgii, Lothar Richter, Ulrich Rückert, and Stefan Kramer. Analyzing microarray data using quantitative association rules. *Bioinformatics*, 21(2):123–129, January 2005.
- [245] Lise Getoor. Link mining: A new data mining challenge. *SIGKDD Explor. Newsl.*, 5(1):84–89, July 2003.
- [246] Lise Getoor and Christopher P. Diehl. Link mining: A survey. *SIGKDD Explor. Newsl.*, 7(2):3–12, December 2005.
- [247] Gad Getz, Erel Levine, and Eytan Domany. Coupled two-way clustering analysis of gene microarray data. *Proceedings of the National Academy of Sciences*, 97(22):12079–12084, 2000.
- [248] Pierre Geurts. Pattern extraction for time series classification. In *Principles of Data Mining and Knowl. Disc.*, volume 2168 of *LNCS*, pages 115–127. Springer, 2001.
- [249] Zoubin Ghahramani. Learning dynamic bayesian networks. In *Adaptive processing of sequences and data structures*, pages 168–197. Springer, 1998.
- [250] Zoubin Ghahramani. Learning dynamic bayesian networks. In *Adaptive processing of sequences and data structures*, pages 168–197. Springer, 1998.
- [251] Debashis Ghosh and Arul M Chinnaiyan. Mixture modelling of gene expression data from microarray experiments. *Bioinformatics*, 18(2):275–286, 2002.
- [252] A. Gionis, H. Mannila, T. Mielikäinen, and P. Tsaparas. Assessing data mining results via swap randomization. *ACM Trans. Knowl. Discov. Data*, 1(3), December 2007.
- [253] Aristides Gionis, Heikki Mannila, Taneli Mielikäinen, and Panayiotis Tsaparas. Assessing data mining results via swap randomization. *ACM Transactions on Knowledge Discovery from Data (TKDD)*, 1(3):14, 2007.
- [254] Federico Girosi. An equivalence between sparse approximation and support vector machines. *Neural computation*, 10(6):1455–1480, 1998.
- [255] Michelle Girvan and Mark EJ Newman. Community structure in social and biological networks. *National Academy of Sciences*, 99(12):7821–7826, 2002.
- [256] Ned Glick. Additive estimators for probabilities of correct classification. *Pattern Recognition*, 10(3):211 – 222, 1978.
- [257] Dmitry Gnatyshak, Dmitry I. Ignatov, Alexander Semenov, and Jonas Poelmans. Gaining insight in social networks with biclustering and triclustering. In *Perspectives in Business Informatics Research*, volume 128 of *Lecture Notes in Business Info. Processing*, pages 162–171. Springer Berlin Heidelberg, 2012.

- [258] Dmitry Gnatyshak, Dmitry I. Ignatov, Alexander Semenov, and Jonas Poelmans. Gaining insight in social networks with biclustering and triclustering. In Natalia Aseeva, Eduard Babkin, and Oleg Kozyrev, editors, *Perspectives in Business Informatics Research*, volume 128 of *Lecture Notes in Business Information Processing*, pages 162–171. Springer Berlin Heidelberg, 2012.
- [259] Antonio Gomariz, Manuel Campos, Roque Marin, and Bart Goethals. Clasp: An efficient algorithm for mining frequent closed sequences. In Jian Pei, Vincent S. Tseng, Longbing Cao, Hiroshi Motoda, and Guandong Xu, editors, *Advances in Knowledge Discovery and Data Mining*, volume 7818 of *Lecture Notes in Computer Science*, pages 50–61. Springer Berlin Heidelberg, 2013.
- [260] Joana Gonçalves. *Integrative mining of gene regulation and its perturbations*. PhD thesis, Instituto Superior Técnico, Universidade de Lisboa, Lisboa, 2013.
- [261] Joana P Gonçalves, Ricardo S Aires, Alexandre P Francisco, and Sara C Madeira. Regulatory snapshots: integrative mining of regulatory modules from expression time series and regulatory networks. *PLoS one*, 7(5):e35977, 2012.
- [262] Joana P Gonçalves and Sara C Madeira. Heuristic approaches for time-lagged biclustering. In *Proceedings of the 12th International Workshop on Data Mining in Bioinformatics*, pages 1–9. ACM, 2013.
- [263] JP Gonçalves and SC Madeira. Latebiclustering: Efficient heuristic algorithm for time-lagged bicluster identification. *IEEE/ACM Transactions on Computational Biology and Bioinformatics*, pages doi-10, 2014.
- [264] A. D. Gordon. 359. note: Classification in the presence of constraints. *Biometrics*, 29(4):pp. 821–827, 1973.
- [265] Osamu Gotoh. Multiple sequence alignment: algorithms and applications. *Advances in biophysics*, 36:159–206, 1999.
- [266] Anuj Goyal, Reede Ren, and Joemon M Jose. Feature subspace selection for efficient video retrieval. In *Advances in Multimedia Modeling*, pages 725–730. Springer, 2010.
- [267] Gösta Grahne and Jianfei Zhu. Efficiently using prefix-trees in mining frequent itemsets. In Bart Goethals and Mohammed Javeed Zaki, editors, *FIMI*, volume 90 of *CEUR Workshop Proceedings*. CEUR-WS.org, 2003.
- [268] Gösta Grahne and Jianfei Zhu. Efficiently using prefix-trees in mining frequent itemsets. In *FIMI*, volume 90, 2003.
- [269] A. Graves. *Supervised Sequence Labelling with Recurrent Neural Networks*. Studies in Computational Intelligence. Springer, 2012.
- [270] Jiajun Gu and Jun Liu. Bayesian biclustering of gene expression data. *BMC Genomics*, 9(Suppl 1):S4, 2008.
- [271] Israel Guerra, Loïc Cerf, João Foscarini, Michel Boaventura, and Wagner Meira. Constraint-based search of straddling biclusters and discriminative patterns. *JIDM*, 4(2):114–123, 2013.
- [272] Gabriela Guimarães. The induction of temporal grammatical rules from multivariate time series. In *Proceedings of the 5th Int. Colloquium on Grammatical Inference: Algorithms and Applications*, pages 127–140, London, UK, 2000. Springer-Verlag.
- [273] Stephan Günemann, Ines Färber, Emmanuel Müller, Ira Assent, and Thomas Seidl. External evaluation measures for subspace clustering. In *Proceedings of the 20th ACM International Conference on Information and Knowledge Management, CIKM '11*, pages 1363–1372, New York, NY, USA, 2011. ACM.
- [274] Tias Guns, Siegfried Nijssen, and Luc de Raedt. k-pattern set mining under constraints. *IEEE Transactions on Knowledge and Data Engineering*, 99(PrePrints), 2011.
- [275] Tias Guns, Siegfried Nijssen, and Luc De Raedt. Itemset mining: A constraint programming perspective. *Artificial Intelligence*, 175(12-13):1951–1983, 2011.
- [276] Yu Guo, Armin Graber, Robert N McBurney, and Raji Balasubramanian. Sample size and statistical power considerations in high-dimensionality data settings: a comparative study of classification algorithms. *BMC Bioinformatics*, 11(1):1–19, 2010.
- [277] Rohit Gupta, Navneet Rao, and Vipin Kumar. Discovery of error-tolerant biclusters from noisy gene expression data. *BMC Bioinformatics*, 12(12):1–17, 2011.
- [278] Valery Guralnik, Duminda Wijesekera, and Jaideep Srivastava. Pattern directed mining of sequence data. In *ACM SIGKDD*, pages 51–57, 1998.
- [279] Dan Gusfield. *Algorithms on strings, trees and sequences: computer science and computational biology*. Cambridge university press, 1997.
- [280] D. Gutierrez-Aviles and C. Rubio-Escudero. Lsl: A new measure to evaluate triclusters. In *Bioinformatics and Biomedicine (BIBM), 2014 IEEE International Conference on*, pages 30–37, Nov 2014.
- [281] David Gutiérrez-Avilés and Cristina Rubio-Escudero. Mining 3d patterns from gene expression temporal data: A new tricluster evaluation measure. *The Scientific World Journal*, 2014, 2014.
- [282] David Gutiérrez-Avilés, Cristina Rubio-Escudero, Francisco Martínez-Álvarez, and José C Riquelme. Trigen: A genetic algorithm to mine triclusters in temporal gene expression data. *Neurocomputing*, 132:42–53, 2014.
- [283] Isabelle Guyon, John Markhoul, Richard Schwartz, and Vladimir Vapnik. What size test set gives good error rate estimates? *IEEE Trans. Pattern Anal. Mach. Intell.*, 20(1):52–64, January 1998.
- [284] Robert Gwadera and Fabio Crestani. Ranking sequential patterns with respect to significance. In *Advances in Knowledge Discovery and Data Mining*, pages 286–299. Springer, 2010.
- [285] Andreas Haag, Silke Goronzy, Peter Schaich, and Jason Williams. Emotion recognition using bio-sensors: First steps towards an automatic system. In *Affective Dialogue Systems*, volume 3068 of *LNCS*, pages 36–48. Springer Berlin-Heidelberg, 2004.
- [286] Mark Hall, Eibe Frank, Geoffrey Holmes, Bernhard Pfahringer, Peter Reutemann, and Ian H. Witten. The weka data mining software: an update. *SIGKDD Explor. Newsl.*, 11(1):10–18, 2009.
- [287] Mark Hall, Eibe Frank, Geoffrey Holmes, Bernhard Pfahringer, Peter Reutemann, and Ian H. Witten. The weka data mining software: an update. *SIGKDD Explor. Newsl.*, 11(1):10–18, 2009.
- [288] W. Hämmäläinen and M. Nykänen. Efficient discovery of statistically significant association rules. In *ICDM*, pages 203–212. IEEE CS, 2008.
- [289] E.-H. Han, G. Karypis, and V. Kumar. Min-apriori: An algorithm for finding association rules in data with continuous attributes, 1997.
- [290] Jiawei Han, Hong Cheng, Dong Xin, and Xifeng Yan. Frequent pattern mining: current status and future directions. *Data Min. Knowl. Discov.*, 15(1):55–86, August 2007.
- [291] Jiawei Han, Jian Pei, Guozhu Dong, and Ke Wang. Efficient computation of iceberg cubes with complex measures. *SIGMOD Rec.*, 30(2):1–12, May 2001.
- [292] Jiawei Han, Jian Pei, and Yiwen Yin. Mining frequent patterns without candidate generation. *SIGMOD Rec.*, 29(2):1–12, 2000.
- [293] Jiawei Han, Qiang Yang, and Edward Kim. Plan mining by divide-and-conquer. In *ACM SIGMOD IW on Research Issues in DMKD*, 1999.
- [294] Lixin Han and Hong Yan. Bsn: An automatic generation algorithm of social network data. *Journal of Systems and Software*, 84(8):1261–1269, 2011.
- [295] Blaise Hanczar and Mohamed Nadif. Ensemble methods for biclustering tasks. *Pattern Recognition*, 45(11):3938–3949, 2012.
- [296] D J Hand. Recent advances in error rate estimation. *Pattern Recogn. Lett.*, 4(5):335–346, November 1986.
- [297] Daniel Hanisch, Alexander Zien, Ralf Zimmer, and Thomas Lengauer. Co-clustering of biological networks and gene expression data. *Bioinformatics*, 18(suppl 1):S145–S154, 2002.
- [298] J. A. Hartigan. Direct Clustering of a Data Matrix. *Journal of the American Statistical Association*, 67(337):123–129, 1972.
- [299] J. A. Hartigan and M. A. Wong. A k-means clustering algorithm. *JSTOR: Applied Statistics*, 28(1):100–108, 1979.

- [300] David Haussler, Michael Kearns, and Robert Schapire. Bounds on the sample complexity of bayesian learning using information theory and the vc dimension. In *Proceedings of the fourth annual workshop on Computational learning theory*, COLT '91, pages 61–74, San Francisco, CA, USA, 1991. Morgan Kaufmann Publishers Inc.
- [301] Trond Helleem, Bjarte Dysvik, and Inge Jonassen. LSimpute: accurate estimation of missing values in microarray data with least squares methods. *Nucleic acids research*, 32(3):34+, February 2004.
- [302] R. Henriques and C. Antunes. Learning predictive models from integrated healthcare data: Extending pattern-based and generative models to capture temporal and cross-attribute dependencies. In *System Sciences (HICSS), 2014 47th Hawaii International Conference on*, pages 2562–2569, Jan 2014.
- [303] R. Henriques and S. Madeira. Biclustering with flexible plaid models to unravel interactions between biological processes. *Computational Biology and Bioinformatics, IEEE/ACM Transactions on*, 2015.
- [304] R. Henriques, A. Paiva, and C. Antunes. Accessing emotion patterns from affective interactions using electrodermal activity. In *Affective Computing and Intelligent Interaction (ACII), 2013 Humaine Association Conference on*, pages 43–48, Sept 2013.
- [305] Rui Henriques, Cláudia Antunes, and Sara C. Madeira. A structured view on pattern mining-based biclustering. *Pattern Recognition*, 2015.
- [306] Rui Henriques and Cláudia Antunes. On the need of new approaches for the novel problem of long-term prediction over multi-dimensional data. In Roger Lee, editor, *Computer and Information Science 2012*, volume 429 of *Studies in Computational Intelligence*, pages 121–138. Springer Berlin Heidelberg, 2012.
