



Progetto FIA **MEDI-AI**

By Andrea Gisolfi, Antonio Merola,
Giovanni Nigro, Giacomo Favale

VI SIETE MAI RITROVATI IN QUESTA SITUAZIONE?

The image shows a Google search results page. At the top left is the Google logo. The search bar contains the query "ho la febbre quali potrebbero essere le conseguenze?". To the right of the search bar are icons for clearing the search, voice search, and a magnifying glass. Below the search bar is a row of category buttons: "Siti di luoghi", "Immagini", "Video", "Notizie", "Libri", "Maps", "Voli", and "Finanza". A status message indicates "Circa 667.000 risultati (0,23 secondi)". The main content area displays a snippet of text from a search result, followed by a large red text overlay.

ho la febbre quali potrebbero essere le conseguenze?

Siti di luoghi Immagini Video Notizie Libri Maps Voli Finanza

Circa 667.000 risultati (0,23 secondi)

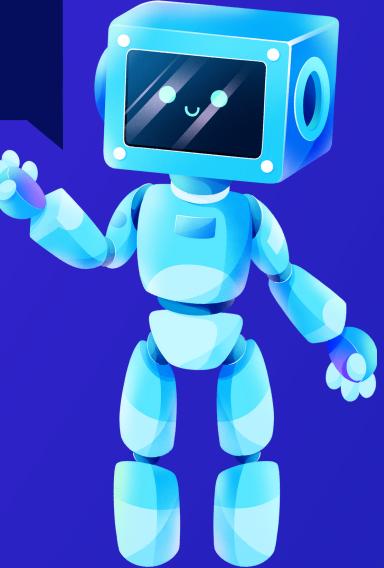
Generalmente la febbre di per sé non è pericolosa, tranne che a temperature che superano 40 gradi. Particolare attenzione va riservata ai bambini e neonati, nei quali temperature troppo alte possono determinare conseguenze estremamente gravi, fino alla

morte.

OBIETTIVI

L'obiettivo del nostro progetto è la creazione di un chatbot che sia in grado di predire la possibile malattia dell'utente, basandosi sui sintomi forniti.

Ciao!





