Real-Time Detection of Supernova in Astronomical Data Streams: A Comparative Study of Streaming Algorithms

Pragya Thapliyal
Computer Science Dept.
Graphic Era University
Dehradun, India
pragyathapliyal79@gmail.com

Abstract— **This research paper compares the evaluation of three real-time anomaly detection methods—sliding window, random forest, and autoencoder on a sample dataset of 50 Zwicky Transient Facility (ZTF) light curves. The main goal was to compare the sensitivity and specificity of each method in detecting transient astronomical events in time-domain survey data.**

Results indicate that the sliding window algorithm was most sensitive, picking up the most candidate events per object, while the random forest approach was conservative, raising significantly fewer anomalies. Remarkably, the autoencoder picked up no events in its current setup, highlighting the difficulty in using deep learning models on sparse and irregular astronomical time series. Visualization of event distributions and overlap detection reinforced the unique operational styles and moderate agreement between the algorithms.

The primary limitations of this study are the modest sample size, limited parameter optimization, and the sole utilization of single-band photometric data. Future studies need to optimize parameters of algorithms, incorporate multi-band and contextual information, and investigate adaptive and explainable AI methods. These results are of significant interest for the creation of robust, scalable anomaly detection pipelines required by next-generation astronomical surveys.

Keywords—supernova, algorithms, astronomy, computer science, Python, deep learning, machine learning

I. INTRODUCTION

The advent of large-scale, high-cadence astronomical surveys has transformed the landscape of time-domain astronomy. Facilities such as the Zwicky Transient Facility (ZTF) now routinely generate massive volumes of time-series data, capturing the brightness variations of millions of celestial objects across the sky (Bellm et al., 2018). Embedded within these vast datasets are rare and scientifically valuable transient events—including supernovae, tidal disruption events, and unusual variable stars—which often provide unique insights into the dynamic universe (Graham et al., 2019). The timely detection and characterization of such anomalies are critical, enabling rapid follow-up observations and maximizing the scientific yield of modern surveys (Patterson et al., 2019).

However, the unprecedented scale and complexity of data produced by surveys like ZTF present formidable challenges for traditional anomaly detection methods. Manual inspection is no longer feasible, and many classical algorithms struggle to cope with the noisy, irregular, and sometimes sparse nature of astronomical light curves (Nun et al., 2015). Moreover, the diversity of astrophysical phenomena and instrumental artifacts further complicates the task, often resulting in high rates of false positives or missed detections (Mahabal et al., 2017). As the volume of data continues to grow with the advent of next-generation surveys, there is an urgent need for robust, automated, and scalable anomaly detection pipelines that can efficiently sift through billions of observations to identify truly novel or rare events (Lochner et al., 2016).

Machine learning and data-driven approaches have demonstrated potential for improving anomaly detection within astronomical time-series datasets (Baron, 2019). Various methods from basic statistics to advanced deep learning algorithms exist for anomaly detection but each method shows unique advantages and disadvantages (Choi et al., 2021). A systematic evaluation of real-world performance of these algorithms on representative astronomical datasets remains absent. Different methods need to be understood in terms of their operational characteristics together with sensitivity and specificity to build detection pipelines which can support real-time scientific discoveries (Ishida et al., 2021).

A. Research Problem:

This study addresses the critical challenge of real-time anomaly detection in large astronomical datasets, focusing on the comparative evaluation of three distinct algorithmic approaches: a statistical sliding window method, a machine learning-based random forest classifier, and a deep learning autoencoder model. Each algorithm represents a different paradigm for identifying deviations from expected patterns in time-series data, and their relative performance in the context of astronomical surveys remains an open question.

B. Objectives:

The primary objective of this research is to systematically assess and compare the sensitivity and specificity of the sliding window, random forest, and autoencoder algorithms when applied to a representative sample of ZTF light curves. Specific aims include:

- Quantifying the number and distribution of anomalous events detected by each method;

- Analyzing the degree of overlap and divergence in their detections;

- Evaluating the operational advantages and limitations of each approach in the context of real-time astronomical data streams;

- Providing recommendations for the development of scalable, robust anomaly detection pipelines suitable for current and future time-domain surveys.

