



Plataforma de Ciência de Dados Aplicada à Saúde

Aprendizagem de máquina supervisionada

Fiocruz
ICICT

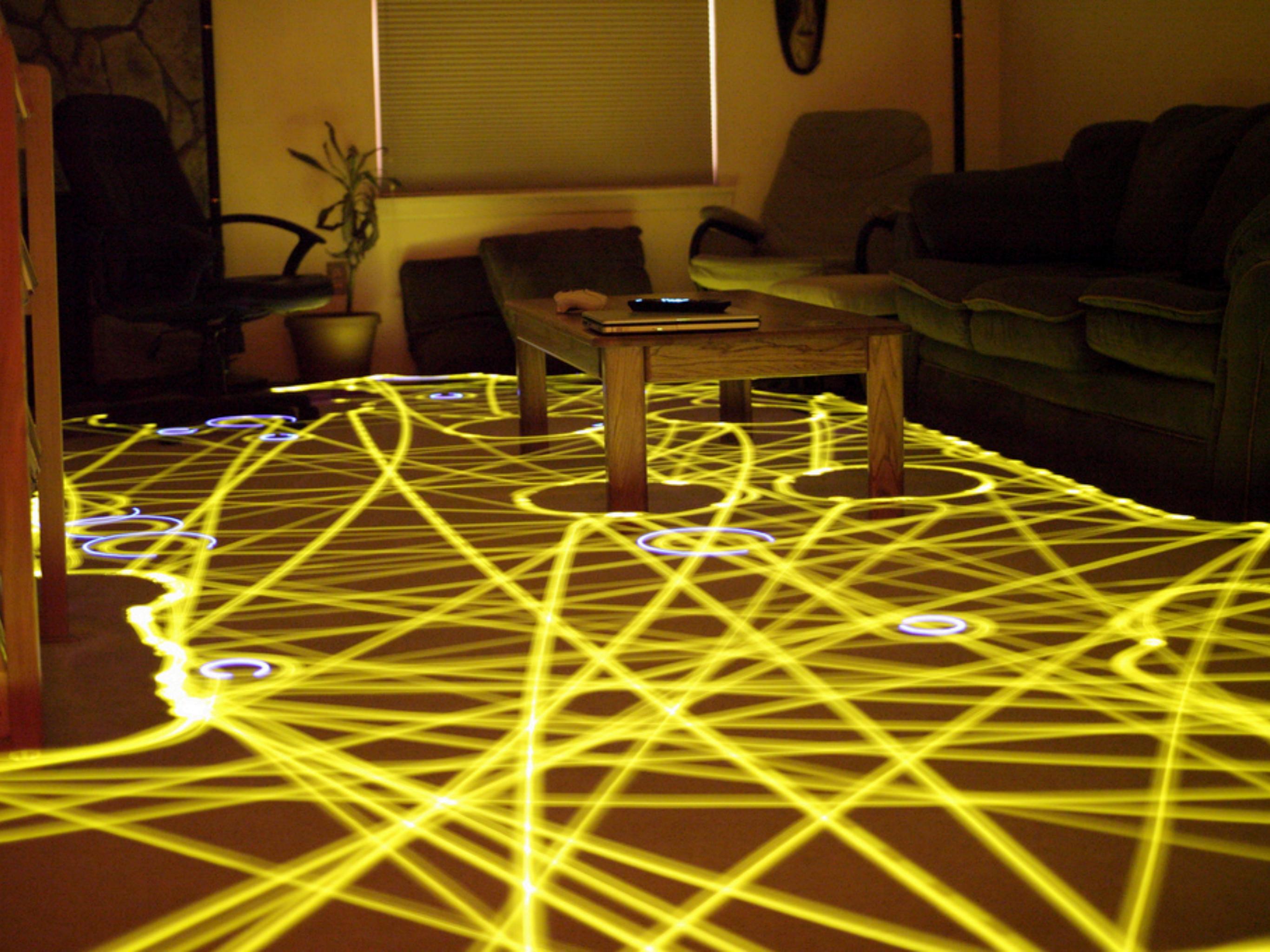
Vanderlei Pascoal
Raphael Saldanha

Roteiro

- ❖ Base teórica
- ❖ Recursos computacionais (softwares e pacotes)
- ❖ Intervalo
- ❖ Prática com R

Machine Learning





Novo algoritmo

- ❖ Meta: ganhar pontos
- ❖ Cada 50g de poeira: +10 pontos
- ❖ Cada 10s sem aspirar poeira: -5 pontos



