

# SNAD: enabling discovery in the era of big data

**Maria Pruzhinskaya,<sup>a,b,\*</sup> Emille E. O. Ishida,<sup>b</sup> Konstantin Malanchev,<sup>c</sup>  
Anastasia Lavrukhina,<sup>a</sup> Etienne Russeil,<sup>b,d</sup> Timofey Semenikhin,<sup>a,e</sup>  
Sreevarsha Sreejith,<sup>f</sup> Emmanuel Gangler,<sup>b</sup> Matvey Kornilov,<sup>a,g</sup> Vladimir  
Korolev and Alina Volnova<sup>a,h</sup>**

<sup>a</sup>*Lomonosov Moscow State University, Sternberg Astronomical Institute,  
Universitetsky pr. 13, Moscow, 119234, Russia*

<sup>b</sup>*Université Clermont Auvergne, CNRS, LPCA,  
63000 Clermont-Ferrand, France*

<sup>c</sup>*McWilliams Center for Cosmology & Astrophysics, Department of Physics, Carnegie Mellon University,  
Pittsburgh, PA 15213, USA*

<sup>d</sup>*The Oskar Klein Centre, Department of Astronomy, Stockholm University,  
AlbaNova, SE-10691 Stockholm, Sweden*

<sup>e</sup>*Lomonosov Moscow State University, Faculty of Physics,  
Leninsky Gori 1, Moscow 119234, Russia*

<sup>f</sup>*Physics department, University of Surrey,  
Stag Hill, Guildford GU2 7XH*

<sup>g</sup>*National Research University Higher School of Economics,  
21/4 Staraya Basmannaya Ulitsa, Moscow, 105066, Russia*

<sup>h</sup>*Space Research Institute of the Russian Academy of Sciences (IKI),  
84/32 Profsoyuznaya Street, Moscow, 117997, Russia*

E-mail: [pruzhinskaya@gmail.com](mailto:pruzhinskaya@gmail.com)

---

\*Speaker

In the era of wide-field surveys and big data in astronomy, the SNAD team (<https://snad.space>) is exploiting the potential of modern datasets for discovering new, unforeseen, or rare astrophysical objects and phenomena with machine learning (ML). The SNAD pipeline was built under the hypothesis that, although automatic ML algorithms have a crucial role to play in this task, the scientific discovery is only completely realized when such systems are designed to boost the impact of domain knowledge experts. Our key contributions include the development of the *Coniferest* Python library, which offers implementations of two active learning algorithms with an “expert in loop”, and the creation of the SNAD Transient Miner, facilitating the search for specific types of transients. We have also developed the SNAD Viewer, a web portal that provides a centralized view of individual objects from the Zwicky Transient Facility’s (ZTF) data releases, making the analysis of potential anomalies more efficient. Finally, when applied to ZTF data, our approach has resulted in more than a hundred new supernova (SN) candidates, along with a few other non-catalogued objects, such as red dwarf flares, superluminous SNe, RS CVn type variables, and young stellar objects.

## 1. Introduction

The SNAD<sup>1</sup> [31] project focuses on applying anomaly detection (AD) techniques to astronomical datasets, combining machine learning (ML) methods with human expertise. In this proceeding, we discuss the key terminology used to describe anomalies in SNAD, introduce our active AD approach, and summarize our recent results.

## 2. Terminology

There is no consensus in the astronomical literature on the definitions of terms such as anomalies, novelties, outliers, abnormal or unusual objects, rare phenomena, or unknown types. While these terms are often treated as synonyms (e.g., [15]), some studies provide more explicit definitions, offering clear criteria for specific cases (e.g., [2], see Section 3).

In SNAD, we distinguish between machine-detected **outliers** and astrophysical **anomalies**, where the latter refers to objects or phenomena<sup>2</sup> of scientific interest to researchers.

### 2.1 Outlier

By outlier we assume a statistical deviation within the data or, in the context of ML anomaly detection algorithms, any object with the high anomaly score. The number of outliers depends on the chosen threshold. While these objects are unique or unusual from the perspective of the algorithm, they are not always of astrophysical interest, e.g. artifacts or misclassifications in considered catalogs. A good example could be a detection of a binary microlensing event in the Open Supernova Catalog [26].

The causes of these outliers can vary: they may result from poor feature engineering, an inappropriate choice of AD algorithm, or the specificities of the considered dataset.