- [307] Rui Henriques and Cláudia Antunes. Generative modeling of itemset sequences derived from real databases. In *IC on Enterprise Information Systems (ICEIS'14)*, pages 264–272, Lisbon, Portugal, 2014. INSTICC.
- [308] Rui Henriques, Cláudia Antunes, and Sara Madeira. Methods for the efficient discovery of large item-indexable sequential patterns. *LNAI*, 7765, 2014.
- [309] Rui Henriques, Cláudia Antunes, and SaraC. Madeira. Generative modeling of repositories of health records for predictive tasks. *Data Mining and Knowledge Discovery*, 29(4):999–1032, 2015.
- [310] Rui Henriques and Sara Madeira. Bicpam: Pattern-based biclustering for biomedical data analysis. *Algorithms for Molecular Biology*, 9(1):27, 2014.
- [311] Rui Henriques and Sara Madeira. Bicspam: Flexible biclustering using sequential patterns. *BMC Bioinformatics*, 15:130, 2014.
- [312] Rui Henriques and Sara C. Madeira. Bip: Effective discovery of overlapping biclusters using flexible plaid models. In *13th IW of Data Mining in Bioinformatics (BIOKDD'14)*, New York, US, 2014. ACM.
- [313] Rui Henriques and Sara C. Madeira. Bicnet: Efficient biclustering of biological networks to unravel non-trivial modules. In *Algorithms in Bioinformatics (WABI), LNCS*. Springer-Verlag, 2015.
- [314] Rui Henriques and Sara C. Madeira. Pattern-based biclustering with constraints for gene expression data analysis. In *Computational Methods in Bioinformatics and Systems Biology (EPIA-CMBSB)*, LNAI. Springer, 2015.
- [315] Rui Henriques, Sara C. Madeira, and Cláudia Antunes. F2g: Efficient discovery of full-patterns. In *ECML/PKDD nfMCP*, Prague, 2013.
- [316] Rui Henriques and SaraC. Madeira. Towards robust performance guarantees for models learned from high-dimensional data. In Aboul Ella Hassanien, Ahmad Taher Azar, Vaclav Snasael, Janusz Kacprzyk, and Jemal H. Abawajy, editors, *Big Data in Complex Systems*, volume 9 of *Studies in Big Data*, pages 71–104. Springer International Publishing, 2015.
- [317] Rui Henriques, SaraC. Madeira, and Cláudia Antunes. Multi-period classification: learning sequent classes from temporal domains. *Data Mining and Knowledge Discovery*, 29(3):792–819, 2015.
- [318] Rui Henriques and Ana Paiva. Learning effective models of emotions from physiological signals: The seven principles. In Hugo Plácido da Silva, Andreas Holzinger, Stephen Fairclough, and Dennis Majoe, editors, *Physiological Computing Systems*, Lecture Notes in Computer Science, pages 137–155. Springer Berlin Heidelberg, 2014.
- [319] Rui Henriques and Ana Paiva. Seven principles to mine flexible behavior from physiological signals for effective emotion recognition and description in affective interactions. In *Physiological Computing Systems (PhyCS'14)*, pages 75–82, Lisbon, Portugal, 2014. SCITEPRESS.
- [320] Rui Henriques, Ana Paiva, and Cláudia Antunes. On the need of new methods to mine electrodermal activity in emotion-centered studies. In Longbing Cao, Yifeng Zeng, AndreasL. Symeonidis, Vladimirl. Gorodetsky, PhilipS. Yu, and MunindarP Singh, editors, *Agents and Data Mining Interaction*, volume 7607 of *Lecture Notes in Computer Science*, pages 203–215. Springer Berlin Heidelberg, 2013.
- [321] Rui Henriques, Silvia Moura Pina, and Cláudia Antunes. Temporal mining of integrated healthcare data: Methods, revealings and implications. In *SDM IW on Data Mining for Medicine and Healthcare*, pages 56–64, Austin, US, 2013. SIAM.
- [322] Philip Hingston. *Using finite state automata for sequence mining*, volume 24. Australian Computer Society, Inc., 2002.
- [323] Dorit S. Hochbaum and Asaf Levin. Approximation algorithms for a minimization variant of the order-preserving submatrices and for biclustering problems. *ACM Trans. Algorithms*, 9(2):19:1–19:12, March 2013.
- [324] Sepp Hochreiter, Ulrich Bodenhofer, Martin Heusel, Andreas Mayr, Andreas Mitterecker, Adetayo Kasim, Tatsiana Khamiakova, Suzy Van Sanden, Dan Lin, Willem Talloen, Luc Bijmens, Hinrich W. H. Göhlmann, Ziv Shkedy, and Djork-Arné Clevert. FABIA: factor analysis for bicluster acquisition. *Bioinformatics*, 26(12):1520–1527, June 2010.
- [325] Sture Holm. A Simple Sequentially Rejective Multiple Test Procedure. *Scandinavian Journal of Stat.*, 6(2):65–70, 1979.
- [326] Petter Holme and Jari Saramäki. Temporal networks. *Physics reports*, 519(3):97–125, 2012.
- [327] Neal S Holter, Amos Maritan, Marek Cieplak, Nina V Fedoroff, and Jayanth R Banavar. Dynamic modeling of gene expression data. *Proceedings of the National Academy of Sciences*, 98(4):1693–1698, 2001.
- [328] Neal S Holter, Amos Maritan, Marek Cieplak, Nina V Fedoroff, and Jayanth R Banavar. Dynamic modeling of gene expression data. *Proceedings of the National Academy of Sciences*, 98(4):1693–1698, 2001.
- [329] Pi-Fuei Hsieh and David Landgrebe. Classification of high dimensional data. *ECE Technical Reports*, page 52, 1998.
- [330] Chih-Wei Hsu, Chih-Chung Chang, Chih-Jen Lin, et al. A practical guide to support vector classification, 2003.
- [331] Bing Hu, Yanping Chen, and Eamonn J Keogh. Time series classification under more realistic assumptions. In *SDM*, pages 578–586. SIAM, 2013.
- [332] Bing Hu, Yanping Chen, and Eamonn J. Keogh. Time series classification under more realistic assumptions. In *SDM*, pages 578–586. SIAM, 2013.
- [333] Jianping Hua, Zixiang Xiong, James Lowey, Edward Suh, and Edward R. Dougherty. Optimal number of features as a function of sample size for various classification rules. *Bioinformatics*, 21(8):1509–1515, April 2005.
- [334] Qinghua Huang. A biclustering technique for mining trading rules in stock markets. In Dehuai Zeng, editor, *Applied Informatics and Communication*, volume 224 of *Communications in Computer and Information Science*, pages 16–24. Springer Berlin Heidelberg, 2011.
- [335] Qinghua Huang. A biclustering technique for mining trading rules in stock markets. In Dehuai Zeng, editor, *Applied Informatics and Communication*, volume 224 of *Communications in Comp. and Inf. Science*, pages 16–24. Springer Berlin Heidelberg, 2011.
- [336] Shoudong Huang and Gamini Dissanayake. Convergence and consistency analysis for extended kalman filter based slam. *Robotics, IEEE Transactions on*, 23(5):1036–1049, 2007.

- [337] Sui Huang. Gene expression profiling, genetic networks, and cellular states: an integrating concept for tumorigenesis and drug discovery. *Journal of Molecular Medicine*, 77(6):469–480, 1999.
- [338] Yaochun Huang, Hui Xiong, Weili Wu, and Sam Y. Sung. Mining quantitative maximal hyperclique patterns: a summary of results. In *Proceedings of the 10th Pacific-Asia conference on Advances in Knowledge Discovery and Data Mining*, PAKDD'06, pages 552–556, Berlin, Heidelberg, 2006. Springer-Verlag.
- [339] Dirk Husmeier. Sensitivity and specificity of inferring genetic regulatory interactions from microarray experiments with dynamic bayesian networks. *Bioinformatics*, 19(17):2271–2282, 2003.
- [340] Trey Ideker, Owen Ozier, Benno Schwikowski, and Andrew F Siegel. Discovering regulatory and signalling circuits in molecular interaction networks. *Bioinformatics*, 18(suppl 1):S233–S240, 2002.
- [341] Trey Ideker, Owen Ozier, Benno Schwikowski, and Andrew F Siegel. Discovering regulatory and signalling circuits in molecular interaction networks. *Bioinformatics*, 18(suppl 1):S233–S240, 2002.
- [342] Jan Ihmels, Sven Bergmann, and Naama Barkai. Defining transcription modules using large-scale gene expression data. *Bioinformatics*, 20(13):1993–2003, September 2004.
- [343] K. Iswandy and A. Koenig. Towards effective unbiased automated feature selection. In *Hybrid Intelligent Systems*, pages 29–29, 2006.
- [344] Tommi Jaakkola, David Haussler, et al. Exploiting generative models in discriminative classifiers. *Advances in neural information processing systems*, pages 487–493, 1999.
- [345] Qiong Jackson, David Landgrebe, et al. An adaptive classifier design for high-dimensional data analysis with a limited training data set. *Geoscience and Remote Sensing, IEEE Transactions on*, 39(12):2664–2679, 2001.
- [346] Stéphanie Jacquemont, François Jacquenet, and Marc Sebban. Mining probabilistic automata: a statistical view of sequential pattern mining. *Mach. Learn.*, 75(1):91–127, April 2009.
- [347] Stéphanie Jacquemont, François Jacquenet, and Marc Sebban. Mining probabilistic automata: a statistical view of sequential pattern mining. *Machine Learning*, 75(1):91–127, 2009.
- [348] Anil K Jain and B Chandrasekaran. 39 dimensionality and sample size considerations in pattern recognition practice. *Handbook of statistics*, 2:835–855, 1982.
- [349] Anil K. Jain and Richard C. Dubes. *Algorithms for clustering data*. Prentice-Hall, Inc., Upper Saddle River, NJ, USA, 1988.
- [350] Nitin Jain, Jayant Thatté, Thomas Braciale, Klaus Ley, Michael O'Connell, and Jae K. Lee. Local-pooled-error test for identifying differentially expressed genes with a small number of replicated microarrays. *Bioinformatics*, 19(15):1945–1951, 2003.
- [351] S. Jaroszewicz and T. Scheffer. Fast discovery of unexpected patterns in data, relative to a bayesian network. In *KDD*, pages 118–127. ACM, 2005.
- [352] Asif Javed and Ashfaq Khokhar. Frequent pattern mining on message passing multiprocessor systems. *Distributed and Parallel Databases*, 16(3):321–334, 2004.
- [353] David D. Jensen and Paul R. Cohen. Multiple comparisons in induction algorithms. *Mach. Learn.*, 38(3):309–338, March 2000.
- [354] S. Jerritta, M. Murugappan, R. Nagarajan, and K. Wan. Physiological signals based human emotion recognition: a review. In *Signal Processing and its Applications (CSPA), 2011 IEEE 7th International Colloquium on*, pages 410–415, 2011.
- [355] Ji, Hao, Reyhani, and Lendasse. Direct and recursive prediction of time series using mutual information selection. In Joan Cabestany, Alberto Prieto, and Francisco Sandoval Hernández, editors, *IWANN*, volume 3512 of *Lecture Notes in Computer Science*, pages 1010–1017. Springer, 2005.
- [356] George H. John and Pat Langley. Estimating continuous distributions in bayesian classifiers. In *Proceedings of the Eleventh conference on Uncertainty in artificial intelligence, UAI'95*, pages 338–345, San Francisco, CA, USA, 1995. Morgan Kaufmann Publishers Inc.
- [357] Laveen Kanal and B. Chandrasekaran. On dimensionality and sample size in statistical pattern classification. *Pattern Recognition*, 3(3):225–234, 1971.
- [358] Zaven A Karian. *Handbook of fitting statistical distributions with R*. CRC, 2010.
- [359] Nikola K Kasabov, Zeke SH Chan, Vishal Jain, Igor Sidorov, and Dimiter S Dimitrov. Gene regulatory network discovery from time-series gene expression data—a computational intelligence approach. In *Neural Information Processing*, pages 1344–1353. Springer, 2004.
- [360] Paul Kellam, Xiaohui Liu, Nigel Martin, Christine Orengo, Stephen Swift, and Allan Tucker. A framework for modelling virus gene expression data. *Intelligent Data Analysis*, 6(3):267–279, 2002.
- [361] Maurice G. Kendall. *Rank correlation methods*. Griffin, London, 1948.
- [362] Eamonn J. Keogh and Michael J. Pazzani. An enhanced representation of time series which allows fast and accurate classification, clustering and relevance feedback. In *KDD*, pages 239–243, 1998.
- [363] Thomas B Kepler and Timothy C Elston. Stochasticity in transcriptional regulation: origins, consequences, and mathematical representations. *Biophysical journal*, 81(6):3116–3136, 2001.
- [364] Mehdi Khiari, Patrice Boizumault, and Bruno Crémilleux. Constraint programming for mining n-ary patterns. In *Principles and practice of constraint programming*, pages 552–567, Berlin, Heidelberg, 2010. Springer-Verlag.
