Personalized Visual-Interactive Music Classification

Christian Ritter[1], Christian Altenhofen[2], Matthias Zeppelzauer[3], Arjan Kuijper[1,2], Tobias Schreck[4], and Jürgen Bernard[1]

[1]TU Darmstadt, Germany
[2]Fraunhofer Institute for Computer Graphics Research, IGD, Darmstadt, Germany
[3]St. Pölten University of Applied Sciences, St. Pölten, Austria
[4]Graz University of Technology, Graz, Austria

Abstract
We present an interactive visual music classification tool that will allow users to automatically structure music collections in a personalized way. With our approach, users play an active role in an iterative process of building classification models, using different interactive interfaces for labeling songs. The interactive tool conflates interfaces for the detailed analysis at different granularities, i.e., audio features, music songs, as well as classification results at a glance. Interactive labeling is provided with three complementary interfaces, combining model-centered and human-centered labeling-support principles. A clean visual design of the individual interfaces depicts complex model characteristics for experts, and indicates our work-in-progress towards the abilities of non-experts. The result of a preliminary usage scenario shows that, with our system, hardly any knowledge about machine learning is needed to create classification models of high accuracy with less than 50 labels.

CCS Concepts
•*Human-centered computing* → *Visualization application domains;* •*Computing methodologies* → *Machine learning;*

1. Introduction

Today, numerous datasets prompt for extraction of potentially useful knowledge. This applies to research, industry, society, and to personal data collections as well. The scope of applied machine learning (ML) techniques is expanding from expert-driven scenarios to situations involving non-experts, e.g., in everyday life. Personal digital music collections are one example where non-expert users may benefit from applying those techniques. Beyond sorting functionality for long lists of music songs, the classification of music or the recommendation of new music are promising to structure and exploit personal music collections in a useful way. Formally, classification finds functions that map instances x to labels y, according to a pre-defined set of labeled training data. In ML, many computational models exist which can be used to facilitate classification tasks, e.g., decision trees, neural networks, support vector machines, and more (see also Section 2.1).

Classification requires labels which have to be provided by users in advance. A way to facilitate labeling during a personalized training process is active learning [Set12], where a user is queried for labels of instances the classifier is most uncertain about. However, solely using active learning as a means for candidate selection confronts users with a constant stream of questions. This may raise frustration and decrease concentration [CT10], rather than assigning users an active role in candidate selection and model building.

ML can be enhanced with visual-interactive means to allow broader user groups to participate in the iterative and interactive learning [ACKK14, SSZ*16, SSZ*17, ERT*17]. The principle to combine model-centered (using active learning) with human-centered labeling (using visual-interactive interfaces) (referred to as visual interactive learning (VIAL) [BZSA18]) conveys this rationale for labeling tasks.

Our goal is to make model learning a more personalized, interactive, and transparent endeavor, not only for experts but - in future - also for non-experts. We present a general interactive ML approach that enables (expert) users to create personalized classifiers. In addition, we shed light on our process to extend and re-design the approach towards the abilities of non-experts, using the example of music classification as a clearly defined, relevant, and end user-centered application domain.

Previous work has addressed visual analytics approaches for analysis and exploration of multimedia content. Zahálka et al. present a literature survey on multimedia analytic and propose a multimedia analytics process model including user interaction and visualization [ZW14]. With a visual analytics technique, van der Corput and Jarke J. van Wijk support the discovery of relations in image collection considering both data content and metadata [vdCvW17]. Related to that, Gao et al. present a human-centered approach for personalized classifier training and image retrieval [GYSF09]. The Map of Mozart is an example for the visual cluster analysis of music songs proposed for interactive comparison of music categories [MLR06]. Finally, we refer to surveys for music classification and recommendation that do not explicitly incorporate visual analytics techniques [HDR*08, FLTZ11, SDP12].