### 2.2 Anomaly

In the context of SNAD, the concept of anomaly includes the following categories:

- **Unknown-unknowns** are completely unforeseen phenomena or objects that are not predicted by any current theory. While their discovery might be driven by the search for anomalies, the exact nature of what is being sought remains entirely unknown. Such findings could range from requiring minor extensions to existing theories to the development of entirely new theoretical models, possibly even introducing new physics. A good historical example is the discovery of GRBs in the late 1960s by the U.S. Vela satellites, which were originally designed to detect gamma radiation from potential secret nuclear tests in space [17].
- **Known-unknowns** are phenomena or objects that have been theoretically predicted but has not yet received observational confirmation, or whose confirmation remains unreliable. Examples include pair-instability SNe (candidates in [27, 28]), dark matter annihilation or decay

<sup>1</sup><https://snad.space/>

<sup>2</sup>The difference between object and phenomenon is essential. Examples of objects include galaxies, neutron stars, black holes, planets, asteroids etc. In contrast, phenomena encompass events with a characteristic timescale, such as supernovae (SNe), gamma-ray bursts (GRBs), solar flares etc.

(candidates discussed in [1, 5–8, 33]), primordial black holes, wormholes, Dyson spheres etc. While some of these discoveries are expected to eventually be made as they arise naturally from our fundamental theories and represent critical tests of our current understanding of the Universe (e.g., the existence of cosmic neutrino background or primordial gravitational waves), their confirmation awaits improvements in observational precision and the development of new instruments and experiments. Others, however, may come as complete surprises and fundamentally change our worldview (e.g., Dyson spheres).

- **Rare-knowns** refer to rare types of astronomical objects or phenomena that have already been observed (e.g., kilonovae, repeating fast radio bursts, superluminous supernovae (SLSNe)). The rarity of such anomalies may arise from their intrinsically small numbers in nature or from observational constraints, including selection effects. In the latter case, they should only be considered as anomalies while the number of identified members remains significantly lower compared to other known astronomical objects and phenomena.

#### Notes:

While belonging to a rare class (e.g., kilonovae, of which only approximately a dozen have been confidently confirmed; see references in [34]) is obvious to any astrophysicist, subtler anomalies may involve unusual behavior within a common class. For example, peculiar modulation cycles of some RR Lyrae stars [22], the appearance of anomalous features in a lensing light curve [14], or spectroscopically peculiar SNe [10]. Such anomalies are more likely to be identified by domain experts. For example, a specialist in SNe may overlook an interesting feature in a Cepheid light curve, while it would be immediately recognized by a Cepheid expert. This makes the definition of an anomaly **expert-dependent**, as it varies based on the expert’s knowledge and field of expertise.

Secondly, the definition of an anomaly depends on the goals of the project. Even though M-dwarf flares are well-known phenomena, identifying them in a specific dataset can still meet the criteria for rarity in that context (e.g., [32]). Another example is the study by [2], where spectroscopic anomalies are defined as any transient whose spectroscopic label is not of Type Ia-normal, Ia-91T-like, II-normal, or IIP supernovae. This makes the definition of anomaly **task-dependent**, as it is tied to the specific scientific objectives or the dataset under consideration.

### 3. Active approach

The goal of “healthy” anomaly detection is *to ensure that a large fraction of outliers are anomalies*. To achieve this, SNAD employs an **active** anomaly detection approach.

Active anomaly detection algorithms adapt to human expert preferences during the learning process. These algorithms iteratively refine their models based on expert feedback, enabling them to better identify objects of scientific interest.

To support this approach, we developed Coniferest<sup>3</sup> [18], a Python library which offers implementations of three anomaly detection algorithms: the Isolation Forest and two active learning algorithms with an “expert in loop” — Active Anomaly Discovery (AAD, [9, 16]) and PineForest, the latter being a method developed by our group (Korolev et al., in prep.).

<sup>3</sup><https://coniferest.snad.space/>

## 4. Projects and main results

As stated in Section 2.2, the definition of an anomaly is both expert- and task-dependent. As a result, several sub-projects have emerged within SNAD, each focusing on different aspects of anomaly detection.

### 4.1 Traditional anomaly detection

By applying different AD algorithms to the photometric data from the Open Supernova Catalog (OSC, [13]) and the Zwicky Transient Facility survey (ZTF, [4]), we identified a few dozen potentially interesting anomalies, as described in [26], [20], and [30].

