THE ASTROPHYSICAL JOURNAL, 975:172 (8pp), 2024 November 10 https://doi.org/10.3847/1538-4357/ad782c

Participatory Science and Machine Learning Applied to Millions of Sources in the Hobby–Eberly Telescope Dark Energy Experiment

Lindsay R. House[1,7], Karl Gebhardt[1], Keely Finkelstein[1], Erin Mentuch Cooper[1,2], Dustin Davis[1], Daniel J. Farrow[3,4], and Donald P. Schneider[5,6]

[1] Department of Astronomy, The University of Texas at Austin, 2515 Speedway Boulevard, Austin, TX 78712, USA; lindsay.r.house@gmail.com
[2] McDonald Observatory, The University of Texas at Austin, 2515 Speedway Boulevard, Austin, TX 78712, USA
[3] Centre of Excellence for Data Science, Artificial Intelligence & Modeling (DAIM), University of Hull, Cottingham Road, Hull, HU6 7RX, UK
[4] E. A. Milne Centre for Astrophysics University of Hull, Cottingham Road, Hull, HU6 7RX, UK
[5] Department of Astronomy and Astrophysics, The Pennsylvania State University, University Park, PA 16802, USA
[6] Institute for Gravitation and the Cosmos, The Pennsylvania State University, University Park, PA 16802, USA
Received 2024 July 11; revised 2024 September 5; accepted 2024 September 6; published 2024 October 31

Abstract

We are merging a large participatory science effort with machine learning to enhance the Hobby–Eberly Telescope Dark Energy Experiment (HETDEX). Our overall goal is to remove false positives, allowing us to use lower signal-to-noise data and sources with low goodness-of-fit. With six million classifications through Dark Energy Explorers, we can confidently determine if a source is not real at over 94% confidence level when classified by at least 10 individuals; this confidence level increases for higher signal-to-noise sources. To date, we have only been able to apply this direct analysis to 190,000 sources. The full sample of HETDEX will contain around 2–3 million sources, including nearby galaxies ([O II] emitters), distant galaxies (Lyα emitters or LAEs), false positives, and contamination from instrument issues. We can accommodate this tenfold increase by using machine learning with visually vetted samples from Dark Energy Explorers. We have already increased by over tenfold the number of sources that have been visually vetted from our previous pilot study where we only had 14,000 visually vetted LAE candidates. This paper expands on the previous work by increasing the visually vetted sample from 14,000 to 190,000. In addition, using our currently visually vetted sample, we generate a real or false positive classification for the full candidate sample of 1.2 million LAEs. We currently have approximately 17,000 volunteers from 159 countries around the world. Thus, we are applying participatory or citizen scientist analysis to our full HETDEX data set, creating a free educational opportunity that requires no prior technical knowledge.

Unified Astronomy Thesaurus concepts: Cosmological constant (334); Dark energy (351); Cosmological parameters (339); Cosmological parameters from large-scale structure (340); Astronomy education (2165); Lyα galaxies (978); Baryon acoustic oscillations (138)

1. Introduction

The continually increasing size of the astronomical data sets requires new analysis techniques to be leveraged in order to handle these efficiently and extract robust scientific results. While the use of machine learning (ML) has been available for decades, its use is now an essential component within the field (C. J. Fluke & C. Jacobs 2020). ML is a subfield of artificial intelligence where algorithms are used to recognize patterns, make predictions, and even apply these results to new applications (C. J. Torney et al. 2019; O. Zawacki-Richter et al. 2019).

One of the primary issues with ML is determining the accuracy for a given application and interpreting the results. There are multiple ways to assign accuracy, including visual vetting on a subset of the sample. Human visual vetting, even for verification, quickly becomes intractable as the data sets increase in size. In fact, many of the data samples have surpassed the ability to use visual vetting within a given collaboration due to the limited number of eyes available.

Many large astronomical experiments have already paved the way in using human classification to reach science goals. The collaboration between scientists and the public is known as participatory science, also called citizen science (A. H. Kimura & A. Kinchy 2016; B. Strasser et al. 2019). There are multiple ways to include these participants within the various pipelines. Large surveys that have done so are the Cosmic Assembly Near-infrared Deep Extragalactic Legacy Survey (CANDELS) and the complementary Galaxy Zoo participatory science project (B. D. Simmons et al. 2016; K. L. Masters & Galaxy Zoo Team 2020). In addition, the Laser Interferometer Gravitational Wave Observatory (LIGO) utilizes the Gravity Spy participatory science project (M. Zevin et al. 2024).

The Hobby–Eberly Telescope Dark Energy Experiment (HETDEX) is designed to study the expansion rate of the universe at $1.9 < z < 3.5$ with an accuracy comparable to even the best low-z experiments (K. Gebhardt et al. 2021; G. J. Hill et al. 2021). To date, this multiyear program has already generated nearly one billion spectra, and one trillion resolution elements. The sheer size of this data set necessitates robust statistical techniques to identify false positives, [O II] contamination, and artifacts. Even with the most detailed analysis, we have not yet reached our design specifications for false positive rate (E. Mentuch Cooper et al. 2023). Therefore, we have created and used the participatory science project Dark Energy Explorers[8] to improve HETDEX with ML techniques.

[7] NSF Graduate Research Fellow.

[8] https://www.zooniverse.org/projects/erinmc/dark-energy-explorers

In L. R. House et al. (2023), we provide the first use of ML to the HETDEX database. There, we find that the Dark Energy Explorers results are accurate to over 95% with the ability to find false positives. This work only considered 10% of the full HETDEX database. Given the success we had in using ML and Dark Energy Explorers, our goal is to apply to the full database, where we need to classify about 10 million spectra.

