

Visual Analytic Process to Familiarize the Average Person with Ways to Apply Machine Learning

Andrew Tran; California State University, Sacramento; Sacramento, CA;
Yamini Dasu; California State University, Sacramento; Sacramento, CA;
Anna Baynes; California State University, Sacramento; Sacramento, CA;
Email: {andrewtran,yaminidasu,abaynes.ecs}.csus.edu

Abstract

The everyday consumer is inundated with applications powered by machine learning. But in an ordinary day, do we encounter situations and choices which could also benefit from machine learning for which there is no specific tool invented yet? We describe scenarios where people without any machine learning background could find it useful to define their own solution which uses machine learning. Although machine learning is becoming ubiquitous, the average person is unaware of the steps involved. This abstraction makes sense, in many situations, such as traffic predictions, it is not necessary for the driver to know what machine learning algorithm is running. However, we consider examples where knowing how to incorporate machine learning into a problem would assist in decision making. We propose a workflow with operations leading to a final application. There are several challenges here, namely, the average consumer is not expected to have a mathematical background, nor is expected to acquire any additional background. To achieve this new utility, we use a visual analytic pipeline which integrates machine learning and the person. We use the IEEE VAST 2018 Challenge as a case study in which the user steps through the workflow. Finally, we envision the resulting application.

Introduction

Machine Learning is everywhere from virtual personal assistants, commuting predictions, online customer support, online product recommendations, and more. Yet, the everyday consumer, not the novice machine learning student, is unaware of the decision trees, clustering algorithms, SVMs, and other machine learning techniques which support these processes. For the most part, this abstraction completely makes sense. A family driving to their daughter's piano recital does not care to know how neural networks were used to predict the traffic they see on their smart phones. But problems do arise where had the user known machine learning was an option, the user would indeed care to reasonably understand the background process. We describe two general motivating examples before giving a specific case. First, a decision-making problem can occur which is a perfect candidate for machine learning techniques. For example, the problem can be formatted as a classic supervised learning problem with a clear training and test datasets. But the people involved in the decision do not have a machine learning background. We can also assume it may not be feasible to hire a data scientist for a consultation. The second general motivating example occurs when you have already used machine learning to solve a problem. But you need to convince an audience of non-experts how you arrived at your

decision. This situation requires leading the audience through the process before they accept your results. Both of these motivating examples are materialized in the IEEE VAST 2018 Challenge [10]. We summarize the challenge's fictional scenario and explain how it presents a specific case for both motivating examples.

The background story for this challenge surrounds the Rose-Crested Blue Pipit, which is the favorite bird of a town called Mistford. The birds live in a nearby nature preserve where their environment has been compromised by a manufacturing company called Kasios. To prove the bird's population is not endangered, Kasios provides a set of Pipit bird calls. Kasios dataset contains mp3 files of birdcall recordings along with the locations of their recordings. Kasios claims that this dataset shows that the Rose-Crested Blue Pipit is thriving across the entire preserve. With the dataset alone, Kasios has convinced the town people that their favorite bird is not endangered. But this dataset is not annotated with any information, i.e. an expert has not listened to the bird calls to describe the bird species. Therefore, in this critical situation, the town may be misled into believing possible false information. If there exist constraints on this problem, such as, no available or unbiased ornithologists to listen to Kasios recording, how will the town verify Kasios' claims? One more data set is available: Over several years, ornithologist and other bird call enthusiasts have recorded and annotated bird calls throughout the park. This dataset is richly annotated with the bird call's: English name, Vocalization type (i.e. a call, a song, or some other sound), Quality score (A, B, C, D, or E), time, and date of where the sound was recorded.

This situation is a prime example where the town and nature preserve officials can employ the power of machine learning to determine if Kasios' claims are true. A supervised machine learning algorithm can be used with Kasios' bird calls as the test data set and the previously richly annotated bird call recordings as the training data set. The town and nature officials can apply this process and determine the validity of Kasios' claims. But this action alone would ignore the town's belief in Kasios. The officials may need to change the town's sentiment. The sentiment of Mistford town describes the second motivating example. An audience of the average consumer will need to understand how machine learning produced its conclusions.

