



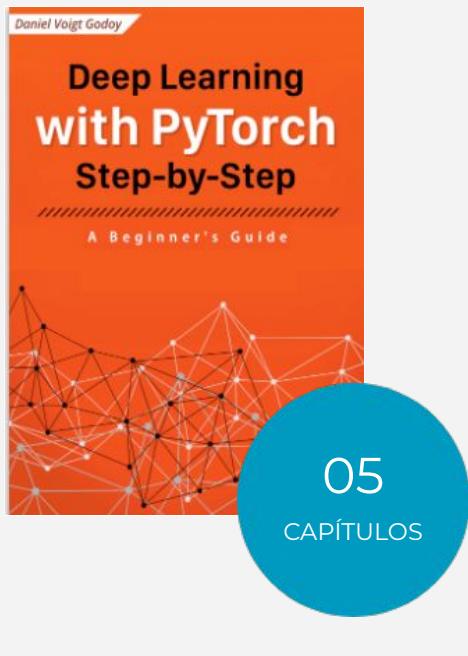
PPGEEC2318

# Projeto #01

Aprendizado de Máquina

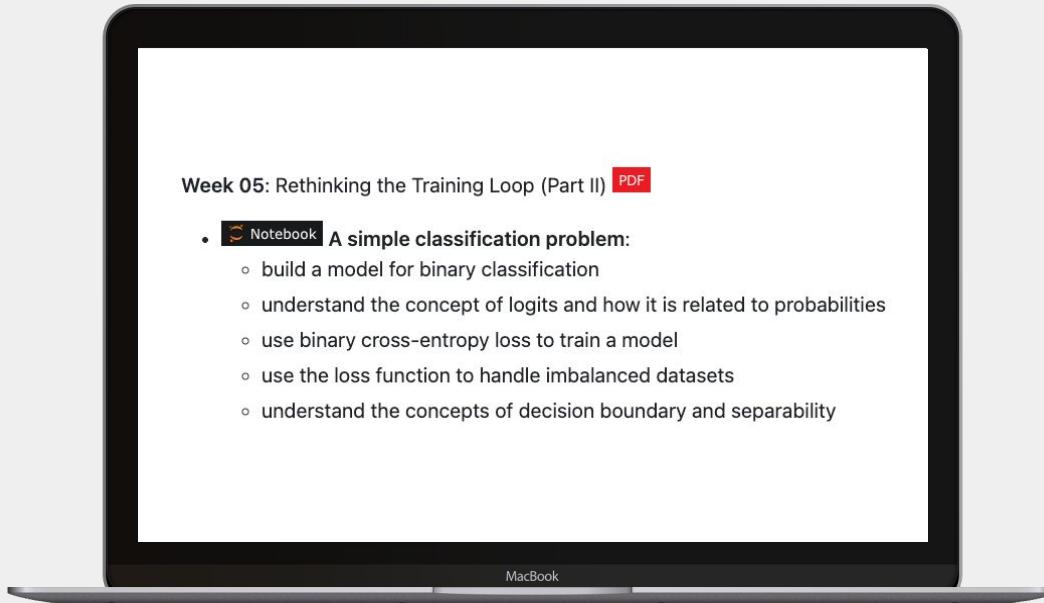
## Teoria e Prática

# O que já foi visto no curso?



1. Fundamentos de ML  
Conceitos de aprendizados, pipeline, separação entre treino/teste, problemas de regressão e classificação.
2. Regressão Linear  
Implementação manual com Pytorch, uso da MSELoss, otimização com tf.optim, normalização de dados e métricas como RMSE.
3. Classificação Binária  
Formulação da função logística (sigmoide), perda de entropia cruzada (BCELoss e BCEWithLogitsLoss), treinamento com nn. Module, uso de DataLoader.
4. Exploração e Pré-Processamento  
Leitura de datasets CSV, tratamento de valores ausentes, normalização, codificação de variáveis categóricas
5. Visualização e Interpretação de Resultados  
Gráficos de dispersão, análise de acurácia, precisão, recall em problemas de classificação binária

# Objetivo: Aplicar em um problema real o que já estudamos



## • Class Architecture

Usar, Modificar, Adaptar, Criar,  
Testar, Avaliar, Experimentar



# Free Data Science Projects

The best way to learn how to complete data projects is by building data projects. Dataquest learners spend their time working through real-world data challenges that teach learners to combine multiple skills and tools to solve a problem or accomplish a task.