- [365] Sunyong Kim, Seiya Imoto, and Satoru Miyano. Dynamic bayesian network and nonparametric regression for nonlinear modeling of gene networks from time series gene expression data. *Biosystems*, 75(1):57–65, 2004.
- [366] Ralph Kimball and Margy Ross. *The Data Warehouse Toolkit: The Complete Guide to Dimensional Modeling*. John Wiley & Sons, Inc., USA, 2nd edition, 2002.
- [367] A. Kirsch, M. Mitzenmacher, A. Pietracaprina, G. Pucci, E. Upfal, and F. Vandin. An efficient rigorous approach for identifying statistically significant frequent itemsets. In *PODS*, pages 117–126. ACM, 2009.
- [368] Sergey Kirshner. *Modeling of Multivariate Time Series Using Hidden Markov Models*. PhD thesis, California S. Univ. Long Beach, 2005. AAI3164062.
- [369] D Kleinfeld and H Sompolinsky. Associative neural network model for the generation of temporal patterns: Theory and application to central pattern generators. *Biophysical Journal*, 54(6):1039–1051, 1988.
- [370] Yuval Kluger, Ronen Basri, Joseph T. Chang, and Mark Gerstein. Spectral Biclustering of Microarray Data: Coclustering Genes and Conditions. *Genome Research*, 13(4):703–716, 2003.
- [371] Inge Koch and Kanta Naito. Prediction of multivariate responses with a selected number of principal components. *Comput. Stat. Data Anal.*, 54:1791–1807, July 2010.
- [372] Judice L. Y. Koh, Huiming Ding, Michael Costanzo, Anastasia Baryshnikova, Kiana Toufighi, Gary D. Bader, Chad L. Myers, Brenda J. Andrews, and Charles Boone. Drygin: a database of quantitative genetic interaction networks in yeast. *Nucleic Acids Research*, 38(suppl 1):D502–D507, 2010.
- [373] Ron Kohavi. The power of decision tables. In *ECML*, pages 174–189, London, UK, UK, 1995. Springer-Verlag.
- [374] Ron Kohavi and George H. John. Wrappers for feature subset selection. *Artif. Intell.*, 97(1-2):273–324, December 1997.
- [375] Ron Kohavi and David H. Wolpert. Bias plus variance decomposition for zero-one loss functions. In *Machine Learning*, pages 275–283. Morgan Kaufmann Publishers, 1996.

- [376] Mehmet Koyuturk, Wojciech Szpankowski, and Ananth Grama. Biclustering gene-feature matrices for statistically significant dense patterns. In *Proceedings of the 2004 IEEE Computational Systems Bioinformatics Conference*, CSB '04, pages 480–484, Washington, DC, USA, 2004. IEEE Computer Society.
- [377] Hans-Peter Kriegel, Peer Kröger, Jörg Sander, and Arthur Zimek. Density-based clustering. *Wiley Interdisc. Rev.: Data Mining and Knowledge Discovery*, 1(3):231–240, 2011.
- [378] Balaji Krishnapuram, AJ Harterink, Lawrence Carin, and Mario AT Figueiredo. A bayesian approach to joint feature selection and classifier design. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 26(9):1105–1111, 2004.
- [379] Hye-Chung Monica Kum, Susan Paulsen, and Wei Wang. Comparative study of sequential pattern mining models. In *Foundations of Data Mining and Knowledge Discovery*, pages 43–70. Springer, 2005.
- [380] Sun-Yuan Kung, Man-Wai Mak, and Ilias Tagkopoulos. Symmetric and asymmetric multi-modality biclustering analysis for microarray data matrix. *Journal of bioinformatics and computational biology*, 4(02):275–298, 2006.
- [381] Sergei O. Kuznetsov and Jonas Poelmans. Knowledge representation and processing with formal concept analysis. *Wiley Interdisc. Reviews: Data Mining and Knowledge Discovery*, 3(3):200–215, 2013.
- [382] Peter A. Lachenbruch. On expected probabilities of misclassification in discriminant analysis, necessary sample size, and a relation with the multiple correlation coefficient. *Biometrics*, 24(4):823–834, 1968.
- [383] Niels Landwehr, Mark Hall, and Eibe Frank. Logistic model trees. *Machine Learning*, 95(1-2):161–205, 2005.
- [384] Kenneth Lange, David R Hunter, and Ilsoon Yang. Optimization transfer using surrogate objective functions. *Journal of computational and graphical statistics*, 9(1):1–20, 2000.
- [385] W John Langlois, Toshiyasu Sasaoka, Alan R Saltiel, and Jerrold M Olefsky. Negative feedback regulation and desensitization of insulin- and epidermal growth factor-stimulated p21ras activation. *Journal of Biological Chemistry*, 270(43):25320–25323, 1995.
- [386] Julia A Lasserre, Christopher M Bishop, and Thomas P Minka. Principled hybrids of generative and discriminative models. In *Computer Vision and Pattern Recognition, 2006 IEEE Computer Society Conference on*, volume 1, pages 87–94. IEEE, 2006.
- [387] Fabien Lauer and Gerard Bloch. Incorporating prior knowledge in support vector machines for classification: A review. *Neurocomputing*, 71(7-9):1578 – 1594, 2008. Progress in Modeling, Theory, and Application of Computational Intelligence 15th European Symposium on Artificial Neural Networks 2007 15th European Symposium on Artificial Neural Networks 2007.
- [388] Pierre-Alain Laur, Jean-Emile Symphor, Richard Nock, and Pascal Poncelet. Statistical supports for mining sequential patterns and improving the incremental update process on data streams. *Intelligent Data Analysis*, 11(1):29–47, 2007.
- [389] S. Laxman, P. Naldurg, R. Sripada, and R. Venkatesan. Connections between mining frequent itemsets and learning generative models. In *Data Mining, 2007. ICDM 2007. Seventh IEEE International Conference on*, pages 571–576, Oct 2007.
- [390] Srivatsan Laxman, P. Sastry, and K. Unnikrishnan. Discovering frequent episodes and learning hidden markov models: A formal connection. *IEEE TKDE*, 17:1505–1517, 2005.
- [391] Srivatsan Laxman, P Shanti Sastry, and KP Unnikrishnan. Fast algorithms for frequent episode discovery in event sequences. In *Proc. 3rd Workshop on Mining Temporal and Sequential Data*. Citeseer, 2004.
- [392] Srivatsan Laxman, Vikram Tankasali, and Ryen W. White. Stream prediction using a generative model based on frequent episodes in event sequences. In *Proceedings of the 14th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, KDD '08, pages 453–461, New York, NY, USA, 2008. ACM.
- [393] L. Lazzeroni and A. Owen. Plaid models for gene expression data. *Statistica Sinica*, 12:61–86, 2002.
- [394] Alison P Lee, Yuchen Yang, Sydney Brenner, and Byrappa Venkatesh. Tfcones: a database of vertebrate transcription factor-encoding genes and their associated conserved noncoding elements. *BMC genomics*, 8(1):441, 2007.
- [395] William Lee, Desiree Tillo, Nicolas Bray, Randall H. Morse, Ronald W. Davis, Timothy R. Hughes, and Corey Nislow. A high-resolution atlas of nucleosome occupancy in yeast. *Nature Genetics*, 39(10):1235–1244, September 2007.
- [396] Neal Lesh, Mohammed J. Zaki, and Mitsunori Ogihara. Mining features for sequence classification. In *KDD*, pages 342–346, New York, NY, USA, 1999. ACM.
- [397] Charles S. Lessard. *Signal Processing of Random Physiological Signals*. Synthesis Lectures on Biomedical Engineering. Morgan and Claypool Publishers, 2006.
- [398] Benjamin Letham, Cynthia Rudin, and David Madigan. Sequential event prediction. *Machine Learning*, 93(2-3):357–380, 2013.
- [399] S. E. Levinson, L. R. Rabiner, and M. M. Sondhi. An Introduction to the Application of the theory of Probabilistic Functions of a Markov Process to Automatic Speech Recognition. *Bell Sys. Tech. Journal*, 62(4), 1983.
- [400] Ao Li and David Tuck. An effective tri-clustering algorithm combining expression data with gene regulation information. *Gene regulation and systems biology*, 3:49, 2009.
- [401] Wenmin Li, Jiawei Han, and Jian Pei. Cmar: Accurate and efficient classification based on multiple class-association rules. In *ICDM*, pages 369–376. IEEE CS, 2001.
- [402] Zhenmin Li, Zhifeng Chen, Sudarshan M Srinivasan, and Yuanyuan Zhou. C-miner: Mining block correlations in storage systems. In *FAST*, volume 4, pages 173–186, 2004.
- [403] Feng Liang, Sheng Ma, and Joseph L Hellerstein. Discovering fully dependent patterns. In *SDM*, pages 511–527. SIAM, 2002.
- [404] Shoudan Liang, Stefanie Fuhrman, Roland Somogyi, et al. Reveal, a general reverse engineering algorithm for inference of genetic network architectures. In *Pacific symposium on biocomputing*, volume 3, pages 18–29, 1998.
- [405] Andy Liaw and Matthew Wiener. Classification and regression by randomforest. *R news*, 2(3):18–22, 2002.
- [406] Dao-I Lin and Zvi M. Kedem. Pincer search: A new algorithm for discovering the maximum frequent set. In *EDBT*, volume 1377 of *LNCS*, pages 105–119. Springer, 1998.
- [407] Jessica Lin, Eamonn J. Keogh, Stefano Lonardi, and Bill Yuan chi Chiu. A symbolic representation of time series, with implications for streaming algorithms. In Mohammed Javeed Zaki and Charu C. Aggarwal, editors, *DMKD*, pages 2–11. ACM, 2003.
- [408] Tsvi Lissack and King-Sun Fu. Error estimation in pattern recognition via l-distance between posterior density functions. *IEEE Transactions on Information Theory*, 22(1):34–45, 1976.
- [409] Bing Liu, Wynne Hsu, and Yiming Ma. Integrating classification and association rule mining. In *KDD*, pages 80–86. ACM, 1998.
- [410] Bing Liu, Wynne Hsu, and Yiming Ma. Pruning and summarizing the discovered associations. In *Proceedings of the fifth ACM SIGKDD international conference on Knowledge discovery and data mining*, pages 125–134. ACM, 1999.
- [411] Guimei Liu, Jinyan Li, Kelvin Sim, and Limsoon Wong. Distance based subspace clustering with flexible dimension partitioning. In Rada Chirkova, Asuman Dogac, M. Tamer Özsu, and Timos K. Sellis, editors, *ICDE*, pages 1250–1254. IEEE, 2007.
- [412] Guimei Liu, Hongjun Lu, Wenwu Lou, and Jeffrey Xu Yu. On computing, storing and querying frequent patterns. In *Proceedings of the ninth ACM SIGKDD international conference on Knowledge discovery and data mining*, KDD '03, pages 607–612, New York, NY, USA, 2003. ACM.

- [413] Hongyan Liu, Jiawei Han, Dong Xin, and Zheng Shao. Top-down mining of interesting patterns from very high dimensional data. In *Proceedings of the 22nd International Conference on Data Engineering, ICDE '06*, pages 114–, Washington, DC, USA, 2006. IEEE Computer Society.
- [414] Huan Liu and Hiroshi Motoda. *Feature Selection for Knowledge Discovery and Data Mining*. Kluwer Academic Publishers, Norwell, MA, USA, 1998.
- [415] Jinze Liu and Wei Wang. Op-cluster: Clustering by tendency in high dimensional space. In *ICDM*, pages 187–, Washington, DC, USA, 2003. IEEE Computer Society.
- [416] Jinze Liu, Jiong Yang, and Wei Wang. Biclustering in gene expression data by tendency. In *CSB*, pages 182–193. IEEE, 2004.
- [417] Jun Liu, Andrew Neuwald, and Charles Lawrence. Bayesian models for multiple local sequence alignment and gibbs sampling strategies. *American Stat. Ass.*, 90(432):1156–1170, 1995.
- [418] Tie-Fei Liu, Wing-Kin Sung, and Ankush Mittal. Model gene network by semi-fixed bayesian network. *Expert Systems with Applications*, 30(1):42–49, 2006.
- [419] Alan J Lockett and Risto Miikkulainen. Temporal convolution machines for sequence learning. Technical Report AI-09-04, Department of Computer Sciences, the University of Texas at Austin, 2009.
- [420] Elsa Loekito and James Bailey. Using highly expressive contrast patterns for classification-is it worthwhile? In *Advances in Knowledge Discovery and Data Mining*, pages 483–490. Springer, 2009.
- [421] Francisco-Javier Lopez, Armando Blanco, Fernando Garcia-Alcalde, Carlos Cano, and Antonio Marin. Fuzzy association rules for biological data analysis: A case study on yeast. *BMC Bioinf.*, 9, 2008.
- [422] FranciscoJ Lopez, Armando Blanco, Fernando Garcia, Carlos Cano, and Antonio Marin. Fuzzy association rules for biological data analysis: A case study on yeast. *BMC Bioinformatics*, 9:1–18, 2008.
- [423] Cécile Low-Kam, Chedy Raïssi, Mehdi Kaytoue, and Jian Pei. Mining statistically significant sequential patterns. In *Data Mining (ICDM), 2013 IEEE 13th International Conference on*, pages 488–497. IEEE, 2013.
