

# CONFIA: Combate a Fake News utilizando IA

Professor Paulo Márcio Souza Freire

**Fundação de Apoio à Escola Técnica / RJ - FAETEC  
(DDE-CPEM-DESUP)**

**Instituto Militar de Engenharia - IME**

3 de junho de 2024

# Tópicos abordados

- 1 Fake News
- 2 CONFIA
  - Educação Midiática com apoio Computacional
  - Combate Automático às Fake News nos Meios Digitais
- 3 Considerações Finais
- 4 Referências

## ▪ Definição de *Fake News* adotada

- Publicações intencionalmente e verificadamente falsas (Disinformation)

## ▪ Comportamento disseminativo das *Fake News*

- Divulgação (Efeito Câmara de Eco)
  - Publicação
  - Propagação

## ▪ Razões para o poder de divulgação

- Fatores inerentes ao homem [Shu et al., 2017]
  - Confirmação da opinião
  - Ganhos x perdas
  - Aceitação dos outros
- Carência de legislação punitiva
- Facilidade de criação de contas [Conroy et al., 2015]

# Exemplos

- Em 2018, Fake News divulgadas em um aplicativo de troca de mensagens podem ter ocasionado linchamentos na Índia<sup>1</sup>
- Mulher foi espancada no litoral de São Paulo após uma Fake News, divulgada em uma rede social, afirmar que a dona de casa sequestrava crianças para rituais de magia negra<sup>2</sup>
- Desinformação sobre COVID-19 a partir de *Fake News* divulgadas em meios digitais [Mejova and Kalimeri, 2020]
- Fake News têm prejudicado a assistência nas enchentes do Rio Grande do Sul<sup>3</sup>

---

<sup>1</sup><https://www.bbc.com>

<sup>2</sup><http://g1.globo.com>

<sup>3</sup><https://netlab.eco.ufrj.br/post/enchentes-norio-grande-do-sul-uma-analise-da-desinformacao-multiplataforma-sobre-o-desastre-climati>

# "Soluções?"

## ▪ Linhas de Ação - Visão Geral

- Criação de Legislação Punitiva<sup>4</sup>
- Serviços de Checagem de Fatos<sup>5</sup>
- Educação Midiática [UNESCO, 2019]
- Combate Automático em Meios Digitais [Freire and Goldschmidt, 2020]

---

<sup>4</sup><https://www12.senado.leg.br/noticias/materias/2020/06/02/nova-versao-de-lei-contra-fake-news-tera-restricoes-a-contas-anonimas-e-mais-poder-a-denuncias-de-usuarios>

<sup>5</sup><https://www.aosfatos.org/>

## ■ Linhas de Ação - CONFIA

- Educação Midiática com apoio Computacional
- Combate Automático em Meios Digitais



# Educação Midiática com apoio Computacional

## Jogos Educacionais Digitais com Tutoria Inteligente (JEDi)

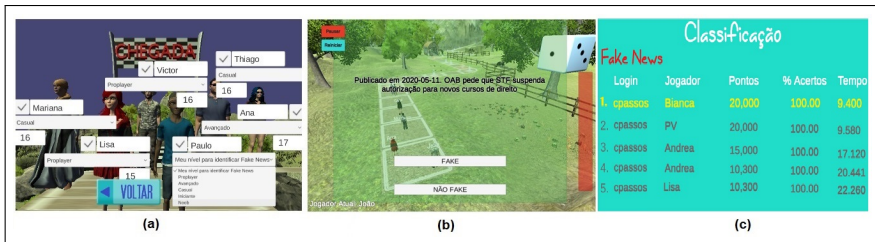
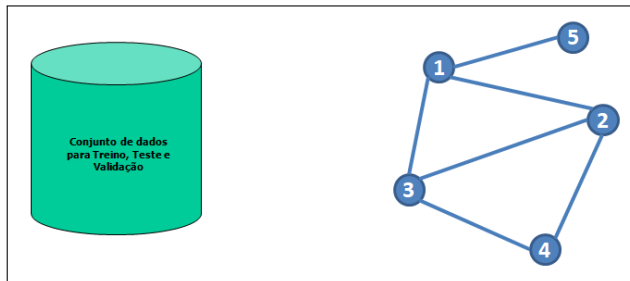


