

IA - 2020.1^{SI} LCC



Prof. Rafael Magalhães

Aprendizagem Supervisionada Classificação



Conteúdo

- Conceitos Teóricos
- Avaliação de Modelos

Conteúdo Programático

Apresentação e ritos iniciais

1. Fundamentos e Métodos

2. Tópicos Avançados

3. Projetos

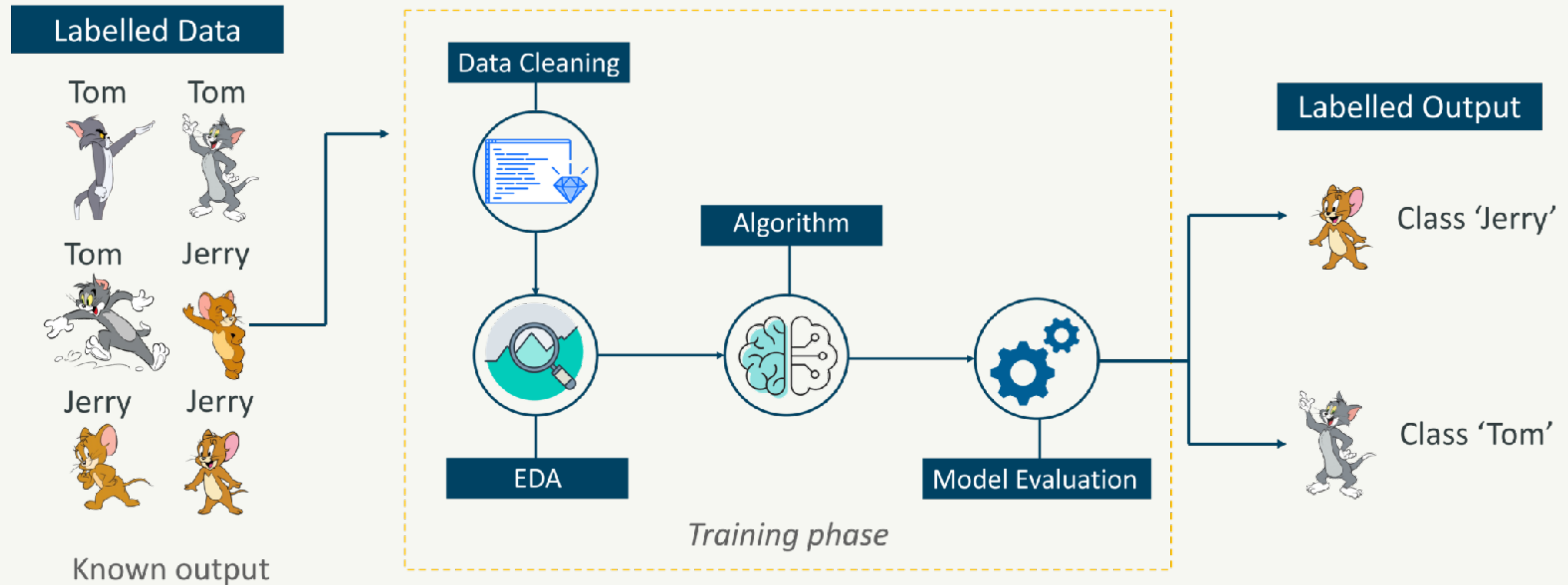
Fechamento e ritos finais



1. Fundamentos e Métodos

- Fundamentos de IA
- Métodos Supervisionados
Classificação
- Métodos Supervisionados
Regressão
- Métodos Não Supervisionados
- Redes Neurais Artificiais

Métodos Supervisionados Classificação



O que é IA?

... a habilidade para **raciocinar, planejar, resolver problemas, pensar abstratamente, compreender ideias complexas, aprender a partir da experiência.**

Métodos Supervisionados

Classificação



Métodos Supervisionados

Classificação



Ideia principal (Intuition):

Classificação == ‘Colocar coisas em uma caixa’

Saídas (Targets): Apenas Categorias

Entradas (Input): Valores numéricos ou categóricos

Métodos Supervisionados

Classificação

Ideia principal (Intuition):

Criar ‘caixas de categorias’ utilizando dados de treinamento:

predict

$$f(x) = ax + b$$

algorithm

→ (a) (b)

fit (X, Y)

