



Aula: Manipulando datasets

Maurício Balboni, Karina Machado

balboni@furg.br

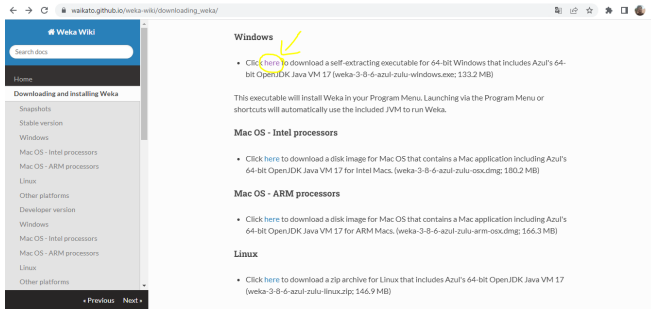
Universidade Federal do Rio Grande - FURG

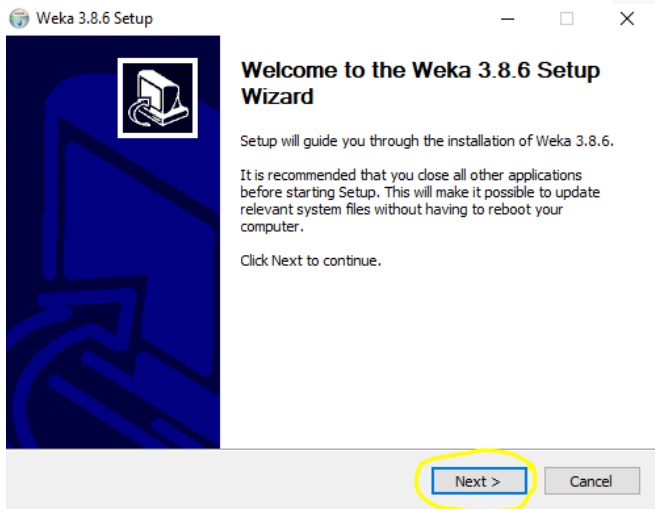
- 1 WEKA
 - Instalação no Windows
 - Instalação no Ubuntu
- 2 Introdução aos Datasets
 - Repositórios de datasets
 - Arquivos .arff
 - Repositorio de datasets .arff
- 3 Manipulação de dataset
- 4 Conclusão

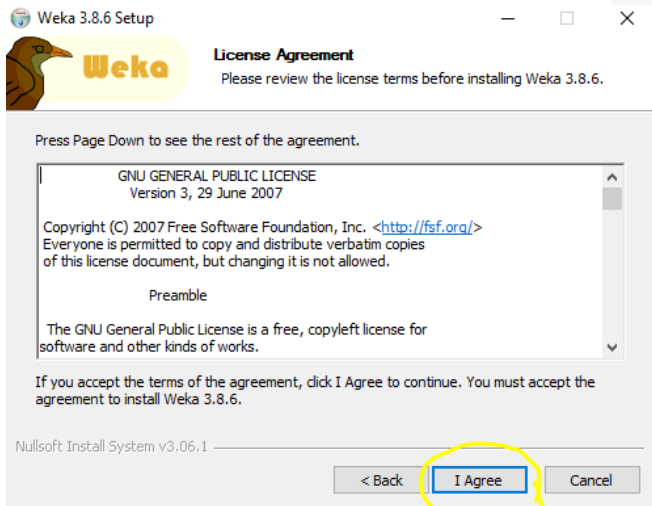


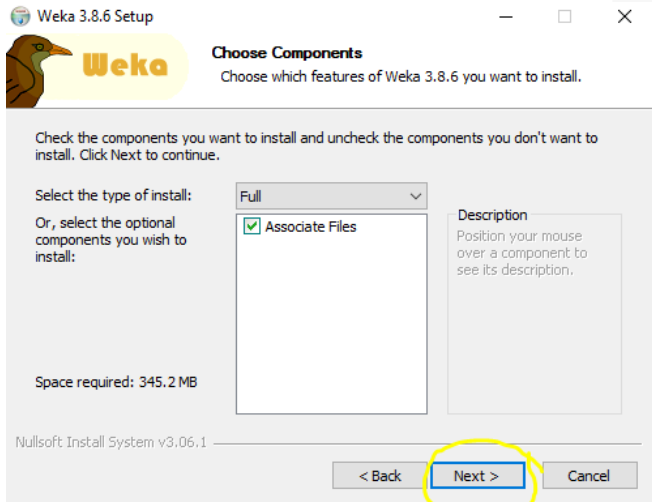
https://waikato.github.io/weka-wiki/downloading_weka/

