# Machine learning techniques for analysis of photometric data from the Open Supernova catalog

M.V. Kornilov[1,2,*], M.V. Pruzhinskaya[1], K.L. Malanchev[1,2], E.E.O. Ishida[3], F. Mondon[3], A.A. Volnova[4], V.S. Korolev[5]

[1]*Sternberg Astronomical Institute, Lomonosov Moscow State University, Russia; matwey.kornilov@gmail.com*
[2]*National Research University Higher School of Economics; Moscow, Russia*
[3]*Université Clermont Auvergne, CNRS/IN2P3, LPC, France*
[4]*Space Research Institute of the Russian Academy of Sciences (IKI), Moscow, Russia*
[5]*Central Aerohydrodynamic Institute, Zhukovsky, Russia*

**Abstract** The next generation of astronomical surveys will revolutionize our understanding of the Universe, raising unprecedented data challenges in the process. One of them is the impossibility to rely on human scanning for the identification of unusual/unpredicted astrophysical objects. Moreover, given that most of the available data will be in the form of photometric observations, such characterization cannot rely on the existence of high resolution spectroscopic observations. The goal of this project is to detect the anomalies in the Open Supernova Catalog (http://sne.space/) with use of machine learning. We will develop a pipeline where human expertise and modern machine learning techniques can complement each other. Using supernovae as a case study, our proposal is divided in two parts: the first developing a strategy and pipeline where anomalous objects are identified, and a second phase where such anomalous objects submitted to careful individual analysis. The strategy requires an initial data set for which spectroscopic is available for training purposes, but can be applied to a much larger data set for which we only have photometric observations. This project represents an effective strategy to guarantee we shall not overlook exciting new science hidden in the data we fought so hard to acquire.

**Keywords:** Machine Learning Techniques, Supernovae, Astronomical Surveys

## 1. Introduction

Supernova stars (SNe) are ones of the most brightness and interesting objects in the Universe. They are responsible for chemical enrichment of interstellar medium; density waves induced by their energetic explosions causes the star formation; SNe are origin of high energy cosmic rays; moreover, thanks to SNe we are studying the composition and distance scale of the Universe which defines its following destiny.

The generation of precise, large, and complete supernova surveys in the last years has increased the need of developing automated analysis tools to process this large amount of data. These scientific observations present both great opportunities and challenges for astronomers and machine learning (ML) researchers.

The lack of spectroscopic support makes the photometrical supernova typing is very required. The analysis of big supernova dataset with ML methods is needed to distinguish the supernova by types on base of N-parameter grid. Such study allows us to purify the considered

SN sample from non-supernova contamination as well — the problem, which is relevant for all large supernova database that collect SN candidates without careful analysis of each candidate and basing on the secondary indicators (proximity to the galaxies, transient behaviour, arise/decline rate on light curves (LCs), absolute magnitude). It is also expected that during such analysis the unknown variable objects or SNe with unusual properties can be detected. As an example of unique objects one can refer to SN 2006jc — SN with very strong but relatively narrow He I lines in early spectra (∼30 similar objects are known, [25]), SN 2005bf — supernova attributed to SN Ib but with two broad maxima on LCs, SN 2010mb — unusual SN Ic with very low decline rate after the maximum brightness that is not consistent with radioactive decay of 56Ni, ASASSN-15lh — for some time it was considered as the most luminous supernova ever observed (two times brighter than super-luminous SNe), later the origin of this object was challenged and now it is considered as a tidal disruption of a main-sequence star by a black hole. Finding such objects (and then studying them more closely) is one of the main aims of the current project. As such sources are typically rare, the task of finding them can be framed as an anomaly detection problem.

Astronomers have already benefited from developments in machine learning [2], in particular for exoplanet search [22, 29, 26], but the synergy is far from that achieved by other endeavours in genetics [17], ecology [9] or medicine [30], where scientific questions drive the development of new algorithms. Moreover, given the relatively recent advent of large data sets, most of the ML efforts in astronomy are concentrated in classification [16, 15, 19] and regression [13, 6] tasks.

Astronomical anomaly detection has not been yet fully implemented in the enormous amount of data that has been gathered. As a matter of fact, barring a few exceptions, most of the previous studies can be divided into only two different trends: clustering [27] and subspace analysis [12] methods. More recently, random forest algorithms have been extensively used by themselves [3] or in hybrid statistical analysis [24]. Although all of this has been done to periodic variables there is not much done for transients and even less for supernova.

