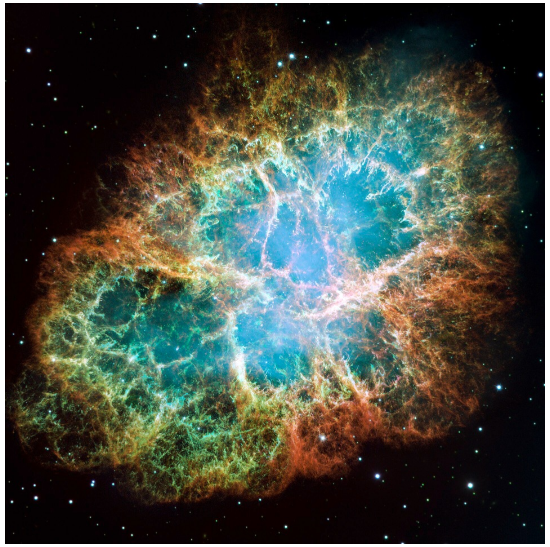
# Feature-Based and Deep Learning Methods for Astronomical Light Curve Classification

Valentin Baymakanov [1], Khakim Khamitov [1], Alina Raimova [1], Anastasia Isaeva [2], Roman Taran [2], Mohammad Zedan [1], Matwey Kornilov [1,3,4], Mariia Demianenko [5, 6], Konstantin Malanchev [7], Mikhail Hushchyn [1]

[1] HSE University, [2] Southern Federal University, [3] Lomonosov Moscow State University, [4] Sternberg astronomical institute, [5] Max-Planck-Institut für Astronomie, [6] Heidelberg University, [7] Carnegie Mellon University

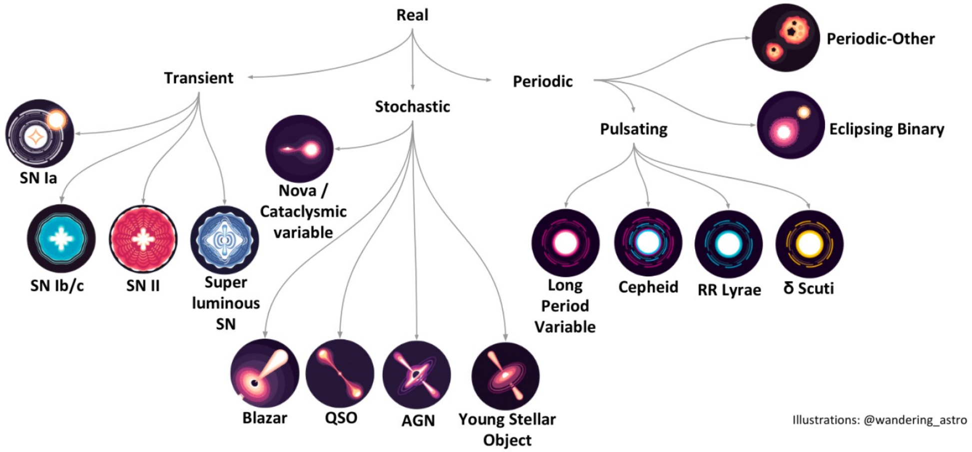
# Introduction

- Modern-day astronomical surveys collect **large amounts of data** every day
- Observed objects are **transient**, **periodic** and **aperiodic** sources
- A wide variety of tasks, including **classification**, **flare detection**, determining **astrophysical properties** of observed objects, etc.
- The amount of data and variety of tasks necessitate the **development of effective methods** to process this data



**Crab Nebula**

**NGC 4414 Galaxy**




**Tree of astronomical classification**

# Objectives and tasks

We explore the classification of astronomical objects from their light curves in several passbands.

Possible routes:

1. Classical ML with physically motivated features
2. Deep learning approach on raw data
3. Foundation models

**Goals:**

- Develop classifiers using feature-based approach
- Develop LSTM and Transformer classifiers on raw light curves

