

# **DETECÇÃO AUTOMÁTICA DE EXOPLANETAS UTILIZANDO ALGORITMOS DE APRENDIZADO DE MÁQUINA.**

**Orientado(a): Bruno Henrique Dourado Macedo**

**Orientador(a): Willian Zalewski**

Voluntário IC;

E-mail: bhd.macedo.2017@aluno.unila.edu.br

E-mail do Orientador: willian.zalewski@unila.edu.br

Grande Área do Conhecimento: Ciências Exatas e da Terra (CET)

Sigla do Instituto: ILATIT

# Sumário

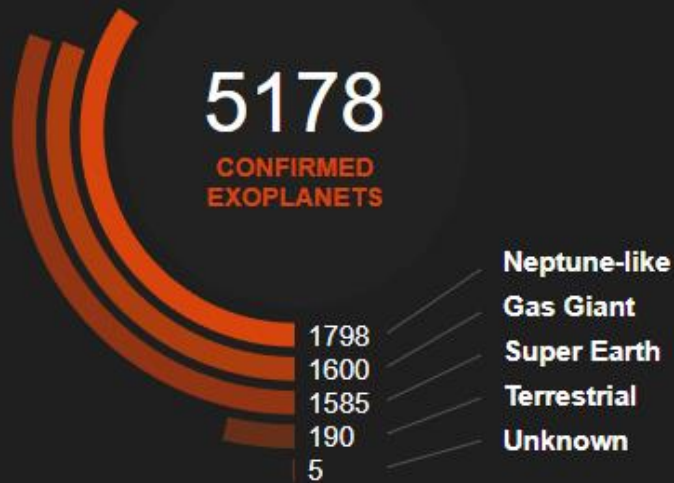
- Introdução
- Objetivos
- Métodos
- Resultados
- Conclusão
- Referências

# Introdução

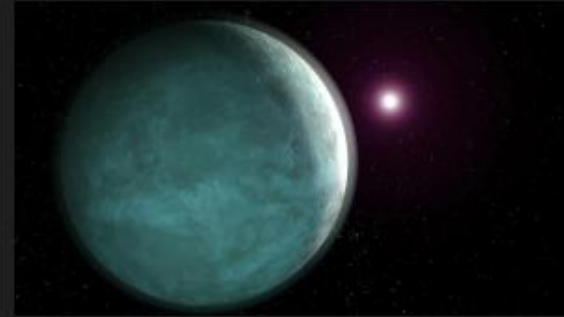
- Exoplanetas;
- Telescópio Espacial Kepler;
- Transito Planetário;
- Serie temporal;