$f(x, y)$

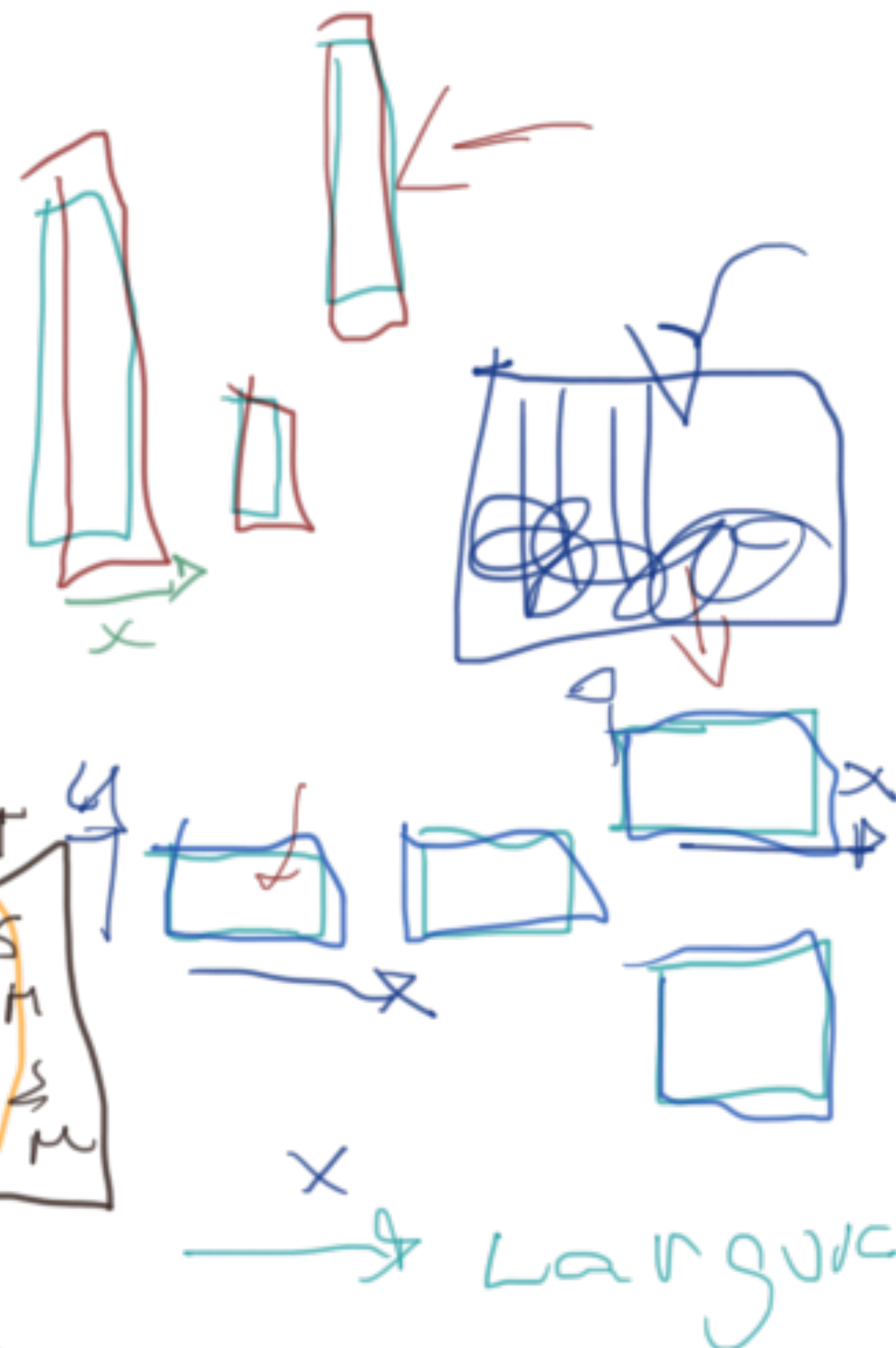


(a) = 0

(b) = 0

$a = 0.5$ (fit) 0.3

$b = 0.5$ 1



y ↑ altura

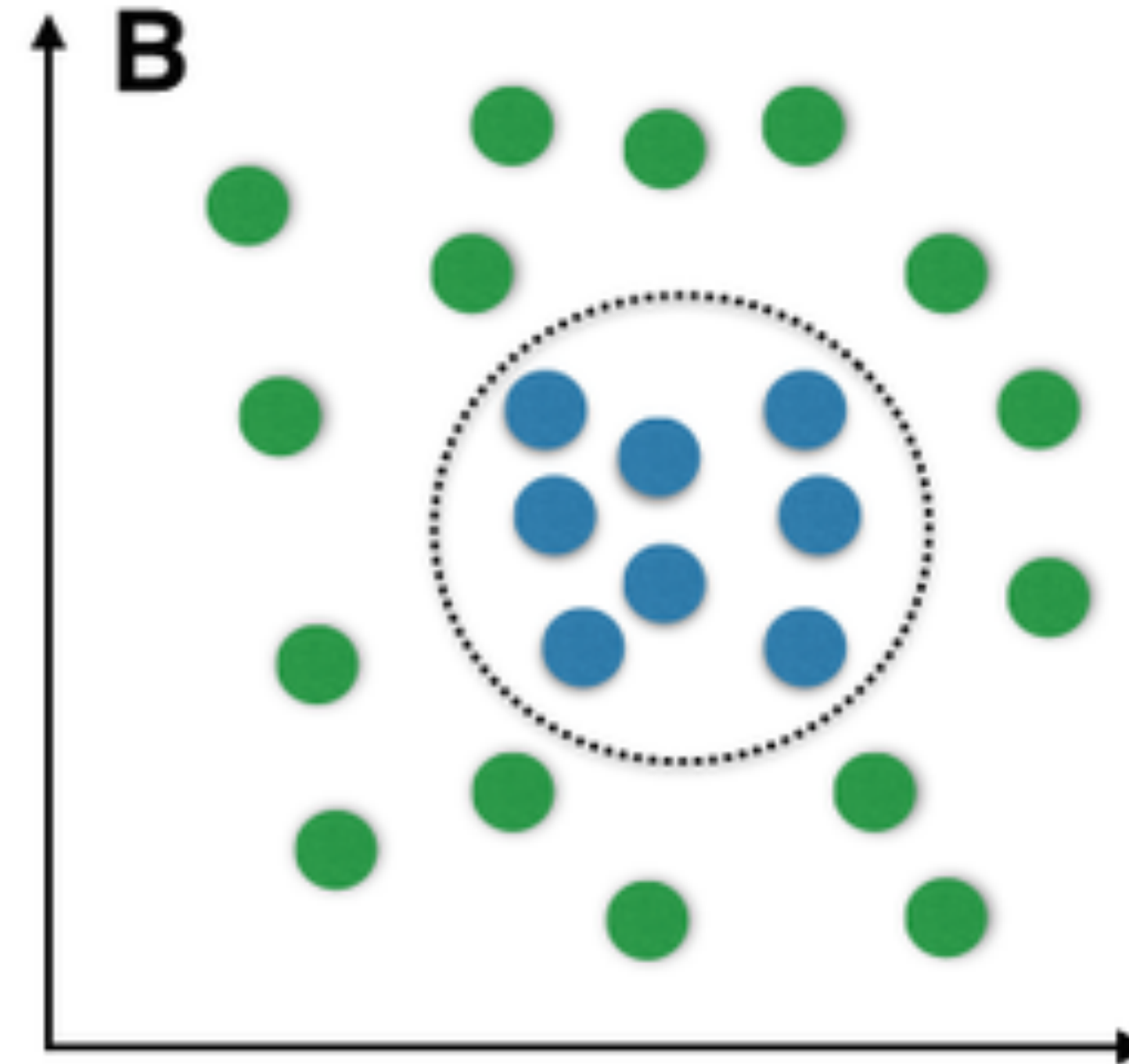
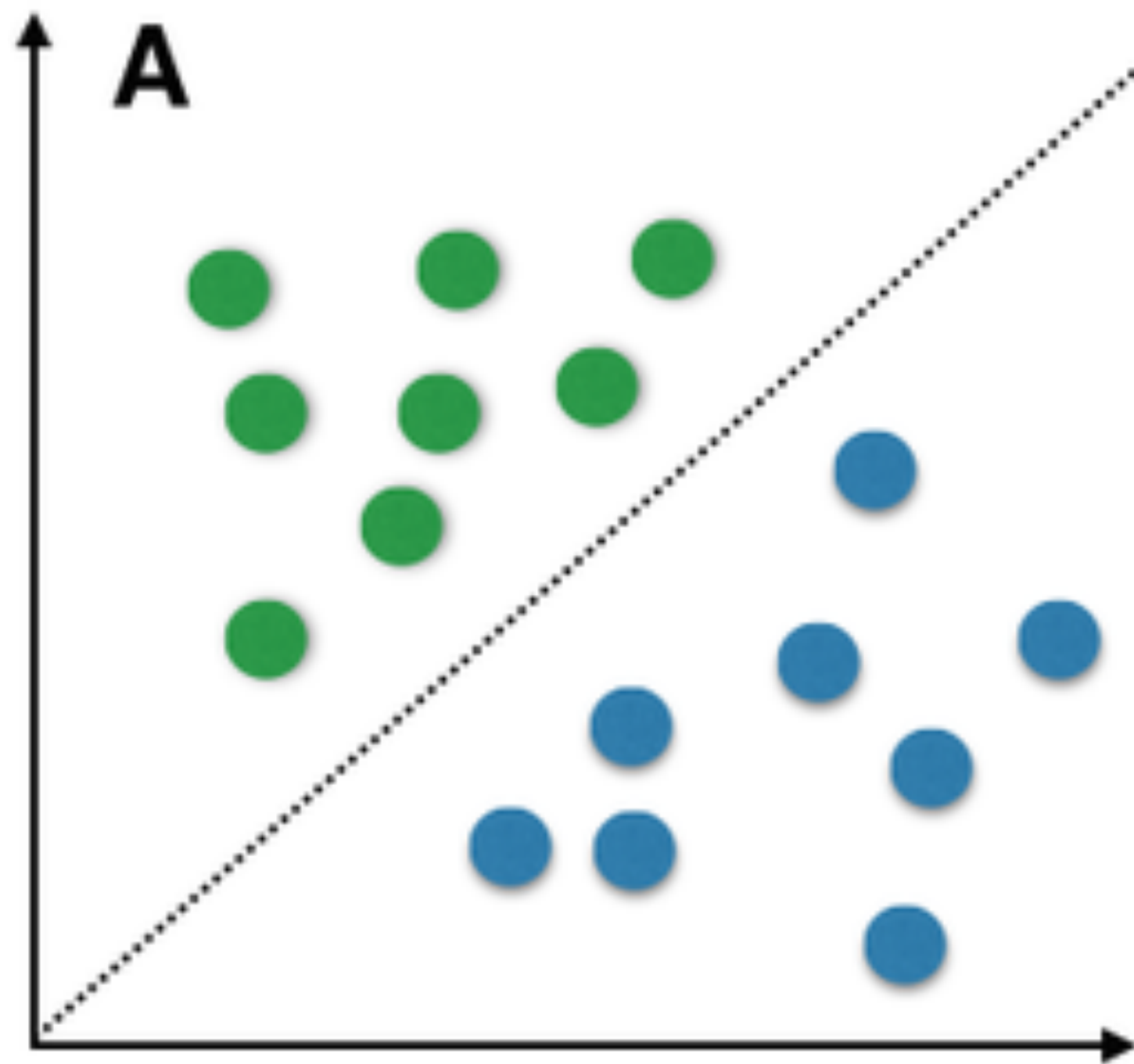
x → largura

