

# Aprendizado de Máquina e Reconhecimento de Padrões 2021.2



## Aula 00 Apresentação da Disciplina

Prof. Samuel Martins (Samuka)

*samuel.martins@ifsp.edu.br*



# Objetivos

---

- Capacitar o aluno para implementar soluções baseadas em métodos de Inteligência Artificial, mais especificamente baseados em Aprendizado de Máquina e Reconhecimento de Padrões, por meio de implementação de algoritmos básicos da literatura.
- Capacitar o aluno a utilizar bibliotecas, preferencialmente de código aberto, voltadas a tarefa de reconhecimento de padrões.
- Estimular o aluno a desenvolver o raciocínio lógico voltado para a aplicação deste tipo de algoritmos à solução de problemas de diversos domínios.

# Programa

---

- Introdução aos Conceitos de Aprendizado de Máquina e Reconhecimento de Padrões;
- Regressão linear e vetores de características;
- Sobreajuste, treinamento, validação e teste;
- Problemas de classificação e limites de decisão;
- Método dos vizinhos mais próximos;
- Classificadores lineares;
- Redução de dimensionalidade de dados;
- Classificação não supervisionada;
- Redes Neurais;
- Máquinas de vetores de suporte.

# Dinâmica do Curso

---

## Aulas

- Quinta-feira: 19:10 – 22:40
- **Aulas ao vivo\***, videoaulas, material de leitura, etc
- Material disponibilizado no **Moodle** *antes* da aula
- As aulas ao vivo serão gravadas e disponibilizadas posteriormente

## \*Google Meet

- Criar conta com o **email institucional** ([\\*@aluno.ifsp.edu.br](mailto:*@aluno.ifsp.edu.br))
- Informações: [ifsp.edu.br/google](http://ifsp.edu.br/google)

Link das aulas: <http://tiny.cc/D2APR-2021-2-aulas>

## Presença

- Questões semanais ao final da aula (via Moodle)
- Tais questões também valerão nota (ver **avaliação**)

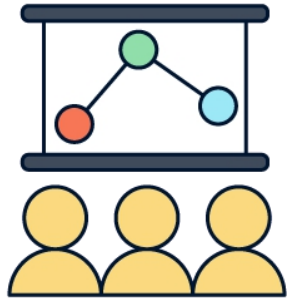
## Comunicação

- Mural de avisos no **Moodle**
- SUAP
- Email: [samuel.martins@ifsp.edu.br](mailto:samuel.martins@ifsp.edu.br)
  - No assunto do email, procure usar a sigla **[D2APR 2021.2] Assunto do Email**

# Aulas guiadas (majoritariamente) por Projetos

---

## Projetos de Portfólio



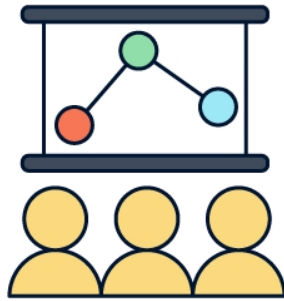
## Projetos de Estudo



# Aulas guiadas (majoritariamente) por Projetos

---

## Projetos de Portfólio



- Problemas mais próximos de **cenários reais** das empresas
- Voltados para a resolução de **problemas de negócio**:
  - Qual o **objetivo de negócio** resolvido?
  - Qual o **retorno** que sua solução traz?
  - *Data-analytic thinking*

Ex:

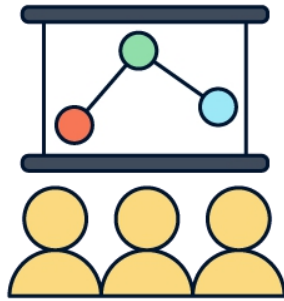
- Previsão de vendas, churn
- Segmentação de clientes
- Classificação de doenças

## Projetos de Estudo



# Aulas guiadas (majoritariamente) por Projetos

## Projetos de Portfólio



- Problemas mais próximos de **cenários reais** das empresas
- Voltados para a resolução de **problemas de negócio**:
  - Qual o **objetivo de negócio** resolvido?
  - Qual o **retorno** que sua solução traz?
  - *Data-analytic thinking*

Ex:

- Previsão de vendas, churn
- Segmentação de clientes
- Classificação de doenças

## Projetos de Estudo



- Problemas **mais simples**
- Voltados para o **entendimento** de **conceitos** e técnicas
- Sem relação com **objetivos de negócio**