- [424] Nizar R. Mabroukeh and C. I. Ezeife. A taxonomy of sequential pattern mining algorithms. *ACM Comput. Surv.*, 43(1):3:1–3:41, December 2010.
- [425] Shev MacNamara, Kevin Burrage, and Roger B Sidje. Multiscale modeling of chemical kinetics via the master equation. *Multiscale Modeling & Simulation*, 6(4):1146–1168, 2008.
- [426] Jamie I. MacPherson, Jonathan Dickerson, John Pinney, and David Robertson. Patterns of hiv-1 protein interaction identify perturbed host-cellular subsystems. *PLoS Comput Biol*, 6(7):e1000863, 2010.
- [427] Sara Madeira, Miguel Nobre Parreira Cacho Teixeira, Isabel Sá-Correia, and Arlindo Oliveira. Identification of regulatory modules in time series gene expression data using a linear time biclustering algorithm. *IEEE/ACM Transactions on Computational Biology and Bioinformatics*, 1:153–165, january 2010.
- [428] Sara C. Madeira. Neuroclinomics: Understanding neurodegenerative diseases through clinical and omics data integration, 2009.
- [429] Sara C. Madeira and Arlindo L. Oliveira. Biclustering algorithms for biological data analysis: A survey. *IEEE/ACM Trans. Comput. Biol. Bioinformatics*, 1(1):24–45, January 2004.
- [430] M. Mahfouz and M. Ismail. Bidens: Iterative density based biclustering algorithm with application to gene expression analysis. In *PWASET*, volume 37, pages 342–348, 2009.
- [431] MohamedA. Mahfouz and MohamedA. Ismail. Soft flexible overlapping biclustering utilizing hybrid search strategies. In *Adv. Mach. Learning Tech. and App.*, volume 322 of *Commun. Comp. and Inf. Science*, pages 315–326. Springer Berlin Heidelberg, 2012.
- [432] Kazuhisa Makino and Takeaki Uno. New algorithms for enumerating all maximal cliques. In Torben Hagerup and Jyrki Katajainen, editors, *SWAT*, volume 3111 of *Lecture Notes in Computer Science*, pages 260–272. Springer, 2004.
- [433] Danilo P. Mandic and Jonathon Chambers. *Recurrent Neural Networks for Prediction: Learning Algorithms, Architectures and Stability*. John Wiley & Sons, Inc., New York, NY, USA, 2001.
- [434] Shawn Mankad and George Michailidis. Biclustering three-dimensional data arrays with plaid models. *Comp. & Graphical Stat.*, 2013.
- [435] Heikki Mannila, Hannu Toivonen, and A Inkeri Verkamo. Discovery of frequent episodes in event sequences. *Data mining and knowledge discovery*, 1(3):259–289, 1997.
- [436] Sabrina Mantaci, Antonio Restivo, and Marinella Sciortino. Distance measures for biological sequences: Some recent approaches. *Int. J. Approx. Reasoning*, 47(1):109–124, 2008.
- [437] Daniel Marbach, Robert J Prill, Thomas Schaffter, Claudio Mattiussi, Dario Floreano, and Gustavo Stolovitzky. Revealing strengths and weaknesses of methods for gene network inference. *Proceedings of the National Academy of Sciences*, 107(14):6286–6291, 2010.
- [438] Daniel Marbach, Thomas Schaffter, Claudio Mattiussi, and Dario Floreano. Generating realistic in silico gene networks for performance assessment of reverse engineering methods. *Journal of computational biology*, 16(2):229–239, 2009.
- [439] Adam A Margolin, Ilya Nemenman, Katia Basso, Chris Wiggins, Gustavo Stolovitzky, Riccardo D Favera, and Andrea Califano. Aracne: an algorithm for the reconstruction of gene regulatory networks in a mammalian cellular context. *BMC bioinformatics*, 7(Suppl 1):S7, 2006.
- [440] David Martin, Christine Brun, Elisabeth Remy, Pierre Mouren, Denis Thieffry, and Bernard Jacq. Gotoolbox: functional analysis of gene datasets based on gene ontology. *Genome Biology*, (12):101, 2004.
- [441] J. Kent Martin and D. S. Hirschberg. Small sample statistics for classification error rates ii: Confidence intervals and significance tests. Technical report, DICS, 1996.
- [442] Ricardo Martinez, Claude Pasquier, and Nicolas Pasquier. Genminer: Mining informative association rules from genomic data. In *BIBM*, pages 15–22. IEEE CS, 2007.
- [443] Eleni Matechou, Ivy Liu, Shirley Pledger, and Richard Arnold. Biclustering models for ordinal data. In *Presentation at the NZ Statistical Assn. Annual Conference. University of Auckland*, 2011.
- [444] U. Maulik, A. Mukhopadhyay, M. Bhattacharyya, L. Kaderali, B. Brors, S. Bandyopadhyay, and R. Eils. Mining quasi-bicliques from hiv-1-human protein interaction network: A multiobjective biclustering approach. *IEEE/ACM TCBB*, 10(2):423–435, 2013.
- [445] Harley H McAdams and Adam Arkin. Stochastic mechanisms in gene expression. *Proceedings of the National Academy of Sciences*, 94(3):814–819, 1997.
- [446] Tara McIntosh and Sanjay Chawla. High confidence rule mining for microarray analysis. *IEEE/ACM Trans. Comput. Biol. Bioinformatics*, 4(4):611–623, October 2007.
- [447] Tara McIntosh and Sanjay Chawla. High confidence rule mining for microarray analysis. *Computational Biology and Bioinformatics, IEEE/ACM Transactions on*, 4(4):611–623, 2007.
- [448] Geoffrey J. McLachlan, RW Bean, and David Peel. A mixture model-based approach to the clustering of microarray expression data. *Bioinformatics*, 18(3):413–422, 2002.
- [449] Edward Meeds, Zoubin Ghahramani, Radford M Neal, and Sam T Roweis. Modeling dyadic data with binary latent factors. *Adv. in Neural Info. Proc. Systems*, 19:977, 2007.

- [450] N. Megiddo and R. Srikant. Discovering predictive association rules. In *KDD and DM*, pages 274–278. AAAI Press, 1998.
- [451] G. Melli. Synthetic classification data sets (SCDS). Technical report, School of Computing Science, Simon Fraser University, Burnaby, British Columbia, Canada, 1997. <http://fas.sfu.ca/cs/people/GradStudents/melli/SCDS/>.
- [452] Pedro Mendes, Wei Sha, and Keying Ye. Artificial gene networks for objective comparison of analysis algorithms. *Bioinformatics*, 19(suppl 2):ii122–ii129, 2003.
- [453] Jia Meng, Shou-Jiang Gao, and Yufei Huang. Enrichment constrained time-dependent clustering analysis for finding meaningful temporal transcription modules. *Bioinformatics*, 25(12):1521–1527, 2009.
- [454] Jean-Philippe Métivier, Patrice Boizumault, Bruno Crémilleux, Mehdi Khiari, and Samir Loudni. A constraint language for declarative pattern discovery. In *Symposium on Applied Computing*, pages 119–125, New York, NY, USA, 2012. ACM.
- [455] Tom Michoel, Riet De Smet, Anagha Joshi, Yves Van de Peer, and Kathleen Marchal. Comparative analysis of module-based versus direct methods for reverse-engineering transcriptional regulatory networks. *BMC systems biology*, 3(1):49, 2009.
- [456] Ron Milo, Shai Shen-Orr, Shalev Itzkovitz, Nadav Kashtan, Dmitri Chklovskii, and Uri Alon. Network motifs: simple building blocks of complex networks. *Science*, 298(5594):824–827, 2002.
- [457] Guy W. Mineau, Akshay Bissoon, and Robert Godin. Simple pre- and post-pruning techniques for large conceptual clustering structures. *Electron. Trans. Artif. Intell.*, 4(C):1–20, 2000.
- [458] Pablo Minguez, Fátima Al-Shahrour, David Montaner, and Joaquín Dopazo. Functional profiling of microarray experiments using text-mining derived bioentities. *Bioinformatics*, 23(22):3098–3099, 2007.
- [459] David Minnen, Charles Isbell, Irfan Essa, and Thad Starner. Detecting subdimensional motifs: An efficient algorithm for generalized multivariate pattern discovery. In *Data Mining, 2007. ICDM 2007. Seventh IEEE International Conference on*, pages 601–606. IEEE, 2007.
- [460] Debahuti Mishra, Kailash Shaw, Sashikala Mishra, Amiya Kumar Rath, and Milu Acharya. Discovering non-exclusive functional modules from gene expression data. *IJ Inf. Commun. Technol.*, 3(3):227–241, 2011.
- [461] Sushmita Mitra and Haider Banka. Multi-objective evolutionary biclustering of gene expression data. *Pattern Recogn.*, 39(12):2464–2477, December 2006.
- [462] Sushmita Mitra, Ranajit Das, Haider Banka, and Subhasis Mukhopadhyay. Gene interaction – an evolutionary biclustering approach. *Information Fusion*, 10(3):242 – 249, 2009. Special Issue on Natural Computing Methods in Bioinformatics.
- [463] Eric Mjolsness, David H Sharp, and John Reinitz. A connectionist model of development. *Journal of theoretical Biology*, 152(4):429–453, 1991.
- [464] Annette M. Molinaro, Richard Simon, and Ruth M. Pfeiffer. Prediction error estimation: a comparison of resampling methods. *Bioinformatics*, 21(15):3301–3307, 2005.
- [465] Carla S Möller-Levet, Simon N Archer, Giselda Bucca, Emma E Laing, Ana Slak, Renata Kabiljo, June CY Lo, Nayantara Santhi, Malcolm von Schantz, Colin P Smith, et al. Effects of insufficient sleep on circadian rhythmicity and expression amplitude of the human blood transcriptome. *Proceedings of the National Academy of Sciences*, 110(12):E1132–E1141, 2013.
- [466] F. Mörchen. *Time Series Feature Extraction for Data Mining Using DWT and DFT*. Reihe Informatik. Univ., 2003.
- [467] F. Mörchen. *Time series knowledge mining*. Wissenschaft in Dissertationen. Görlich & Weiershäuser, 2006.
- [468] Karima Mouhoubi, Lucas Létocart, and Céline Rouveirol. A knowledge-driven bi-clustering method for mining noisy datasets. In *Neural Info. Processing*, volume 7665 of *LNCS*, pages 585–593. Springer, 2012.
- [469] Abdullah Mueen, Eamonn J Keogh, Qiang Zhu, Sydney Cash, and M Brandon Westover. Exact discovery of time series motifs. In *SDM*, pages 473–484. SIAM, 2009.
- [470] Abdullah Mueen, Eamonn J. Keogh, Qiang Zhu, Sydney Cash, and M. Brandon Westover. Exact discovery of time series motifs. In *SDM*, pages 473–484. SIAM, 2009.
- [471] Sayan Mukherjee, Pablo Tamayo, Simon Rogers, Ryan Rifkin, Anna Engle, Colin Campbell, Todd R. Golub, and Jill P. Mesirov. Estimating dataset size requirements for classifying dna microarray data. *Journal of Computational Biology*, 10:119–142, 2003.
- [472] Anirban Mukhopadhyay, Ujjwal Maulik, and Sanghamitra Bandyopadhyay. A novel coherence measure for discovering scaling biclusters from gene expression data. *Journal of bioinformatics and computational biology*, 7(05):853–868, 2009.
- [473] Anirban Mukhopadhyay, Ujjwal Maulik, and Sanghamitra Bandyopadhyay. A novel biclustering approach to association rule mining for predicting hiv-1–human protein interactions. *PLoS ONE*, 7(4):e32289, 04 2012.
- [474] Emmanuel Müller, Ira Assent, Ralph Krieger, Stephan Günemann, and Thomas Seidl. Densest: Density estimation for data mining in high dimensional spaces. In *SDM*, pages 173–184. SIAM, 2009.
- [475] Emmanuel Müller, Stephan Günemann, Ira Assent, and Thomas Seidl. Evaluating clustering in subspace projections of high dimensional data. *Proc. VLDB Endow.*, 2(1):1270–1281, August 2009.
- [476] James Munkres. Algorithms for the Assignment and Transportation Problems. *Society for Ind. and Applied Math.*, 5(1):32–38, 1957.
- [477] T. M. Murali and Simon Kasif. Extracting conserved gene expression motifs from gene expression data. In *Pacific Symposium on Biocomputing*, pages 77–88, 2003.
- [478] Kevin Murphy. *Dynamic Bayesian Networks: Representation, Inference and Learning*. PhD thesis, UC Berkeley, CS, 2002.
- [479] Kevin Murphy, Saira Mian, et al. Modelling gene expression data using dynamic bayesian networks. Technical report, Technical report, Computer Science Division, University of California, Berkeley, CA, 1999.
- [480] Youji Nakagawa, Naoko Sakumoto, Yoshinobu Kaneko, and Satoshi Harashima. Mga2p is a putative sensor for low temperature and oxygen to induce *ic1* transcription in *Saccharomyces cerevisiae*. *Biochemical and biophysical research communications*, 291(3):707–713, 2002.