Métodos Supervisionados

Classificação

Ideia principal (Intuition):

Criar 'caixas de categorias' utilizando dados de treinamento:



Revendo o Problema Íris

- Variáveis ou Atributos do Preditor / Variáveis Objetivo
- Objetivo:
Predizer a variável objetivo (classe), dado as variáveis atributos
- Classificação:
a variável objetivo é categórica
- Regressão:
a variável objetivo é contínua

Dados de Entrada [X]				Categoria
CS	LS	CP	LP	
5.0	2.0	3.5	1.0	Iris-versicolor
6.0	2.2	4.0	1.0	Iris-versicolor
6.2	2.2	4.5	1.5	Iris-versicolor
6.0	2.2	5.0	1.5	Iris-virginica
4.5	2.3	1.3	0.3	Iris-setosa
5.0	2.3	3.3	1.0	Iris-versicolor
5.5	2.3	4.0	1.3	Iris-versicolor
6.7	2.5	5.8	1.8	Iris-virginica
4.9	3.1	1.5	0.1	Iris-setosa
4.6	3.1	1.5	0.2	Iris-setosa
				Dados de Saída [y]

Revendo o Problema Íris

Convenções de NOMES

- Feature
 - = Atributos
 - = Variáveis de predição
 - = Variáveis independentes
- Target variable
 - = Variáveis dependentes
 - = Variável de resposta

Dados de Entrada [X]				Categoria
CS	LS	CP	LP	
5.0	2.0	3.5	1.0	Iris-versicolor
6.0	2.2	4.0	1.0	Iris-versicolor
6.2	2.2	4.5	1.5	Iris-versicolor
6.0	2.2	5.0	1.5	Iris-virginica
4.5	2.3	1.3	0.3	Iris-setosa
5.0	2.3	3.3	1.0	Iris-versicolor
5.5	2.3	4.0	1.3	Iris-versicolor
6.7	2.5	5.8	1.8	Iris-virginica
4.9	3.1	1.5	0.1	Iris-setosa
4.6	3.1	1.5	0.2	Iris-setosa
				Dados de Saída [y]

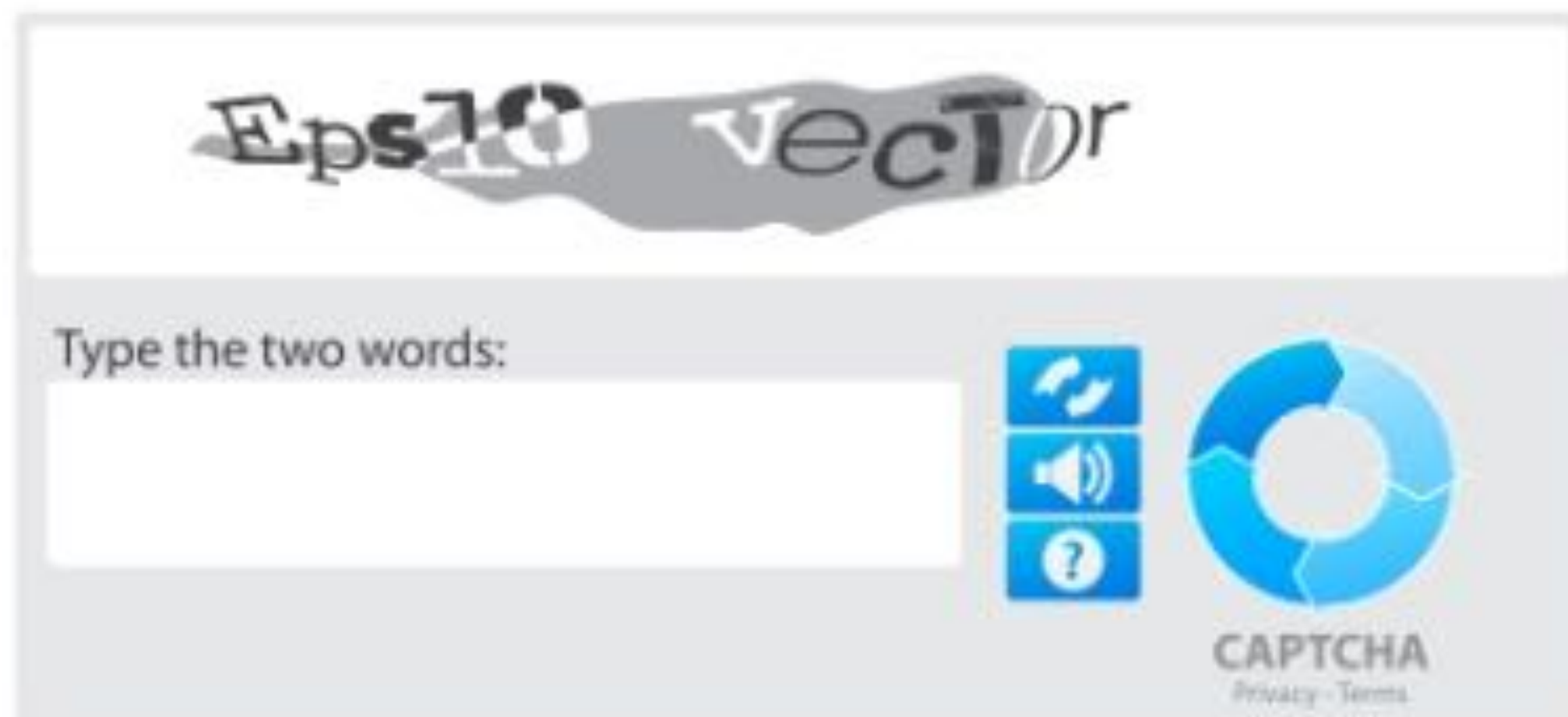
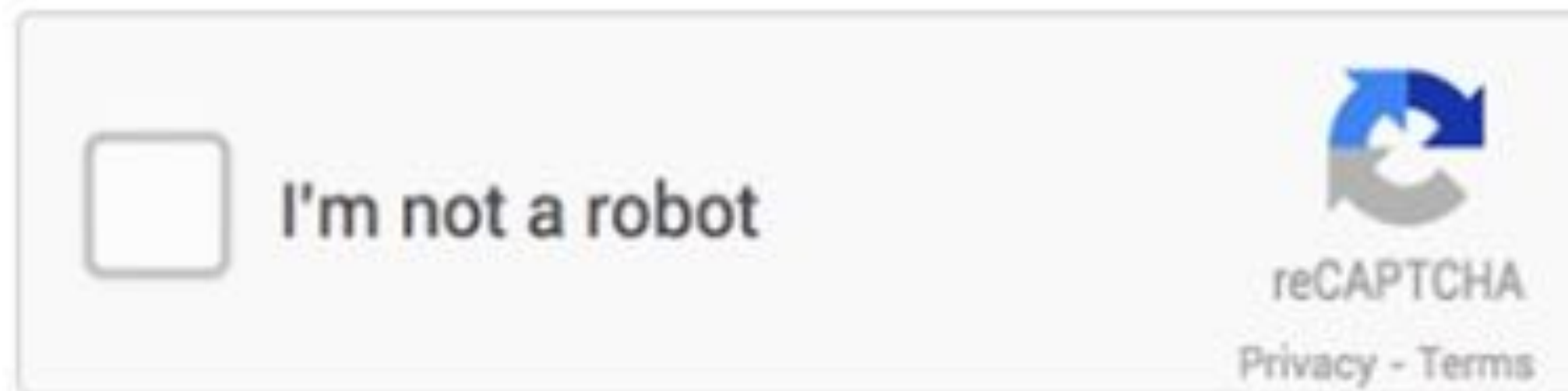
Métodos Supervisionados

