Monthly Notices of the ROYAL ASTRONOMICAL SOCIETY

MNRAS **538,** 301–311 (2025) https://doi.org/10.1093/mnras/staf262
Advance Access publication 2025 February 12

Decision tree ensembles for automatic spectroscopic classification of tidal disruption events

Andreas Humpe [1,2★] Paolo A Mazzali,[1,3] Avishay Gal-Yam[4] and Ivo Siekmann[5,6,7]

[1]*Astrophysics Research Institute, Liverpool John Moores University, Liverpool L3 5RF, UK*
[2]*Institute for Applications of Machine Learning and Intelligent Systems (IAMLIS), Munich University of Applied Sciences, Lothstr. 34, D-80335 Munich, Germany*
[3]*Max Planck Institute for Astrophysics, Karl-Schwarzschildstr. 1, D-85748 Garching, Germany*
[4]*Department of Particle Physics and Astrophysics, Weizmann Institute of Science, 76100 Rehovot, Israel*
[5]*School of Computer Science and Applied Mathematics, Liverpool John Moore University, Byrom Way, Liverpool L3 3AF, UK*
[6]*Data Science Research Centre, Liverpool L3 3AF, UK*
[7]*Liverpool Centre for Cardiovascular Science (LCCS), Liverpool L69 7TX, UK*

Accepted 2025 February 6. Received 2025 January 25; in original form 2024 November 19

ABSTRACT

The principal objective of this study was to develop a reliable model for the automatic classification of tidal disruption events (TDEs) using spectroscopic data. A total of 147 TDE spectra and 3626 spectra of various supernova types and AGNs were included in the data, sourced from PESSTO-SSDR1-4. An ensemble learning approach was employed using bagging with decision trees as base learners, optimized through cost-sensitive analysis and Bayesian hyperparameter tuning. A high test accuracy of 97.67 per cent, with balanced precision and recall, was achieved by the optimized model. To enhance TDE detection, a dynamic threshold adjustment was applied, prioritizing recall, which increased from 47.22 per cent to 83.33 per cent. Most TDEs were correctly identified due to this adjustment, with a reduction in precision from 85.00 per cent to 22.22 per cent and a decrease in overall accuracy from 97.67 per cent to 88.23 per cent, reflecting the prioritization of recall over precision. Relative to their occurrence in our data set, SN IIn, SN IIP, SN II, and AGNs are the most likely objects to be misclassified as TDEs. The effectiveness of the proposed methodology in accurately classifying TDEs while managing the rate of false positives is demonstrated by these results. This approach is particularly valuable in TDE detection, where minimizing false negatives is crucial to ensuring these rare events are not missed. The potential of ensemble learning, combined with cost-sensitive analysis and threshold optimization, in handling data sets in astrophysical research is highlighted by the study, offering a robust tool for future TDE classifications. The proposed method could be particularly beneficial for upcoming large-scale surveys.

Key words: techniques: spectroscopic – software: machine learning – transients: tidal disruption events.

1 INTRODUCTION

Tidal disruption events (TDEs) take place when a star gets too close to a supermassive black hole (SMBH) and is consequently torn apart by the tidal forces imposed by the SMBH's gravity. When the gravitational pull of the SMBH on the orbiting star exceeds the self-gravity that holds the star together, the star is not suddenly swallowed by the black hole but is instead spaghettified, resulting in a spinning accretion disc of hot gas surrounding the black hole (e.g. Rees 1988; Gezari 2021). The formation of the accretion disc releases a large amount of energy, causing a radiative flare and emissions across the wavelength spectrum, from gamma-rays and X-rays to ultraviolet, optical light, and radio waves (Komossa 2015). These emissions finally make the TDE detectable with telescopes from Earth and space. As the SMBH consumes the material over time, the original bright flare gradually becomes fainter and fades away over several months or even years. By analysing the resulting emission spectra and light curve, it is possible to determine the properties of the black hole and the tidally disrupted star.

Since black holes do not emit electromagnetic radiation and thus cannot be observed directly, TDEs provide an important source for astronomers to indirectly study and explore their characteristics. Research suggests that SMBHs can be found in the centres of most, if not all, large galaxies (Ferrarese 2006). TDEs play a key role in galaxy evolution, and TDEs can therefore provide important insights into the population of stars and the environment at the centres of galaxies. Furthermore, TDEs can be used to study and analyse predictions made by general relativity, such as relativistic precession due to the curvature of space–time around massive objects, and gravitational lensing effects (Stone et al. 2019).

According to Gezari (2021), the cumulative number of discovered TDEs remained below 60 until 2020, with recent discoveries possibly bringing the total to just over 100 by 2024 (Hammerstein et al. 2023; Yao et al. 2023; Masterson et al. 2024). This illustrates the rarity of TDEs, which constitute only a small fraction of the events identified

★ E-mail: ariahump@ljmu.ac.uk

in large transient surveys such as the Public ESO Spectroscopic Survey of Transient Objects (PESSTO) or the Zwicky Transient Facility (ZTF). The spectra of TDEs exhibit a variety of emission lines, including hydrogen, helium, and Bowen fluorescence lines (Charalampopoulos et al. 2022). However, due to the small sample of reported TDEs and the wide range of spectral characteristics, there is an ongoing debate about whether these represent distinct subtypes of TDEs or if the observed variety arises from factors such as the viewing angle, or the properties of the black hole and the disrupted star.

To identify TDEs and distinguish them from other transients, such as supernovae, the light curve can be used. While the light curve of a TDE rises rapidly, similar to that of a supernova, it decays more gradually over a prolonged period, lasting months or even years. This characteristic shape of the light curve can be employed to identify TDEs. However, TDEs and supernova subtypes are normally identified and classified using their spectra. Although more costly, this method is much more efficient and rapid because spectra are directly observed by the transient survey, eliminating the need for prolonged follow-up observations to obtain a light curve. Automatic spectroscopic TDE classification could therefore be particularly beneficial for upcoming large-scale surveys.

The continuously increasing number of observations, driven by more and larger surveys, has also resulted in the need to automate the identification of transients, like supernovae, using quantitative methods. With the upcoming large-scale surveys such as the Vera C. Rubin Observatory's Legacy Survey of Space and Time (LSST), automated classification will be essential for the efficient allocation of follow-up resources. According to Bricman & Gomboc (2020), LSST is expected to discover between 35 000 and 80 000 TDEs in 10 yr of operation, far exceeding the capacity for manual classification.

During the 1990s, the first automated tools for feature identification in e.g. supernova spectra emerged. For instance, SYNOW (Jeffery & Branch 1990) is a highly parametrized spectrum synthesis code used for the empirical analysis of supernova spectra. Later, and still frequently used today, the cross-correlation-based identification software SNID (Blondin & Tonry 2007) was developed. It calculates the correlation between an input spectrum to be classified and a variety of template spectra representing different supernova types. By identifying the template with the highest correlation to the input spectrum, it yields the most likely supernova subtype. Similarly, Next Generation SuperFit (NGSF) is a template-matching software tool available on Weizmann Interactive Supernova Data Repository (WISeREP) for the spectral classification of major supernova types associated with host galaxies (Goldwasser et al. 2022), based on IDL Superfit (Howell et al. 2005). This tool employs chi-squared minimization to identify the optimal combination of host galaxy and supernova templates from a spectral library, effectively matching the spectrum of the target object.

However, in recent years, more powerful algorithms from data science have been employed for spectra classification. As an example, the Deep Automated Supernova and Host (DASH) classifier implements a deep CNN to train a matching algorithm (Muthukrishna, Parkinson & Tucker 2019b). According to the authors, this approach is orders of magnitude faster and more accurate than previous classification tools. Although a deep-learning-based method for the spectroscopic classification of Type Ia supernovae (SNIASCORE) has included TDEs in its large transient training set, it only performs binary classifications for Ia and non-Ia supernovae (Fremling et al. 2021). An automatic spectroscopic classification tool specifically for TDEs is therefore still lacking in the literature.