We propose a visual analytic workflow which can both permit the officials a seamless entry to utilize machine learning in their decision making and secondly, provide the townspeople through visualization validation ways to interpret the machine learning results. Beyond this fictional scenario, the benefit of having the everyday user involved in the machine learning process is

it can democratize machine learning so that anyone can define a problem which benefits from machine learning techniques.

However, the problem is that current machine learning tools and techniques are designed for either machine learning experts or at the very least, a person with a novice data science background. This assumption does not hold for the everyday consumer.

Given our listed benefits of having the everyday user participate in the machine learning process, the problem statement is: how can we make machine learning accessible and understandable to the everyday user? In this paper, we tackle this problem by describing interactive visualization techniques to turn machine learning into an educational visual analytic experience. The user is connected throughout the process, which benefits both the machine learning technique and the end results. In this work, we assume an SVM (Support Vector Machine) machine learning classifier will be used. The users are not changing the machine learning technique nor is the purpose to teach the user about the classifier mechanisms. Instead we want the user to understand the workflow and understand and inspect the solutions (i.e. classifications) made. We propose the following interactive visualization techniques: multi-view visualization validation views and annotating with context-based examples. We present a case study showcasing our techniques based on the IEEE VAST 2018 Challenge on the Kasios bird calls. The rest of the paper is as follows: we describe related works in information visualization, machine learning, and visual analytic systems. Next, we state the machine learning model we utilize and our assumptions to focus the elements on this problem. Then we present our proposed user involved workflow with the interactive visualization techniques. We present our case study utilizing these techniques. Finally, we conclude with analysis and directions for future work.

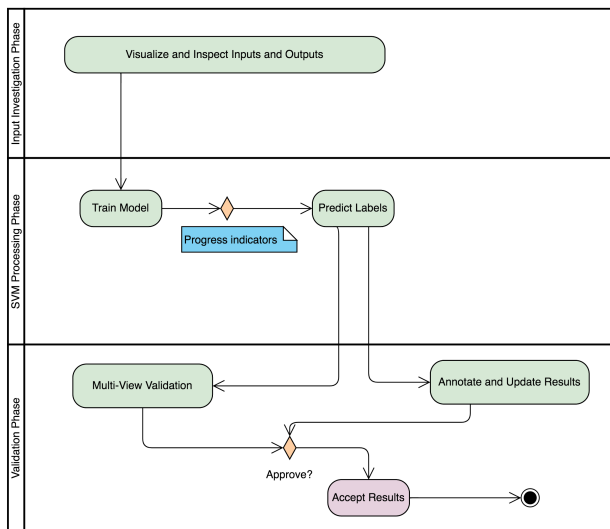


Figure 1. Flowchart of including the everyday user into the machine learning analytics process.

Related Works

The work presented lies in the intersection of information visualization, human computer interaction, and machine learning. Therefore, in this section we present relevant works in these areas.

Machine Learning and Human Computer Interaction

There have been several key movements in the area to combine human computer interaction benefits with machine learning [2], [20], [16]. The workshop [15] examined human computer interaction impacts on machine learning. Recognizing that machine learning is not an autonomous process on passively collected data, but users must tune algorithms, gather data, and decide what to model. Also, machine learning has the potential to democratize human computer interaction work by having people program interactions based on learning from behaviors. For example, [9] argues that artificial intelligence’s goals of powerful and intelligent interfaces can benefit from the principles used in human computer interaction design. Furthermore, human computer interaction can benefit from learning about users to meet the needs of users through artificial intelligence methods. [4] discusses concepts in usability and user experience for successful user interfaces and the impact on information retrieval interfaces. There are many tools intended to allow novice data scientists to apply machine learning techniques on their datasets. [12] presents the machine learning and analytic tool, Orange, which allows for both experienced and novice data mining users to perform analytic tasks and design visual data mining applications using Orange’s component-based framework.

