

# A machine-learning approach for a CRIS research outputs' SDG classifications

euroCRIS Strategic Membership  
Meeting 2023, Pamplona, Spain

**António Luís Lopes** IT-IUL (Iscte)  
**Catarina Roseta-Palma** BRU-IUL (Iscte)  
**Ana Simaens** BRU-IUL (Iscte)





01

About Iscte and our  
CRIS

02

SDGs in the CRIS

03

Applying Machine  
Learning

04

Evaluation

05

Conclusions



# About Iscte and our CRIS



# 01



# About Iscte

University Institute of  
Lisbon

The Iscte - University Institute of Lisbon was founded in 1972 as one of Portugal's first modern universities, with the central aim of studying labour and social dynamics in a rapidly changing world.

Since its establishment, the university has expanded its disciplinary reach into five schools in the following areas: Business, Sociology and Public Policy, Technology and Architecture, Social Sciences and Humanities, and Applied Digital Technologies.

5

Schools

8

Research  
Centres

100+

Degrees

11k+

Students

57%

MSc, PhD  
Students

22%

International  
Students



## About our CRIS

### Ciência-IUL

The work on Iscte's CRIS started in 2010 as part of the development of an evaluation system for the University. In 2013 it spun off as its own individual system separated from the evaluation system.

From then on, it evolved to gather all kinds of relevant scientific and research information and to focus on internal and external integrations to avoid any duplicate efforts in collecting or providing data to the stakeholders.

**2000+**

**Researchers**

**1600+**

**Projects**

**54k+**

**Publications**

**300k+**

**Citations**

**3**

**Internal  
Integrations**

**8**

**External  
Integrations**

02



**SDGs in the  
CRIS**



# Sustainable Development Goals

Iscte's Publications and Projects that contribute to the SDGs



Nr. publications: 245  
Nr. Projects: 16



Nr. publications: 60  
Nr. Projects: 4



Nr. publications: 1126  
Nr. Projects: 125



Nr. publications: 1859  
Nr. Projects: 317



Nr. publications: 585  
Nr. Projects: 61



Nr. publications: 55  
Nr. Projects: 6



Nr. publications: 168  
Nr. Projects: 18



Nr. publications: 1300  
Nr. Projects: 138



Nr. publications: 1835  
Nr. Projects: 74



Nr. publications: 1779  
Nr. Projects: 238



Nr. publications: 1468  
Nr. Projects: 94



Nr. publications: 691  
Nr. Projects: 16



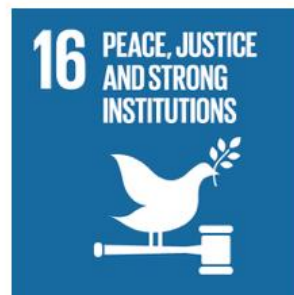
Nr. publications: 435  
Nr. Projects: 26



Nr. publications: 64  
Nr. Projects: 5



Nr. publications: 195  
Nr. Projects: 13



Nr. publications: 1103  
Nr. Projects: 113



Nr. publications: 291  
Nr. Projects: 17

# Sustainable Development Goals

Iscte's Publications and Projects that contribute to the SDGs



## Quality education

Ensure inclusive and equitable quality education and promote lifelong learning opportunities for all

[Know More](#)

Apply filter

Iscte - Instituto Universitário de Lisboa

Filter

### Publications (1859)

Type	Title	Quartile
Scientific journal paper	<a href="#">Tackling regional skill shortages: From single employer strategies to local partnerships</a>	Q1
Scientific journal paper	<a href="#">Lisbon, the Portuguese Erasmus city? Mis-match between representation in urban policies and international student experiences</a>	Q1
Scientific journal paper	<a href="#">Confidence intervals for means and variances of nonnormal distributions</a>	Q3
Scientific journal paper	<a href="#">Memories of (Un)Freire literacy policies in Southern Africa from the 1970s on: telling the (hi)story through life histories and photography of (dis)empowerment in Mozambique</a>	--
Scientific journal paper	<a href="#">Internet of things and consumer engagement on retail: State-of-the-art and future directions</a>	Q1
Scientific journal paper	<a href="#">The impacts of animal farming: A critical review of secondary and high school textbooks</a>	Q1