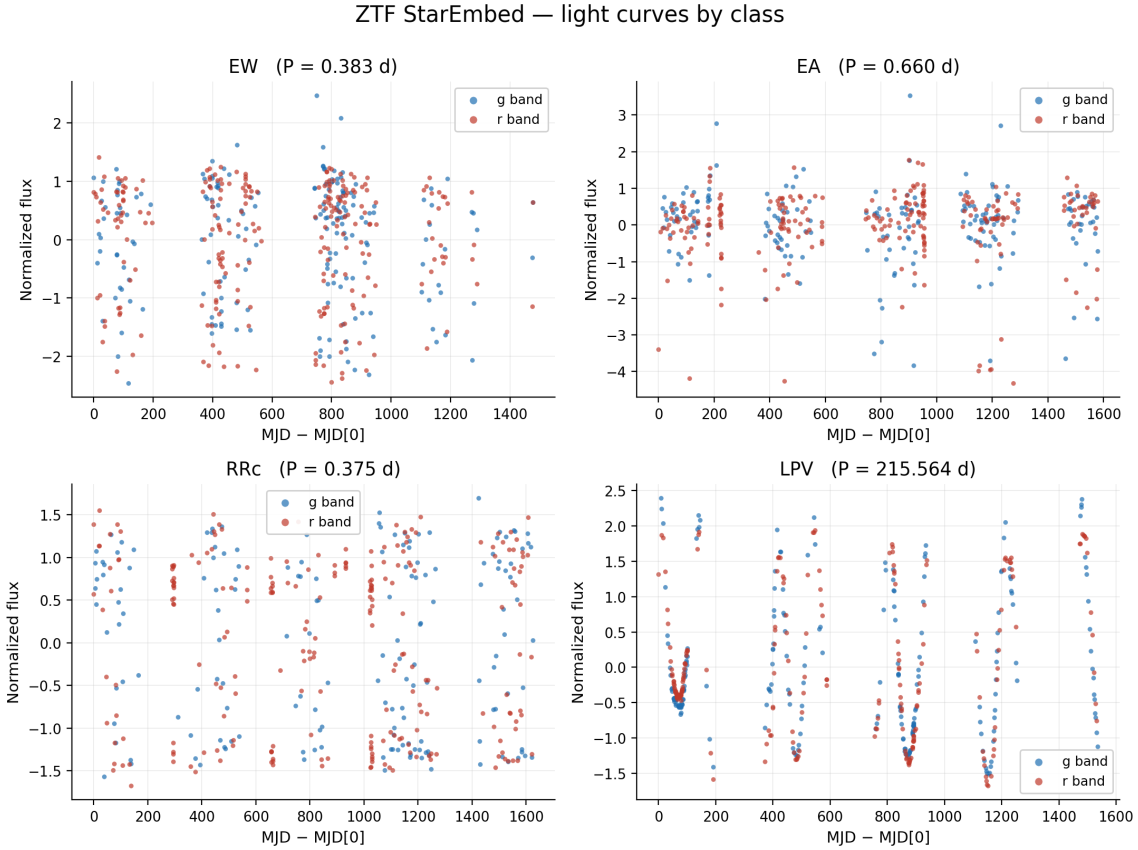
**Vera C. Rubin Observatory**

# What is a light curve?

A **light curve** is a graph of light intensity of an astronomical object as a function of time

Light curves collected by observatories have the following characteristics:

- Irregular cadence (sampling)
- Missing observations
- Measurement uncertainties
- Multiple passbands



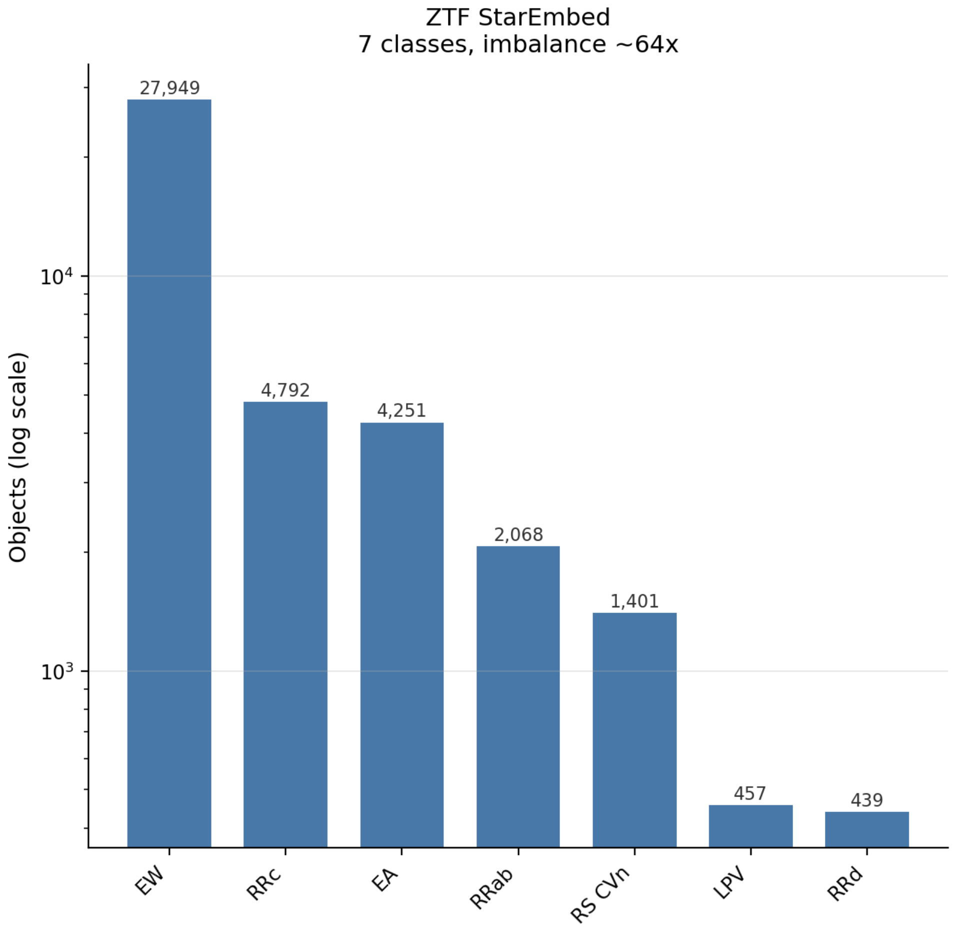

**Light curve examples**

# Dataset

**ZTF StarEmbed** is a curated subset of the Zwicky Transient Facility astronomical survey (Palomar Observatory, California) designed to evaluate machine learning models on astronomical light curves. It combines multi-band ZTF observations with expert-vetted labels across seven astrophysical classes.