Information Visualization and Machine Learning

Combining machine learning and information visualization creates powerful tools for countless applications [23],[19], [3], [1]. The seminar [5] calls for bridging information visualization with machine learning by addressing several questions, including: 1) The role of the user in a mixed visual analytic and machine learning system. 2) What interactive visualization tasks could benefit from machine learning (and vice versa)? 3) What are machine learning methods which can help scale up to information visualization demands? For example, [24] applies visual

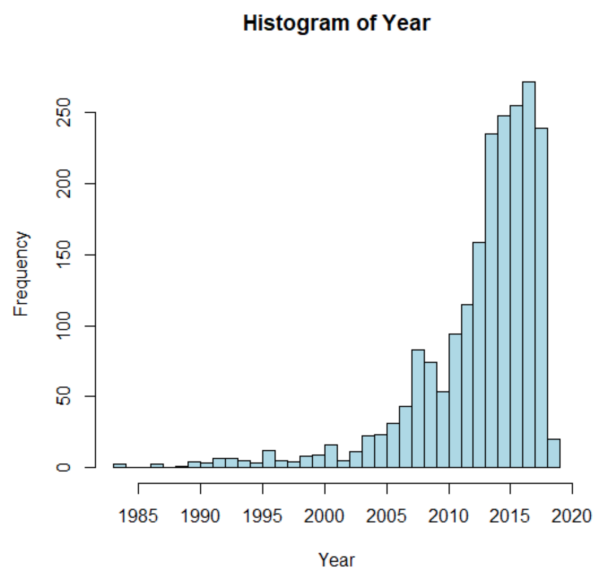


Figure 2. The frequency of recordings for each year is shown.

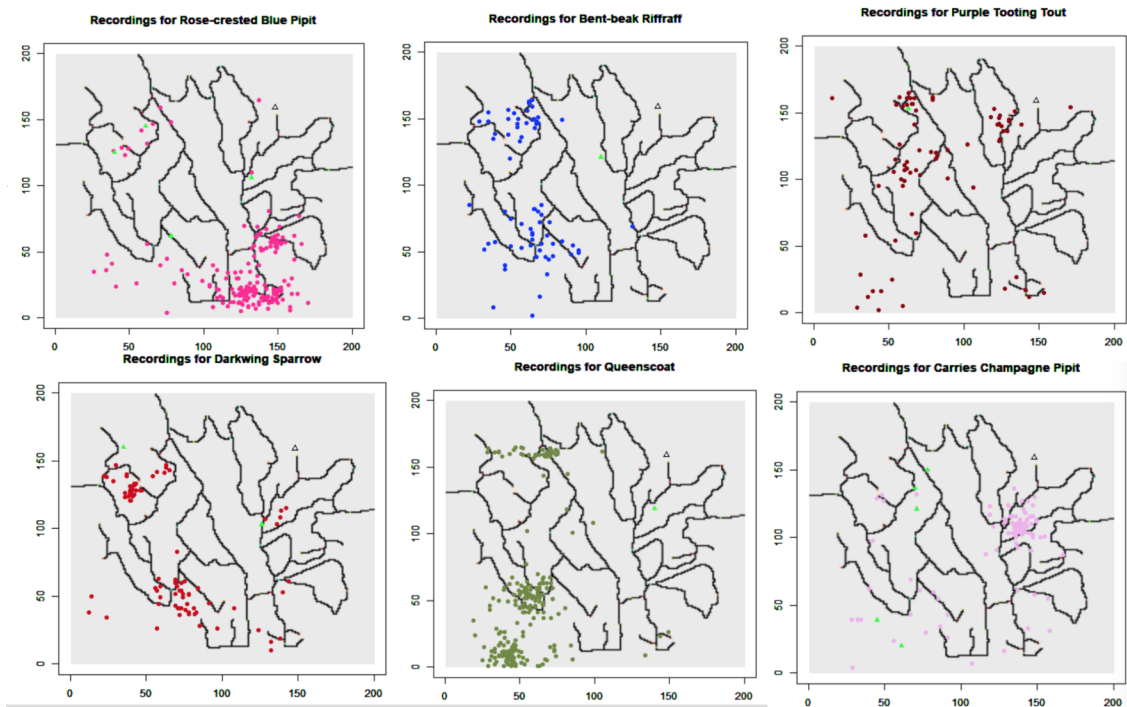


Figure 3. Multiple views of the predictions on the maps of the preserve are shown along with the expert annotated recordings of the same predicted specie. These visualizations help the users explore how confident they are in the predictions made. If they see a prediction of a specie but no other previously recording specie of the same type is in the same location, then the user will question the accuracy of the prediction.