### Projects (317)

Title
<a href="#">Women architects in former Portuguese colonial Africa: gender and struggle for professional recognition (1953-1985)</a>
<a href="#">Cooperation Project between the Youth Employment Observatory and the Portuguese Public Employment Services (IEFP)</a>
<a href="#">Concepção e Implementação do Laboratório de Competências Transversais na Universidade Amílcar Cabral a partir do caso Iscte</a>
<a href="#">Master's Degree of Managing Digital Transformation in the Health Sector</a>
<a href="#">Profissionalização artística e formação superior jazzística: a inserção profissional de jovens diplomados em Portugal</a>
<a href="#">Trajetórias biográficas, percursos profissionais e inscrição urbana de estudantes internacionais brasileiros em Lisboa</a>
<a href="#">Flipping Learning Internationally in a Post Pandemic Era</a>
<a href="#">11 TEIP - 2º Programa Territorialização de Políticas Educativas de Intervenção Prioritária</a>





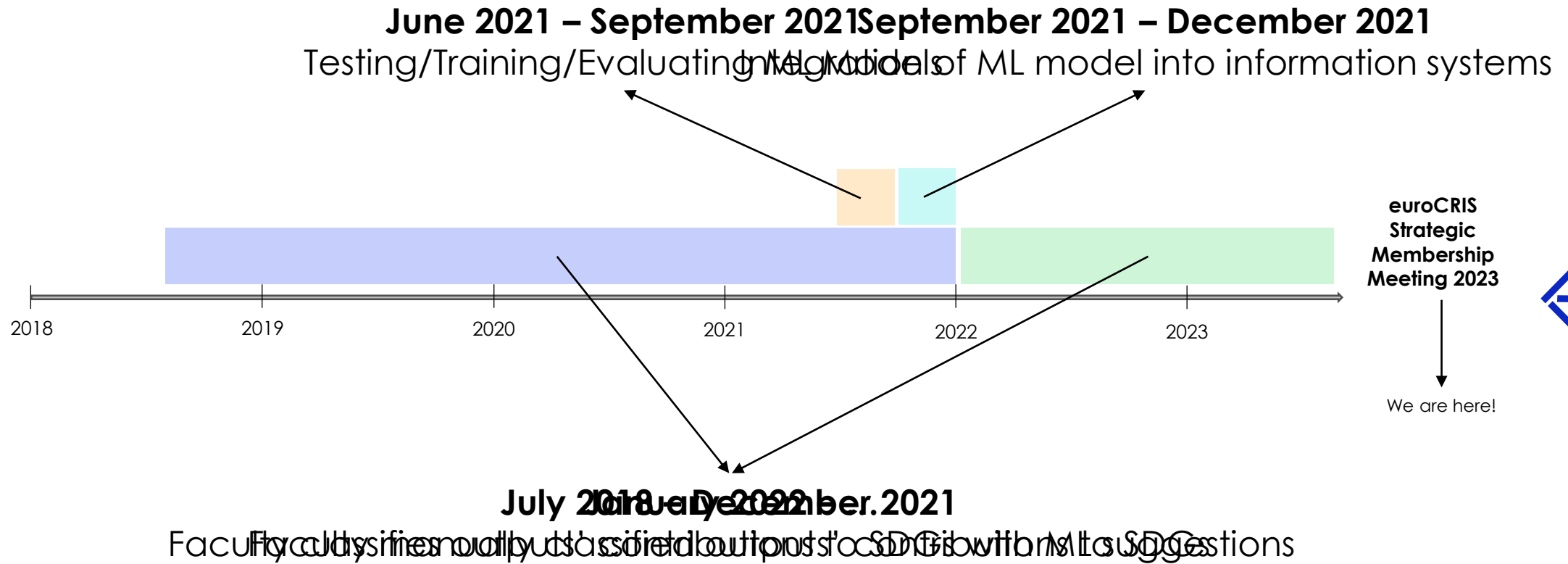
# Applying Machine Learning



# 03

# SDG Classification at Iscte

Timeline of our approach



# SDG Classification at Iscte

Suggestions in our CRIS (Current Research Information System)

