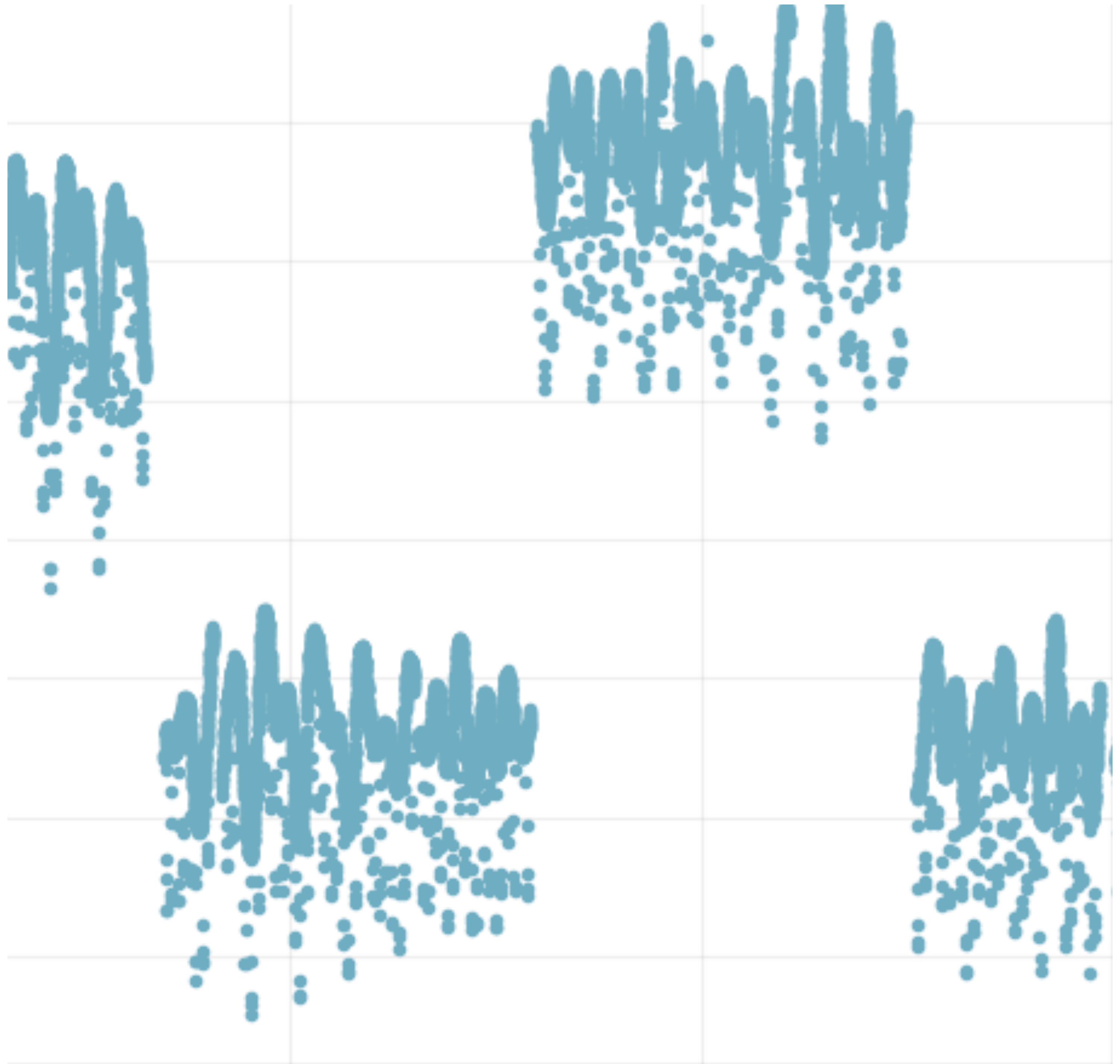
Detection of exoplanets in transit light curves with Conditional Flow Matching and XGBoost

Stefano Fiscale[1,2,3], PhD student

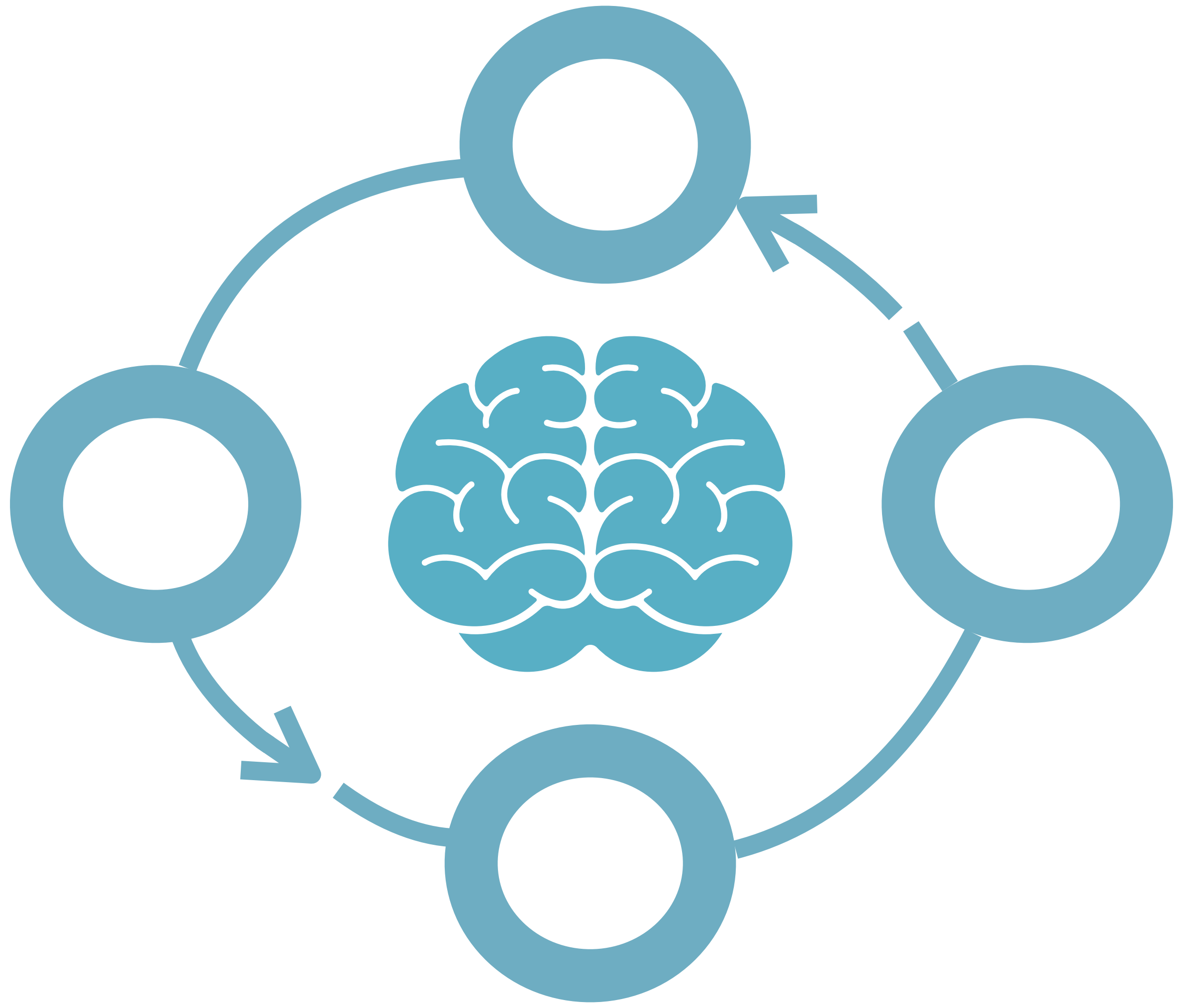
PLATO - ESP2025, Planets throughout the Habitable Zone, Marseille 23-25 June 2025

[1] UNESCO Chair "Environment, Resources and Sustainable Development", Department of Science and Technology, Parthenope University of Naples, Naples, Italy
[2] INAF, Osservatorio Astronomico di Capodimonte, Naples, Italy
[3] Aix Marseille Univ, CNRS, CNES, Institut Origines, LAM, Marseille, France