The objective of this article is to fill this research gap by developing a machine learning (ML) algorithm that can spectroscopically distinguish TDEs from supernovae, thereby enabling the automatic identification of TDEs in large spectroscopic surveys. The goal is to develop an algorithm that achieves high classification accuracy and effectively identifies most TDEs while maintaining a low false positive rate. Currently, TDE identification often relies on manual spectral inspection and light curve analysis, a process that is time-consuming and potentially subjective. Our approach aims to surpass this traditional method by providing a faster and automated solution.

Decision tree ensembles are applied as they are known for handling imbalanced data effectively, reducing overfitting, and improving generalization (Dietterich 2000 and Zhou 2012). The bagging approach enhances stability and accuracy while reducing variance (Sutton 2005). To prioritize the minority class, cost-sensitive learning assigns a higher penalty for TDE misclassification, thereby reducing false negatives. For efficient hyperparameter tuning that balances complexity and performance, a Bayesian optimization algorithm with five-fold cross-validation is employed (Stuke, Rinke & Todorović 2021). Finally, dynamic threshold selection adjusts the decision probability to maximize the number of correctly identified TDEs (Lipton, Elkan & Naryanaswamy 2014). Overall, this approach is expected to yield a robust and reliable model for effective TDE identification, with strong generalization to out-of-sample data and applicability in real-world scenarios.

The results of our analysis show that the decision tree ensembles, optimized with Bayesian hyperparameter tuning and dynamic threshold selection, achieve robust performance on our data set. With an 83 per cent TDE detection rate on unseen data, the model with dynamic threshold selection identifies most TDEs while maintaining a low false positive rate and experiencing only a small loss in accuracy compared to the performance without dynamic threshold selection. The algorithm demonstrates high reliability and effectiveness, offering a significant improvement in TDE detection for transient surveys.

2 LITERATURE REVIEW

TDEs were first theoretically predicted in the 1970s and 1980s. Building on Einstein's general theory of relativity (Einstein 1916), which describes gravity as the result of the curvature of space–time caused by mass and energy. Misner, Thorne & Wheeler (1973) theoretically explored the tidal forces around black holes and how these forces could stretch and compress nearby objects. The more detailed description of TDEs is linked to works by Hills (1975), Frank & Rees (1976), and Carter & Luminet (1982), who hypothesized that a star orbiting a black hole might be torn apart by the strong tidal forces if it comes too close. As a consequence, half of the star's mass is expected to be ejected at high velocity, while the other half would be accreted into a spinning disc of hot gas surrounding the black hole. The formation of the accretion disc then causes a radiative flare and observable emissions across the wavelength spectrum, ranging from gamma-rays and X-rays to ultraviolet, optical light, and radio waves (Komossa 2015). When TDEs were theoretically predicted in the 1970s, astrophysicists at the time doubted that such a TDE could ever be observed. However, the first TDEs were detected in archival searches of the soft X-ray *ROSAT* All-Sky Survey in the 1990s, and by 2020, a total of 56 TDEs had been discovered (Gezari 2021). With recent additions to the list, we may have surpassed a cumulative total of 100 reported TDEs in the literature by 2024 (Masterson et al. 2024). The occurrence of a TDE is estimated to happen only once every 10^4 to 10^5 yr per galaxy (Velzen & Farrar 2014; Holoien et al. 2016). However, with upcoming large-scale surveys such as LSST,

the number of TDE discoveries is expected to rise rapidly in the coming years.

The mass of the SMBH and the characteristics of the disrupted star mainly determine the properties of the TDE. The majority of stars involved in TDEs are main-sequence stars, although giant stars and white dwarfs can also be tidally disrupted by black holes (Wang, Perna & Armitage 2021). The mass, composition, and structure of the star influence the fallback rate of material, the accretion process, and the spectroscopic properties of the TDE (Komossa 2015). Whether the star is fully or partially disrupted primarily depends on the mass of the black hole and how close the star passes by it. This is because, as the mass of a black hole increases, its Schwarzschild radius grows faster than the tidal radius at which a star is tidally disrupted. Thus, a partial disruption might occur when a star passes within the tidal disruption radius of the black hole, but not close enough for a full disruption, and may therefore lose significant mass but survive the encounter. In contrast, a full disruption, where the entire material of the star is accreted, might occur with less massive SMBHs that have a tidal radius much larger than their Schwarzschild radius (Wang et al. 2021).

The light curve of TDEs is characterized by a rapid rise followed by a gradual decline over an extended period of time. The time-scale of the decay has been found to correlate with the mass of the host galaxy, while the decay rate correlates with the peak luminosity of the light curve (Hammerstein et al. 2023). In comparison to other transients like supernovae, TDEs typically exhibit a prolonged decay and often show bluer and brighter emission peaks (Kawana et al. 2020). This is due to disparities in the underlying physical processes, which may also explain why some TDEs exhibit differences in rise time and blackbody radius compared to supernovae.

Emission lines and spectral signatures associated with TDEs are observed across various wavelengths. In the optical spectrum, Balmer lines, along with He II and He I lines, are frequently detected, indicating the presence of high-velocity gas (van Velzen et al. 2021; Zhu et al. 2023). Bowen fluorescence is also observed in some TDEs, leading to enhanced emission of helium lines, which are often seen alongside hydrogen lines. Highly ionized lines from Fe VII and Fe X have also been observed in the aftermath of TDEs (Gezari et al. 2003). Soft X-ray flares originating from the inner regions of the accretion disc have been reported. Additionally, the ultraviolet spectrum of some TDEs, characterized by a blackbody continuum, is consistent with temperatures of approximately 10 000 to 30 000 K. In dust-obscured regions, infrared emissions have also been associated with certain TDEs.

The specific shape and peak of the light curve of TDEs, as previously discussed, can be used for their identification and to distinguish them from supernovae and Active Galactic Nuclei (AGN). While supernovae have more rapidly declining light curves compared to TDEs, AGNs typically exhibit more irregular variability due to their ongoing accretion processes (Chan et al. 2019). Also, soft X-ray and ultraviolet emissions are indicative for stellar material accretion and thus might point to TDEs. While AGNs are typically found in actively star-forming and massive elliptical galaxies, TDEs are primarily detected in quiescent Balmer-strong galaxies or post-starburst galaxies with distinctive star formation histories (French, Arcavi & Zabludoff 2016; Law-Smith et al. 2017; Graur et al. 2018). Although more commonly associated with AGNs, mid-infrared outbursts in dust-obscured environments might also indicate the presence of TDEs. However, many observational characteristics of TDEs are not unique, making their identification challenging. Since many TDEs may also be obscured by dust or intrinsically faint, especially at large distances, detections are biased toward particularly luminous events. Faint TDEs with more rapidly declining light curves, make follow up observations difficult and might even be overlooked. Moreover, TDEs have been primarily reported in rare post-starburst galaxies rather than in the more common star-forming galaxies, which may further bias the current sample of reported TDEs. Originally, TDEs were identified by the manual inspection of the light curve and spectrum. More recently, ML tools and simulation software has been suggested for TDE classification.