Ex:

- Classificação de flores (iris)
- Sobreviventes do Titanic
- Aluguéis de Boston

# Desenvolvimento em Ciclos (“Sprints”)

---

## Sprint 1

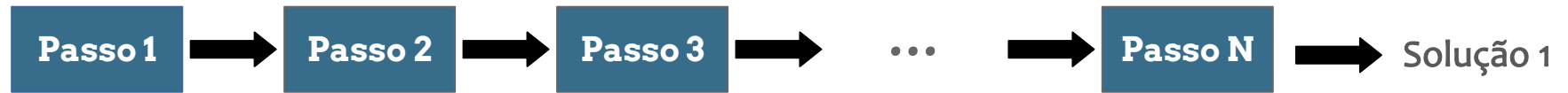




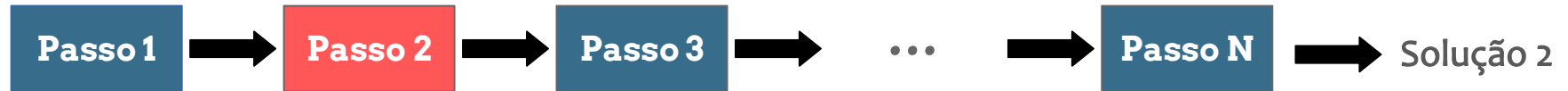
# Desenvolvimento em Ciclos (“Sprints”)

---

## Sprint 1



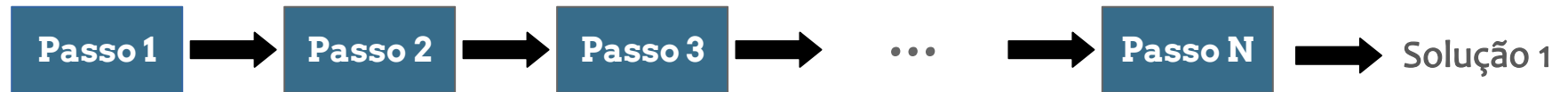
## Sprint 2



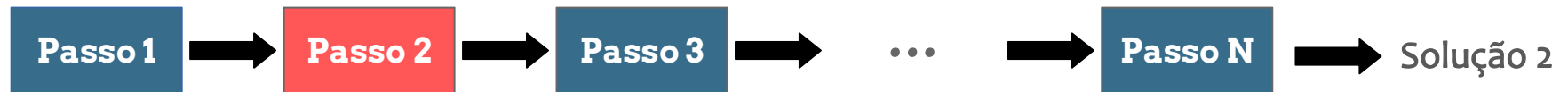
# Desenvolvimento em Ciclos (“Sprints”)

---

## Sprint 1



## Sprint 2



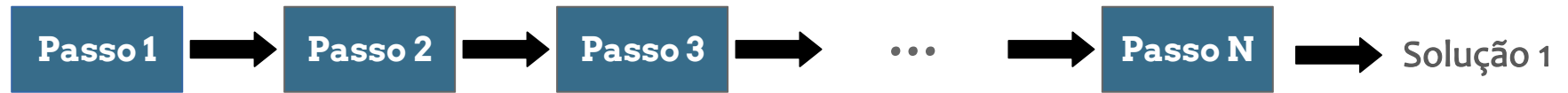
## Sprint 3



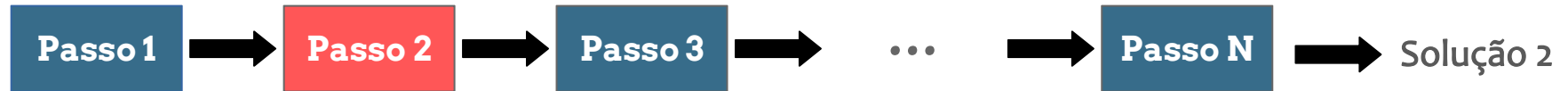
# Desenvolvimento em Ciclos (“Sprints”)

---

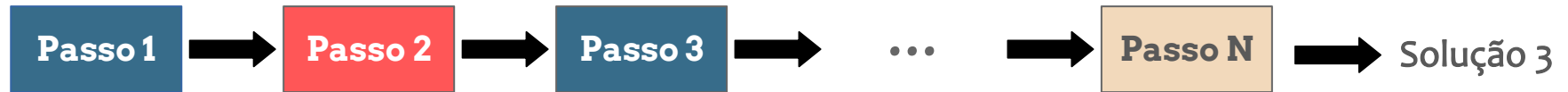
## Sprint 1



## Sprint 2



## Sprint 3



...