Overview



Data Preparation



Machine Learning workflow

Experiments and Future Directions

Fiscale et al., Electronics 2025, 14(9), 1738

Limitation of current vetting algorithms

Processing features relevant for humans in classification

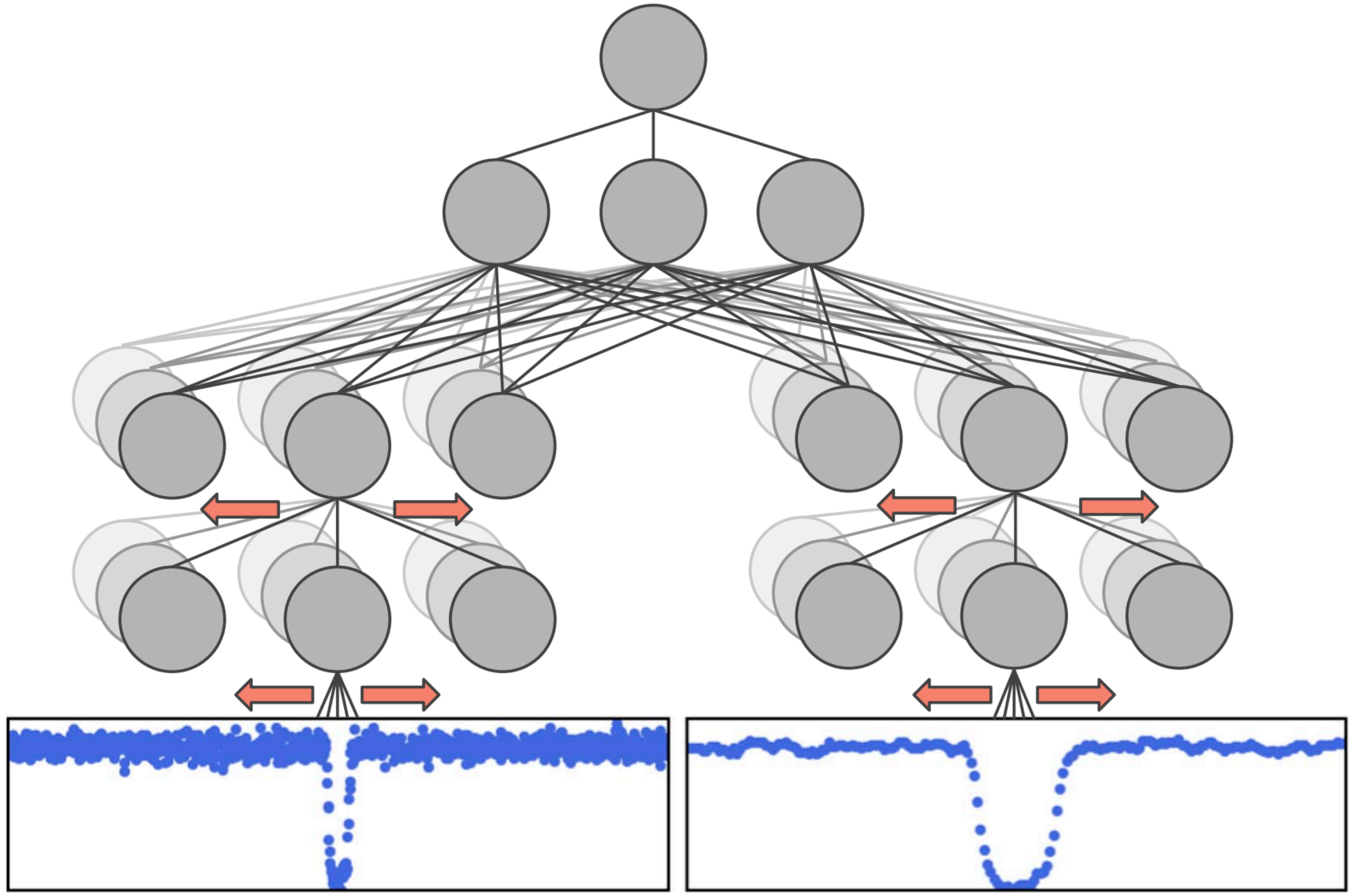


Figure credit: Shallue & Vanderburg, AJ, 155:94, 2018

Limitation of current vetting algorithms
Processing features relevant for humans in classification

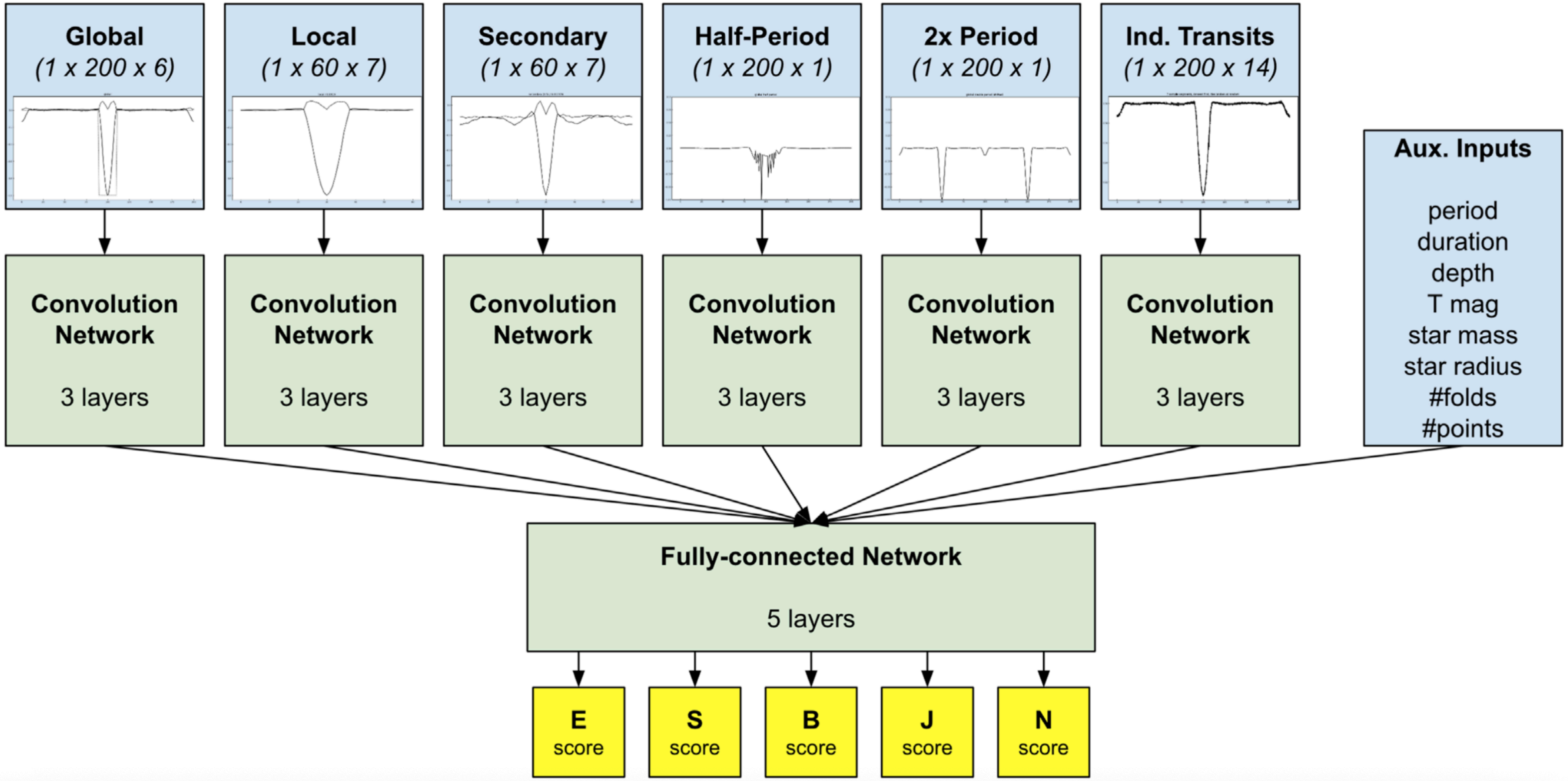


Figure credit: Tey E. et al., AJ, 165:95, 2023

Limitation of current vetting algorithms

Processing features relevant for humans in classification

Figure credit **EfficientNet: Rethinking Model Scaling for Convolutional Neural Networks**

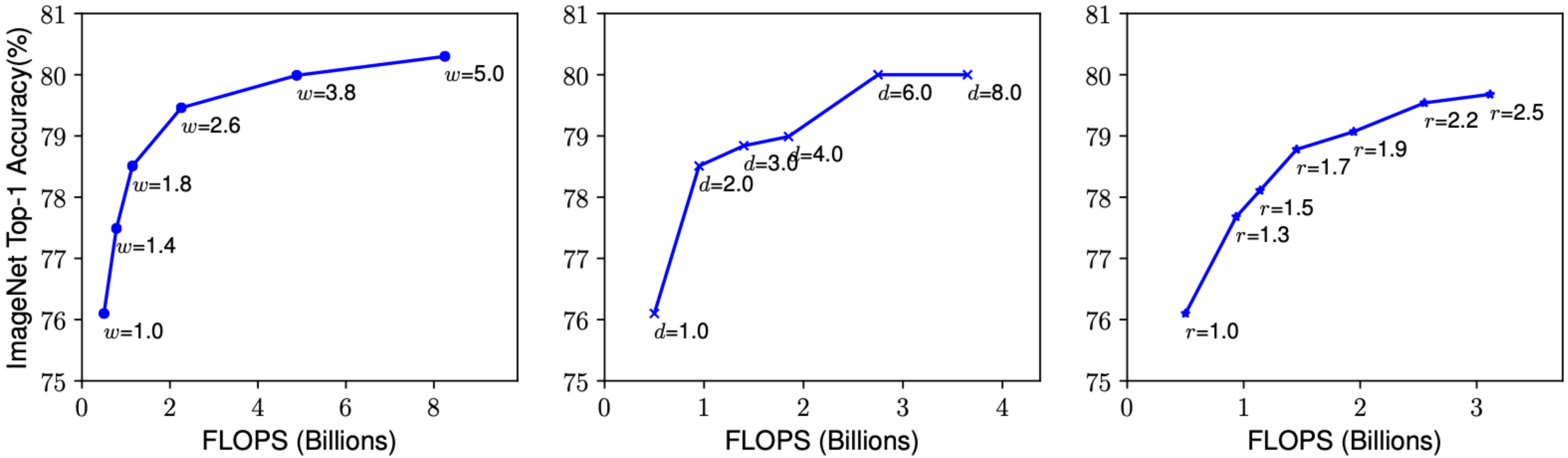


Figure 3. **Scaling Up a Baseline Model with Different Network Width (w), Depth (d), and Resolution (r) Coefficients.** Bigger networks with larger width, depth, or resolution tend to achieve higher accuracy, but the accuracy gain quickly saturate after reaching 80%, demonstrating the limitation of single dimension scaling. Baseline network is described in Table 1.