- [481] Alex Nanopoulos, Rob Alcock, and Yannis Manolopoulos. Information processing and technology. In *Feature-based classification of time-series data*, pages 49–61. Nova Science Publishers, Commack, NY, USA, 2001.
- [482] Alex Nanopoulos, Rob Alcock, and Yannis Manolopoulos. Information processing and technology. chapter Feature-based classification of time-series data, pages 49–61. Nova Science Publishers, Inc., Commack, NY, USA, 2001.
- [483] Krishnamurthy Natarajan, Michael R Meyer, Belinda M Jackson, David Slade, Christopher Roberts, Alan G Hinnebusch, and Matthew J Marton. Transcriptional profiling shows that *gc4p* is a master regulator of gene expression during amino acid starvation in yeast. *Molecular and cellular biology*, 21(13):4347–4368, 2001.
- [484] Gerard J Nau, Joan FL Richmond, Ann Schlesinger, Ezra G Jennings, Eric S Lander, and Richard A Young. Human macrophage activation programs induced by bacterial pathogens. *Proceedings of the National Academy of Sciences*, 99(3):1503–1508, 2002.
- [485] Juan A. Nepomuceno, Alicia Troncoso Lora, Jesús S. Aguilar-Ruiz, and Jorge García-Gutiérrez. Biclusters evaluation based on shifting and scaling patterns. In *IDEAL*, pages 840–849, Berlin, Heidelberg, 2007. Springer-Verlag.
- [486] Juan A Nepomuceno, Alicia Troncoso, Isabel A Nepomuceno-Chamorro, and Jesús S Aguilar-Ruiz. Integrating biological knowledge based on functional annotations for biclustering of gene expression data. *Computer Methods and Programs in Biomedicine*, 2015.
- [487] Mark EJ Newman. Communities, modules and large-scale structure in networks. *Nature Physics*, 8(1):25–31, 2012.
- [488] AA Neznanov and AA Parinov. Analyzing social networks services using formal concept analysis research toolbox. In *SNAFCA*, 2015.

- [489] Raymond T. Ng, Laks V. S. Lakshmanan, Jiawei Han, and Alex Pang. Exploratory mining and pruning optimizations of constrained associations rules. *SIGMOD R.*, 27(2):13–24, 1998.
- [490] Hai-Long Nguyen, Wee-Keong Ng, and Yew-Kwong Woon. Closed motifs for streaming time series classification. *KAIS*, pages 1–25, 2013.
- [491] Hung S. Nguyen. Approximate Boolean Reasoning: Foundations and Applications in Data Mining. *Transactions on Rough Sets V*, 4100:334–506, 2006.
- [492] Partha Niyogi and Federico Girosi. On the relationship between generalization error, hypothesis complexity, and sample complexity for radial basis functions. *Neural Comput.*, 8(4):819–842, May 1996.
- [493] Amina Noor. *Modeling Gene Regulatory Networks from Time Series Data using Particle Filtering*. PhD thesis, Texas A&M University, 2011.
- [494] G. Norén, Johan Hopstadius, Bate, Star, and I. Edwards. Temporal pattern discovery in longitudinal electronic patient records. *Data Min. Knowl. Discov.*, 20(3):361–387, May 2010.
- [495] Noa Novershtern, Aviv Regev, and Nir Friedman. Physical module networks: an integrative approach for reconstructing transcription regulation. *Bioinformatics*, 27(13):i177–i185, 2011.
- [496] O. Odibat and C. Reddy. *A Generalized Framework for Mining Arbitrarily Positioned Overlapping Co-clusters*, chapter 30, pages 343–354. SIAM, 2011.
- [497] Omar Odibat and Chandan K Reddy. Efficient mining of discriminative co-clusters from gene expression data. *Knowledge and Information Systems*, pages 1–30, 2013.
- [498] Omar Odibat and Chandan K. Reddy. Efficient mining of discriminative co-clusters from gene expression data. *Knowledge and Information Systems*, pages 1–30, 2013.
- [499] Markus Ojala, Niko Vuokko, Aleksi Kallio, Niina Haiminen, and Heikki Mannila. Randomization of real-valued matrices for assessing the significance of data mining results. In *SDM*, volume 8, pages 494–505. SIAM, 2008.
- [500] Yoshifumi Okada, Wataru Fujibuchi, and Paul Horton. A biclustering method for gene expression module discovery using closed itemset enumeration algorithm. *IPSJ Transactions on Bioinformatics*, 48(SIG5):39–48, 2007.
- [501] Yoshifumi Okada, Kosaku Okubo, Paul Horton, and Wataru Fujibuchi. Exhaustive search method of gene expression modules and its application to human tissue data. *IAENG IJ of Comp. Science*, 34(1):119–126, 2007.
- [502] M Oppen, W Kinzel, J Kleinz, and R Nehl. On the ability of the optimal perceptron to generalise. *Journal of Physics A: Mathematical and General*, 23(11):L581, 1990.
- [503] Patryk Orzechowski. Proximity measures and results validation in biclustering – a survey. In *Artificial Intelligence and Soft Computing*, volume 7895 of *LNCS*, pages 206–217. Springer Berlin Heidelberg, 2013.
- [504] Michael R Osborne, Brett Presnell, and Berwin A Turlach. On the lasso and its dual. *Journal of Computational and Graphical statistics*, 9(2):319–337, 2000.
- [505] Jason Palmer, Kenneth Kreutz-Delgado, Bhaskar D Rao, and David P Wipf. Variational em algorithms for non-gaussian latent variable models. In *Advances in neural information processing systems*, pages 1059–1066, 2005.
- [506] Feng Pan, Gao Cong, Anthony K. H. Tung, Jiong Yang, and Mohammed Javeed Zaki. Carpenter: finding closed patterns in long biological datasets. In *KDD*, pages 637–642. ACM, 2003.
- [507] Feng Pan, A.K.H. Tung, Gao Cong, and Xin Xu. Cobbler: combining column and row enumeration for closed pattern discovery. In *Sci. Stat. Data. Man.*, pages 21 – 30, 2004.
- [508] Wei Pan, Jizhen Lin, Chap T Le, et al. Model-based cluster analysis of microarray gene-expression data. *Genome Biol*, 3(2):1–0009, 2002.
- [509] Gaurav Pandey, Gowtham Atluri, Michael Steinbach, Chad L. Myers, and Vipin Kumar. An association analysis approach to biclustering. In *KDD*, pages 677–686, New York, NY, USA, 2009. ACM.
- [510] Jong Soo Park, Ming-Syan Chen, and Philip S Yu. *An effective hash-based algorithm for mining association rules*, volume 24. ACM, 1995.
- [511] Jong Soo Park, Ming-Syan Chen, and Philip S Yu. Efficient parallel data mining for association rules. In *Proceedings of the fourth international conference on Information and knowledge management*, pages 31–36. ACM, 1995.
- [512] Lance Parsons, Ehtesham Haque, and Huan Liu. Subspace clustering for high dimensional data: A review. *SIGKDD Exp. Newsl.*, 6(1):90–105, 2004.
- [513] Nicolas Pasquier, Yves Bastide, Rafik Taouil, and Lotfi Lakhal. Discovering frequent closed itemsets for association rules. In *Proceedings of the 7th International Conference on Database Theory*, ICDT ’99, pages 398–416, London, UK, UK, 1999. Springer-Verlag.
- [514] Nicolas Pasquier, Yves Bastide, Rafik Taouil, and Lotfi Lakhal. Efficient mining of association rules using closed itemset lattices. *Inf. Syst.*, 24(1):25–46, 1999.
- [515] Dhaval Patel, Wynne Hsu, and Mong Li Lee. Mining relationships among interval-based events for classification. In *SIGMOD*, pages 393–404, New York, NY, USA, 2008. ACM.
- [516] Anne Patrikainen and Marina Meila. Comparing subspace clusterings. *IEEE TKDE*, 18(7):902–916, July 2006.
- [517] Judea Pearl. *Probabilistic reasoning in intelligent systems: networks of plausible inference*. Morgan Kaufmann Publishers Inc., San Francisco, CA, USA, 1988.
- [518] René Peeters. The maximum edge biclique problem is np-complete. *Discrete Appl. Math.*, 131(3):651–654, September 2003.
- [519] Jian Pei, Guozhu Dong, Wei Zou, and Jiawei Han. On computing condensed frequent pattern bases. In *Proceedings of the 2002 IEEE International Conference on Data Mining*, ICDM ’02, pages 378–, Washington, DC, USA, 2002. IEEE Computer Society.
- [520] Jian Pei and Jiawei Han. Can we push more constraints into frequent pattern mining? In *KDD*, pages 350–354, New York, NY, USA, 2000. ACM.
- [521] Jian Pei and Jiawei Han. Constrained frequent pattern mining: a pattern-growth view. *SIGKDD Explor. Newsl.*, 4(1):31–39, 2002.
- [522] Jian Pei, Jiawei Han, Hongjun Lu, Shojiro Nishio, Shiwei Tang, and Dongqing Yang. H-mine: Hyper-structure mining of frequent patterns in large databases. In *Data Mining, 2001. ICDM 2001, Proceedings IEEE International Conference on*, pages 441–448. IEEE, 2001.
- [523] Jian Pei, Jiawei Han, Behzad Mortazavi-Asl, Jianyong Wang, Helen Pinto, Qiming Chen, Umeshwar Dayal, and Mei-Chun Hsu. Mining sequential patterns by pattern-growth: The prefixspan approach. *IEEE Trans. on Knowl. and Data Eng.*, 16(11):1424–1440, November 2004.
- [524] Jian Pei, Jiawei Han, Behzad Mortazavi-Asl, and Hua Zhu. Mining access patterns efficiently from web logs. In *PAKDD*, pages 396–407, London, UK, UK, 2000. Springer-Verlag.
- [525] Jian Pei, Jiawei Han, and Wei Wang. Mining sequential patterns with constraints in large databases. In *Proceedings of the eleventh international conference on Information and knowledge management*, pages 18–25. ACM, 2002.
- [526] Jian Pei, Jiawei Han, and Wei Wang. Constraint-based sequential pattern mining: the pattern-growth methods. *Journal of Intelligent Information Systems*, 28(2):133–160, 2007.
- [527] Liuqing Peng and Junying Zhang. An entropy weighting mixture model for subspace clustering of high-dimensional data. *Pattern Recognition Letters*, 32(8):1154–1161, 2011.
- [528] Jose B Pereira-Leal, Anton J Enright, and Christos A Ouzounis. Detection of functional modules from protein interaction networks. *PROTEINS: Structure, Function, and Bioinformatics*, 54(1):49–57, 2004.

- [529] Gianvito Pio, Michelangelo Ceci, Domenica D'Elia, and Donato Malerba. Network reconstruction for the identification of mirna:mrna interaction networks. In Toon Calders, Floriana Esposito, Eyke Hüllermeier, and Rosa Meo, editors, *Machine Learning and Knowledge Discovery in Databases*, volume 8726 of *Lecture Notes in Computer Science*, pages 508–511. Springer Berlin Heidelberg, 2014.
- [530] Clara Pizzuti. Ga-net: A genetic algorithm for community detection in social networks. In Gunter Rudolph, Thomas Jansen, Nicola Beume, Simon Lucas, and Carlo Poloni, editors, *Parallel Problem Solving from Nature: PPSN X*, volume 5199 of *Lecture Notes in Computer Science*, pages 1081–1090. Springer Berlin Heidelberg, 2008.
- [531] Marc Plantevit and Bruno Cremlieux. Condensed representation of sequential patterns according to frequency-based measures. In *Adv. in Int. Data Analysis*, volume 5772 of *LNCS*, pages 155–166. Springer Berlin Heid., 2009.
- [532] Beatriz Pontes, Raúl Giráldez, and Jesús S Aguilar-Ruiz. Configurable pattern-based evolutionary biclustering of gene expression data. *Algorithms for Molecular Biology*, 8(1):4, 2013.
- [533] Ignacio Ponzoni, Francisco Azuaje, Juan Augusto, and David Glass. Inferring adaptive regulation thresholds and association rules from gene expression data through combinatorial optimization learning. *IEEE/ACM Trans. Comput. Biology Bioinform.*, 4(4):624–634, 2007.
- [534] R. Potharst and A. J. Feelders. Classification trees for problems with monotonicity constraints. *SIGKDD Explor. Newsl.*, 4(1):1–10, June 2002.
- [535] Richard J. Povinelli, Michael T. Johnson, Andrew C. Lindgren, and Jinjin Ye. Time series classification using gaussian mixture models of reconstructed phase spaces. *IEEE Trans. on Knowl. and Data Eng.*, 16(6):779–783, June 2004.
- [536] Amela Prelič, Stefan Bleuler, Philip Zimmermann, Anja Wille, Peter Bühlmann, Wilhelm Gruissem, Lars Hennig, Lothar Thiele, and Eckart Zitzler. A systematic comparison and evaluation of biclustering methods for gene expression data. *Bioinformatics*, 22(9):1122–1129, 2006.
- [537] Bernard Prum, François Rodolphe, and Élisabeth de Turckheim. Finding words with unexpected frequencies in deoxyribonucleic acid sequences. *Journal of the Royal Statistical Society. Series B (Methodological)*, pages 205–220, 1995.