Among the diverse findings, there are candidates in SNe (e.g., `oid`<sup>4</sup> = 795209200003484), pre-main-sequence stars (e.g., `oid` = 807210200026027), novae (e.g., `oid` = 695211100022045), active galactic nuclei (AGNs; e.g., `oid` = 718211400012193), luminous blue variables (e.g., `oid` = 695211100131796), a binary microlensing event AT 2021luey (`oid` = 643210400013909), a radio source NVSS J080730+755017 with an optical counterpart (`oid` = 858205100001741), and a RS Canum Venaticorum star (`oid` = 695211200019653).

Furthermore, we also discovered numerous new variable stars (e.g., `oid` = 807212100012737) and identified several cases of misclassification in the OSC, such as SN 2006kg, which is an AGN, and Gaia16aye, a binary microlensing event misclassified as supernova.

### 4.2 Supernovae as anomalies

This study represents the first science case testing the SNAD active learning pipeline [25].

We analyzed 70 ZTF fields at high galactic latitudes, comprising ~26.5 million light curves, and visually inspected 2100 outliers identified by the AAD algorithm. This resulted in 104 SN-like objects, 57 of which were newly reported to the Transient Name Server (TNS), while 47 had already been mentioned in other catalogs as either confirmed SNe or candidates. The light curves of unreported transients were visually inspected and fitted with supernova models, assigning to them a probable photometric classification: Ia, Ib/c, IIP, IIL, or IIn. Additionally, we identified slow-evolving transients as promising SLSNe, along with a few other non-catalogued events such as M-dwarf flares and AGNs.

This work demonstrates the effectiveness of the human-machine interaction underlying the AAD strategy. The SNe discovered with AAD, along with those identified by the SNAD Transient Miner [3], form the basis of the SNAD catalog of transients: <https://snad.space/catalog/>.

### 4.3 Superluminous supernovae as anomalies

Superluminous supernovae are transients significantly brighter ( $L_{\max} \gtrsim 10^{44}$  erg s<sup>-1</sup>) than ordinary supernovae, representing a relatively rare class of objects<sup>5</sup> [23]. The exact mechanisms of SLSN explosions, which explain their observational properties, remain unknown. The explanations of their luminosity include the production of large amounts of radioactive <sup>56</sup>Ni (exceeding several solar masses, sometimes even dozens), interaction of the ejecta with a dense circumstellar medium

<sup>4</sup>ZTF object identifier; <https://ztf.snad.space/>

<sup>5</sup>As of January 2025, TNS lists 278 classified SLSNe. Also, a catalog specifically focused on SLSNe I includes 262 events [12].

(CSM), magnetar-powered energy deposition, and fallback accretion onto a newly formed black hole (for a review, see [11]). SLSNe can be detected at high redshifts ( $z \geq 2$ ), making them valuable for studying star formation in the early Universe, the interstellar medium, and even for cosmological applications [24].

The first SNAD SLSN candidates were identified during a supernova search using active learning applied to ZTF DR3 photometric data [25]. One of the candidates, SNAD160, turned out to be a possible pair-instability SN, as its light curve exhibits a slow rise and decay consistent with predictions for such events [27].

In a follow-up study by [19], the PineForest active AD algorithm was used to search for SLSNe in ZTF data. Starting with a dataset of  $\sim 14$  million objects and using 8 previously confirmed SLSN light curves as priors, we scrutinized 120 candidates. This led to the identification of 8 SLSN candidates, 2 of which (AT 2018moa and AT 2018mob) had not been reported previously.

#### 4.4 M-dwarf flares as anomalies

In this project, we focus on searching for M-dwarf flares in ZTF data. M-dwarfs are main candidates for hosting extraterrestrial life, making it crucial to understand their activity. Moreover, such phenomena provide valuable insights into the magnetic activity of astronomical objects, including the Sun.

In [32] we analyzed over 35 million ZTF light curves and constructed the largest catalogue of ZTF flaring stars to date, identifying 134 flares with amplitudes ranging from  $-0.2$  mag to  $-4.6$  mag, including repeated flares.

#### 4.5 Artefacts as anomalies

A significant portion of the outliers identified by AD algorithms consists of artefacts. For example, 68 percent of all outliers analyzed in [20] were found to be bogus light curves. Even with an active learning approach, it is impossible to reject all artefacts immediately, as they are highly diverse and occupy various regions of the parameter space.

During the expert analysis of outliers, we created the SNAD Knowledge Database [21], which contains a few thousand labeled ZTF objects, approximately half of which are artefacts. We also stored the most interesting artefacts in the SNAD artefact catalog<sup>6</sup>, initially without specific scientific goals. Using this small but well-labeled sample, we developed a real-bogus classifier and integrated its results into the active anomaly detection pipeline as an additional feature. This significantly reduced the number of artefacts among outliers in regions of the feature space densely populated with artefacts [29].