In this study we search the anomalies in photometrical data of the Open Supernova Catalog[a] [11]. We use the Isolation Forest as an outlier detection algorithm that identifies anomalies instead of normal observations [18]. This technique is based on the fact that anomalies are data points that are few and different. Similarly to Random Forest it is built on an ensemble of binary (isolation) trees.

## 2 Data

### 2.1 The Open Supernova Catalog

The data are drawn from the Open Supernova Catalog [11]. The catalog is constructed by combining many publicly available data sources (such as Asiago Supernova Catalog, Carnegie Supernova Project, Gaia Photometric Science Alerts, Nearby Supernova Factory, Panoramic Survey Telescope & Rapid Response System (Pan-STARRS), SDSS Supernova Survey, Sternberg Astronomical Institute Supernova Light Curve Catalogue, Supernova Legacy Survey (SNLS), MASTER, All-Sky Automated Survey for Supernovae (ASAS-SN), iPTF, etc.) and from individual publications. It represents an open repository for supernova metadata,

---

[a] https://sne.space/

light curves, and spectra in an easily downloadable format. This catalog also includes some contamination from non-SN objects.

Our choice is justified by the fact that the catalog incorporates the data for more than $5 \times 10^4$ SNe/SNe candidates ($\sim 1.2 \times 10^4$ of SNe have > 10 photometrical observations and $\sim 5\times10^3$ of SNe have spectra). For comparison, SDSS supernova catalog contains only $\sim 4 \times 10^3$ of SNe LCs and $\sim$ 600 SNe with spectra.

The catalog contains the data in different photometrical passbands. To have a more homogeneous data sample, we chose only those SNe that have LCs in $g'r'i'$, *gri* or BRI filters. We assume that $g'r'i'$ filters are close enough to gri and transform BRI to *gri* (see Sect. 2.2). We require >= 3 photometrical points in each filter with a 3-day binning. After this cut, our sample contains 3197 objects (2026 objects in $g'r'i'$, 767 objects in *gri*, and 404 objects in BRI).

### 2.2 Transformation between *BRI* and *gri*

To increase the sample we convert the Bessel's BRI into *gri* filters using the Lupton's (2005) transformation equations[b]. These equations are derived by matching SDSS DR4 photometry to Peter Stetson's published photometry for stars:

$$\begin{cases} B = u - 0.8116\,(u - g) + 0.1313 \\ B = g + 0.3130\,(g - r) + 0.2271 \\ V = g - 0.2906\,(u - g) + 0.0885 \\ V = g - 0.5784\,(g - r) - 0.0038 \\ R = r - 0.1837\,(g - r) - 0.0971 \\ R = r - 0.2936\,(r - i) - 0.1439 \end{cases} \tag{1}$$

## 3 Anomaly detection

### 3.1. LCs fit

It is more convenient to implement the ML algorithm to the data with uniform time grid which is unfortunately not the case with supernovae. Commonly used technique to transform unevenly distributed data onto uniform grid is to fit them with Gaussian processes (GP). Usually, each light curve is fitted by GP independently. However, in this study we developed the MULTIVARIATE GAUSSIAN PROCESS[c] interpolation that allows correlating multi-color LCs and approximates the data by GP in all filters in a one global fit (for details see Kornilov et. 2019, in prep.).

When the fit by MULTIVARIATE GAUSSIAN PROCESS was done, we checked the results of approximation by eye. Those SNe with unsatisfactory fit were removed from the further consideration (mainly the objects with bad photometrical quality). We also extrapolated the fit to have a bigger temporal coverage. In the end we got a sample that consists of 1999 objects.

---

[b] http://www.sdss3.org/dr8/algorithms/sdssUBVRITransform.php
[c] https://github.com/matwey/gp-multistate-kernel

Based on the results of approximation we extracted photometry (in flux) in the range of [−20, 100] days with 1-day bin relative to the LC maximum in *r* filter and the kernel parameters.

After the approximation procedure, each object has 373 features: 121 × 3 fluxes in three bands, 9 fitted parameters of Gaussian Process kernel, and logarithm of likelihood of the fit. We examine two cases of outliers search: with all features and with smaller number of features obtained by dimensionality reduction.