- It builds confidence and experience.
- It trains learners for professional data work.
- It helps learners build a portfolio of projects.

A screenshot of a Dataquest project interface. On the left, there's a code editor window titled "script.py" containing the following Python code:

```
script.py    top20_deathtoll.csv
1 import pandas as pd
2 import matplotlib.pyplot as plt
3
4 top20_deathtoll = pd.read_csv('top20_deathtoll.csv')
5 plt.barh(top20_deathtoll['Country_Other'],
6          top20_deathtoll['Total_Deaths'])
7 plt.show()
```

To the right of the code editor is a "Output" window displaying a horizontal bar chart. The chart shows the total number of deaths for various countries. The top 10 countries listed are: USA, Brazil, India, Mexico, UK, Italy, France, Spain, Argentina, Colombia. The x-axis represents the total deaths, ranging from 0 to 300,000. The y-axis lists the countries. A callout box points to the first line of code with the text: "Read CSV file into a pandas DataFrame and assign the result to a variable." Another callout box points to the fifth line of code with the text: "Creates a horizontal bar chart using the Matplotlib library, with the countries on the y-axis and the total number of deaths on the x-axis." At the bottom of the interface, there are buttons for "Run Code", "Submit Answer", and "Next Screen".

## The place to talk about data Global online community of data enthusiasts

Subscribe to our weekly newsletter and join our Slack.  
We'll keep you informed about everything happening in the Club.

Email address

Join

You'll get an invite within 3 minutes



Talk about data, machine learning, and engineering



Attend weekly events and learn from free courses



Ask career questions and discuss career options

AQUELES SUGERIDOS PELOS PROFESSOR

# Escolher 01 **Problema**



### Heart Disease Prediction



### Admission Prediction



### Customer Purchase Pred.



### Alzheimer Risk Prediction



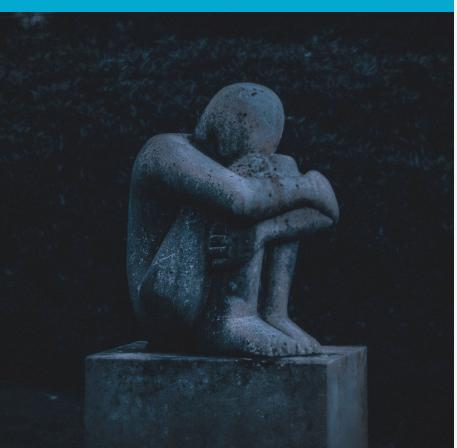
### Card Fraud Prediction



### School Dropout Prediction



### Depression Risk Prediction



### Customer Churn Pred.





maxim-eyengue / Heart-Disease-App



Type / to search



1



Code

Issues

Pull requests

Actions

Projects

Security

Insights



Heart-Disease-App

Public



1



0



1

main

1 Branch

0 Tags

Go to file

Add file

Code

About

maxim-eyengue Update README.md

03add8c · last month

49 Commits

.elasticbeanstalk

forced config files upload

2 months ago

data

Add files via upload

5 months ago

images

removed previous illustrations

2 months ago

slides

added presentation slides

last month

Dockerfile

update for betetr readability

2 months ago

LICENSE.txt

Rename LICENSE to LICENSE.txt

5 months ago

Pipfile

update for betetr readability

2 months ago

Pipfile.lock

update for betetr readability

2 months ago

README.md

Update README.md

last month

heart-dataset.zip

Add files via upload

5 months ago

midterm\_project.ipynb

update for betetr readability

2 months ago

no\_app\_predict\_test.py

update for betetr readability

2 months ago

An application to predict heart disease using patient data.