Recognizing the utility of this approach, we started a dedicated campaign to systematically search for artefacts in ZTF using the PineForest algorithm – where artefacts were treated as anomalies. The resulting artefact dataset includes 1127 FITS images of two sizes ( $28 \times 28$  and  $63 \times 63$  pixels), along with a corresponding set of images of nominal objects. This was performed to support machine learning applications in astronomy, offering a real human-labeled dataset for the development and testing of algorithms (Sreejith et al., in preparation).

---

<sup>6</sup><https://snad.space/art/>

## 5. Conclusions

- **Well-defined concepts are a fundamental part of scientific progress.** In SNAD, we distinguish machine-detected outliers from astrophysical anomalies and classify anomalies into unknown-unknowns, known-unknowns, and rare-knowns. Their interpretation is inherently expert-dependent and task-dependent, emphasizing the role of context in anomaly detection.
- **Active anomaly detection is an efficient approach** that significantly improves the identification of scientifically interesting objects and phenomena by incorporating expert feedback into the learning process.
- **The SNAD transient catalog has been created**, with some supernova candidates missed in the official ZTF alert stream. These findings can help avoid similar losses in future large-scale astronomical surveys.
- **Several interesting anomalies were identified**, including SLSN candidates, one PISN candidate (SNAD160), binary microlensing events, previously uncatalogued AGNs, cataclysmic variables, and other rare variable stars.
- **The largest catalog of ZTF flaring stars to date was compiled**, representing 134 M-dwarf flares with amplitudes ranging from  $-0.2$  mag to  $-4.6$  mag.
- **Outliers in real data are often artifacts**, which can either be studied as a separate class or effectively filtered out to improve anomaly detection efficiency.
- **Anomaly detection systems must be built for experts**, ensuring interpretability and adaptability to domain-specific needs.

## Acknowledgements

We acknowledge Patrick Aleo, Boris Demkov, Vadim Krushinskiy, Anastasia Malancheva, Alexandra Novinskaya and Anastasia Voloshina for their contribution to SNAD. M. Kornilov, A. Lavrukina, M. Pruzhinskaya, T. Semenikhin acknowledge support from a Russian Science Foundation grant 24-22-00233, <https://rscf.ru/en/project/24-22-00233/>.

## References

- [1] O. Adriani, G. C. Barbarino, G. A. Bazilevskaya, R. Bellotti, M. Boezio, E. A. Bogomolov, L. Bonechi, M. Bongi, V. Bonvicini, S. Bottai, A. Bruno, F. Cafagna, D. Campana, P. Carlson, M. Casolino, G. Castellini, M. P. de Pascale, G. de Rosa, N. de Simone, V. di Felice, A. M. Galper, L. Grishantseva, P. Hofverberg, S. V. Koldashov, S. Y. Krutkov, A. N. Kvashnin, A. Leonov, V. Malvezzi, L. Marcelli, W. Menn, V. V. Mikhailov, E. Mocchiutti, S. Orsi, G. Osteria, P. Papini, M. Pearce, P. Picozza, M. Ricci, S. B. Ricciarini, M. Simon, R. Sparvoli, P. Spillantini, Y. I. Stozhkov, A. Vacchi, E. Vannuccini, G. Vasilyev, S. A. Voronov, Y. T. Yurkin, G. Zampa, N. Zampa, and V. G. Zverev. An anomalous positron abundance in cosmic rays with energies 1.5-100GeV. , 458(7238):607–609, April 2009.

