EuroVis Workshop on Visual Analytics (2019)
C. Turkay and T. von Landesberger (Editors)

Visual Analysis of Degree-of-Interest Functions to Support Selection Strategies for Instance Labeling

Jürgen Bernard [1], Marco Hutter[1], Christian Ritter [1], Markus Lehmann[1], Michael Sedlmair[2], and Matthias Zeppelzauer[3]

[1]TU Darmstadt, Darmstadt, Germany
[2] University of Stuttgart, Stuttgart, Germany
[3]St. Pölten University of Applied Sciences, Austria

Abstract

Manually labeling data sets is a time-consuming and expensive task that can be accelerated by interactive machine learning and visual analytics approaches. At the core of these approaches are strategies for the selection of candidate instances to label. We introduce degree-of-interest (DOI) functions as atomic building blocks to formalize candidate selection strategies. We introduce a taxonomy of DOI functions and an approach for the visual analysis of DOI functions, which provide novel complementary views on labeling strategies and DOIs, support their in-depth analysis and facilitate their interpretation. Our method shall support the generation of novel and better explanation of existing labeling strategies in future.

1. Introduction

Labeling data sets is of utmost importance to conduct supervised machine learning. However, labeling is an expensive and time-consuming task, requiring humans in-the-loop. To automate and accelerate this process, different approaches have been introduced. One class of approaches is active learning (AL) [Set12] where the machine learner proactively asks the user for labels and semi-supervised learning, where known labels are propagated to unlabeled instances [BTF18, LWG19]. In contrast, user-centered approaches employ visual interfaces [SG10, HKBE12, HNH*12, BLBC12], enabling humans to identify, select, and label interesting instances. Finally, there are attempts to combine the strengths of both in a visual-interactive labeling (VIAL) process [BZSA18].

At the core of all above approaches are *labeling strategies*, which are functional rules that determine which instance shall be selected for labeling next. The quality of the strategy strongly influences the learning and labeling progress. Bringing different selection strategies down to the same level of description would enable first to formalize selection strategies and second to systematically analyze selection strategies, their similarities and unique characteristics. A formalization of selection strategies would further help to combine existing approaches and to facilitate the development of novel strategies. A basic requirement for such a formalization is the identification of common *building blocks* for selection strategies which has not been investigated so far and is the topic of this paper.

A recent observational study revealed that selection strategies seem to be composed of a variety of more basic building blocks [BHL*18]. The same appears to apply to selection strategies observed from user behavior [BZL*18]. Based on these results, we introduce a simple type of building block, i.e., *degree-of-interest* (DOI) functions that enable the composition of complex labeling strategies through, e.g., meaningful linear combination. From the large set of possible DOIs, each focusing on different criteria of the underlying data or the involved classification model, combinations can be formed to represent either existing selection strategies or to yield completely new selection strategies. To ease the design of novel strategies and to facilitate their analysis, comparison, and interpretation by analysts, this paper provides two contributions:

- We present a taxonomy of DOIs for labeling in Section 2, based on a review of existing strategies. The taxonomy is hierarchically organized, according to the characteristics captured by the DOIs, and provides a basic organizing principle for DOIs. This taxonomy supports the design and comparison of DOIs with the goal to create novel, more effective strategies for instance selection.
- We present a visual-interactive approach for the structural analysis of DOIs as a first approach to (visually) investigate and explain the characteristics and behavior of DOIs in Section 3. The aim of our visualization approach is threefold and targets the identification of interesting (a) data-related, (b) model-related, and (c) temporal characteristics of DOIs in the labeling process.

The proposed approach is meant to enable a deeper understanding of labeling strategies in AL and VA and foster explainability of approaches building upon them.