Characteristics of the dataset:

- 41,357 objects
- 7 classes of variable stars
- Median length ~770 observations
- Severe class imbalance across classes




Class distribution

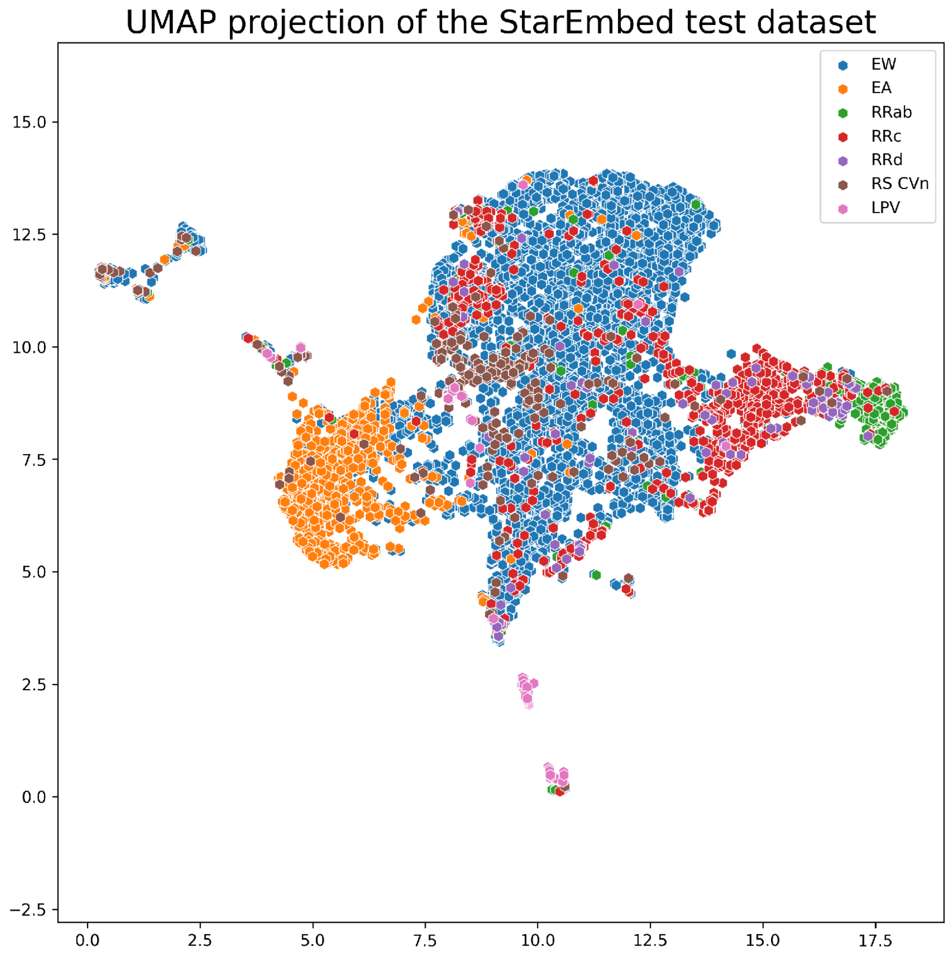
# Feature-based approach

From light curves we computed 26 features* *(Amplitude, Mean, MeanVariance, Median, etc.)* per passband.

We then concatenated them across passbands and used this representation as features.

Explored methods:

1. Logistic regression
2. CatBoost
3. MLP




**UMAP Projection of StarEmbed Objects Using Computed Features**

** Computed with the light-curve library – https://github.com/light-curve/light-curve-python*

# LSTM

Input: 512×4 sequence [Δt, lg λ, flux, flux_err] + padding mask

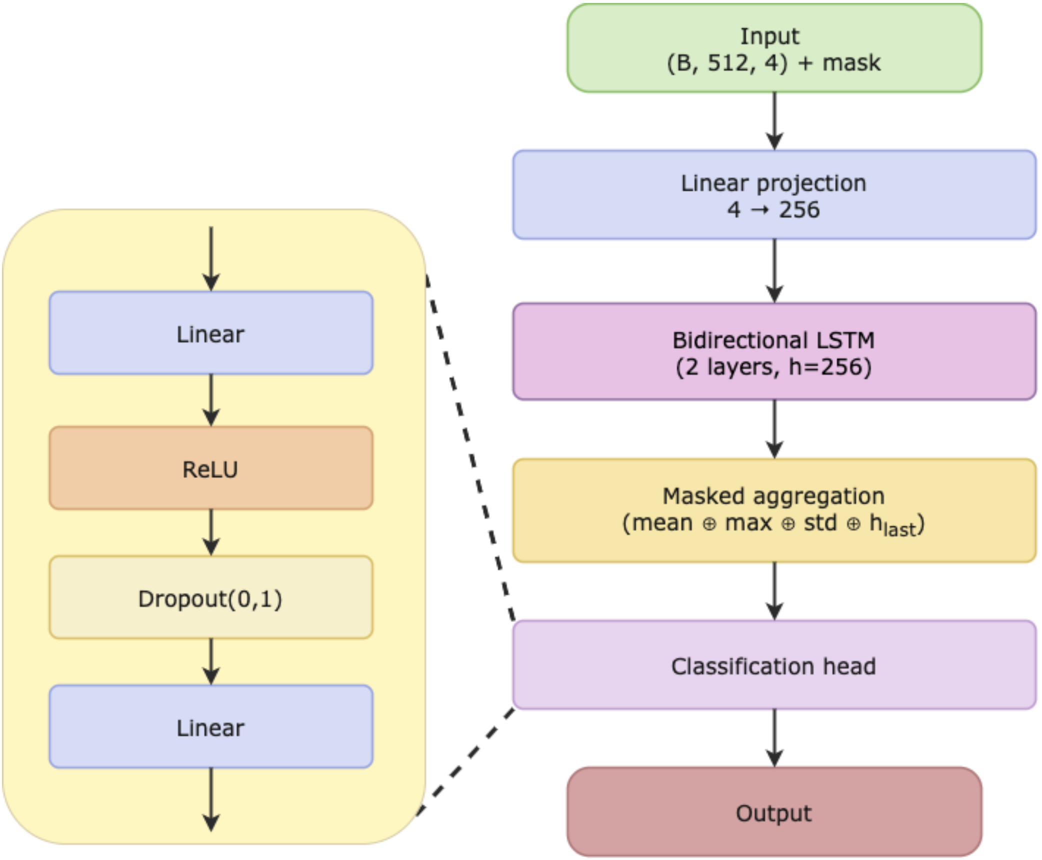
**LSTM + Linear:** 1 unidirectional layer (h=128), final hidden state → linear classifier