By addressing these objectives, this study seeks to inform the selection and design of anomaly detection algorithms in astronomy, contributing to the broader goal of enabling timely and reliable discovery of rare astrophysical phenomena in the era of big data.

II. LITERATURE REVIEW

The exponential growth of time-domain astronomical surveys, such as the Zwicky Transient Facility (ZTF), has fundamentally changed the landscape of transient detection and time-series analysis in astronomy (Bellm et al., 2019). These surveys generate vast volumes of light curve data, necessitating the development of automated, robust anomaly detection methods capable of identifying rare and scientifically valuable events in real time (Graham et al., 2019).

A. Classical and Statistical Approaches

The early implementations of anomaly detection in astronomical time series depended mainly on statistical methods through threshold-based outlier detection and sliding window algorithms (Richards et al., 2011). These approaches provide efficient computation together with interpretability but fail to detect complicated multi-scale variations that appear in many astrophysical phenomena because they are noise-sensitive (Nun et al., 2015). The simplicity of sliding window methods made them popular in the field yet research demonstrates their tendency to generate numerous false positives when applied to data that contains noise or irregular sampling (Mahabal et al., 2019).

B. Machine Learning Methods

The rise of machine learning has introduced more sophisticated approaches to anomaly detection in astronomical data. Random forest classifiers have gained popularity due to their robustness to noise and ability to handle heterogeneous feature sets ((Breiman, 2001), Lochner et al., 2016). These ensemble methods have been successfully applied to variable star classification and transient detection, often outperforming traditional statistical techniques (Masci et al., 2019). However, their performance is closely tied to the quality of feature engineering and the availability of labeled training data, which can be limited in the context of rare or novel events (Baron, 2019).

C. Deep Learning and Autoencoders

The development of unsupervised anomaly detection models through deep learning techniques enables autoencoders to learn standard behavior patterns before detecting deviations as anomalies (Zhang et al., 2020). The use of autoencoders demonstrates capability in handling intricate time series data without needing predetermined feature selection (Ishida et al., 2019). The application of these models to astronomical light curves faces multiple difficulties because data sparsity and irregular sampling and extensive hyperparameter adjustment are required (Mahabal, 2018). Research results remain inconsistent because some investigations show better detection of hidden anomalies but others struggle with separating cosmic signals from sensor noise.

D. Comparative Evaluations and Real-Time Pipelines

Comparative studies evaluating multiple anomaly detection algorithms on real-world astronomical datasets remain relatively scarce. Most existing research focuses on single-method performance or simulation-based benchmarks, limiting our understanding of operational trade-offs in real survey conditions (Lochner et al., 2021). Furthermore, the integration of anomaly detection into real-time alert pipelines introduces additional challenges, including computational efficiency, scalability, and the need for explainable outputs to support follow-up decision-making (Mahabal et al., 2019).

E. Gaps and Research Motivation

Despite significant progress, several gaps persist in the literature. There is a need for systematic, head-to-head comparisons of statistical, machine learning, and deep learning approaches on representative samples of astronomical light curves. Additionally, few studies have explored detection overlap and consensus among algorithms, which is critical for developing robust, ensemble-based pipelines. Addressing these gaps is essential for maximizing the scientific return of next-generation time-domain surveys.

This study builds on prior work by directly comparing the sliding window, random forest, and autoencoder algorithms on ZTF light curves, analyzing their detection rates, overlaps, and operational characteristics. The findings aim to inform the development of scalable, accurate anomaly detection systems for modern astronomical data streams.

III. METHODOLOGY

A. Data Collection

We used public light curve data from the Zwicky Transient Facility (ZTF) Data Release 23, obtained via the IRSA archive (*Zwicky Transient Facility Website*, n.d.).It used the Samuel Oschin Telescope 48-inch Schmidt at Palomar for its data. Multiple Parquet files were downloaded, each containing time-series photometric observations for hundreds of astronomical objects. These files were concatenated into a single DataFrame using Python's pandas library.