Figura: Telas do JEDi: (a) Configuração; (b) Principal; (c) Fim de partida

- Publicação: **SBIE 2020 - Paper Award (First place)**  
**Revista Brasileira de Informática na Educação**  
**2021**  
**Education and Information Technologies**  
**2023**

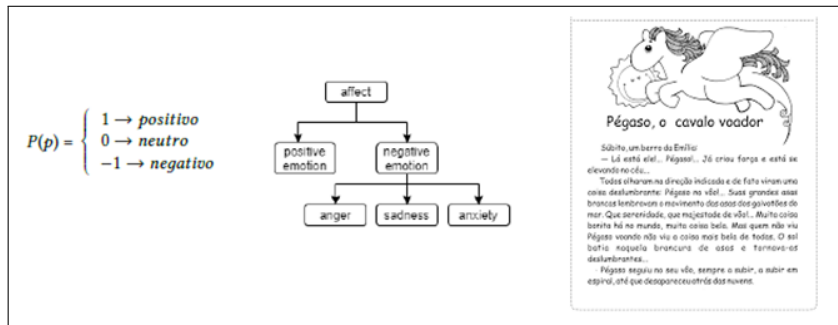
# Combate Automático às Fake News

## Construção de Datasets em português - dados de propagação



- Publicação: WebMedia 2020 - Paper Award (First place)  
SBBD 2024 (Submetido)

## Detecção de Fake News via Classes gramaticais e Análise de Sentimentos em textos e imagens



- Publicação: WebMedia 2020  
BraSNAM 2021  
WebMedia 2021

## Checagem de Fatos Automática via WebSemântica

### "It is impossible for journalists to enter Tibetan areas"

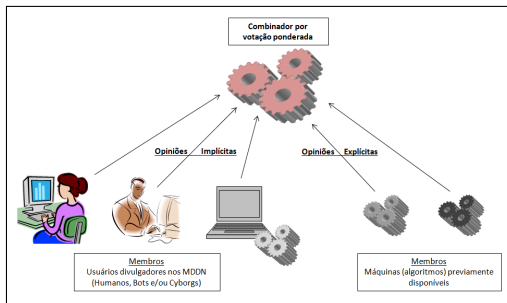
Philip Bruno, correspondent for "World" in China, said that journalists of the AFP who have been deported from the Tibetan province of Qinghai "were not illegal."

**Facts** The Dalai Lama denounces the "hell" imposed since he fled Tibet in 1959

**Video** Anniversary of the Tibetan rebellion: China on guard



## Detecção de Fake News via Crowd



- Publicação: WebMedia 2019  
ENIAC 2019 - Paper Award (2nd place)  
Expert Systems with Applications 2021  
WebMedia 2022 - CTD Award (First place)  
Expert Systems with Applications 2024  
SBSI 2024

## Detecção de Fake News via Ensemble



- Publicação: SBBD 2023  
WebMedia 2023

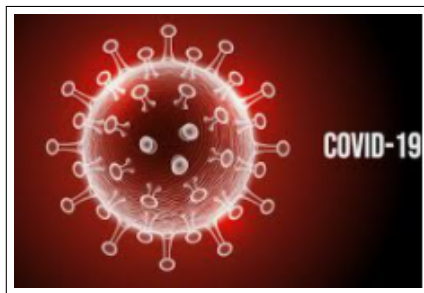


## Estado da arte sobre Combate Automático às Fake News



- Publicação: SBBD 2019 (Minicurso)  
RMCT 2021  
ERSI 2021 (Minicurso)