# Exoplaneta

## Planet Types



## New Discovery



PLANET NAME  
**TOI-5238 b**

PLANET TYPE  
**Neptune-like**

DISCOVERY DATE  
**2022**

DETECTION METHOD  
**Transit**

Fonte: <https://exoplanets.nasa.gov/>

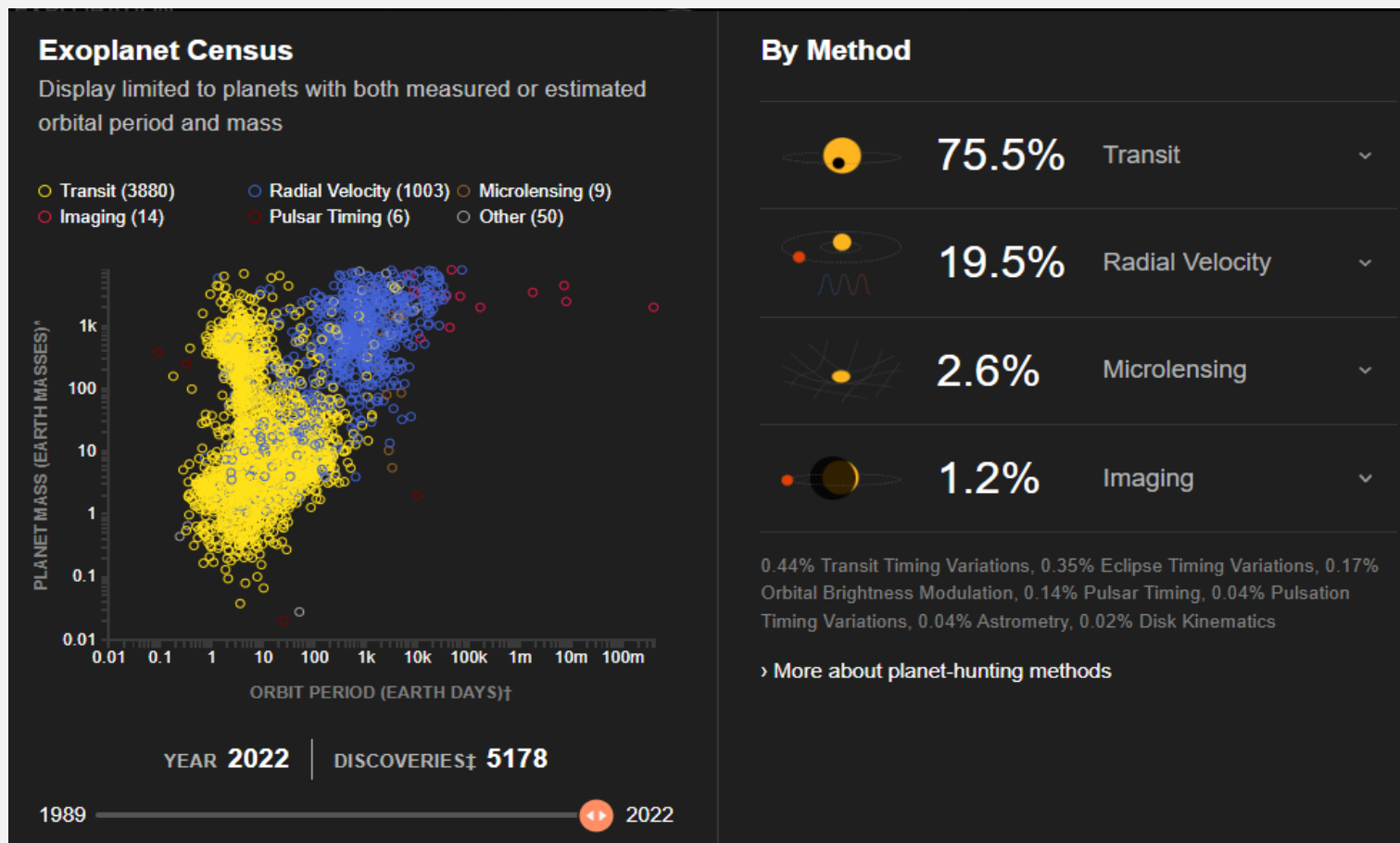


**4ª SIEPE** SEMANA INTEGRADA  
ENSINO | PESQUISA | EXTENSÃO

**EICTI**

XI ENCONTRO ANUAL DE INICIAÇÃO  
CIENTÍFICA E VII ENCONTRO DE INICIAÇÃO EM  
DESENVOLVIMENTO; TECNOLÓGICO E INOVAÇÃO





Fonte: <https://exoplanets.nasa.gov/>



4ª SIEPE

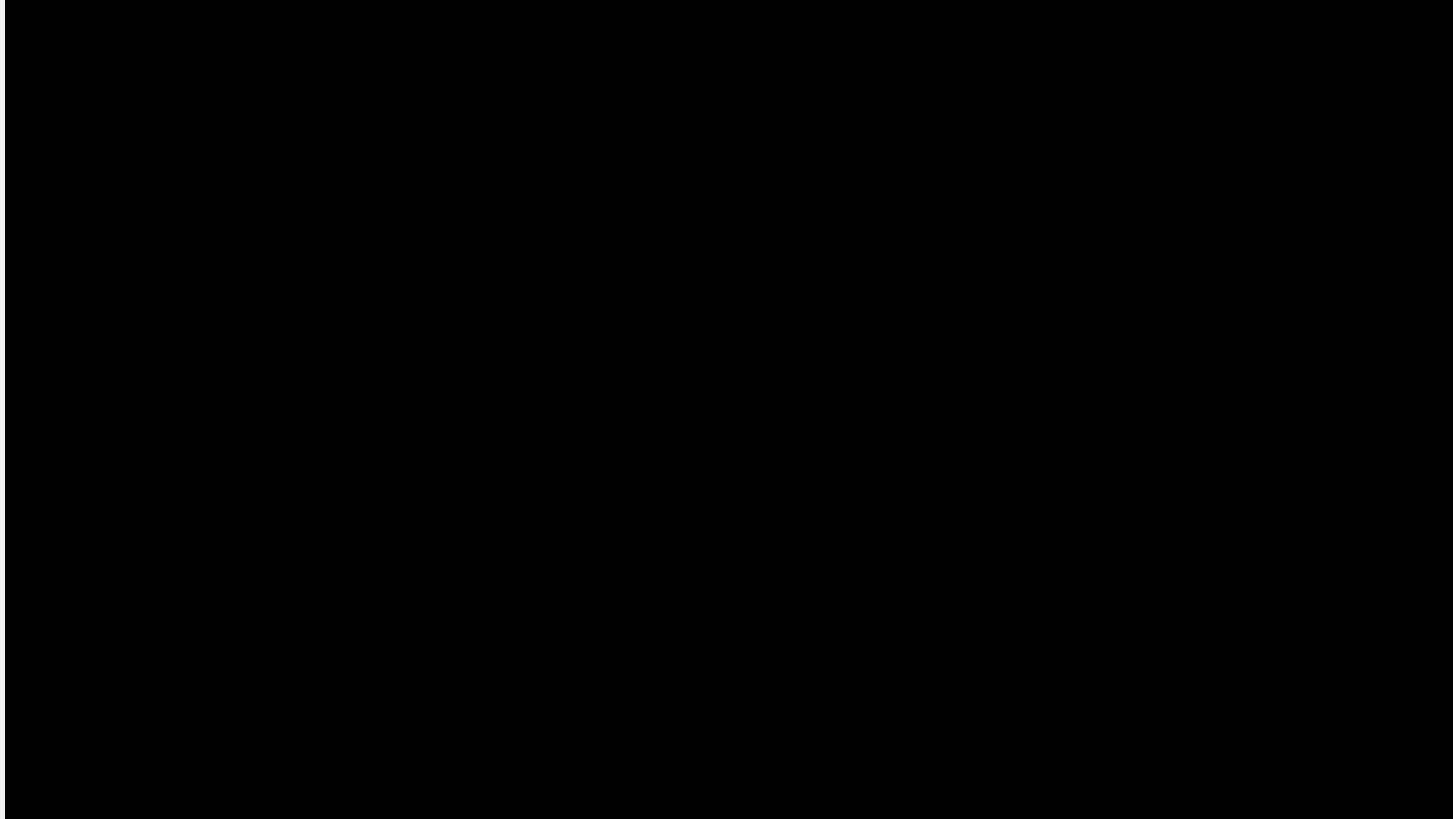
SEMANA INTEGRADA  