- [538] John R Punin, Mukkai S Krishnamoorthy, and Mohammed J Zaki. Web usage mining: Languages and algorithms. In *Exploratory Data Analysis in Empirical Research*, pages 266–281. Springer, 2003.
- [539] Gengsheng Qin and Lejla Hotilovac. Comparison of non-parametric confidence intervals for the area under the ROC curve of a continuous-scale diagnostic test. *Stat. Methods Med. Res.*, 17(2):207–221, 2008.
- [540] Yong qing Wei, Dong Liu, and Lin shan Duan. Distributed prefixspan algorithm based on mapreduce. In *Inf. Tech. in Medicine and Education*, volume 2, pages 901–904, 2012.
- [541] J. R. Quinlan and R. M. Cameron-Jones. Oversearching and layered search in empirical learning. In *Proceedings of the 14th international joint conference on Artificial intelligence - Volume 2*, IJCAI'95, pages 1019–1024, San Francisco, CA, USA, 1995. Morgan Kaufmann Publishers Inc.
- [542] Ross Quinlan. *C4.5: Programs for Machine Learning*. Morgan Kaufmann Publishers, San Mateo, CA, 1993.
- [543] L. R. Rabiner. A tutorial on hidden Markov models and selected applications in speech recognition. *Proceedings of the IEEE*, 77(2):257–286, February 1989.
- [544] Troy Raeder, T. Ryan Hoens, and Nitesh V. Chawla. Consequences of variability in classifier performance estimates. In *ICDM*, pages 421–430, 2010.
- [545] Luc De Raedt, Tias Guns, and Siegfried Nijssen. Constraint programming for data mining and machine learning. In *AAAI*. AAAI Press, 2010.
- [546] ShuzlinaAbdul Rahman, A. Bakar, and Z. Hussein. A review on protein sequence clustering research. In *ICBE*, volume 21 of *IFMBE Proceedings*, pages 275–278. Springer Berlin-Heidelberg, 2008.
- [547] Naren Ramakrishnan, Satish Tadepalli, Layne T Watson, Richard F Helm, Marco Antonioti, and Bud Mishra. Reverse engineering dynamic temporal models of biological processes and their relationships. *Proceedings of the National Academy of Sciences*, 107(28):12511–12516, 2010.
- [548] Ganesh Ramesh, William A. Maniatty, and Mohammed J. Zaki. Feasible itemset distributions in data mining: theory and application. In *Symposium on Princ. of data. sys.*, pages 284–295. ACM Press, 2003.
- [549] Marco F Ramoni, Paola Sebastiani, and Isaac S Kohane. Cluster analysis of gene expression dynamics. *Proceedings of the National Academy of Sciences*, 99(14):9121–9126, 2002.
- [550] Claudia Rangel, John Angus, Zoubin Ghahramani, Maria Lioumi, Elizabeth Sotheran, Alessia Gaiba, David L Wild, and Francesco Falciani. Modeling t-cell activation using gene expression profiling and state-space models. *Bioinformatics*, 20(9):1361–1372, 2004.
- [551] Chotirat (Ann) Ratanamahatana and Eamonn J. Keogh. Making time-series classification more accurate using learned constraints. In Michael W. Berry, Umeshwar Dayal, Chandrika Kamath, and David B. Skillicorn, editors, *SDM*. SIAM, 2004.
- [552] S. Raudys. On dimensionality, sample size, and classification error of nonparametric linear classification algorithms. *IEEE Trans. Pattern Anal. Mach. Intell.*, 19(6):667–671, June 1997.
- [553] Sarunas J. Raudys and Anil K. Jain. Small sample size effects in statistical pattern recognition: Recommendations for practitioners. *IEEE Trans. Pattern Anal. Mach. Intell.*, 13(3):252–264, March 1991.
- [554] David J Reiss, Nitin S Baliga, and Richard Bonneau. Integrated biclustering of heterogeneous genome-wide datasets for the inference of global regulatory networks. *BMC bioinformatics*, 7(1):280, 2006.
- [555] Kristoffer TG Rigbolt, Tatyana A Prokhorova, Vyacheslav Akimov, Jeanette Henningsen, Pia T Johansen, Irina Kratchmarova, Moustapha Kassem, Matthias Mann, Jesper V Olsen, and Blagoy Blagoev. System-wide temporal characterization of the proteome and phosphoproteome of human embryonic stem cell differentiation. *Science signaling*, 4(164):rs3, 2011.
- [556] Stéphane Robin, J-J Daudin, Hugues Richard, M-F Sagot, and Sophie Schbath. Occurrence probability of structured motifs in random sequences. *Journal of Computational Biology*, 9(6):761–773, 2002.
- [557] Peter M Robinson. *Time series with long memory*. Oxford University Press, 2003.
- [558] M. Rocha, R. Mendes, P. Maia, D. Glez-Pena, and F. Fdez-Riverola. A platform for the selection of genes in dna microarray data using evolutionary algorithms. In *GECCO*. ACM Press, 2007.
- [559] John F. Roddick and Myra Spiliopoulou. A survey of temporal knowledge discovery paradigms and methods. *IEEE Trans. on Knowl. and Data Eng.*, 14(4):750–767, July 2002.
- [560] Andreas Rosenwald and *dbbc team*. The use of molecular profiling to predict survival after chemotherapy for diffuse large-B-cell lymphoma. *N Engl J Med*, 346(25):1937–1947, 2002.
- [561] Peter Rousseeuw. Silhouettes: a graphical aid to the interpretation and validation of cluster analysis. *J. Comput. Appl. Math.*, 20(1):53–65, November 1987.
- [562] Davide Roverso. Multivariate temporal classification by windowed wavelet decomposition and recurrent neural networks. In *ANS Int. Topical Meeting on NPICHHI*, 2000.
- [563] Swarup Roy, Dhruva K Bhattacharyya, and Jugal K Kalita. Cobi: Pattern based co-regulated biclustering of gene expression data. *Pattern Recognition Letters*, 34(14):1669–1678, 2013.

- [564] Marilyn Safran, Irina Dalah, Justin Alexander, Naomi Rosen, Tsippi Iny Stein, Michael Shmoish, Noam Nativ, Iris Bahir, Tirza Doniger, Hagit Krug, A Sirota-Madi, T Olender, Y Golan, G Stelzer, A Harel, and D Lancet. Genecards version 3: the human gene integrator. *Database*, 2010:baq020, 2010.
- [565] Lucas Sanchez and Denis Thieffry. A logical analysis of the drosophila gap-gene system. *Journal of theoretical Biology*, 211(2):115–141, 2001.
- [566] Ashok Savasere, Edward Robert Omiecinski, and Shamkant B Navathe. An efficient algorithm for mining association rules in large databases, 1995.
- [567] Christopher J Savoie, Sachiyo Aburatani, Shouji Watanabe, Yoshihiro Eguchi, Shigeru Muta, Seiya Imoto, Satoru Miyano, Satoru Kuhara, and Kosuke Tashiro. Use of gene networks from full genome microarray libraries to identify functionally relevant drug-affected genes and gene regulation cascades. *DNA research*, 10(1):19–25, 2003.
- [568] Sophie Schbath. An efficient statistic to detect over-and under-represented words in dna sequences. *Journal of Computational Biology*, 4(2):189–192, 1997.
- [569] T. Scheffer. Finding association rules that trade support optimally against confidence. *Intell. Data Anal.*, 9(4):381–395, July 2005.
- [570] Alexander Schliep, Ivan G Costa, Christine Steinhoff, and Alexander Schonhuth. Analyzing gene expression time-courses. *Computational Biology and Bioinformatics, IEEE/ACM Transactions on*, 2(3):179–193, 2005.
- [571] J. Scott. *Social Network Analysis*. SAGE Publications, 2012.
- [572] Paola Sebastiani, Marco Ramoni, Vikki Nolan, Clinton Baldwin, and Martin Steinberg. Genetic dissection and prognostic modeling of overt stroke in sickle cell anemia. *Nature Genetics*, 37(4):435–440, March 2005.
- [573] Eran Segal, Alexis Battle, and Daphne Koller. Decomposing gene expression into cellular processes. In *Pacific Symp. on Biocomputing*, volume 8, pages 89–100, 2003.
- [574] Eran Segal, Dana Pe’er, Aviv Regev, and Daphne Koller. Learning module networks. In *Proceedings of the Nineteenth conference on Uncertainty in Artificial Intelligence*, pages 525–534. Morgan Kaufmann Publishers Inc., 2002.
- [575] Eran Segal, Ben Taskar, Audrey Gasch, Nir Friedman, and Daphne Koller. Rich probabilistic models for gene expression. *Bioinformatics*, 17(suppl 1):S243–S252, 2001.
- [576] Eran Segal, Haidong Wang, and Daphne Koller. Discovering molecular pathways from protein interaction and gene expression data. *Bioinformatics*, 19(suppl 1):i264–i272, 2003.
- [577] Karlton Sequeira and Mohammed Zaki. Schism: a new approach to interesting subspace mining. *Int. J. Bus. Intell. Data Min.*, 1(2):137–160, December 2005.
- [578] Akdes Serin and Martin Vingron. Debi: Discovering differentially expressed biclusters using a frequent itemset approach. *Algorithms for Molecular Biology*, 6:1–12, 2011.
- [579] Juliet Popper Shaffer. Multiple hypothesis testing. *Annual Review of Psychology*, 46(1):561–584, 1995.
- [580] Juliet Popper Shaffer. Multiple hypothesis testing. *Annual review of psychology*, 46(1):561–584, 1995.
- [581] Fanhua Shang, L.C. Jiao, and Fei Wang. Graph dual regularization non-negative matrix factorization for co-clustering. *Pattern Recognition*, 45(6):2237 – 2250, 2012. Brain Decoding.
- [582] Roded Sharan, Igor Ulitsky, and Ron Shamir. Network-based prediction of protein function. *Molecular systems biology*, 3(1), 2007.
- [583] Shai S Shen-Orr, Ron Milo, Shmoolik Mangan, and Uri Alon. Network motifs in the transcriptional regulation network of escherichia coli. *Nature genetics*, 31(1):64–68, 2002.
- [584] Qizheng Sheng, Yves Moreau, and Bart De Moor. Biclustering microarray data by gibbs sampling. In *ECCB*, pages 196–205, 2003.
- [585] Ilya Shmulevich, Edward R Dougherty, Seungchan Kim, and Wei Zhang. Probabilistic boolean networks: a rule-based uncertainty model for gene regulatory networks. *Bioinformatics*, 18(2):261–274, 2002.
- [586] A. Silberschatz and A. Tuzhilin. What makes patterns interesting in knowledge discovery systems. *IEEE TKDE*, 8(6):970–974, December 1996.
- [587] C. Silverstein, S. Brin, and R. Motwani. Beyond market baskets: Generalizing association rules to dependence rules. *Data Min. Knowl. Discov.*, 2(1):39–68, January 1998.
- [588] Kelvin Sim, Vivekanand Gopalkrishnan, Arthur Zimek, and Gao Cong. A survey on enhanced subspace clustering. *Data Mining and Knowledge Discovery*, 26(2):332–397, 2013.
- [589] Surendra K. Singhi and Huan Liu. Feature subset selection bias for classification learning. In *IC on Machine Learning*, pages 849–856, New York, NY, USA, 2006. ACM.
- [590] Atish P. Sinha and Huimin Zhao. Incorporating domain knowledge into data mining classifiers: An application in indirect lending. *Decision Support Systems*, 46(1):287 – 299, 2008.
- [591] Kimmen Sjolander. Hidden markov models in computational biology: applications to protein modeling. Technical report, Univ of California at Santa Cruz, 1994.
- [592] Roland Somogyi and Carol Ann Sniegowski. Modeling the complexity of genetic networks: understanding multigenic and pleiotropic regulation. *Complexity*, 1(6):45–63, 1996.
- [593] Eduardo Sontag, Anatoly Kiyatkin, and Boris N Kholodenko. Inferring dynamic architecture of cellular networks using time series of gene expression, protein and metabolite data. *Bioinformatics*, 20(12):1877–1886, 2004.
- [594] Antti Sorjamaa, Jin Hao, Nima Reyhani, Yongnan Ji, and Amaury Lendasse. Methodology for long-term prediction of time series. *Neurocomput.*, 70:2861–2869, October 2007.
- [595] Antti Sorjamaa and Amaury Lendasse. Time series prediction using dirrec strategy. In *ESANN’06*, pages 143–148, 2006.
- [596] Victor Spirin and Leonid A Mirny. Protein complexes and functional modules in molecular networks. *National Academy of Sciences*, 100(21):12123–12128, 2003.
- [597] R. Srikant and R. Agrawal. Mining quantitative association rules in large relational tables. *SIGMOD Rec.*, 25(2):1–12, June 1996.
- [598] Ramakrishnan Srikant and Rakesh Agrawal. Mining sequential patterns: Generalizations and performance improvements. In *Proceedings of the 5th International Conference on Extending Database Technology: Advances in Database Technology, EDBT ’96*, pages 3–17, London, UK, UK, 1996. Springer-Verlag.
- [599] Ramakrishnan Srikant, Quoc Vu, and Rakesh Agrawal. Mining association rules with item constraints. In *KDD*, volume 97, pages 67–73, 1997.