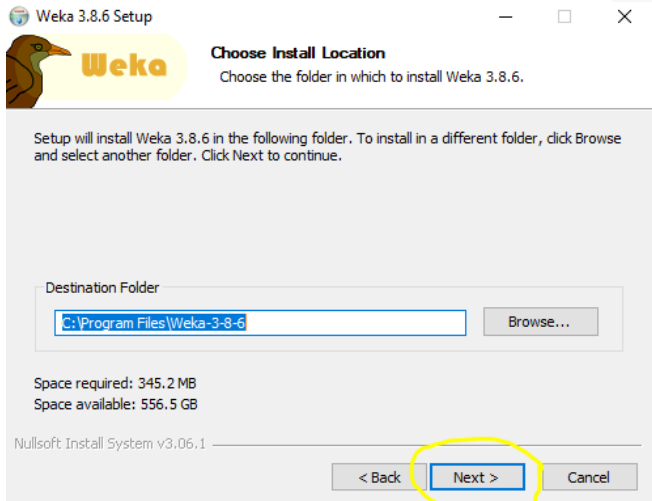
- 1 WEKA
Instalação no Windows
Instalação no Ubuntu
- 2 Introdução aos Datasets
Repositórios de datasets
Arquivos .arff
Repositorio de datasets .arff
- 3 Manipulação de dataset
- 4 Conclusão

