Interpretable, Redshift-Free Photometric Typing of Type Ia Supernovae for the Rubin LSST

Anurag Garg[1], Artyom Kabanov[2, 3], Mikhail Mozikov[4, 5], Ilya Makarov[3, 4, 5]

[1]Ministry of Education, Abu Dhabi, UAE
[2]AI Talent Hub, ITMO University, Russia
[3]Research Center of the Artificial Intelligence Institute, Innopolis University, Innopolis, Russia
[4]AI Research Institute, Russia
[5]ISP RAS, Moscow, Russia
agkabanov@niuitmo.ru

Abstract

Time-domain surveys increasingly outpace spectroscopic follow-up, making fast and accurate photometric classification of supernovae a critical bottleneck for cosmology and transient science. We present a lightweight, interpretable machine-learning pipeline for *redshift-free* identification of Type Ia supernovae tailored for operational use with the Rubin Observatory's Legacy Survey of Space and Time (LSST). The approach couples a broker-friendly, light-curve preprocessing stage with a tuned Random Forest ensemble (Random Search and Bayesian Optimization). Because the task is strongly class-imbalanced and errors have asymmetric costs, we prioritize precision–recall metrics and report the SPCC figure-of-merit (F1). On the post-challenge SPCC dataset of 21,319 simulated light curves, our final model (Bayesian optimization) attains F1 (SPCC) $= 0.941$, precision $= 0.958$, recall $= 0.924$, and ROC-AUC $= 0.886$, matching strong baselines while remaining transparent and resource-efficient. Requiring only multi-band photometry—no host information or redshifts—our method can triage alerts early, reduce spectroscopic load, and accelerate construction of high-purity SN Ia samples for distance-ladder and dark-energy analyses. We discuss integration into real-time broker pipelines and show that emphasizing interpretable ensembles and precision–recall metrics yields a robust, scalable path to AI-accelerated discovery in time-domain astronomy.

Introduction

Supernovae (SNe) are broadly divided into hydrogen-poor Type I and hydrogen-rich Type II (Minkowski 1979; Morrell et al. 2024). Thermonuclear Type Ia SNe are standardizable candles and underpin modern measurements of cosmological distances and expansion. As wide-field surveys increase event rates by orders of magnitude, spectroscopy cannot keep up (Bloom and Richards 2011); photometric typing must carry more of the load. While Bayesian hierarchical approaches can marginalize over uncertain types (Newling et al. 2011; Knights et al. 2013; Rubin et al. 2015), they still benefit from accurate probabilistic classifications.

Machine learning (ML) provides a practical route to scalable photometric classification and has seen growing success across surveys (Möller et al. 2024, 2016; Reza, Wang, and Hu 2025; Garretson et al. 2021; Möller et al. 2022; Dobryakov et al. 2021). However, many pipelines rely on host-galaxy information or redshifts, which are not always available or reliable in low S/N, early-time, or crowded-field conditions. We target the stricter and operationally valuable setting of redshift-free classification using photometry alone.

The Rubin LSST will produce multi-band time-series photometry and *millions of alerts per night*, which makes scalable photometric typing and broker triage operationally essential (Ivezić et al. 2019). Community-scale simulations such as PLAsTiCC established a standard benchmark for LSST-like photometric classification and informed feature/metric choices (Kessler et al. 2019).

Our goal is photometric typing of SNe Ia without any redshift or host-galaxy information. This choice mirrors operational constraints in Rubin/LSST alert brokers, where latency and missing host metadata are common. We therefore ask: how far can a carefully engineered classical model go under these constraints? This motivates a broker-ready, redshift-free classifier that consumes photometry alone and optimizes the SPCC F1.

Contributions.

- We analyze redshift-free photometric typing with no host-galaxy metadata and demonstrate strong performance with a tuned Random Forest on SPCC.
- We prioritize PR metrics and the SPCC figure-of-merit F1 (alongside Precision/Recall and ROC–AUC), which are more informative under class imbalance.
- Documented a simple, reproducible preprocessing recipe (time re-zeroing to first detection, sub-hour grouping, per-filter handling, and uncertainty propagation) suitable for alert-stream deployment.

Related Work

Early work established the spectroscopic Type I/II dichotomy (Minkowski 1979). For photometric typing, model-dependent features (e.g., SALT2 fits) combined with boosted trees achieved strong AUC on simulated datasets (Lochner et al. 2016; Guy et al. 2007). Gradient-boosted trees, CNNs, and sequence models have all been explored (Möller et al. 2016; Möller and de Boissière 2020; Garretson et al. 2021). Redshift-agnostic feature extraction using SALT2 and related formulations has proven particularly useful (Lochner et al. 2016). Large-scale LSST simulations and SNN-based approaches report near-saturated AUC on balanced tasks (Möller and de Boissière 2020). Recent work demonstrates the feasibility of training with simulations and predicting on real photometry using only light curves (Dobryakov et al. 2021).