- [600] Jaideep Srivastava, Robert Cooley, Mukund Deshpande, and Pang-Ning Tan. Web usage mining: Discovery and applications of usage patterns from web data. *ACM SIGKDD Explorations Newsletter*, 1(2):12–23, 2000.
- [601] Michael Steinbach, Pang-Ning Tan, Hui Xiong, and Vipin Kumar. Generalizing the notion of support. In *Proceedings of the tenth ACM SIGKDD international conference on Knowledge discovery and data mining, KDD ’04*, pages 689–694, New York, NY, USA, 2004. ACM.
- [602] Michael Steinbach, Haoyu Yu, Gang Fang, and Vipin Kumar. Using constraints to generate and explore higher order discriminative patterns. In *PAKDD*, volume 6634 of *LNCS*, pages 338–350. Springer, 2011.

- [603] Andreas Stolcke and Stephen M. Omohundro. Hidden markov model: Induction by bayesian model merging. In *Advances in NIPS*, pages 11–18, San Francisco, CA, USA, 1993. Morgan Kaufmann Publishers Inc.
- [604] Gustavo Stolovitzky, DON Monroe, and Andrea Califano. Dialogue on reverse-engineering assessment and methods. *Annals of the New York Academy of Sciences*, 1115(1):1–22, 2007.
- [605] Catherine A. Sugar, Gareth, and M. James. Finding the number of clusters in a data set: An information theoretic approach. *American Statistical Association*, 98:750–763, 2003.
- [606] Jimeng Sun, Yinglian Xie, Hui Zhang, and Christos Faloutsos. Less is more: Sparse graph mining with compact matrix decomposition. *Statistical Analysis and Data Mining: The ASA Data Science Journal*, 1(1):6–22, 2008.
- [607] Xin Sun, Qiming Hou, Zhong Ren, Kun Zhou, and Baining Guo. Radiance transfer biclustering for real-time all-frequency biscale rendering. *Visualization and Computer Graphics, IEEE Transactions on*, 17(1):64–73, 2011.
- [608] B. Surendiran and A. Vadivel. Feature selection using stepwise anova discriminant analysis for mammogram mass classification. *IJ on Signal Image Proc.*, 2(1):4, 2011.
- [609] Stephen Swift and Xiaohui Liu. Predicting glaucomatous visual field deterioration through short multivariate time series modelling. *Artificial Intelligence in Medicine*, 24(1):5–24, 2002.
- [610] Damian Szklarczyk, Andrea Franceschini, Stefan Wyder, Kristoffer Forslund, Davide Heller, Jaime Huerta-Cepas, Milan Simonovic, Alexander Roth, Alberto Santos, Kalliopi P Tsafou, et al. String v10: protein–protein interaction networks, integrated over the tree of life. *Nucleic acids research*, page gku1003, 2014.
- [611] I. Tagkopoulos, N. Slavov, and S.Y. Kung. Multi-class biclustering and classification based on modeling of gene regulatory networks. In *Bioinformatics and Bioengineering, 2005. BIBE 2005. Fifth IEEE Symposium on*, pages 89–96, Oct 2005.
- [612] Ilias Tagkopoulos, Nikolai Slavov, and Sun-Yuan Kung. Multi-class biclustering and classification based on modeling of gene regulatory networks. In *Bioinformatics and Bioengineering, 2005. BIBE 2005. Fifth IEEE Symposium on*, pages 89–96. IEEE, 2005.
- [613] Souhaib Ben Taieb, Gianluca Bontempi, Antti Sorjamaa, and Amaury Lendasse. Long-term prediction of time series by combining direct and mimo strategies. In *Proceedings of the 2009 Int. Joint Conf. on Neural Networks, IJCNN'09*, pages 1559–1566, Piscataway, NJ, USA, 2009. IEEE Press.
- [614] Pang-Ning Tan, Vipin Kumar, and Jaideep Srivastava. Selecting the right interestingness measure for association patterns. In *Proceedings of the eighth ACM SIGKDD international conference on Knowledge discovery and data mining, KDD '02*, pages 32–41, New York, NY, USA, 2002. ACM.
- [615] A. Tanay, R. Sharan, and R. Shamir. Biclustering Algorithms: A Survey. *Handbook of Computational Molecular Biology*, 2004.
- [616] Amos Tanay, Roded Sharan, and Ron Shamir. Discovering statistically significant biclusters in gene expression data. *Bioinformatics*, 18:136–144, 2002.
- [617] Chun Tang, Li Zhang, Murali Ramanathan, and Aidong Zhang. Interrelated two-way clustering: An unsupervised approach for gene expression data analysis. In *BIBE*, pages 41–, Washington, DC, USA, 2001. IEEE Computer Society.
- [618] Lei Tang and Huan Liu. Graph mining applications to social network analysis. In *Managing and Mining Graph Data*, pages 487–513. Springer, 2010.
- [619] Saeed Tavazoie, Jason D Hughes, Michael J Campbell, Raymond J Cho, and George M Church. Systematic determination of genetic network architecture. *Nature genetics*, 22(3):281–285, 1999.
- [620] M.C. Teixeira, P.T. Monteiro, J. Guerreiro, J.P. Gonçalves, N.P. Mira, S.C. dos Santos, T.R. Cabrito, M. Palma, C. Costa, A.P. Francisco, S.C. Madeira, A.L. Oliveira, A.T. Freitas, and I. Sá-Correia. The yeastract database: an upgraded information system for the analysis of gene and genomic transcription regulation in *Saccharomyces cerevisiae*. *Nucleic Acids Research*, (Database issue), 2014.
- [621] Mukund Thattai and Alexander Van Oudenaarden. Stochastic gene expression in fluctuating environments. *Genetics*, 167(1):523–530, 2004.
- [622] Tianhai Tian and Kevin Burrage. Stochastic neural network models for gene regulatory networks. In *Evolutionary Computation, 2003. CEC'03. The 2003 Congress on*, volume 1, pages 162–169. IEEE, 2003.
- [623] Tianhai Tian and Kevin Burrage. Stochastic neural network models for gene regulatory networks. In *Evolutionary Computation, 2003. CEC'03. The 2003 Congress on*, volume 1, pages 162–169. IEEE, 2003.
- [624] Michael E Tipping. Sparse bayesian learning and the relevance vector machine. *The journal of machine learning research*, 1:211–244, 2001.
- [625] Peter Toft, Egill Rostrup, Finn A. Nielsen, Finn Arup Nielsen, Lars Kai Hansen, Cyril Goutte, and Cyril Goutte. On clustering fmri time series. *Neuroimage*, 9:298–310, 1998.
- [626] Hannu Toivonen. Sampling large databases for association rules. In *Proceedings of the 22th International Conference on Very Large Data Bases, VLDB '96*, pages 134–145, San Francisco, CA, USA, 1996. Morgan Kaufmann Publishers Inc.
- [627] Hannu Toivonen et al. Sampling large databases for association rules. In *VLDB*, volume 96, pages 134–145, 1996.
- [628] Vera Tomaino, Pietro H. Guzzi, Mario Cannataro, and Pierangelo Veltri. Experimental comparison of biclustering algorithms for ppi networks. In *BCB*, pages 671–676. ACM, 2010.
- [629] Yongxin Tong, Lei Chen, Yurong Cheng, and Philip S. Yu. Mining frequent itemsets over uncertain databases. *Proc. VLDB Endow.*, 5(11):1650–1661, July 2012.
- [630] G. Toussaint. Bibliography on estimation of misclassification. *IEEE Transactions on Information Theory*, 20(4):472–479, 1974.
- [631] O. Troyanskaya, M. Cantor, G. Sherlock, P. Brown, T. Hastie, R. Tibshirani, D. Botstein, and R. B. Altman. Missing value estimation methods for DNA microarrays. *Bioinformatics*, 17(6):520–525, June 2001.
- [632] Duy Tin Truong, R. Battiti, and M. Brunato. Discovering non-redundant overlapping biclusters on gene expression data. In *ICDM*, pages 747–756, 2013.
- [633] Vincent Tseng and Chao-Hui Lee. Effective temporal data classification by integrating sequential pattern mining and probabilistic induction. *Expert Sys.App.*, 36(5):9524–9532, 2009.
- [634] Vincent Tseng and Chao-Hui Lee. Effective temporal data classification by integrating sequential pattern mining and probabilistic induction. *Expert Sys.App.*, 36(5):9524–9532, July 2009.
- [635] Vincent S. Tseng and Chao-Hui Lee. Effective temporal data classification by integrating sequential pattern mining and probabilistic induction. *Expert Systems with Applications*, 36(5):9524 – 9532, 2009.
- [636] G. Tsoumakas and I. Katakis. Multi Label Classification: An Overview. *International Journal of Data Warehouse and Mining*, 3(3):1–13, 2007.
- [637] Allan Tucker, Stephen Swift, and Xiaohui Liu. Variable grouping in multivariate time series via correlation. *Systems, Man, and Cybernetics, Part B: Cybernetics, IEEE Transactions on*, 31(2):235–245, 2001.
- [638] Heather Turner, Trevor Bailey, and Wojtek Krzanowski. Improved biclustering of microarray data demonstrated through systematic performance tests. *Comp. Stat. & Data Analysis*, 48(2):235–254, 2005.

- [639] Petre Tzvetkov, Xifeng Yan, and Jiawei Han. Tsp: Mining top-k closed sequential patterns. *Knowledge and Information Systems*, 7(4):438–457, 2005.
- [640] Miranda van Uitert, Wouter Meuleman, and Lodewyk Wessels. Biclustering sparse binary genomic data. *Journal of Computational Biology*, 15(10):1329–1345, 2008.
- [641] Igor Ulitsky and Ron Shamir. Identification of functional modules using network topology and high-throughput data. *BMC systems biology*, 1(1):8, 2007.
- [642] Takeaki Uno, Masashi Kiyomi, and Hiroki Arimura. Lcm ver.3: collaboration of array, bitmap and prefix tree for frequent itemset mining. In *OSDM*, pages 77–86. ACM, 2005.
- [643] Tomohiro Ushikubo, Wataru Inoue, Mitsumasa Yoda, and Masaki Sasai. Testing the transition state theory in stochastic dynamics of a genetic switch. *Chemical physics letters*, 430(1):139–143, 2006.
- [644] John W. van Ness and Cary Simpson. On the effects of dimension in discriminant analysis. *Technometrics*, 18(2):175–187, 1976.
- [645] Vladimir Vapnik. *Estimation of Dependences Based on Empirical Data: Springer Series in Statistics*. Springer-Verlag New York, Inc., Secaucus, NJ, USA, 1982.
- [646] Vladimir N. Vapnik. *Statistical Learning Theory*. Wiley-Interscience, 1998.
- [647] Vladimir Naumovich Vapnik and Vladimir Vapnik. *Statistical learning theory*, volume 2. Wiley New York, 1998.
- [648] Nicolas Vayatis and Robert Azencott. Distribution-dependent vapnik-chervonenkis bounds. In Paul Fischer and HansUlrich Simon, editors, *Computational Learning Theory*, volume 1572 of *LNCS*, pages 230–240. Springer Berlin Heidelberg, 1999.
- [649] Adriano Veloso, Jr. Wagner Meira, and Mohammed J. Zaki. Lazy associative classification. In *ICDM*. IEEE, 2006.
- [650] Alessia Visconti, Francesca Cordero, and Ruggero G. Pensa. Leveraging additional knowledge to support coherent bicluster discovery in gene expression data. *Intell. Data Anal.*, 18(5):837–855, 2014.
- [651] Aleksandra M Walczak, Andrew Mugler, and Chris H Wiggins. A stochastic spectral analysis of transcriptional regulatory cascades. *Proceedings of the National Academy of Sciences*, 106(16):6529–6534, 2009.
- [652] E.A. Wan. Temporal backpropagation for fir neural networks. In *IJC on Neural Networks*, pages 575–580 vol.1, 1990.
- [653] Haixun Wang, Wei Wang, Jiong Yang, and Philip S. Yu. Clustering by pattern similarity in large data sets. In *SIGMOD*, pages 394–405, New York, NY, USA, 2002. ACM.
- [654] Jianyong Wang and Jiawei Han. Bide: Efficient mining of frequent closed sequences. In *ICDE*, pages 79–, Washington, DC, USA, 2004. IEEE Computer Society.
- [655] Jianyong Wang, Jiawei Han, and Jian Pei. Closet+: searching for the best strategies for mining frequent closed itemsets. In *Proceedings of the ninth ACM SIGKDD international conference on Knowledge discovery and data mining*, KDD '03, pages 236–245, New York, NY, USA, 2003. ACM.
- [656] Jianyong Wang and George Karypis. Harmony: Efficiently mining the best rules for classification. In *SDM*, volume 5, pages 205–216. SIAM, 2005.
- [657] Ke Wang, Yu He, and Jiawei Han. Pushing support constraints into association rules mining. *Knowledge and Data Engineering, IEEE Transactions on*, 15(3):642–658, 2003.
- [658] Ke Wang, Yuelong Jiang, and Laks V. S. Lakshmanan. Mining unexpected rules by pushing user dynamics. In *Proceedings of the ninth ACM SIGKDD international conference on Knowledge discovery and data mining*, KDD '03, pages 246–255, New York, NY, USA, 2003. ACM.