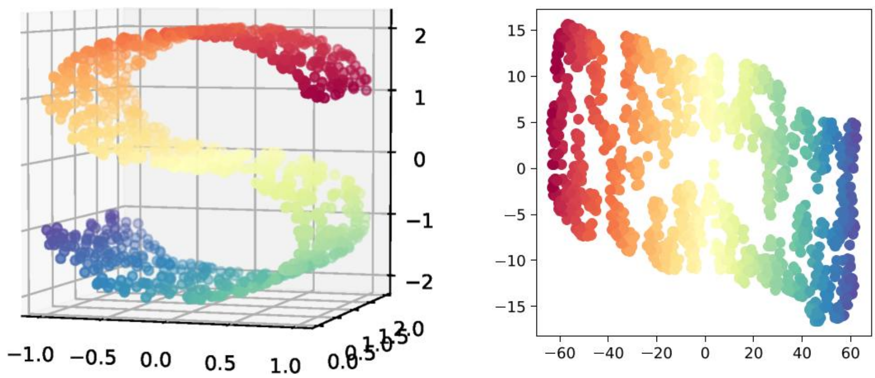


***Fig1.*** *Left panel: sample three-dimensional set of labeled data. Right panel: the same data set reduced into two-dimensional space by t-SNE algorithm.*

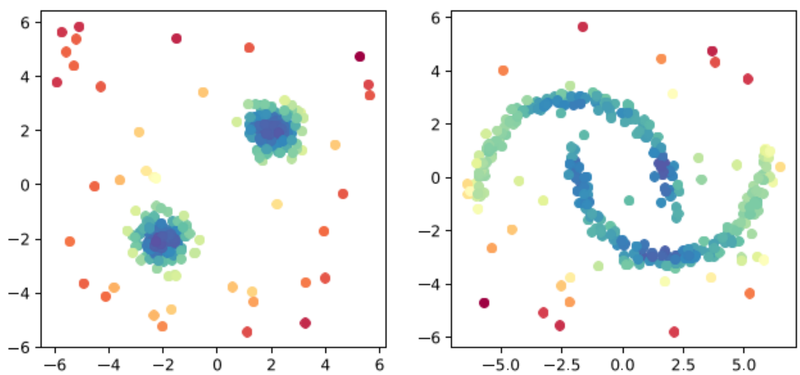


***Fig2.*** *Isolation forest applied to the two different sample data sets (left panel and right panel respectively). Redder points are ranked as anomalies.*

### 3.2. Dimensionality Reduction

Each object has its own flux scale due to the different origin and different distance. So, before the dimensionality reduction procedure we normalized each vector of 363 photometrical points by its maximum value and used the maximum value as one more feature. Then, we applied t-SNE [21] for dimensionality reduction of the data with 374 features: we obtained 7 feature reduced data sets: from 2 to 8 features.

In Fig. 1 we show t-SNE applied to sample data set in three-dimensional space. One may see that t-SNE is a nonlinear dimensionality reduction technique keeping vicinity of adjacent points.

***Table1.*** *List of found anomalies.*

| Name | Coordinates | Object type | Ref. |
|---|---|---|---|
| SN2016bln | 13 34 45.49 +13 51 14.3 | Ia-91T | [7] |
| SN2013cv | 16 22 43.16 +18 57 35.6 | SN Ia-pec | [35, 5] |
| SN1000+0216 | 10 00 05.87 +02 16 23.6 | SLSN | [8] |
| SN2006kg | 01 04 16.98 +00 46 08.9 | AGN | [4, 34, 28] |
| Gaia16aye | 19 40 01.13 +30 07 53.4 | Binary microlensing event | [1,33] |

### 3.3. Isolation Forest

Isolation forest is an ensemble of random isolation trees. Each isolation tree is a space partitioning tree similar to a widely-know Kd-tree. However, in contrast to Kd-tree, space coordinate (a feature) and a split value are selected at random for every node of the isolation tree. This algorithm leads to an unbalanced tree unusable for spatial search, but the tree has the following important property. A path distance between the root and a leaf is shorter on average for points distanced in space from “normal” data. This allows us to construct enough random trees to estimate average root-leaf path distance for every data sample that we have, and then rank the data samples based on the path length.

In Fig. 2 we show isolation forest applied to the different sample data sets. Note that the major advantage of the isolation forest is that it doesn’t make any assumptions on normal data distribution. At the left panel of Fig. 2 we could fit the data by two normal probability distribution function and then find outliers. This approach fails for the right panel of Fig. 2 where the isolation forest still succeeded.

We run the Isolation Forest algorithm on each data set (see Sect. 3.2) and obtained a list of anomalies.

## 4. Results

We visually inspected $\sim 100$ outliers among a total 1999 objects. Using the publicly available sources we checked what kind of astrophysical objects they are. The most prominent outliers are listed in Table 1 and described below, the rest are still being studied.

### 4.1. Peculiar SNe Ia

Type Ia supernova phenomenon is an explosion of a carbon-oxygen white dwarf that exceed the Chandrasekhar limit either by matter accretion from a companion star or by merging with another white dwarf [32, 14, 31]. SNe Ia are used as universal distance ladder since their luminosity at maximum light is approximately the same. However, SNe Ia can be divided by subtypes and not all of them are suitable for cosmology.