Drawbacks of these approaches

Model complexity

- **100,000,000** model parameters to be optimized
- **overfitting** issues

Application to new surveys

Development of new architectures **from scratch**

Occam’s razor heuristic

“…solving principle that recommends searching for explanations constructed with the smallest possible set of elements”

Occam's razor heuristic

Deep Residual Learning for Image Recognition

Kaiming He Xiangyu Zhang Shaoqing Ren Jian Sun
Microsoft Research
{kahe, v-xiangz, v-shren, jiansun}@microsoft.com

EfficientNet: Rethinking Model Scaling for Convolutional Neural Networks

---

Mingxing Tan [1] **Quoc V. Le** [1]

Size does matter: Exoplanet detection with a sparse convolutional neural network

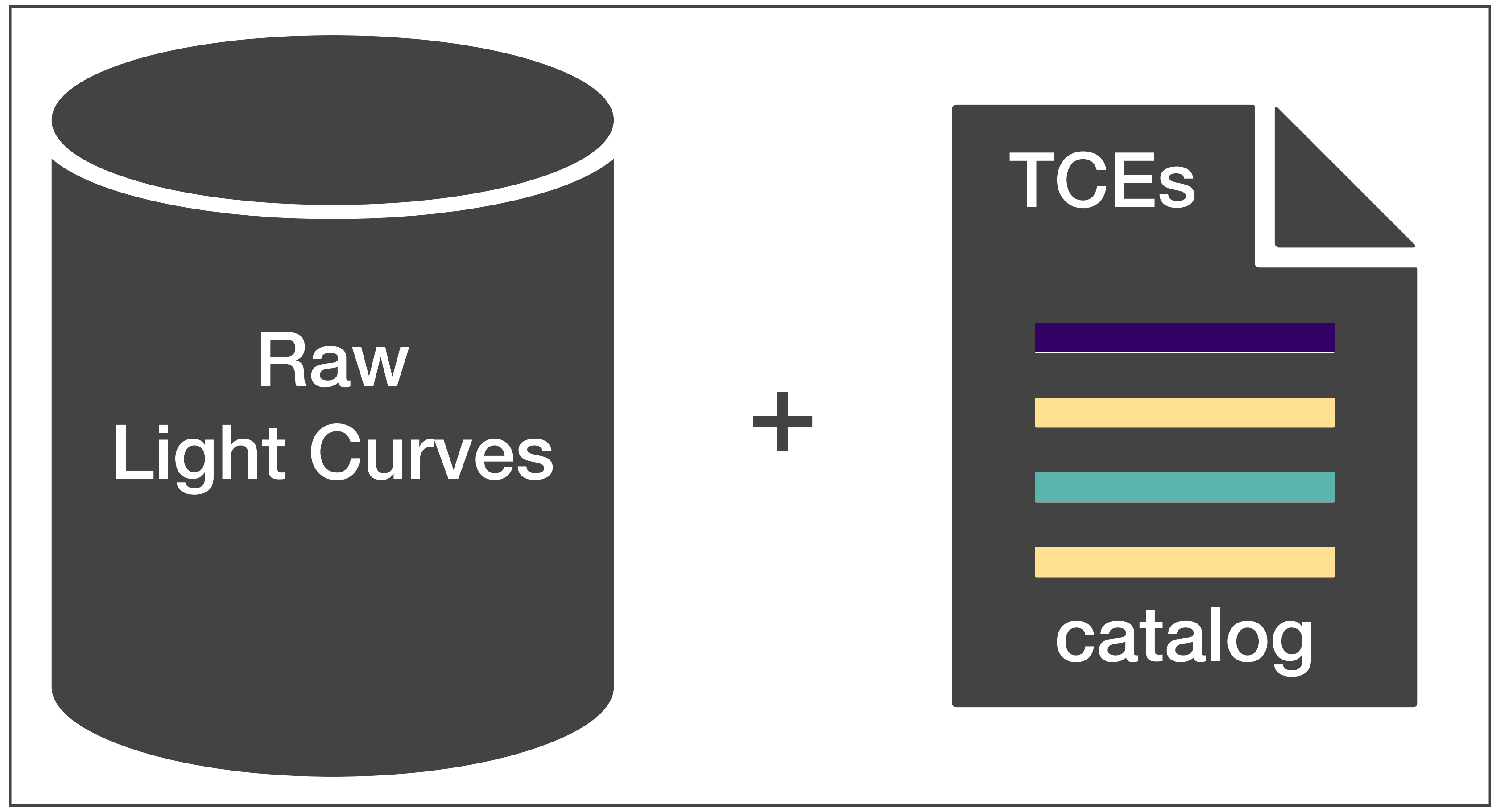
K. Visser [a,*], B. Bosma [b], E. Postma [a,c]

[a] *Jheronimus Academy of Data Science, Sint Janssingel 92, 5211 DA, 's-Hertogenbosch, The Netherlands*
[b] *Vrije Universiteit Amsterdam, De Boelelaan 1105, 1081 HV, Amsterdam, The Netherlands*
[c] *CSAI, Tilburg University, Warandelaan 2, 5037 AB, Tilburg, The Netherlands*

Data preparation

Data Preparation

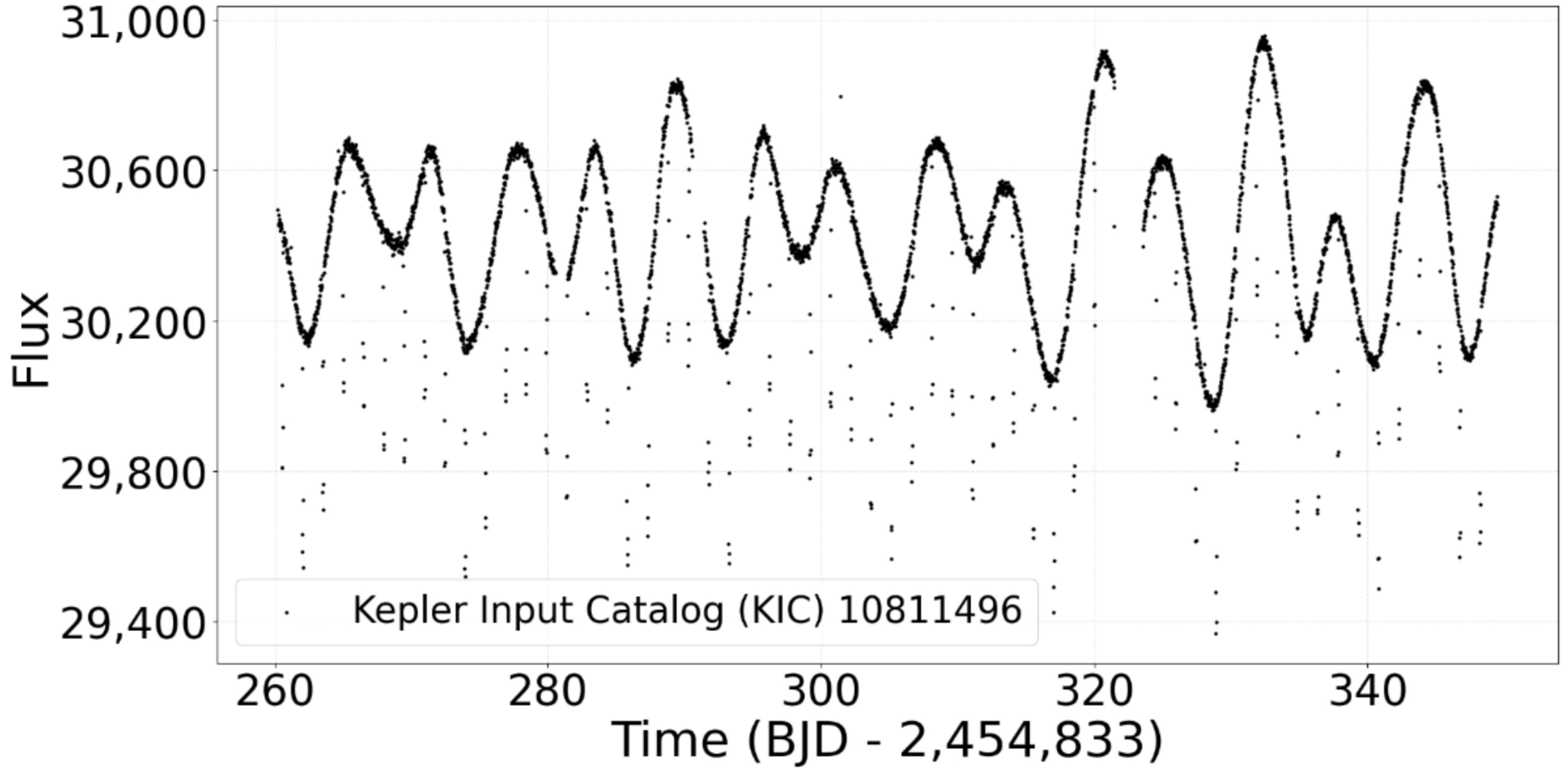
From raw light curves to input representation