ENSINO | PESQUISA | EXTENSÃO

EICTI

XI ENCONTRO ANUAL DE INICIAÇÃO  
CIENTÍFICA E VII ENCONTRO DE INICIAÇÃO EM  
DESENVOLVIMENTO; TECNOLÓGICO E INOVAÇÃO



# Trânsito planetário



Fonte: NASA Vídeo: <https://www.youtube.com/watch?v=BFi4HB UdWkk>



**4ª SIEPE**

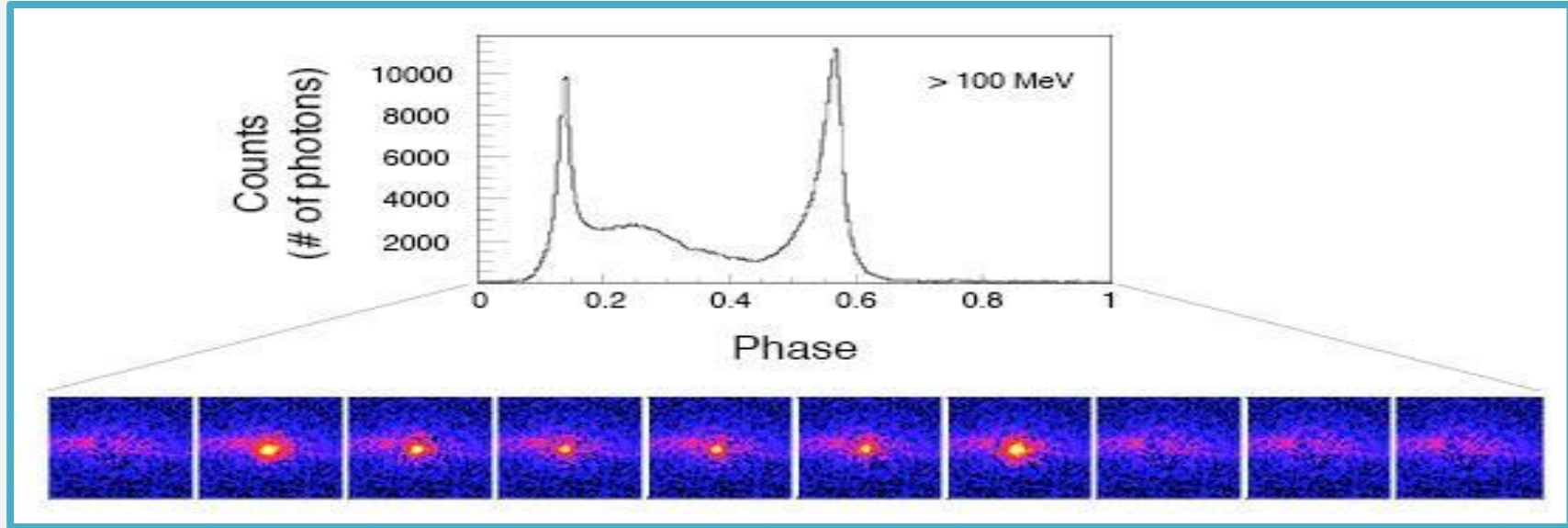
**SEMANA INTEGRADA**  
ENSINO | PESQUISA | EXTENSÃO

**EICTI**

XI ENCONTRO ANUAL DE INICIAÇÃO  
CIENTÍFICA E VII ENCONTRO DE INICIAÇÃO EM  
DESENVOLVIMENTO; TECNOLÓGICO E INOVAÇÃO



# Trânsito planetário



Fonte: Kepler and K2 Science Center, 2018.