Classificação

- Automatizar tarefas que demandam muito tempo ou tarefas que demandam atividade manual
 - ▶ Exemplo: Diagnóstico médico,
- Fazer previsões sobre algo futuro
 - ▶ Exemplo: Um cliente irá clicar na propaganda ou não?
- É necessário dados já etiquetados (labeled data)
 - ▶ Dados históricos já classificados
 - ▶ Experimentos para criar um dataset e as classificações
 - ▶ ‘Crowd-sourcing’ para criar dados categorizados

Métodos Supervisionados

Classificação



Revendo o Problema Íris

Ver no Colab

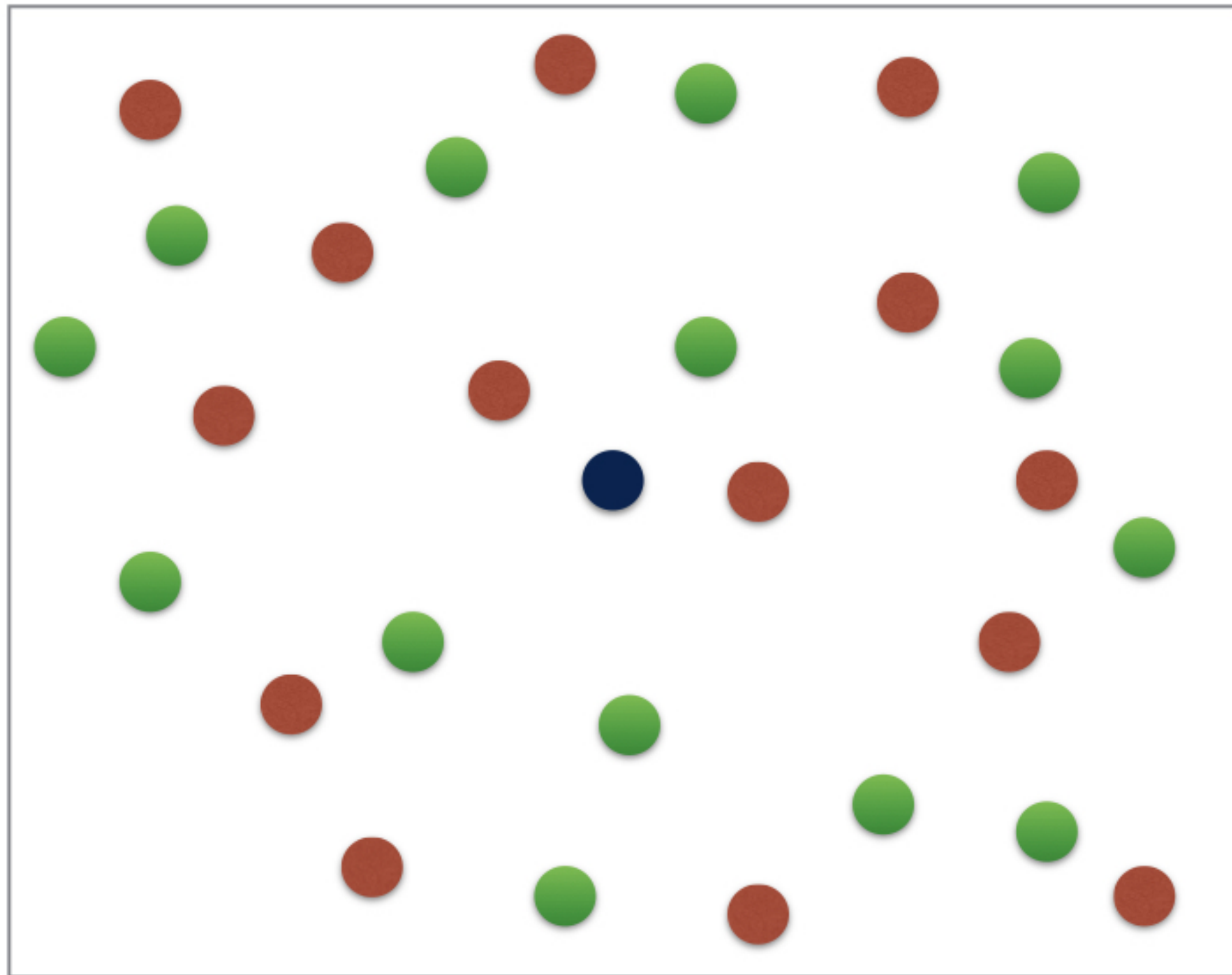
Dados de Entrada [X]					
CS	LS	CP	LP	Categoria	
5.0	2.0	3.5	1.0		Iris-versicolor
6.0	2.2	4.0	1.0		Iris-versicolor
6.2	2.2	4.5	1.5		Iris-versicolor
6.0	2.2	5.0	1.5		Iris-virginica
4.5	2.3	1.3	0.3		Iris-setosa
5.0	2.3	3.3	1.0		Iris-versicolor
5.5	2.3	4.0	1.3		Iris-versicolor
6.7	2.5	5.8	1.8		Iris-virginica
4.9	3.1	1.5	0.1		Iris-setosa
4.6	3.1	1.5	0.2		Iris-setosa
				Dados de Saída [y]	

Classificador K-Nearest Neighbors

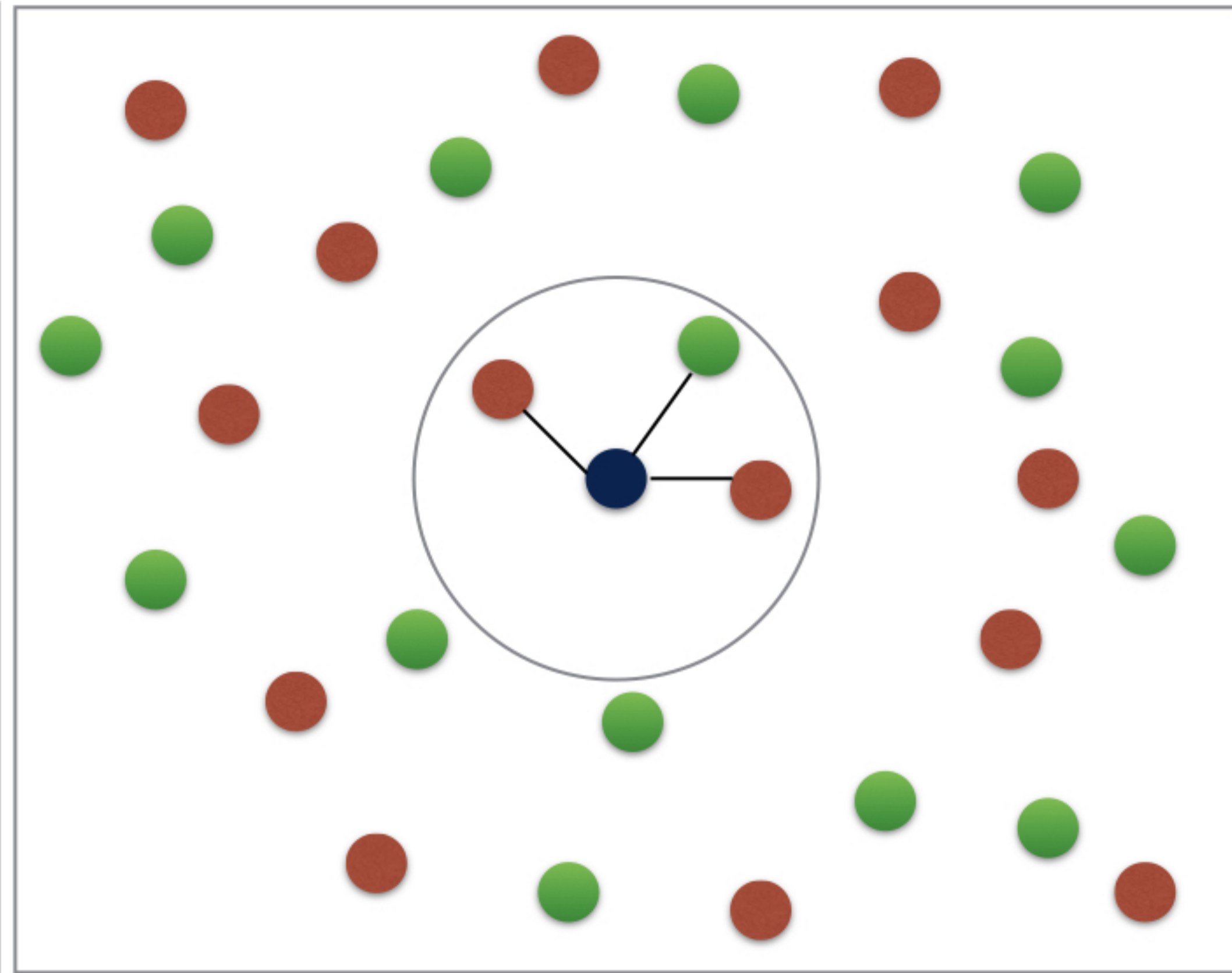
Predizer a categoria de um ponto de dado, como:

- Observando os 'K' vizinhos mais próximos desse ponto
- Aceitando a votação da maioria

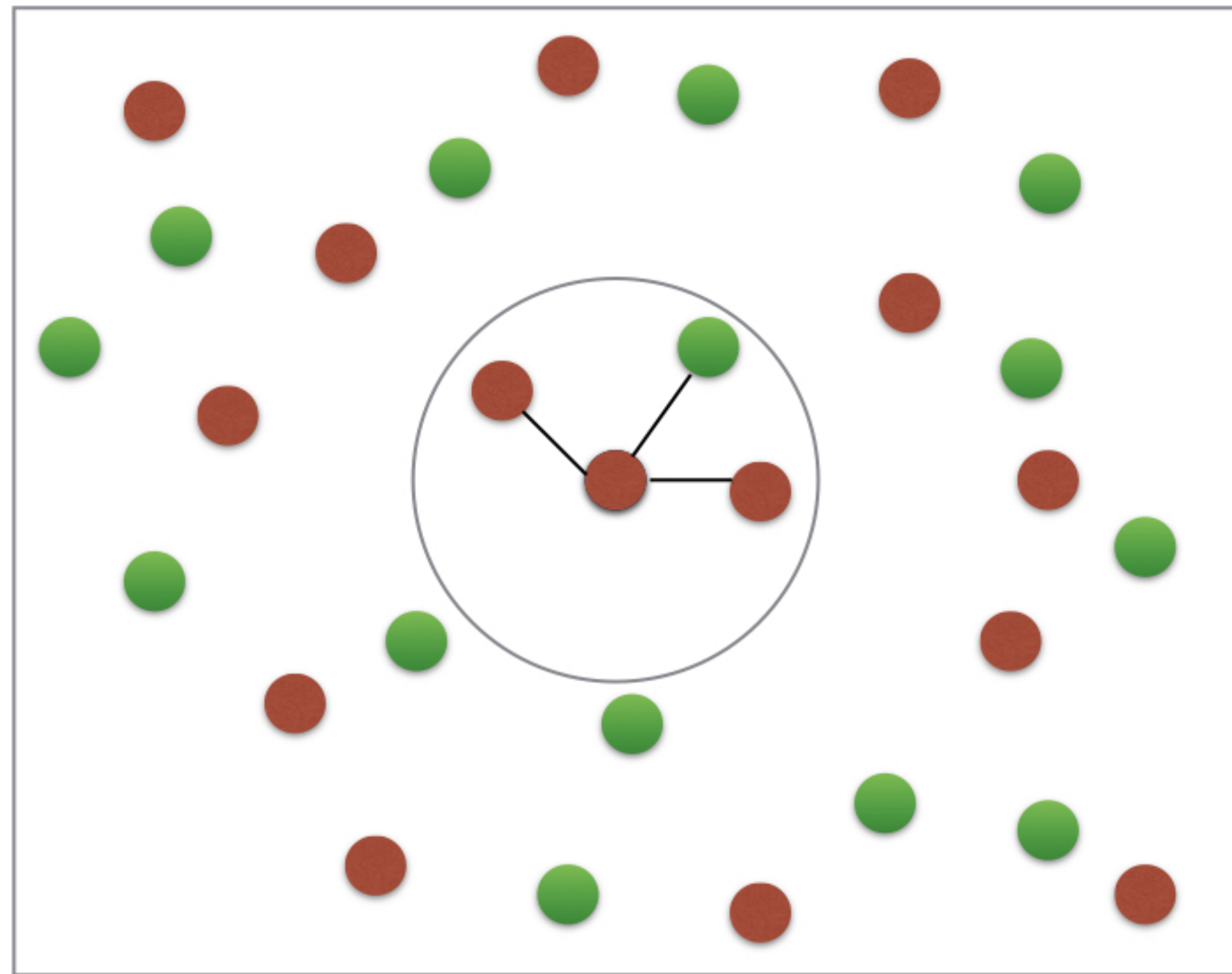
Classificador K-Nearest Neighbors



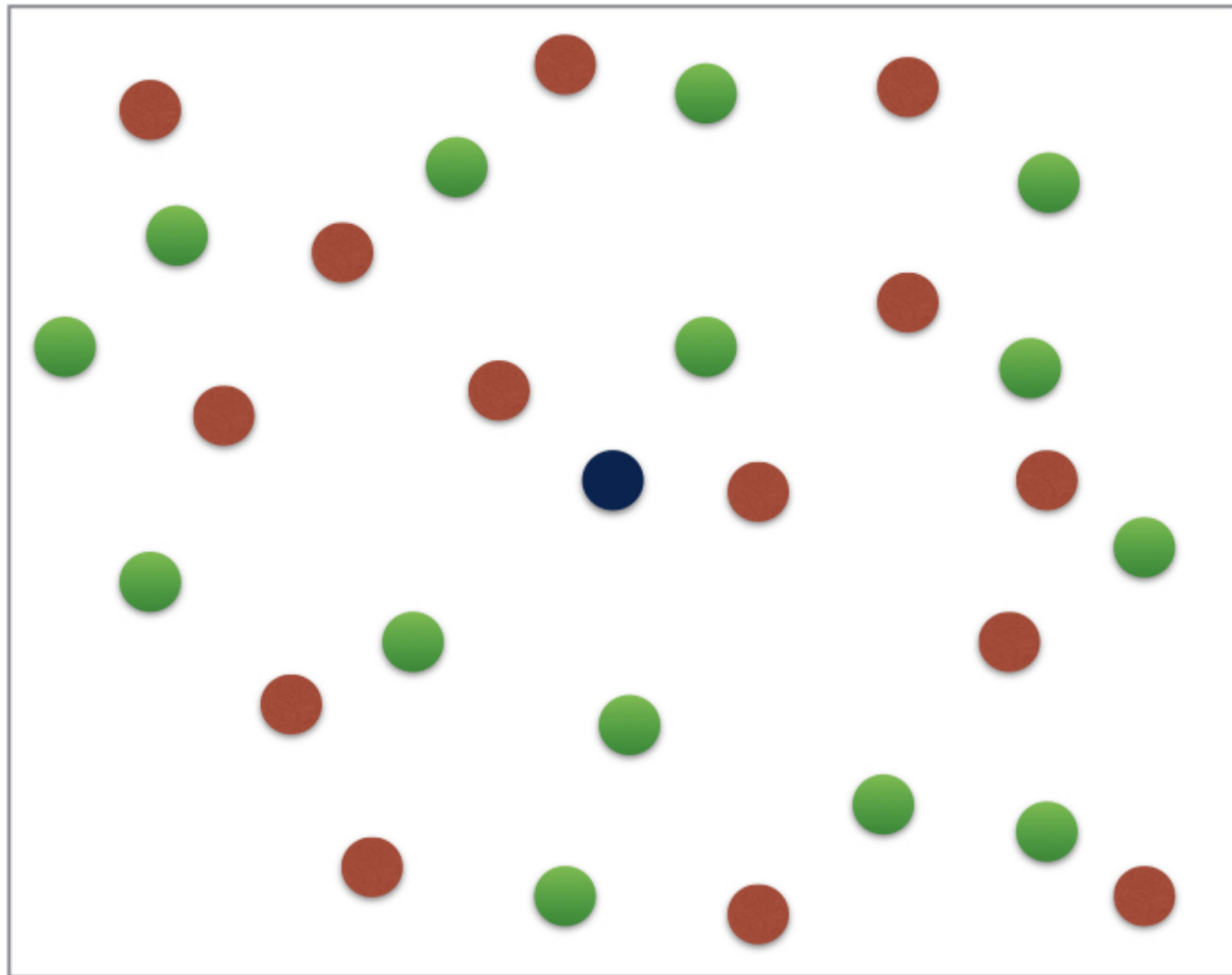
Classificador K-Nearest Neighbors



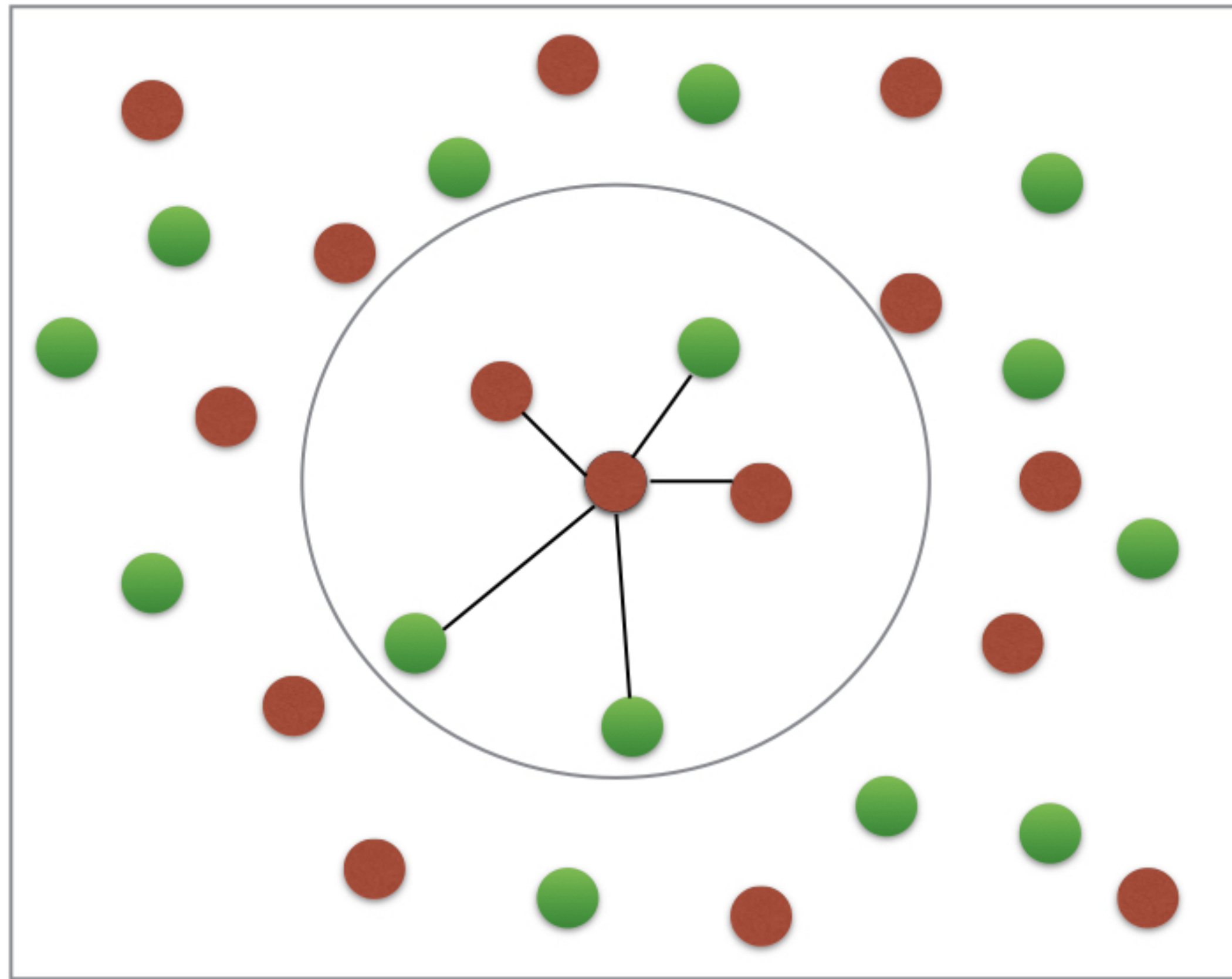
Classificador K-Nearest Neighbors



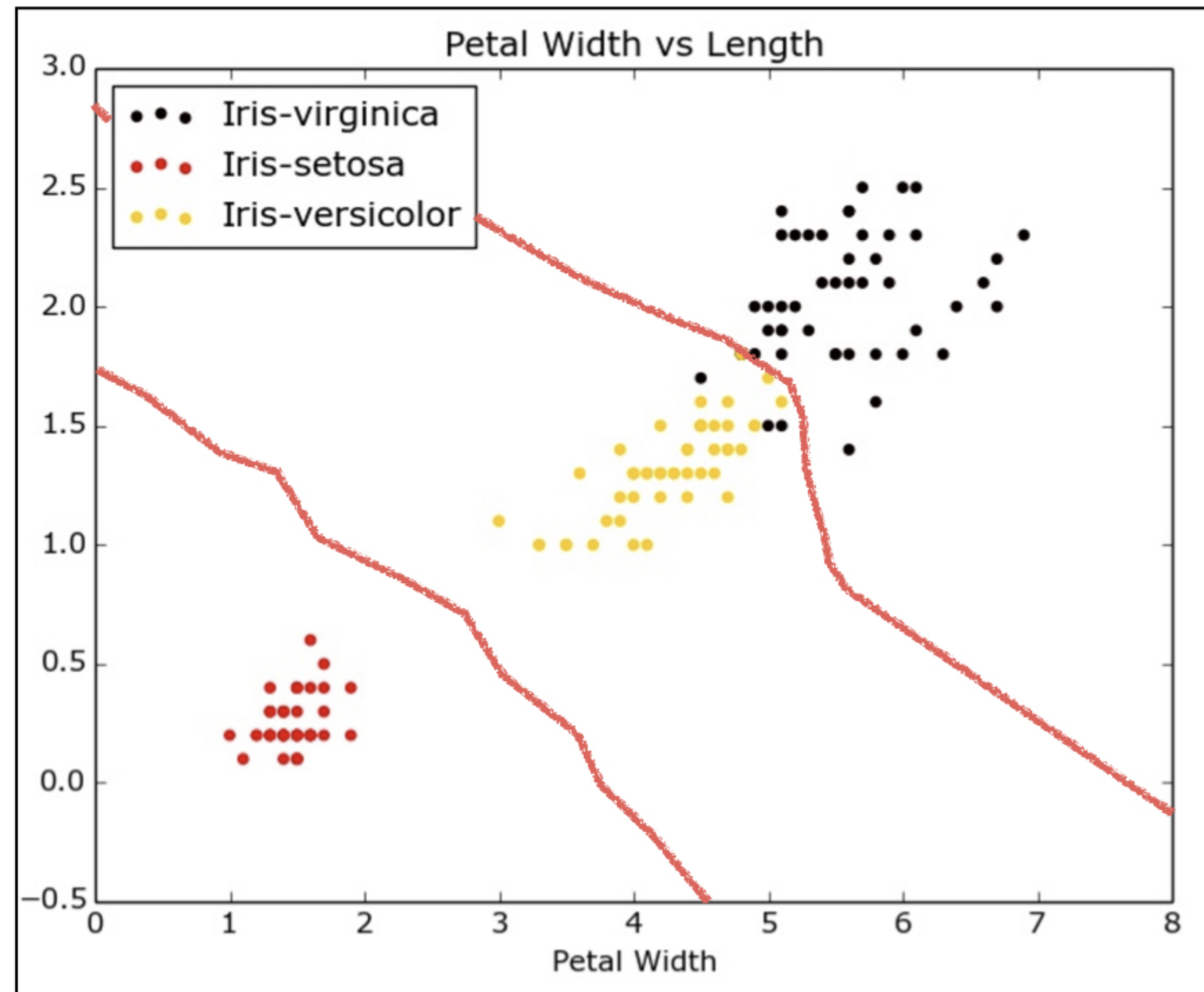
Classificador K-Nearest Neighbors



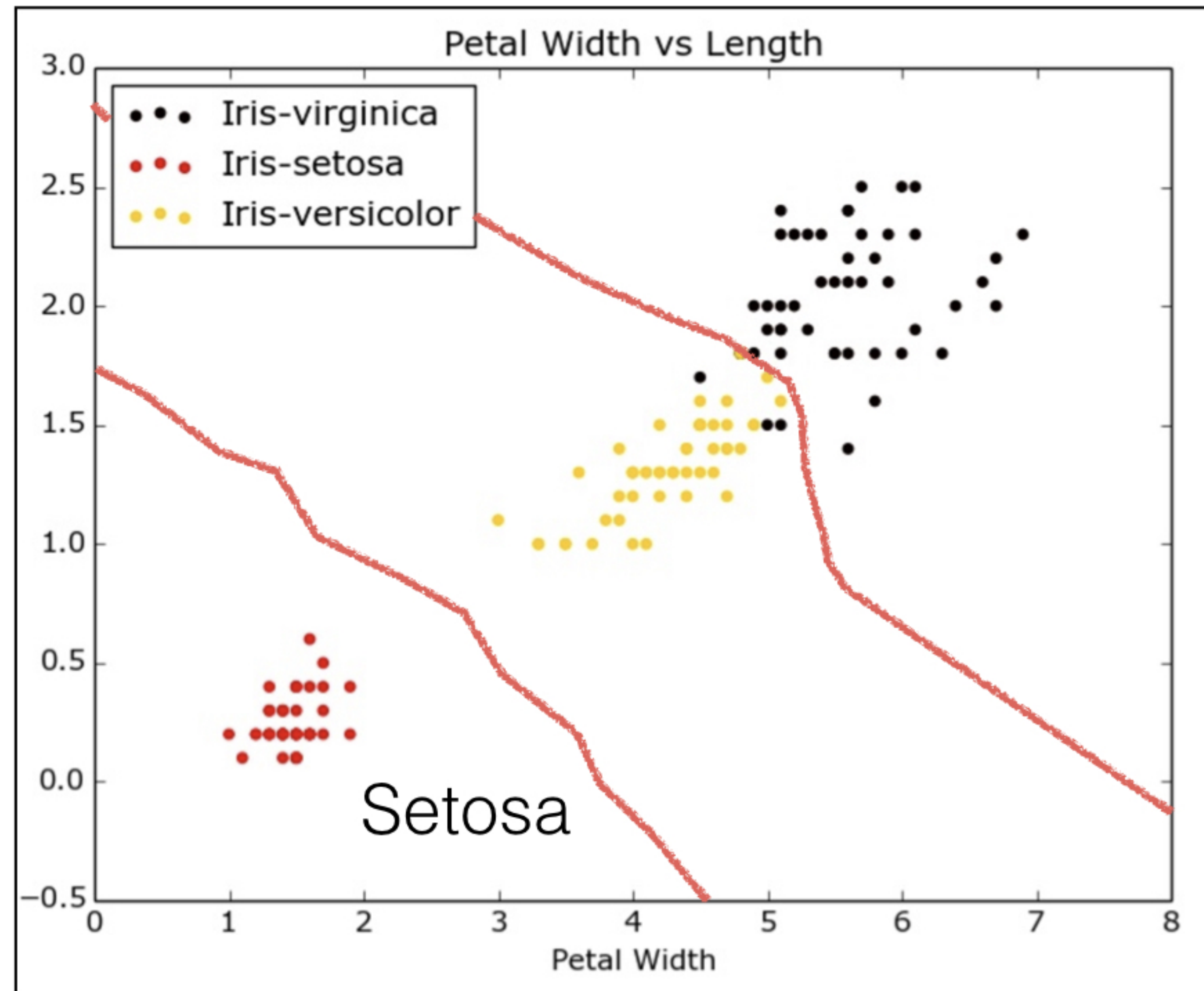
Classificador K-Nearest Neighbors



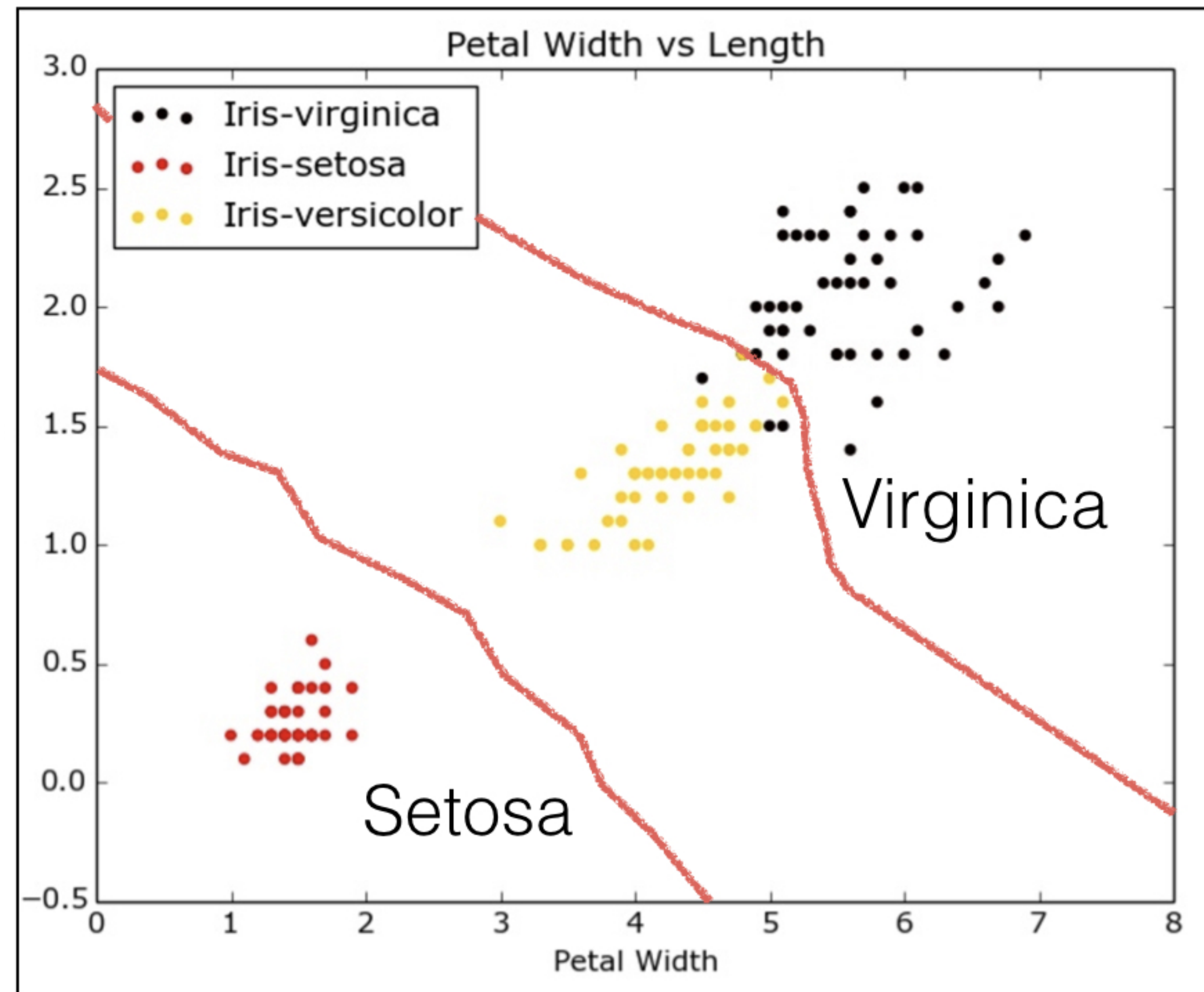