The current database for HETDEX has about 1.2 million Lyα emitters (LAEs) and about 1 million OII emitting galaxies. These numbers come from a down selection of spectra based on the signal-to-noise (SN) and goodness-of-fit of the emission line. The initial sample is about 10 million spectra. We realize that our selection, as it is, fails to exclude all false positives and fails to include all true positives. When we can robustly identify more sources, our cosmological constraints on the expansion of the universe will increase by about the square root of the numbers (A. S. Leung et al. 2017; D. J. Farrow et al. 2021; K. Gebhardt et al. 2021; D. Davis et al. 2023; E. Mentuch Cooper et al. 2023). Similarly, the false positive rate makes our accuracy worse. Therefore, the larger the total number of LAE sources and the less false positives due to noise (FP) and [O II] contamination will allow us to more accurately measure the expansion rate of the universe.

HETDEX is a untargeted survey, which refers to observations that are conducted without a specific predetermined focus on particular objects, eliminating selection bias, yet creating difficulties when classifying. In this paper we discuss an innovative approach to a data pipeline which will allow accurate and efficient labeling. HETDEX uses the LAEs as a cosmological tracer. Thus, our objects of interest are the LAE galaxies, which must be sorted and selected from the whole sample of objects collected. Just as important in selecting these galaxies is the removal of artifacts, false positives, [O II] emitting galaxies, meteor trails, and other emission-like features. LAEs have been detected over an extensive redshift range and their redshifted 1215.67Å line can easily be detected using low-resolution spectroscopy or narrow-band imaging (C. Gronwall et al. 2007). The large-scale clustering of the LAEs will allow us to see the effects of dark energy on the universe and yield the cosmological parameters of scientific interest.

We must differentiate between false positives and galaxy misclassifications since these two errors affect the HETDEX correlation analysis differently. For false positives, we are referring to noise or pixel defects that are confused as an apparent emission line. The misidentification of galaxies is a larger issue, especially when [O II] emitters are designated as LAEs. In this case, the clustering signal of the [O II] galaxies will leave an imprint on the clustering of the LAEs. Therefore, this demonstrates the importance of a robust, clean catalog and will showcase how Dark Energy Explorers and ML are powerful techniques for addressing this issue. Here, we establish a method of how visual vetting has the ability to significantly improve on the HETDEX algorithms.

Zooniverse,[9] the world's largest participatory science platform, allows us to make progress on visually vetting a large subset of the HETDEX data set that can be applied to the full sample. Classifying the whole HETDEX catalog, with many millions of sources labeled, would enable significant improvements toward the measurements of the HETDEX cosmological parameters (L. R. House et al. 2023).

We show here how to use participatory science and ML to classify 1.2 million HETDEX sources. In Section 2, we discuss participatory science and the success of Dark Energy Explorers to explore various HETDEX regimes and the statistics we use for the classification. Section 3 focuses on the ML algorithm and how we incorporate millions of sources with visually vetted sources. Finally, section 4 shows how we utilize the algorithm and Dark Energy Explorers in the creation of a data pipeline for the current HETDEX sample.

2. Dark Energy Explorers: A Participatory Science Campaign

Participatory science has continued to grow with organizations such as Zooniverse, SciStarter,[10] and CitizenScience.gov.[11] To date, Zooniverse alone has garnered millions of participants and has launched close to 500 projects. This demonstrates the clear demand for visual classification in research and shows that combining human and machine classifications can efficiently produce results superior to those of either one alone (L. Trouille et al. 2019). This combination of techniques is what we aim to accomplish with Dark Energy Explorers.

2.1. Developing Dark Energy Explorers

We launched Dark Energy Explorers in 2021 February and the project has had an impact of roughly six million classifications in the project's lifetime so far. Zooniverse is a host platform that has cultivated a participatory science community and has launched hundreds of projects. Once on Zooniverse, either app or website, participants must login or create a new account to save their classifications. As an official NASA participatory science project, we can also be found under the Citizen Science Section of NASA.gov.[12] Once on the Dark Energy Explorers home page, choose the active workflow, "Fishing for Signal in a Sea of Noise," which will prompt a tutorial if you are a new visitor. The tutorial demonstrates how to classify the HETDEX data (discussed in more detail in L. R. House et al. 2023).

When creating this tool to be used by the general public, it was essential to simplify the classification process into digestible, jargon-free tasks. The tutorial and field guide serve as the mechanism to do just that. The tutorial walks participants through what to look for in the HETDEX data in just a few easy-to-understand steps. The primary criteria users consider have not changed from the prior work in L. R. House et al. (2023). As a reminder, for the "Fishing for Signal in a Sea of Noise" the workflow participants classify sources based on:

1. The quality of the data collected,
2. The strength of the emission line, and
3. The appearance of the emission line in at least one or more of the fiber spectrum.

Once trained, the participants are provided deidentified, processed data in the form of user-friendly imaging. The volunteers must select between two options: Keep this Galaxy