While early tools for classifying transients relied on similarity analyses, such as correlation analysis between templates and the spectrum to be classified, more recent approaches have suggested the use of ML algorithms for transient classification. Traditional statistical methods, such as correlations and t-tests, originated during a time when data collection and analysis were expensive, resulting in typically small data sets. These methods offer robust techniques for inference and hypothesis testing with small samples, allowing researchers to draw conclusions about a larger population. However, in the era of big data, where data acquisition costs are lower and data collection is often automated, large data sets have become common. ML algorithms are particularly powerful in this context, as they can leverage vast amounts of data to learn complex, generalizable patterns and achieve high predictive accuracy (Bzdok, Altman & Krzywinski 2018). Nevertheless, the increased predictive performance and complexity of ML models, compared to traditional statistical methods, often come at the cost of reduced interpretability and a higher risk of overfitting.

Convolutional Neural Networks (CNNs), a deep learning method widely used in computer vision and image analysis, have been employed for the image-based classification of transients, often to distinguish between real and false signals or to classify various types of supernovae. In the context of stamp classifiers, where a small image cut-out (stamp) of a new astronomical observation is compared to a previous image of the same area of the sky to detect transients, CNNs have been used by researchers such as Carrasco-Davis et al. (2021) and Reyes-Jainaga et al. (2023). Additionally, CNNs have been applied to astronomical image time series, where images are taken over time of the same field of view to observe spatio-temporal evolution (Bairouk et al. 2023). When images of objects are compared over time, CNNs have also been used for transient classification based on the resulting light curves (Förster et al. 2021). Furthermore, CNNs have been employed in spectral data analysis for transient classification. For instance, Muthukrishna et al. (2019b) developed the DASH classifier, which uses spectroscopic data for supernova classification. Overall, CNNs have been reported in the literature as effective in rapidly and accurately classifying astronomical transients across various data types, including image stamps, image time series, and spectroscopic data.

In contrast to CNNs, which are well suited for pattern recognition in spatial data like images, recurrent neural networks (RNNs) are typically used for sequential data to capture temporal dependencies. Consequently, RNNs have been used for transient classification based on light curves. For instance, Muthukrishna et al. (2019a) developed the Real-time Automated Photometric IDentification (RAPID) tool, which utilizes astronomical time-series data to classify 12 different transient classes. More recently, Gagliano et al. (2023) implemented a neural network with a single recurrent layer to leverage host galaxy photometry and early light curve data for supernova classification. They reported an accuracy of 82 per cent $\pm$ 2 per cent within three days of event discovery and 87 per cent $\pm$ 5 per cent within 30 d. Additionally, Fremling et al. (2021) applied an RNN to low-resolution spectroscopic data for the binary classification of Type Ia

and non-Ia supernovae, achieving 90 per cent accuracy in correctly classifying SN Ia spectra. In general, RNNs have been particularly effective in handling time-series data, allowing early and accurate transient classification.

Although neural network-based methods are powerful in handling complex spatial and temporal data, tree-based methods like random forests and boosted decision trees also offer robust alternatives for classification tasks. For instance, Goldstein et al. (2015) applied random forests to identify various point-source transients in reference-subtracted optical images. The algorithm was trained using data from the Dark Energy Survey Supernova programme and supplemented with artificial Type Ia supernovae. According to the authors, the random forest achieved high performance in detecting Type Ia supernovae, with a low false positive rate of only 2.5 per cent. Similarly, Santos, Quartin & Reis (2020) and Villar et al. (2020) used decision tree ensembles and random forests, respectively, to classify supernovae based on photometric light curves. Villar et al. (2020) reported an overall accuracy of 87 per cent across five different supernova classes. Boosted decision trees have also been successfully employed for photometric supernova classification by Lochner et al. (2016), further supporting the effectiveness of tree-based classifiers for transient classification. Finally, Richards et al. (2011) used random forest classification on supernova light curves in a semi-supervised learning setting, reporting high performance with a Type Ia supernova purity of 95 per cent. In summary, tree-based algorithms have proven to be highly effective and reliable for supernova classification, delivering strong results across various studies.

Some authors have also applied support vector machines (SVMs) for supernova classification. The basic idea of SVMs is to find the optimal hyperplane that best separates different classes in the feature space. SVMs can be particularly powerful in high-dimensional spaces when there are clear margins between classes. For instance, Bailey et al. (2007) and Lochner et al. (2016) employed SVMs for supernova classification tasks. While Lochner et al. (2016) applied the method to light curve data, Bailey et al. (2007) used difference images for their analysis. Although both studies generally support the usefulness of SVMs for supernova classification, they conclude that neural network or decision tree-based algorithms tend to perform better overall. In conclusion, while SVMs show potential for supernova classification, neural networks and decision tree-based algorithms generally outperform them in terms of overall accuracy and performance.

While neural networks, decision tree-based methods, and SVMs rely on data-driven pattern recognition, a fundamentally different approach is based on Bayesian methods, which rely on probabilistic inference and the integration of prior knowledge for classification tasks. Poznanski, Maoz & Gal-Yam (2007) implemented an automatic Bayesian classification algorithm for supernovae, based on single-epoch multiband photometry and host galaxy redshift information, to sort supernova candidates into Type Ia and core-collapse supernovae. The approach is based on a prior redshift estimate and template fitting to maximize the likelihood of the object being a Type Ia or core-collapse supernova. They conclude that the Bayesian methodology can correctly classify up to 97 per cent of the data.

For automatic TDE classification, various software tools have been developed that use light-curve information to select TDEs in transient data sets. For instance, Stein et al. (2024) introduced TDESCORE, a binary photometric classifier for TDEs. They used approximately 3000 nuclear transients from ZTF, of which about 2 per cent are TDEs, to train a ML algorithm for binary TDE classification. As input for the gradient-boosted decision tree (XGBoost), three-dimensional information from flux, wavelengths, and time is used. From the light-curve data, parameters such as WISE W1–W2 host colour, colour at *g*-band peak, and rate of colour change are extracted and fed to the algorithm. Likewise, Gomez et al. (2023) developed a random forest based model for TDE classification based on light curve and host galaxy information. It is an expansion of the Finding Luminous and Exotic Extragalactic Transients (FLEET) machine-learning algorithm for identifying hydrogen-poor superluminous supernovae. Although, the original FLEET algorithm classifies various supernova types, it can be adjusted for the binary classification of TDEs. Finally, the Automatic Learning for the Rapid Classification of Events broker has extended its light-curve classifier to include TDE classification (Sánchez-Sáez et al. 2021). It applies a random forest algorithm that is fed with variability features computed from the ZTF alert stream and colours obtained from AllWISE and ZTF photometry. Overall, studies on automatic TDE classification based on light-curve features and host galaxy information conclude that automatic TDE detection in large transient surveys is feasible and can be successfully implemented. However, although photometry is inexpensive and abundantly available, the use of light-curve data relies on multiple observations over time, and rapid identification at an early stage, for example, when the tidal disruption begins, is not possible.

The literature on automatic transient classification reveals a lack of studies focusing on spectroscopic TDE detection. Although Fremling et al. (2021) included 0.4 per cent TDEs in their data set, they only performed binary classification for Ia and non-Ia supernovae. To address this gap, we have developed a ML-based algorithm for binary TDE and non-TDE classification based on spectroscopy. With upcoming large-scale surveys such as LSST, thousands of new TDE discoveries are expected, making automatic classification essential for managing the data volume and rapidly deciding on follow-up resources. According to the literature, neural network and tree-based methodologies are the most frequently used algorithms for transient classification tasks, demonstrating high performance. However, as demonstrated by earlier TDE classification tools based on light-curve information, decision tree ensembles may be better suited for handling imbalanced data due to their flexibility in class weighting, sampling techniques, and their ability to focus on harder-to-classify instances from the minority class.

3 DATA

For our data analysis, we make use of spectroscopic data from PESSTO SSDR1-4. The PESSTO is a programme that systematically obtains spectroscopic data for supernovae and optical transients with a magnitude limit for classification of 20.5 in the *B*, *V*, and *R* bands. The first four data releases (SSDR1-4) include a large collection of spectroscopic observations of various transient events, including supernovae, TDEs, and AGNs (Smartt et al. 2015). The data used in this study were downloaded from the WISeREP, a data base that archives spectral and photometric data for supernovae and other transients (Yaron & Gal-Yam 2012).