Assigning users an active role in the labeling and classification process comes with four main challenges. First, most ML models run in a black-box manner. Building such models is a complex task which requires a transparent way to access the modeling process [SSZ*17]. Second, the ML process is incremental, requiring an interactive role of the user to train, analyze, and improve the model. Thereby, users have to make informed decisions in the modeling process [FO03]. Third, to support individual classification tasks, an explicit and user-centric (instead of only learner-centric)
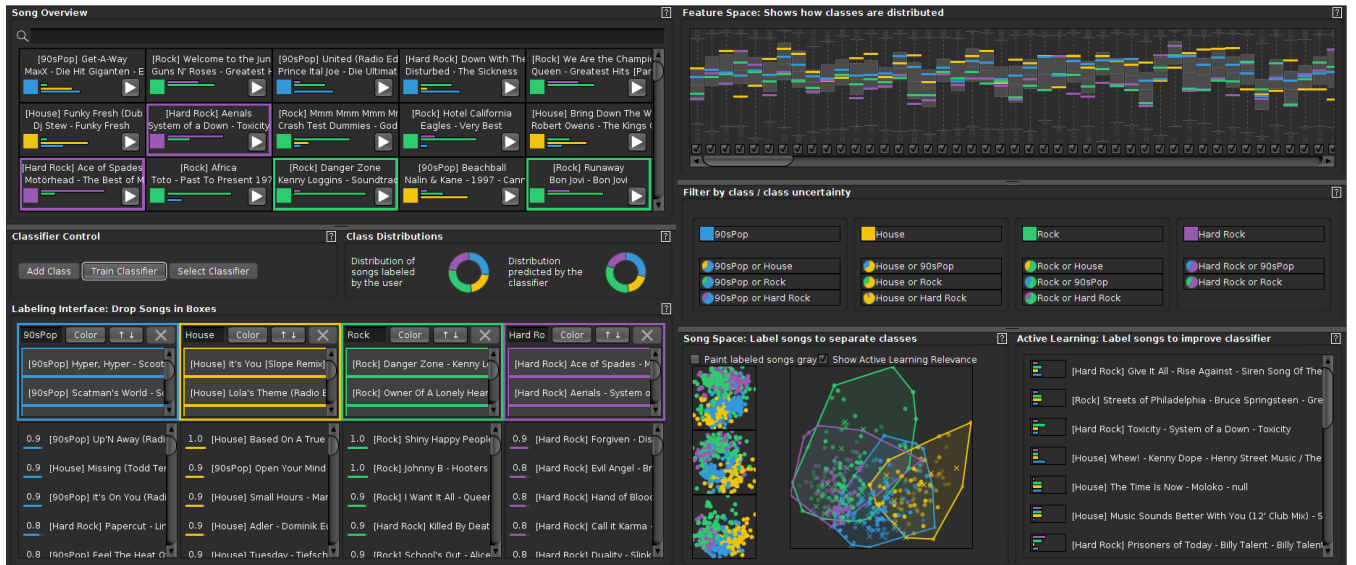
Figure 1: *Overview of our system for the personalized visual classification of music collections. Upper left: querying and browsing interface for all songs. Lower left: class creation, labeling, and class prediction interface. Upper right: analysis of features and class distributions. Center right: filtering interface for classes and class collisions. Lower right: two interfaces for the meaningful selection of labeling candidates (based on dimensionality reduction and active learning). Throughout this work, we apply a test dataset for genre classification. Note in our examples, we give the true class label in [brackets] in the song names for verification.*

way to give feedback is needed [SRL*07], coupled with the ability to have the control over the model [CT10]. Finally, classification approaches come with the challenge of a tedious labeling process to create training data [Set12].