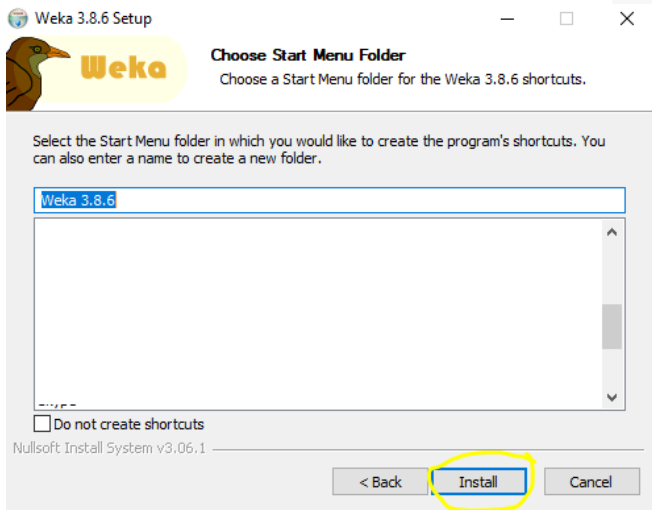


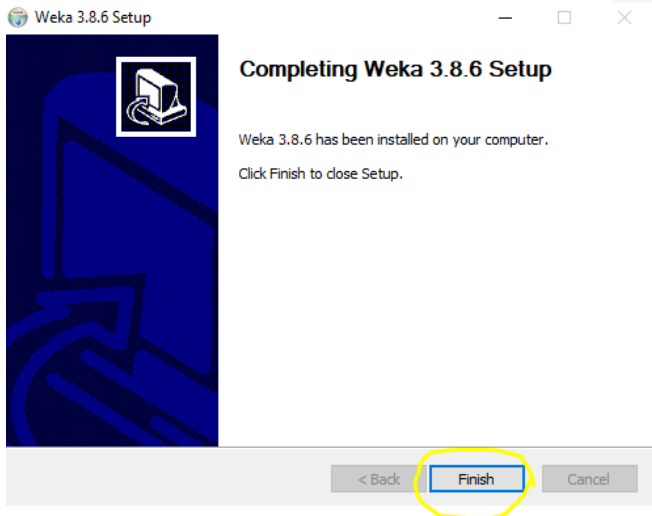












Interface no Windows



1 WEKA

Instalação no Windows

Instalação no Ubuntu

2 Introdução aos Datasets

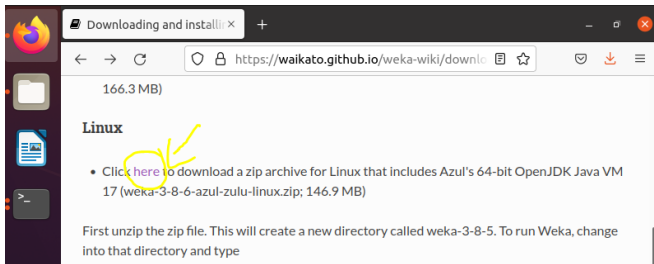
Repositórios de datasets

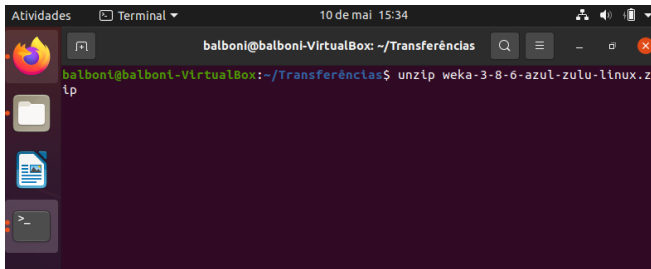
Arquivos .arff

Repositorio de datasets .arff

3 Manipulação de dataset

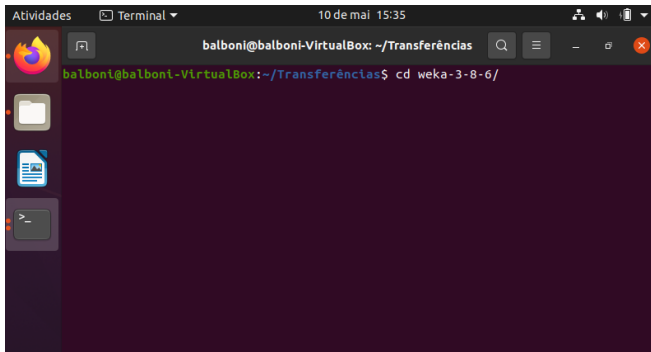
4 Conclusão





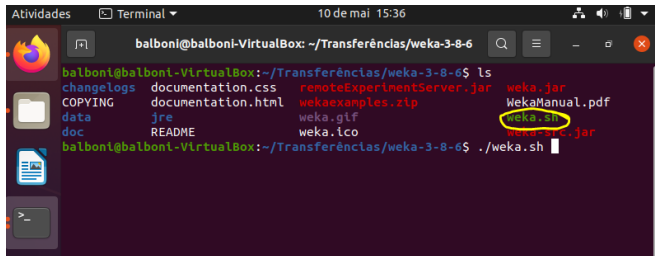
The screenshot shows a terminal window titled "Terminal" with the date and time "10 de mai 15:34". The terminal prompt is "balboni@balboni-VirtualBox: ~/Transferências". The command "unzip weka-3-8-6-azul-zulu-linux.zip" has been entered, and the output "ip" is visible on the next line. The terminal window is part of the Ubuntu desktop environment, with the "Atividades" (Activities) overview visible on the left side.

```
balboni@balboni-VirtualBox: ~/Transferências
balboni@balboni-VirtualBox:~/Transferências$ unzip weka-3-8-6-azul-zulu-linux.z
ip
```

The screenshot shows a terminal window titled "Terminal" with the date and time "10 de mai 15:35". The window's title bar also includes "Atividades" and "balboni@balboni-VirtualBox: ~/Transferências". The terminal prompt is "balboni@balboni-VirtualBox:~/Transferências\$". The command "cd weka-3-8-6/" has been entered and is highlighted in blue. The terminal window has a dark purple background and a sidebar on the left with icons for Firefox, Files, and the terminal itself.

```
balboni@balboni-VirtualBox: ~/Transferências
balboni@balboni-VirtualBox:~/Transferências$ cd weka-3-8-6/
```



```
balboni@balboni-VirtualBox: ~/Transferências/weka-3-8-6
balboni@balboni-VirtualBox:~/Transferências/weka-3-8-6$ ls
changelogs  documentation.css  remoteExperimentServer.jar  weka.jar
COPYING     documentation.html  wekaexamples.zip           WekaManual.pdf
data        jre                 weka.gif                   weka.sh
doc         README             weka.ico                   weka-vc.jar
balboni@balboni-VirtualBox:~/Transferências/weka-3-8-6$ ./weka.sh
```

Interface no Ubuntu



Dataset Iris

	A	B	C	D	E	F
1	Id	SepalLengthCm	SepalWidthCm	PetalLengthCm	PetalWidthCm	Species
2	1	5.1	3.5	1.4	0.2	Iris-setosa
3	2	4.9	3.0	1.4	0.2	Iris-setosa
4	3	4.7	3.2	1.3	0.2	Iris-setosa
5	4	4.6	3.1	1.5	0.2	Iris-setosa
6	5	5.0	3.6	1.4	0.2	Iris-setosa
7	6	5.4	3.9	1.7	0.4	Iris-setosa
8	7	4.6	3.4	1.4	0.3	Iris-setosa
9	8	5.0	3.4	1.5	0.2	Iris-setosa
10	9	4.4	2.9	1.4	0.2	Iris-setosa
11	10	4.9	3.1	1.5	0.1	Iris-setosa
12	11	5.4	3.7	1.5	0.2	Iris-setosa
13	12	4.8	3.4	1.6	0.2	Iris-setosa
14	13	4.8	3.0	1.4	0.1	Iris-setosa
15	14	4.3	3.0	1.1	0.1	Iris-setosa
16	15	5.8	4.0	1.2	0.2	Iris-setosa
17	16	5.7	4.4	1.5	0.4	Iris-setosa
18	17	5.4	3.9	1.3	0.4	Iris-setosa
19	18	5.1	3.5	1.4	0.3	Iris-setosa
20	19	5.7	3.8	1.7	0.3	Iris-setosa
21	20	5.1	3.8	1.5	0.3	Iris-setosa