2. A Taxonomy of Degree-of-Interest Functions

2.1. Degree-of-Interest Functions

Inspired by the work by Bernard et al. [BZL*18] who created low-level building blocks for ten high-level user strategies, we introduce the general concept of *degree-of-interest* (DOI) functions. DOI functions are functional components that capture basic aspects of labeling strategies. A DOI receives a set of instances $X \in \mathbb{R}^d$ and an index k to one instance X_k as input and assigns an interestingness score i to X_k, e.g., in the range $[0,...,1]$. A DOI function f is thus formally defined as $i = f(X,k)$. When f is applied to all instances of set X, a ranking by interestingness of the instances in X with respect to the DOI f can be computed. One DOI alone does not necessarily yield a good labeling performance since it focuses only on

DOI: 10.2312/eurova.20191116

DOIs	Description: DOIs based on...	Surveys & References
Data-Based	**characteristics of the data**	
Clustering	**the results of clustering algorithms**	[Jai10]
Single Clustering	**the result of a single clustering algorithm**	[HBV02]
Cluster Characterist.	**characteristics of relations of instance to**	[HBV02, BZL*18, STMT12]
Cluster Compactness	**within-cluster compactness (lower values are better)**	[HBV02, Dun74, Rou87]
Cluster Separation	**between-cluster separation (higher values are better)**	[HBV02, Dun74, Rou87]
Committee Results	**the results of multiple clustering algorithms (ensemble)**	
Density	**the local data density in the vicinity of an instance**	[BSB*15, BZL*18]
Outliers	**outlier detection**	[KN98, BKNS00a, RRS00, CBK09]
Model-Based	**characteristics of the underlying model (classifier)**	
Uncertainty	**probability distributions for instances assigned by the classifier**	[Set12, WKBD06, SC08]
Relevance	**the probability distributions for instances assigned by the classifier**	[VPS*02]
Spatialization	**spatial information and relations between high-dimensional data**	
Class Relations	**relations of instances to class characteristics (centroids, spread, etc.)**	
Class Characteristics	**uncertainty caused by class spatialization**	[STMT12]
Class Compactness	**within-class compactness (lower values are better)**	[STMT12, Dun74, Rou87]
Class Separation	**between-class separation (higher values are better)**	[STMT12, Dun74, Rou87]
Neighbor Relations	**neighbor instances**	
Neighbor Votes	**the diversity of winning class labels (votes) of k nearest neighbors**	[Sha48, Sim49]
Neighbor Probabilities	**the comparison of probability distributions among k-NN**	[Kol33, Smi48, KL*51, FT04]
Neighbor Prob. Aggr.	**aggregated probability distributions among k-NN**	
Error Reduction	**heuristics to reduce classification errors**	[Set12]
Expected model change	**selecting the instance that would change the model most**	[SCR08]
Risk reduction	**selecting the instance that reduces generalization error most**	[QHR*09]
Variance reduction	**selecting the instance that reduces output variance most**	[HJL06]
Committees	**a committee of classification models (ensemble)**	[SOS92, Mam98, Set12]
Votes	**the diversity of winning class labels (votes) of the committee**	[SOS92, Sha48, Sim49]
Probabilities	**the divergence of probability distributions proposed by the committee**	[KL*51, Kol33, Smi48, FT04]

Table 1: *The taxonomy of DOI functions with a short explanation of each category and related references.*

one aspect. A *strategy* is usually composed of multiple DOIs, e.g., combined by a weighted linear combination, i.e. i = $\sum_{j=1}^{n} w_j * f_j(.)$.

2.2. Overview of the DOI Taxonomy

DOIs represent common building blocks for the formal description of selection strategies and thereby a vocabulary for their definition. The proposed taxonomy tries to organize the broad set of DOIs in a common structure, see Table 1. **Data-based** and **Model-based** DOIs form the highest level. Data-based DOIs represent criteria for interestingness such as density, outlierness, and cluster-based measures and can be computed in a completely unsupervised way. Model-based DOIs, compute the interestingness of instances with respect to a (pre-trained or just initialized) classification model M, which extends their formal definition to $i = f(X,k,M)$.

2.3. Data-Based DOIs

Data-based DOIs comprise three subgroups targeting different data characteristics: **Clustering**, **Density**, and **Outliers** (cf. Table 1).