**Deep BiLSTM:** 2 bidirectional layers (h=256), pooling mean ⊕ max ⊕ std ⊕ final state (~3.7M params)

Head variants: GELU head, wide head, GELU + wide

Augmentations: random subsampling, random bins, random window

Training: AdamW, class-weighted cross-entropy, early stopping




**Deep BiLSTM architecture**

# Transformer

Input: 512×4 sequence [Δt, lg λ, flux, flux_err] + padding mask

Encoder: d_model=128, 4 heads, FFN 512, fixed sinusoidal positional encoding

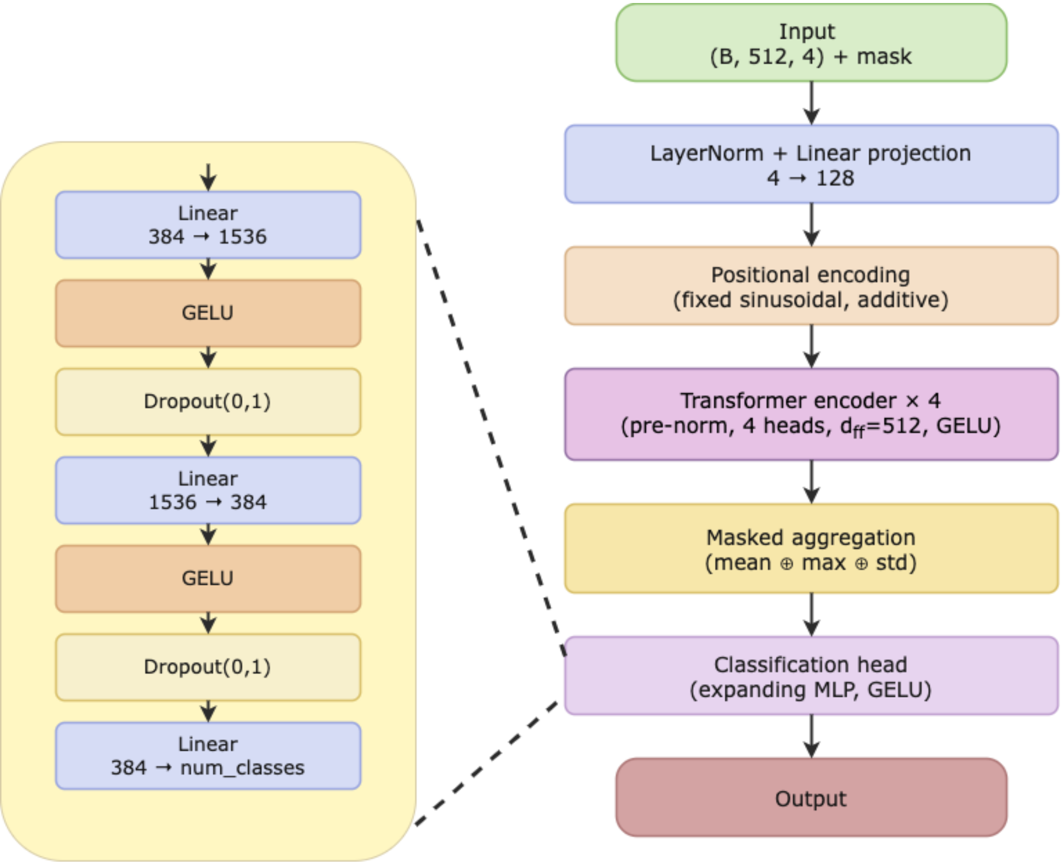
**Transformer + Linear:** 2 layers, mean pooling, linear head (0.27M params)