Model-dependent light-curve fits such as SALT2 remain a de facto standard for SN Ia analyses (Guy et al. 2007), and recent revisions improve calibration surfaces and UV behavior (Taylor et al. 2021). For early-time operation and real-time streams, deep sequence models (e.g., RAPID) and production brokers (Fink, ALeRCE, ANTARES, Lasair) demonstrate scalable photometric classification on ZTF/LSST-like alerts (Möller et al. 2021; Förster et al. 2021; Smith et al. 2019; Bellm et al. 2019; Masci et al. 2019).

Rubin LSST and Dataset

The Vera C. Rubin Observatory's Legacy Survey of Space and Time (LSST) will repeatedly image a wide area of the southern sky with a 3.2 Gpix camera and a large field of view, producing time-series photometry in $griz$ bands (among others)(Ivezić et al. 2019). For development and benchmarking we use the post-challenge Supernova Photometric Classification Challenge (SPCC) dataset (Kessler et al. 2010), which provides 21,319 simulated supernova light curves with fluxes, uncertainties, band identifiers, sky positions, and Milky Way extinction.

We follow LSST system nomenclature and anticipated data products as summarized by Ivezić et al. (2019), and we reference PLAsTiCC for LSST-like simulation and benchmarking protocols (Kessler et al. 2019).

Preprocessing and Features

We follow a cosmology-aware preprocessing pipeline inspired by Charnock and Moss (2017). Times are re-zeroed to the day of first detection per object (no negative phases), close-in-time observations (within 1 hour) are grouped, and per-band sequences are constructed. We retain flux and corresponding uncertainty at each step and compute compact summary features per band and globally (e.g., number of detections, rise/decline proxies). This yields a low-dimensional, interpretable representation suitable for classical ML.

Although our features are model-agnostic, they are compatible with SALT2-derived summaries commonly used for SN Ia light curves (Guy et al. 2007; Taylor et al. 2021).

Preprocessing details. For each object, let t_i be the observation time (MJD), $b_i \in \{g, r, i, z\}$ the band, f_i the (calibrated) flux and σ_i its reported uncertainty. We apply a broker-friendly, minimal pipeline:

1. **Time re-zeroing.** Define $t_0 = \min_i t_i$ (first detection). Use $\tau_i = t_i - t_0 \geq 0$ to remove negative phases and align light curves at first detection.
2. **Sub-hour grouping.** Within each band b independently, sort observations by τ_i and form contiguous groups G such that consecutive points satisfy $\Delta\tau \leq 1/24$ day (1 hour).
3. **Group averaging with uncertainty propagation.** For each group G use inverse-variance weights $w_i = \sigma_i^{-2}$ and compute
$$\bar{\tau}_G = \frac{\sum_{i\in G} w_i \tau_i}{\sum_{i\in G} w_i}, \bar{f}_G = \frac{\sum_{i\in G} w_i f_i}{\sum_{i\in G} w_i}, \bar{\sigma}_G^2 = \frac{1}{\sum_{i\in G} w_i}.$$
This reduces irregular sampling and noise while preserving temporal shape.
4. **Per-band sequences and summaries.** Build per-band sequences $\{(\bar{\tau}_G, \bar{f}_G, \bar{\sigma}_G)\}$ and derive compact, interpretable summaries (e.g., number of detections, rise/decline proxies, per-band amplitude and timescale), which feed the RF classifier.

If magnitudes are provided, we first convert to fluxes before averaging; grouping is then performed in flux space to avoid bias from magnitude nonlinearity.

Methods

We evaluate several families of models and select a Random Forest (RF) classifier for its robustness on tabular, heterogeneous features and its interpretability via feature importances. Hyperparameters are tuned with Random Search and Bayesian Optimization (Bergstra and Bengio 2012; Snoek, Larochelle, and Adams 2012) (Table 1). Given the class imbalance (SN Ia minority), we optimize directly for the SPCC figure-of-merit F1 and monitor precision, recall, and ROC-AUC.

Model selection. Under the same preprocessing, we compared a linear perceptron, shallow RNNs (1–2 layers, up to 10 hidden units), and tree ensembles. Across random seeds, Random Forest consistently achieved the highest SPCC F1 and more favorable precision–recall curves. Given broker constraints (latency, interpretability) and the workshop's emphasis on deployable AI, we adopt RF as the primary model.