Clustering involves DOIs that compute interestingness with respect to a previous clustering in the feature space X. Clustering reveals basic structures in the data and can be useful for the selection of candidates, e.g., representative instances for a class. Clustering DOIs can be computed from *single clustering* as well as from a *committee* result. Another differentiation is the type of characteristics used to calculate the interestingness for an instance. Classes of criteria include (i) *cluster characteristics* (e.g., size of clusters), (ii) *cluster compactness* (e.g., the modified Hubert statistic or the Dunn family of indices [Dun74, HBV02]), and (iii) *cluster separation* (e.g. used for the Silhouettes index [Rou87]). In Figure 3, we use a Cluster Centroid Distances DOI based on a hierarchical clustering result and k-Means.

Density follows the objective to assign high interestingness scores to instances in highly dense areas in the feature space. This is important since selecting instances in dense areas could provide good representatives of the data. Density measures can be used for implementation (e.g. criteria used in density-based clustering [EKS*96]) as well as distance measures to nearby samples [WKBD06]. An example DOI is counting the number of instances in the ε range of an instance [BZL*18] used in Figure 3.

Outliers includes criteria to highlight anomalies in the sample distribution in X. Outliers may represent unusual or extreme cases for a certain class and may help to delimit class boundaries [CBK09]. A variety of algorithms can be used, such as statistics based on angles [KShZ08] or neighborhood distances [KN98] like the local outlier factor [BKNS00b] (cf. Figure 3).

2.4. Model-Based DOIs

Model-based DOIs take the underlying classification model into account. They can be partitioned into five subclasses (cf. Table 1).

2.4.1. Uncertainty-based DOIs

This group of DOIs estimate the interestingness of an instance w.r.t. the uncertainty of a given classifier. According to the AL principles [Set12], those instances for which a classifier is most uncertain about are with high probability those the classifier can learn most from. The uncertainty of a classifier can be assessed from its (probabilistic) outputs (class scores) (e.g. the variance of class scores obtained for an instance) or from metric measurements in feature space, e.g. the distance of an instance from the closest decision boundary [WKBD06, VPS*02]. Example implementations adopted from AL strategies are *Least Significant Confidence* [Set12], *Smallest Margin* [WKBD06] (cf. Figure 1), as well as *Entropy* [VPS*02].

2.4.2. Relevance-based DOIs

Interestingness is related to the relevance of an instance for a class, i.e., typical and highly representative samples for a class shall be selected. Relevance fosters the identification of positive examples for a class instead of difficult instances and is thus complementary to uncertainty-based DOIs. Relevance-based selection fosters the construction of (initial) class models and is particularly useful in early phases of the labeling process [BHL*18] as well as in ranking-based retrieval settings [WH11]. Relevance criteria can be derived by looking for samples the classifier is most certain about [AQ07]. In absence of a trained model, similar instances to already labeled ones can be considered relevant. An example DOI is to select the most probable positive instances for a given class [AQ07]

2.4.3. Spatialization-based DOIs

Interestingness scores are computed from spatial relations of instances and modeled classes in the feature space, as well as from predictions for neighboring instances. This results in more general DOIs complementary to uncertainty- and relevance-based DOIs. DOIs in this group help to identify interesting areas in the feature space, e.g. where ambiguities exist. Subgroup *Class relations* captures relations between instances and classes and is similarly organized as subgroup *Clustering* of data-based DOIs. Subgroup *Neighbor relations* compares the predicted labels of instances to their neighbors on the basis of class predictions (*neighbor votes*), class probability distributions (*neighbor probabilities*), and aggregations of probability distributions from neighbors (*neighbor prob. aggr.*). A DOI could, for example, estimate interestingness as the entropy of predicted labels for the k nearest neighbors, i.e., instances surrounded by differently labeled instances are considered interesting.