4<sup>a</sup> SIEPE

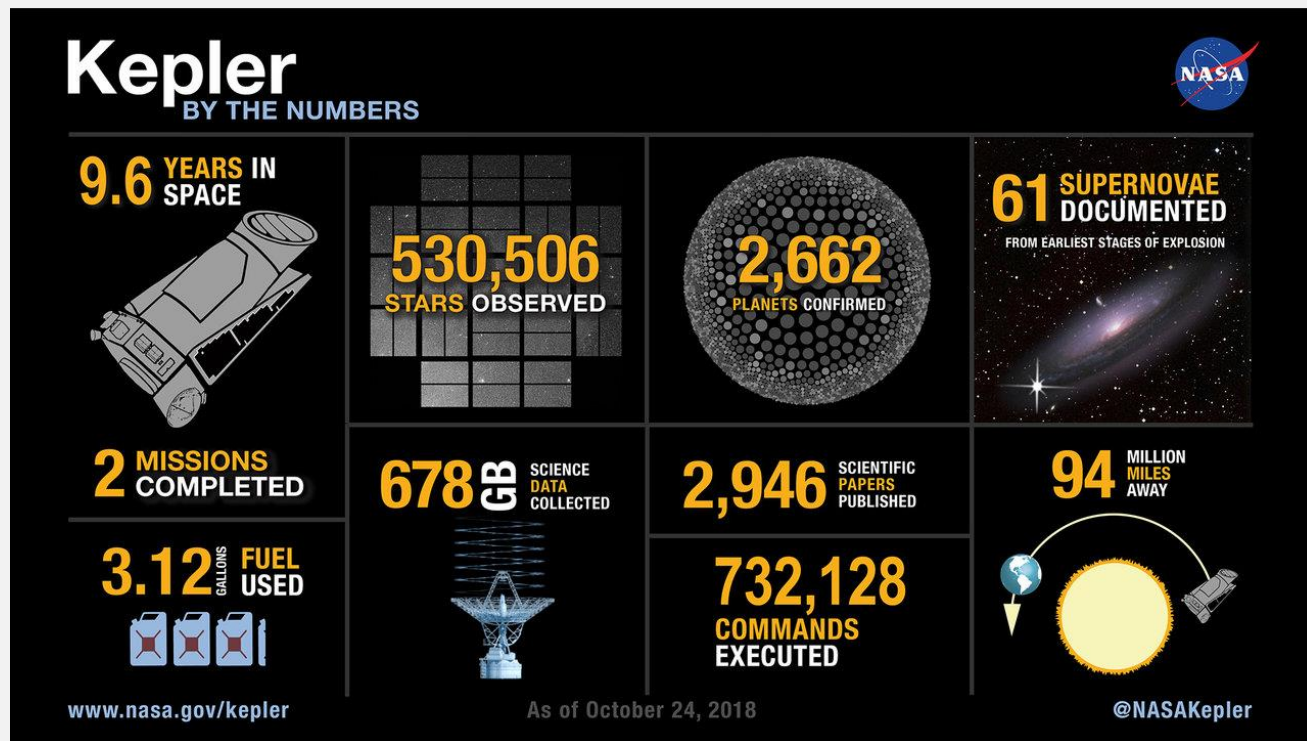
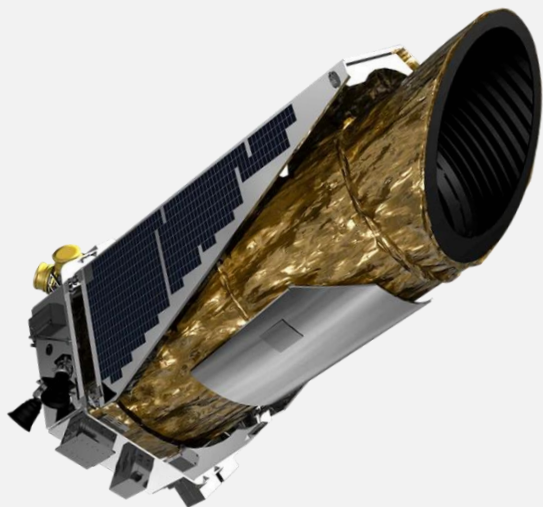
SEMANA INTEGRADA  
ENSINO | PESQUISA | EXTENSÃO