Type	Publication
Title	The Sustainable Development Goals in higher education: a machine-learning approach


Considering the text content of this record, it is suggested that you choose this Sustainable Development Goal.

1 NO POVERTY



✘ Unselected

2 ZERO HUNGER



✘ Unselected

3 GOOD HEALTH AND WELL BEING



✘ Unselected

**Suggested**

4 QUALITY EDUCATION



✘ Unselected

5 GENDER EQUALITY



✘ Unselected

7 AFFORDABLE AND CLEAN ENERGY



8 DECENT WORK AND ECONOMIC GROWTH



9 INDUSTRY, INNOVATION AND INFRASTRUCTURE



10 REDUCED INEQUALITIES

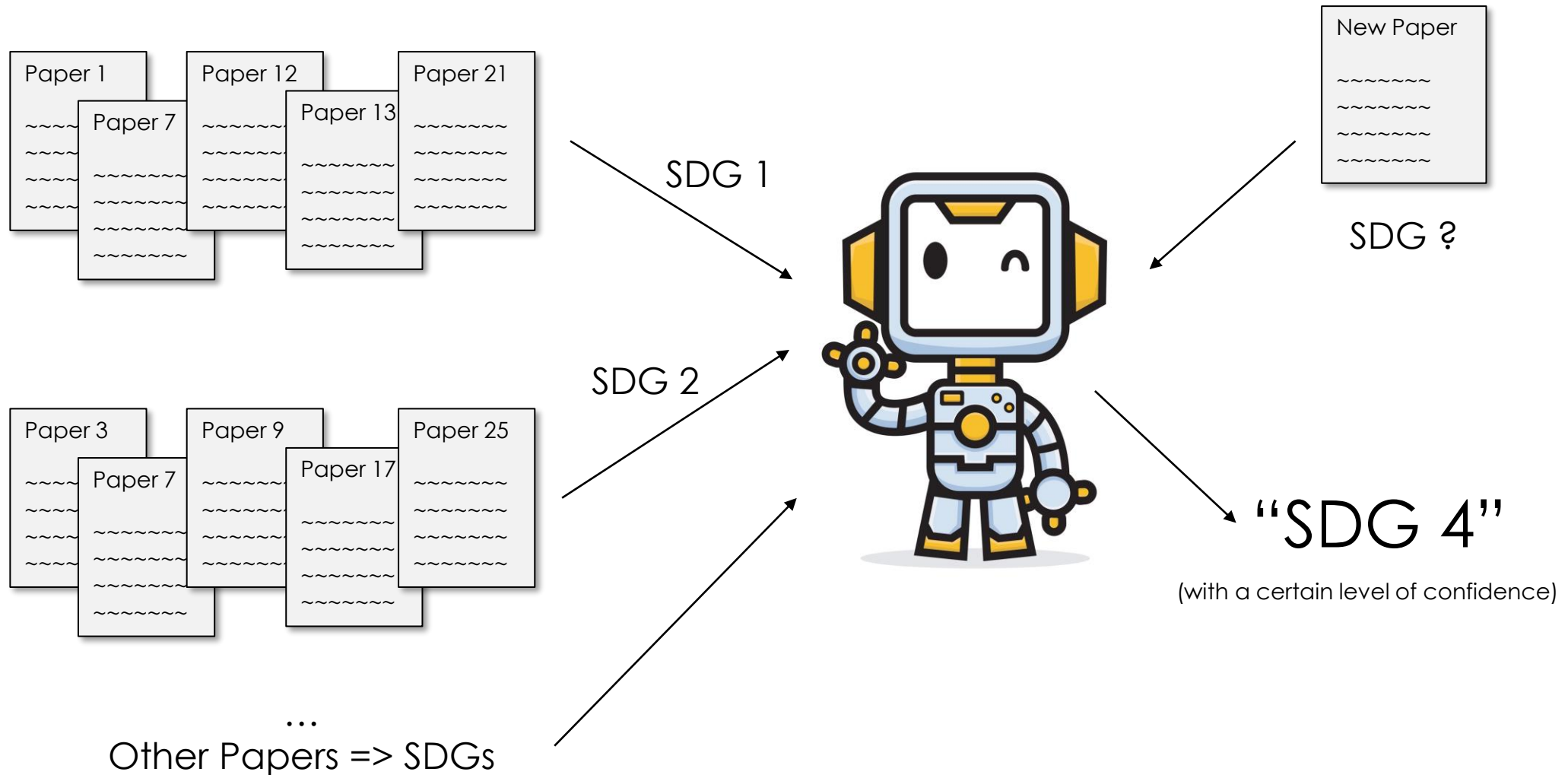


11 SUSTAINABLE CITIES AND COMMUNITIES



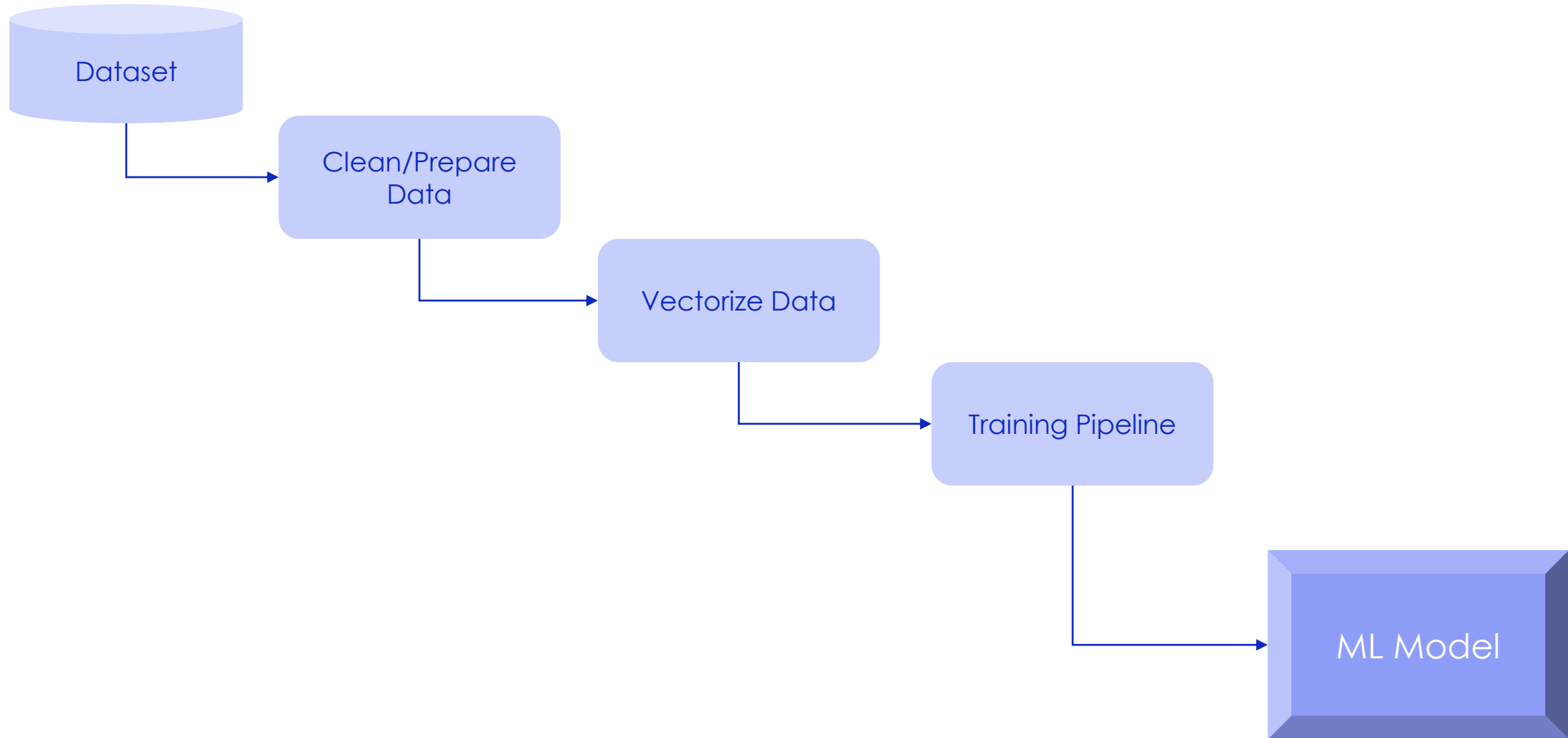
# Machine Learning Model for SDG Classification