We present a visual analytics approach that allows expert users to label music collections and to build personalized classification models. The approach builds upon requirements to interactive visual ML models and implements the principles of visual interactive learning (VIAL) concept [BZSA18]. Our primary contributions are:

- A novel music classification system which learns personal preferences of users
- Visualization techniques that allow the analysis of data and model characteristics at three different granularities (features, instances, and classes)
- Three visualization techniques that enable the user to label instances in different ways and to iteratively learn the classifier, depending on their information need, expertise, and preference

We also share our work-in-progress from an expert system towards a VIAL approach for end users following the goal to open ML to broader user groups. After presenting the system in Section 2, we demonstrate the approach in a usage scenario on genre classification (Section 3) and conclude in Section 4.

2. Approach

We present the technical background (Section 2.1) and requirements (Section 2.2) for our approach, followed by an in-depth description of the system and individual views in Section 2.3.

2.1. Technical Background

Music Feature Extraction: With the ESSENTIA [BWG*13] library, we use a standard approach for music feature extraction. ES-

SENTIA combines spectral, time-domain, tonal, rhythm, SFX and other high-level descriptors to yield a compact feature representation. ESSENTIA is open source, in contrast to projects that have become commercial such as the Echo Nest API [EWJL10].

Classification: We use general purpose classifiers which are fast, robust, and have become popular across several application domains. Users can select from the following algorithms: Random Forest [Bre01], C4.5 [Qui93], KStar [CT95], NaiveBayes [HKP12], SimpleLogistic [FHT00], MultilayerPerceptron [RHW86], SVM [CV95], and AdaBoost [Sch99], all provided with the WEKA [SF16] framework.

Active Learning: According to surveys on active learning [Set09, Set12], existing algorithms can be divided into uncertainty sampling, query-by-committee, error reduction, relevance-based, and data-driven variants. We use Smallest Margin [SDW01] as a prominent example for uncertainty sampling, but point out that our approach does not depend on the choice of one algorithm.

2.2. Requirements to Personalized Classification Approaches

We utilized a set of requirements to build and iteratively refine our tool. The requirements were the result of a review of related work, experiences of previous approaches [will be added], as well as interviews and formative evaluation rounds with musicians (and drummers) conducted in early design phases.

R_1 *Granularities*: information about instances, features, and classes.
R_2 *Transparency*: inference about data and model output.
R_3 *Guidance*: emphasis on important aspects of data and model.
R_4 *Efficiency*: reducing effort for labeling and model building.
R_5 *Uncertainty*: assessment of the probabilistic model output.
R_6 *Drill-Down*: focus on (semantic) aspects of interest.
R_7 *Instant Feedback*: ad-hoc representation of results for validation.
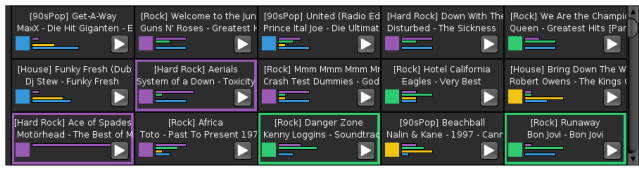R_8 *Simple UI*: the user interface should be both intuitive and usable.

Figure 2: *The Song Overview allows searching and browsing the personal music collection. Color-coding depicts information about class assignment as well as the uncertainty of the classifier.*

2.3. System Overview and General Design Principles

With the different views of the system (see Figure 1), we provide a means to (i) analyze music collections at three different granularities ($\mathbf{R}_1$): features, instances, and classes and to (ii) label individual songs with three different and complementing visual interfaces.

General visualization and interaction designs of the system are as follows. Color is always used to depict class information [STMT12]. As such, color also serves as a means to link class information across different views. *Drag-and-drop* is the main interaction technique for labeling instances in an intuitive way ($\mathbf{R}_8$). This allows picking instances in several views and dropping these instances into the class boxes in the labeling interface (see Figure 4). Filtering ($\mathbf{R}_6$) allows to drill-down to interesting songs, either from a particular class or with conflicts between two classes (see Figure 3). In addition to the general use of best practices, many visualization and interaction designs of individual views are the result of an iterative design process, including observations of non-experts and informal interviews ($\mathbf{R}_8$). One concrete outcome was to include help buttons at the upper right of every view to explain the usage of respective views in detail. In the description of individual views, we will refer to the *design process* to indicate simplifications of the design made for non-experts.