9 https://www.zooniverse.org

10 https://scistarter.org/

11 https://www.citizenscience.gov/

12 https://science.nasa.gov/citizen-science/dark-energy-explorers/

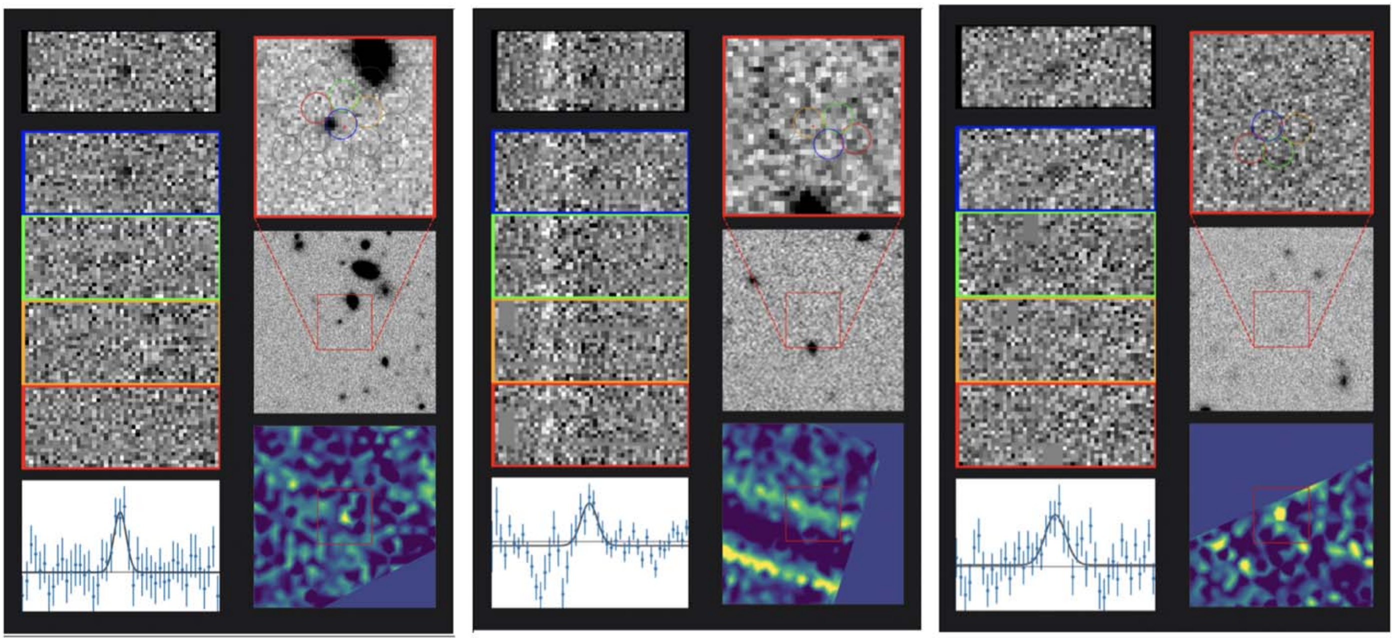

Figure 1. Examples of the "mini's" from Dark Energy Explorers "Fishing for Signal in a Sea of Noise" workflow. From left to right: keep (real galaxy/emission line), throwback (bad detection), and a tricky case that might need more information.

or Throwback. The field guide is a shortened tutorial that allows quick access to classification reminders. See Figure 1 for a comparison. Figure 1 shows an example of three sources and what they would look like to the participants on Dark Energy Explorers. In brief, these show a combination of the various types of data we collect with the HET, i.e., fiber cutouts, imaging, flux map, and emission line. See L. R. House et al. (2023) for a detailed description of each Dark Energy Explorers input image. On the left is an example of a real galaxy or LAE; in the middle is a source that is an artifact; and on the right is a source that is possibly real is a possibility to be real, but the HETDEX team would need to explore more information to decide. This is one of the key advantages of Dark Energy Explorers because it allows for a triage of sources that lie in this unclear regime and the team can quickly identify areas that need a closer look; rather than sifting through all the sources ourselves, this technique creates a smaller, more manageable small subset.

From K. Gebhardt et al. (2021), D. Davis et al. (2023), E. Mentuch Cooper et al. (2023), and L. R. House et al. (2023) it is essential to explore lower SN regimes to maximize the number of LAE candidates that will lead to a more precise HETDEX cosmology. The previous work began with $\mathrm{SN} > 6$ sources when using Dark Energy Explorers, but this work has shown we can progress to lower SN regimes while still retaining accuracy with visual vetting. So far, ≈190,000 LAE candidates have been classified by the Dark Energy Explorers down to a SN ratio of 4.8. This includes HETDEX Data Release 2, 3, 4, and the COSMOS field.

2.2. Using Dark Energy Explorers as an Education and Public Engagement Tool

The research results and improvements to HETDEX would not be possible without the classifications from our Dark Energy Explorers participants. While the primary goal of Dark Energy Explorers is to improve the accuracy of the HETDEX cosmology, it has quickly demonstrated the impact it has as an extraordinary educational tool. Collaboration with McDonald Observatory, where the Hobby–Eberly Telescope (HET) is based, has allowed us to grow our volunteer base of Dark Energy Explorers while also allowing unique opportunities at the Observatory. We currently have an exhibit at the HET visitors center, which allows a rare experience of visiting a telescope and then being able to classify the data as an amateur researcher. In addition, McDonald Observatory has assisted in developing worksheets and videos for educators to use in traditional classrooms, libraries, and museums. We have continued engagement with our volunteers through Teleconferencing/Zoom nights, our Dark Energy Explorers YouTube Channel,[13] design competitions, and blog posts. The following section discusses the overall results of implementing these public engagement efforts, and we hope it continues to be a rewarding educational outlet for our dedicated participants.

2.3. Results and Impact of Dark Energy Explorers

Our overall goal is to vet all HETDEX sources visually. By the end of HETDEX, we will have nearly 10 million sources; this is effectively impossible to do within the team. The only possibility is to include many participatory scientists in classifying the sources. We have been doing this through the Dark Energy Explorers since 2021 February. Dark Energy Explorers has already proven incredibly successful at accomplishing this goal, with roughly six million classifications in the project's lifetime. For the current workflow, that is approximately 190,000 unique spectra that are identified by a minimum of 10 individuals. Figure 2 shows our classifications as a function of time, where the purple line demonstrates the total number of classifications each month since Dark Energy Explorers launched with the first workflow, "Nearby versus Distant," which has since been retired (L. R. House et al. 2023).