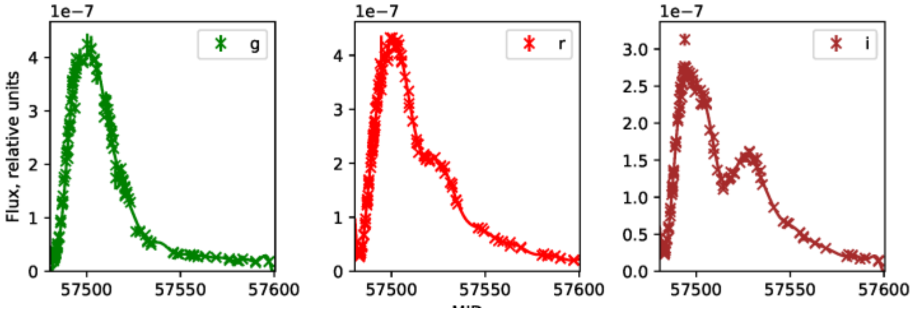


***Fig3***. *Light curves in gri filters of SN Ia-91T 2016bln [23].*

SN2016bln [7], classified by our code as anomaly, belongs to the so-called 1991T-like-supernovae subtype (see Fig. 3). SNe Ia-91T are characterized by higher peak luminosity and broader LCs than a normal SN Ia, and different early spectrum evolution.

Another novelty is SN2013cv ([35], see Fig. 4). This peculiar supernova has large peak optical and UV luminosity and show an absence of iron absorption lines in the early spectra. Cau at al. suggest that SN2013cv is an intermediate case between the normal and super-Chandrasekhar events [5].

### 4.2. Superluminous SNe

Superluminous SNe (SLSN) are supernovae with an absolute peak magnitude $M < -21$ mag in any band. According to [10] SLSN can be divided into three broad classes: SLSN-I without hydrogen in their spectra, hydrogen-rich SLSN-II that often show signs of interaction with circum-stellar material (CSM), and finally, SLSN-R, a rare class of hydrogen-poor events with slowly evolving LCs, powered by the radioactive decay of $^{56}Ni$.

SN 1000+0216 (Fig. 5) was discovered in the framework of the Canada-France-Hawaii Telescope Legacy Survey Deep Fields and has a redshift z = 3.9. It may be an example of a pulsational pair-instability SN or a SLSN-II which extreme optical emission is explained by the strong interaction between the expanding ejecta and massive CSM [8].

### 4.3. AGN

SN2006kg was erroneously classified as Type II supernovae ([4], see Fig. 6). The following studies identified it as an active galactic nucleus (AGN [34, 28]).