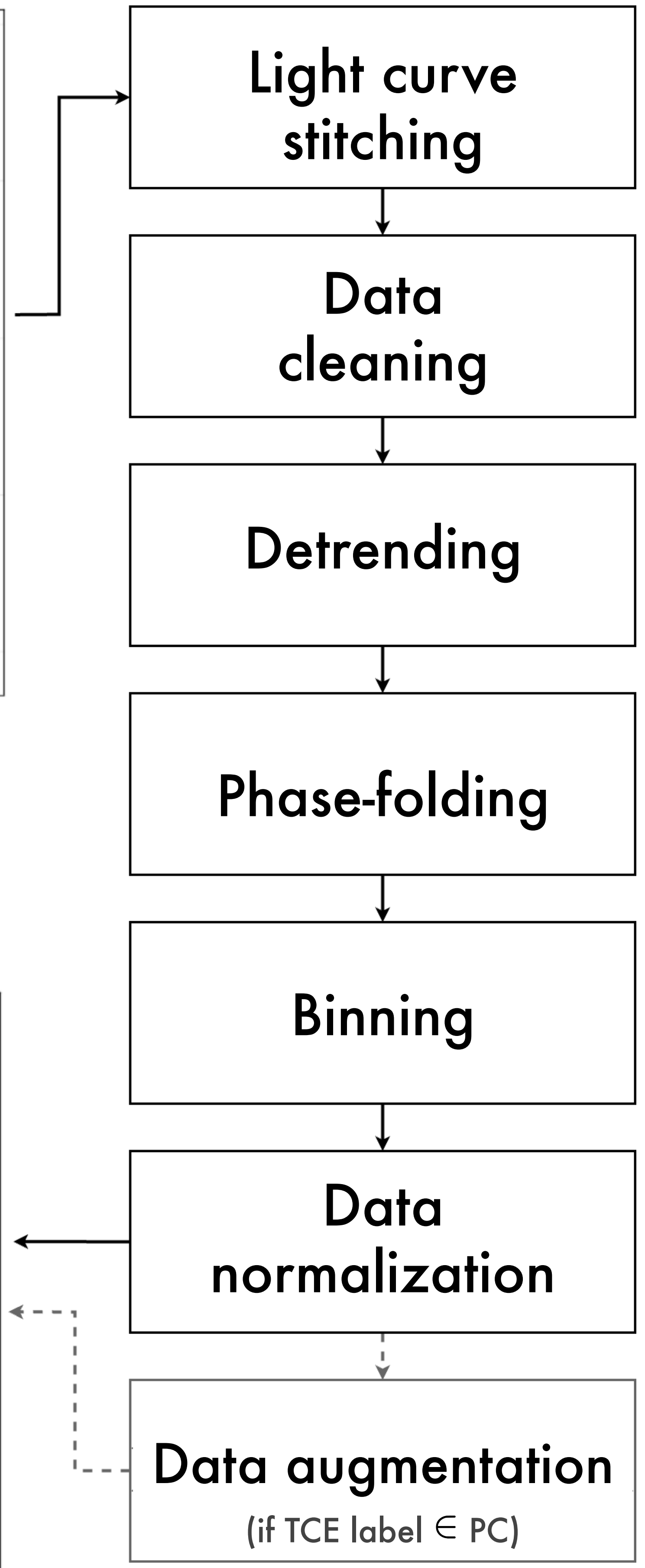


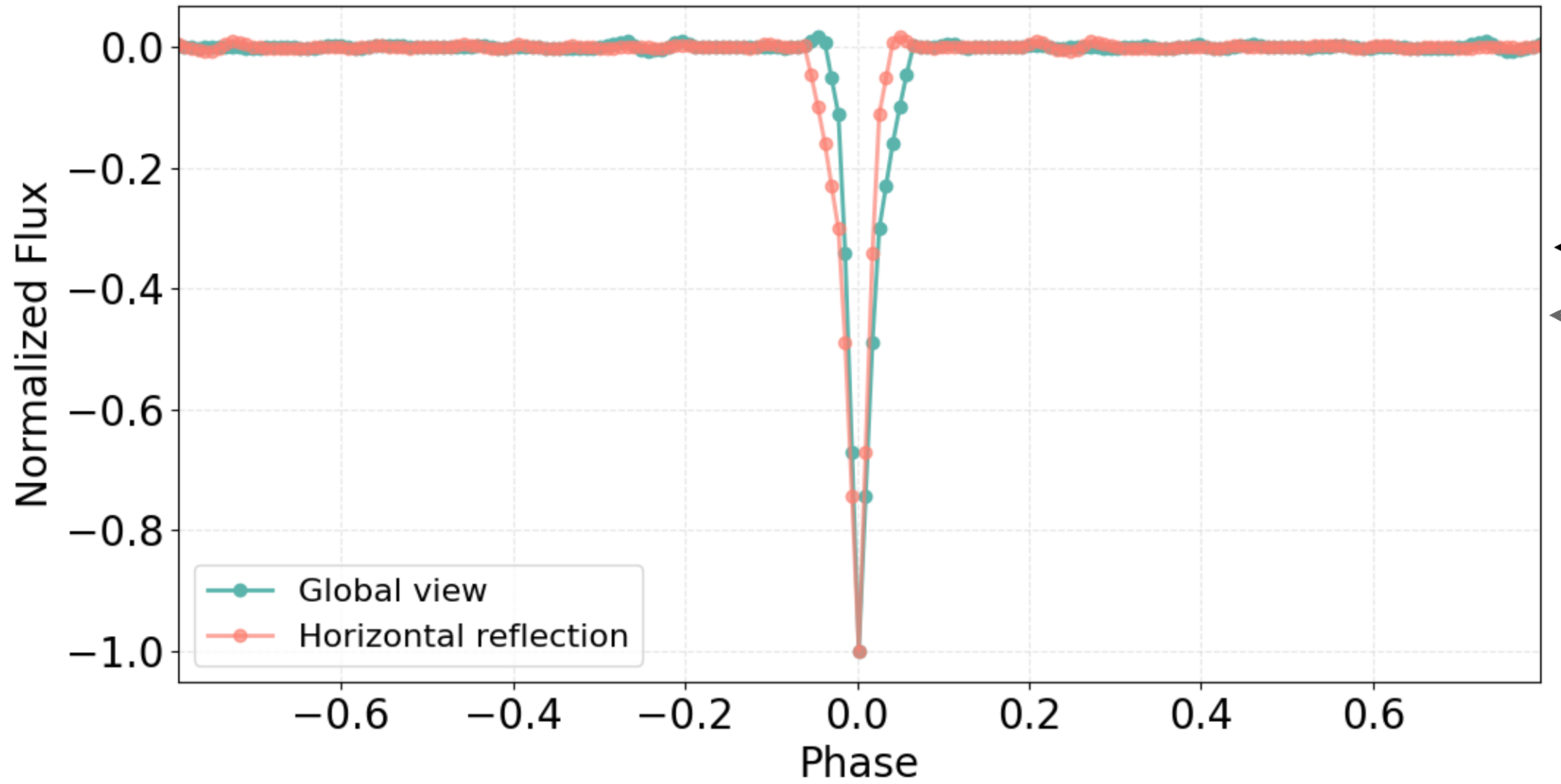
Kepler, TESS

3 classes

AFP
PC
J



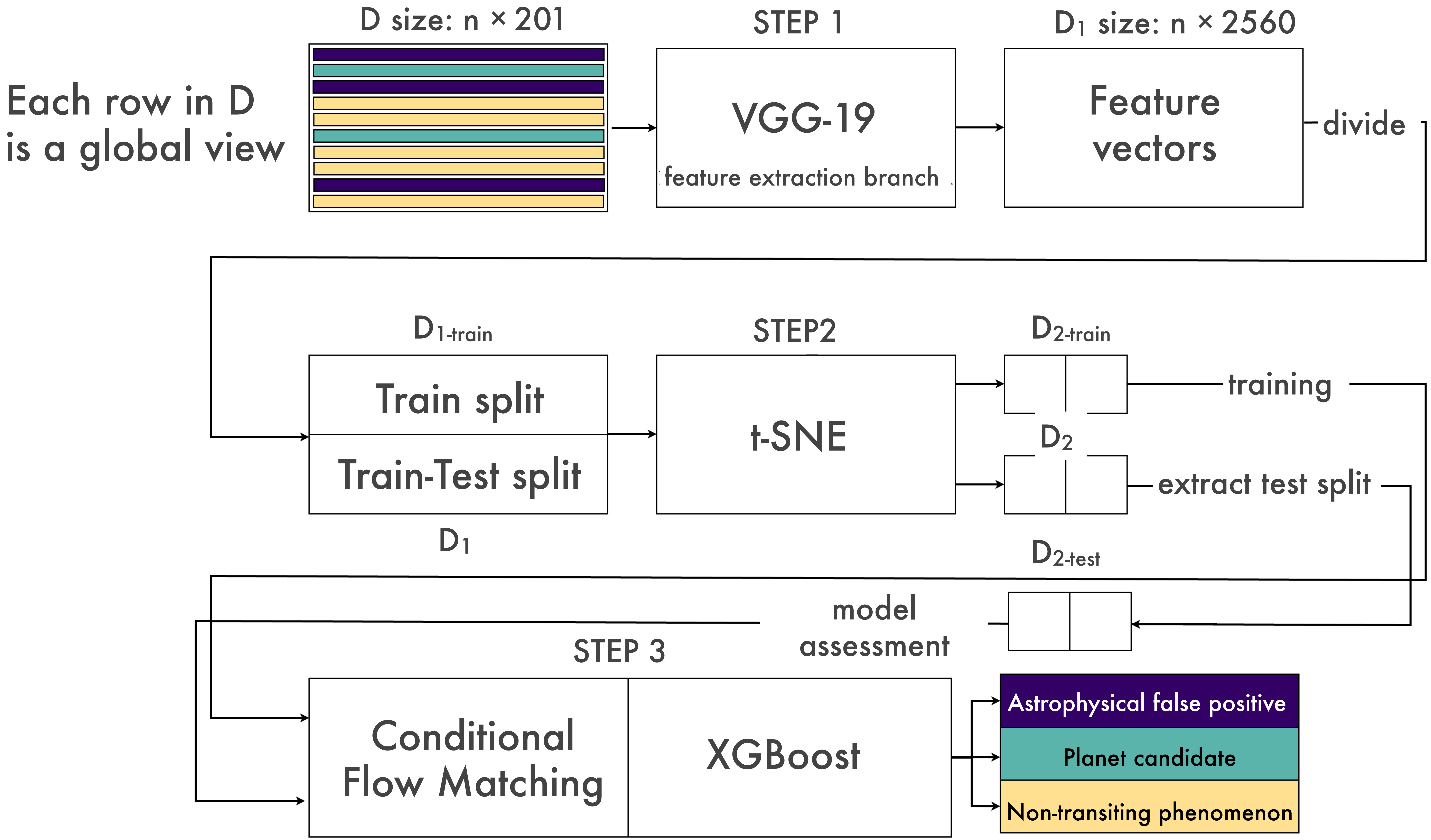




Machine Learning workflow

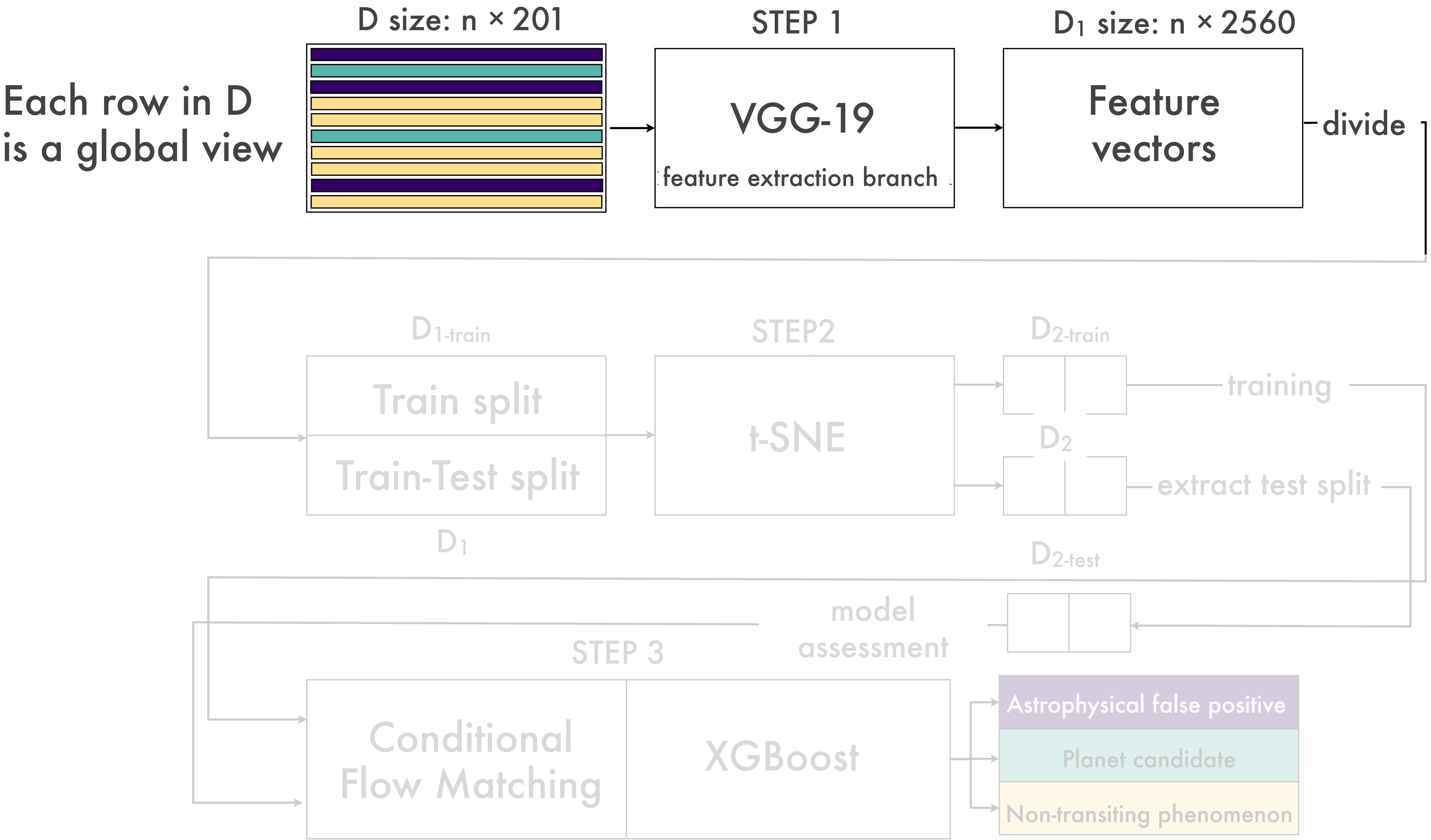
Machine Learning workflow

Feature Extraction, Dimensionality Reduction and Classification

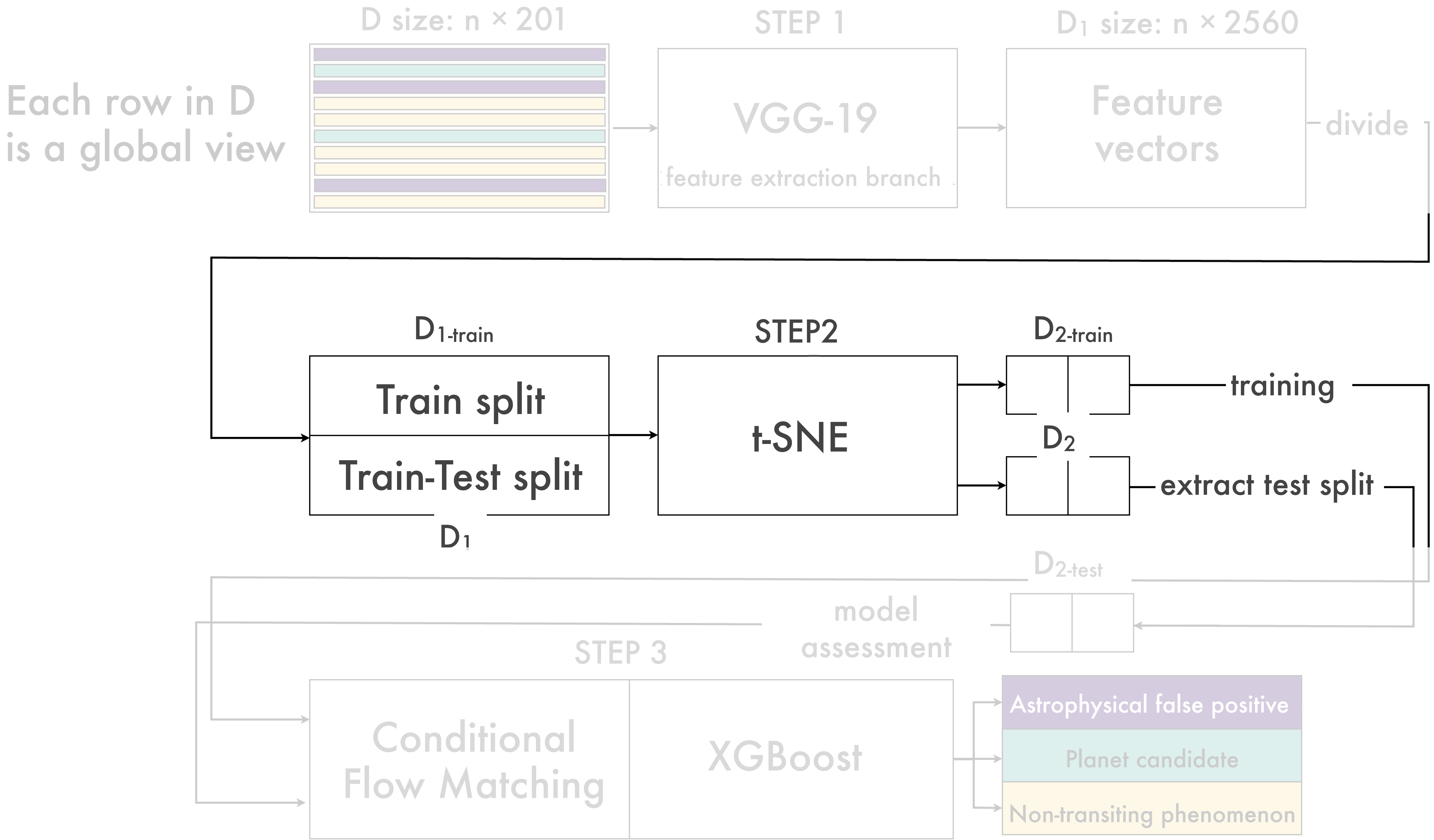


Machine Learning workflow

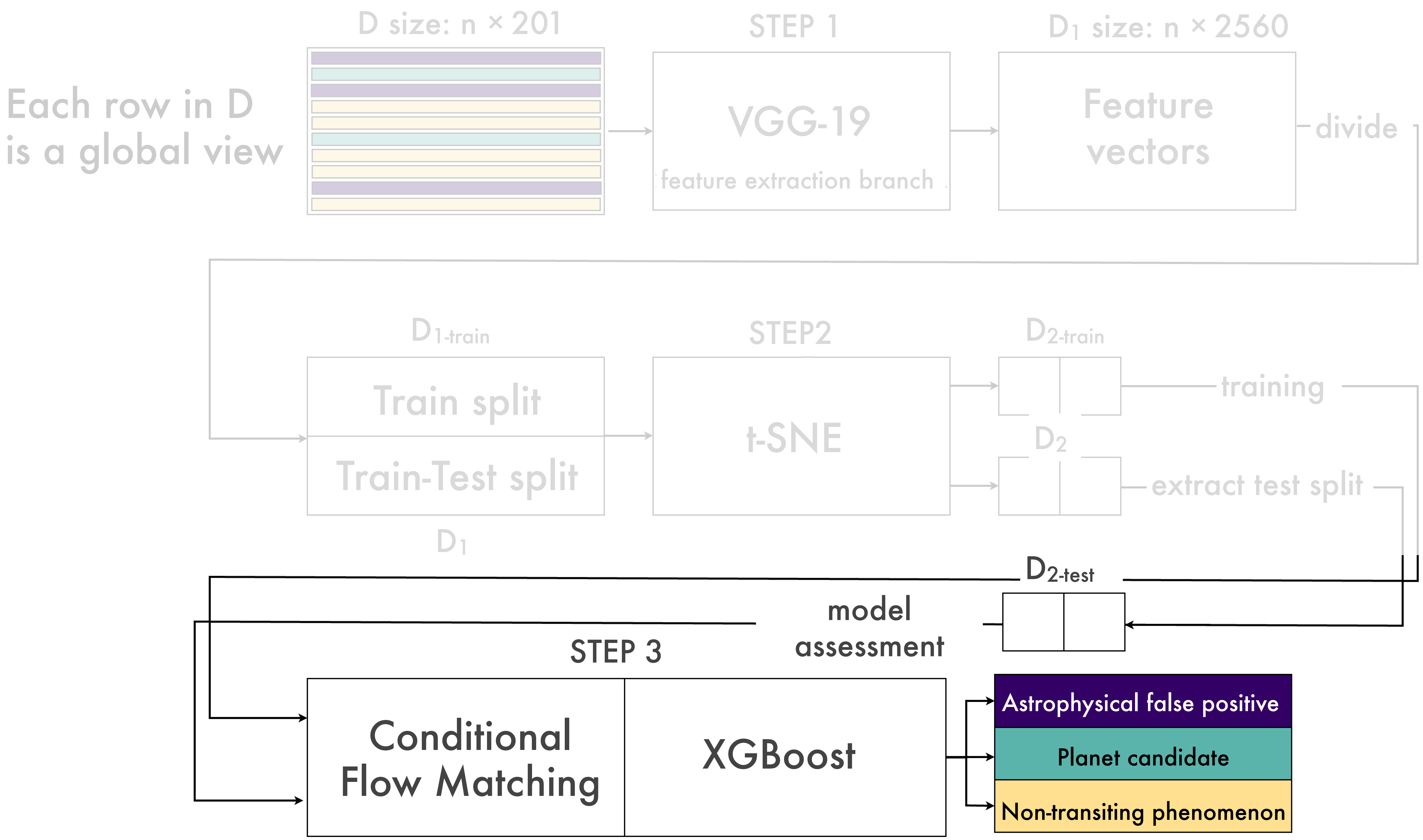
STEP 1: Feature Extraction



Machine Learning workflow
STEP 2: Dimensionality Reduction



Machine Learning workflow
STEP 3: Classification



Results