PESSTO SSDR1-4 observations were acquired using the European Southern Observatory's New Technology Telescope (ESO-NTT), located at the La Silla Observatory in Chile. It is a 3.58-m telescope with an altitude-azimuth mount. The primary instrument employed for optical spectroscopy and imaging is the ESO Faint Object Spectrograph and Camera version 2 (EFOSC2-NTT). EFOSC2-NTT is a focal reducer with a focal length of 200 mm that facilitates the acquisition of high-quality spectral data across a broad wavelength range (Smartt et al. 2015).

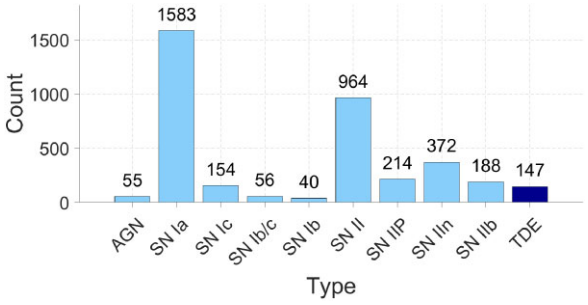

Figure 1. Distribution of Supernovae, AGNs, and TDEs in data set.

The selection of transients for PESSTO SSDR1-4 is based on real-time alerts from transient surveys, such as ZTF and the Asteroid Terrestrial-impact Last Alert System. Upon detection, candidate transients are prioritized based on factors such as brightness, proximity to host galaxies, and scientific interest. The selected transients are then spectroscopically followed up using EFOSC2-NTT. The spectroscopic data are reduced and analysed to identify characteristic spectral features that facilitate accurate classification. After the reduction of raw spectroscopic data by the PESSTO reduction and analysis team, automated transient classification tools such as SNID (Blondin & Tonry 2007), GELATO (Harutyunyan et al. 2008), and SUPERFIT (Howell et al. 2005) are used. Although automated classification provides rapid results, PESSTO experts verify these classifications to confirm their validity.

As of 2024 December 22, the downloaded data set includes a total of 147 TDEs in accordance with PESSTO classification. In addition to the TDE data, spectra for various supernovae (types II, IIP, IIb, IIn, Ia, Ib, Ib/c, and Ic), and AGNs were also downloaded for our data set. Overall, 3773 spectra were included in the raw data set before pre-processing. With only 147 TDEs, the proportion of TDEs in our data set, as illustrated in Fig. 1, is consistent with the rarity of TDE detections in transient surveys compared to supernovae (Perley et al. 2020). Although there are other sources for TDEs, we deliberately did not include them in our data set to maintain comparability, as different telescopes may have specific characteristics that could distort the analysis.

3.1 Pre-processing of the data

In line with earlier studies for supernova classification (Davison, Parkinson & Tucker (2022), Muthukrishna et al. (2019a), Blondin & Tonry (2007)) we pre-process the spectral data before data analysis. As the raw spectra do not have a common wavelength grid, the raw flux data must first be binned to a common grid between 3500 and 9000 angstroms. The grid is then used as a reference for interpolating the flux data to a uniform wavelength scale. The flux data are interpolated on to the common wavelength grid using linear interpolation.

Next, we apply a median filter with a window size of 6 bins to the spectra data. A median filter is a non-linear technique for noise reduction in the data. In contrast to linear filters, it is robust to outliers but preserves edges e.g. sharp transitions in the spectra. Further, the spectral data are adjusted to their rest frame. This de-redshifting is necessary to account for the expansion of the universe and the resulting stretching of emitted light, which causes a shift toward the red end of the spectrum. Next, for continuum removal, a 13-point cubic spline interpolation was applied as described in DASH (Muthukrishna et al. 2019b). This method models the continuum of the transient spectra, after which the transient spectra are divided by the continuum to obtain the relative strengths of spectral features. Finally, the flux data are normalized between zero and one and log-binned to a common grid with 1024 spectral bins, each representing a feature. Fig. 2 shows the median spectra for all supernovae subtypes, AGNs, and TDEs before continuum removal.

4 METHODOLOGY

The data set after pre-processing now consists of 3773 labelled spectra with 1024 features, each spectrum having a binary label of either TDE or non-TDE. To maintain the original distribution of the data set between TDE and non-TDE, a stratified sample consisting of 75 per cent of the data are drawn. This sample is used for training and validation of the decision tree, whereas the remaining 25 per cent is held back for out-of-sample testing.

For the classification task we are using decision tree ensembles. A decision tree is a non-parametric supervised learning method that recursively splits data into subsets based on predictor values, aiming to increase the purity of the subsets until a final classification is reached at the leaf nodes. At each node, the algorithm selects the feature and threshold that best splits the data by maximizing the reduction in impurity. The impurity measure used here is the *Gini impurity*G, defined as:

$$G = 1 - \sum_{i=1}^{C} p_i^2, \tag{1}$$

where p_i is the the probability that a randomly selected sample belongs to class i at the node, and C is the number of classes. Specifically, there are two classes, TDEs and non-TDEs; thus, $C = 2$. Therefore, the Gini impurity formula simplifies to:

$$G = 1 - (p_{\text{TDE}}^2 + p_{\text{non-TDE}}^2), \tag{2}$$

where p_{TDE} is the proportion of TDE samples at the node and $p_{\text{non-TDE}}$ is the proportion of non-TDE samples at the node. The decision tree algorithm chooses the split that results in the largest decrease in impurity, calculated as:

$$\Delta G = G_{\text{parent}} - \sum_{k=1}^{K} \left(\frac{N_k}{N_{\text{parent}}} G_k \right), \tag{3}$$

where G_{parent} is the Gini impurity of the parent node; G_k is the Gini impurity of child node k; N_{parent} is the number of samples in the parent node; N_k is the number of samples in child node k; and K is the number of child nodes resulting from the split. With continuous independent variables, such as measures of flux at different wavelengths, the algorithm splits each parent node into two child nodes; therefore, $K = 2$. All possible split points (thresholds) for all variables are evaluated to determine the best binary split that maximizes the decrease in impurity. The splitting process is applied recursively to each node to build a hierarchical tree structure, with subsequent splits that further reduce impurity by selecting the most suitable variables and thresholds.

To reduce the risk of overfitting and increase accuracy, an ensemble model that combines multiple decision trees is incorporated. The bootstrap aggregation, or bagging, is used as an ensemble method to find a robust model with multiple decision trees. Bagging uses bootstrapping to generate multiple subsets of the original data set. Each subset is a bootstrap sample where samples are randomly selected with replacement. Thus, in some subsets, certain samples