- [2] P. D. Aleo, A. W. Engel, G. Narayan, C. R. Angus, K. Malanchev, K. Auchettl, V. F. Baldassare, A. Berres, T. J. L. de Boer, B. M. Boyd, K. C. Chambers, K. W. Davis, N. Esquivel, D. Farias, R. J. Foley, A. Gagliano, C. Gall, H. Gao, S. Gomez, M. Grayling, D. O. Jones, C. C. Lin, E. A. Magnier, K. S. Mandel, T. Matheson, S. I. Raimundo, V. G. Shah, M. D. Soraisam, K. M. de Soto, S. Vicencio, V. A. Villar, and R. J. Wainscoat. Anomaly Detection and Approximate Similarity Searches of Transients in Real-time Data Streams. , 974(2):172, October 2024.
- [3] P. D. Aleo, K. L. Malanchev, M. V. Pruzhinskaya, E. E. O. Ishida, E. Russeil, M. V. Kornilov, V. S. Korolev, S. Sreejith, A. A. Volnova, and G. S. Narayan. SNAD transient miner: Finding missed transient events in ZTF DR4 using k-D trees. , 96:101846, October 2022.
- [4] Eric C. Bellm, Shrinivas R. Kulkarni, Matthew J. Graham, Richard Dekany, Roger M. Smith, Reed Riddle, Frank J. Masci, George Helou, Thomas A. Prince, Scott M. Adams, C. Barbarino, Tom Barlow, James Bauer, Ron Beck, Justin Belicki, Rahul Biswas, Nadejda Blagorodnova, Dennis Bodewits, Bryce Bolin, Valery Brinnel, Tim Brooke, Brian Bue, Mattia Bulla, Rick Burruss, S. Bradley Cenko, Chan-Kao Chang, Andrew Connolly, Michael Coughlin, John Cromer, Virginia Cunningham, Kishalay De, Alex Delacroix, Vandana Desai, Dmitry A. Duev, Gwendolyn Eadie, Tony L. Farnham, Michael Feeney, Ulrich Feindt, David Flynn, Anna Franckowiak, S. Frederick, C. Fremling, Avishay Gal-Yam, Suvi Gezari, Matteo Giomi, Daniel A. Goldstein, V. Zach Golkhou, Ariel Goobar, Steven Groom, Eugene Hacquard, David Hale, John Henning, Anna Y. Q. Ho, David Hover, Justin Howell, Tiara Hung, Daniela Huppenkothen, David Imel, Wing-Huen Ip, Željko Ivezić, Edward Jackson, Lynne Jones, Mario Juric, Mansi M. Kasliwal, S. Kaspi, Stephen Kaye, Michael S. P. Kelley, Marek Kowalski, Emily Kramer, Thomas Kupfer, Walter Landry, Russ R. Laher, Chien-De Lee, Hsing Wen Lin, Zhong-Yi Lin, Ragnhild Lunnan, Matteo Giomi, Ashish Mahabal, Peter Mao, Adam A. Miller, Serge Monkenwitz, Patrick Murphy, Chow-Choong Ngeow, Jakob Nordin, Peter Nugent, Eran Ofek, Maria T. Patterson, Bryan Penprase, Michael Porter, Ludwig Rauch, Umaa Rebbapragada, Dan Reiley, Mickael Rigault, Hector Rodriguez, Jan van Roestel, Ben Rusholme, Jakob van Santen, S. Schulze, David L. Shupe, Leo P. Singer, Maayane T. Soumagnac, Robert Stein, Jason Surace, Jesper Sollerman, Paula Szkody, F. Taddia, Scott Terek, Angela Van Sistine, Sjoert van Velzen, W. Thomas Vestrand, Richard Walters, Charlotte Ward, Quan-Zhi Ye, Po-Chieh Yu, Lin Yan, and Jeffry Zolkower. The zwicky transient facility: System overview, performance, and first results. *Publications of the Astronomical Society of the Pacific*, 131(995):018002, dec 2018.
- [5] A. Boyarsky, O. Ruchayskiy, D. Iakubovskiy, and J. Franse. Unidentified Line in X-Ray Spectra of the Andromeda Galaxy and Perseus Galaxy Cluster. , 113(25):251301, December 2014.
- [6] Torsten Bringmann, Xiaoyuan Huang, Alejandro Ibarra, Stefan Vogl, and Christoph Weniger. Fermi LAT search for internal bremsstrahlung signatures from dark matter annihilation. , 2012(7):054, July 2012.
- [7] Esra Bulbul, Maxim Markevitch, Adam Foster, Randall K. Smith, Michael Loewenstein, and