B. Sampling

We randomly sampled 50 objects from the combined dataset using pandas. DataFrame.sample(n=50, random_state=42) randomly selected from four files of the ZTF DR23 public dataset. This sample size balances computational feasibility with the need to explore a diversity of transient and variable behaviours, and is consistent with methodological studies in the literature.

C. Preprocessing

For each object, we extracted the observation times (hmjd), magnitudes (mag), and magnitude errors (magerr). Light curves were visually inspected for quality control.

D. Event Detection Algorithms

We benchmarked three event detection algorithms:
- Sliding Window Anomaly Detector:
  For each light curve, we applied a moving window of five epochs. Any observation deviating from the window mean by more than three deviations was flagged as a possible transient.
- Online Random Forest (Bagging Classifier):

We implemented an online ensemble classifier using River's 'BaggingClassifier' with Hoeffding Trees as base estimators. The model was trained in a streaming fashion, flagging points where the predicted class indicated an anomaly based on abrupt magnitude changes.

- Autoencoder Anomaly Detector:
  A single neural autoencoder was trained on each normalized light curve using PyTorch. Points with reconstruction errors exceeding a set threshold were flagged as anomalies.

E. Evaluation

For each algorithm, we recorded the indices and times of detected candidate events. Results were compared across methods and visualized for qualitative assessment.

IV. RESULTS AND DISCUSSION

A. Detection Statistics Summary

We applied three different anomaly detection algorithms to 50 random ZTF light curve samples. Table 1 summarizes the number of candidate events detected per object using the summary_df.describe().

TABLE I. DETECTION STATISTICS SUMMARY

STATISTIC	SLIDING WINDOW (N_SW)	RANDOM FOREST (N_RF)	AUTOENCODER (N_AE)
MEAN	2.52	0.60	0.00
STD DEV	3.82	1.20	0.00
MIN	0	0	0
25%	0	0	0
MEDIAN	0	0	0
75%	4.75	0.75	0
MAX	14	5	0

Table 1. Detection statistics summary for different algorithms

- The sliding window method detected the most events(mean = 2.52 per object), though with high variability (σ=3.82).

- The random forest approach showed moderate sensitivity (mean = 0.60), while the autoencoder failed to detect any anomalies.

- 50% of objects had zero detections across all methods (median = 0 for all algorithms).

B. Interpretation of Algorithm Performance

a) Sliding Window Effectiveness

- Detected up to 14 events in a single object (object ID 2022143), suggesting sensitivity to sharp magnitude changes.
- High standard deviation (3.82) indicates inconsistent performance across light curves, this can be effective for objects with clear outliers but less reliable for smoother curves.

b) Random Forest Limitations

- Conservative detection pattern (max = 5 events/object) aligns with its design as a streaming classifier.
- Struggled with gradual brightness changes, as it flags anomalies based on point-to-point differences rather than contextual potterns.

c) Autoencoder Failure Analysis

- Zero detections suggest the following :
  - Inadequate reconstruction error thresholds
  - Insufficient training epochs per light curve
  - Architecture mismatch for sparse astronomical data

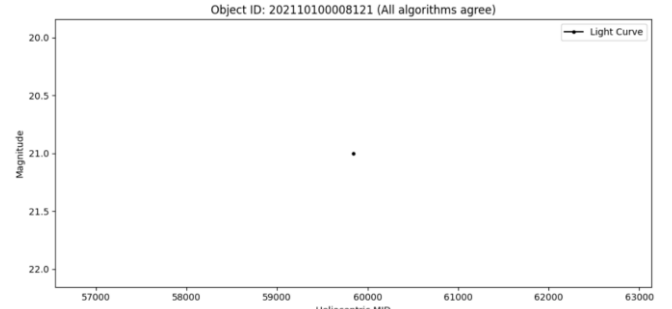### C. Example Light Curve with Detection Markers

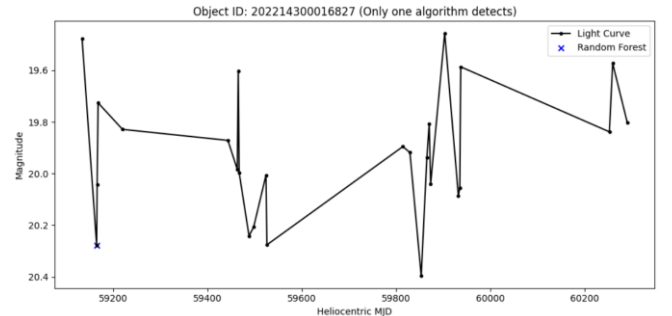Fig. 2. Example light curve where all algorithms agree