[13] https://www.youtube.com/@DarkEnergyExplorers

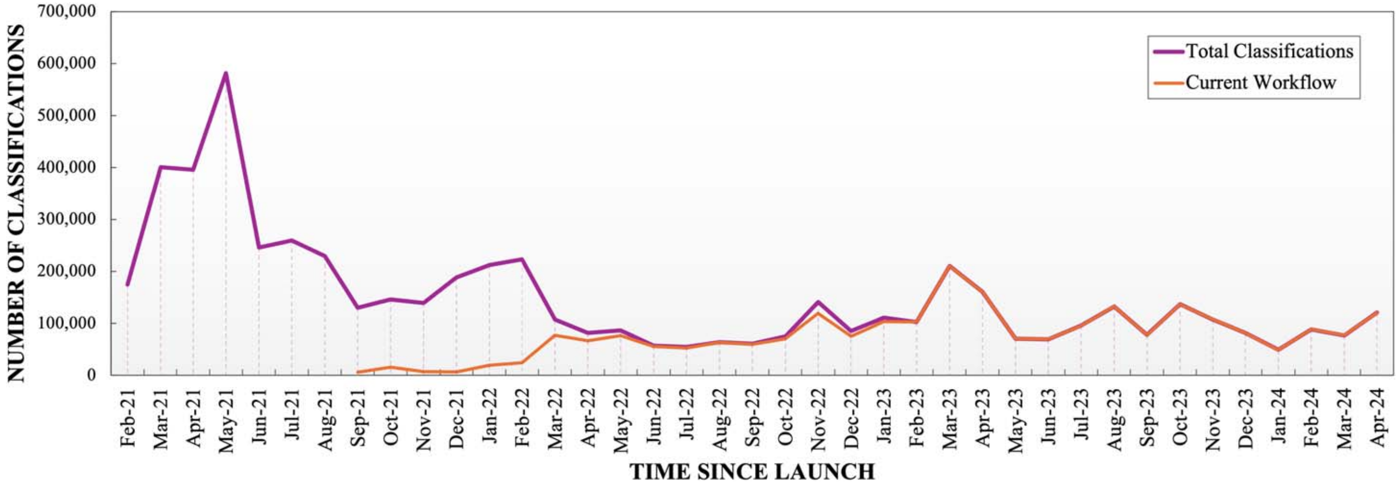

Figure 2. The total classifications, where each source is classified by a minimum of 10 volunteers, collected on Dark Energy Explorers since launch. The orange line represents the current active workflow classifications "Fishing for a Signal in a Sea of Noise" and the purple represents the total classifications by month.

The current workflow "Fishing for Galaxies in a Sea of Noise" has 2.1 million classifications resulting in 190,000 completed LAE candidates, which are vetted by at least 10 different participants, giving us the confidence to rely on this average (L. R. House et al. 2023). Figure 2 represents the current workflow in orange and the total classifications since its launch in September of 2021. These millions of classifications have been done by approximately 17,000 volunteers that represent over 159 countries all over the world, with the top three being in the United States, United Kingdom, and India. The Dark Energy Explorers average 153,000 classifications per month and we hope to increase this through outreach and engagement efforts to reach our goal of the entire HETDEX catalog being classified.

3. Visualization in Machine Learning

While ML has been shown to be effective, many aspects of research are still met with obstacles that require human verification (S. Amershi et al. 2014). This combination of machines and humans, we argue, is a good way to make a large problem more efficient and accurate. In particular, it provides necessary insight to the interpretation of the data, in the case of HETDEX the high-dimensional spectral elements (K. Crowston et al. 2017; C. J. Torney et al. 2019). Other participatory science projects have led the way in imaging classification with both ML and human vetting, yet we discuss how this can be done with spectroscopic inputs (B. D. Simmons et al. 2016; K. L. Masters & Galaxy Zoo Team 2020).

For the Dark Energy Explorers project utilizing ML has proven to be most useful for identifying artifacts. We have done this through the "Fishing for Signal in a Sea of Noise" workflow discussed in Section 2. Using the same DEE_probability from prior work, we assign each classification a value of 1.0 meaning an object is a real LAE detection or a 0.0 which identifies a false positive or artifact. Aggregating over a minimum of 10 Dark Energy Explorers participants, the values are then averaged to generate a DEE_probability. Once those labels have been acquired and the DEE_probability determined for each visually vetted source, we can apply an ML algorithm.

3.1. Algorithm: t-SNE

The ML algorithm discussed and used in this work is known as t-distributed stochastic neighbor embedding (or t-SNE). t-SNE serves as a statistical method for visualizing high-dimensional data by giving each data point a location in a two or three-dimensional map (L. van der Maaten & G. Hinton 2008; L. van der Maaten 2015). For the plethora of high-dimensional data that we get from HETDEX, there will be millions of elements with 1036 dimensions, this will result in billions of spectral elements that t-SNE can handle exceptionally well.

Using unsupervised ML, specifically t-SNE, for large data sets has proven effective when the parameters are optimized accordingly (A. C. Belkina et al. 2019). t-SNE has a cost function that with different initializations, such as the HETDEX spectra and algorithm parameters, we can get different results. The results depend on the random seed, the data input, and the hyperparameters chosen. Despite this, t-SNE visualizations can be effective in grouping together sources with similar spectra, especially artifacts (e.g., L. R. House et al. 2023).

The results are not reproducible but when using tuned hyperparameters it will keep the global aspects of the data, and in the case of cleaning the catalog of artifacts, such as in HETDEX, this works well. t-SNE and many dimensionality reduction algorithms can be tricky to interpret (L. van der Maaten & G. Hinton 2008) and the data from Dark Energy Explorers provides context to interpret this data in a scientifically meaningful way. First, we set the dimensionality reduction to reduce from 50 to 2. Importantly, the perplexity parameter has been shown to give the best results with values of 5–50 (L. van der Maaten 2015). Therefore, the perplexity was found to be optimized at 30 and combining this with an initial iteration parameter of 1000 ensured that the algorithm reached a stable configuration. Utilizing the `scikit-learn` Python package, the left-hand side of Figure 3 displays the results of HETDEX LAE candidates in black after using t-SNE (F. Pedregosa et al. 2012).