- Scott W. Randall. Detection of an Unidentified Emission Line in the Stacked X-Ray Spectrum of Galaxy Clusters. , 789(1):13, July 2014.
- [8] J. Chang, J. H. Adams, H. S. Ahn, G. L. Bashindzhagyan, M. Christl, O. Ganel, T. G. Guzik, J. Isbert, K. C. Kim, E. N. Kuznetsov, M. I. Panasyuk, A. D. Panov, W. K. H. Schmidt, E. S. Seo, N. V. Sokolskaya, J. W. Watts, J. P. Wefel, J. Wu, and V. I. Zatsepin. An excess of cosmic ray electrons at energies of 300-800GeV. , 456(7220):362–365, November 2008.
- [9] Shubhomoy Das, Weng-Keen Wong, Alan Fern, Thomas G. Dietterich, and Md Amran Siddiqui. Incorporating Feedback into Tree-based Anomaly Detection. In *Workshop on Interactive Data Exploration and Analytics (IDEA'17)*, KDD workshop, page arXiv:1708.09441, Aug 2017.
- [10] Kishalay De, U. Christoffer Fremling, Avishay Gal-Yam, Ofer Yaron, Mansi M. Kasliwal, and S. R. Kulkarni. The Peculiar Ca-rich SN2019ehk: Evidence for a Type IIb Core-collapse Supernova from a Low-mass Stripped Progenitor. , 907(1):L18, January 2021.
- [11] Avishay Gal-Yam. The Most Luminous Supernovae. , 57:305–333, August 2019.
- [12] Sebastian Gomez, Matt Nicholl, Edo Berger, Peter K. Blanchard, V. Ashley Villar, Sofia Rest, Griffin Hosseinzadeh, Aysha Aamer, Yukta Ajay, Wasundara Athukoralalage, David C. Coulter, Tarraneh Eftekhari, Achille Fiore, Noah Franz, Ori Fox, Alexander Gagliano, Daichi Hiramatsu, D. Andrew Howell, Brian Hsu, Mitchell Karmen, Matthew R. Siebert, Réka Könyves-Tóth, Harsh Kumar, Curtis McCully, Craig Pellegrino, Justin Pierel, Armin Rest, and Qinan Wang. The Type I superluminous supernova catalogue I: light-curve properties, models, and catalogue description. , 535(1):471–515, November 2024.
- [13] James Guillochon, Jerod Parrent, Luke Zoltan Kelley, and Raffaella Margutti. An Open Catalog for Supernova Data. , 835(1):64, January 2017.
- [14] Cheongho Han, Andrzej Udalski, Ian A. Bond, Chung-Uk Lee, Andrew Gould, Michael D. Albrow, Sun-Ju Chung, Kyu-Ha Hwang, Youn Kil Jung, Hyoun-Woo Kim, Yoon-Hyun Ryu, Yossi Shvartzvald, In-Gu Shin, Jennifer C. Yee, Hongjing Yang, Weicheng Zang, Sang-Mok Cha, Doeon Kim, Dong-Jin Kim, Seung-Lee Kim, Dong-Joo Lee, Yongseok Lee, Byeong-Gon Park, Richard W. Pogge, Przemek Mróz, Michał K. Szymański, Jan Skowron, Radosław Poleski, Igor Soszyński, Paweł Pietrukowicz, Szymon Kozłowski, Krzysztof A. Rybicki, Patryk Iwanek, Krzysztof Ulaczyk, Marcin Wrona, Mariusz Gromadzki, Mateusz J. Mróz, Fumio Abe, Richard Barry, David P. Bennett, Aparna Bhattacharya, Hirosame Fujii, Akihiko Fukui, Ryusei Hamada, Yuki Hirao, Stela Ishitani Silva, Yoshitaka Itow, Rintaro Kirikawa, Naoki Koshimoto, Yutaka Matsubara, Shota Miyazaki, Yasushi Muraki, Greg Olmschenk, Clément Ranc, Nicholas J. Rattenbury, Yuki Satoh, Takahiro Sumi, Daisuke Suzuki, Mio Tomoyoshi, Paul J. Tristram, Aikaterini Vandenbroucke, Hibiki Yama, and Kansuke Yamashita. OGLE-2018-BLG-0971, MOA-2023-BLG-065, and OGLE-2023-BLG-0136: Microlensing events with prominent orbital effects. , 686:A234, June 2024.