Hyperparameter	Random Search	Bayesian Opt.
n_estimators	1400	175
max_features	auto	0.6777
max_samples	None	0.9222
bootstrap	False	True
max_depth	40	None

Table 1: Hyperparameters explored for RF. The two best settings are shown.

Experimental Setup

Given class imbalance, we report Precision, Recall, F1 (SPCC), and ROC–AUC. Definitions:

$$\text{Purity or Precision} = \frac{TP}{TP + FP}, \tag{1}$$

$$\text{Completeness or Recall} = \frac{TP}{TP + FN}, \tag{2}$$

$$\text{F1} = 2 \cdot \frac{\text{Precision} \cdot \text{Recall}}{\text{Precision} + \text{Recall}}, \tag{3}$$

$$\mathrm{F1}_{\mathrm{macro}} = \frac{1}{K}\sum_{k=1}^{K} F1_k. \quad (4)$$

Macro-averaging treats each class equally. In addition, we report the SPCC figure-of-merit (eq. 5) for direct comparability with the original challenge (Kessler et al. 2010).

We frame the task as binary classification (SN Ia vs. non-SN Ia) and report the SPCC figure-of-merit

$$F1 = \frac{1}{TP + FN} \cdot \frac{TP^2}{TP + 3\,FP}, \quad (5)$$

with precision $\frac{TP}{TP+FP}$ and recall $\frac{TP}{TP+FN}$. We also provide ROC-AUC for comparison with prior work (Kessler et al. 2010).

Unless noted otherwise, 'F1' refers to the SPCC figure-of-merit (5).

Given class imbalance, we emphasize precision–recall evaluation following established guidance on PR vs ROC analysis for skewed datasets (Davis and Goadrich 2006; Saito and Rehmsmeier 2015).

Results

Table 2 summarizes the best RF models after tuning. Bayesian optimization yields slightly higher ROC-AUC and precision at a small cost in recall; Random Search offers the converse. Both reach essentially the same SPCC F1.

Metric	Random Search	Bayesian Opt.
ROC-AUC	0.857	0.886
F1 (SPCC)	0.944	0.941
Precision	0.925	0.958
Recall	0.963	0.924

Table 2: Random Forest performance on SPCC (post-challenge split).

Consistent with the model selection in Methods, we take the Bayesian-optimized configuration (right column in Table 2) as our final model due to higher precision at essentially the same SPCC F1.

Results indicate that a tuned RF over minimal preprocessing remains a competitive baseline in a strictly redshift-free regime. This is encouraging for real-time brokers where latency and missing host information are typical.

Discussion

Why F1 over ROC-AUC? With skewed classes, ROC-AUC can look strong even when the positive class (SN Ia) suffers many false positives. The SPCC F1 explicitly penalizes such errors and better aligns with the operational objective: maximizing the purity and completeness of the SN Ia set used for cosmology (Jeni, Cohn, and De La Torre 2013; Davis and Goadrich 2006; Saito and Rehmsmeier 2015).

Why not Deep Learning? While lightweight deep learning (DL) architectures, such as shallow RNNs or 1D-CNNs, offer the potential for end-to-end learning from raw time series (Charnock and Moss 2017; Möller and de Boissière 2020), they present distinct challenges in this constrained setting. First, the highly irregular cadence and large gaps characteristic of ground-based photometry often necessitate complex imputation methods or specialized architectures (e.g., Phased LSTMs) that undermine the computational efficiency required for low-latency brokers. Second, on dataset sizes typical of early survey phases or specific challenges like SPCC, deep models are often more prone to overfitting and instability across random seeds compared to ensemble methods (Lochner et al. 2016). Finally, the opaque nature of neural networks complicates the interpretability required for vetting alert streams, whereas our Random Forest approach provides transparent feature importance and robust performance without extensive architectural tuning.

Interpretability and operations. RF feature importances and saliency at the level of light-curve summaries provide transparency useful for vetting candidates and diagnosing failure modes in a real-time broker. Because the model consumes only photometry, it can run early in the alert stream to prioritize scarce follow-up resources.

For real-time operations, our pipeline can be integrated into community brokers that already implement large-scale alert ingestion, annotation and filtering (Fink, ALeRCE, ANTARES, Lasair) (Möller et al. 2021; Förster et al. 2021; Narayan et al. 2018; Smith et al. 2019; Sánchez-Sáez et al. 2021).