Application on Kepler and TESS, comparison with other models, the contribution of VGG and t-SNE

Results

Application on Kepler and TESS data

Dataset	Class	Precision	Recall	F1-Score	Misclass. Rate (%)
Kepler Q1–Q17 DR24	AFP	0.9943	0.9932	0.9937	1.25
	PC	0.9972	0.9986	0.9979	0.42
	NTP	0.9803	0.9803	0.9803	3.93
Kepler Q1–Q17 DR25	AFP	0.910	0.985	0.946	2.1
	PC	0.971	0.996	0.983	
	NTP	0.997	0.972	0.984	
TESS TEY23	B	1.000	1.000	1.000	0.0
	E	1.000	1.000	1.000	
	J	1.000	1.000	1.000	

Table 1. Classification performance of the model across three datasets: Kepler Q1–Q17 Data Release (DR) 24, Kepler Q1-Q17 Data Release 25, and TESS TEY23 (Evan Tey *et al* 2023 *AJ* **165** 95). The metrics have been computed on test samples.

Results

Comparison with state-of-the art vetting models

Model [Ref.]	Survey	Precision	Recall	F1-Score
SOM [39]	Kepler	0.864	0.865	0.864
SOM [39]	K2	0.945	0.972	0.958
RFC + SOM [30]	NGTS	0.901	0.914	0.907
Exominer [21]	Kepler	0.968	0.974	0.971
Exominer-Basic [21]	TESS	0.88	0.73	0.79
Astronet-Triage-v2 [22]	TESS	0.84	0.99	0.909
Transformer [42]	TESS	0.809	0.8	0.805
This work	Kepler	**0.974**	**0.987**	**0.980**
This work	TESS	**1.0**	**1.0**	**1.0**