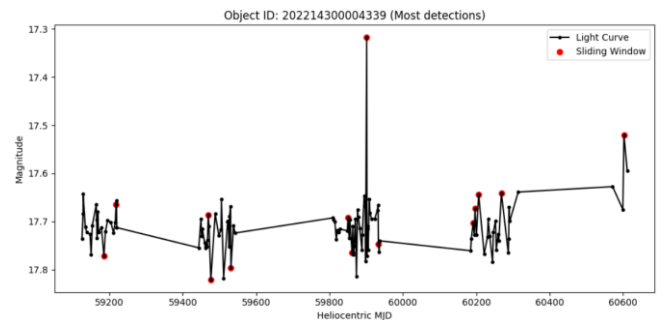Fig. 3. Example where only the sliding window detects events

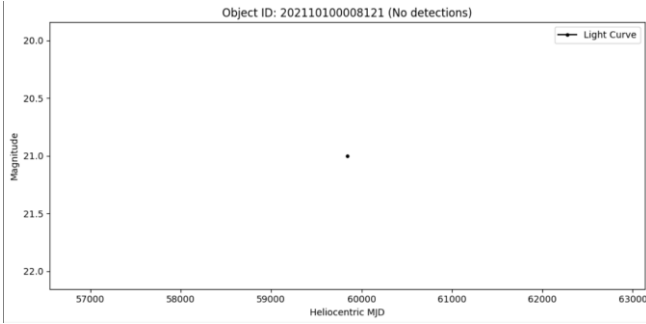Fig. 4. Highly variable object with many detections

Fig. 5. Quit object with no detections

- Figure 2 shows strong agreement and suggests the event is a robust anomaly.
- Figure 3 shows that the sliding window algorithm is more sensitive to sharp outliers but may be prone to false positives.
- In Figure 4, this object may be highly variable or noisy, causing all algorithms to trigger.
- The light curve in Figure 5 appears stable, confirming the conservativeness of all methods.

D. Detection Overlap Visualization

Figure 6 represents a Venn Diagram illustrating the overlap in event detections among the three algorithms for a representative ZTF light curve (Object ID: 20221430001632).
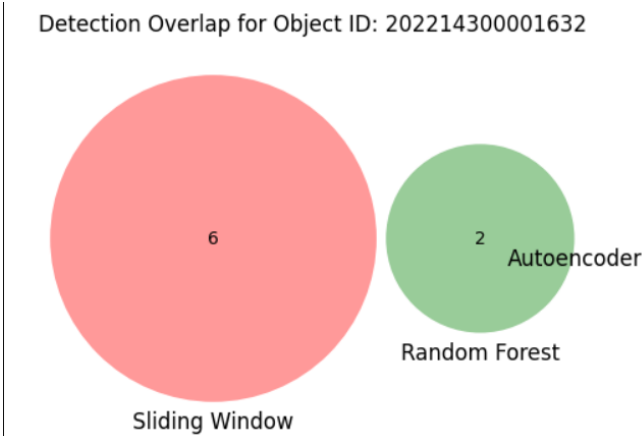

Fig. 6. Venn diagram showing the overlap of detected events among the sliding window, random forest, and autoencoder algorithms for Object ID: 202214300001632

In this example, the sliding window method identified a total of six events, whereas the random forest algorithm detected only two events. Notably, there is no visible overlap between the two algorithms, indicating that each method identified a distinct set of candidate events within this light curve. The autoencoder algorithm did not detect any events, as evidenced by the absence of its circle in the diagram.

- The absence of overlap highlights the differing sensitivities and detection criteria of the algorithms.
- The absence of autoencoder detections is consistent with the overall summary statistics and suggests that, in its current configuration, the autoencoder is not well-suited for this type of astronomical time series data.