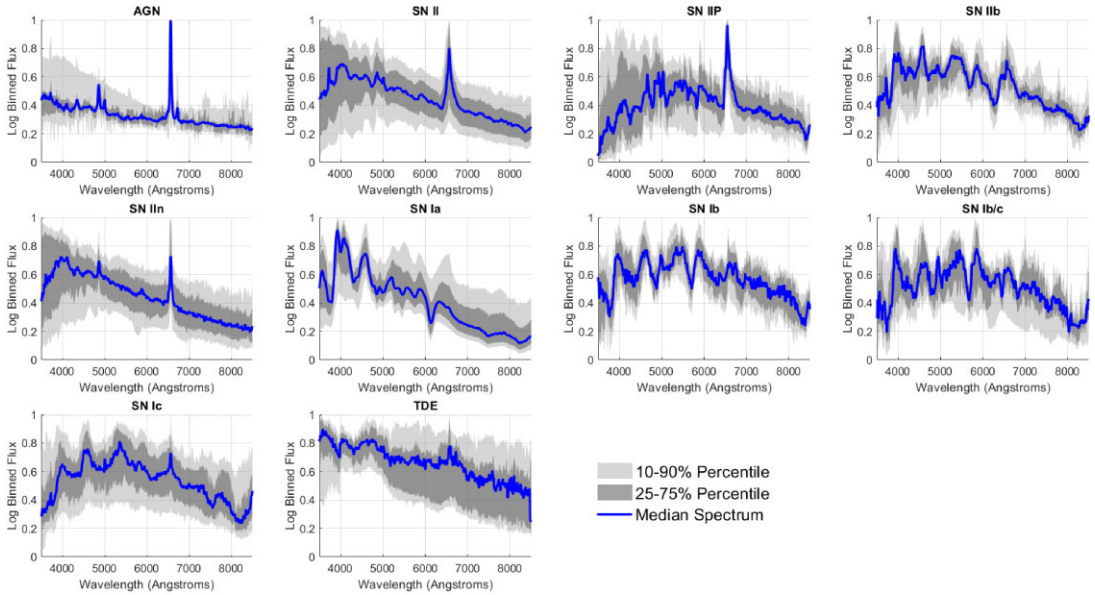

Figure 2. Median spectra for all supernova subtypes, AGNs, and TDEs before continuum removal.

might appear multiple times, whereas they are missing in others. These different bootstrap samples are then used to train multiple decision trees. This approach leads to higher predictability and better generalization abilities.

Each decision tree in the ensemble outputs a posterior probability $p_t(y = 1 \mid \mathbf{x})$ for the positive class. The ensemble aggregates these probabilities by averaging them across all trees:

$$p_{\text{ensemble}}(y = 1 \mid \mathbf{x}) = \frac{1}{T} \sum_{t=1}^{T} p_t(y = 1 \mid \mathbf{x}), \tag{4}$$

where T is the total number of trees in the ensemble.

Due to the imbalanced data set, in which TDEs constitute only 3.9 per cent, we implemented a cost-sensitive learning approach using a cost matrix that reflects the imbalance ratio between TDEs and other transients, thereby appropriately penalizing misclassifications of the minority class:

$$\text{Imbalance Ratio} = \frac{\text{Number of Majority Class}}{\text{Number of Minority Class}} = \frac{3626}{147} \approx 25. \tag{5}$$

The cost matrix $\mathbf{C}$ is then defined as:

$$\mathbf{C} = \begin{bmatrix} C_{00} & C_{01} \\ C_{10} & C_{11} \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ 25 & 0 \end{bmatrix}, \tag{6}$$

where C_{ij} represents, e.g., the cost of predicting class i when the true class is j. In the matrix, the cost for misclassifying a non-TDE as a TDE is set at 1, whereas the cost for misclassifying a TDE as a non-TDE is set at 25. As the TDE class is heavily underrepresented in the data set, it is important to penalize the misclassification of a TDE significantly more than that of a non-TDE. For each instance $\mathbf{x}$, the expected misclassification cost for each class is computed. The expected cost of predicting class 0 (non-TDE) is:

$$\text{Cost}(0 \mid \mathbf{x}) = 1 \times p(y = 1 \mid \mathbf{x}) \tag{7}$$

and the expected cost of predicting class 1 (TDE) is:

$$\text{Cost}(1 \mid \mathbf{x}) = 25 \times (1 - p(y = 1 \mid \mathbf{x})). \tag{8}$$

By selecting the class with the lower expected cost, the model prioritizes correctly identifying TDEs, thereby improving its performance on the underrepresented class. The class with the minimum expected cost is selected as follows:

$$\hat{y}(\mathbf{x}) = \arg \min_{k \in \{0,1\}} \text{Cost}(k \mid \mathbf{x}). \tag{9}$$

However, this approach also has the drawback of accepting a higher rate of false positives. To further increase the performance on out-of-sample data, Bayesian optimization with five-fold cross-validation is used for hyperparameter tuning. The number of decision trees in the model is optimized to balance accuracy and complexity while avoiding overfitting. The objective of the Bayesian optimization is to to find the optimal hyperparameter vector $\boldsymbol{\lambda} = (T, L, S)$ that minimizes the cross-validation loss $f(\boldsymbol{\lambda})$. The vector $\boldsymbol{\lambda}$ consists of the number of decision trees T, the minimum leaf size L, and the maximum number of splits S, which are tuned simultaneously to balance accuracy and complexity while avoiding overfitting. The minimum leaf size L controls the bias–variance trade-off, and the maximum number of splits S influences the depth of the trees. The search space of λ is defined by the Cartesian product Λ of all allowable ranges for each hyperparameter:

$$\Lambda = \Lambda_T \times \Lambda_L \times \Lambda_S. \tag{10}$$

After n evaluations, we have a set of observations:

$$D_{1:n} = \{(\lambda_i, f(\lambda_i))\}_{i=1}^{n}, \tag{11}$$

where $D_{1:n}$ represents the set of pairs consisting of hyperparameter vectors and their corresponding function values (e.g. the cross-validation losses). Thus, λ_i is the vector of hyperparameters at iteration i and $f(\lambda_i)$ is the corresponding function value (average cross-validation loss) evaluated at λ_i. Bayesian optimization models $f(\boldsymbol{\lambda})$ as a random function using a probabilistic surrogate model in form of a Gaussian Process. At a new hyperparameter combination λ^*, the Gaussian Process provides a posterior predictive distribution for $f(\lambda^*)$ with mean $\mu(\lambda^*)$ and variance $\sigma^2(\lambda^*)$.

An acquisition function $a(\lambda)$ is used that balances exploration (searching unexplored regions with high uncertainty in the surrogate

model) and exploitation (searching near promising hyperparameters). It directs the selection of new hyperparameter combinations by evaluating how each combinations might lead to an improvement over the current best solution. The acquisition function therefore evaluates the expected improvement (denoted as $\Gamma(\lambda)$) over the current best observed loss $f_{\text{best}} = \min_{1\leq i\leq n} f(\lambda_i)$. We get:

$$a(\lambda) := \Gamma(\lambda) = \mathbb{E}[\max(0,\ f_{\text{best}} - f(\lambda))]. \tag{12}$$

Since, the posterior predictive distribution is typically assumed to be normal, $\mathcal{N}\big(\mu(\boldsymbol{\lambda}_*),\ \sigma^2(\boldsymbol{\lambda}_*)\big)$, the closed-form of expected improvement supports sampling in regions likely to improve over the current best loss while taking uncertainty into account. At iteration $n+1$, the next hyperparameter vector λ_{n+1} is chosen by maximizing the acquisition function:

$$\lambda_{n+1} = \arg\max_{\lambda\in\Lambda} a\big(\lambda \mid D_{1:n}\big). \tag{13}$$

Once λ_{n+1} is selected, we run five-fold cross-validation again to obtain $f(\lambda_{n+1})$ and update:

$$D_{1:n+1} = D_{1:n} \cup \{(\lambda_{n+1},\ f(\lambda_{n+1}))\}. \tag{14}$$

In our model, we performed 25 iterations (objective evaluations) for the Bayesian optimization procedure to ensure a thorough but computationally feasible hyperparameter search.