**Deep Transformer:** 4 layers, pooling mean ⊕ max ⊕ std

Head variants: ReLU/GELU, narrow/wide heads

Augmentations: random subsampling, random bins, random window

Training: AdamW, warmup + cosine schedule, class-weighted cross-entropy




**Deep Transformer architecture**

# Model comparisons

| Architecture | Acc | Bal. Acc | Macro F1 | Macro ROC-AUC |
|---|---|---|---|---|
| Logistic Regression | 0.787 | **0.801** | 0.661 | 0.955 |
| CatBoost | **0.907** | 0.742 | 0.771 | **0.981** |
| MLP | 0.903 | 0.762 | **0.779** | **0.981** |
| LSTM + Linear | 0.572 | 0.577 | 0.435 | 0.886 |
| Deep BiLSTM | 0.634 | 0.631 | 0.495 | 0.909 |
| Deep BiLSTM + GELU | 0.517 | 0.619 | 0.421 | 0.901 |
| Deep BiLSTM + wide head | 0.523 | 0.620 | 0.472 | 0.914 |
| Deep BiLSTM + GELU + wide | 0.487 | 0.598 | 0.392 | 0.886 |
| Deep BiLSTM + Rnd. window | 0.751 | 0.731 | 0.621 | 0.953 |
| Transformer + Linear | 0.517 | 0.524 | 0.374 | 0.857 |
| Deep Transformer | 0.529 | 0.595 | 0.442 | 0.878 |
| Deep Transformer + GELU | 0.600 | 0.609 | 0.477 | 0.895 |
| Deep Transformer + wide head | 0.595 | 0.629 | 0.495 | 0.907 |
| Deep Transf. + GELU + wide | 0.616 | 0.631 | 0.488 | 0.909 |
| Deep Transf. + GELU + wide + Rnd. bins | 0.711 | 0.645 | 0.555 | 0.916 |
| *TCN + LSTM + Features (agent)* | *0.896* | *0.784* | *0.784* | *0.979* |

# Conclusions

Feature-based approaches perform best: the pipeline is **time-tested** and the physically motivated features encode strong domain knowledge about variable objects.

Deep learning on raw light curves is **data-limited**: ~40k labeled objects are not enough to learn representations from scratch. Augmentations effectively expand the training set, and we see consistently better results.

The next logical step of scaling data and architectures is **foundation models** pre-trained on large unlabeled surveys and fine-tuned for classification.

Thank you for your attention!

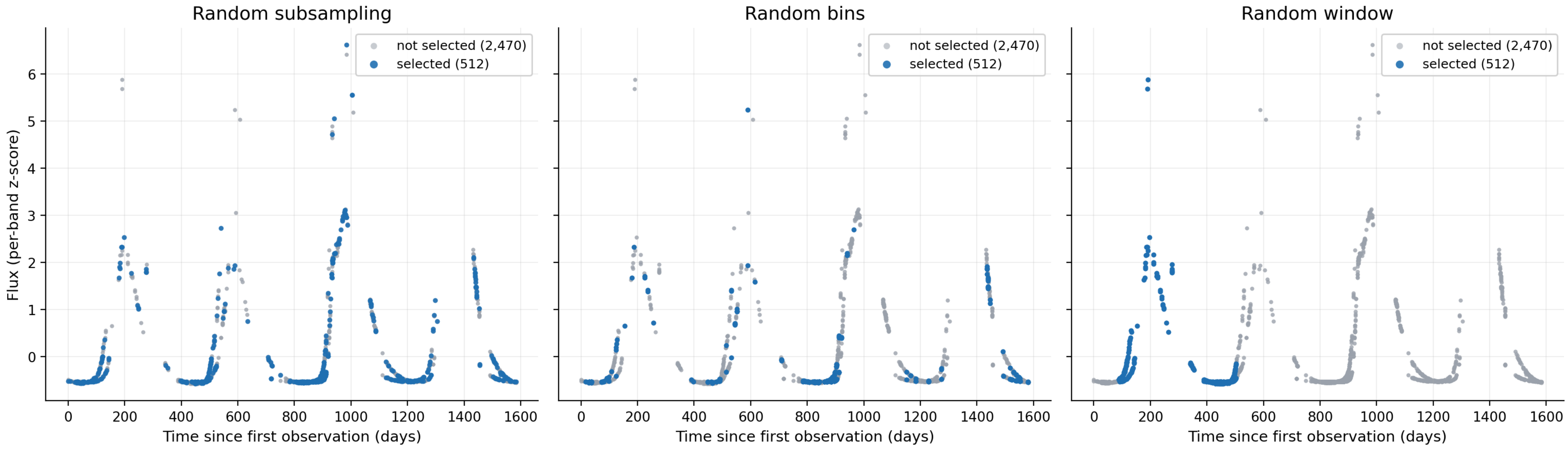
# Prepared dataset

| | ZTFID | lightcurve | | | | | type | redshift | SN_Ia |
|---|---|---|---|---|---|---|---|---|---|
| | | mjd | passband | flux | flux_err | lgwave | | | |
| 0 | ZTF25aaabezb | 60674.120903 | g | 27984.668455 | 6651.685384 | 3.676372 | SN Ia | 0.073 | 1 |
| | | *+14 rows* | ... | ... | ... | ... | | | |
| | | mjd | passband | flux | flux_err | lgwave | | | |
| 1 | ZTF19aapxocr | 58587.320463 | g | 65355.223466 | 9908.537578 | 3.676372 | SN Ia | 0.071 | 1 |
| | | *+9 rows* | ... | ... | ... | ... | | | |
| | | mjd | passband | flux | flux_err | lgwave | | | |
| 2 | ZTF20acbovrt | 59111.366053 | g | 19684.309197 | 4691.026098 | 3.676372 | SN Ia | 0.05 | 1 |
| | | *+63 rows* | ... | ... | ... | ... | | | |
| | | mjd | passband | flux | flux_err | lgwave | | | |
| 3 | ZTF18aatezaj | 59297.395255 | r | 39387.643642 | 9109.136374 | 3.803893 | SN Ia | 0.03419 | 1 |
| | | *+28 rows* | ... | ... | ... | ... | | | |
| | | mjd | passband | flux | flux_err | lgwave | | | |
| 5 | ZTF20aaurucm | 58954.175984 | g | 172599.725267 | 10669.445367 | 3.676372 | SN Ia | 0.09 | 1 |
| | | *+27 rows* | ... | ... | ... | ... | | | |