2.3.1. Song Overview

With the Song Overview (see Figure 2), users have a means to search for songs and browse through the music collection. The visual representation of songs (title, artists, and current classifier uncertainty) provides detailed information at the granularity of individual instances. Users can directly deduce the uncertainty of the classifier by analyzing horizontal bars depicting the probability distribution ($\mathbf{R}_5$). Using horizontal bars is the result of design process. This encoding is now recurring in three views to foster intuitiveness and familiarization. Similar to other encodings, uncertainty bars are automatically updated when the classifier is re-trained ($\mathbf{R}_7$). At start, when no classification model is trained, the Song View can be used to define initial labels and, thus, resolve bootstrap problems known from active learning [AP11]. Bold outlines indicate songs assigned to classes (two songs labeled green and purple in Figure 2).
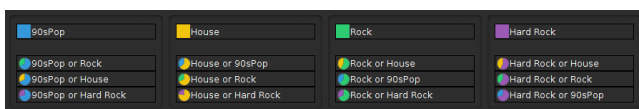

Figure 3: *Filtering interface to facilitate drill-down to important instances. Classes and class conflicts can be selected, e.g., to focus on semantically similar classes that are more difficult to learn.*
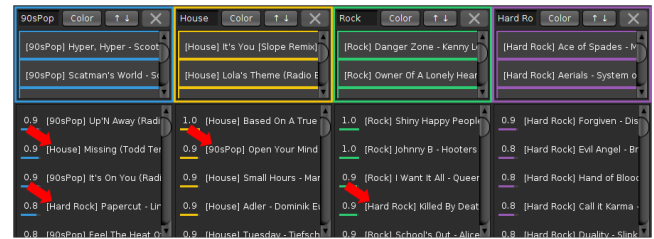
Figure 4: *The Labeling interface allows defining classes, dropping songs into class boxes (labeling), and validating class predictions with the list-based interface. In our usage scenario, the user re-assigned misclassified songs, resulting in an increase of accuracy.*

2.3.2. Labeling Interface

Core functionalities of the Labeling Interface (see Figure 4) are the definition of classes which, in turn, act as containers to drop songs for interactive labeling. In the lower part, users can validate the classifier's assignments of songs to classes in a list-based interface. By default, the interface lists songs by the degree of certainty of the classifier's class prediction (represented with horizontal bars and score labels for each song), helping users to analyze the classification uncertainty by the output of the classifier per-class ($\mathbf{R}_2$, $\mathbf{R}_5$). In addition, users can re-label misclassified songs with the drag-and-drop interaction to improve the classifier ($\mathbf{R}_4$). Inverting the list allows the assessment of most uncertain songs, e.g., to confirm or adapt class memberships.

2.3.3. Feature Space

The Feature Space interface (see Figure 6) is especially designed for expert users to discover relationships between features, data, and classes to facilitate more informed decisions ($\mathbf{R}_2$). Users can analyze every feature included in the dataset, sorted by a feature selection criterion, e.g., mutual information [Bat94]. Boxplots depict the distribution of songs in every feature. In addition, the interface allows the analysis of the discrimination strength of features for the different classes, depicted with colored lines for every class. Users can manually exclude features with checkboxes at the bottom. As a result of the design process, we added a detail view for selected features at the right (Figure 6). To further ease the use for non-experts, we added a mouse-over functionality to explain the analysis capability in detail.

2.3.4. Active Learning

The overarching goal of the Active Learning interface (see Figure 7) is to highlight candidate instances the model is most uncertain about. As such, users have an effective tool to improve the classification model with only few labels ($\mathbf{R}_4$). Our special requirement to the active learner is providing a sorted list of most uncertain instances, rather than only a single (worst) instance. With the list-based interface, users can select preferred songs from the suggestions for labeling, thus combining model-based and user-based preferences [BZSA18]. The list-based interface has a simple design ($\mathbf{R}_8$) and guides users in the labeling process ($\mathbf{R}_3$). To inform users about the degree of uncertainty, the interface again uses horizontal barcharts to show the probability distribution for any given song ($\mathbf{R}_5$), which may be beneficial especially for experts.