The models are then evaluated based on accuracy, precision, recall, and the F1 score. Accuracy is the ratio of correct classifications (true positives and true negatives) to the total number of cases, calculated as:

$$\text{Accuracy} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{TN} + \text{FP} + \text{FN}}. \tag{15}$$

In this equation, TP (true positives) refers to the number of positive cases that are correctly classified by the model. Likewise, TN (true negatives) represents the number of negative cases that are correctly classified. The term FP (false positives) denotes the number of negative cases incorrectly classified as positive, while FN (false negatives) is the number of positive cases that the model incorrectly classifies as negative. While accuracy indicates the proportion of correct classifications, it can be misleading when dealing with highly unbalanced data. For instance, in our data set, where only 3.9 per cent of cases are TDEs, a model that classifies all instances as non-TDEs would still achieve an accuracy of 0.961, or 96.1 per cent.

In contrast, precision is the proportion of true positives among all positive classifications (i.e. true positives and false positives). In our case, precision indicates how many of the predicted TDEs were actually TDEs, and it is calculated as:

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}}. \tag{16}$$

Recall is computed as the number of true positives divided by the number of true positives plus false negatives. It therefore measures the true positive rate, that is, the proportion of actual TDEs correctly identified by the model. Recall is calculated as:

$$\text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}}. \tag{17}$$

Finally, the F1 score combines recall and precision to provide a balance between these two metrics. It is calculated as the harmonic mean of precision and recall:

$$\text{F1 Score} = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}}. \tag{18}$$

Given the objective of maximizing the identification of TDEs while minimizing false positives, all these metrics are relevant.

Table 1. Optimized parameter values.

Parameter	Tested range	Optimal value found
Number of trees	10–500	146
Minimum leaf size	1–50	42
Max number of splits	10–2000	1954
Probability threshold	0.1–0.9	0.1

A dynamic threshold selection process is implemented to find a good balance between precision and recall. Different thresholds for the predicted probability of classifying a spectrum as a TDE are tested. The ensemble model provides the probabilities of an individual spectrum being classified as a TDE. In the dynamic threshold selection, a range between 0.1 and 0.9 with 100 intervals is examined. For each threshold, the recall, precision, and F1 score are calculated. The optimal threshold is then set at the value with the highest F1 score.

The final model is then tested with the 25 per cent of held-back data that has not been seen by the model before. To evaluate the performance of the model on out-of-sample data, accuracy, precision, recall, and the F1 score is calculated again and compared to the results of the model validation results.

5 RESULTS

The results of the Bayesian optimization for hyperparameter tuning of the decision tree ensembles indicate an optimal number of trees of 146, a minimum leaf size of 42, and a maximum number of splits at 1954 (Table 1). With a tested search range between 10 and 500 for the number of trees, the optimal value of 146 lies towards the lower part of the range, suggesting a balance between too few and too many trees. The optimal minimum leaf size of 42 is near the upper end of the tested range of 1 to 50, indicating that larger leaves might be sufficient for good decision making. The optimal maximum number of splits, at 1954, is toward the upper end of the range of 10 to 2000, suggesting that a higher level of detail and depth in the decision trees is required.

The results of the dynamic threshold selection process for optimizing the F1 score yielded a threshold value of 0.1, at the lower end of the range between 0.1 and 0.9. This suggests that balancing recall and precision, where both false positives and false negatives have significant implications, requires a low probability threshold (Table 1).

The in-sample data set can be divided into a training set and a validation set during the cross-validation model tuning process. Table 2 presents the average results for training and validation during cross-validation. While accuracy remains relatively stable at 99.98 per cent for the training set and 97.88 per cent for the validation set, the precision, recall, and F1 score decrease in the validation set compared to the training set. For instance, the F1 score drops from 0.998 in the training set to 0.674 in the validation set. However, the recall of 0.559 indicates that more than half of the TDEs in the validation set are successfully identified. Given the highly unbalanced nature of the data set, the performance of the algorithm appears satisfactory (for the in-sample confusion matrix see upper row of Fig. 3). The confusion matrix shows the number of predictions by actual class labels (as found in the data set) and predicted class labels (as assigned by our model). The rows indicate the actual class, with categories for TDEs and non-TDEs, while the columns indicate the predicted class as either TDE or non-TDE. Correct predictions are shown along the main diagonal, running from the top left to the

Table 2. In- and out-of-sample (held back data) evaluation metrics.

Metric	Accuracy	Precision	Recall	F1 score	Confusion matrix
Training set	0.9998	0.9955	1.0000	0.9978	TN: 2174.8, FP: 0.4, FN: 0.0, TP: 88.8
Validation set	0.9788	0.8493	0.5586	0.6739	TN: 541.6, FP: 2.2, FN: 9.8, TP: 12.4
Out-of-sample (test set)	0.9767	0.8500	0.4722	0.6071	TN: 904.0, FP: 3.0, FN: 19.0, TP: 17.0
Out-of-sample (threshold optimized)	0.8823	0.2222	0.8333	0.3509	TN: 802.0, FP: 105.0, FN: 6.0, TP: 30.0

bottom right, where true positives and true negatives appear. Incorrect predictions, such as false positives and false negatives, are shown in the off-diagonal cells.

In the out-of-sample (held back data) application of the optimized model, the evaluation metrics remain fairly stable compared to the in-sample results. While accuracy and precision are consistent between in-sample and out-of-sample data sets, recall and the F1 score decrease by 9 per cent and 7 per cent, respectively. However, with dynamic threshold optimization, the recall increases to 0.83, successfully identifying most TDEs out-of-sample. Although the false positive rate also increases, overall, 83 per cent of TDEs could be identified by examining only 15 per cent of the entire data set. Given that only 3.9 per cent of the sample consists of TDEs, these out-of-sample results can be considered a positive outcome (for the out-of-sample confusion matrix see lower row of Fig. 3). Thus, depending on the primary objective, dynamic threshold optimization yields higher recall but at the cost of more false positives, whereas the model without dynamic thresholding results in fewer false positives but at the cost of lower recall. However, considering both in-sample and out-of-sample results, the model appears stable.

Table 3 shows that, relative to their occurrence in the data set, the objects most frequently misclassified as TDEs are SN IIn (17.5 per cent), SN IIP (17.5 per cent), SN II (14.4 per cent), and AGNs (14.3 per cent). Similar to TDEs, these objects often exhibit prominent H α emission features around 6500 Å, as illustrated by the median spectra in Fig. 2. This spectral overlap may lead the algorithm to confuse them with TDEs when classification is based on a single spectrum.

When observing the feature importance plot (Fig. 4), the first range of high importance features lies between 3900 and 4000 angstroms. In this region, two absorption lines of single ionized calcium (Ca II) are found at 3933 angstroms (K line) and 3968 angstroms (H line). In particular, SN II and AGNs frequently show absorption around these lines, whereas TDEs might typically display weaker absorption in the Ca II H and K region or a different appearance due to their high blackbody temperature (Hammerstein et al. 2023). Consequently, the algorithm likely distinguishes TDEs from SN II and AGNs using these features.

Additionally, the Ca II triplet lines at 8498, 8542, and 8662 angstroms are characteristic of SN II in the near-infrared, once strong emission has developed in this region (Andronova 1990). Similarly, Ca II triplet lines have also been associated with AGNs (Vargas et al. 1993). In contrast, TDEs may show weaker or absent emission here, helping the algorithm separate SNe II from TDEs.

Another cluster of high importance features appears around 5900 angstroms, where Na I D and He I (5876) are found. TDEs are often helium-rich, exhibiting lines such as He II (4686) and He I (5876), whose appearance and strength may further distinguish them from other transients in the data set (Zabludoff et al. 2021).

Finally, Balmer lines at H α (6563) and H β (4861) also emerge among the high importance features. TDEs often show extremely broad, smooth Balmer emission (Zabludoff et al. 2021), which the algorithm likely detects. With these features, the algorithm might separate TDEs from Type I supernovae.

Overall, the high importance features in the Ca II H and K region, the Ca II triplet, and strong He and H lines appear to be key in differentiating TDEs from supernovae and AGNs.