2.4.4. Error-reduction-based DOIs

One group of DOIs focuses on instances that are expected to reduce the error of a classifier most (or maximize its generalization ability). Such DOIs foster the identification of the most valuable instances for the model (without knowing their true label). Evaluating the impact of an instance to a classification model is, however, time-consuming as it may require retraining. Selection criteria are *expected model change* [SCR08] and *risk reduction* [QHR*09]. A more efficient alternative is *variance reduction* [HJL06] which can avoid retraining for certain classifiers. An example DOI function would be to estimate interestingness of an instance from the expected reduction of the classification error that this instance yields when added to the training set (with different label hypotheses).

2.4.5. Committee-based DOIs

This group of DOIs derive interestingness scores from a committee of classifiers. The use of a committee reduces the possible bias on model-based DOIs from a particular classifier [Set12]. We differentiate between DOIs that simply measure the (dis)agreement of the committee (*Votes*) and DOIs relying on the (dis)agreement of the predicted probabilities (*Probabilities*). An example DOI function is to measure the divergence [KL*51] of the probability distributions of class scores of the committee for every instance. High divergence thereby yields a high interestingness score [Set12].

3. Visual Analysis of DOI Characteristics

To better understand and explain the behavior and characteristics of different DOIs, we introduce a VA approach for their detailed analysis. The approach supports two primary goals 1) the validation of DOIs *designed* by an analyst as well as 2) understanding the characteristics of DOIs *applied* by an analyst. To address these goals, we design VA approaches to unveil DOI characteristics with respect to three complementary perspectives, each of which forms an individual analysis task: The DOIs can be analyzed with respect to:

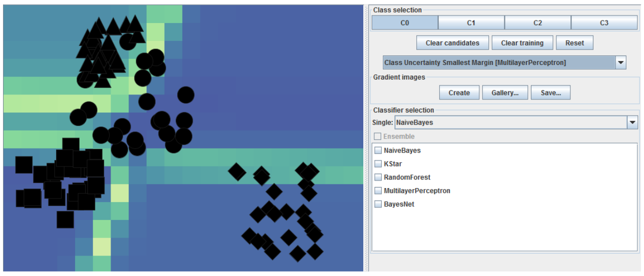

Figure 1: *Prototype for the interactive analysis of DOIs. DOIs and underlying classifiers can be selected by the analyst. Data points (black) can be moved (drag-and-drop) to conduct what-if-analysis. In the example, the "Dense Coloring" mode is active (representing interestingness distribution in the data space). The selected Smallest Margin DOI clearly unveils the decision boundary of the Naive Bayes classifier which empirically validates its correct implementation and supports the interpretation of DOI and classifier.*

- Data characteristics ($\mathbf{T}_1$)
- Model (classifier) characteristics ($\mathbf{T}_2$)
- Changes in the labeling process ($\mathbf{T}_3$)

3.1. Visual-Interactive Analysis of DOIs

DOIs are the starting point of the analysis. After DOI selection, analysts may want to explore the interestingness of areas in the feature space ($\mathbf{T}_1$) and how the DOI relates to the classifier and its predictions ($\mathbf{T}_2$). Furthermore, an analyst may be interested in how DOIs change their interestingness scores in response to changes in the labeling process, i.e., an increased number of labeled instances ($\mathbf{T}_3$). Figure 1 shows our VA prototype.

Data Space Coloring versus Instance Coloring We provide two visualization techniques, both of which use colors to encode interestingness scores (bright means most interesting) [TFS08]. In the *Dense Coloring* mode, we compute the DOI in the visual (output) space, based on screen coordinates (e.g., Figure 2). As a result, the entire screen space will be colored, similar to techniques for the visualization of classifier decision boundaries [MHT18]. Dense Coloring is suitable either for 2D data sets (instances mapped to pixels), or for multivariate data in combination with dimensionality reduction techniques that exhibit an inverse mapping [ERT19]. In the *Instance Coloring* mode, we directly show individual instances colored w.r.t. their interestingness scores [BHL*18] (e.g., Figure 3). This is possible for 2D data sets (e.g., for DOI validation) and for multivariate data in combination with dimensionality reduction.