- 1 WEKA
Instalação no Windows
Instalação no Ubuntu
- 2 **Introdução aos Datasets**
Repositórios de datasets
Arquivos .arff
Repositório de datasets .arff
- 3 Manipulação de dataset
- 4 Conclusão

kaggle

1 WEKA

Instalação no Windows

Instalação no Ubuntu

2 Introdução aos Datasets

Repositórios de datasets

Arquivos .arff

Repositorio de datasets .arff

3 Manipulação de dataset

4 Conclusão



iris - Notepad

File Edit Format View Help

@RELATION iris

@ATTRIBUTE sepallength REAL

@ATTRIBUTE sepalwidth REAL

@ATTRIBUTE petallength REAL

@ATTRIBUTE petalwidth REAL

@ATTRIBUTE class {Iris-setosa,Iris-versicolor,Iris-virginica}

@DATA

5.1,3.5,1.4,0.2,Iris-setosa

4.9,3.0,1.4,0.2,Iris-setosa

4.7,3.2,1.3,0.2,Iris-setosa

4.6,3.1,1.5,0.2,Iris-setosa

5.0,3.6,1.4,0.2,Iris-setosa

5.4,3.9,1.7,0.4,Iris-setosa

4.6,3.4,1.4,0.3,Iris-setosa

5.0,3.4,1.5,0.2,Iris-setosa

4.4,2.9,1.4,0.2,Iris-setosa

4.9,3.1,1.5,0.1,Iris-setosa

5.4,3.7,1.5,0.2,Iris-setosa

4.8,3.4,1.6,0.2,Iris-setosa

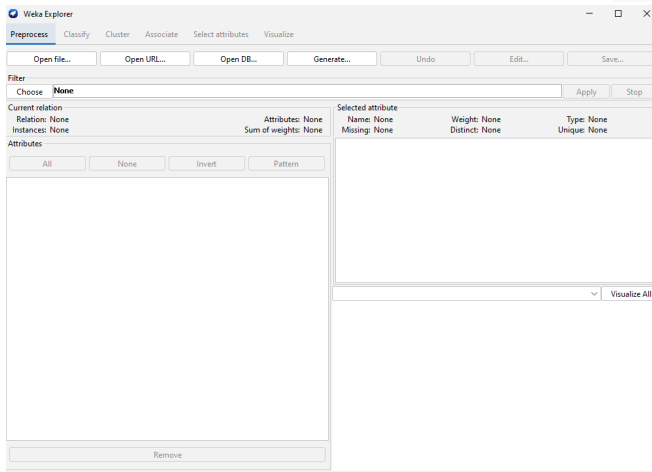
4.8,3.0,1.4,0.1,Iris-setosa

- 1 WEKA
Instalação no Windows
Instalação no Ubuntu
- 2 **Introdução aos Datasets**
Repositorios de datasets
Arquivos .arff
Repositorio de datasets .arff
- 3 Manipulação de dataset
- 4 Conclusão

<https://github.com/renatopp/arff-datasets>

<https://storm.cis.fordham.edu/~gweiss/data-mining/datasets.html>

<https://waikato.github.io/weka-wiki/datasets/>



	A	B	C	D	E	F	G	H	I
1	Idade	genero	febre	tosse	Pressão	falta de ar	positivo covid	vacinação	Emergencia
2		20 M	não	sim	normal	não	sim	2 dose	leve
3		35 F	alta	sim	normal	sim	sim	nenhuma	alta
4		80 M	baixa	sim	normal	moderada	não	2 dose	media
5		75 O	alta	sim	alta	sim	sim	nenhuma	urgencia
6		40 M	não	sim	normal	não	não	2 dose	leve
7		50 F	sim	sim	alta	sim	sim	nenhuma	urgencia
8		70 F	não	sim	normal	não	sim	3 dose	media
9		35 F	não	sim	normal	não	sim	3 dose	leve

contato: balboni@furg.br