EICTI

XI ENCONTRO ANUAL DE INICIAÇÃO  
CIENTÍFICA E VII ENCONTRO DE INICIAÇÃO EM  
DESENVOLVIMENTO; TECNOLÓGICO E INOVAÇÃO



# Telescópio Espacial Kepler



Fonte: <https://exoplanets.nasa.gov/>



**4ª SIEPE**

SEMANA INTEGRADA  
ENSINO | PESQUISA | EXTENSÃO

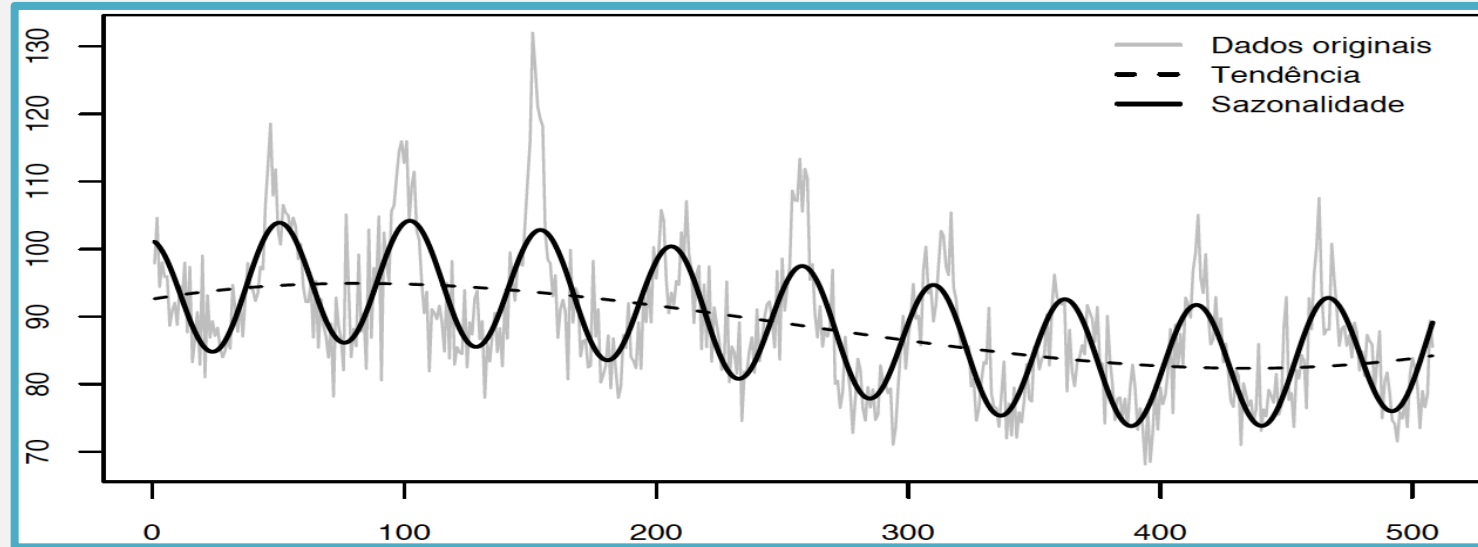
**EICTI**

XI ENCONTRO ANUAL DE INICIAÇÃO  
CIENTÍFICA E VII ENCONTRO DE INICIAÇÃO EM  
DESENVOLVIMENTO; TECNOLÓGICO E INOVAÇÃO





# Série Temporal



Fonte: Ferrero (2009).



**4ª SIEPE**

**SEMANA INTEGRADA**  
ENSINO | PESQUISA | EXTENSÃO

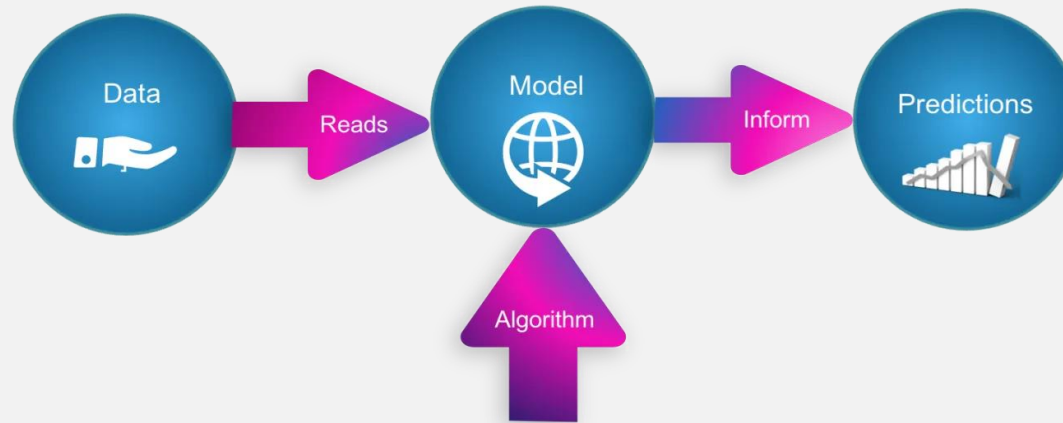
**EICTI**

XI ENCONTRO ANUAL DE INICIAÇÃO  
CIENTÍFICA E VII ENCONTRO DE INICIAÇÃO EM  
DESENVOLVIMENTO; TECNOLÓGICO E INOVAÇÃO



# Objetivo

- Aplicação de IA para auxiliar no processo de detecção automática de exoplanetas;



**4ª SIEPE**

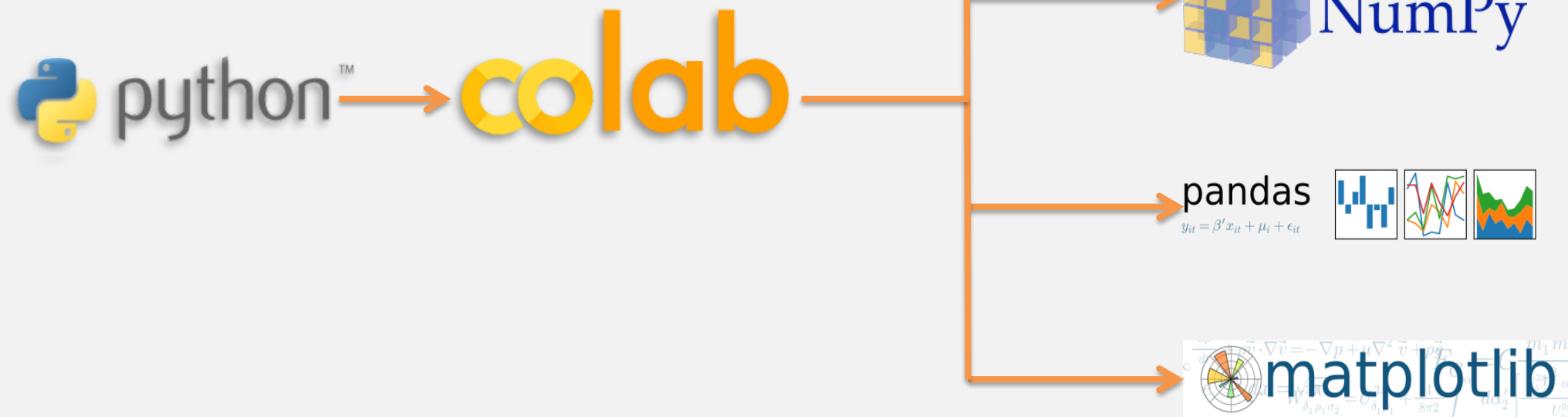
**SEMANA INTEGRADA**  
ENSINO | PESQUISA | EXTENSÃO