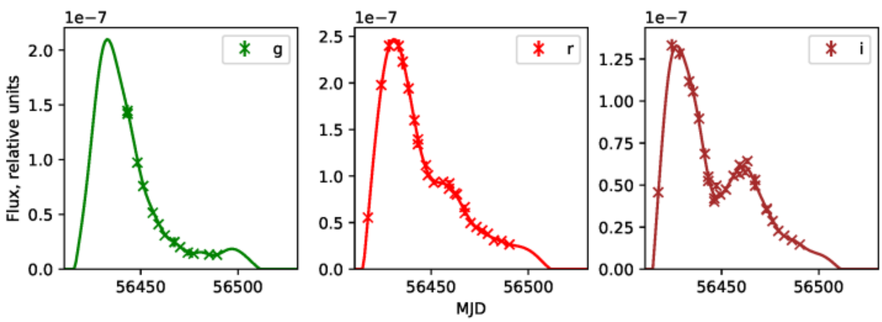


***Fig4.*** *Light curves in gri filters of peculiar SN2013cv [5, 34].*

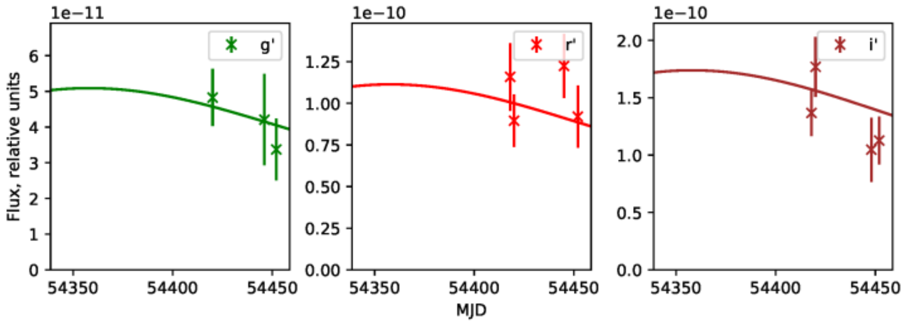


***Fig5***. *Light curves in gri filters of superluminous SN1000+0216 [8].*

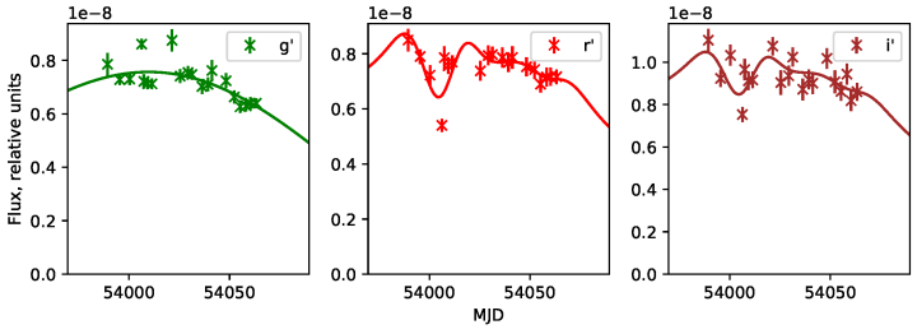


***Fig6.*** *Light curves in gri filters of SN2006kg [28].*

## 4.4. Binary microlensing event

Gaia16aye [1] is an object with the most non-SN behaviour in our set of outliers (Fig. 7). In [33] it was reported that Gaia16aye is a binary microlensing event – gravitational microlensing by binary systems — the first ever discovered towards the Galactic Plane.

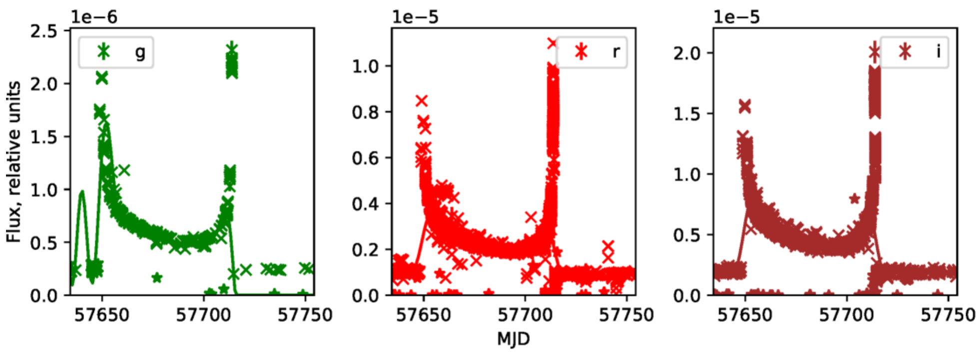


***Fig7.*** *Light curves in gri filters of binary microlensing event Gaia16aye (http://gsaweb.ast.cam.ac.uk/alerts/alert/Gaia16aye/followup)*

## 5. Conclusions

The development of large synoptic sky surveys has led to a discovery of huge number of supernovae and supernova candidates. Among the SN discovered every year, only 10% have spectroscopic confirmation. The amount of astronomical data increases dramatically with time and already beyond human capabilities. While now community has dozens of thousands SN candidates, during ten-year survey Large Synoptic Sky Telescope (LSST, [20]) will discover over ten million supernovae (and only a small fraction of them will receive a spectroscopic confirmation). The LSST cadence will allow receiving the light curves for ∼ 105 SNe, but before these SNe will be used in any physical analysis, they must be classified by types. In order to process this information and to extract all possible knowledge, machine learning techniques become necessary. Such approach will allow not only to classify supernova candidates by known types, but to reveal other variable objects (novae, counterparts of GW alerts, kilonovae, GRB afterglows) that were mistakenly classified as SN and what is even more important to detect astronomical objects with strange physical properties – anomalies. Finding such objects (and then studying them more closely) is of high priority and one of the main aims of the current study.

We used the Isolation Forest algorithm to search the anomalies in the Open Supernova Catalog. During the data pre-processing we fitted the supernova LCs in three (*gri*) filters by Gaussian processes. The GP-MULTISTATE-KERNEL[d] (Kornilov et al. 2019, in prep.) was specially developed to introduce the correlation between the filters. As a result, we found ∼ 100 anomalies, among which peculiar Type Ia SNe, SLSN, AGN, binary microlensing event.