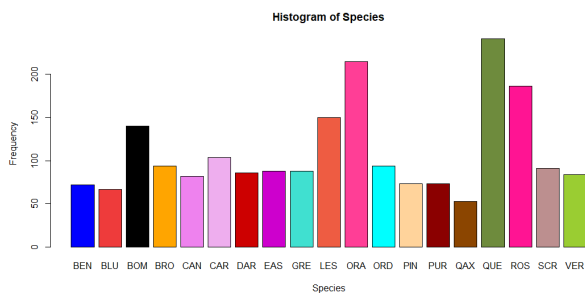


Figure 4. The different species populations are depicted.

analytic and machine learning techniques to solve urban problems by creating an interactive visualization tool for urban planners to analyze public transport commuters. The tool trains a classifier using the planners' classifications and in next iterations of the tool use, provides data on public transport commuters. There also exists the train of thought to understand how to include the user's explicit knowledge in the visual analytic process. In [6], the authors propose a conceptual model which incorporates user's tacit and explicit knowledge in the analytical reasoning during the visual analytic process. It is based on the visualization model of van Wijk [22].

Bird Call Analysis and Visualizations

As part of the IEEE VAST 2018 Challenge, bird call analysis must be researched. Bird call analysis is a well-studied field. Including their biology and physiology, for example, [14] studies

the biology and physiology of how birds sing through several experiments, including placing singing birds in helium air, to identify a new model in songbird phonation. In [13] propose methods to efficiently extract single sound events and connected sequences of bird calls in noisy environments. They use randomized decision trees to learn and classify sounds. In [21] describe methods to visualize long-duration recordings of the environment using spectrogram zooming, extracting acoustic features, and applying statistics reveal the acoustic features important to an ecologist. Spectrograms are also used to classify sounds. In [7] use feature extraction to create spectrograms and finally k-nearest neighbor classifier to classify frog calls.

Interactive Visualization Techniques for Machine Learning Assistance

In this work, we assume an SVM machine learning classifier will be used. The users are not changing the machine learning technique nor is the purpose to teach the user about the classifier mechanisms. Instead we want the user to understand the workflow and understand and inspect the solutions (i.e. classifications) made. We propose the following interactive visualization techniques: multi-view visualization validation views and annotating with context-based examples.

Figure 1 introduces the abstracted visual analytic flow chart for the user to interact with the classifier. The pipeline will start with a clear indication of how to tie in the user's inputs and outputs to the process. The user will have a set of inputs, for example, items with missing values which represent the classifications needed. The user will define what the output should be. There will also be an ability to visualize and examine their input dataset.

The better the user knows their own inputs, then they will be better at assessing the final results. They can see clearly what is the missing label which needs to be determined by the classifier. Therefore, the user understands the purpose behind employing the machine learning method. We are abstracting the details of the model, the refining process, and the math behind it. As the classifier begins refining its model based on the training data set, the system communicates to the user the progress for each label. For example, if there are three types of labels A, B, C. We show the user progress indicators when label A's model is updating and complete. This indicator allows the user to see that there is processing in operation. They do not need to know the exact details, nor do they probably want to know the computational level understanding. Instead they want to know how they are using the machine learning tool and that it is providing them a solution to their end goal.

Finally, when the classifier begins to predict the labels for the test data set, the user is notified of the start. Next, the user is given the results. This is when the interactive visualization tools come into play. First, the user is shown all the data. They must go to the next phase in the flow chart to validate the results. In the validation step, they are given several methods to validate the method. If other meta data is available, they are shown to the user along with the prediction made. The user then has a chance to flag or adjust the predicted values. The user is shown multiple views to validate results. Once they are satisfied with the validation, they can annotate the data. Through this process, the user has basically been involved in a machine learning technique on their personal data set.