2.3.5. Song Space

The Song Space interface (see Figure 5) allows the informed selection of labels based on spatial characteristics of the
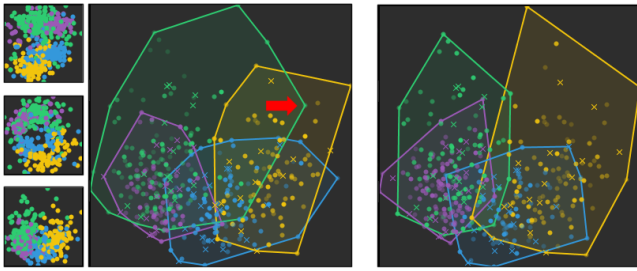
Figure 5: *Projection of songs to 2D plots using PCA. Color-coding helps to asses the separation of classes. Labeling between the two displayed states (large images) included a falsely assigned green song, resulting in more compact classes after re-training.*

dataset [BHZ*17] ($\mathbf{R}_2$). Three dimensionality reduction techniques (tSNE [vdMH08], PCA [Jol02], MDS [Kru64]) are used to map songs into the visual space, revealing the spatial distribution of the songs. Typical patterns to be found are dense regions, local clusters, or outliers. In addition, coloring of classes allows the assessment of class distributions ($\mathbf{R}_7$). The application of convex hulls [SP07] further supports the analysis of intersecting classes. Given that, users can identify conflicting regions in the song space [BHZ*17]. Hovering songs provides tool-tips and fills a text box at the bottom. Drag-and-drop is again used to support selecting and labeling songs ($\mathbf{R}_4$). In the design process, we observed that non-experts need additional guidance towards most relevant instances. Accordingly, we use an active learning model to facilitate algorithmic guidance. Transparency is now used to deemphasize instances that are less relevant for model improvement.

3. Usage Scenario

In this section, we showcase the usage of our approach by a scenario. As we aim to extend the system towards new user groups, we recruited a musician with a high interest in music history, experience in visualization and HCI, but also basic knowledge in ML. The user's task is to create a genre classifier with four classes (90s Pop, House, Rock, and Hard Rock). For every class 100 candidate songs are provided. Our focus is on the demonstration of the labeling process, not on the (semantically) correct assignment of labels [BHZ*17]. To exclude this human factor, we add a tag with ground truth information to each song (e.g., [Rock]). In parallel, we employ the ground truth information to assess the gain of the classifier's accuracy during the labeling process. During the interview, notes and screenshots were taken for documentation.

At start, the musician chooses two representative candidates for each genre from the Song Overview (Figure 2) and builds an initial classifier. The Song Space in Figure 5 now shows the first classification result which still has large spatial overlaps between the four classes, though. He drags some songs from overlapping regions into
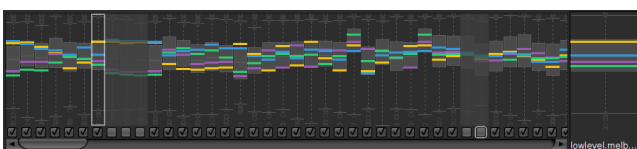

Figure 6: *The Feature Space interface shows the distribution of songs with boxplots and the class discrimination with colored lines for every feature. Interesting features can be analyzed in detail (right). Checkboxes allow the deselection of features.*
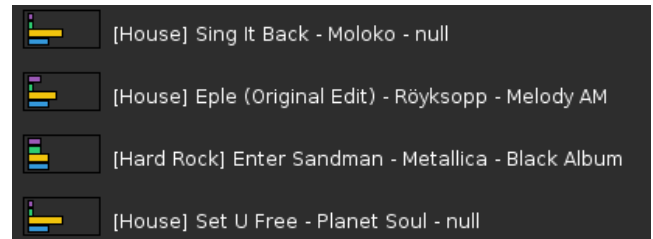