## Acknowledgements

M. Pruzhinskaya and M. Kornilov are supported by RFBR grant according to the research project 18-32-0042618 for anomaly analysis and LCs interpolation. K. Malanchev is supported by RBFR grant 18-502-12025 for preparing Open Supernova Catalog data. E. E. O. Ishida acknowledges support from CNRS 2017 MOMENTUM grant and Foundation for the advancement of theoretical physics and Mathematics “BASIS”. A. Volnova acknowledges support from RSF grant 18-12-00522 for analysis of interpolated LCs. We used the equipment funded by the Lomonosov Moscow State University Program of Development. K.L.M. and M.V.P. acknowledge the Lomonosov Moscow State University Program of Development ”Leading Science Schools of MSU: Physics of Stars, Relativistic Compact Objects and Galaxies” for LCs extrapolation.

## References

[1] V. Bakis, U. Burgaz, T. Butterley, J. M. Carrasco, V. S. Dhillon, M. Dominik, A. Floers, L. K. Hardy, G. Leto, S. P. Littlefair, J. R. Maund, P. Ochner, A. Pastorello, A. Piascik, L. Rhodes, R. Z. Sanchez, K. V. Sokolovsky, I. Steele, S. Taubenberger, G. Terreran, L. Tomasella, R. W. Wilson, L. Wyrzykowski, and A. M. Zubareva. Gaia16aye: a flaring object of uncertain nature in Cygnus. *The Astronomer's Telegram*, 9376, Aug. 2016.

[d] https://github.com/matwey/gp-multistate-kernel

[2] N. M. Ball and R. J. Brunner. Data Mining and Machine Learning in Astronomy. International Journal of Modern Physics D, **19**, 1049–1106,2010.

[3] D. Baron and D. Poznanski. The weirdest SDSS galaxies: results from an outlier detection algorithm. MNRAS, **465,** 4530–4555, Mar. 2017.

[4] B. Bassett, A. Becker, H. Brewington, C. Choi, D. Cinabro, F. Dejongh, J. Dembicky, D. L. Depoy, B. Dilday, M. Doi, J. Frieman, P. Garnavich, M. Harvanek, C. Hogan, J. Holtzman, M. Im, S. Jha, K. Konishi, J. Krzesinski, H. Lampeitl, R. Kessler, B. Ketzeback, D. Long, O. Malanushenko, J. Marriner, D. McGinnis, R. McMillan, G. Miknaitis, T. Morokuma, R. Nichol, K. Pan, J. L. Prieto, M. Richmond, A. Riess, R. Romani, M. Sako, G. Saurage, D. Schneider, M. Smith, S. Snedden, N. Takanashi, K. Tokita, K. van der Heyden, S. Watters, C. Wheeler, N. Yasuda, C. Zheng, A. Aragon-Salamanca, M. Bremer, M. Turatto, A. Goobar, J. Sollerman, P. Ruiz-Lapuente, F. Castander, A. Romer, C. Collins, J. Lucey, A. Edge, M. Stritzinger, U. Hopp, W. Kollatschny, J. Eastman, L. Watson, R. Assef, K. Schlesinger, and A. Crotts. Supernovae 2006kg-2006lc. Central Bureau Electronic Telegrams, 688, Oct. 2006.

[5] [5] Y. Cao, J. Johansson, P. E. Nugent, A. Goobar, J. Nordin, S. R. Kulkarni, S. B. Cenko, O. D. Fox, M. M. Kasliwal, C. Fremling, R. Amanullah, E. Y. Hsiao, D. A. Perley, B. D. Bue, F. J. Masci, W. H. Lee, and N. Chotard. Absence of fast-moving iron in an intermediate type ia supernova between normal and super-chandrasekhar. The Astrophysical Journal, **823(2)**, 147, 2016.

[6] S. Cavuoti, M. Brescia, C. Tortora, G. Longo, N. R. Napolitano, M. Radovich, F. La Barbera, M. Capaccioli, J. T. A. de Jong, F. Getman, A. Grado, and M. Paolillo. Machine-learning-based photometric redshifts for galaxies of the ESO Kilo-Degree Survey data release 2. MNRAS, **452**, 3100–3105, Sept. 2015.

[7] S. B. Cenko, Y. Cao, M. Kasliwal, A. A. Miller, C. Fremling, M. West, M. Gregg, and S. R. Kulkarni. DCT and Gemini Spectroscopic Classification of AT 2016bln (=iPTF 16abc). The Astronomer's Telegram, 8909, Apr. 2016.

[8] J. Cooke, M. Sullivan, A. Gal-Yam, E. J. Barton, R. G. Carlberg, E. V. Ryan-Weber, C. Horst, Y. Omori, and C. G. D´ıaz. Superluminous supernovae at redshifts of 2.05 and 3.90. Nature, **491,** 228–231, Nov. 2012.