Or how to teach a robot choose which SDG an output contributes to



# Machine Learning Model for SDG Classification

The actual machine learning training stage



04

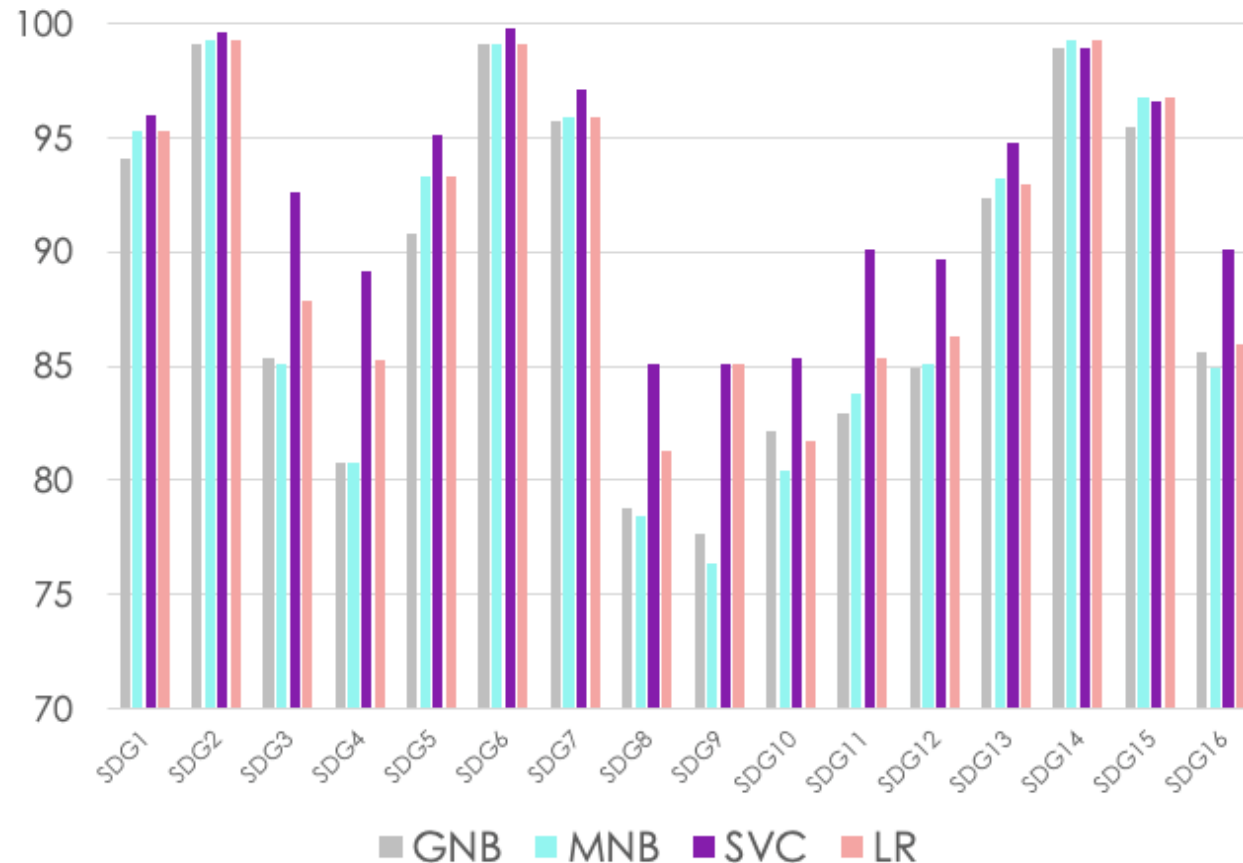


**Evaluation**



# Evaluation of Machine Learning Algorithms

## Algorithms' Accuracy per SDG



We tested the following algorithms with a dataset of 9665 annotated records:

- Gaussian Naive Bayes (GNB)
- Multinomial Naive Bayes (MNB)
- Linear Support Vector Classification (SVC)
- Logistic Regression (LR)

We used the following metrics:

- ACCURACY – ratio of correctly classified samples
- RECALL - ability to find all the positive samples
- PRECISION – ability to identify all the positive samples without accidentally marking too many negative samples as positive
- F-MEASURE - harmonic mean of the Precision and Recall

# Comparison between Users Choices and Model Predictions

How much did the robot get right using the SVC-based Model?

## → Methodology

For each output, the user chose up to 3 SDGs.