Classificador K-Nearest Neighbors



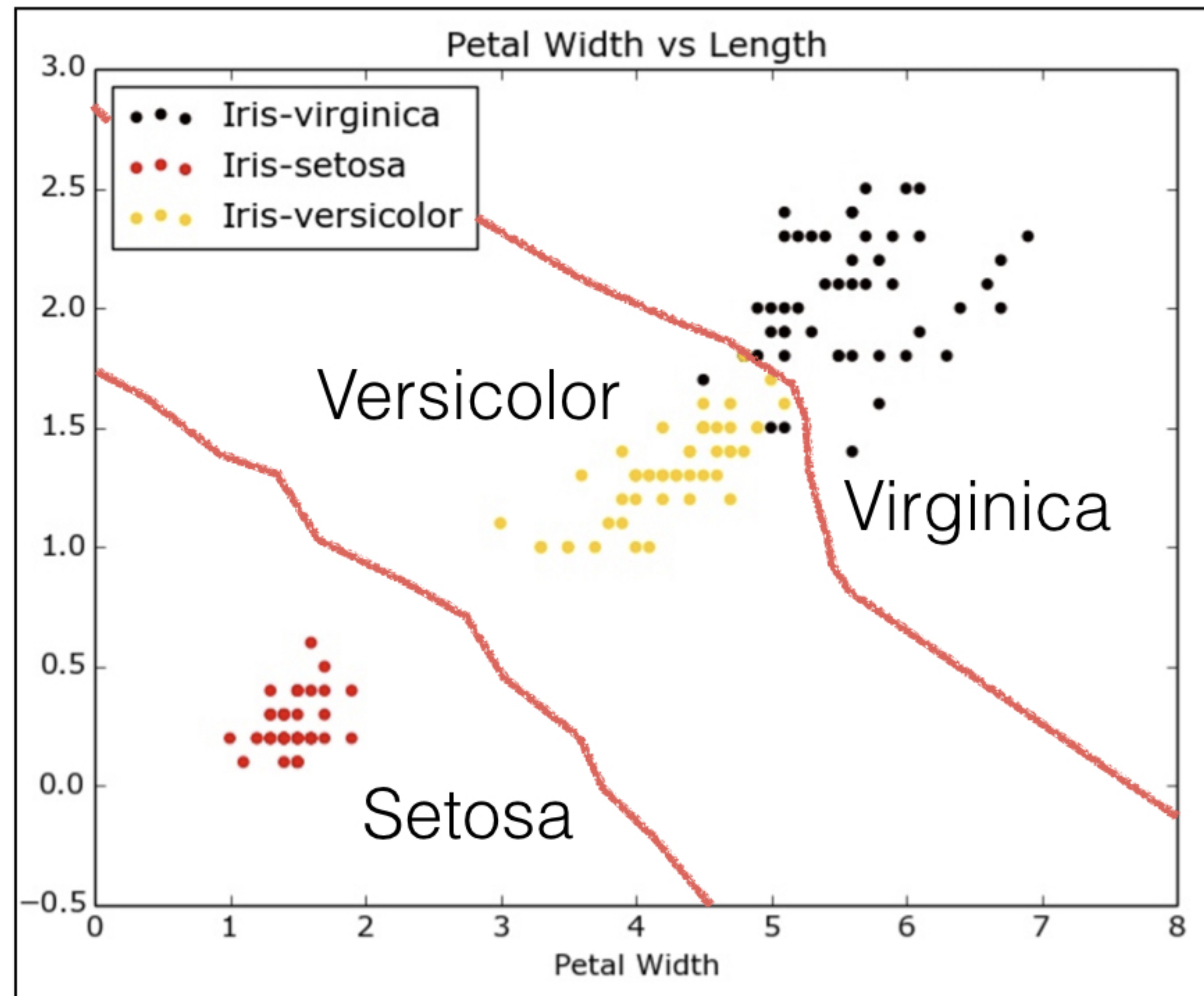
Classificador K-Nearest Neighbors



Classificador K-Nearest Neighbors



Classificador K-Nearest Neighbors



Avaliando o Modelo

E se o problema for avaliar diabetes:

	Tem Diabetes	Não Tem
Prediz Diabetes	True Positives	
Prediz Não Tem		

Avaliando o Modelo

E se o problema for avaliar diabetes:

	Tem Diabetes	Não Tem
Prediz Diabetes	True Positives	
Prediz Não Tem		True Negatives

Avaliando o Modelo

E se o problema for avaliar diabetes:

	Tem Diabetes	Não Tem
Prediz Diabetes	True Positives	False Positives
Prediz Não Tem		True Negatives

Avaliando o Modelo

E se o problema for avaliar diabetes:

	Tem Diabetes	Não Tem
Prediz Diabetes	True Positives	False Positives
Prediz Não Tem	False Negatives	True Negatives

Avaliando o Modelo

E se o problema for avaliar diabetes:

	Tem Diabetes	Não Tem
Prediz Diabetes	True Positives	False Positives
Prediz Não Tem	False Negatives	True Negatives

Erro Tipo I



Erro Tipo I



Avaliando o Modelo

E se o problema for avaliar diabetes:

	Tem Diabetes	Não Tem
Prediz Diabetes	True Positives	False Positives
Prediz Não Tem	False Negatives	True Negatives

Acurácia (Accuracy) : “Quantas vezes eu acerto a predição True e False”

Precisão (Precision) : “Quantas vezes eu acerto corretamente quando digo que alguém tem diabetes?”

Recall: “Qual o percentual de verdadeiros casos eu consegui detectar?”

Dúvidas ???

IA - 2020.1^{SI} LCC



Prof. Rafael Magalhães

Aprendizagem Supervisionada Classificação