## Um Ambiente Computacional para Combate automático a Fake News sobre COVID-19 e outras SARS em Redes Sociais Virtuais



- Fomento: CNPq 401662/2020-9
- Publicação: WebMedia 2022

## Um Ambiente Computacional para Combate automático a Fake News sobre Óleo e Gás em Redes Sociais Virtuais



- Projeto: Centro de Inteligência Artificial do Estado do Rio de Janeiro (CIA-Rio)

# Considerações Finais

# Bibliografia I



Conroy, N., Rubin, V., and Chen, Y. (2015).

Automatic deception detection: Methods for finding fake news.  
*Association for Information Science and Technology*, 52:1–4.



Freire, P. and Goldschmidt, R. (2020).

Combatendo fake news nas redes sociais via crowd signals implícitos.  
*In Anais do XVI Encontro Nacional de Inteligência Artificial e Computacional*, pages 424–435, Porto Alegre, RS, Brasil. SBC.



Mejova, Y. and Kalimeri, K. (2020).

Advertisers jump on coronavirus bandwagon: Politics, news, and business.  
*ArXiv*, abs/2003.00923.



Shu, K., Sliva, A., Wang, S., Tang, J., and Liu, H. (2017).

Fake news detection on social media: A data mining perspective.  
*SIGKDD Explor. Newsl.*, 19(1):22–36.

## Bibliografia II



UNESCO (2019).

*Manual para Educação e Treinamento em Jornalismo.*

UNESCO.