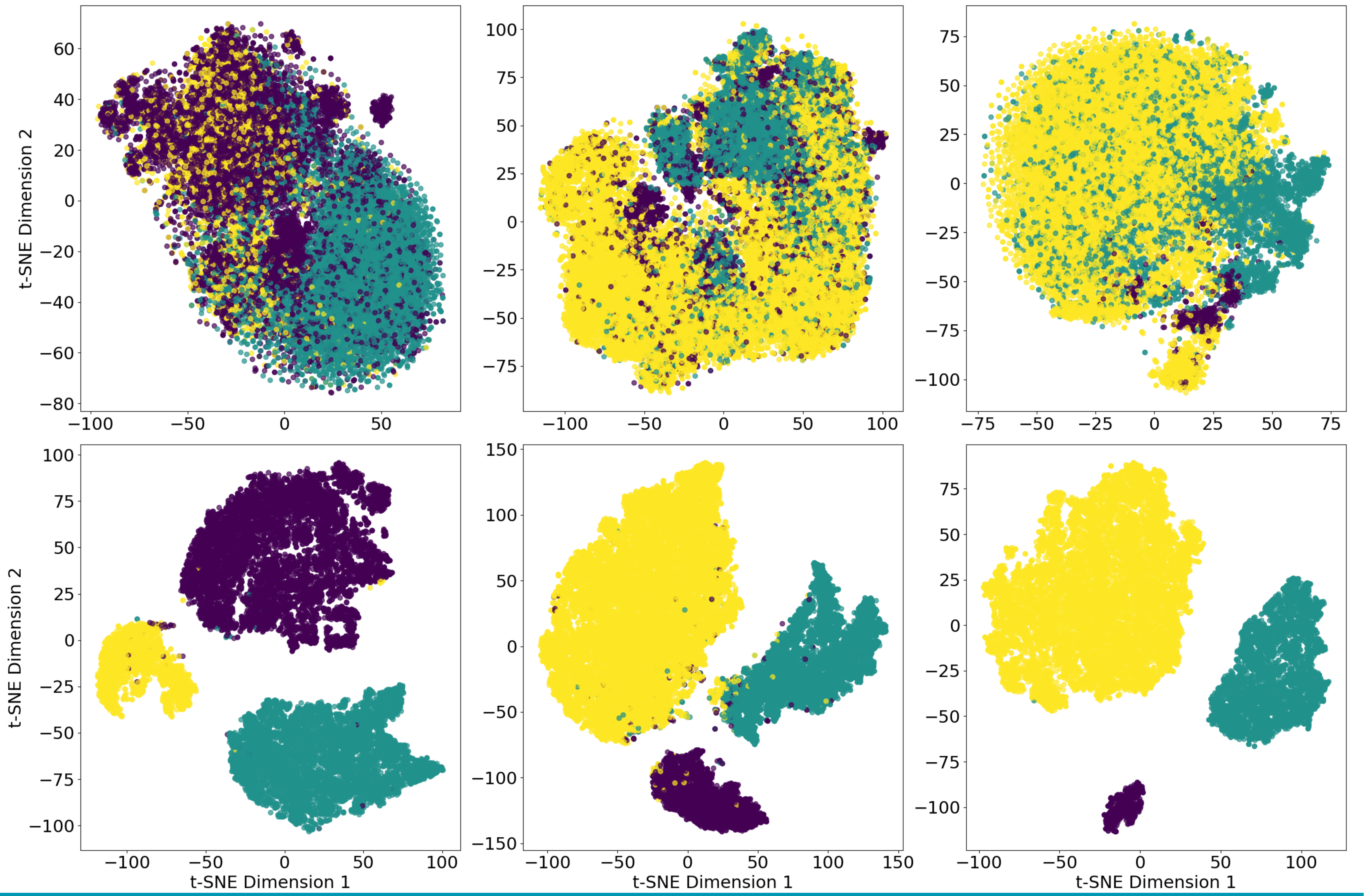
Table 2. Performance of different vetting models. Our precision, recall, and F1-scores for Kepler data are computed by averaging the scores of Table 1 obtained on each class. Other model scores are taken from the reference manuscripts. The best results on Kepler and TESS datasets are highlighted in boldface.

Results

The power of VGG-19 in extracting relevant patterns from data

Features extracted by DART-Vetter (top row) and VGG19 (bottom row) in the two dimensional embedding defined by t-SNE.

(Left) Kepler Q1-Q17 DR24
(Middle) Kepler Q1-Q17 DR25
(Right) TESS TEY23

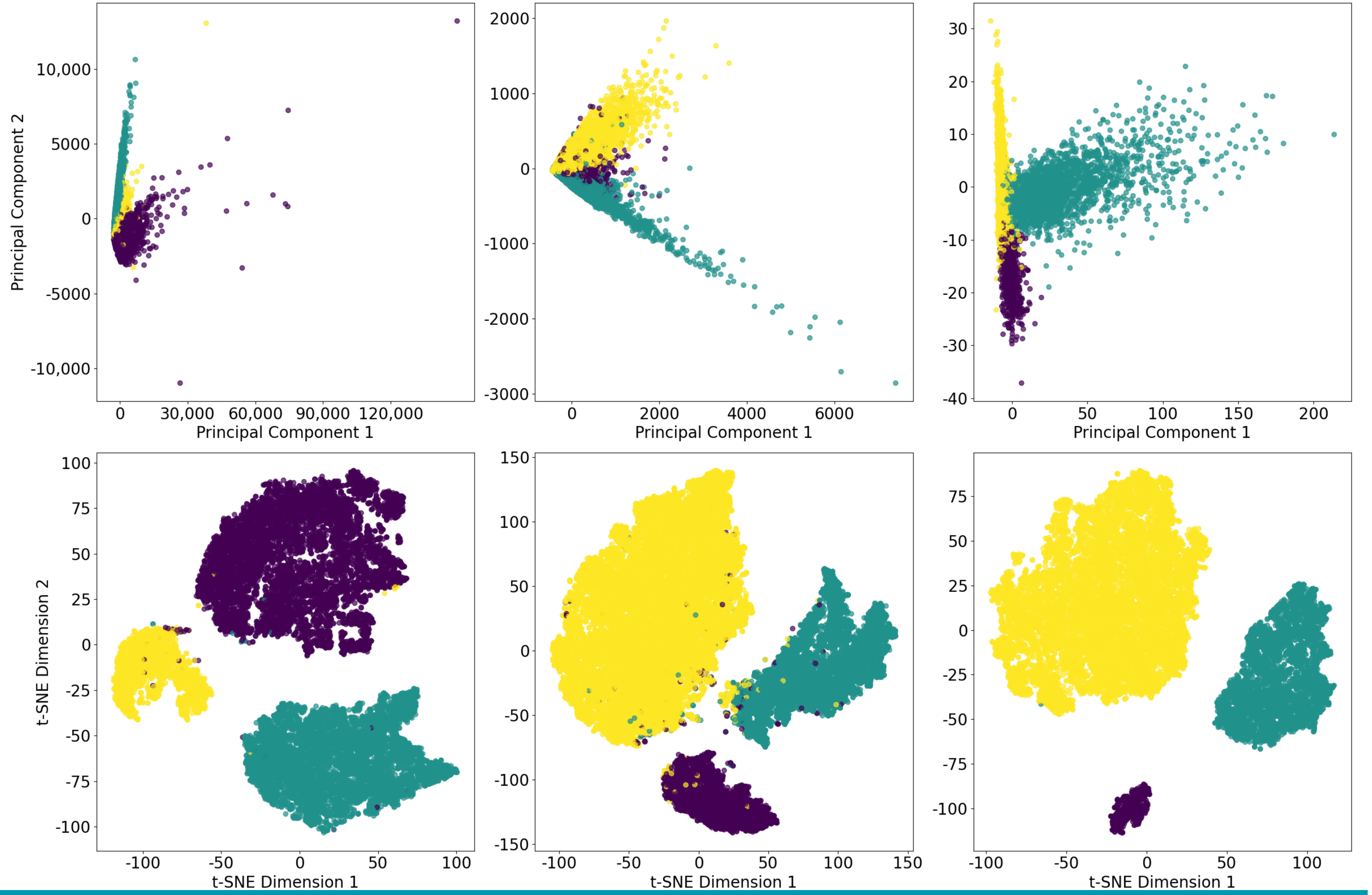


Results

The power of t-SNE in enhancing class separability

Comparison between PCA (top row) and t-SNE (bottom row) in dimensionality reduction.

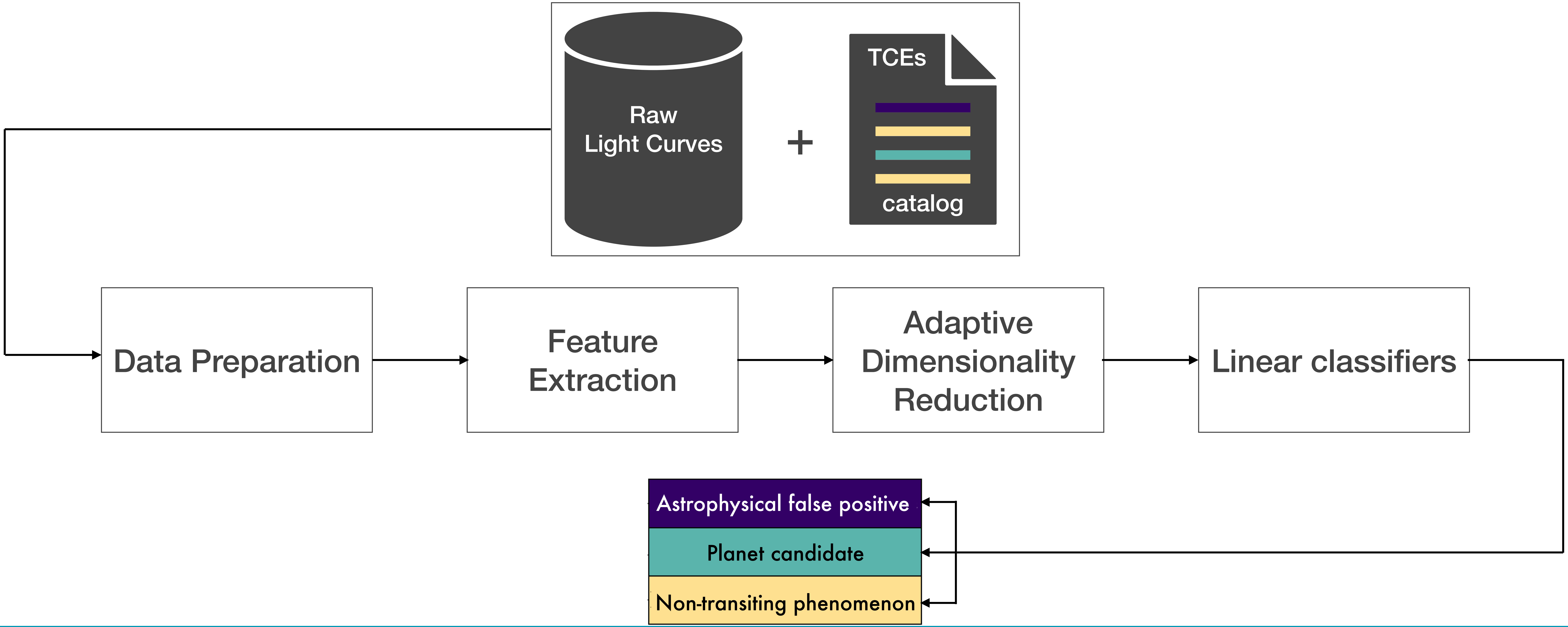
(Left) Kepler Q1-Q17 DR24
(Middle) Kepler Q1-Q17 DR25
(Right) TESS TEY23