Our Machine Learning model made the prediction for each combination of output/SDG, providing the corresponding score. The scores for each SDG were then sorted and only the top 3 were considered.

We then compared the two sets to see how much they matched.

Users' Choices ↔ Model's Choices

[3, 4, 9]

100%

[3, 4, 9]

[1, 2, 13]

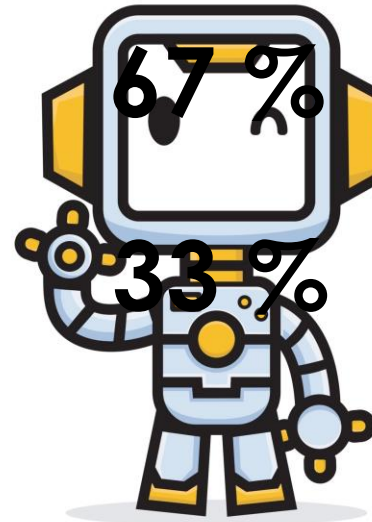
67%

[1, 5, 13]

[3, 7, 11]

33%

[1, 7, 17]





## Conclusions



05

# Conclusions

Machine Learning Model is successful in suggesting SDGs that should be associated with outputs.

This work is not specific to SDGs. The same approach can be applied to other classification efforts.

We “automated” our users’ choices. That doesn’t mean they’re right. Expert review is needed.





# iscte

INSTITUTO  
UNIVERSITÁRIO  
DE LISBOA



[ciencia.iscte-iul.pt/sustainable-development-goals](https://ciencia.iscte-iul.pt/sustainable-development-goals)

**António Luís Lopes** (alsl@iscte-iul.pt)

**Catarina Roseta-Palma** (catarina.roseta@iscte-iul.pt)

**Ana Simaens** (ana.simaens@iscte-iul.pt)

eiic  
eie

iscte

INSTITUTO  
UNIVERSITÁRIO  
DE LISBOA

st