Figure 7: *An active learning model provides a list of candidate instances the classifier is most uncertain about. Our interactive filter reduced songs towards conflicts between class blue and orange.*

the labeling interface to improve the classifiers' performance at this location of the song space. In addition, the musician selects songs at the outbound of the class distributions (such as the green song marked with a red arrow). After labeling 16 instances, the accuracy of the classifier is now 49%. Next, the musician decides to use the Active Learning interface (Figure 7) to further improve the classifier. With another eight labeled songs, the accuracy increases to 53%. The musician recognizes that there is still confusion between 90s Pop and House as well as Rock and Hard Rock. This analysis task is supported with the Song Space interface showing the overlaps of classes (Figure 1). Therefore, he applies the Filter interface (Figure 3) to drill-down the music collection. The Active Learning interface now shows only songs with conflicting predictions. After another eight labels the accuracy is now 63%. In the final step, the musician wants to validate the performance of the classifier. For that purpose, he inspects the lists of classified songs in the Labeling interface. Figure 4 shows how he identifies misclassified songs among the top-ranked songs in several classes (red arrows). The re-assignment of these instances leads to a gain of accuracy to 72%. With only 49 labeled songs, the musician has created a personal classifier (here: for genres) that is now able to automatically predict nearly three quarters of his music collection correctly.

4. Conclusion and Future Work

We presented a visual analytics tool for the personalized classification of music collections. The system integrates several visual interactive labeling (VIAL) techniques and mechanisms to support interactive machine learning. In particular, different linked views allow users to label and classify music songs, as well as to validate the classification model. With the system, experts can effectively create personal classifiers. To open interactive machine learning for non-experts, we also disclosed our work-in-progress towards simplistic and usable visualization and interaction designs.

Future work includes further evaluation with non-expert users and the refinement of the interactive interfaces. It will be particularly interesting to quantify the effect of including users in the classification loop in comparison with automatic approaches, and how the visual and interaction design influences this comparison. Our approach is a basis for further experiments in this direction. Finally, we plan to extend the techniques to other recommendation tasks to investigate their general applicability.

Acknowledgements

References

[ACKK14] AMERSHI S., CAKMAK M., KNOX W. B., KULESZA T.: Power to the people: The role of humans in interactive machine learning. *AI Magazine 35*, 4 (2014), 105–120. 1

[AP11] ATTENBERG J., PROVOST F.: Inactive Learning?: Difficulties Employing Active Learning in Practice. *SIGKDD Explor. Newsl. 12*, 2 (Mar. 2011), 36–41. doi:10.1145/1964897.1964906. 3

[Bat94] BATTITI R.: Using mutual information for selecting features in supervised neural net learning. *IEEE Transactions on Neural Networks 5*, 4 (Jul 1994), 537–550. doi:10.1109/72.298224. 3

[BHZ*17] BERNARD J., HUTTER M., ZEPPELZAUER M., FELLNER D., SEDLMAIR M.: Comparing visual-interactive labeling with active learning: An experimental study. *TVCG 24*, 1 (2017). accepted for TVCG. doi:10.1109/TVCG.2017.2744818. 4

[Bre01] BREIMAN L.: Random forests. *Machine Learning 45*, 1 (2001), 5–32. doi:10.1023/A:1010933404324. 2

[BWG*13] BOGDANOV D., WACK N., GÓMEZ E., GULATI S., HERRERA P., MAYOR O., ROMA G., SALAMON J., ZAPATA J. R., SERRA X.: Essentia: an audio analysis library for music information retrieval. In *International Society for Music Information Retrieval Conference (ISMIR)* (2013), pp. 493–498. URL: http://hdl.handle.net/10230/32252. 2