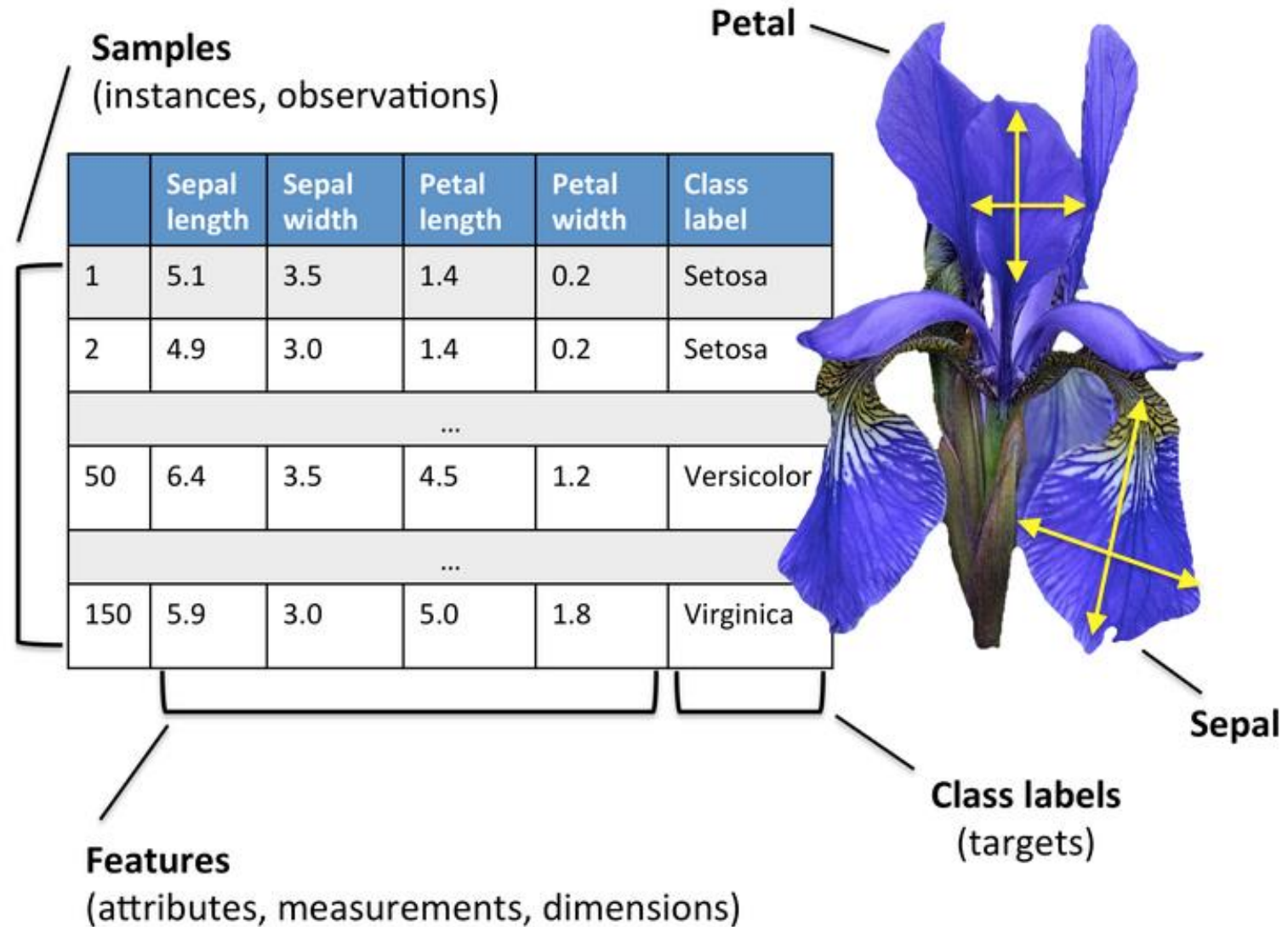
# Exemplo Teórico – IA

---

**Prof. Paulo Márcio Souza Freire  
CPEM-DESUP / DDE  
FAETEC**

**Instituto Militar de Engenharia  
IME**

# Dataset 1 - Iris





# Dataset 2 – Jogar Tênis

Aparência	Temperatura	Umidade	Vento	Jogar Tênis
Ensolarado	Quente	Alta	Fraco	Não
Ensolarado	Quente	Alta	Forte	Não
Nublado	Quente	Alta	Fraco	Sim
Chuvoso	Moderado	Alta	Fraco	Sim
Chuvoso	Fresco	Normal	Fraco	Sim
Chuvoso	Fresco	Normal	Forte	Não
Nublado	Fresco	Normal	Forte	Sim
Ensolarado	Moderado	Alta	Fraco	Não
Ensolarado	Fresco	Normal	Fraco	Sim
Chuvoso	Moderado	Normal	Fraco	Sim
Ensolarado	Moderado	Normal	Forte	Sim
Nublado	Moderado	Alta	Forte	Sim
Nublado	Quente	Normal	Fraco	Sim
Chuvoso	Moderado	Alta	Forte	Não

# Aplicação

---

- Uso de Ferramentas
- Programação
  - Bibliotecas
  - “Manual”

# Aplicação

---

- Uso de Ferramentas
- Programação
  - **Bibliotecas**
  - “Manual”

# Mineração de Dados

## Classificador Bayesiano (Exemplo)

- Um dia ensolarado, quente, de alta umidade e vento fraco é adequado para jogar tênis ?

	Aparência	Temperatura	Umidade	Vento	Jogar Tênis
1	Ensolarado	Quente	Alta	Fraco	Não
2	Ensolarado	Quente	Alta	Forte	Não
3	Nublado	Quente	Alta	Fraco	Sim
4	Chuvoso	Moderado	Alta	Fraco	Sim
5	Chuvoso	Fresco	Normal	Fraco	Sim
6	Chuvoso	Fresco	Normal	Forte	Não
7	Nublado	Fresco	Normal	Forte	Sim
8	Ensolarado	Moderado	Alta	Fraco	Não
9	Ensolarado	Fresco	Normal	Fraco	Sim
10	Chuvoso	Moderado	Normal	Fraco	Sim
11	Ensolarado	Moderado	Normal	Forte	Sim
12	Nublado	Moderado	Alta	Forte	Sim
13	Nublado	Quente	Normal	Fraco	Sim
14	Chuvoso	Moderado	Alta	Forte	Não