- [659] Liang Wang, Reynold Cheng, Sau Dan Lee, and David Cheung. Accelerating probabilistic frequent itemset mining: A model-based approach. In *Proceedings of the 19th ACM International Conference on Information and Knowledge Management*, CIKM '10, pages 429–438, New York, NY, USA, 2010. ACM.
- [660] Miao Wang, Xuequn Shang, Xiaoyuan Li, Zhanhuai Li, and Wenbin Liu. Efficient mining differential co-expression constant row bicluster in real-valued gene expression datasets. *Applied Mathematics & Information Sciences*, 7(2), 2013.
- [661] Miao Wang, Xuequn Shang, Shaohua Zhang, and Zhanhuai Li. Fdcluster: mining frequent closed discriminative bicluster without candidate maintenance in multiple microarray datasets. In *Data Mining Workshops (ICDMW), 2010 IEEE International Conference on*, pages 779–786. IEEE, 2010.
- [662] Shu Wang, Robin R Gutell, and Daniel P Miranker. Biclustering as a method for rna local multiple sequence alignment. *Bioinformatics*, 23(24):3289–3296, 2007.
- [663] Zhiguan Wang, Chi Wai Yu, Ray CC Cheung, and Hong Yan. Hypergraph based geometric biclustering algorithm. *Pattern Recognition Letters*, 33(12):1656–1665, 2012.
- [664] Zidong Wang, Huijun Gao, Jinde Cao, and Xiaohui Liu. On delayed genetic regulatory networks with polytopic uncertainties: robust stability analysis. *NanoBioscience, IEEE Transactions on*, 7(2):154–163, 2008.
- [665] Zidong Wang, Xiaohui Liu, Yurong Liu, Jinling Liang, and Veronica Vinciotti. An extended kalman filtering approach to modeling nonlinear dynamic gene regulatory networks via short gene expression time series. *IEEE/ACM Transactions on Computational Biology and Bioinformatics (TCBB)*, 6(3):410–419, 2009.
- [666] Zidong Wang, Fuwen Yang, Daniel WC Ho, Stephen Swift, Allan Tucker, and Xiaohui Liu. Stochastic dynamic modeling of short gene expression time-series data. *NanoBioscience, IEEE Transactions on*, 7(1):44–55, 2008.
- [667] Joe H. Ward. Hierarchical grouping to optimize an objective function. *Journal of the American Statistical Association*, 58(301):236–244, 1963.
- [668] Takashi Washio and Hiroshi Motoda. State of the art of graph-based data mining. *SIGKDD Explor. Newsl.*, 5(1):59–68, July 2003.
- [669] Ted Way, Berkman Sahiner, Lubomir M. Hadjiiski, and Heang-Ping Chan. Effect of finite sample size on feature selection and classification: A simulation study. *Medical Physics*, 37(2):907–920, 2010.
- [670] G. Webb. Discovering significant patterns. *Machine Learning*, 68(1):1–33, 2007.
- [671] GI Webb. *Magnum opus version 3.0. 1. Software, GI Webb & Associates, Melbourne, Aust*, 2005.
- [672] Peter H. Westfall and S. Stanley Young. *Resampling-Based Multiple Testing: Examples and Methods for p-Value Adjustment*. John Wiley & Sons, 1993.
- [673] TomF. Wilderjans, Dirk Depril, and Iven Van Mechelen. Additive biclustering: A comparison of one new and two existing als algorithms. *Journal of Classification*, 30(1):56–74, 2013.
- [674] Donald Wlodkovic, Joanna Skommer, Dagmara McGuinness, Chris Hillier, and Zbigniew Darzynkiewicz. Er-golgi network: A future target for anti-cancer therapy. *Leukemia research*, 33(11):1440–1447, 2009.
- [675] Fang-Xiang Wu, Wen-Jun Zhang, and Anthony J Kuslik. Modeling gene expression from microarray expression data with state-space equations. In *Pacific Symposium on Biocomputing*, volume 9, pages 581–592, 2004.
- [676] Wei-Sheng Wu and Wen-Hsiung Li. Identifying gene regulatory modules of heat shock response in yeast. *BMC genomics*, 9(1):439, 2008.
- [677] Xiaopeng Xi, Eamonn Keogh, Christian Shelton, Li Wei, and Chotirat Ann Ratanamahatana. Fast time series classification using numerosity reduction. In *ICML*, pages 1033–1040. ACM, 2006.

- [678] Hu Xia, Jian Zhuang, and Dehong Yu. Novel soft subspace clustering with multi-objective evolutionary approach for high-dimensional data. *Pattern Recognition*, 46(9):2562–2575, 2013.
- [679] Rongjing Xiang, Jennifer Neville, and Monica Rogati. Modeling relationship strength in online social networks. In *IC on World wide web, WWW*, pages 981–990, New York, USA, 2010. ACM.
- [680] Dong Xin, Hong Cheng, Xifeng Yan, and Jiawei Han. Extracting redundancy-aware top-k patterns. In *Proceedings of the 12th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, KDD '06*, pages 444–453, New York, NY, USA, 2006. ACM.
- [681] Dong Xin, Zheng Shao, Jiawei Han, and Hongyan Liu. C-cubing: Efficient computation of closed cubes by aggregation-based checking. In Ling Liu, Andreas Reuter, Kyu-Young Whang, and Jianjun Zhang, editors, *ICDE*, page 4. IEEE Computer Society, 2006.
- [682] Eric P Xing, Michael I Jordan, Richard M Karp, et al. Feature selection for high-dimensional genomic microarray data. In *ICML*, volume 1, pages 601–608. Citeseer, 2001.
- [683] Hui Xiong, Xiao-Feng Heb, Chris Ding, Ya Zhang, Vipin Kumar, and Stephen R Holbrook. Identification of functional modules in protein complexes via hyperclique pattern discovery. *Pac. Symp. Biocomput.*, 2005.
- [684] Hui Xiong, Pang-Ning Tan, and Vipin Kumar. Hyperclique pattern discovery. *DMKD*, 13(2):219–242, 2006.
- [685] X. Yan and Jiawei Han. gspan: graph-based substructure pattern mining. In *Data Mining, 2002. ICDM 2003. Proceedings. 2002 IEEE International Conference on*, pages 721–724, 2002.
- [686] Xifeng Yan, Hong Cheng, Jiawei Han, and Dong Xin. Summarizing itemset patterns: A profile-based approach. In *KDD*, pages 314–323, New York, NY, USA, 2005. ACM.
- [687] Xifeng Yan, Jiawei Han, and Ramin Afshar. CloSpan: Mining Closed Sequential Patterns in Large Datasets. In *In SDM*, pages 166–177, 2003.
- [688] Cheng Yang, Usama Fayyad, and Paul S. Bradley. Efficient discovery of error-tolerant frequent itemsets in high dimensions. In *Proceedings of the seventh ACM SIGKDD international conference on Knowledge discovery and data mining, KDD '01*, pages 194–203, New York, NY, USA, 2001. ACM.
- [689] Jiong Yang, Haixun Wang, Wei Wang, and Philip Yu. Enhanced biclustering on expression data. In *Bioinformatics and Bioengineering, 2003. Proceedings. Third IEEE Symposium on*, pages 321–327. IEEE, 2003.
- [690] Jiong Yang, Wei Wang, Haixun Wang, and P. Yu. delta-clusters: capturing subspace correlation in a large data set. In *ICDE*, pages 517–528, 2002.
- [691] Jiong Yang, Wei Wang, Philip S. Yu, and Jiawei Han. Mining long sequential patterns in a noisy environment. In *SIGMOD*, pages 406–417, New York, NY, USA, 2002. ACM.
- [692] Jiong Yang, Wei Wang, Philip S Yu, and Jiawei Han. Mining long sequential patterns in a noisy environment. In *Proceedings of the 2002 ACM SIGMOD international conference on Management of data*, pages 406–417. ACM, 2002.
- [693] Qiang Yang and Xindong Wu. 10 challenging problems in data mining research. *International Journal of Information Technology and Decision Making*, 5(4):597–604, 2006.
- [694] Wen-Hui Yang, Dao-Qing Dai, and Hong Yan. Finding correlated biclusters from gene expression data. *Knowledge and Data Engineering, IEEE Transactions on*, 23(4):568–584, April 2011.
- [695] Zhenglu Yang, Yitong Wang, and Masaru Kitsuregawa. Lapin: effective sequential pattern mining algorithms by last position induction for dense databases. In *DASFAA*, pages 1020–1023, Berlin, Heidelberg, 2007. Springer-Verlag.
- [696] Xiaoxin Yin and Jiawei Han. Cpar: Classification based on predictive association rules. In *SDM*, volume 3, pages 331–335. SIAM, 2003.
- [697] K.Y. Yip, Ben Kao, Xinjie Zhu, Chun Kit Chui, Sau Dan Lee, and D.W. Cheung. Mining order-preserving submatrices from data with repeated measurements. *Knowledge and Data Engineering, IEEE Transactions on*, 25(7):1587–1600, July 2013.
- [698] Illhoi Yoo, Patricia Alafaireet, Miroslav Marinov, Keila Pena-Hernandez, Rajitha Gopidi, Jia-Fu Chang, and Lei Hua. Data mining in healthcare and biomedicine: A survey of the literature. *Journal of Medical Systems*, 36(4):2431–2448, 2012.
- [699] Kun-Ming Yu, Jiayi Zhou, and WeiChen Hsiao. Load balancing approach parallel algorithm for frequent pattern mining. In *Parallel Computing Technologies*, volume 4671 of *LNCS*, pages 623–631. Springer, 2007.
- [700] Mohammed J. Zaki. Spade: An efficient algorithm for mining frequent sequences. *Mach. Learn.*, 42(1-2):31–60, January 2001.
- [701] Mohammed J. Zaki and Karam Gouda. Fast vertical mining using diffsets. In *KDD*, pages 326–335, New York, NY, USA, 2003. ACM.
- [702] Mohammed J. Zaki and Ching-Jui Hsiao. Efficient algorithms for mining closed itemsets and their lattice structure. *IEEE TKDE*, 17(4):462–478, April 2005.
- [703] Mohammed J Zaki, Srinivasan Parthasarathy, Mitsunori Ogihara, and Wei Li. Parallel algorithms for discovery of association rules. *Data Mining and Knowledge Discovery*, 1(4):343–373, 1997.
- [704] H. Zhang, B. Padmanabhan, and A. Tuzhilin. On the discovery of significant statistical quantitative rules. In *KDD*, pages 374–383. ACM, 2004.
- [705] Jing Zhang, Bo Jiang, Ming Li, John Tromp, Xuegong Zhang, and Michael Q Zhang. Computing exact p-values for dna motifs. *Bioinformatics*, 23(5):531–537, 2007.
- [706] Lijun Zhang, Chun Chen, Jiajun Bu, Zhengguang Chen, Deng Cai, and Jiawei Han. Locally discriminative coclustering. *Knowledge and Data Engineering, IEEE Transactions on*, 24(6):1025–1035, 2012.
- [707] Min-Ling Zhang and Zhi-Hua Zhou. A k-nearest neighbor based algorithm for multi-label classification. *IEEE International Conference on Granular Computing*, 2:718–721 Vol. 2, 2005.
- [708] Yuji Zhang, Jianhua Xuan, Benildo G de los Reyes, Robert Clarke, and Habtom W Resson. Reverse engineering module networks by pso-rnn hybrid modeling. *BMC genomics*, 10(Suppl 1):S15, 2009.
- [709] Hongya Zhao, K Chan, L Cheng, and H Yan. A probabilistic relaxation labeling framework for reducing the noise effect in geometric biclustering of gene expression data. *Pattern Recognition*, 42(11):2578–2588, 2009.
- [710] Lizhuang Zhao and Mohammed J. Zaki. tricluster: An effective algorithm for mining coherent clusters in 3d microarray data. In *In Proc. of the 2005 ACM SIGMOD international conference on Management of data*, pages 694–705. ACM Press, 2005.
- [711] Yu Zheng, Lizhu Zhang, Xing Xie, and Wei-Ying Ma. Mining interesting locations and travel sequences from gps trajectories. In *WWW*, pages 791–800. ACM, 2009.
- [712] Feida Zhu, Xifeng Yan, Jiawei Han, and Philip S Yu. *Mining Frequent Approximate Sequential Patterns*, pages 66–87. Chapman&Hall, 2009.
- [713] Feida Zhu, Xifeng Yan, Jiawei Han, and PhilipS. Yu. gprune: A constraint pushing framework for graph pattern mining. In Zhi-Hua Zhou, Hang Li, and Qiang Yang, editors, *Advances in Knowledge Discovery and Data Mining*, volume 4426 of *Lecture Notes in Computer Science*, pages 388–400. Springer Berlin Heidelberg, 2007.
- [714] Feida Zhu, Xifeng Yan, Jiawei Han, P.S. Yu, and Hong Cheng. Mining colossal frequent patterns by core pattern fusion. In *Data Engineering, 2007. ICDE 2007. IEEE 23rd International Conference on*, pages 706–715, April 2007.
- [715] Albrecht Zimmermann and Björn Bringmann. Aggregated subset mining. In *Advances in Knowledge Discovery and Data Mining*, pages 664–672. Springer, 2009.