[BZSA18] BERNARD J., ZEPPELZAUER M., SEDLMAIR M., AIGNER W.: Vial: a unified process for visual interactive labeling. *The Visual Computer* (Mar 2018). doi:10.1007/s00371-018-1500-3. 1, 2, 3

[CT95] CLEARY J. G., TRIGG L. E.: K*: An instance-based learner using and entropic distance measure. In *International Conference on International Conference on Machine Learning (ICML)* (San Francisco, CA, USA, 1995), Morgan Kaufmann Publishers Inc., pp. 108–114. URL: http://dl.acm.org/citation.cfm?id=3091622.3091636. 2

[CT10] CAKMAK M., THOMAZ A. L.: Optimality of human teachers for robot learners. In *International Conference on Development and Learning (ICDL)* (2010), IEEE. doi:10.1109/DEVLRN.2010.5578865. 1, 3

[CV95] CORTES C., VAPNIK V.: Support-vector networks. *Machine Learning 20*, 3 (1995), 273–297. URL: https://doi.org/10.1007/BF00994018, doi:10.1007/BF00994018. 2

[ERT*17] ENDERT A., RIBARSKY W., TURKAY C., WONG B. W., NABNEY I., BLANCO I. D., ROSSI F.: The state of the art in integrating machine learning into visual analytics. *Computer Graphics Forum* (2017). doi:10.1111/cgf.13092. 1

[EWJL10] ELLIS D. P., WHITMAN B., JEHAN T., LAMERE P.: The echo nest musical fingerprint. In *International Society for Music Information Retrieval Conference* (2010), vol. 32. 2

[FHT00] FRIEDMAN J., HASTIE T., TIBSHIRANI R.: Additive logistic regression: a statistical view of boosting. *The Annals of Statistics 28*, 2 (04 2000), 337–407. URL: https://doi.org/10.1214/aos/1016218223, doi:10.1214/aos/1016218223. 2

[FLTZ11] FU Z., LU G., TING K. M., ZHANG D.: A survey of audio-based music classification and annotation. *Trans. Multi. 13*, 2 (Apr. 2011), 303–319. doi:10.1109/TMM.2010.2098858. 1

[FO03] FAILS J. A., OLSEN JR. D. R.: Interactive machine learning. In *Conference on Intelligent User Interfaces (IUI)* (New York, NY, USA, 2003), ACM, pp. 39–45. doi:10.1145/604045.604056. 2

[GYSF09] GAO Y., YANG C., SHEN Y., FAN J.: Incorporate visual analytics to design a human-centered computing framework for personalized classifier training and image retrieval. In *Advances in Information and Intelligent Systems*. 2009, pp. 165–187. doi:10.1007/978-3-642-04141-9_8. 1

[HDR*08] HUANG T. S., DAGLI C. K., RAJARAM S., CHANG E. Y., MANDEL M. I., POLINER G. E., ELLIS D. P. W.: Active learning for interactive multimedia retrieval. *Proceedings of the IEEE 96*, 4 (April 2008), 648–667. doi:10.1109/JPROC.2008.916364. 1

[HKP12] HAN J., KAMBER M., PEI J.: *Data Mining – Concepts and Techniques*. Morgan Kaufmann Publishers, 2012. 3rd edition. 2

[Jol02] JOLLIFFE I. T.: *Principal Component Analysis*, 3rd ed. Springer, 2002. 4

[Kru64] KRUSKAL J. B.: Multidimensional scaling by optimizing goodness of fit to a nonmetric hypothesis. *Psychometrika 29*, 1 (1964), 1–27. 4

[MLR06] MAYER R., LIDY T., RAUBER A.: The map of mozart. In *International Conference on Music Information Retrieval (ISMIR)* (2006), pp. 351–352. 1

[Qui93] QUINLAN J. R.: *C4.5: Programs for Machine Learning*. Morgan Kaufmann Publishers Inc., San Francisco, CA, USA, 1993. 2