**EICTI**

XI ENCONTRO ANUAL DE INICIAÇÃO  
CIENTÍFICA E VII ENCONTRO DE INICIAÇÃO EM  
DESENVOLVIMENTO; TECNOLÓGICO E INOVAÇÃO



# Método



4ª SIEPE

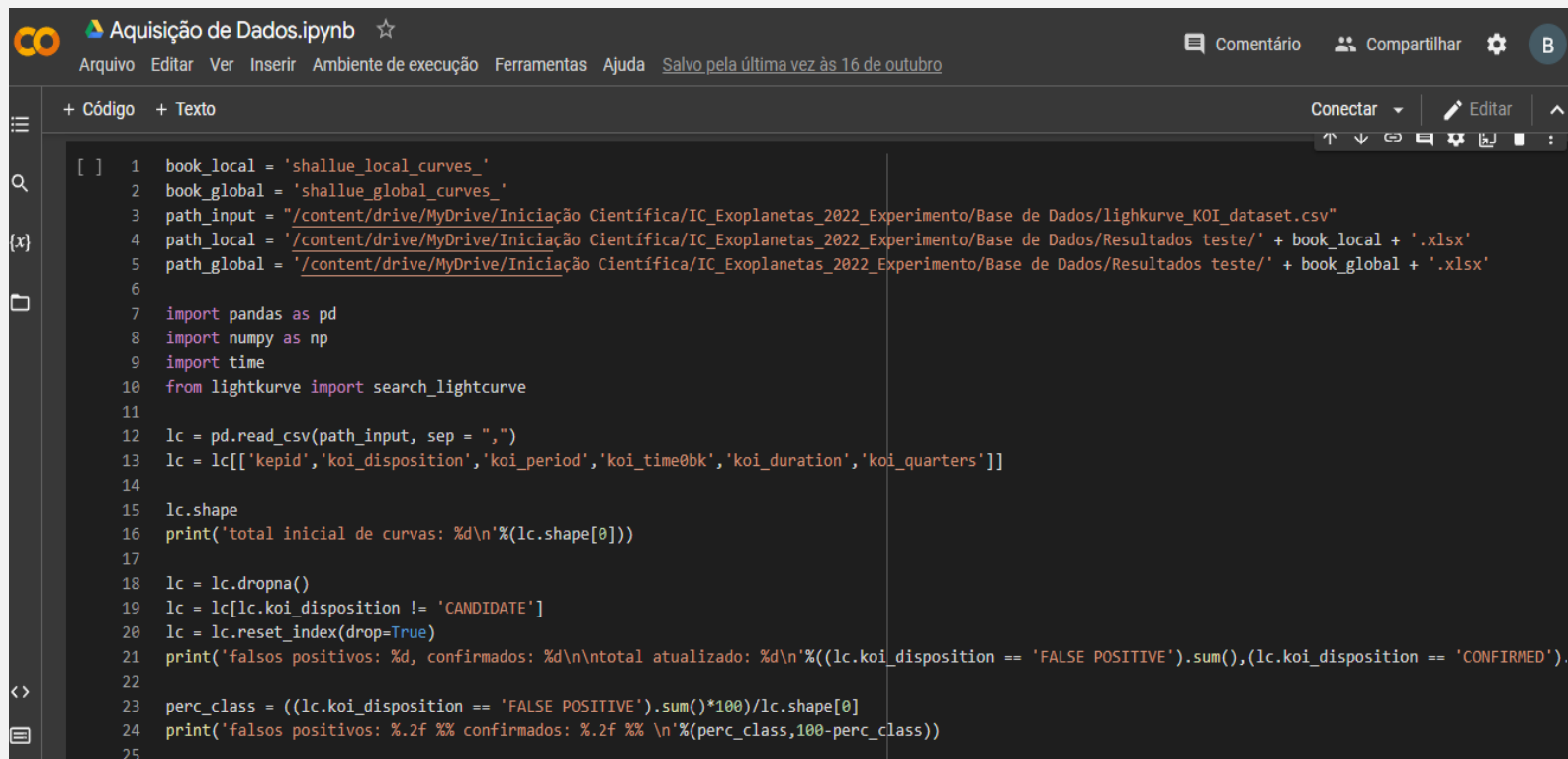
SEMANA INTEGRADA  
ENSINO | PESQUISA | EXTENSÃO

EICTI

XI ENCONTRO ANUAL DE INICIAÇÃO  
CIENTÍFICA E VII ENCONTRO DE INICIAÇÃO EM  
DESENVOLVIMENTO; TECNOLÓGICO E INOVAÇÃO



# Aquisição e Pré-processamento



```
[ ] 1 book_local = 'shallue_local_curves_'
2 book_global = 'shallue_global_curves_'
3 path_input = "/content/drive/MyDrive/Iniciação Científica/IC_Exoplanetas_2022_Experimento/Base de Dados/ligthkurve_KOI_dataset.csv"
4 path_local = '/content/drive/MyDrive/Iniciação Científica/IC_Exoplanetas_2022_Experimento/Base de Dados/Resultados teste/' + book_local + '.xlsx'
5 path_global = '/content/drive/MyDrive/Iniciação Científica/IC_Exoplanetas_2022_Experimento/Base de Dados/Resultados teste/' + book_global + '.xlsx'
6
7 import pandas as pd
8 import numpy as np
9 import time
10 from lightkurve import search_lightcurve
11
12 lc = pd.read_csv(path_input, sep = ",")
13 lc = lc[['kepid', 'koi_disposition', 'koi_period', 'koi_time0bk', 'koi_duration', 'koi_quarters']]
14
15 lc.shape
16 print('total inicial de curvas: %d\n'%(lc.shape[0]))
17
18 lc = lc.dropna()
19 lc = lc[lc.koi_disposition != 'CANDIDATE']
20 lc = lc.reset_index(drop=True)
21 print('falsos positivos: %d, confirmados: %d\n\ntotal atualizado: %d\n'%((lc.koi_disposition == 'FALSE POSITIVE').sum(), (lc.koi_disposition == 'CONFIRMED').
22
23 perc_class = ((lc.koi_disposition == 'FALSE POSITIVE').sum()*100)/lc.shape[0]
24 print('falsos positivos: %.2f %% confirmados: %.2f %% \n'%(perc_class, 100-perc_class))
25
```

Fonte: Autoria Própria (2022).



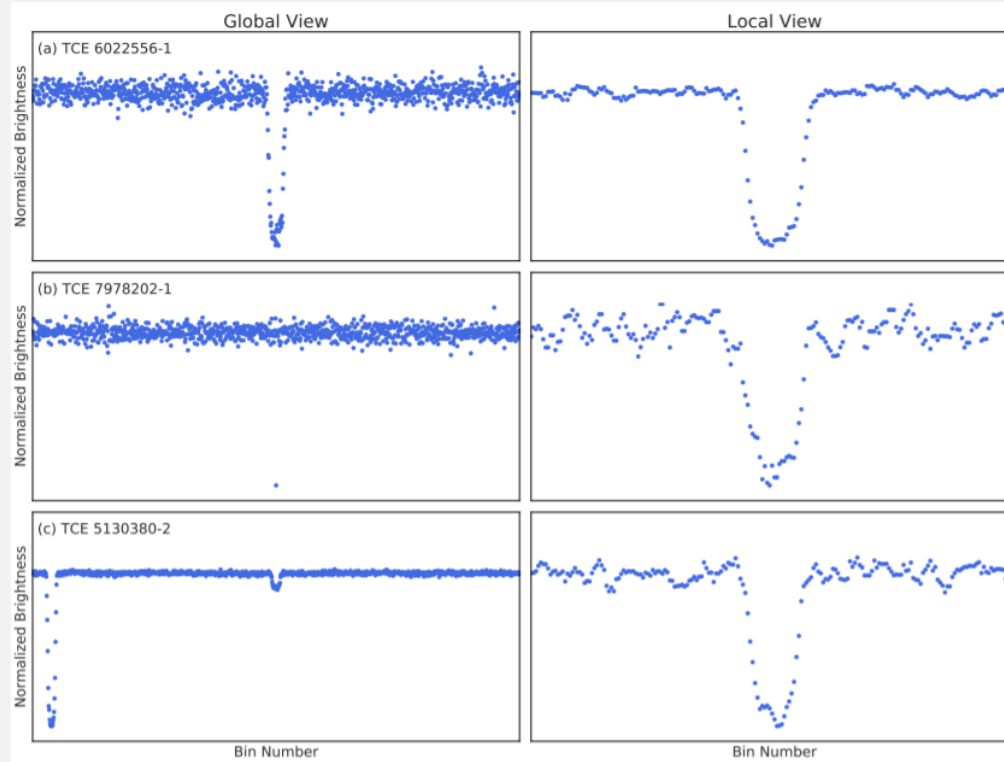
**4ª SIEPE** SEMANA INTEGRADA  
ENSINO | PESQUISA | EXTENSÃO