What-If-Analysis We provide an interaction technique that supports the assessment of model and DOI changes triggered by data manipulations. To facilitate what-if-analysis, we enable data editing via interactively dragging instances. Instant recalculation of models and DOIs allows the immediate assessment of effects caused by data changes. We demonstrate the effect in a supplemental video and in Figure 2 for an artificial 2D data set where the Smallest Margin DOI reveals changes in the decision boundary caused by moving an instance. Figure 2 further illustrates how

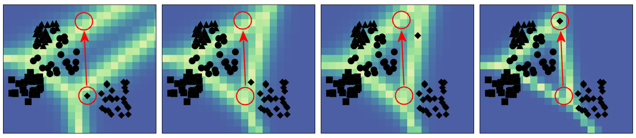

Figure 2: *What-if-analysis to better understand the interplay of DOI and model (Smallest Margin DOI applied with a simple logistic classifier): moving a single data element causes a considerable change of the decision boundaries of the classifier.*

dense mapping in combination with an appropriate DOI (smallest margin in this case) can be exploited to visualize classifier internals such as the decision boundaries. Note that this approach is classifier agnostic and can be used to visualize classification boundaries of arbitrary classifiers.

3.2. Analyzing DOIs w.r.t. data characteristics ($\mathbf{T}_1$)

Data-based DOIs are useful to support analysts in exploratory data analysis. Instance Coloring provides a useful visual interface for this purpose. The example in Figure 3 shows how different DOIs behave for a given data distribution (artificial 2D data set).

3.3. Analyzing DOIs w.r.t. model characteristics ($\mathbf{T}_2$)

For $\mathbf{T}_2$ we present a visual interface that facilitates the exploration of DOIs and their interplay with the model they are bound to. Given a model-based DOI, analysts can compare its interestingness estimates for different underlying classifiers. Figure 4 shows the relations of an uncertainty-based DOI to six different classifiers and how this reveals intrinsic properties of classifiers.

3.4. Analyzing DOIs w.r.t. changes in the labeling process ($\mathbf{T}_3$)

The third analysis task supports the assessment of changes caused by iteratively increasing the number of labeled training instances. Given some instance selection strategy, the interestingness values of its DOIs can be analyzed over time. Individual iterations of the labeling process are aligned next to each other in a small-multiples manner. Analysts can explore the changing DOI scores during the labeling process to better understand their behavior. In Figure 5 a labeling process is analyzed for a Spatial Balancing DOI. Analysts may further use this visual interface to conduct what-if analysis, i.e. what happens if a particular instance is labeled next?

4. Conclusions

We introduced *Degree-of-Interest* (DOI) functions as basic building blocks of instance selection strategies together with a taxonomy that systematically organizes them. To foster interpretability,

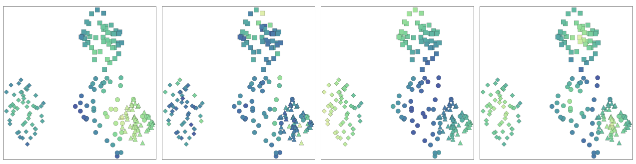

Figure 3: *Analysis of data-related DOI characteristics (Instance Coloring mode). A density-based, an outlier-based and two clustering-based DOIs are compared. These characteristics may play an important role in the design of future labeling strategies.*

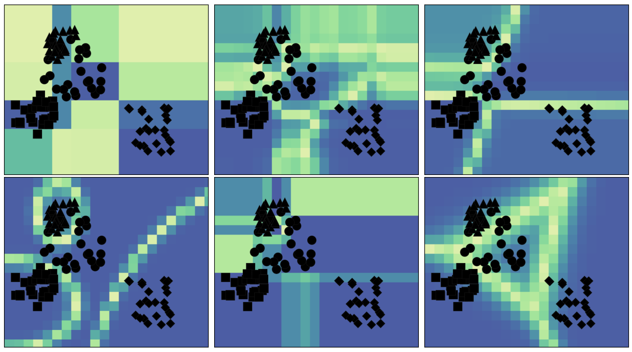

Figure 4: *Dense Coloring mode used for the analysis of a model-based DOI (Smallest Margin) and its relation to six different classifiers (from left to right and top to bottom: BayesNet, KStar, Multilayer Perceptron, Naive Bayes, Random Forest, Simple Logistic). The DOI reveals regions in the data space where classifiers are unsure, i.e., the decision boundaries in this case. The visualization helps to understand the nature of the DOI and here, as a byproduct, unveils interesting intrinsic properties of the classifiers.*