Future Directions

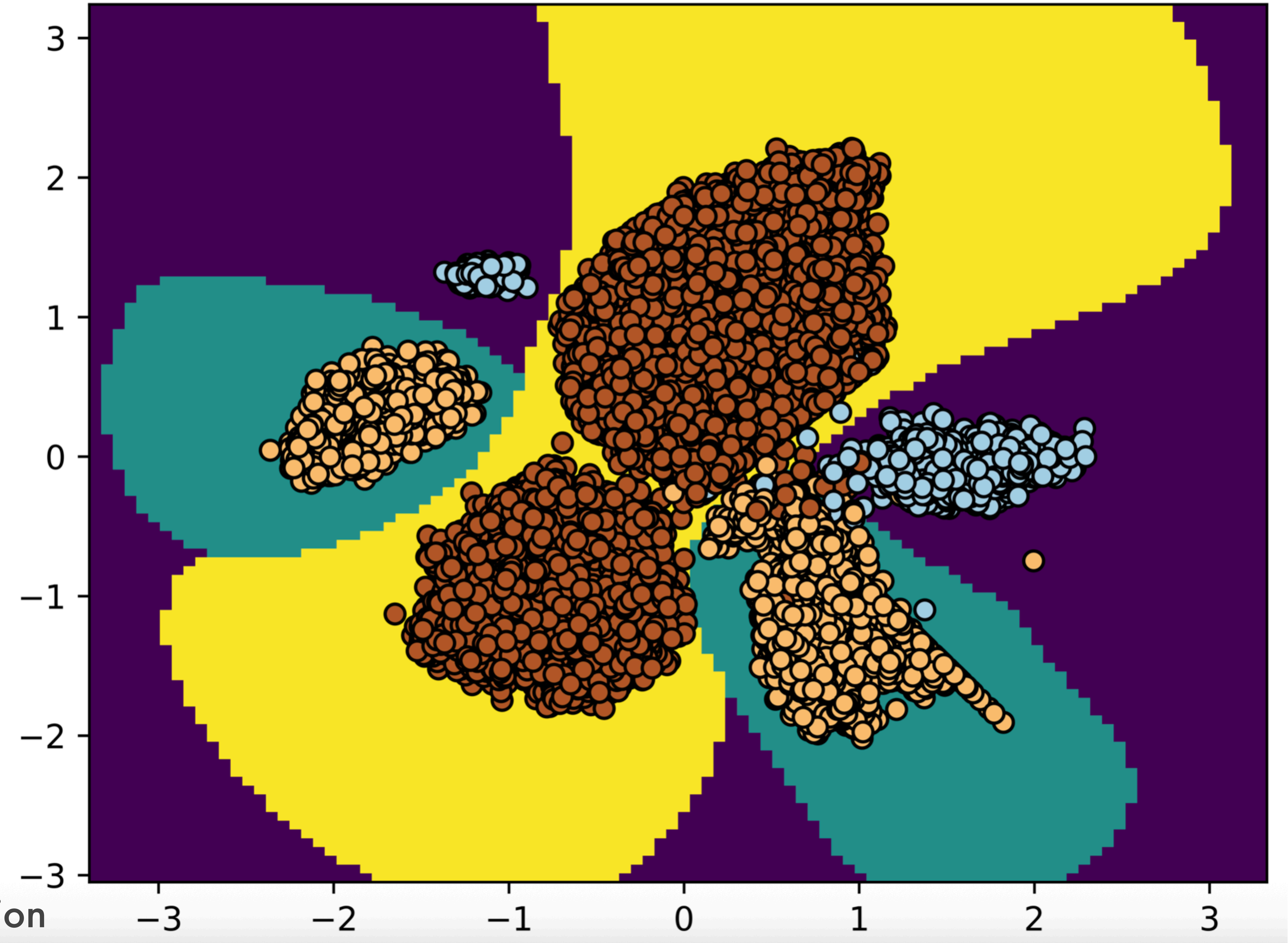
Future Directions

A. Application to simulated PLATO data (ongoing work at Laboratoire d'Astrophysique de Marseille);
B. Development of an end-to-end pipeline for processing data from *Kepler*, TESS and PLATO missions.



Application on multiple domains

Kepler DR25 & TESS TEY23



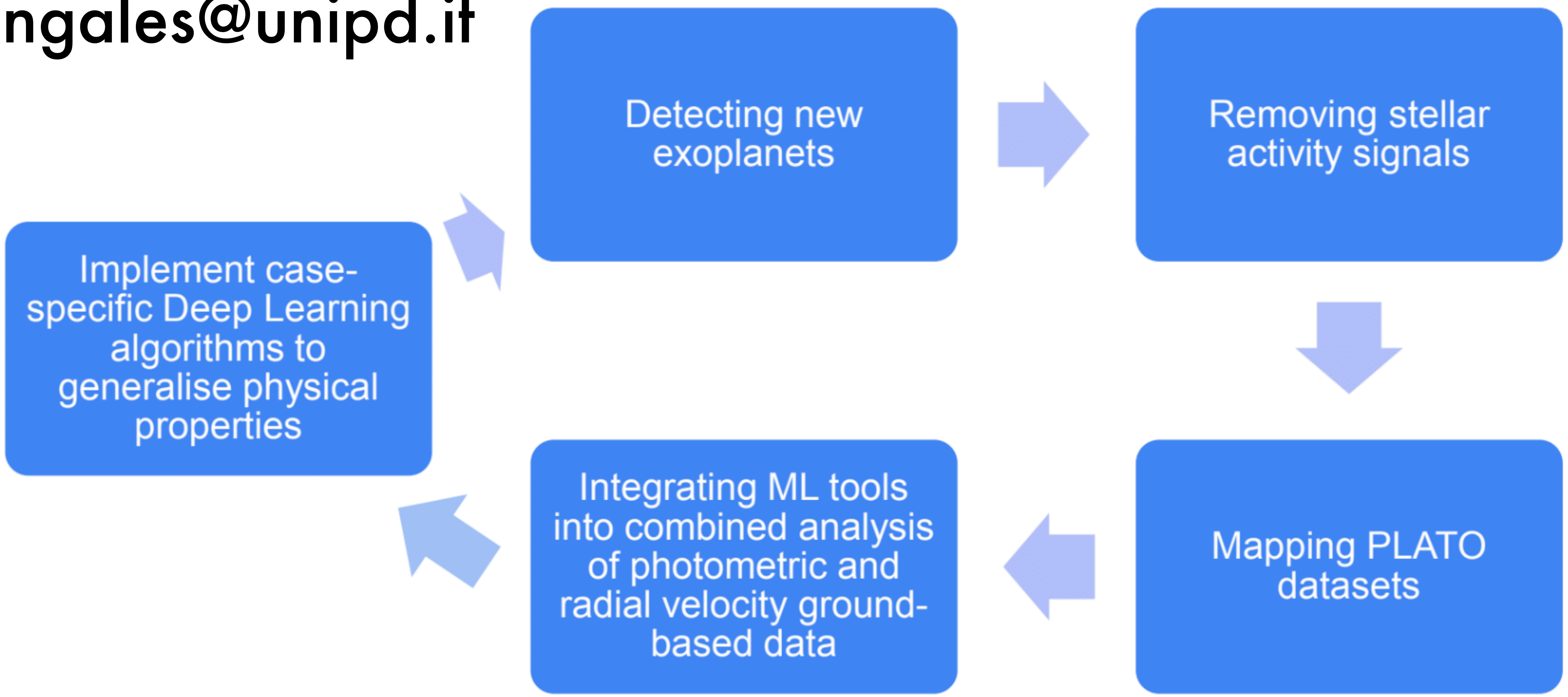
Fiscale et al., in preparation

PLATO Machine Learning WG

Aim: Development of Machine Learning tools to support PLATO Working Packages focused on data analysis.

WG Leader: Tiziano Zingales

tiziano.zingales@unipd.it



Stefano Fiscale

Third-year PhD student

PhD supervisor: Alessandra Rotundi
PhD tutor: Laura Inno, Alessio Ferone

Mail: stefano.fiscale001@studenti.uniparthenope.it
Website: https://stefanofisc.github.io/

esa plato

UNIVERSITA' DEGLI STUDI DI NAPOLI PARTHENOPE

INAF
ISTITUTO NAZIONALE
DI ASTROFISICA

LAM
LABORATOIRE D'ASTROPHYSIQUE
DE MARSEILLE