[Readme](#)[MIT license](#)[Activity](#)[1 star](#)[1 watching](#)[0 forks](#)[Report repository](#)

## Releases

No releases published

## Packages

No packages published

## Languages

 Jupyter Notebook 98.9% Python 1.0%



KatePril / admission-prediction

 Type / to search

Code

Issues

Pull requests

Actions

Projects

Security

Insights



admission-prediction

Public



1



0



Star



main



1 Branch



0 Tags

 Go to file

Add file



Code

## About

No description, website, or topics provided.

Readme

Activity

0 stars

1 watching

0 forks

Report repository

## Releases

No releases published

## Packages

No packages published

## Languages

Jupyter Notebook 99.8%

Other 0.2%

KatePril Update README.md

8f8d4f6 · 3 months ago

28 Commits

dataset

Add dataset and data exploration

6 months ago

gradio

[Add gradio interface](#)

5 months ago

Dockerfile

Add Dockerfile

[Add gradio interface](#)

5 months ago

Pipfile

Add gradio interface

5 months ago

Pipfile.lock

Add gradio interface

5 months ago

README.md

Update README.md

3 months ago

dv.bin

Add gradio interface

5 months ago

model.bin

Add gradio interface

5 months ago

notebook.ipynb

Add gradient boost

6 months ago

predict.py

Add gradio interface

5 months ago

supervisord.conf

Add gradio interface

5 months ago

train.py

Add gradio interface

5 months ago



Pei-Tong / Customer-Purchase-Predict\_ML

Q Type  to search



<> Code

Issues

Pull requests

Actions

Projects

Security

Insights



Customer-Purchase-Predict\_ML

Public

Watch

1

Fork

0

Star

0

main

1 Branch

0 Tags

Go to file

Add file

<> Code

Pei-Tong Add notebook.py

378dbd5 · 5 months ago

17 Commits

.devcontainer

Added Dev Container Folder

5 months ago

notebook\_files

mid project version1

5 months ago

.gitignore

Initial commit

5 months ago

README.md

Update README.md

5 months ago

app.py

mid project version1

5 months ago

customer\_purchase\_data.csv

mid project version1

5 months ago

dockerfile

mid project version1

5 months ago

notebook.ipynb

mid project version1

5 months ago

notebook.py

Add notebook.py

5 months ago

predict.ipynb

mid project version1

5 months ago

predict.py

mid project version1

5 months ago

random\_forest\_model.pkl

mid project version1

5 months ago

## About

Use random forest to predict customer purchase status

Readme

Activity

0 stars

1 watching

0 forks

Report repository

## Releases

No releases published

## Packages

No packages published

## Languages

Jupyter Notebook 97.6%

Python 2.4%



PATRICK079 / Operation-Mind-Shield

 Type / to search

Code

Issues

Pull requests

Actions

Projects

Security

Insights



Operation-Mind-Shield

Public

Watch

1



Fork

0



Star

3



main



1 Branch



0 Tags

 Go to file

Add file

&lt;&gt; Code



PATRICK079 Update README.md

66542cb · 5 months ago

70 Commits



.devcontainer

Added Dev Container Folder

7 months ago



Dockerfile

Add files via upload

5 months ago



Dockerrun.aws.json

Add files via upload

5 months ago



README.md

Update README.md

5 months ago



alzer\_Api.py

Add files via upload

5 months ago



alzheimer image.png

Add files via upload

7 months ago



alzheimers.py

Update alzheimers.py

5 months ago



alzheimers\_disease\_data.csv

Add files via upload

7 months ago



alzheimers\_notebook-2.ipynb

Add files via upload

7 months ago



col\_names\_cat.pk1

Add files via upload

7 months ago



final\_model\_cat.pk1

Add files via upload

7 months ago



final\_scaler\_cat.pk1

Add files via upload

7 months ago

## About

No description, website, or topics provided.

Readme

Activity

3 stars

1 watching

0 forks

Report repository

## Releases

No releases published

## Packages

No packages published

## Languages

Jupyter Notebook 99.7%

Other 0.3%

 karinait / credit\_card\_fraud\_predictor

Code Issues Pull requests Actions Projects Security Insights

 Type / to search

 Watch 1

 Fork 0

 Star 1

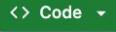


## credit\_card\_fraud\_predictor Public

 main  1 Branch  0 Tags

 Go to file 

 Add file

 Code

### About

No description, website, or topics provided.