# Onde encontrar datasets?

kaggle

## Datasets

Explore, analyze, and share quality data. [Learn more](#) about data types, creating, and collaborating.

<https://www.kaggle.com/datasets>

Datasets

4,554 machine learning datasets

<https://paperswithcode.com/datasets>

Dataset Search

<https://datasetsearch.research.google.com/>



UC Irvine  
Machine Learning  
Repository

Datasets

<https://archive-beta.ics.uci.edu/ml/datasets>



Microsoft | Microsoft Research Open Data

<https://msropendata.com/datasets>

roboflow

<https://public.roboflow.com/>



dados.gov.br

PORTAL BRASILEIRO DE DADOS ABERTOS

<https://dados.gov.br/>



OpenDataNI  
SonraíOscailteTÉ

<https://www.opendatani.gov.uk/>



VisualData

<https://visualdata.io/discovery>



DATA.GOV

<https://www.data.gov/>

Awesome Public Datasets

<https://github.com/awesomedata/awesome-public-datasets>



EU Open Data Portal  
Access to European Union open data

<https://data.europa.eu/euodp/data/dataset>

# Onde encontrar datasets?

com questões de negócio

kaggle

## Datasets

Explore, analyze, and share quality data. [Learn more](#) about data types, creating, and collaborating.

<https://www.kaggle.com/datasets>

ZIND!

<https://zindi.africa/>

dockship

<https://dockship.io/>

com desafios de  
contratação

CrowdANALYTIX

<https://www.crowdanalytix.com/community>

innocentive

<https://www.innocentive.com/>

DRIVEN DATA

<https://www.drivendata.org/>

DATASOURCE.AI

<https://www.datasource.ai/en>

# Atendimento

---

Quintas-feiras: 18:00 – 19:00

- **Google Meet:** <http://tiny.cc/atendimento-2021-2-samuka>

# Avaliação

---

A avaliação será baseada na participação ativa em sala de aula, laboratórios práticos e projetos, sendo:

- Ao final de cada aula, a(o) aluna(o) deverá realizar/responder uma **atividade** sobre o assunto da aula, via Moodle.
  - A(o) aluna(o) terá até o **início da próxima aula** para submeter a atividade.
  - Caso o(a) estudante não entregue ou copie a atividade, ela será **zerada**.
- Teremos uma série de **tarefas/atividades** com no máximo 2 integrantes ao longo do curso.
- Teremos um **projeto prático** interdisciplinar em grupo

# Avaliação

---

$$NF: 0.45 * T + 0.5 * P + 0.05 * Q$$

Aprovação  
(se  $NF \geq 6.0$  e Frequência  $\geq 75\%$ )

T: Média das tarefas do curso [0, 10]

P: Nota do projeto interdisciplinar [0, 10]