# All features

| | | | |
|---|---|---|---|
| **Amplitude** | Half of the flux/brightness range; measures how strongly the object varies between minimum and maximum. | **MeanVariance** | Standard deviation divided by the mean; relative variability amplitude. |
| **AndersonDarlingNormal** | Anderson-Darling normality test statistic; shows how far the value distribution is from Gaussian. | **Median** | Median flux/brightness value. |
| **BazinFit** | Feature from fitting the Bazin function, a light-curve shape model often used for transients and supernovae. | **MedianAbsoluteDeviation** | Median absolute deviation from the median; robust measure of scatter. |
| **BeyondNStd** | Fraction of points farther than N standard deviations from the mean; measures outliers or strong deviations. | **MedianBufferRangePercentage** | Fraction of points near the median within a specified range; measures concentration around the typical level. |
| **Cusum** | Range of cumulative sums of deviations from the mean; captures systematic or asymmetric temporal changes. | **PercentAmplitude** | Maximum deviation from the median; robust amplitude estimate. |
| **Eta** | Von Neumann index; measures how different neighboring observations are, sensitive to light-curve smoothness. | **PercentDifferenceMagnitudePercentile** | Inter-percentile range normalized by the median; relative width of the distribution. |
| **EtaE** | Modified Eta for unevenly sampled time series; accounts for time gaps between observations. | **RainbowFit** | Experimental physically motivated multiband transient-model fit using bolometric flux and temperature evolution. |
| **ExcessVariance** | Excess variance; estimates intrinsic variability after accounting for measurement errors. | **ReducedChi2** | Reduced chi-squared relative to the weighted mean; shows how much variability exceeds expected measurement noise. |
| **FluxNNotDetBeforeFd** | Number of non-detections before the first detection in flux | **Roms** | Robust median statistic: average absolute deviation from the median normalized by measurement errors. |
| **InterPercentileRange** | Inter-percentile range; width of the central part of the value distribution. | **Skew** | Skewness of the value distribution. |
| **Kurtosis** | Excess kurtosis; describes tail heaviness and peak sharpness of the distribution. | **StandardDeviation** | Standard deviation of flux/brightness. |
| **MaximumSlope** | Maximum rate of change between consecutive observations. | **StetsonK** | Robust light-curve shape index accounting for measurement errors; related to peakedness and variability structure. |
| **Mean** | Mean flux/brightness value. | **WeightedMean** | Weighted mean, where more precise measurements receive larger weight. |

# Light curve augmentations




**Three types of augmentation for LPV from ZTF Starembed**

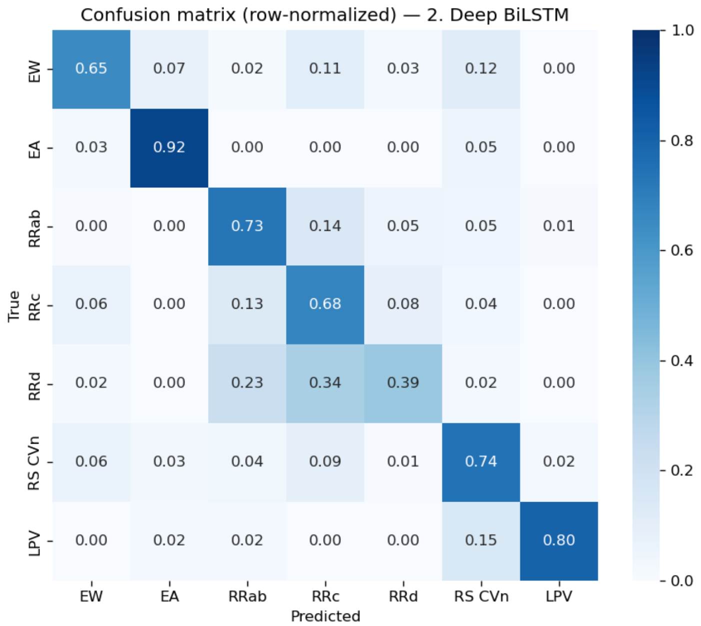
**Random subsampling**: keep 512 points drawn uniformly at random from the whole curve, in time order

**Random bins**: randomly pick 24h bins and keep all their points until 512 are collected. Drop oversampled points from the last bin