 Readme

 Activity

 1 star

 1 watching

 0 forks

Report repository

### Releases

No releases published

### Packages

No packages published

### Languages



Language	Percentage
Jupyter Notebook	99.7%
Other	0.3%

 README



Maxkaizo / ---\_ML-Zoomcamp-2024

 Type / to search

Code

Issues

Pull requests

Actions

Projects

Security

Insights



---\_ML-Zoomcamp-2024

Public



1



0



0



main



1 Branch



0 Tags

 Go to file

Add file



Code

## About

No description, website, or topics provided.

Readme

Activity

0 stars

1 watching

0 forks

Report repository

## Releases

No releases published

## Packages

No packages published

## Languages

Jupyter Notebook 99.4%

Other 0.6%

Maxkaizo modified: 3\_modeling.ipynb

8fe33eb · 5 months ago

47 Commits

Dataset Documentation

modified: 3\_modeling.ipynb

5 months ago

Dataset ENAPE Statisticc

renamed: enape\_2021\_fd.pdf.zip -> Dataset Documentat...

5 months ago

Dataset Raw Download

renamed: enape\_2021\_fd.pdf.zip -> Dataset Documentat...

5 months ago

\_\_pycache\_\_

new file: Dockerfile

5 months ago

.dockerignore

new file: .dockerignore

5 months ago

.gitignore

new file: .gitignore

5 months ago

1\_pre\_processing.ipynb

modified: 1\_pre\_processing.ipynb

5 months ago

2\_edo.ipynb

modified: 1\_pre\_processing.ipynb

5 months ago

3\_modeling.ipynb

modified: 3\_modeling.ipynb

5 months ago

Dockerfile

new file: Dockerfile

5 months ago

Pipfile

new file: .dockerignore

5 months ago

Pipfile.lock

new file: .dockerignore

5 months ago



Wali-Mohamed / DepressionRiskPredictor

Q Type / to search



Code

Issues

Pull requests

Actions

Projects

Security

Insights



DepressionRiskPredictor

Public

Watch

1



Fork

0



Star

1



main

1 Branch

0 Tags

Go to file

Add file

Code

About

No description, website, or topics provided.

Readme

Activity

1 star

1 watching

0 forks

Report repository

Releases

No releases published

Packages

No packages published

Languages

Jupyter Notebook 99.8%

Other 0.2%

Wali-Mohamed	Update README.md	8f387ea · 5 months ago	37 Commits
.ipynb_checkpoints	app is working	5 months ago	
data	app is working	5 months ago	
depression_predictor	app is on cloud	5 months ago	
images	model fine tuning with LR done	5 months ago	
plots	app is on cloud	5 months ago	
Dockerfile	final	5 months ago	
Pipfile	app is working	5 months ago	
Pipfile.lock	app is working	5 months ago	
README.md	Update README.md	5 months ago	
note.md	app is working	5 months ago	
notebook.ipynb	app is on cloud	5 months ago	
predict.py	app is working	5 months ago	



dmytrovotko / ml-churn-prediction

Type to search



Code

Issues

Pull requests

Actions

Projects

Security

Insights



ml-churn-prediction

Public

Watch 1

Fork 0

Star 1

main

1 Branch 0 Tags

Go to file

Add file

Code

dmytrovotko Readme

061f2a5 · 5 months ago 18 Commits



Dataset

5 months ago



Dataset

5 months ago



app.py add route '/'

5 months ago



Readme, screenshots

5 months ago



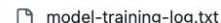
Readme

5 months ago



EDA notebook

5 months ago



Model training log

5 months ago



EDA notebook

5 months ago



## ML project Predicting Customer Churn for a Gym

A machine learning project for DataTalks Club Machine Learning ZoomCamp '24.

### About

No description, website, or topics provided.