SPECIFICA P.E.A.S





SPECIFICA P.E.A.S

Performance





SPECIFICA P.E.A.S

Performance

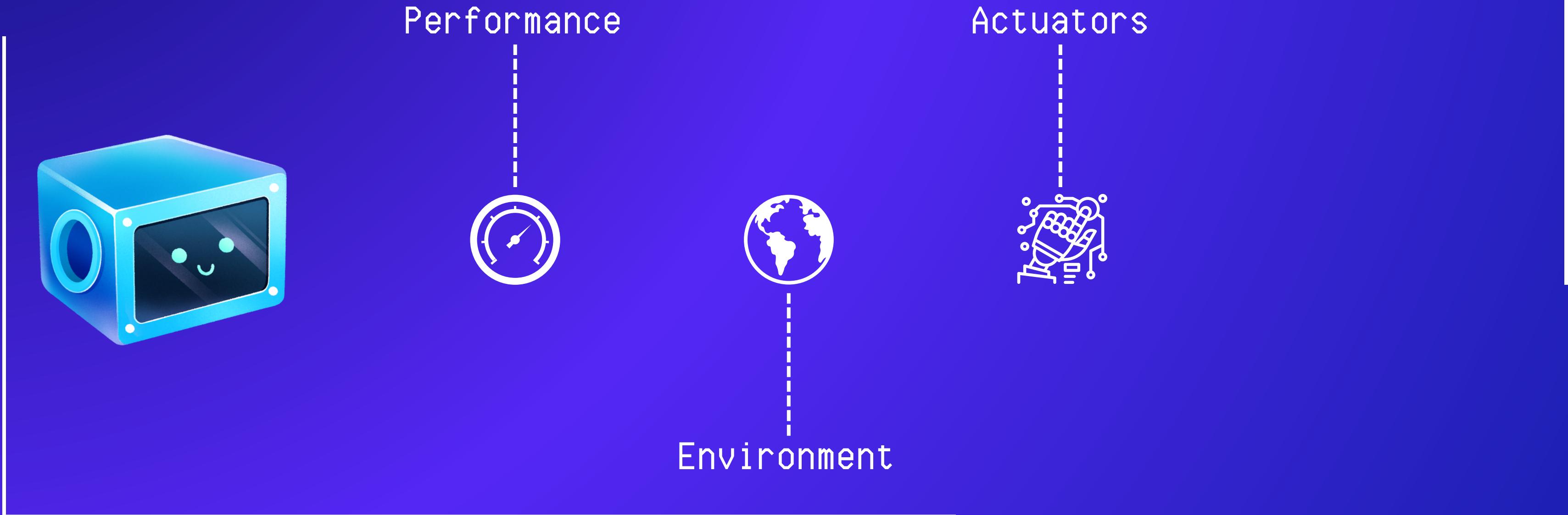


Environment



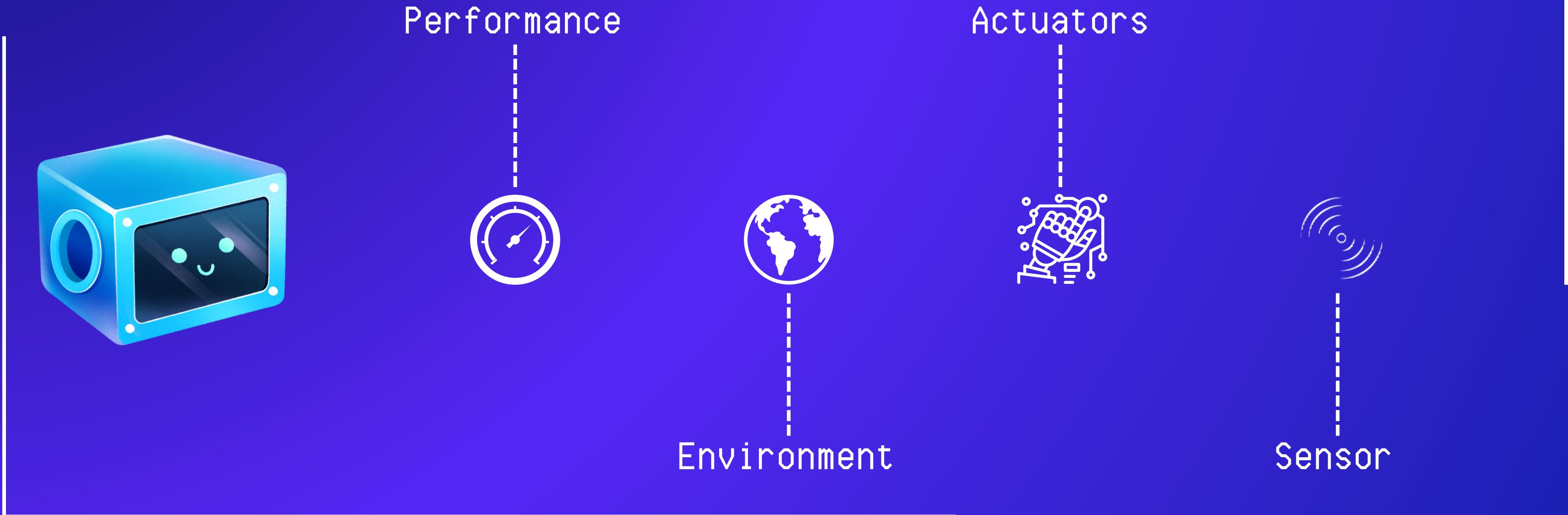


SPECIFICA P.E.A.S





SPECIFICA P.E.A.S



SPECIFICHE DELL'AMBIENTE

SPECIFICHE DELL'AMBIENTE



Parzialmente osservabile

SPECIFICHE DELL'AMBIENTE



SPECIFICHE DELL'AMBIENTE



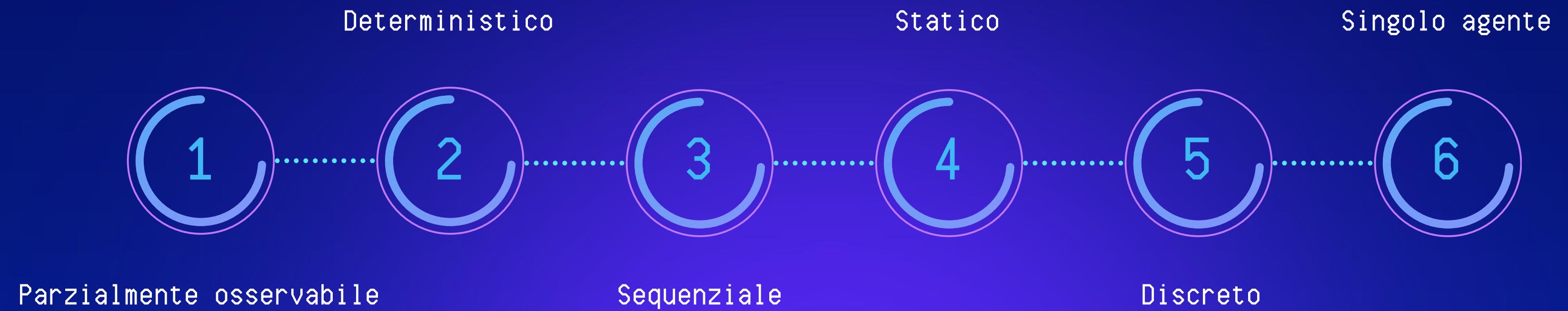
SPECIFICHE DELL'AMBIENTE



SPECIFICHE DELL'AMBIENTE



SPECIFICHE DELL'AMBIENTE



Data Understanding

Data Preparation



Scelta del dataset

Disease Symptom Prediction

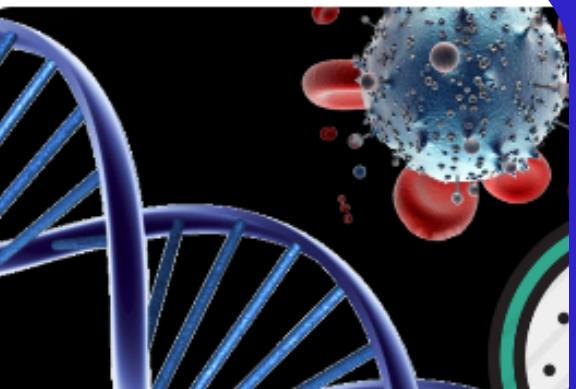
helps to create a disease prediction or healthcare system

Data Card Code (43) Discussion (4)

About Dataset

Context

A dataset to provide the students a source to create a healthcare related system.



Usability ⓘ
9.12

License
[CC BY-SA 4.0](#)

Molti dati mancanti, ma ...

```
RangeIndex: 4920 entries, 0 to 4919
Data columns (total 18 columns):
 #   Column      Non-Null Count Dtype  
--- 
 0   Malattia    4920 non-null   object 
 1   Sintomo 1   4920 non-null   object 
 2   Sintomo 2   4920 non-null   object 
 3   Sintomo 3   4920 non-null   object 
 4   Sintomo 4   4572 non-null   object 
 5   Sintomo 5   3714 non-null   object 
 6   Sintomo 6   2934 non-null   object 
 7   Sintomo 7   2268 non-null   object 
 8   Sintomo 8   1944 non-null   object 
 9   Sintomo 9   1692 non-null   object 
 10  Sintomo 10  1512 non-null   object 
 11  Sintomo 11  1194 non-null   object 
 12  Sintomo 12  744 non-null   object 
 13  Sintomo 13  504 non-null   object 
 14  Sintomo 14  306 non-null   object 
 15  Sintomo 15  240 non-null   object 
 16  Sintomo 16  192 non-null   object 
 17  Sintomo 17  72 non-null   object
```