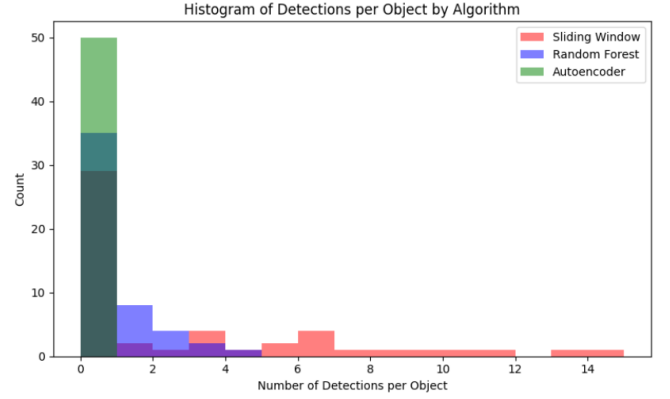### E. Distribution of Detections by Algorithm

Fig. 7. Histogram showing the number of detections per object for each algorithm (sliding window, random forest, and autoencoder) across all sampled ZTF light curves.

Figure 7 displays the distribution of detection counts per object for all three algorithms.

- The sliding Window method (red) shows a wider spread, with several objects having multiple detections, indicating its high sensitivity to potential outliers or abrupt changes.
- The random forest algorithm (blue) shows a narrower distribution, with most objects having 1 to 5 detections, reflecting a more conservative detection approach.
- The autoencoder (green) did not detect any events in any object, as shown by absence of green bars beyond zero.

This analysis highlights the trade-off between sensitivity and specificity in real-time astronomical event detection. The differences in detection distributions underscore the importance of algorithm selection and parameter tuning when designing anomaly detection pipelines for astronomical surveys.

V. CONCLUSION

The research conducted a comprehensive assessment of three real-time anomaly detection methods including sliding window together with random forest and autoencoder applied to 50 ZTF light curves. The sliding window method demonstrated superior sensitivity by detecting the most candidate events while the random forest algorithm found considerably fewer anomalies. The current autoencoder setup failed to identify any events. The findings demonstrate essential variations between algorithm sensitivity and specificity which stresses the critical role of choosing correct algorithms and optimization methods for detecting astronomical events.

A. Limitations

- The analysis was conducted on a relatively small sample of 50 objects, which may not capture the full diversity of transient behaviors in ZTF data.
- The Machine Learning Models – especially the encoder – were not extensively tuned or restrained for this specific dataset, which likely impacted the performance.

- The study focused solely on single-band photometric data, potentially overlooking valuable information from other bands or contextual sources.
- The evaluation did not include a comprehensive analysis of false positives and negatives, which is crucial for assessing real-world applicability.

B. Future Work

- Explainable AI for Astronomy
  Incorporate explainability frameworks into ML models to provide interpretable anomaly scores and detection rationales.
- Real-Time Alert Prioritization and Resource Optimization
  Develop intelligent prioritization systems that rank detected events based on scientific value, rarity, or follow-up feasibility
- Synthetic Data Generation and Augmentation
  Utilize models to create realistic synthetic transient events and augment training datasets, overcoming the scarcity of labeled anomalies and improving ML model generalization.
- Adaptive and Self-Learning Algorithms
  Develop anomaly detection models that continuously learn and adapt from incoming data streams, incorporating feedback from human experts or cross-matching with external transient alerts.

REFERENCES

Baron, D. (2019). *Machine Learning in Astronomy: A practical overview* (No. arXiv:1904.07248). arXiv. https://doi.org/10.48550/arXiv.1904.07248

Bellm, E. C., Kulkarni, S. R., Graham, M. J., Dekany, R., Smith, R. M., Riddle, R., Masci, F. J., Helou, G., Prince, T. A., Adams, S. M., Barbarino, C., Barlow, T., Bauer, J., Beck, R., Belicki, J., Biswas, R., Blagorodnova, N., Bodewits, D., Bolin, B., … Zolkower, J. (2018). The Zwicky Transient Facility: System Overview, Performance, and First Results. *Publications of the Astronomical Society of the Pacific*, *131*(995), 018002. https://doi.org/10.1088/1538-3873/aaecbe