**Random window**: keep 512 consecutive observations from a random start point

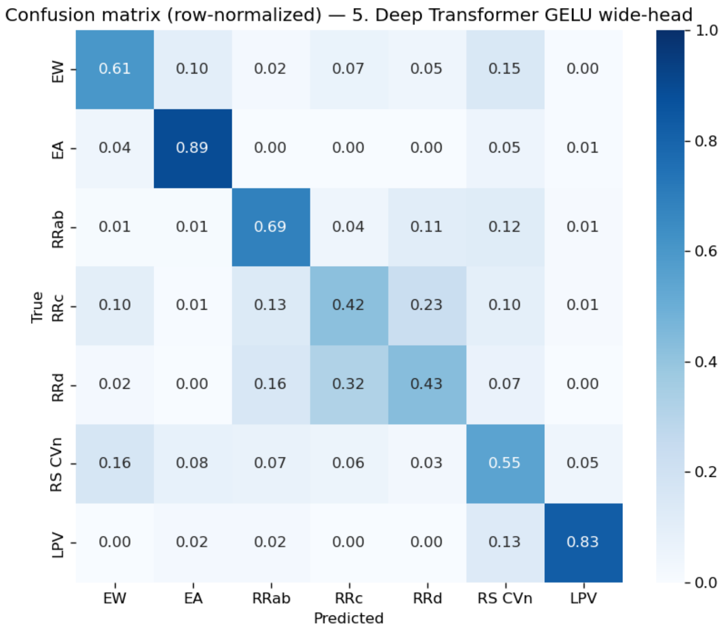
# Light curve augmentation results

### *Deep BiLSTM architecture augmentations*



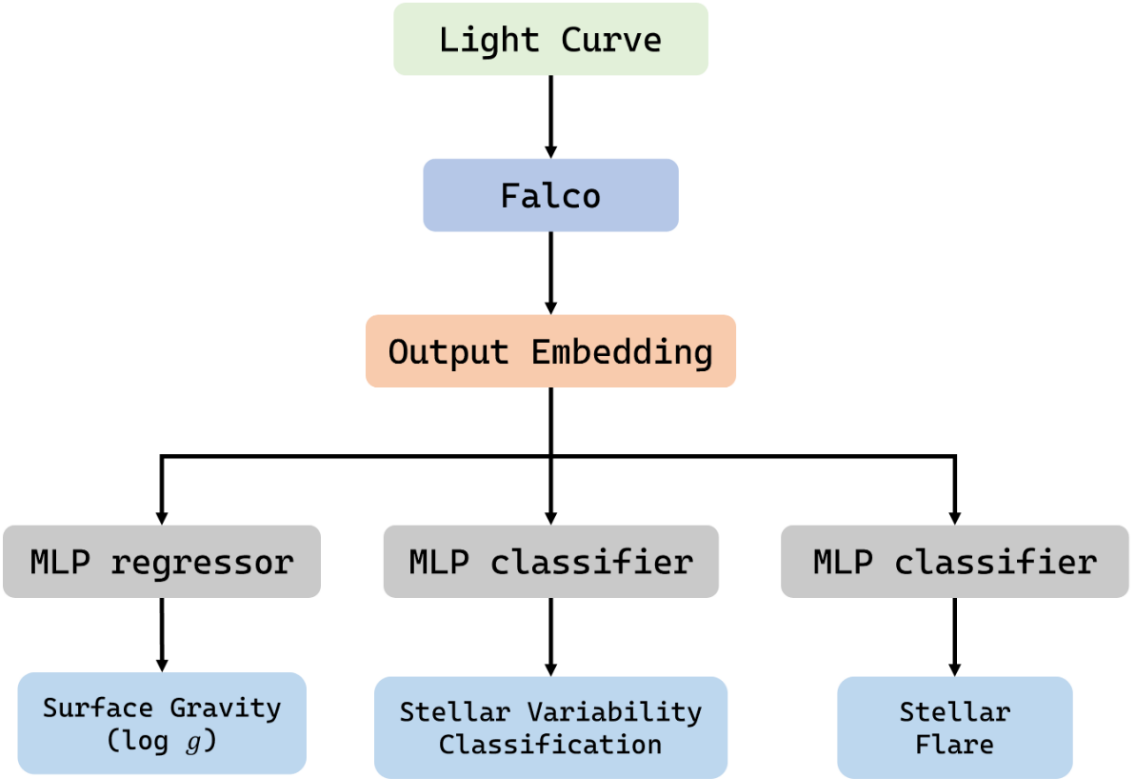

| Experiment | Acc | Bal. Acc | Macro F1 | Wtd F1 | AUC mi | AUC ma |
|---|---|---|---|---|---|---|
| Rnd. subsampling | 0.663 | 0.647 | 0.518 | 0.703 | 0.928 | 0.924 |
| Rnd. bins | 0.689 | 0.653 | 0.556 | 0.732 | 0.935 | 0.931 |
| Rnd. window | **0.751** | **0.731** | **0.621** | **0.780** | **0.960** | **0.953** |

### *Deep Transformer GELU wide-head architecture augmentations*




| Experiment | Acc | Bal. Acc | Macro F1 | Wtd F1 | AUC mi | AUC ma |
|---|---|---|---|---|---|---|
| Rnd. subsampling | 0.601 | 0.632 | 0.487 | 0.657 | 0.900 | 0.905 |
| Rnd. bins | **0.711** | **0.645** | **0.555** | **0.736** | **0.941** | **0.916** |
| Rnd. window | 0.651 | 0.640 | 0.506 | 0.691 | 0.925 | 0.913 |