Case Study: Kasios Bird Call Dataset

In this section, we walk through the analysis for the Kasios Bird Call Dataset. We first review the problem statement and present some basic metadata visual analysis to understand the background of the dataset. Then we look at the classification problem and apply the workflow and visualization techniques de-

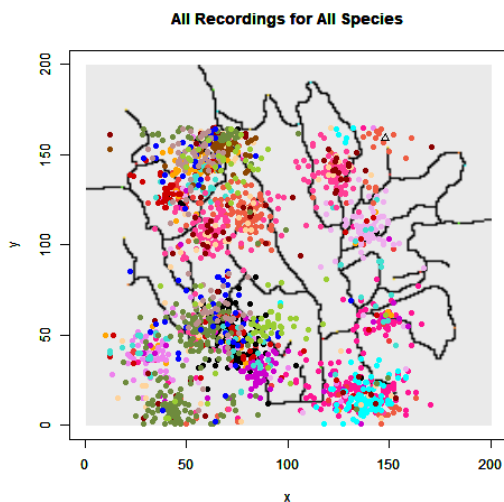


Figure 5. The bird call recordings are drawn on top of the map of the preserve.

scribed in the subsection above.

Kasios Bird Call Dataset Background

To understand the background of the bird species population in wildlife preserve, we perform some analysis on the bird call data which was collected over several years and annotated by experts. We also utilize the map of the preserve in our analysis. After performing preliminary analysis, we will include the Kasios bird call dataset, which is under question, into the analysis. The challenge makes an assumption that the annotated bird call dataset reasonably represents the bird patterns distribution across the park. We abbreviate each of the bird species' names with the first three letters as follows: Bent-beak Riffraff as BEN, Blue-collared Zipper as BLU, Bombadil as BOM, Broad-winged Jojo as BRO, Canadian Cootamum as CAN, Carries Champagne Pipit as CAR, Darkwing Sparrow as DAR, Eastern Corn Skeet as EAS, Green-tipped Scarlet Pipit as GRE, Lesser Birchbeere as LES, Orange Pine Plover as ORA, Ordinary Snape as ORD, Pinkfinch as PIN, Purple Tooting Tout as PUR, Queenscoat as QUE, Rose-crested Blue Pipit as ROS, Scrawny Jay as SCR, and Vermillion Trillian as VER.

A few patterns emerge in the distribution and number of recordings: the most obvious pattern that can be seen in the data is the increase in the number of recordings over time. There is an exponential pattern of growth when observing the frequency of the year the recordings were made, shown in Figure 2. This increase in recordings may reflect an increase in the population of the birds over time until 2016, which shows a sudden drop. This change suggests that the number of recordings for 2018 will also be dropping. We also consider the location and per specie population trends of the birds shown in Figure 4. A majority of the bird recordings were taken in the western half of the map. The eastern half of the map, where the alleged chemical dump site of Kasios is located, has much fewer bird recordings. We also present the visualizations shown in Figure 5 which shows the location of all the bird calls in the preserve where the colors represent the different species. A temporal analysis of the metadata also illuminates key characteristics of the bird activity in the region. Figure 6 shows the number of recordings seem to peak in May and again in October. This result is not surprising because the birds would appear more in Spring and Summer.

Bird Call Classification

Next, we use the methods introduced in the section above to analyze the Kasios' bird call dataset and determine if Kasios' claims that the Pipit is thriving throughout the park is correct. First, we developed the machine learning mechanism whose details stay abstracted and hidden from the user. We briefly discuss the design of the classifier for the bird calls, as follows. To study the bird calls supplied by Kasios, it is necessary to first understand some background on bird calls. The Mel-frequency cepstral coefficients (MFCCs) are used to characterize sounds [18]. MFCCs are the coefficients which represent a short-term spectrum of a sound, based on a linear cosine transform of a log power spectrum on a nonlinear mel scale of frequency. MFCCs are typically used to compare audio files. To identify the birds in Kasios' file, first we convert each of the known bird calls and Kasios' bird calls into MFCCs. Then we build an SVM classifier based on the known bird calls. Then we use the built SVM to predict the bird

name for the unknown bird calls in the Kasios file. In order to build MFCCs, we use python libraries [11]. We convert each mp3 file into a wav file using sox. This component described will be hidden from the user.