[9] C. Criscia, B. Ghattasb, and G. Pererac. A review of supervised machine learning algorithms and their applications to ecological data. Ecological Modelling, **240,** 113–122, 2012.

[10] A. Gal-Yam. Luminous Supernovae. Science, **337**, 927, Aug. 2012.

[11] J. Guillochon, J. Parrent, L. Z. Kelley, and R. Margutti. An Open Catalog for Supernova Data. ApJ, **835**, 64, Jan. 2017.

[12] M. Henrion, D. J. Hand, A. Gandy, and D. J. Mortlock. Casos: a subspace method for anomaly detection in high dimensional astronomical databases. Statistical Analysis and Data Mining: The ASA Data Science Journal, **6(1)**, 53–72.

[13] H. Hildebrandt, S. Arnouts, P. Capak, L. A. Moustakas, C. Wolf, F. B. Abdalla, R. J. Assef, M. Banerji, N. Benítez, G. B. Brammer, T. Budavári, S. Carliles, D. Coe, T. Dahlen, R. Feldmann, D. Gerdes, B. Gillis, O. Ilbert, R. Kotulla, O. Lahav, I. H. Li, J.-M. Miralles, N. Purger, S. Schmidt, and J. Singal. PHAT: PHoto-z Accuracy Testing. A&A, **523**, A31, Nov. 2010.

[14] I. Iben, Jr. and A. V. Tutukov. Supernovae of type I as end products of the evolution of binaries with components of moderate initial mass (M not greater than about 9 solar masses). ApJS, **54**, 335–372, Feb. 1984.

[15] E. E. O. Ishida and R. S. de Souza. Kernel PCA for Type Ia supernovae photometric classification.

MNRAS, **430**, 509–532, Mar. 2013.

[16] R. Kessler, B. Bassett, P. Belov, V. Bhatnagar, H. Campbell, A. Conley, J. A. Frieman, A. Glazov, S. Gonz´alez-Gait´an, R. Hlozek, S. Jha, S. Kuhlmann, M. Kunz, H. Lampeitl, A. Mahabal, J. Newling, R. C. Nichol, D. Parkinson, N. Sajeeth Philip, D. Poznanski, J. W. Richards, S. A. Rodney, M. Sako, D. P. Schneider, M. Smith, M. Stritzinger, and M. Varughese. Results from the Supernova Photometric Classification Challenge. PASP, **122**, 1415, Dec. 2010.

[17] M. W. Libbrecht and W. S. Noble. Machine learning applications in genetics and genomics. Nature Reviews Genetics, **16**, 321–332, 2015.

[18] F. T. Liu, K. M. Ting, and Z.-H. Zhou. Isolation-based anomaly detection. ACM Trans. Knowl. Discov. Data, **6(1)**, 3:1–3:39, Mar. 2012.

[19] M. Lochner, J. D. McEwen, H. V. Peiris, O. Lahav, and M. K. Winter. Photometric Supernova Classification with Machine Learning. ApJS, **225**, 31, Aug. 2016.

[20] LSST Science Collaboration, P. A. Abell, J. Allison, S. F. Anderson, J. R. Andrew, J. R. P. Angel, L. Armus, D. Arnett, S. J. Asztalos, T. S. Axelrod, and et al. LSST Science Book, Version 2.0. ArXiv e-prints, Dec. 2009.

[21] L. v. d. Maaten and G. Hinton. Visualizing data using t-sne. Journal of machine learning research, **9(Nov)**, 2579–2605, 2008.

[22] S. D. McCauliff, J. M. Jenkins, J. Catanzarite, C. J. Burke, J. L. Coughlin, J. D. Twicken, P. Tenenbaum, S. Seader, J. Li, and M. Cote. Automatic Classification of Kepler Planetary Transit Candidates. ApJ, **806**, 6, June 2015.

[23] A. A. Miller, Y. Cao, A. L. Piro, N. Blagorodnova, B. D. Bue, S. B. Cenko, S. Dhawan, R. Ferretti, O. D. Fox, C. Fremling, A. Goobar, D. A. Howell, G. Hosseinzadeh, M. M. Kasliwal, R. R. Laher, R. Lunnan, F. J. Masci, C. McCully, P. E. Nugent, J. Sollerman, F. Taddia, and S. R. Kulkarni. Early Observations of the Type Ia Supernova iPTF 16abc: A Case of Interaction with Nearby, Unbound Material and/or Strong Ejecta Mixing. ApJ, **852**, 100, Jan. 2018.