**EICTI**

XI ENCONTRO ANUAL DE INICIAÇÃO  
CIENTÍFICA E VII ENCONTRO DE INICIAÇÃO EM  
DESENVOLVIMENTO; TECNOLÓGICO E INOVAÇÃO



# Pré-processamento



Fonte: SHALLUE & VANDERBURG (2017)



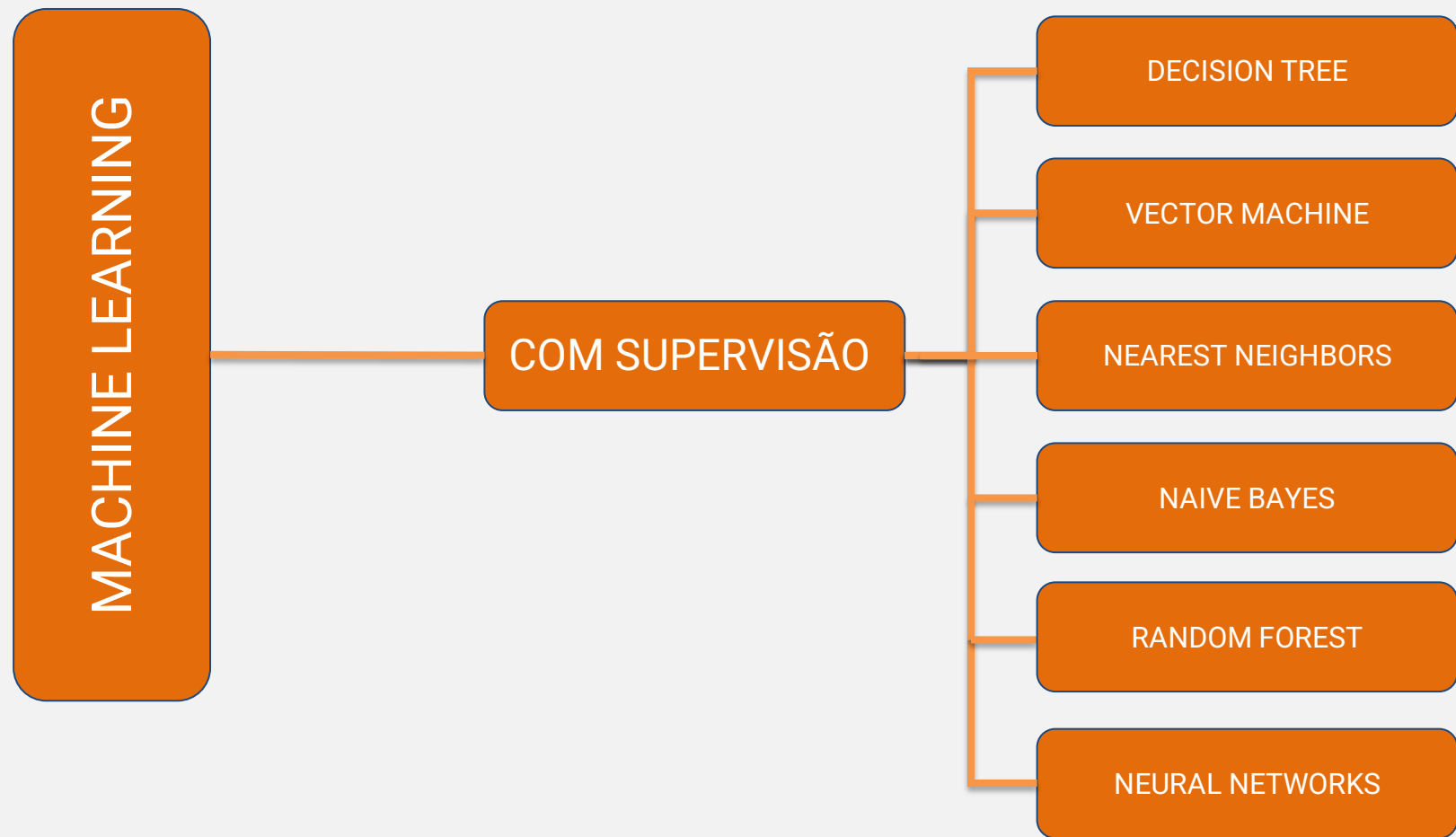
4<sup>a</sup> SIEPE

SEMANA INTEGRADA  
ENSINO | PESQUISA | EXTENSÃO

EICTI

XI ENCONTRO ANUAL DE INICIAÇÃO  
CIENTÍFICA E VII ENCONTRO DE INICIAÇÃO EM  
DESENVOLVIMENTO; TECNOLÓGICO E INOVAÇÃO





4<sup>a</sup> SIEPE

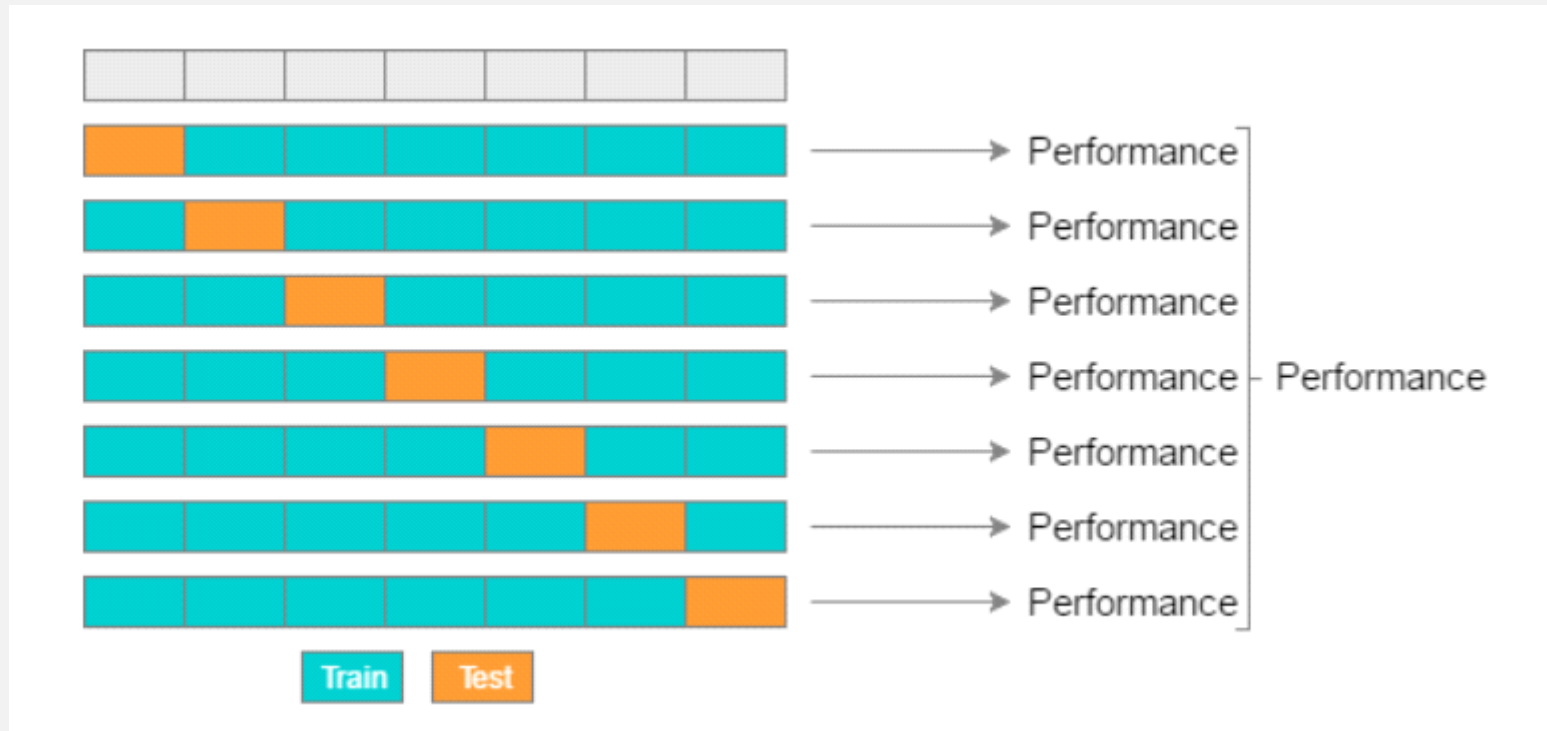
SEMANA INTEGRADA  
ENSINO | PESQUISA | EXTENSÃO

EICTI

XI ENCONTRO ANUAL DE INICIAÇÃO  
CIENTÍFICA E VII ENCONTRO DE INICIAÇÃO EM  
DESENVOLVIMENTO; TECNOLÓGICO E INOVAÇÃO



# Avaliação dos modelos



Fonte: Eduard Bonada i Cruells, Cross-Validation Strategies.



4ª SIEPE

SEMANA INTEGRADA  
ENSINO | PESQUISA | EXTENSÃO

EICTI

XI ENCONTRO ANUAL DE INICIAÇÃO  
CIENTÍFICA E VII ENCONTRO DE INICIAÇÃO EM  
DESENVOLVIMENTO; TECNOLÓGICO E INOVAÇÃO



# Resultados

Para a precisão dos algoritmos escolhemos a acurácia como o melhor indicador.

$$ac = \frac{VP + VN}{n}$$

Verdadeiros Positivos (VP)

Verdadeiros Negativos (VN);

$n = VP + VN + FP + FN$

(FACELI et al., 2011)



4<sup>a</sup> SIEPE

SEMANA INTEGRADA  
ENSINO | PESQUISA | EXTENSÃO

EICTI

XI ENCONTRO ANUAL DE INICIAÇÃO  
CIENTÍFICA E VII ENCONTRO DE INICIAÇÃO EM  
DESENVOLVIMENTO; TECNOLÓGICO E INOVAÇÃO





# Resultados

Foram divididos em duas tabelas, local e global os resultados.

Global								
Modelo	Acurácia		Precisão		Recall		F1	
	Média	Desvio	Média	Desvio	Média	Desvio	Média	Desvio
SVM	74,06%	3,73%	78,55%	2,21%	76,65%	7,85%	77,39%	4,18%
RF	72,20%	3,03%	75,59%	1,85%	77,39%	6,82%	76,48%	3,54%
DT	68,13%	1,87%	72,99%	1,91%	72,03%	5,63%	72,54%	2,60%
NB	70,54%	7,96%	79,93%	3,16%	66,47%	17,55%	71,30%	10,97%
NN	71,01%	2,33%	79,93%	2,72%	67,69%	5,21%	73,15%	2,87%
MLP	77,65%	4,17%	82,42%	4,12%	78,26%	10,62%	80,39%	4,83%

Local								
Modelo	Acurácia		Precisão		Recall		F1	
	Média	Desvio	Média	Desvio	Média	Desvio	Média	Desvio
SVM	68,85%	2,21%	74,57%	1,46%	71,59%	5,15%	72,52%	2,83%
RF	66,01%	1,22%	70,58%	0,73%	71,98%	1,79%	71,26%	1,27%
DT	63,38%	1,10%	68,80%	0,64%	68,39%	2,96%	68,68%	1,62%
NB	68,38%	1,95%	80,31%	1,58%	60,97%	3,21%	69,27%	2,41%
NN	66,64%	1,45%	69,05%	1,61%	78,23%	2,61%	73,31%	1,14%
MLP	70,32%	1,46%	74,90%	1,13%	74,14%	3,84%	74,17%	1,76%

Fonte: Autoria Própria (2022).