6 DISCUSSION

The results of this study demonstrate that the implemented decision tree ensemble method, enhanced with Bayesian hyperparameter tuning and dynamic threshold selection, provides a robust tool for TDE classification. The algorithm achieves a recall of 83 per cent on the out-of-sample data set, successfully identifying most TDEs while maintaining a low false positive rate. This approach offers a significantly faster and more automated solution to TDE identification compared to traditional light curve analysis, advancing current methods. Cost-sensitive learning was essential for addressing the highly imbalanced data set, while dynamic threshold optimization helped maximize recall with low false positives. The findings fill a gap in the literature on automatic TDE classification using ML, as no previous study has analysed automated spectral classification of TDEs. This contribution is timely, given the anticipated increase in TDE discoveries from upcoming surveys like LSST. As large data sets with transients become more common, the need for automated and reliable classification methods will be crucial for effective resource allocation for follow-up observations.

Compared to previous methods for transient classification, such as the cross-correlation-based SNID (Blondin & Tonry 2007) or the CNN-based DASH (Muthukrishna et al. 2019b), this study implements a novel approach specifically focused on the binary spectroscopic classification of TDEs and non-TDEs. While previous studies have successfully applied ML algorithms to supernova classification, they have either excluded TDEs or included them as a minor part of a non-Ia supernova subclass (Fremling et al. 2021). This study advances the subject by developing a dedicated TDE classifier, which can distinguish between rare TDEs and non-TDE with high accuracy and

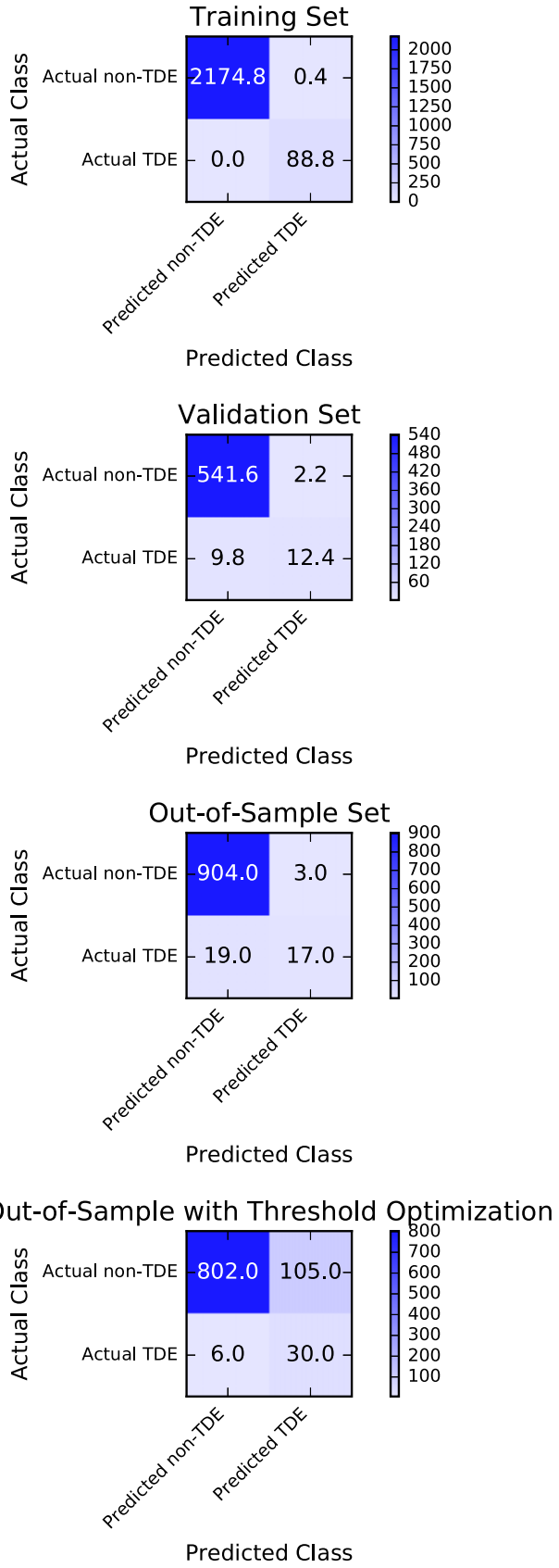

Figure 3. Confusion matrices for cross-validation and test set.

Table 3. Misclassification results.

Type	Misclassified	Total in test data set	Per cent misclassified
AGN	2	14	14.3 per cent
SN Ia	37	380	9.7 per cent
SN Ib	0	11	0.0 per cent
SN Ib/c	3	21	14.3 per cent
SN Ic	1	47	2.1 per cent
SN II	33	229	14.4 per cent
SN IIb	2	51	3.9 per cent
SN IIn	17	97	17.5 per cent
SN IIP	10	57	17.5 per cent
Total	105	907	11.6 per cent

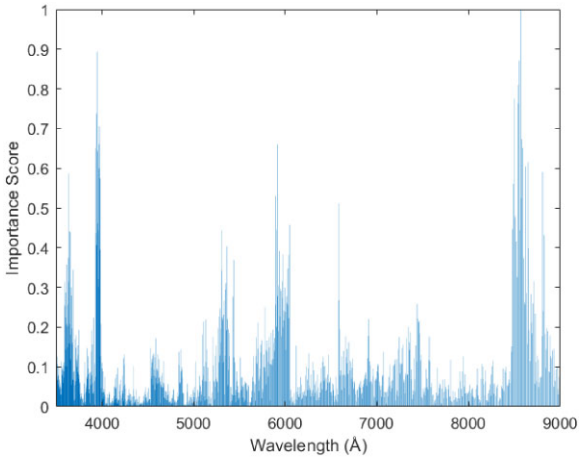

Figure 4. Importance scores by wavelength.

recall. Although CNNs have been successfully used in classifying transient data, decision tree ensembles might offer advantages in handling imbalanced data sets, as supported by our results. This aligns with earlier studies, such as Goldstein et al. (2015) and Villar et al. (2020), which highlight the flexibility of tree-based methods in focusing on underrepresented classes within imbalanced data sets, such as TDEs.

The results of our study also have theoretical implications. They demonstrate that decision tree ensembles with Bayesian hyperparameter tuning and dynamic threshold selection can effectively classify rare astronomical transients like TDEs spectroscopically. This suggests that tree-based methods may also be useful in similar astrophysical contexts where data imbalances cause challenges. Furthermore, the model's ability to generalize well to out-of-sample data highlights the potential of ML algorithms to identify patterns in complex astronomical data sets, a task that is exceedingly difficult with traditional manual inspection techniques. The results also indicate that TDEs can be automatically identified spectroscopically by distinct spectral features compared to supernovae and AGNs.

The practical implications of our results are important for large-scale transient surveys, where thousands of new TDEs might be detected annually. Automatic TDE identification will advance the classification process and enable rapid decision-making on follow-up resources. This approach could be integrated into real-time pipelines for surveys like LSST to offer near real-time classification of incoming data.

Despite the strong performance of our algorithm, there are limitations that should be acknowledged. One limitation is the relatively high false positive rate in the out-of-sample data, which is acceptable given the rarity of TDEs and our objective, but could still lead to unnecessary follow-up observations in practice. In particular, the model with dynamic threshold optimization shows a high recall but at the cost of higher false positives. Depending on the objective, this trade-off between recall and false positives is crucial. Either 53 per cent of TDEs are missed with the non-threshold-optimized model, or 78 per cent are false positives with the threshold-optimized model. However, identifying most TDEs with minimal false positives is not possible. With more TDE discoveries in the future, we expect the algorithm's performance to improve substantially, and future research should investigate this further. While the decision tree ensemble was selected for its ability to handle imbalanced data, other methods like RNNs might also be tested. RNNs could leverage time-series data from light curves, potentially enhancing classification accuracy further by combining spectral and temporal information. Finally, although our approach performs well on the data set used, it remains to be seen how it will generalize to new survey data with different characteristics. Future research could test the method on data from upcoming large-scale surveys with richer data sets.