We present the validation and multi-view interactive visualization utilities for this example. This stage depends on the context of the problem. Therefore, we use some of the knowledge about the locations of the species, based on the expert annotated dataset. We create a multi-view scatterplot to show the locations of the birds in relation to the predicted label for the Kasios' bird call recording. This visualization, shown in Figure 3 allows the user to validate if they agree with the predicted results. Functionality such as filtering, zooming, and enhancing with annotations will allow users to investigate the results. The SVM classifier prediction is shown in bright green and the other color shows the prediction. For example, the blue shows the known recordings of the Bent-beak Riffraff. Based on the locations of the known Bent-beak Riffraff, we are not very confident in the result of the classifier. There are not any Bent-beak Riffraff recordings in the region that the classifier labeled this bird call. Based on visualizations of the bird mappings, we guess this might be a Carries Champagne Pipit. When we browse through the all the predicted versus known bird sounds above, we do not see any bright green values (which would be a bird sound from the Kasios set) in the region we see the Rose-crested Blue. This realization leads us to doubt Kasios' claim that the Rose-crested Blue Pipits are thriving throughout the park. However, our classifier did find a few recordings for the Rose-crested Blue Pipits. Only three of the bird calls were identified as Rose-crested Blue Pipit. This result does not support Kasios' claim that the Pipits are seen throughout the preserve.

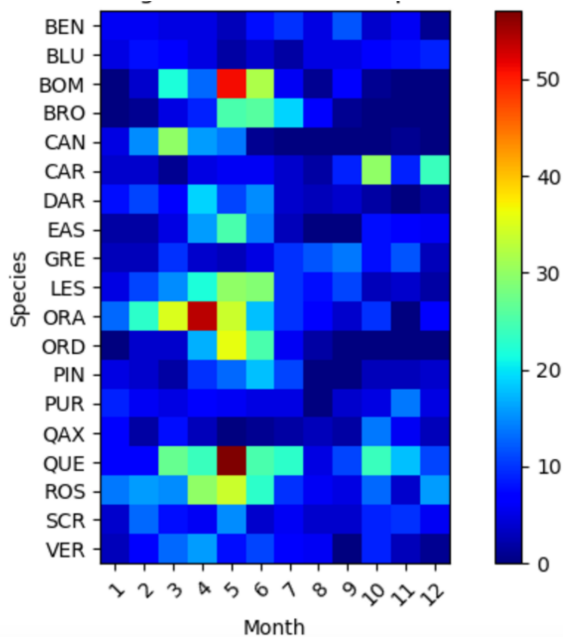


Figure 6. The heatmap represents the temporal activity of the birds in the nature preserve.

Understanding the Accuracy of the Machine Learning Algorithm

In this section, we describe scaffolding techniques to help understand the machine learning algorithm and the results generated. The first one is to allow the user to examine the machine learning predictions versus the expert annotated birdcalls side-by-side to compare accuracy. In 7, users can see an overall view of how closely the machine learning predictions compare with the expert's annotations. From the overall view, users can select each colored dot, which represents a bird, to compare colors, which represent the species. Users can also focus on a particular species to investigate further questions. For example, given the expert classifications, how many of those birds did the machine learning algorithm correctly classify? Figure 8 shows that the Rose-crested Blue Pipit has several inaccuracies in the lower left map of the park. This side-by-side interactive comparison helps users to understand the level of accuracy in the results of the machine learning algorithm. As a result, users gain confidence in the tool.

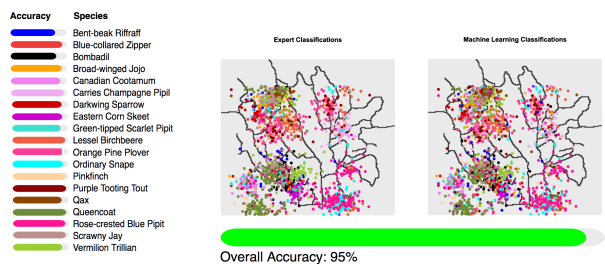


Figure 7. Side-by-side comparison of all the species classifications by the experts and the machine learning techniques.

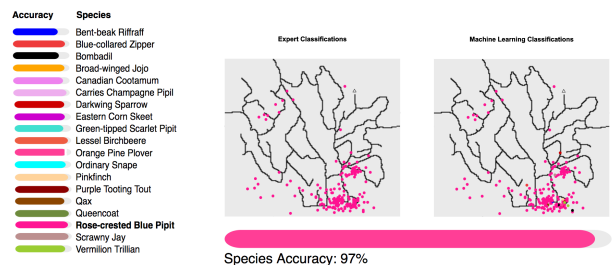


Figure 8. The user can focus on a specific species to investigate how the machine learning results compare with the expert classifications.