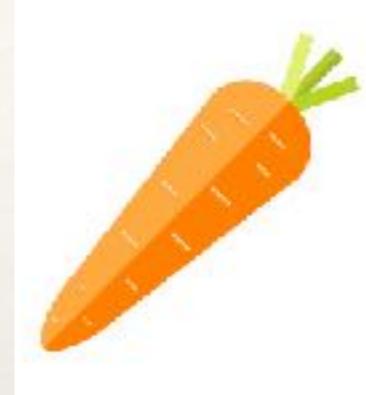
O que lembrar

- ❖ Tarefa de aprendizagem bem definida
- ❖ Parâmetros claros e abrangentes

Aprendizagem supervisionada

- ❖ Detecção de padrões
- ❖ Associação de padrões a um resultado
- ❖ Predição de um resultado

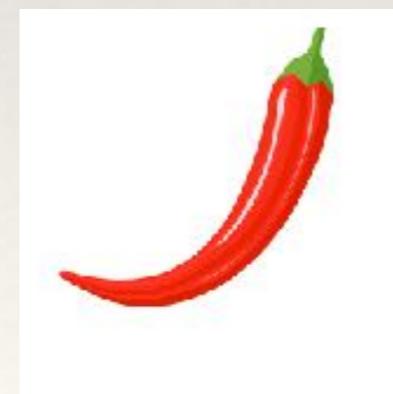
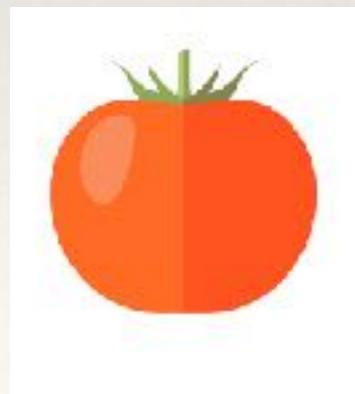
Classificação de Receitas



Italiana



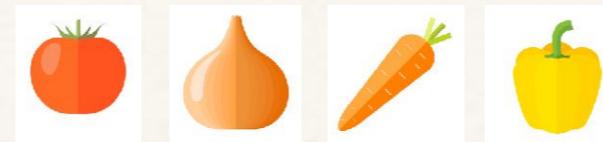
Mexicana



Mexicana

Receitas

Treino



Italiana



Mexicana

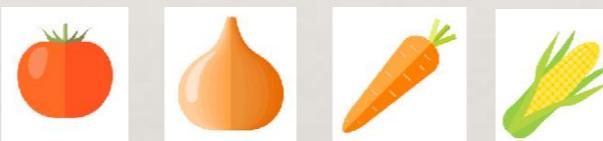


Italiana

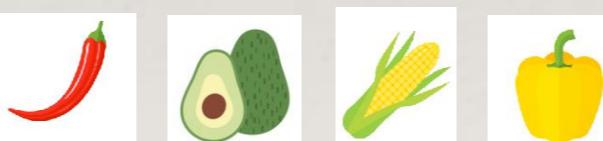


Mexicana

Teste



Mexicana



Mexicana



Italiana



Mexicana



Italiana

Um (mal) exemplo

- ❖ Tarefa: eliminar o ENEM, selecionar estudantes à partir de informações pré-existentes
- ❖ Base de dados: cadastro de estudantes universitários existentes
- ❖ Qual o padrão existente?

Um (mal) exemplo

- ❖ Homens, brancos, classe média...
- ❖ Pelo algoritmo, quem é selecionado? Homens, brancos, classe média...

Métricas de performance do algoritmo

- ❖ Matriz de confusão
- ❖ Verdadeiros positivos, falsos positivos, verdadeiros negativos, faltos positivos
- ❖ Sensibilidade, especificidade
- ❖ Acurácia
- ❖ F1

Matriz de confusão

		Real	
		Positivo	Negativo
Predito	Positivo	Verdadeiro positivo	Falso positivo
	Negativo	Falso negativo	Verdadeiro negativo

Matriz de confusão

	P	N
P	VP	FP
N	FN	VN

Real



Acurácia

	P	N
P	VP	FP
N	FN	VN

$$\frac{VP + VN}{VP + VN + FP + FN}$$

	P	N
P	40	10
N	10	40

$$= \frac{40 + 40}{40 + 40 + 10 + 10} = \frac{80}{100} = 0,80$$

Prática com R

Não se aborreça com um erro, ele pode te lembrar uma música...

```
> as.numeric(nodup>All$Start.Num)
```

```
[1] NA  
NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA  
[47] NA  
NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA  
[93] NA  
NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA  
[139] NA  
NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA  
[185] NA NA NA NA
```

Warning message:

NAs introduced by coercion