7 CONCLUSION

This study presents a robust and reliable ML model for automatic classification of TDEs. This is a step forward in the automation of rare transient spectroscopic identification. By implementing decision tree ensembles optimized with Bayesian hyperparameter tuning and dynamic threshold selection, the model achieves high recall rates, successfully identifying most TDEs while maintaining a low false positive rate. The implications of this work offer a practical solution for managing the large volumes of data expected from upcoming large-scale surveys, while also contributing to the broader understanding of TDEs and how they are spectroscopically distinguished from, e.g. supernovae. Future research should focus on further refining the model, potentially integrating additional data types, and testing its applicability on larger and more diverse data sets.

ACKNOWLEDGEMENTS

This research did not receive any funding. During the preparation of this work the authors used CHATGPT to check grammar, readability, and spelling. After using this tool, the authors reviewed and edited the content as needed and take full responsibility for the content of the published article. PYTHON and MATLAB were employed to perform the analysis. We would like to thank the anonymous referee for the valuable comments and suggestions, which have contributed to improving the quality of this manuscript.

DATA AVAILABILITY

The authors do not have permission to share data.

REFERENCES

Andronova A. A., 1990, Astrophysics, 32, 235
Bailey S., Aragon C., Romano R., Thomas R. C., Weaver B. A., Wong D., 2007, ApJ, 665, 1246
Bairouk A., Chaumont M., Fouchez D., Paquet J., Comby F., Bautista J., 2023, A&A, 673, A141
Blondin S., Tonry J. L., 2007, ApJ, 666, 1024
Bricman K., Gomboc A., 2020, ApJ, 890, 73
Bzdok D., Altman N., Krzywinski M., 2018, Nat. Methods, 15, 233
Carrasco-Davis R. et al., 2021, AJ, 162, 231
Carter B., Luminet J., 1982, Nature, 296, 211
Chan C.-H., Piran T., Krolik J. H., Saban D., 2019, ApJ, 881, 113
Charalampopoulos P. et al., 2022, A&A, 659, A34
Davison W., Parkinson W., Tucker B. E., 2022, ApJ, 925, 186
Dietterich T. G., 2000, Lecture Notes in Computer Science, Vol. 1857, Multiple Classifier Systems, MCS 2000. Springer, Berlin, Heidelberg, p. 1
Einstein A., 1916, Ann. Phys., Lpz., 354, 769
Ferrarese L., 2006, Proc. Int. Astron. Union, 1, S232
Frank J., Rees M. J., 1976, MNRAS, 176, 633
Fremling C. et al., 2021, ApJ, 917, L2
French K. D., Arcavi I., Zabludoff A., 2016, ApJ, 818, L21
Förster F. et al., 2021, AJ, 161, 242
Gagliano A., Contardo G., Foreman-Mackey D., Malz A. I., Aleo P. D., 2023, ApJ, 954, 6
Gezari S., 2021, ARA&A, 59, 21
Gezari S., Halpern J. P., Komossa S., Grupe D., Leighly K. M., 2003, ApJ, 592, 375553
Goldstein D. A. et al., 2015, AJ, 150, 82
Goldwasser S., Yaron O., Sass A., Irani I., Gal-Yam A., Howell D. A., 2022, Trans. Name Server AstroNote, 191, 1
Gomez S., Villar V. A., Berger E., Gezari S., van Velzen S., Nicholl M., Blanchard P. K., Alexander K. D., 2023, ApJ, 949, 113
Graur O., Decker French K., Zahid H. J., Guillochon J., Mandel K. S., Auchettl K., Zabludoff A. I., 2018, ApJ, 853, 39
Hammerstein E. et al., 2023, ApJ, 942, aca283
Harutyunyan A. H. et al., 2008, A&A, 488, 383
Hills J., 1975, Nature, 254, 295
Holoien T. W.-S. et al., 2016, MNRAS, 455, 2918
Howell D. A. et al., 2005, ApJ, 634, 1190
Jeffery D. J., Branch D., 1990, in Wheeler J. C., Piran T., Weinberg S., eds, Jerusalem Winter School for Theoretical Physics. World Scientific, Singapore, p. 149
Kawana K., Maeda K., Yoshida N., Tanikawa A., 2020, ApJ, 890, L26
Komossa S., 2015, J. High Energy Astrophys., 7, 148
Law-Smith J., Ramirez-Ruiz E., Ellison S. L., Foley R. J., 2017, ApJ, 850, 22
Lipton Z. C., Elkan C., Naryanaswamy B., 2014, in Calders T., Esposito F., Hüllermeier E., Meo R., eds, Lecture Notes in Computer Science, Vol. 8725, Machine Learning and Knowledge Discovery in Databases. ECML PKDD 2014. Springer, Berlin, p. 15
Lochner M., McEwen J. D., Peiris H. V., Lahav O., Winter M. K., 2016, ApJS, 225, 31
Masterson M. et al., 2024, ApJ, 961, 211
Misner C. W., Thorne K. S., Wheeler J. A., 1973, Gravitation. W.H. Freeman and Company, New York
Muthukrishna D., Narayan G., Mandel K. S., Biswas R., Hložek R., 2019a, PASP, 131, 118002
Muthukrishna D., Parkinson D., Tucker B. E., 2019b, ApJ, 885, 85
Perley D. A. et al., 2020, ApJ, 904, 35
Poznanski D., Maoz D., Gal-Yam A., 2007, AJ, 134, 1285
Rees M., 1988, Nature, 333, 523
Reyes-Jainaga I. et al., 2023, ApJ, 952, 12
Richards J. W., Homrighausen D., Freeman P. E., Schafer C. M., Poznanski D., 2011, MNRAS, 419, 1121
Sánchez-Sáez P. et al., 2021, AJ, 161, 141
Santos M. V. D., Quartin M., Reis R. R. R., 2020, MNRAS, 497, 2974
Smartt S. J. et al., 2015, A&A, 579, A40
Stein R. et al., 2024, ApJ, 965, L14
Stone N. C., Kesden M., Cheng R. M., van Velzen S., 2019, Gen. Relativ. Gravit., 51, 30

Stuke A., Rinke P., Todorović M., 2021, Mach. Learn.: Sci. Technol., 2, 035022
Sutton C. D., 2005, in Rao C. R., Wegman E. J., Solka J. L., eds, Handbook of Statistics, Vol. 24, Classification and Regression Trees, Bagging, and Boosting. Elsevier, Amsterdam, p. 303
van Velzen S. et al., 2021, ApJ, 908, 4
Vargas M. L. G., Díaz A. I., Terlevich E., Terlevich R., 1993, Ap&SS, 205, 85
Velzen S. v., Farrar G. R., 2014, ApJ, 792, 53
Villar V. A. et al., 2020, ApJ, 905, ace326
Wang Y. H., Perna R., Armitage P. J., 2021, MNRAS, 503, 6005
Yao Y. et al., 2023, ApJ, 955, L6
Yaron O., Gal-Yam A., 2012, PASP, 124, 668
Zabludoff A. et al., 2021, Space Sci. Rev., 217, 54
Zhou Z. H., 2012, Ensemble Methods: Foundations and Algorithms. Chapman and Hall/CRC, New York
Zhu J., Jiang N., Wang T., Huang S., Lin Z., Wang Y., Wang J. G., 2023, ApJ, 952, 7

This paper has been typeset from a TEX/LATEX file prepared by the author.