Molti dati mancanti, ma ...

```
RangeIndex: 4920 entries, 0 to 4919
Data columns (total 18 columns):
 #   Column      Non-Null Count Dtype  
--- 
 0   Malattia    4920 non-null   object 
 1   Sintomo 1   4920 non-null   object 
 2   Sintomo 2   4920 non-null   object 
 3   Sintomo 3   4920 non-null   object 
 4   Sintomo 4   4572 non-null   object 
 5   Sintomo 5   3714 non-null   object 
 6   Sintomo 6   2934 non-null   object 
 7   Sintomo 7   2268 non-null   object 
 8   Sintomo 8   1944 non-null   object 
 9   Sintomo 9   1692 non-null   object 
 10  Sintomo 10  1512 non-null   object 
 11  Sintomo 11  1194 non-null   object 
 12  Sintomo 12  744 non-null   object 
 13  Sintomo 13  504 non-null   object 
 14  Sintomo 14  306 non-null   object 
 15  Sintomo 15  240 non-null   object 
 16  Sintomo 16  192 non-null   object 
 17  Sintomo 17  72 non-null   object
```

... there's nothing we can do



Tanti dati duplicati?

| index | itching | skin_rash | nodal_skin_eruptions | continuous_sneezing | shivering |
|-------|---------|-----------|----------------------|---------------------|-----------|
| 4919 | 0 | 1 | 0 | 0 | 0 |
| 4918 | 0 | 1 | 0 | 0 | 0 |
| 4917 | 0 | 0 | 0 | 0 | 0 |
| 4916 | 0 | 1 | 0 | 0 | 0 |
| 4915 | 0 | 0 | 0 | 0 | 0 |
| 4914 | 0 | 0 | 0 | 0 | 0 |
| 4913 | 0 | 0 | 0 | 0 | 0 |
| 4912 | 0 | 0 | 0 | 0 | 0 |
| 4911 | 0 | 0 | 0 | 0 | 0 |
| 4910 | 0 | 0 | 0 | 0 | 0 |

Tanti dati duplicati?



| index | itching | skin_rash | nodal_skin_eruptions | continuous_sneezing | shivering |
|-------|---------|-----------|----------------------|---------------------|-----------|
| 4919 | 0 | 1 | | 0 | 0 |
| 4918 | 0 | 1 | | 0 | 0 |
| 4917 | 0 | 0 | | 0 | 0 |
| 4916 | 0 | 1 | | 0 | 0 |
| 4915 | 0 | 0 | | 0 | 0 |
| 4914 | 0 | 0 | | 0 | 0 |
| 4913 | 0 | 0 | | 0 | 0 |
| 4912 | 0 | 0 | | 0 | 0 |
| 4911 | 0 | 0 | | 0 | 0 |
| 4910 | 0 | 0 | | 0 | 0 |

| index | itching | skin_rash | nodal_skin_eruptions | continuous_sneezing | shivering |
|-------|---------|-----------|----------------------|---------------------|-----------|
| 303 | 0 | 1 | 0 | 0 | 0 |
| 302 | 0 | 1 | 0 | 0 | 0 |
| 301 | 0 | 1 | 0 | 0 | 0 |
| 300 | 0 | 1 | 0 | 0 | 0 |
| 299 | 0 | 0 | 0 | 0 | 0 |
| 298 | 0 | 1 | 0 | 0 | 0 |
| 297 | 0 | 1 | 0 | 0 | 0 |
| 296 | 0 | 1 | 0 | 0 | 0 |
| 295 | 0 | 1 | 0 | 0 | 0 |
| 294 | 0 | 1 | 0 | 0 | 0 |
| 293 | 0 | 1 | 0 | 0 | 0 |

...NON PIÙ



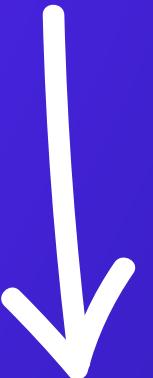
Data Modeling



MACHINE LEARNING



Apprendimento supervisionato



Classificazione

SCELTA DELL'ALGORITMO



Decision Tree

Basato sulla costruzione
di un albero



Naive Bayes

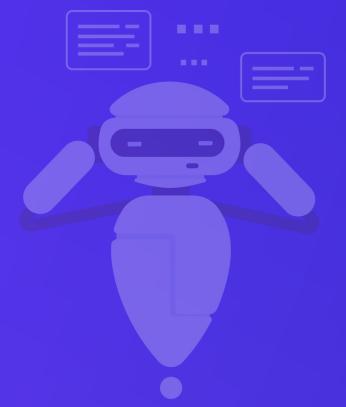
Basato sull'utilizzo del
Teorema di Bayes