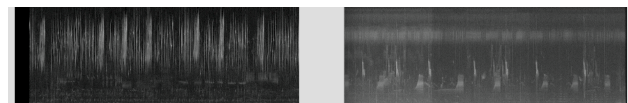


Figure 9. The spectrogram of the Rose-crested Blue Pipit (left) and Bent-beak Riffraff (right).

An additional way to provide users tools to comprehend the inner workings of the machine learning algorithm is to utilize contextual awareness. Each recording can be converted to a visual representation called a spectrogram. Spectrograms depict the frequency of sound as they vary with sound. We generate the spectrogram for the different species. In Figure 9, we see the spectrogram of the Rose-crested Blue Pipit versus the Bent-beak

Riffraff. The user can distinguish the two for the several varying visual characteristics. Through this process, the user is building a mental visual model of the Rose-crested Blue Pipit. Based on this side-by-side spectrogram comparison, the user can guess classifications for unknown bird recordings. This process can be formed as a tutorial which steps the user to classify similar to supervised machine learning. From this tutorial, the user has a better sense of the inner workings of the machine learning techniques. This technique requires an understanding of the context and background of the domain.

Conclusion

In this paper, we make the following contributions: a visual analytic workflow methodology to incorporate the everyday user into understanding and utilizing machine learning. We show how to apply this workflow to the IEEE VAST 2018 Challenge to understand the bird population changes in a nature preserve. For future work, we would like to build a complete system as part of an application and conduct a user study. We would like to see what type of everyday situations and datasets people would choose to solve with machine learning. We would like to see what features help people to understand machine learning. Machine learning is becoming increasingly prevalent in the lives of people around the world. A better understanding of the machine learning and AI that is present in these devices will lead to a relatively smoother, more pleasant experience when using them.