3.2. Input Selection Criteria

To better differentiate our sources of interest, i.e., the Lyα emitting galaxies, we select for the Lyα emission line in our

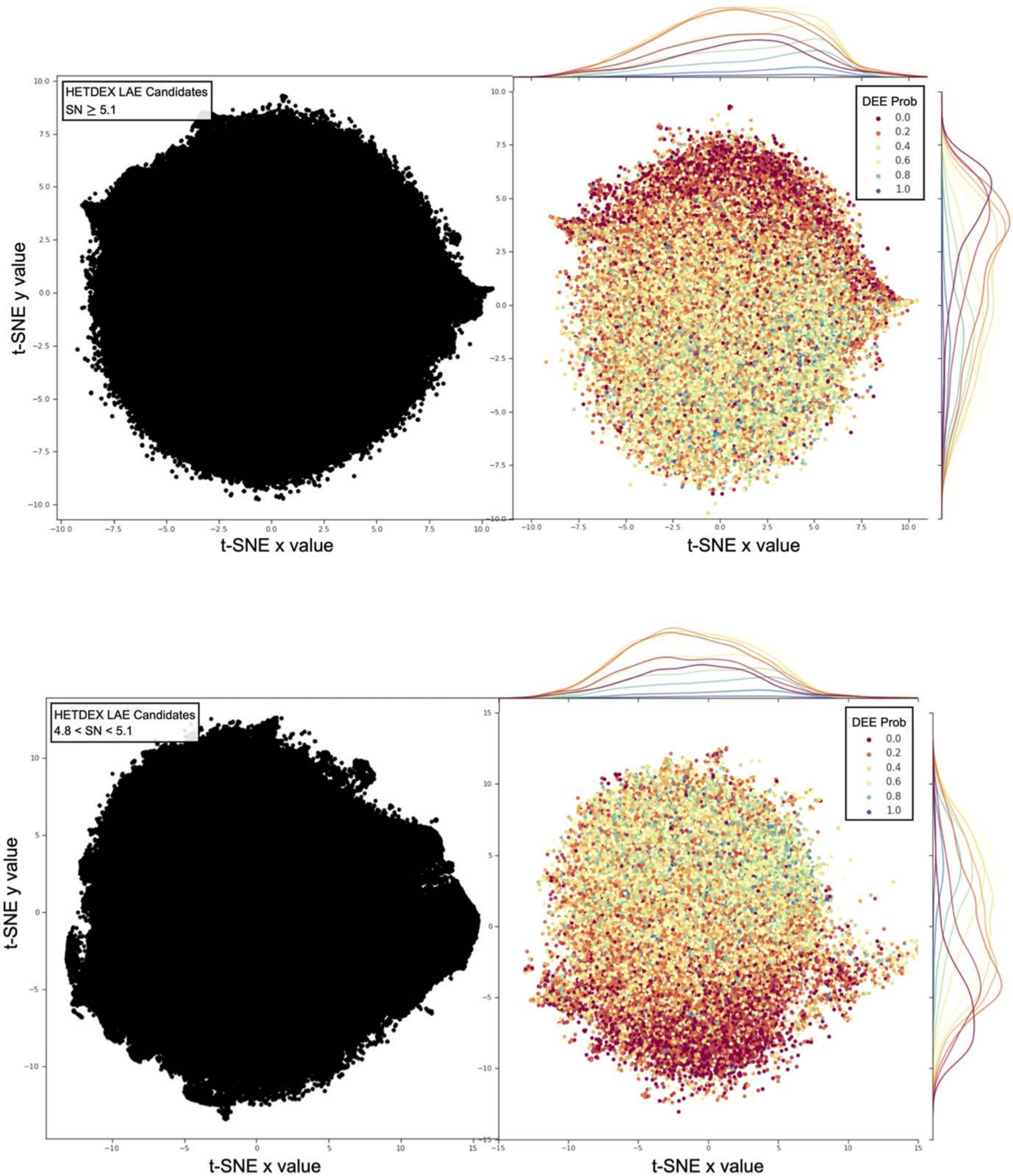

Figure 3. The results of the t-SNE ML algorithm, each with the same hyperparameters and each row has the same t-SNE projection. On the left-hand side, the plots shown in black are ≃1.2 million HETDEX LAE candidates. On the right-hand side, the colored plots shown are a result of the ≃190,000 LAE candidates that have been labeled by Dark Energy Explorers and assigned a DEE_probability. The top two plots show the LAE candidates above SN ratio $\geqslant 5.1$ and the two bottom plots show the LAE candidates with $4.8 < \mathrm{SN} < 5.1$. Note: The orientation of the t-SNE axes relative to the data points has no inherent meaning or significance beyond the visualization itself. This is simply a characteristic of the dimensionality reduction algorithm.

HETDEX catalog. The spectral elements of these LAE candidates are then used to cut around the emission line of the one-dimensional spectra. We select for 50 Å on either side of the Lyα emission. This results in a 100 Å cut, in which HETDEX captures flux in 2 Å bins, resulting in 50 dimensions used for the t-SNE input for each LAE candidate. We further distribute into two sub-samples by SN ratio to downsize our sample of 1.2 million sources. The result of the t-SNE runs can be shown in Figure 3. Shown at the top of Figure 3 is the SN ratio range $\geqslant 5.1$ with $\sim$600,000 sources. Similarly, shown at the bottom of Figure 3 is the range of $5.1 > \mathrm{SN} > 4.8$ with another $\sim$600,000 sources for a total of 1.2M LAE candidate detections. Our labels from Dark Energy Explorers that are fed into t-SNE include the low-SN sources ($6 > \mathrm{SN} > 4.8$) to ensure the algorithm can train on a similar low-SN sample and eliminate bias. The full SN range from Dark Energy Explorers is then visualized in color, labeled, and shown in Figure 3 for each SN bin, respectively.