Classe Não (5/14)

Classe Sim (9/14)

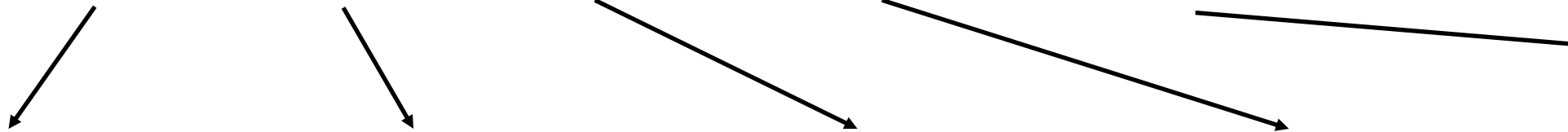
Atributos Aparência,  
Temperatura, Umidade e  
Vento

# Mineração de Dados

## Classificador Bayesiano (Exemplo)

- ensolarado, quente, alta umidade, vento fraco, sim

$$(2/9) * (2/9) * (3/9) * (6/9) * (9/14) = 0,007$$



	Aparência	Jogar Tênis
1	Nublado	Sim
2	Chuvoso	Sim
3	Chuvoso	Sim
4	Nublado	Sim
5	<b>Ensolarado</b>	Sim
6	Chuvoso	Sim
7	<b>Ensolarado</b>	Sim
8	Nublado	Sim
9	Nublado	Sim

	Temperatura	Jogar Tênis
1	<b>Quente</b>	Sim
2	Moderado	Sim
3	Fresco	Sim
4	Fresco	Sim
5	Fresco	Sim
6	Moderado	Sim
7	Moderado	Sim
8	Moderado	Sim
9	<b>Quente</b>	Sim

	Umidade	Jogar Tênis
1	<b>Alta</b>	Sim
2	<b>Alta</b>	Sim
3	Normal	Sim
4	Normal	Sim
5	Normal	Sim
6	Normal	Sim
7	Normal	Sim
8	<b>Alta</b>	Sim
9	Normal	Sim

	Vento	Jogar Tênis
1	<b>Fraco</b>	Sim
2	<b>Fraco</b>	Sim
3	<b>Fraco</b>	Sim
4	Forte	Sim
5	<b>Fraco</b>	Sim
6	<b>Fraco</b>	Sim
7	Forte	Sim
8	Forte	Sim
9	<b>Fraco</b>	Sim

	Jogar Tênis
1	Não
2	Não
3	<b>Sim</b>
4	<b>Sim</b>
5	<b>Sim</b>
6	Não
7	<b>Sim</b>
8	Não
9	<b>Sim</b>
10	<b>Sim</b>
11	<b>Sim</b>
12	<b>Sim</b>
13	<b>Sim</b>
14	Não

# Mineração de Dados

## Classificador Bayesiano (Exemplo)

- $P(\text{Jogar} = \text{Sim} \mid \text{ensolarado, quente, alta umidade, vento fraco}) = 0,0071$

$$(2/9) * (2/9) * (3/9) * (6/9) * (9/14) = 0,007$$

- $P(\text{Jogar} = \text{Não} \mid \text{ensolarado quente, alta umidade, vento fraco}) = 0,0274$

$$(3/5) * (2/5) * (4/5) * (2/5) * (5/14) = 0,027$$

**Resposta do Classificador Bayesiano: Jogar = Não**

# Linguagem R

# Ambiente de Desenvolvimento

---

## Instalação da Linguagem R

- <https://cran.r-project.org/mirrors.html>

## Instalação da IDE

- <https://www.rstudio.com/products/rstudio/download/>