[24] I. Nun, K. Pichara, P. Protopapas, and D.-W. Kim. Supervised Detection of Anomalous Light Curves in Massive Astronomical Catalogs. ApJ, **793**, 23, Sept. 2014.

[25] A. Pastorello, X.-F. Wang, F. Ciabattari, D. Bersier, P. A. Mazzali, X. Gao, Z. Xu, J.-J. Zhang, S. Tokuoka, S. Benetti, E. Cappellaro, N. Elias-Rosa, A. Harutyunyan, F. Huang, M. Miluzio, J. Mo, P. Ochner, L. Tartaglia, G. Terreran, L. Tomasella, and M. Turatto. Massive stars exploding in a He-rich circumstellar medium - IX. SN 2014av, and characterization of Type Ibn SNe. MNRAS, **456**, 853–869, Feb. 2016.

[26] K. A. Pearson, L. Palafox, and C. A. Griffith. Searching for exoplanets using artificial intelligence. MNRAS, **474**, 478–491, Feb. 2018.

[27] U. Rebbapragada, P. Protopapas, C. E. Brodley, and C. Alcock. Finding anomalous periodic time series. Machine Learning, **74(3)**, 281–313, Mar 2009.

[28] M. Sako, B. Bassett, A. C. Becker, P. J. Brown, H. Campbell, R. Wolf, D. Cinabro, C. B. D'Andrea, K. S. Dawson, F. DeJongh, D. L. Depoy, B. Dilday, M. Doi, A. V. Filippenko, J. A. Fischer, R. J. Foley, J. A. Frieman, L. Galbany, P. M. Garnavich, A. Goobar, R. R. Gupta, G. J. Hill, B. T. Hayden, R. Hlozek, J. A. Holtzman, U. Hopp, S. W. Jha, R. Kessler, W. Kollatschny, G. Leloudas, J. Marriner, J. L. Marshall, R. Miquel, T. Morokuma, J. Mosher, R. C. Nichol, J. Nordin, M. D. Olmstead, L. Östman, J. L. Prieto, M. Richmond, R. W. Romani, J. Sollerman, M. Stritzinger, D. P. Schneider, M. Smith, J. C. Wheeler, N. Yasuda, and C. Zheng. The Data Release of the Sloan Digital Sky Survey-II Supernova Survey. PASP, **130(6)**, 064002, June 2018.

[29] S. E. Thompson, F. Mullally, J. Coughlin, J. L. Christiansen, C. E. Henze, M. R. Haas, and C. J. Burke. A Machine Learning Technique to Identify Transit Shaped Signals. ApJ, **812**, 46, Oct. 2015.

[30] M. Vidyasagar. Identifying predictive features in drug response using machine learning: Opportunities and challenges. Annual Review of Pharmacology and Toxicology, **55(1)**, 15–34, 2015. PMID: 25423479.

[31] R. F. Webbink. Double white dwarfs as progenitors of R Coronae Borealis stars and Type I supernovae. ApJ, **277**, 355–360, Feb. 1984.

[32] J. Whelan and I. Iben, Jr. Binaries and Supernovae of Type I. ApJ, **186**, 1007–1014, Dec. 1973.

[33] L. Wyrzykowski, G. Leto, G. Altavilla, V. Bakis, N. Britavskiy, U. Burgaz, T. Butterley, J. M. Carrasco, V. S. Dhillon, M. Dominik, A. Gomboc, L. K. Hardy, S. P. Littlefair, J. R. Maund, A. Piascik, L. Rhodes, R. Z. Sanchez, K. V. Sokolovsky, I. Steele, R. W. Wilson, A. Hamanowicz, P. Mroz, M. Pawlak, K. Rybicki, M. Sitek, P. Mikolajczyk, Z. Kolaczkowski, R. Street, P. Bendjoya, V. Bozza, J. Dziedzic, K. Niczyj, R. Nowicki, and M. Porebski. Gaia16aye is a binary microlensing event and is crossing the caustic again. The Astronomer's Telegram, **9507**, Sept. 2016.

[34] O. Yaron and A. Gal-Yam. WISeREP — An Interactive Supernova Data Repository. PASP, **124**, 668, July 2012.

[35] L. Zhou, X. Wang, K. Zhang, J. Chen, J. Liang, T. Zhang, X. Zhou, F. Huang, X. sZhao, X. Wang, T. Zhang, D. D. Balam, M. L. Graham, and E. Y. Hsiao. Supernova 2013cv = Psn J16224316+1857356. Central Bureau Electronic Telegrams, **3543**, June 2013.