Q: Média das questões/atividades semanais

NF: Nota final



# Calendário

## AGOSTO

D	S	T	Q	Q	S	S
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	31				

## SETEMBRO

D	S	T	Q	Q	S	S
			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		

## OUTUBRO

D	S	T	Q	Q	S	S
					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

## NOVEMBRO

D	S	T	Q	Q	S	S
	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				

## DEZEMBRO

D	S	T	Q	Q	S	S
			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	31	



Aulas

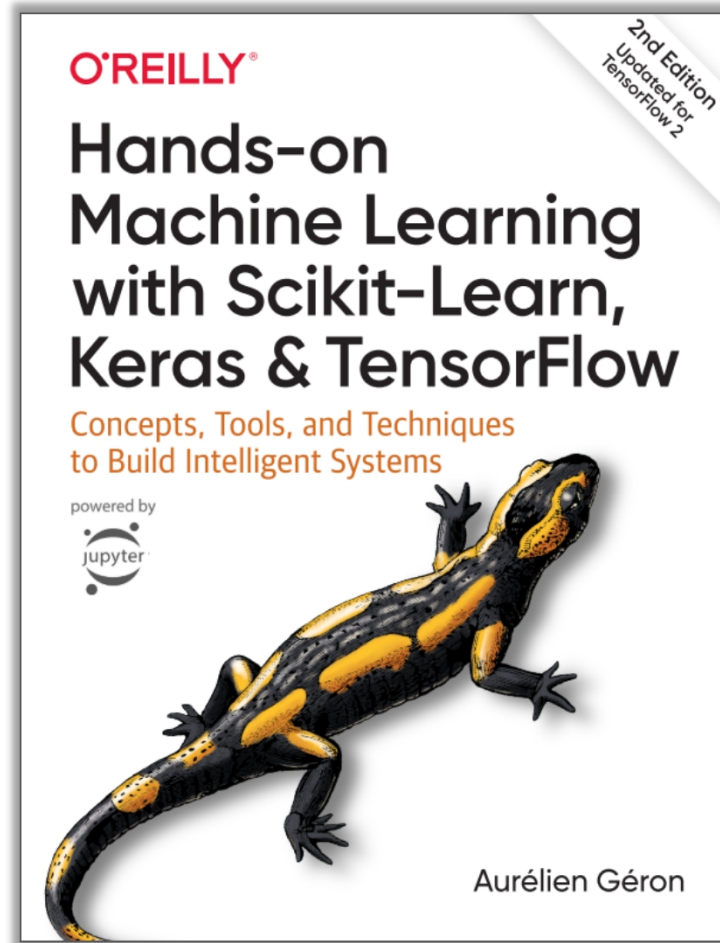


Feriados / Recesso

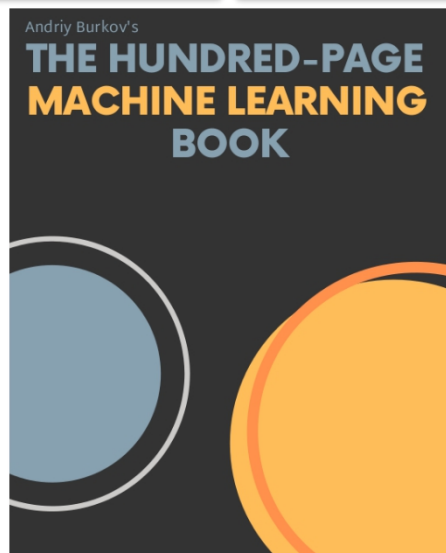
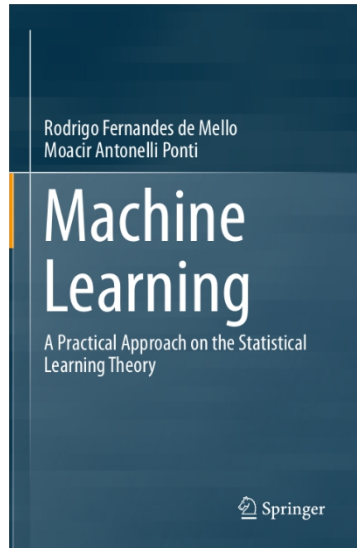
Avaliação/Entregas

# Livro Base

---



# Outros Materiais



<https://www.coursera.org/learn/machine-learning>



<https://towardsdatascience.com/>



<https://machinelearningmastery.com/blog/>



<https://sejaumdatascientist.com/>

# Referências

---

## **Bibliografia:**

- HASTIE, T.; TIBSHIRANI, R.; FRIEDMAN, J. The elements of statistical learning: data mining, inference, and prediction. Springer Science & Business Media, 2009.
- DUDA, R. O.; HART P. E.; STORK D. G. Pattern classification. John Wiley & Sons, 2012.
- CARVALHO, A. et al. Inteligência Artificial–uma abordagem de aprendizado de máquina. Rio de Janeiro: LTC. 2011.

## **Bibliografia Complementar:**

- GOODFELLOW, I.; BENGIO, Y.; COURVILLE, A. Deep learning. MIT press, 2016.
- BISHOP, C. M. Pattern recognition and machine learning. Springer, 2006.
- MITCHELL, T. Machine Learning. McGraw-Hill, 1997.
- GÉRON A. Hands-on machine learning with Scikit-Learn, Keras, and TensorFlow: Concepts, tools, and techniques to build intelligent systems. O'Reilly Media, 2019.

# Aprendizado de Máquina e Reconhecimento de Padrões 2021.2



## Aula 00 Apresentação da Disciplina

Prof. Samuel Martins (Samuka)

*samuel.martins@ifsp.edu.br*