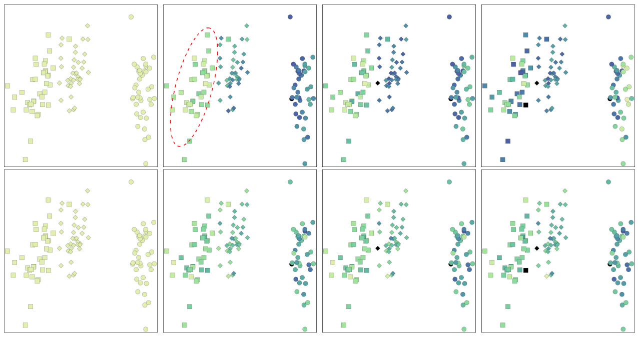

Figure 5: *Observation of a labeling process for the Iris data set [Lic13] (0-3 instances labeled, black colors) (*$\mathbf{T}_3$*). Top: a Spatial Balancing DOI, i.e., instances in uncovered regions are highlighted, such as the left part of the manifold after the first iteration (marked red). Bottom: divergence of Class Probabilities. The diversity seems to increase with the distance to labeled instances.*

we discussed a visual analysis appraoch of DOIs with respect to data and model characteristics, and to the temporal progression of the labeling process. We envision that this approach will enable others to gain a deeper understanding and better explainability of instance selection strategies and the design of novel strategies. Future work includes the formal description of classes of DOIs to ease their re-implementation. We also plan to investigate commonalities and differences between instance selection strategies and label propagation strategies from semi-supervised learning. Finally, we aim for the generation of novel and better labeling strategies.

References

[AQ07] Ayache S., Quénot G.: Evaluation of active learning strategies for video indexing. *Signal Processing: Image Communication 22*, 7-8 (2007), 692–704. 3

[BHL*18] Bernard J., Hutter M., Lehmann M., Müller M., Zeppelzauer M., Sedlmair M.: Learning from the Best - Visual Analysis of a Quasi-Optimal Data Labeling Strategy. In *Conference on Visualization (EuroVis)* (2018), Eurographics. doi:10.2312/eurovisshort.20181085. 1, 3

[BKNS00a] Breunig M. M., Kriegel H.-P., Ng R. T., Sander J.: Lof: Identifying density-based local outliers. In *Int. Conf. On Management of Data (SIGMOD)* (2000), ACM. 2

[BKNS00b] Breunig M. M., Kriegel H.-P., Ng R. T., Sander J.: Lof: identifying density-based local outliers. In *ACM sigmod record* (2000), vol. 29, ACM, pp. 93–104. 2

[BLBC12] Brown E. T., Liu J., Brodley C. E., Chang R.: Dis-function: Learning distance functions interactively. In *IEEE Visual Analytics Science and Technology (VAST)* (2012), IEEE, pp. 83–92. 1

[BSB*15] Bernard J., Sessler D., Bannach A., May T., Kohlhammer J.: A visual active learning system for the assessment of patient well-being in prostate cancer research. In *IEEE VIS WS on Visual Analytics in Healthcare (VAHC)* (2015), ACM, pp. 1–8. doi:10.1145/2836034.2836035. 2

[BTF18] Benato B. C., Telea A. C., Falcão A. X.: Semi-supervised learning with interactive label propagation guided by feature space projections. In *IEEE Conference on Graphics, Patterns and Images (SIBGRAPI)* (2018), IEEE, pp. 392–399. 1

[BZL*18] Bernard J., Zeppelzauer M., Lehmann M., Müller M., Sedlmair M.: Towards User-Centered Active Learning Algorithms. *Computer Graphics Forum (CGF)* (2018). doi:10.1111/cgf.13406. 1, 2