SCELTA DELL'ALGORITMO



Decision Tree

Basato sulla costruzione
di un albero

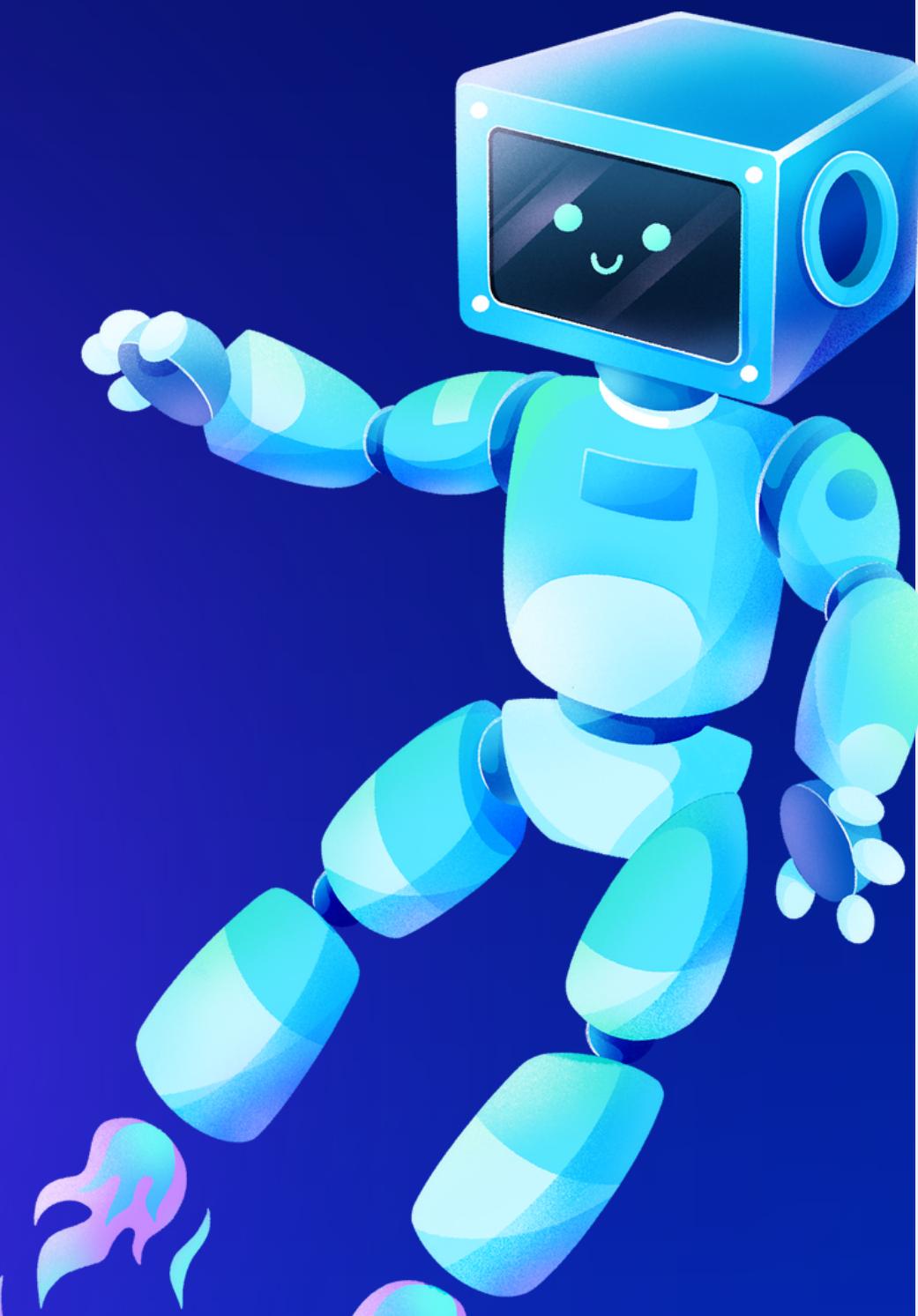


Naive Bayes

Basato sull'utilizzo del