4. Incorporating the Dark Energy Explorers Results with Machine Learning

As large astronomical surveys grow, so do the artifacts and contamination. These artifacts are either found manually (E. Mentuch Cooper et al. 2023) or with artificial intelligence (K. Gebhardt et al. 2021). While artificial intelligence can be useful, the elements from HETDEX are better originally identified by visual vetting. Here we discuss how we analyze the visual classifications from the participants of Dark Energy Explorers and then use the classifications to expand to the full 1.2 million LAE candidates.

4.1. Analysis of Visual Vetting Statistics

Following the original work, we focus on the false positives. Using the same methods from prior work, a DEE_probability of 1.0 means that an object is a real LAE detection, and a probability of 0.0 identifies a false positive or artifact. Therefore, every source is identified by a minimum of 10 Dark Energy Explorers participants and averaged together to get a DEE_probability.

Previously, in L. R. House et al. (2023) a DEE_probability $\leqslant$0.3 results in 92% accuracy across SN. In addition, L. R. House et al. (2023), demonstrated a Dark Energy Explorers probability of below 0.1 gave 98% agreement, which allowed for very efficient and accurate identification of false positives. All of these sources from the pilot sample were cross-examined by members of the HETDEX team, which resulted in the development of the DEE_probability accuracy. Given this high accuracy for identifying false positives, we could confidently move forward with applying these results to a broader full sample with ML. Those are the results explained here.

4.2. Application to the Full HETDEX LAE Catalog using Nearest Neighbors Technique

Following the visual vetting analysis, our next goal is to apply this approach to the full HETDEX LAE catalog as opposed to just a subset. Figure 3 shows the distribution in t-SNE space for the full sample and for those classified by Dark Energy Explorers. The top left-hand panel shows $\sim$600,000 sources for $S/N > 5.1$. The bottom left-hand panel shows the $\sim$600,000 sources that have $4.8 < \mathrm{SN} < 5.1$. The right-hand panels are the corresponding t-SNE distributions for those sources that have a DEE_probability, with the points colored by their probability.

For each source in the panels on the left-hand side, we find the 50 nearest sources in t-SNE space that have a DEE_probability. The average of these 50 is then applied to each of the sources, which we call the DEE_mean. Following the same logic as the DEE_probability, the DEE_mean ranges from 0 to 1.0, from lowest to highest probability of being a real galaxy detection. The DEE_mean becomes an additional statistical method for interpreting the t-SNE analysis.

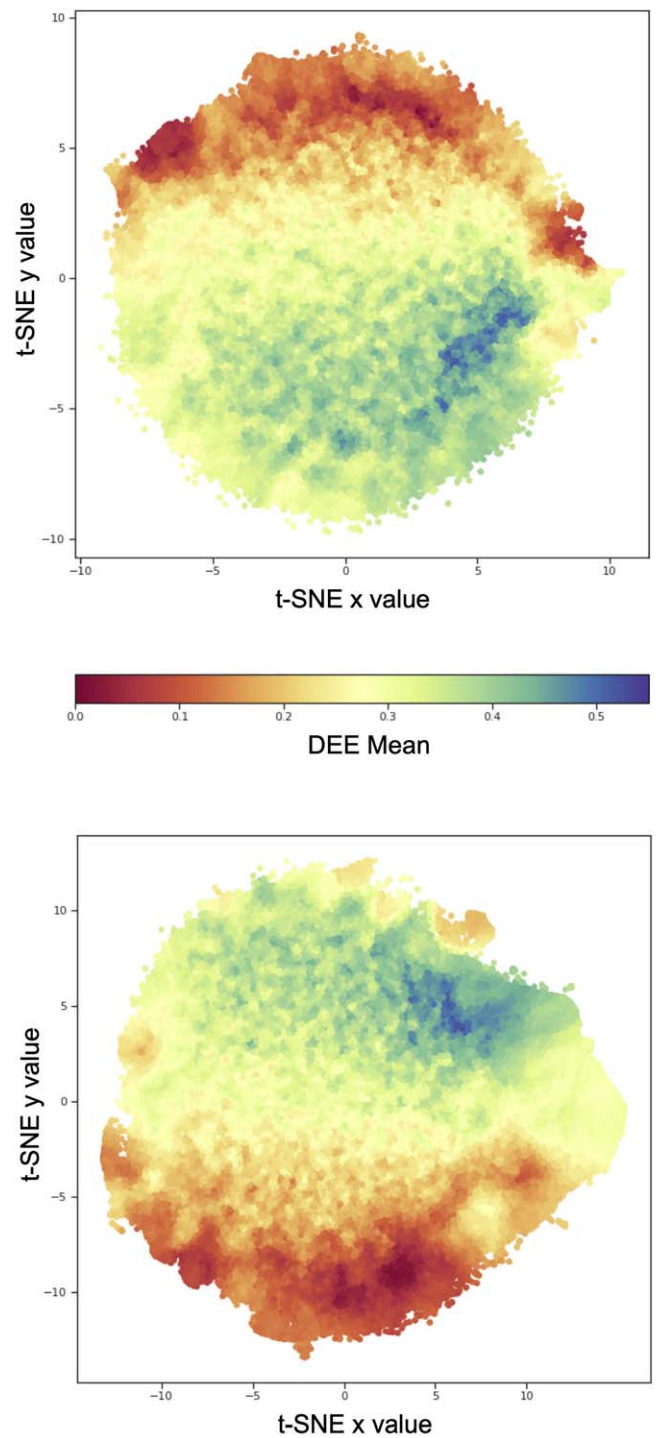

Figure 4. The results of the t-SNE ML algorithm, each with the same hyperparameters. The colors represent the DEE_mean which was developed using the nearest neighbors method. Note: The orientation of the t-SNE axes relative to the data points has no inherent meaning or significance beyond the visualization itself. This is simply a characteristic of the dimensionality reduction algorithm.