[BZSA18] Bernard J., Zeppelzauer M., Sedlmair M., Aigner W.: Vial: a unified process for visual interactive labeling. *The Visual Computer (TVCJ) 34*, 9 (2018), 1189–1207. doi:10.1007/s00371-018-1500-3. 1

[CBK09] Chandola V., Banerjee A., Kumar V.: Anomaly detection: A survey. *ACM Comput. Surv. 41*, 3 (2009), 15:1–15:58. doi:10.1145/1541880.1541882. 2

[Dun74] Dunn J. C.: Well-separated clusters and optimal fuzzy partitions. *Cybernetics and Systems 4*, 1 (1974), 95–104. 2

[EKS*96] Ester M., Kriegel H.-P., Sander J., Xu X., et al.: A density-based algorithm for discovering clusters in large spatial databases with noise. In *KDD* (1996), vol. 96, pp. 226–231. 2

[ERT19] Espadoto. M., Rodrigues. F. C. M., Telea. A. C.: Visual analytics of multidimensional projections for constructing classifier decision boundary maps. In *INSTICC IVAPP* (2019), SciTePress, pp. 28–38. doi:10.5220/0007260800280038. 3

[FT04] Fuglede B., Topsoe F.: Jensen-shannon divergence and hilbert space embedding. In *International Symposium on Information Thsome refeory, 2004. ISIT 2004. Proceedings.* (2004). doi:10.1109/ISIT.2004.1365067. 2

[HBV02] Halkidi M., Batistakis Y., Vazirgiannis M.: Clustering validity checking methods: Part ii. *SIGMOD Rec. 31*, 3 (2002), 19–27. doi:10.1145/601858.601862. 2

[HJL06] Hoi S. C., Jin R., Lyu M. R.: Large-scale text categorization by batch mode active learning. In *World Wide Web* (2006), ACM, pp. 633–642. doi:10.1145/1135777.1135870. 2, 3

[HKBE12] Heimerl F., Koch S., Bosch H., Ertl T.: Visual classifier training for text document retrieval. *IEEE Transactions on Visualization and Computer Graphics (TVCG) 18*, 12 (2012), 2839–2848. doi:10.1109/TVCG.2012.277. 1

[HNH*12] Höferlin B., Netzel R., Höferlin M., Weiskopf D., Heidemann G.: Inter-active learning of ad-hoc classifiers for video visual analytics. In *Visual Analytics Science and Technology (VAST)* (2012), IEEE, pp. 23–32. doi:10.1109/VAST.2012.6400492. 1

[Jai10] Jain A. K.: Data clustering: 50 years beyond k-means. *Pattern Recognition Letters 31*, 8 (2010), 651–666. 2

[KL*51] Kullback S., Leibler R., et al.: On information and sufficiency. *The Annals of Mathematical Stat. 22*, 1 (1951), 79–86. 2, 3

[KN98] Knorr E. M., Ng R. T.: Algorithms for mining distance-based outliers in large datasets. In *Conference on Very Large Data Bases (VLDB)* (1998), Morgan Kaufmann Publishers Inc., pp. 392–403. 2

[Kol33] Kolmogorov A.: Sulla determinazione empirica di una legge di distribuzione. *G. Ist. Ital. Attuari 4* (1933), 83–91. 2

[KShZ08] Kriegel H.-P., S hubert M., Zimek A.: Angle-based outlier detection in high-dimensional data. In *SIGKDD Conference on Knowledge Discovery and Data Mining* (2008), ACM, pp. 444–452. doi:10.1145/1401890.1401946. 2

[Lic13] Lichman M.: UCI machine learning repository, 2013. URL: http://archive.ics.uci.edu/ml. 4

[LWG19] Li Q., Wu X.-M., Guan Z.: Generalized label propagation methods for semi-supervised learning. *arXiv preprint arXiv:1901.09993* (2019). 1

[Mam98] Mamitsuka N. A. H.: Query learning strategies using boosting and bagging. In *International Conference on Machine Learning (ICML)* (1998), vol. 1, Morgan Kaufmann Pub. 2