- [15] Yang Han, Zhiqiang Zou, Nan Li, and Yanli Chen. Identifying Outliers in Astronomical Images with Unsupervised Machine Learning. *Research in Astronomy and Astrophysics*, 22(8):085006, August 2022.
- [16] E. E. O. Ishida, M. V. Kornilov, K. L. Malanchev, M. V. Pruzhinskaya, A. A. Volnova, V. S. Korolev, F. Mondon, S. Sreejith, A. A. Malancheva, and S. Das. Active anomaly detection for time-domain discoveries. , 650:A195, June 2021.
- [17] Ray W. Klebesadel, Ian B. Strong, and Roy A. Olson. Observations of Gamma-Ray Bursts of Cosmic Origin. , 182:L85, June 1973.
- [18] M. V. Kornilov, V. S. Korolev, K. L. Malanchev, A. D. Lavrukhina, E. Russeil, T. A. Semenikhin, E. Gangler, E. E. O. Ishida, M. V. Pruzhinskaya, A. A. Volnova, and S. Sreejith. Coniferest: a complete active anomaly detection framework. *arXiv e-prints*, page arXiv:2410.17142, October 2024.
- [19] T. Majumder, M. V. Pruzhinskaya, E. E. O. Ishida, K. L. Malanchev, and T. A. Semenikhin. Superluminous supernova search with PineForest. *arXiv e-prints*, page arXiv:2410.21077, October 2024.
- [20] K. L. Malanchev, M. V. Pruzhinskaya, V. S. Korolev, P. D. Aleo, M. V. Kornilov, E. E. O. Ishida, V. V. Krushinsky, F. Mondon, S. Sreejith, A. A. Volnova, A. A. Belinski, A. V. Dodin, A. M. Tatarnikov, S. G. Zheltoukhov, and (The SNAD Team). Anomaly detection in the Zwicky Transient Facility DR3. , 502(4):5147–5175, April 2021.
- [21] Konstantin Malanchev, Matvey V. Kornilov, Maria V. Pruzhinskaya, Emille E. O. Ishida, Patrick D. Aleo, Vladimir S. Korolev, Anastasia Lavrukhina, Etienne Russeil, Sreevarsha Sreejith, Alina A. Volnova, Anastasiya Voloshina, and Alberto Krone-Martins. The SNAD Viewer: Everything You Want to Know about Your Favorite ZTF Object. , 135(1044):024503, February 2023.
- [22] László Molnár, Emese Plachy, Attila Bódi, András Pál, Meridith Joyce, Csilla Kalup, Christian I. Johnson, Zoltán Dencs, Szabolcs Mészáros, Henryka Netzel, Karen Kinemuchi, Juna A. Kollmeier, Jose Luis Prieto, and Aliz DEREKAS. To grow old and peculiar: Survey of anomalous variable stars in M80 with age determinations using K2 and Gaia. , 678:A104, October 2023.
- [23] Takashi J. Moriya. Superluminous supernovae. *arXiv e-prints*, page arXiv:2407.12302, July 2024.
- [24] Takashi J. Moriya, Elena I. Sorokina, and Roger A. Chevalier. Superluminous Supernovae. , 214(2):59, March 2018.
- [25] M. V. Pruzhinskaya, E. E. O. Ishida, A. K. Novinskaya, E. Russeil, A. A. Volnova, K. L. Malanchev, M. V. Kornilov, P. D. Aleo, V. S. Korolev, V. V. Krushinsky, S. Sreejith, and E. Gangler. Supernova search with active learning in ZTF DR3. , 672:A111, April 2023.