References

- [1] B. Broeksema, T. Baudel, A. Telea, and P. Crisafulli. Decision exploration lab: A visual analytics solution for decision management. *IEEE Transactions on Visualization and Computer Graphics*, 19(12):1972-1981, Dec 2013. doi: 10.1109/TVCG.2013.146.
- [2] E. Brown, J. Liu, C. Brodley, and R. Chang. Disfunction: Learning distance functions interactively. In *2012 IEEE Conference on Visual Analytics Science and Technology, VAST 2012, Seattle, WA, USA, October 14-19, 2012*, pages 8392, 2012.
- [3] M. Chen and A. Golan. What may visualization processes optimize? *IEEE Transactions on Visualization and Computer Graphics*, 22(12):26192632, Dec 2016. doi: 10.1109/TVCG.2015.2513410.
- [4] Costabile, M. F., and Buono, P. (2013). Principles for human-centred design of IR interfaces. In A. M. Ferro, N. Former, P. Mueller, and G. Santucci (Eds.), *Information retrieval meets information visualization* (pp. 2847). Berlin: Springer.
- [5] D. A. Keim, T. Munzner, F. Rossi, M. Verleysen, Bridging Information Visualization with Machine Learning (Dagstuhl Seminar 15101), *Dagstuhl Reports* 5 (3) (2015) 127. doi:10.4230/DagRep.5.3.1.
- [6] P. Federico, M. Wagner, A. Rind, A. Amor-Amors, S. Miksch, and W. Aigner. The role of explicit knowledge: A conceptual model of knowledge-assisted visual analytics. In *Proc. IEEE Conference on Visual Analytics Science and Technology (VAST)*. IEEE, 2017.
- [7] Han, N.C., Muniandy, S.V., Dayou, J., 2011. Acoustic classification of Australian anurans based on hybrid spectral-entropy approach. *Appl. Acoust.* 72, 639645.
- [8] Heer, J., and Perer, A. Orion: A system for modeling, transformation and visualization of multidimensional heterogeneous networks. In *VAST* (2011).
- [9] Lieberman, H. User Interface Goals, AI Opportunities. *AI Magazine* 30(4).
- [10] <http://vacommunity.org/VAST+Challenge+2018+MC1>, Last Accessed: 6/22/2018.
- [11] https://github.com/jameslyons/python_speech_features, Last Accessed: 6/22/2018.
- [12] Demar, J., B. Zupan, and G. Leban. "From experimental machine learning to interactive data mining, a white paper." AI Lab, Faculty of Computer and Information Science, Ljubljana (2004).
- [13] Lasseck, Mario. "Bird song classification in field recordings: winning solution for NIPS4B 2013 competition." *Proc. of int. symp. Neural Information Scaled for Bioacoustics*, sabiod. org/nips4b, joint to NIPS, Nevada. 2013.
- [14] Nowicki, Stephen, and Peter Marler. "How do birds sing?." *Music Perception: An Interdisciplinary Journal* 5.4 (1988): 391-426.
- [15] D. Sacha, M. Sedlmair, L. Zhang, J. A. Lee, D. Weiskopf, S. C. North, and D. A. Keim. Human-Centered Machine Learning Through Interactive Visualization: Review and Open Challenges. *Proceedings of the 24th European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning*, 2016.
- [16] D. Sacha, A. Stoffel, F. Stoffel, BC Kwon, G. Ellis, and D. Keim. Knowledge generation model for visual analytics. *IEEE Trans. Vis. Comput. Graph.*, 20(12):16041613, 2014.
- [17] KO S., AFZAL S., WALTON S., YANG Y., CHAE J., MALIK A., JANG Y., CHEN M., EBERT D.: Analyzing highdimensional multivariate network links with integrated anomaly detection, highlighting and exploration. In *Proc. of IEEE VAST'14* (2014), pp. 8392. 15.
- [18] Dave, N. (2013) Feature Extraction Methods LPC, PLP and MFCC in Speech Recognition. *International Journal for Advance Research in Engineering and Technology*, 1.
- [19] B.Saket,H.Kim,E.T.Brown,andA.Endert.Visualizationbydemonstration: An interaction paradigm for visual data exploration. *IEEE Transactions on Visualization and Computer Graphics*, 23(1):331340, Jan 2017. doi: 10.1109/TVCG.2016.2598839.
- [20] J. Talbot, B. Lee, A. Kapoor, and D. Tan. Ensemblematrix: interactive visualization to support machine learning with multiple classifiers. In *Proceedings of the 27th International Conference on Human Factors in Computing Systems, CHI 2009, Boston, MA, USA, April 4-9, 2009*, pages 12831292, 2009.
- [21] Towsey, Michael W., Anthony M. Truskinger, and Paul Roe. "The navigation and visualisation of environmental audio using zooming spectrograms." *Data Mining Workshop (ICDMW)*, 2015 IEEE International Conference on. IEEE, 2015.
- [22] J.J.vanWijk.Thevalueofvisualization.In*Proc.IEEEVisualization (VIS 05)*, pp. 7986, 2005. doi: 10.1109/VISUAL.2005.1532781.
- [23] X.Wang,D.H.Jeong,W.Dou,S.W.Lee,W.Ribarsky, and R.Chang. Defining and applying knowledge conversion processes to a visual analytics system. *Computers & Graphics*, 33(5):616-623, Oct. 2009.
- [24] L. Yu, W. Wu, X. Li, G. Li, W. S. Ng, S.-K. Ng, Z. Huang, A. Arunan, and H. M. Watt. iviztrans: Interactive visual learning for home and work place detection from massive public transportation data. In *Proc. IEEE VAST*, pages 4956. IEEE, 2015.

Author Biography

Andrew Tran is an undergraduate student in the Computer Science department at California State University, Sacramento.

Yamini Dasu is an undergraduate student in the Computer Science department at California State University, Sacramento.

Anna Baynes is an Assistant Professor at California State University, Sacramento.

JOIN US AT THE NEXT EI!

IS&T International Symposium on

Electronic Imaging

SCIENCE AND TECHNOLOGY

Imaging across applications . . . Where industry and academia meet!



- **SHORT COURSES • EXHIBITS • DEMONSTRATION SESSION • PLENARY TALKS •**
- **INTERACTIVE PAPER SESSION • SPECIAL EVENTS • TECHNICAL SESSIONS •**

www.electronicimaging.org