Readme

Activity

1 star

1 watching

0 forks

Report repository

### Releases

No releases published

### Packages

No packages published

### Languages



## PESQUISAR DADOS

Ex.: cursos Etiquetas mais comuns   

## GRUPOS



Biblioteca

Comunicados e  
DocumentosContratos e  
ConvêniosDespesas e  
Orçamento

Ensino



Extensão



Institucional



Materiais



Patrimônio



Pesquisa



Pessoas



Processos

## DADOS ABERTOS

O Portal de Dados Abertos da UFRN tem como objetivo disponibilizar para toda a comunidade dados e informações públicas da Instituição, que podem ser usadas no desenvolvimento de aplicativos e ações.

## ESTATÍSTICAS

89

conjuntos de dados

12

grupos

## Escolha Opcional

A partir dos dados abertos da UFRN escolher um problema em aberto.

Validar a ideia antes com o professor.

Por exemplo:

Previsão de aprovação de auxílio financeiro, bolsa de monitoria, aprovação em disciplina, concessão de auxílio moradia, etc

# Requisitos

- Análise Exploratória dos dados

Compreender a distribuição dos dados, se há classes desbalanceadas, estatística das variáveis.

- Preparação e Engenharia de Atributos

Realizar a limpeza (remoção ou imputando registros incompletos) e transformar atributos para um formato adequado ao modelo.

- Implementação do Modelo em Pytorch

Utilizando a classe base fornecida pelo professor em aula, implementar a regressão logística.

- Treinamento do Modelo

Dividir o dataset em conjunto de treino e teste, configurar um otimizador, iterar por múltiplas épocas, calcular a loss e gradientes.

- Avaliação do Modelo

Após treinado, avaliar o desempenho do classificador no conjunto de teste.

- Documentação dos Resultados

Apresentar de forma clara os resultados obtidos, incluindo as métricas calculadas, alguma discussão sobre o que elas indicam, e possivelmente gráficos ou tabelas que ilustram o desempenho.

# Requisitos

- Trabalho Individual ou em Dupla
- Repositório no Github

Arquivo README, descrição dos resultados, códigos, etc.

Submissão no Sigaa indicar o link do repositório do Github.

Exemplo de repositório no estilo model card:

[https://github.com/ivanovitchm/mlops\\_nd\\_c3/tree/main/pipeline](https://github.com/ivanovitchm/mlops_nd_c3/tree/main/pipeline)

Exemplo de repositório geral:

[https://github.com/ivanovitchm/mlops\\_nd\\_c3](https://github.com/ivanovitchm/mlops_nd_c3)

- Nota: 10 pontos com peso 1.

		ABRIL						
		DOMINGO	SEGUNDA	TERÇA	QUARTA	QUINTA	SEXTA	SÁBADO
		04 CRESC.	12 CHEIA	1	2	3	4	5
	6	7	8	9	10	11	12	
	13	14	15	16	17	18	19	
	20	21	22	23	24	25	26	
	27	28	29	30		20 MING.		27 NOVA

18 - SEXTA-FEIRA SANTA  
20 - PÁSCOA  
21 - TIRADENTES  
22 - DESCOBRIMENTO DO BRASIL

		MAIO						
		DOMINGO	SEGUNDA	TERÇA	QUARTA	QUINTA	SEXTA	SÁBADO
		18	04 CRESC.	12 CHEIA	20 MING.	27 NOVA	1	2
	19	4	5	6	7	8	9	10
	20	11	12	13	14	15	16	17
	21	18	19	20	21	22	23	24
	22	25	26	27	28	29	30	31

18 - SEXTA-FEIRA SANTA  
20 - PÁSCOA  
21 - TIRADENTES  
22 - DESCOBRIMENTO DO BRASIL

ABRIL 2025

DOM	SEG	TER	QUA	QUI	SEX	SAB
1	2	3	4	5	6	7
8	9	10	11	12	13	14
15	16	17	18	19	20	21
22	23	24	25	26	27	28
29	30					

LEMBRETE

1	2	3	4	5	6	7
8	9	10	11	12	13	14
15	16	17	18	19	20	21
22	23	24	25	26	27	28
29	30					

JUNHO 2025

1	2	3	4	5	6	7
8	9	10	11	12	13	14
15	16	17	18	19	20	21
22	23	24	25	26	27	28
29	30					

01-DIA DO TRABALHO  
11-DIA DAS MÃES