Bellm, E. C., Kulkarni, S. R., Graham, M. J., Dekany, R., Smith, R. M., Riddle, R., Masci, F. J., Helou, G., Prince, T. A., Adams, S. M., Barbarino, C., Barlow, T., Bauer, J., Beck, R., Belicki, J., Biswas, R., Blagorodnova, N., Bodewits, D., Bolin, B., … Zolkower, J. (2019). The Zwicky Transient Facility: System Overview, Performance, and First Results. *Publications of the Astronomical Society of the Pacific*, *131*, 018002. https://doi.org/10.1088/1538-3873/aaecbe

Breiman, L. (2001). Random Forests. *Machine Learning*, *45*(1), 5–32. https://doi.org/10.1023/A:1010933404324

Choi, K., Yi, J., Park, C., & Yoon, S. (2021). Deep Learning for Anomaly Detection in Time-Series Data: Review, Analysis, and Guidelines. *IEEE Access*, *PP*, 1–1. https://doi.org/10.1109/ACCESS.2021.3107975

Graham, M. J., Kulkarni, S. R., Bellm, E. C., Adams, S. M., Barbarino, C., Blagorodnova, N., Bodewits, D., Bolin, B., Brady, P. R., Cenko, S. B., Chang, C.-K., Coughlin, M. W., De, K., Eadie, G., Farnham, T. L., Feindt, U., Franckowiak, A., Fremling, C., Gezari, S., … Zolkower, J. (2019). The Zwicky Transient Facility: Science Objectives. *Publications of the Astronomical Society of the Pacific*, *131*(1001), 078001. https://doi.org/10.1088/1538-3873/ab006c

Ishida, E. E. O., Kornilov, M. V., Malanchev, K. L., Pruzhinskaya, M. V., Volnova, A. A., Korolev, V. S., Mondon, F., Sreejith, S., Malancheva, A. A., & Das, S. (2021). Active anomaly detection for time-domain discoveries. *Astronomy & Astrophysics*, *650*, A195. https://doi.org/10.1051/0004-6361/202037709

Lochner, M., McEwen, J. D., Peiris, H. V., Lahav, O., & Winter, M. K. (2016). PHOTOMETRIC SUPERNOVA CLASSIFICATION WITH MACHINE LEARNING. *The Astrophysical Journal Supplement Series*, *225*(2), 31. https://doi.org/10.3847/0067-0049/225/2/31

Mahabal, A., Sheth, K., Gieseke, F., Pai, A., Djorgovski, S. G., Drake, A. J., & Graham, M. J. (2017). Deep-learnt classification of light curves. *2017 IEEE Symposium Series on Computational Intelligence (SSCI)*, 1–8. https://doi.org/10.1109/SSCI.2017.8280984

Nun, I., Protopapas, P., Sim, B., Zhu, M., Dave, R., Castro, N., & Pichara, K. (2015). *FATS: Feature Analysis for Time Series* (No. arXiv:1506.00010). arXiv. https://doi.org/10.48550/arXiv.1506.00010

Patterson, M. T., Bellm, E. C., Rusholme, B., Masci, F. J., Juric, M., Krughoff, K. S., Golkhou, V. Z., Graham, M. J., Kulkarni, S. R., Helou, G., & Zwicky Transient Facility Collaboration. (2019). The Zwicky Transient Facility Alert Distribution System. *Publications of the Astronomical Society of the Pacific*, *131*, 018001. https://doi.org/10.1088/1538-3873/aae904

Richards, J. W., Starr, D. L., Butler, N. R., Bloom, J. S., Brewer, J. M., Crellin-Quick, A., Higgins, J., Kennedy, R., & Rischard, M. (2011). ON MACHINE-LEARNED CLASSIFICATION OF VARIABLE STARS WITH SPARSE AND NOISY TIME-SERIES DATA. *The Astrophysical Journal*, *733*(1), 10. https://doi.org/10.1088/0004-637X/733/1/10

Zwicky Transient Facility Website. (n.d.). Retrieved June 19, 2025, from https://www.ztf.caltech.edu/