[MHT18] M. Rodrigues F. C., Hirata R., Telea A. C.: Image-based visualization of classifier decision boundaries. In *IEEE Conference on Graphics, Patterns and Images (SIBGRAPI)* (2018), pp. 353–360. doi:10.1109/SIBGRAPI.2018.00052. 3

[QHR*09] Qi G.-J., Hua X.-S., Rui Y., Tang J., Zhang H.-J.: Two-dimensional multilabel active learning with an efficient online adaptation model for image classification. *IEEE Transactions on Pattern Analysis and Machine Intelligence (TPAMI) 31*, 10 (2009), 1880–1897. doi:10.1109/TPAMI.2008.218. 2, 3

[Rou87] Rousseeuw P. J.: Silhouettes: A graphical aid to the interpretation and validation of cluster analysis. *Journal of Computational and Applied Mathematics 20* (1987), 53 – 65. doi:https://doi.org/10.1016/0377-0427(87)90125-7. 2

[RRS00] Ramaswamy S., Rastogi R., Shim K.: Efficient algorithms for mining outliers from large data sets. *SIGMOD Rec. 29*, 2 (2000), 427–438. doi:10.1145/335191.335437. 2

[SC08] Settles B., Craven M.: An analysis of active learning strategies for sequence labeling tasks. In *Empirical Methods in Natural Language Processing (EMNLP)* (2008), Computational Linguistics, pp. 1070–1079. 2

[SCR08] Settles B., Craven M., Ray S.: Multiple-instance active learning. In *Advances in neural information processing systems* (2008), pp. 1289–1296. 2, 3

[Set12] Settles B.: Active learning. *Synthesis Lectures on Artif. Intell. and Machine Learning 6*, 1 (2012), 1–114. 1, 2, 3

[SG10] Seifert C., Granitzer M.: User-based active learning. In *IEEE Conference on Data Mining Workshops (ICDMW)* (2010), pp. 418–425. doi:10.1109/ICDMW.2010.181. 1

[Sha48] Shannon C. E.: A mathematical theory of communication. *The Bell System Technical Journal 27*, 3 (1948), 379–423. doi:10.1002/j.1538-7305.1948.tb01338.x. 2

[Sim49] Simpson E. H.: Measurement of diversity. *Nature 163*, 4148 (1949), 688. 2

[Smi48] Smirnov N.: Table for estimating the goodness of fit of empirical distributions. *The Annals of Mathematical Statistics 19*, 2 (1948), 279–281. 2

[SOS92] Seung H. S., Opper M., Sompolinsky H.: Query by committee. In *Worksh. on Comput. Learning Theory (COLT)* (1992), ACM, pp. 287–294. doi:10.1145/130385.130417. 2

[STMT12] Sedlmair M., Tatu A., Munzner T., Tory M.: A taxonomy of visual cluster separation factors. *Computer Graphics Forum (CGF) 31*, 3pt4 (2012), 1335–1344. doi:10.1111/j.1467-8659.2012.03125.x. 2

[TFS08] Tominski C., Fuchs G., Schumann H.: Task-driven color coding. In *2008 12th International Conference Information Visualisation* (July 2008), pp. 373–380. doi:10.1109/IV.2008.24. 3

[VPS*02] Vendrig J., Patras I., Snoek C., Worring M., den Hartog J., Raaijmakers S., van Rest J., van Leeuwen D. A.: Trec feature extraction by active learning. In *TREC* (2002). 2

[WH11] Wang M., Hua X.-S.: Active learning in multimedia annotation and retrieval: A survey. *ACM Trans. Intell. Syst. Technol. 2*, 2 (2011), 10:1–10:21. doi:10.1145/1899412.1899414. 3

[WKBD06] Wu Y., Kozintsev I., Bouguet J.-Y., Dulong C.: Sampling strategies for active learning in personal photo retrieval. In *IEEE International Conference on Multimedia and Expo* (2006), IEEE, pp. 529–532. doi:10.1109/ICME.2006.262442. 2