4ª SIEPE

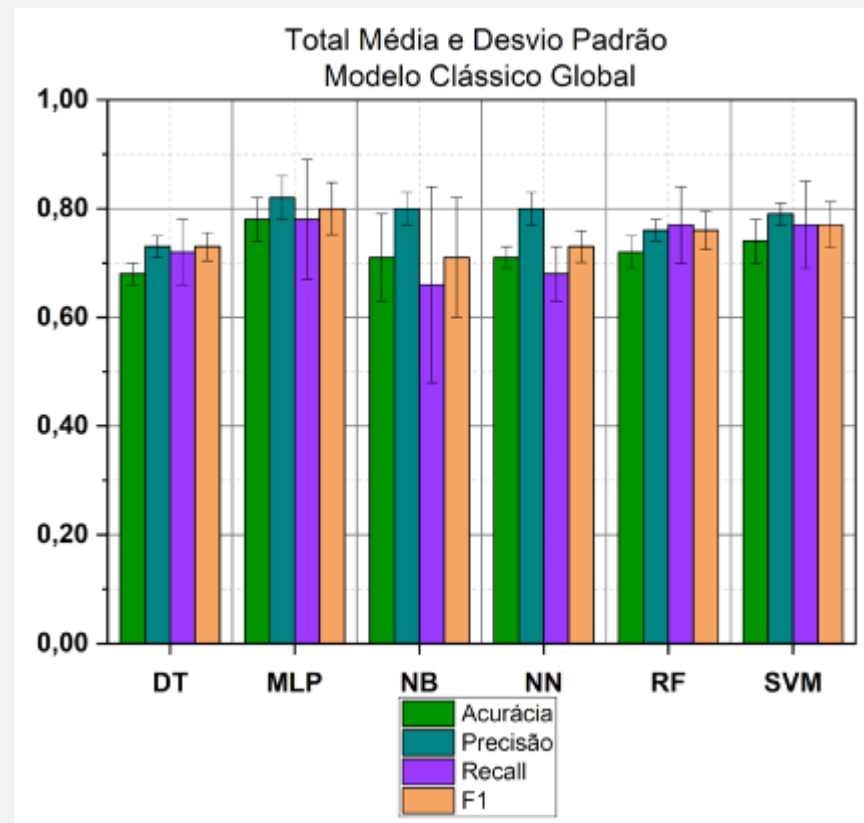
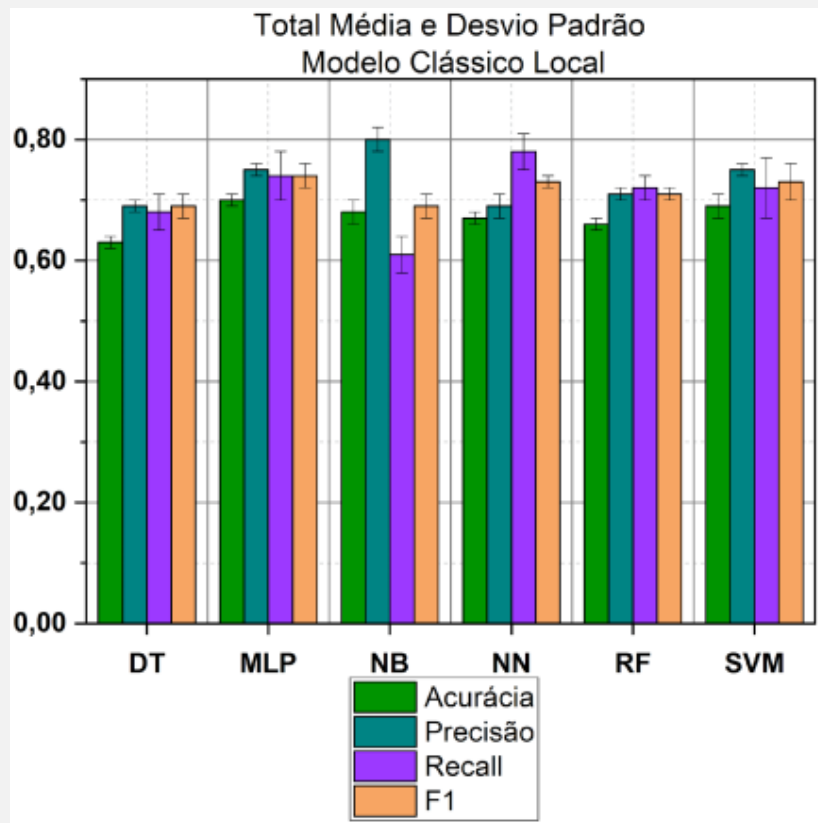
SEMANA INTEGRADA  
ENSINO | PESQUISA | EXTENSÃO

EICTI

XI ENCONTRO ANUAL DE INICIAÇÃO  
CIENTÍFICA E VII ENCONTRO DE INICIAÇÃO EM  
DESENVOLVIMENTO; TECNOLÓGICO E INOVAÇÃO



# Resultados



Fonte: Autoria Própria (2022).



4ª SIEPE

SEMANA INTEGRADA  
ENSINO | PESQUISA | EXTENSÃO

EICTI

XI ENCONTRO ANUAL DE INICIAÇÃO  
CIENTÍFICA E VII ENCONTRO DE INICIAÇÃO EM  
DESENVOLVIMENTO; TECNOLÓGICO E INOVAÇÃO



# Conclusão

Foi possível observar que os dois melhores resultados para acurácia foram *Neural Networks* com 77,65% e *SVM* 74,06% para os dados globais, no caso dos dados locais foram *Neural Networks* com 70,32% e *SVM* 68,85%.

Em trabalhos futuros pretendemos utilizar uma base de dados com uma menor divergência na quantidade de itens e na elaboração de novos algoritmos que se baseiam em redes neurais e *Deep Learning*.



4<sup>a</sup> SIEPE

SEMANA INTEGRADA  
ENSINO | PESQUISA | EXTENSÃO

EICTI

XI ENCONTRO ANUAL DE INICIAÇÃO  
CIENTÍFICA E VII ENCONTRO DE INICIAÇÃO EM  
DESENVOLVIMENTO; TECNOLÓGICO E INOVAÇÃO



# Agradecimentos

Agradeço a UNILA por abrir as portas da universidade, ao professor Willian Zalewski pela orientação neste trabalho, e aos professores da engenharia física e aos meus ex-colegas Patricia e Victor pela ajuda e a minha companheira Renata Benedet pelo apoio.



4<sup>a</sup> SIEPE

SEMANA INTEGRADA  
ENSINO | PESQUISA | EXTENSÃO

EICTI

XI ENCONTRO ANUAL DE INICIAÇÃO  
CIENTÍFICA E VII ENCONTRO DE INICIAÇÃO EM  
DESENVOLVIMENTO; TECNOLÓGICO E INOVAÇÃO



# Referências

1. EXOPLANETS NASA. **Discovery Fast Facts**. 2020. Disponível em: [https://exoplanets.nasa.gov/discovery/missions/#otp\\_fast\\_facts](https://exoplanets.nasa.gov/discovery/missions/#otp_fast_facts). Acesso em: 16 outubro 2022.
2. FACELI, K. et al. **Inteligência artificial: uma abordagem de aprendizado de máquina**. [s.l.: s.n.].
3. MONTANGER, P. O.; ZALEWSKI, W. Classificação automática de objetos astronômicos por meio da análise de séries temporais, 2019.
4. MONTANGER, P. O.; ZALEWSKI, W. Programa computacional para a identificação automática de exoplanetas, 2020.
5. SHALLUE, C. J.; VANDERBURG, A. Identifying exoplanets with deep learning: A five planet resonant chain around kepler-80 and an eighth planet around kepler-90arXiv, 2017.
6. FERRERO, C. A. **Algoritmo kNN para previsão de dados temporais: funções de previsão e critérios de seleção de vizinhos próximos aplicados a variáveis ambientais em limnologia**. Dissertação (Mestrado) — Universidade de São Paulo - USP, 2009.
7. Notebook criados em Python - IC 2021 IA para Exoplanetas. Disponível em: <https://github.com/brunohdmacedo/IC-2021-IA-para-Exoplanetas>.



4ª SIEPE

SEMANA INTEGRADA  
ENSINO | PESQUISA | EXTENSÃO

EICTI

XI ENCONTRO ANUAL DE INICIAÇÃO  
CIENTÍFICA E VII ENCONTRO DE INICIAÇÃO EM  
DESENVOLVIMENTO; TECNOLÓGICO E INOVAÇÃO