Teorema di Bayes

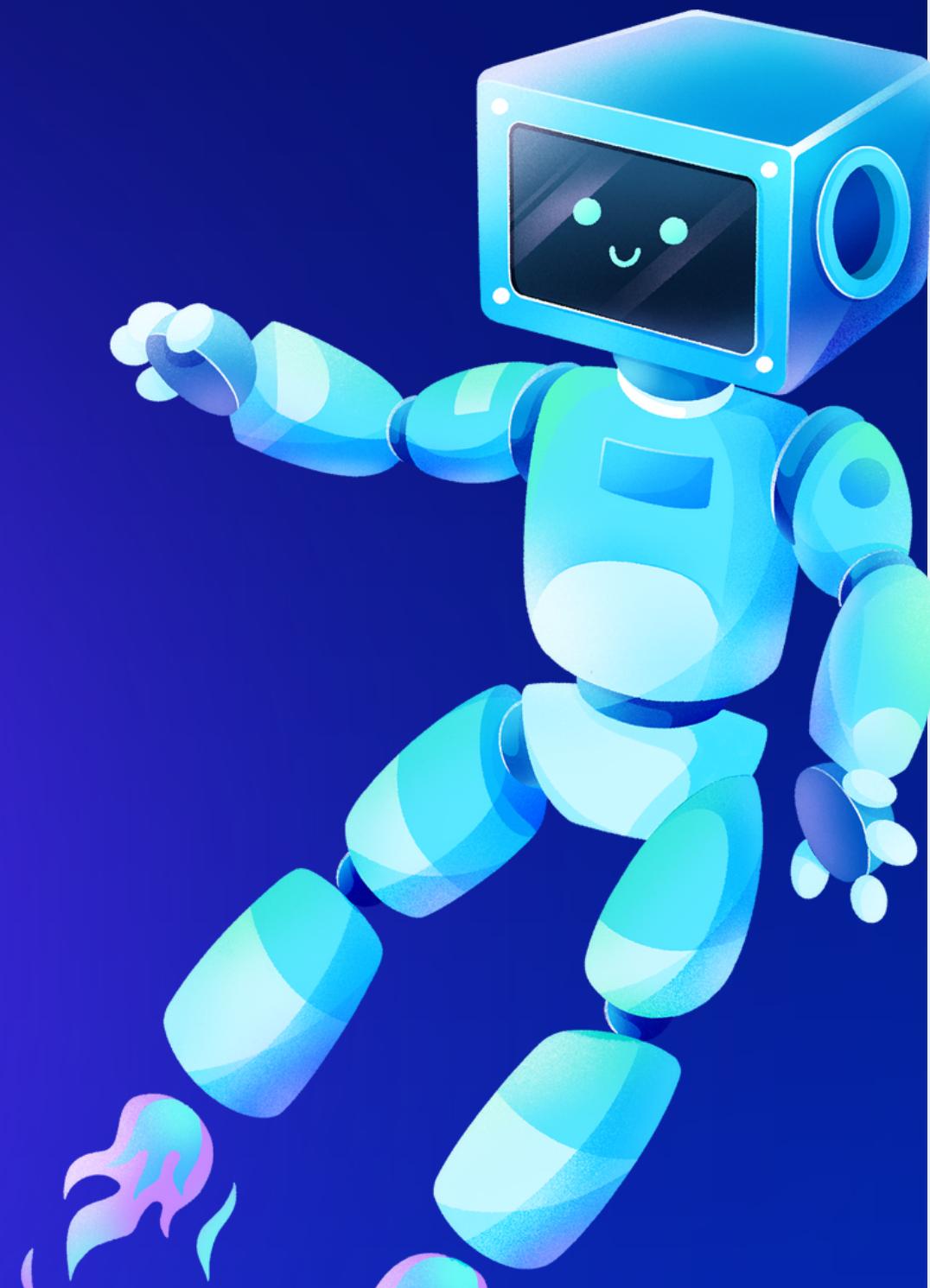
Evaluation

K-FOLD CROSS VALIDATION



K-FOLD CROSS VALIDATION

5



K-FOLD CROSS VALIDATION

5

Accuracy:

0.7128

Precision:

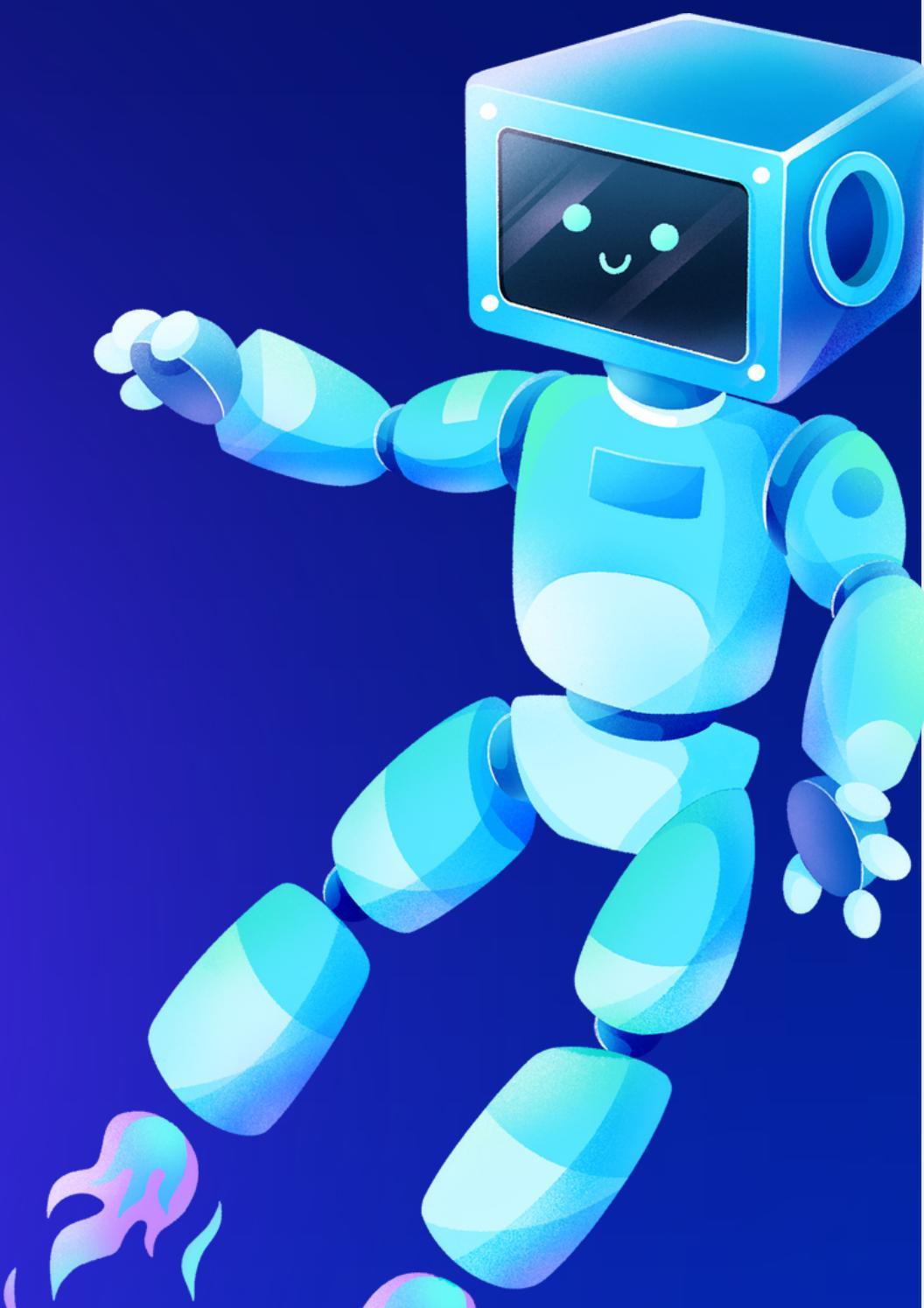
0.7085

Recall:

0.6288

F1-score:

0.6411



THANK YOU!