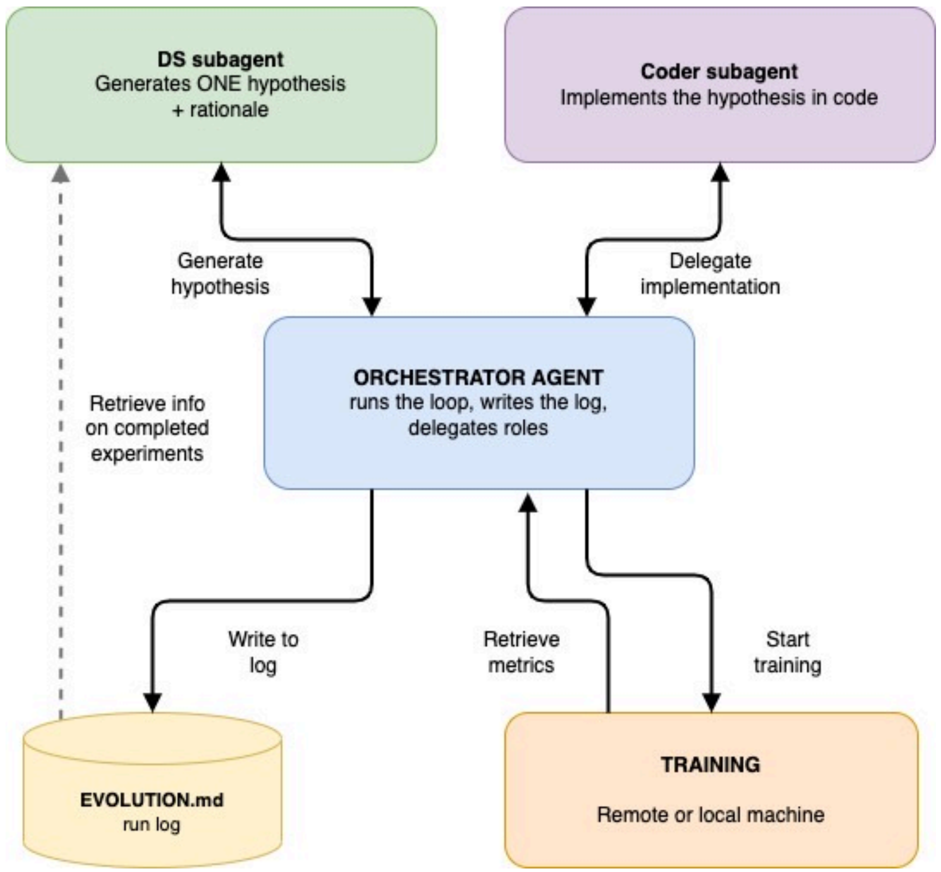
# Foundation model



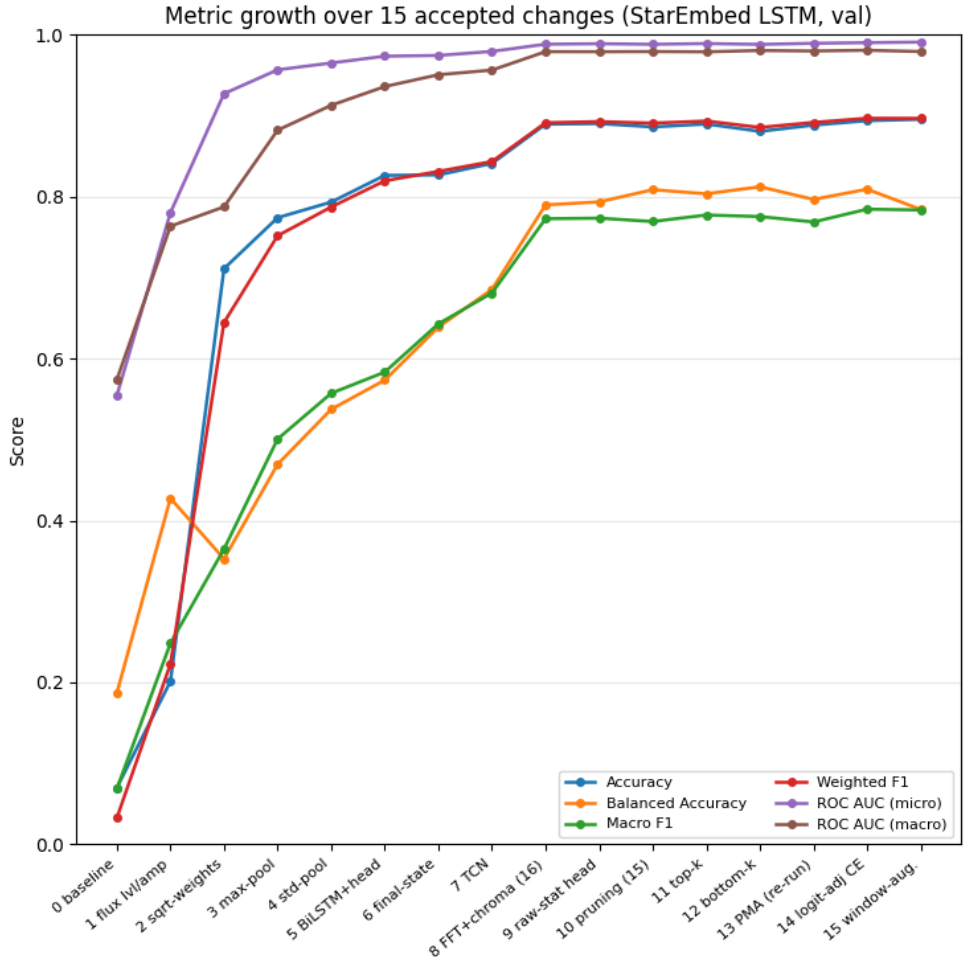

Foundation model approach illustrated with Falco*

*https://arxiv.org/pdf/2504.20290

# Agentic approach




*Agentic pipeline architecture*




*Metric improvements over successful iterations*