Figure 4 shows the same t-SNE distribution as in Figure 3, except we now color the points by their DEE_mean, generated from the nearest neighbors method. It is clear that there is strong clustering for both sources with high and low probabilities of being real. As stated, we find that the Dark

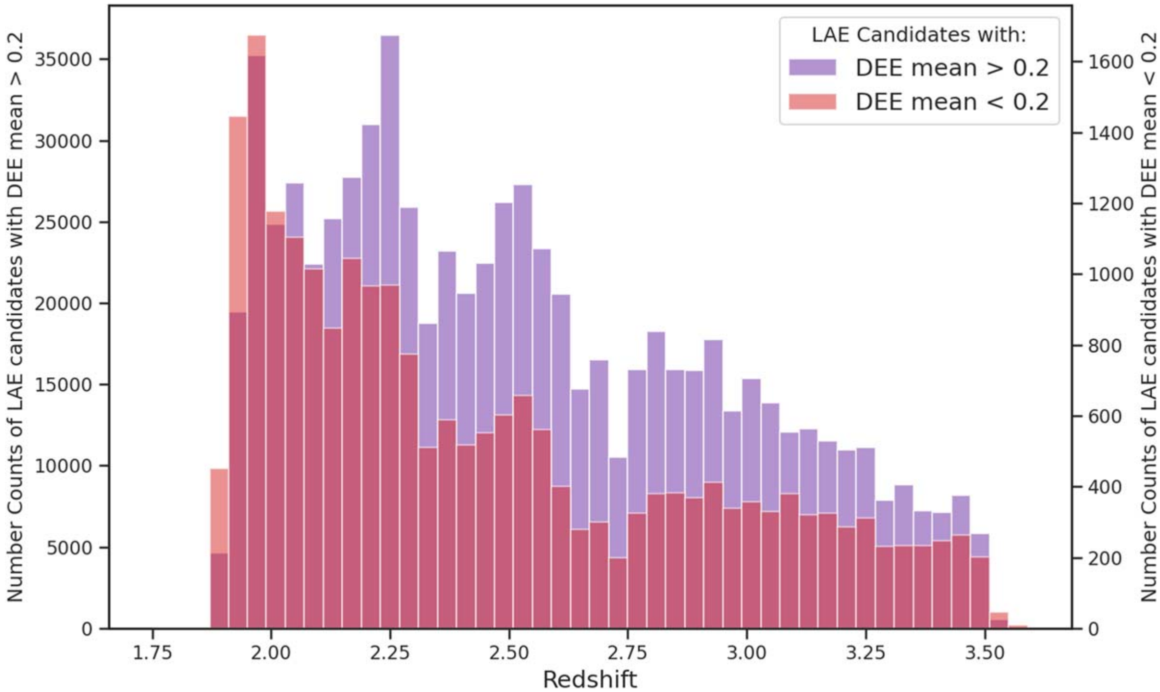

Figure 5. The redshift distribution of the 1.2 million HETDEX LAE candidates with the cut on DEE_mean. DEE_mean is created using the nearest neighbors technique. The histogram in pink identifies the false positives with a DEE_mean < 0.2 and the purple represents a DEE_mean > 0.2, each cut is scaled for comparison, differing by approximately a factor of 20.

Energy Explorers are best at identifying false positive due to instrumental and reduction artifacts, and we use the DEE_mean to further remove these sources. Thus, the red points in Figure 4 are considered false detections and will be removed from subsequent analysis. We find that a cut of DEE_mean <0.2 provides an adequate cut.

5. Overall Implications to HETDEX and the Cosmology

Similar to L. R. House et al. (2023), we rely on visually vetting from the HETDEX team in order to determine accuracy for both DEE_probability and DEE_mean. We previously determined an accuracy of 92% for DEE_probability <0.3, implying that when 7 out of 10 Dark Energy Explorers call a source false, they are correct 92% of the time.

For DEE_mean, the HETDEX team visually vets about 300 sources at random, and we find that for DEE_mean <0.1 we agree 94% of the time and DEE_mean <0.2 we agree 91% of the time. We use this cut as a further technique to remove false positives from the HETDEX data set. In this way, we are able to use the participatory scientists for the full sample. Eventually, we plan to visually vet all sources, without having to use a nearest neighbor approach.

For the current data set of 1.2 million sources, we have 62,000 with DEE_mean <0.2. These will be removed from the sample. While they only represented 5%, they may have certain aspects that could bias the cosmological analysis and therefore their removal is essential. The most obvious concern is shown in Figure 5. In this figure, we show the histogram of the full sample and a scaled (roughly 20-fold) histogram of the sources that we remove based on the DEE_mean. There is clearly a bias toward low redshifts for the sources that are being removed. The point is that there appears to be an increase in the contamination rate as a function of redshift. The overall contamination remains small, and we will study whether this bias could have implications for the cosmology in future work.

An important aspect is that we do not necessarily know the truth. For example, the HETDEX team, while having the deepest understanding of the data set, can make misclassifications. It is possible that having multiple individuals, as we do in Dark Energy Explorers, provides a more robust result. Additionally, using a nearest neighbor approach could be more robust than the individual measures because it relies on averaging within t-SNE space. In the end, HETDEX will use multiple measures of the source classification to understand the influence on the cosmological analysis. The DEE_prob and DEE_mean values and recommendations will be included in the future HETDEX Data Release 3 (HDR3) catalog.