Limitations and Future Work

Our evaluation is simulation-based (SPCC); domain shift to real surveys remains the main limitation. In particular, calibration systematics, irregular cadences, and non-stationary alert qualities can degrade performance. Future work will validate on ZTF-like real data and explore controlled use of host metadata as an optional feature to quantify the trade-off between realism and accuracy.

Conclusion

Under broker-like constraints (no redshift, no host), a lightweight RF with minimal preprocessing yields strong F1 (SPCC) while remaining simple to deploy, making it suitable for Rubin-era alert triage. Emphasizing precision–recall metrics and transparent ensembles offers a robust, scalable route to AI-accelerated discovery in time-domain astronomy.

Acknowledgments

The work of I. Makarov was supported by the Ministry of Economic Development of the Russian Federation (agreement No. 139-10-2025-034 dd. 19.06.2025, IGK 000000C313925P4D0002)

References

Bellm, E. C.; Kulkarni, S. R.; Graham, M. J.; and et al. 2019. The Zwicky Transient Facility: System Overview, Performance, and First Results. *Publications of the Astronomical Society of the Pacific*, 131(995): 018002.

Bergstra, J.; and Bengio, Y. 2012. Random Search for Hyper-Parameter Optimization. *Journal of Machine Learning Research*, 13: 281–305.

Bloom, J.; and Richards, J. 2011. Data mining and machine-learning in time-domain discovery & classification. *Advances in Machine Learning and Data Mining for Astronomy*, 89–112.

Charnock, T.; and Moss, A. 2017. Deep recurrent neural networks for supernovae classification. *The Astrophysical Journal Letters*, 837(2): L28.

Davis, J.; and Goadrich, M. 2006. The Relationship Between Precision-Recall and ROC Curves. In *Proceedings of the 23rd International Conference on Machine Learning (ICML)*, 233–240.

Dobryakov, S.; Malanchev, K.; Derkach, D.; and Hushchyn, M. 2021. Photometric data-driven classification of type Ia supernovae in the open Supernova Catalog. *Astronomy and Computing*, 35: 100451.

Förster, F.; Cabrera-Vives, G.; Castillo-Navarrete, E.; and et al. 2021. The Automatic Learning for the Rapid Classification of Events (ALeRCE) Alert Broker. *The Astronomical Journal*, 161(5): 242.

Garretson, B.; Milisavljevic, D.; Reynolds, J.; Weil, K. E.; Subrayan, B.; Banovetz, J.; and Lee, R. 2021. Supernova host galaxy association and photometric classification of over 10,000 light curves from the Zwicky Transient Facility. *Research Notes of the AAS*, 5(12): 283.

Guy, J.; Astier, P.; Baumont, S.; and et al. 2007. SALT2: Using Distant Supernovae to Improve the Use of Type Ia Supernovae as Distance Indicators. *Astronomy & Astrophysics*, 466(1): 11–21.

Ivezić, Ž.; Kahn, S. M.; Tyson, J. A.; and et al. 2019. LSST: From Science Drivers to Reference Design and Anticipated Data Products. *The Astrophysical Journal*, 873(2): 111.

Jeni, L. A.; Cohn, J. F.; and De La Torre, F. 2013. Facing imbalanced data–recommendations for the use of performance metrics. In *2013 Humaine association conference on affective computing and intelligent interaction*, 245–251. IEEE.

Kessler, R.; Conley, A.; Jha, S.; and Kuhlmann, S. 2010. Supernova photometric classification challenge. *arXiv preprint arXiv:1001.5210*.

Kessler, R.; et al. 2019. Models and Simulations for the Photometric LSST Astronomical Time Series Classification Challenge. *Publications of the Astronomical Society of the Pacific*, 131(1003): 095001.

Knights, M.; Bassett, B. A.; Varughese, M.; Hlozek, R.; Kunz, M.; Smith, M.; and Newling, J. 2013. Extending BEAMS to incorporate correlated systematic uncertainties. *Journal of Cosmology and Astroparticle Physics*, 2013(01): 039.

Lochner, M.; McEwen, J. D.; Peiris, H. V.; Lahav, O.; and Winter, M. K. 2016. Photometric supernova classification with machine learning. *The Astrophysical Journal Supplement Series*, 225(2): 31.

Masci, F. J.; Laher, R. R.; Rusholme, B.; and et al. 2019. The Zwicky Transient Facility: Data Processing, Products, and Archive. *Publications of the Astronomical Society of the Pacific*, 131(995): 018003.

Minkowski, R. 1979. Spectra of supernovae. In *A Source Book in Astronomy and Astrophysics, 1900–1975*, 478–480. Harvard University Press.