[RHW86] RUMELHART D. E., HINTON G. E., WILLIAMS R. J.: Parallel distributed processing: Explorations in the microstructure of cognition, vol. 1. MIT Press, Cambridge, MA, USA, 1986, ch. Learning Internal Representations by Error Propagation, pp. 318–362. URL: http://dl.acm.org/citation.cfm?id=104279.104293. 2

[Sch99] SCHAPIRE R. E.: A brief introduction to boosting. In *International Joint Conference on Artificial Intelligence (IJCAI)* (San Francisco, CA, USA, 1999), Morgan Kaufmann Publishers Inc., pp. 1401–1406. URL: http://dl.acm.org/citation.cfm?id=1624312.1624417. 2

[SDP12] SONG Y., DIXON S., PEARCE M.: A survey of music recommendation systems and future perspectives. In *International Symposium on Computer Music Modeling and Retrieval* (2012), vol. 4. 1

[SDW01] SCHEFFER T., DECOMAIN C., WROBEL S.: Active hidden markov models for information extraction. In *Conference on Advances in Intelligent Data Analysis (IDA)* (London, UK, UK, 2001), Springer-Verlag, pp. 309–318. 2

[Set09] SETTLES B.: *Active Learning Literature Survey*. Tech. Report 1648, Univ. of Wisconsin–Madison, 2009. 2

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

[SF16] SMITH T. C., FRANK E.: *Statistical Genomics: Methods and Protocols*. Springer, New York, NY, 2016, ch. Introducing Machine Learning Concepts with WEKA, pp. 353–378. 2

[SP07] SCHRECK T., PANSE C.: A new metaphor for projection-based visual analysis and data exploration. In *Visualization and Data Analysis 2007* (2007), no. 6495 in SPIE Proceedings, SPIE, pp. 64950L–64950L–12. doi:10.1117/12.697879. 4

[SRL*07] STUMPF S., RAJARAM V., LI L., BURNETT M., DIETTERICH T., SULLIVAN E., DRUMMOND R., HERLOCKER J.: Toward harnessing user feedback for machine learning. In *Conference on Intelligent User Interfaces (IUI)* (New York, NY, USA, 2007), ACM, pp. 82–91. doi:10.1145/1216295.1216316. 2

[SSZ*16] SACHA D., SEDLMAIR M., ZHANG L., LEE J. A., WEISKOPF D., NORTH S. C., KEIM D. A.: Human-Centered Machine Learning Through Interactive Visualization: Review and Open Challenges. In *Artificial Neural Networks, Computational Intelligence and Machine Learning* (2016). 1

[SSZ*17] SACHA D., SEDLMAIR M., ZHANG L., LEE J. A., PELTONEN J., WEISKOPF D., NORTH S. C., KEIM D. A.: What you see is what you can change: Human-centered machine learning by interactive visualization . *Neurocomputing* (2017). ISSN = 0925-2312. doi:10.1016/j.neucom.2017.01.105. 1

[STMT12] SEDLMAIR M., TATU A., MUNZNER T., TORY M.: A taxonomy of visual cluster separation factors. *Computer Graphics Forum 31*, 3pt4 (2012), 1335–1344. doi:10.1111/j.1467-8659.2012.03125.x. 3

[vdCvW17] VAN DER CORPUT P., VAN WIJK J. J.: Comparing personal image collections with picturevis. *Comput. Graph. Forum 36*, 3 (2017), 295–304. doi:10.1111/cgf.13188. 1

[vdMH08] VAN DER MAATEN L., HINTON G. E.: Visualizing high-dimensional data using t-sne. *Journal of Machine Learning Research 9* (2008), 2579–2605. 4

[ZW14] ZAHÁLKA J., WORRING M.: Towards interactive, intelligent, and integrated multimedia analytics. In *2014 IEEE Conference on Visual Analytics Science and Technology, VAST 2014, Paris, France, October 25-31, 2014* (2014), pp. 3–12. doi:10.1109/VAST.2014.7042476. 1