Dark Energy Explorers has had an extremely positive impact on our informal science community and on our overall goal to improve the accuracy of HETDEX. In our future work we aim to continue both of these efforts to reach a HETDEX catalog that has been completely visually classified. This will be roughly a factor of 10 more than what we have now, and we will continue to pave the way for a successful method to use ML and participatory science.

Acknowledgments

The results of the Dark Energy Explorers would not be as robust and useful if not for the care and dedication of the volunteers. We are extremely grateful for their work. It is having a large impact and is motivational.

L.H. acknowledges support from NSF GRFP DGE 2137420 and NASA 21-CSSFP21-0009. K.G. acknowledges support from the NSF-2008793 and from NASA 21-CSSFP21-0009.

Dark Energy Explorers is recognized as an official NASA Citizen Science partner. This publication uses data generated via the Zooniverse.org platform[14], development of which is funded

[14] Zooniverse.org

by generous support, including a Global Impact Award from Google, and by a grant from the Alfred P. Sloan Foundation.

HETDEX is led by the University of Texas at Austin McDonald Observatory and Department of Astronomy with participation from the Ludwig-Maximilians-Universität München, Max-Planck-Institut für Extraterrestrische Physik (MPE), Leibniz-Institut für Astrophysik Potsdam (AIP), Texas A&M University, the Pennsylvania State University, Institut für Astrophysik Göttingen, the University of Oxford, Max-Planck-Institut für Astrophysik (MPA), the University of Tokyo, and Missouri University of Science and Technology. In addition to institutional support, HETDEX is funded by the National Science Foundation (grant AST-0926815), the State of Texas, the US Air Force (AFRL FA9451-04-2-0355), and generous support from private individuals and foundations.

The Hobby–Eberly Telescope (HET) is a joint project of the University of Texas at Austin, the Pennsylvania State University, Ludwig-Maximilians-Universität München, and Georg-August-Universität Göttingen. The HET is named in honor of its principal benefactors, William P. Hobby and Robert E. Eberly. The Institute for Gravitation and the Cosmos is supported by the Eberly College of Science and the Office of the Senior Vice President for Research at the Pennsylvania State University.

The authors acknowledge the Texas Advanced Computing Center (TACC)[15] at The University of Texas at Austin for providing high performance computing, visualization, and storage resources that have contributed to the research results reported within this paper. This research made use of `scikit-learn`, an open-source Python package for machine learning.

ORCID iDs

Lindsay R. House https://orcid.org/0000-0002-1496-6514
Karl Gebhardt https://orcid.org/0000-0002-8433-8185
Keely Finkelstein https://orcid.org/0000-0003-0792-5877
Erin Mentuch Cooper https://orcid.org/0000-0002-2307-0146
Dustin Davis https://orcid.org/0000-0002-8925-9769
Daniel J. Farrow https://orcid.org/0000-0003-2575-0652
Donald P. Schneider https://orcid.org/0000-0001-7240-7449

References

Amershi, S., Cakmak, M., Knox, W. B., & Kulesza, T. 2014, AI Magazine, 35, 105
Belkina, A. C., Ciccolella, C. O., Anno, R., et al. 2019, NatCo, 10, 5415
Crowston, K., Àsterlund, C., & Lee, T. K. 2017, in Hawaii Int. Conf. Proc. on System Sciences (Honolulu: University of Hawaii at Manoa), http://hdl.handle.net/10125/41159
Davis, D., Gebhardt, K., Mentuch Cooper, E., et al. 2023, ApJ, 946, 86
Farrow, D. J., Sá nchez, A. G., Ciardullo, R., et al. 2021, MNRAS, 507, 3187
Fluke, C. J., & Jacobs, C. 2020, WDMKD, 10, e1349
Gebhardt, K., Mentuch Cooper, E., Ciardullo, R., et al. 2021, ApJ, 923, 217
Gronwall, C., Ciardullo, R., Hickey, T., et al. 2007, ApJ, 667, 79
Hill, G. J., Lee, H., MacQueen, P. J., et al. 2021, AJ, 162, 298
House, L. R., Gebhardt, K., Finkelstein, K., et al. 2023, ApJ, 950, 82
Kimura, A. H., & Kinchy, A. 2016, Engag. Sci. Technol. So., 2, 331
Leung, A. S., Acquaviva, V., Gawiser, E., et al. 2017, ApJ, 843, 130
Masters, K. L. & Galaxy Zoo Team 2020, in IAU Symp. 353, Galactic Dynamics in the Era of Large Surveys (Cambridge: Cambridge Univ. Press), 205
Mentuch Cooper, E., Gebhardt, K., Davis, D., et al. 2023, ApJ, 943, 177
Pedregosa, F., Varoquaux, G., Gramfort, A., et al. 2012, JMLR, 12, 2825
Simmons, B. D., Lintott, C., Willett, K. W., et al. 2016, MNRAS, 464, 4420
Strasser, B., Baudry, J., Mahr, D., Sanchez, G., & Tancoigne 2019, Sci. Technol. Stud., 32, 52
Torney, C. J., Lloyd-Jones, D. J., Chevallier, M., et al. 2019, MEcEv, 10, 779
Trouille, L., Lintott, C. J., & Fortson, L. F. 2019, PNAS, 116, 1902
van der Maaten, L. 2015, JMLR, 15, 3221
van der Maaten, L., & Hinton, G. 2008, JMLR, 9, 2579
Zawacki-Richter, O., Marín, V. I., Bond, M., & Gouverneur, F. 2019, Int. J. Educ.Technol. High. Educ., 16, 39
Zevin, M., Jackson, C. B., Doctor, Z., et al. 2024, EPJP, 139, 100

[15] http://www.tacc.utexas.edu