Möller, A.; and de Boissière, T. 2020. SuperNNova: an open-source framework for Bayesian, neural network-based supernova classification. *Monthly Notices of the Royal Astronomical Society*, 491(3): 4277–4293.

Möller, A.; Peloton, J.; Ishida, E. E. O.; and et al. 2021. Fink: A New Generation of Broker for the LSST Community. *Monthly Notices of the Royal Astronomical Society*, 501(3): 3272–3288.

Möller, A.; Ruhlmann-Kleider, V.; Leloup, C.; Neveu, J.; Palanque-Delabrouille, N.; Rich, J.; Carlberg, R.; Lidman, C.; and Pritchet, C. 2016. Photometric classification of type Ia supernovae in the SuperNova Legacy Survey with supervised learning. *Journal of Cosmology and Astroparticle Physics*, 2016(12): 008.

Möller, A.; Smith, M.; Sako, M.; Sullivan, M.; Vincenzi, M.; Wiseman, P.; Armstrong, P.; Asorey, J.; Brout, D.; Carollo, D.; et al. 2022. The dark energy survey 5-yr photometrically identified type Ia supernovae. *Monthly Notices of the Royal Astronomical Society*, 514(4): 5159–5177.

Möller, A.; Wiseman, P.; Smith, M.; Lidman, C.; Davis, T.; Kessler, R.; Sako, M.; Sullivan, M.; Galbany, L.; Lee, J.; et al. 2024. The Dark Energy Survey 5-yr photometrically classified type Ia supernovae without host-galaxy redshifts. *Monthly Notices of the Royal Astronomical Society*, 533(2): 2073–2088.

Morrell, N.; Phillips, M.; Folatelli, G.; Stritzinger, M.; Hamuy, M.; Suntzeff, N. B.; Hsiao, E.; Taddia, F.; Burns, C. R.; Hoeflich, P.; et al. 2024. Optical Spectroscopy of Type Ia Supernovae by the Carnegie Supernova Projects I and II. *The Astrophysical Journal*, 967(1): 20.

Narayan, G.; Zaidi, T.; Soraisam, M. D.; and et al. 2018. Machine-learning-based Brokers for Real-time Classification of the LSST Alert Stream. *The Astrophysical Journal Supplement Series*, 236(1): 9.

Newling, J.; Varughese, M.; Bassett, B.; Campbell, H.; Hlozek, R.; Kunz, M.; Lampeitl, H.; Martin, B.; Nichol, R.; Parkinson, D.; et al. 2011. Statistical classification techniques for photometric supernova typing. *Monthly Notices of the Royal Astronomical Society*, 414(3): 1987–2004.

Reza, M.; Wang, L.; and Hu, L. 2025. An FPCA-Enhanced Ensemble Learning Framework for Photometric Identification of Type Ia Supernovae. *arXiv preprint arXiv:2510.09990*.

Rubin, D.; Aldering, G.; Barbary, K.; Boone, K.; Chappell, G.; Currie, M.; Deustua, S.; Fagrelius, P.; Fruchter, A.; Hayden, B.; et al. 2015. UNITY: confronting supernova cosmology's statistical and systematic uncertainties in a unified Bayesian framework. *The Astrophysical Journal*, 813(2): 137.

Saito, T.; and Rehmsmeier, M. 2015. The Precision-Recall Plot Is More Informative than the ROC Plot When Evaluating Binary Classifiers on Imbalanced Datasets. *PLOS ONE*, 10(3): e0118432.

Sánchez-Sáez, P.; Reyes, I.; Valenzuela, C.; Förster, F.; Eyheramendy, S.; Elorrieta, F.; Bauer, F.; Cabrera-Vives, G.; Estévez, P.; Catelan, M.; et al. 2021. Alert classification for the ALeRCE broker system: The light curve classifier. *The Astronomical Journal*, 161(3): 141.

Smith, K. W.; Williams, R. D.; Young, D. R.; and et al. 2019. Lasair: The Transient Alert Broker for LSST:UK. *Research Notes of the AAS*, 3(1): 26.

Snoek, J.; Larochelle, H.; and Adams, R. P. 2012. Practical Bayesian Optimization of Machine Learning Algorithms. In *Advances in Neural Information Processing Systems 25 (NeurIPS)*, 2951–2959.

Taylor, G.; Hinton, S.; Scolnic, D.; and et al. 2021. A Revised SALT2 Surface for Fitting Type Ia Supernova Light Curves. *Monthly Notices of the Royal Astronomical Society*, 504(3): 4111–4130.