- [26] M. V. Pruzhinskaya, K. L. Malanchev, M. V. Kornilov, E. E. O. Ishida, F. Mondon, A. A. Volnova, and V. S. Korolev. Anomaly detection in the Open Supernova Catalog. , 489(3):3591–3608, November 2019.
- [27] Maria Pruzhinskaya, Alina Volnova, Matwey Kornilov, Konstantin Malanchev, Patrick D. Aleo, Emille E. O. Ishida, Vladimir Korolev, Alexandra Novinskaya, Etienne Russeil, Sreevarsha Sreejith, Stéphane Blondin, Alexandra Kozyreva, and SNAD Team. Could SNAD160 be a Pair-instability Supernova? *Research Notes of the American Astronomical Society*, 6(6):122, June 2022.
- [28] Steve Schulze, Claes Fransson, Alexandra Kozyreva, Ting-Wan Chen, Ofer Yaron, Anders Jerkstrand, Avishay Gal-Yam, Jesper Sollerman, Lin Yan, Tuomas Kangas, Giorgos Leloudas, Conor M. B. Omand, Stephen J. Smartt, Yi Yang, Matt Nicholl, Nikhil Sarin, Yuhan Yao, Thomas G. Brink, Amir Sharon, Andrea Rossi, Ping Chen, Zhihao Chen, Aleksandar Cikota, Kishalay De, Andrew J. Drake, Alexei V. Filippenko, Christoffer Fremling, Laurane Fréour, Johan P. U. Fynbo, Anna Y. Q. Ho, Cosimo Inserra, Ido Irani, Hanindyo Kuncarayakti, Ragnhild Lunnan, Paolo Mazzali, Eran O. Ofek, Eliana Palazzi, Daniel A. Perley, Miika Pursiainen, Barry Rothberg, Luke J. Shingles, Ken Smith, Kirsty Taggart, Leonardo Tartaglia, WeiKang Zheng, Joseph P. Anderson, Letizia Cassara, Eric Christensen, S. George Djorgovski, Lluís Galbany, Anamaria Gkini, Matthew J. Graham, Mariusz Gromadzki, Steven L. Groom, Daichi Hiramatsu, D. Andrew Howell, Mansi M. Kasliwal, Curtis McCully, Tomás E. Müller-Bravo, Simona Paiano, Emmanouela Paraskeva, Priscila J. Pessi, David Polishook, Arne Rau, Mickael Rigault, and Ben Rusholme. 1100 days in the life of the supernova 2018ibb. The best pair-instability supernova candidate, to date. , 683:A223, March 2024.
- [29] T. A. Semenikhin, M. V. Kornilov, M. V. Pruzhinskaya, A. D. Lavrukhina, E. Russeil, E. Gangler, E. E. O. Ishida, V. S. Korolev, K. L. Malanchev, A. A. Volnova, and S. Sreejith. Real-bogus scores for active anomaly detection. *Astronomy and Computing*, 51:100919, April 2025.
- [30] Alina Volnova, Patrick D. Aleo, Emmanuel Gangler, Emille E. O. Ishida, Matwey Kornilov, Vladimir Korolev, Vadim Krushinsky, Anastasia Lavrukhina, Konstantin Malanchev, Maria Pruzhinskaya, Etienne Russeil, Timofey Semenikhin, Sreevarsha Sreejith, and SNAD Team. The Most Interesting Anomalies Discovered in ZTF DR17 from the SNAD-VI Workshop. *Research Notes of the American Astronomical Society*, 7(7):155, July 2023.
- [31] Alina A. Volnova, Patrick D. Aleo, Anastasia Lavrukhina, Etienne Russeil, Timofey Semenikhin, Emmanuel Gangler, Emille E. O. Ishida, Matwey V. Kornilov, Vladimir Korolev, Konstantin Malanchev, Maria V. Pruzhinskaya, and Sreevarsha Sreejith. Exploring the Universe with SNAD: Anomaly Detection in Astronomy. *arXiv e-prints*, page arXiv:2410.18875, October 2024.
- [32] A. S. Voloshina, A. D. Lavrukhina, M. V. Pruzhinskaya, K. L. Malanchev, E. E. O. Ishida, V. V. Krushinsky, P. D. Aleo, E. Gangler, M. V. Kornilov, V. S. Korolev, E. Russeil, T. A. Semenikhin, S. Sreejith, A. A. Volnova, and SNAD Team. SNAD catalogue of M-dwarf flares from the Zwicky Transient Facility. , 533(4):4309–4323, October 2024.

- [33] Christoph Weniger. A tentative gamma-ray line from Dark Matter annihilation at the Fermi Large Area Telescope. , 2012(8):007, August 2012.
- [34] Jin-Ping Zhu, Shichao Wu, Yuan-Pei Yang, Chang Liu, Bing Zhang, Hao-Ran Song, He Gao, Zhoujian Cao, Yun-Wei Yu, Yacheng Kang, and Lijing Shao. Kilonovae and Optical Afterglows from Binary Neutron Star Mergers. II. Optimal Search Strategy for Serendipitous Observations and Target-of-opportunity Observations of Gravitational Wave Triggers. , 942(2):88, January 2023.

## DISCUSSION

**MARIA DAINOTTI:** Do you plan to have a citizen science project so that you can have on independent way to identify the bogus?

**MARIA PRUZHINSKAYA:** Currently, we do not have plans for a citizen science project, but it is certainly an interesting possibility to consider for the future.

**CLAUDIA MARASTON:** Do you have an estimate of the metallicity of the progenitor star of the PISN you detected? Can you extend your approach to a search for the most massive stars?

**MARIA PRUZHINSKAYA:** We obtained the spectrum of the host of SNAD160 and estimated its metallicity. So far, our anomaly detection pipeline has not identified any other PISN candidates. In principle, active anomaly detection can be applicable to the search for any type of object, as long as they share common features present in the considered dataset.

**INGYIN ZAW:** Description of how filtering out uninteresting candidates (to humans) that the machine thinks are interesting is clear. How do you deal with the possibility that the machine is missing candidates that humans would have found interesting?

**MARIA PRUZHINSKAYA:** It is possible that the machine learning algorithm is missing really interesting objects, as there are many types of astrophysical phenomena. However, we believe that while truly anomalous objects are rare, they are not singular occurrences – there are multiple anomalies waiting to be discovered. At this stage, we prioritize identifying unexpected and scientifically valuable cases rather than ensuring completeness. Finding even a single anomaly can be a significant breakthrough.