

Um evento com propósito

Embarcadero Conference 2020 Online

todos conectados



Um evento com propósito
Embarcadero Conference
2020 Online

Machine Learning em Delphi

Apresentação do algoritmo K-NN

Rodolpho Nascimento

Sumário

- Introdução
- Definição de Machine Learning
- Tipos de aprendizado em ML
- O algoritmo K-NN
- Exemplo prático
- Encerramento

Sobre

Rodolpho Nascimento

Rio de Janeiro-RJ

- Mestrando em Ciência da Computação - CEFET/RJ
 - Linha de pesquisa: Computação Afetiva
 - Especialidade: Mineração de Textos
- Formação: Análise e desenvolvimento de sistemas (UNESA)
- Trabalhando com Delphi desde a versão 5.
- Coordenador de Produtos na GKO Informática (RJ)
- Além de Delphi, desenvolvo programas em outras linguagens: C#, Javascript, Dart e Python

Introdução

Introdução

- Objetivo desta palestra é aplicar uma técnica simples de ML, com o código puramente escrito em Delphi
- O algoritmo abordado, será o K-NN
- De forma simplista alguns conceitos teóricos serão abordados
- Nesta apresentação será demonstrada um exemplo aplicação do K-NN em sistemas de concessão de crédito



Definição de Machine Learning

Definição de Machine Learning

Definição de Machine Learning (em português, Aprendizado de Máquina):

“Um campo de estudo que capacita os computadores para a habilidade de aprender sem serem explicitamente programados”

Arthur Samuel, “Some studies in machine learning using the game of checkers. IBM J Res Dev . 1959;3:535–554”

Definição de Machine Learning

- ML não é algo novo. Alan Turing foi um dos pioneiros dos estudos de ML publicando em **1950** o artigo COMPUTING MACHINERY AND INTELLIGENCE, com o famoso “O jogo da imitação” (*the imitation game*).
- O primeiro modelo de neurônio artificial foi proposto pelos pesquisadores Warren McCulloch e Walter Pitts, em **1943**, no artigo *A Logical Calculus of the Ideas Immanent in Nervous Activity*.
- O atual momento (grande quantidade de dados + grande poder de processamento) contribuiu para a popularização da aplicação de ML.

Definição de Machine Learning

Frase para reflexão:



**“Os dados hoje em dia,
representam o novo petróleo”**

Kai-fu Lee, no livro Inteligência Artificial, 2019

Tipos de aprendizado em ML

Tipos de aprendizado em ML

Em ML existem alguns tipos de aprendizado, sendo os mais utilizados:

- **Aprendizado não-supervisionado:** Quando se fornece apenas o exemplo ao algoritmo.

Ex: K-means, DBSCAN.

- **Aprendizado supervisionado:** Quando se fornece o exemplo e as respostas para o algoritmo.

Ex: **K-NN**, Naive Bayes, SVM, Decision Tree

Tipos de aprendizado em ML

Exemplo de um conjunto de dados voltado para aprendizado supervisionado

Tem pelo longo?	Come peixe?	Come osso?	Late ou mia?	Resposta
Sim	Não	Sim	Late	Cachorro
Não	Sim	Não	Mia	Gato
Não	Não	Sim	Late	Cachorro
Sim	Sim	Não	Mia	Gato

Esta etapa é muito importante para o aprendizado supervisionado!

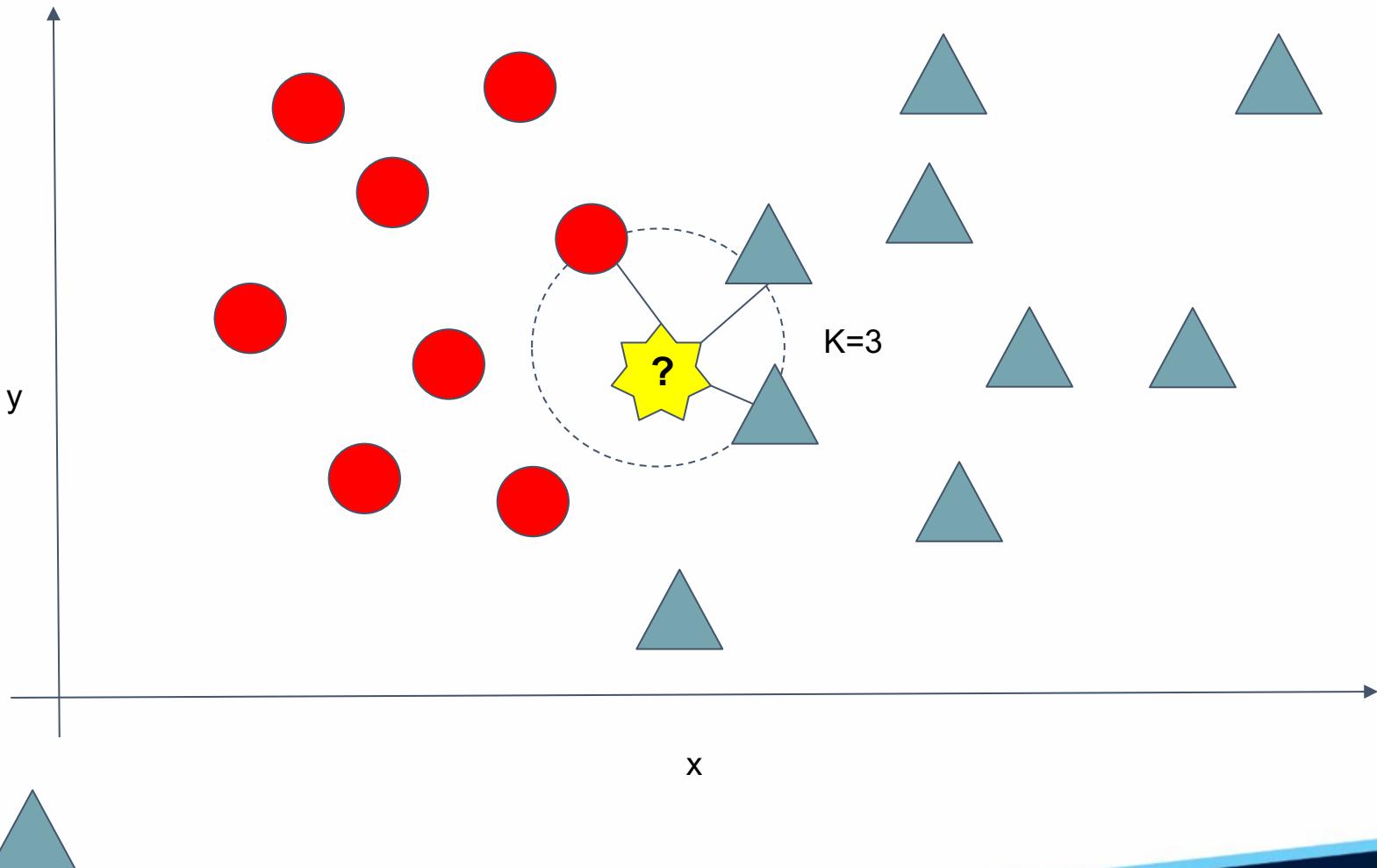


O algoritmo K-NN

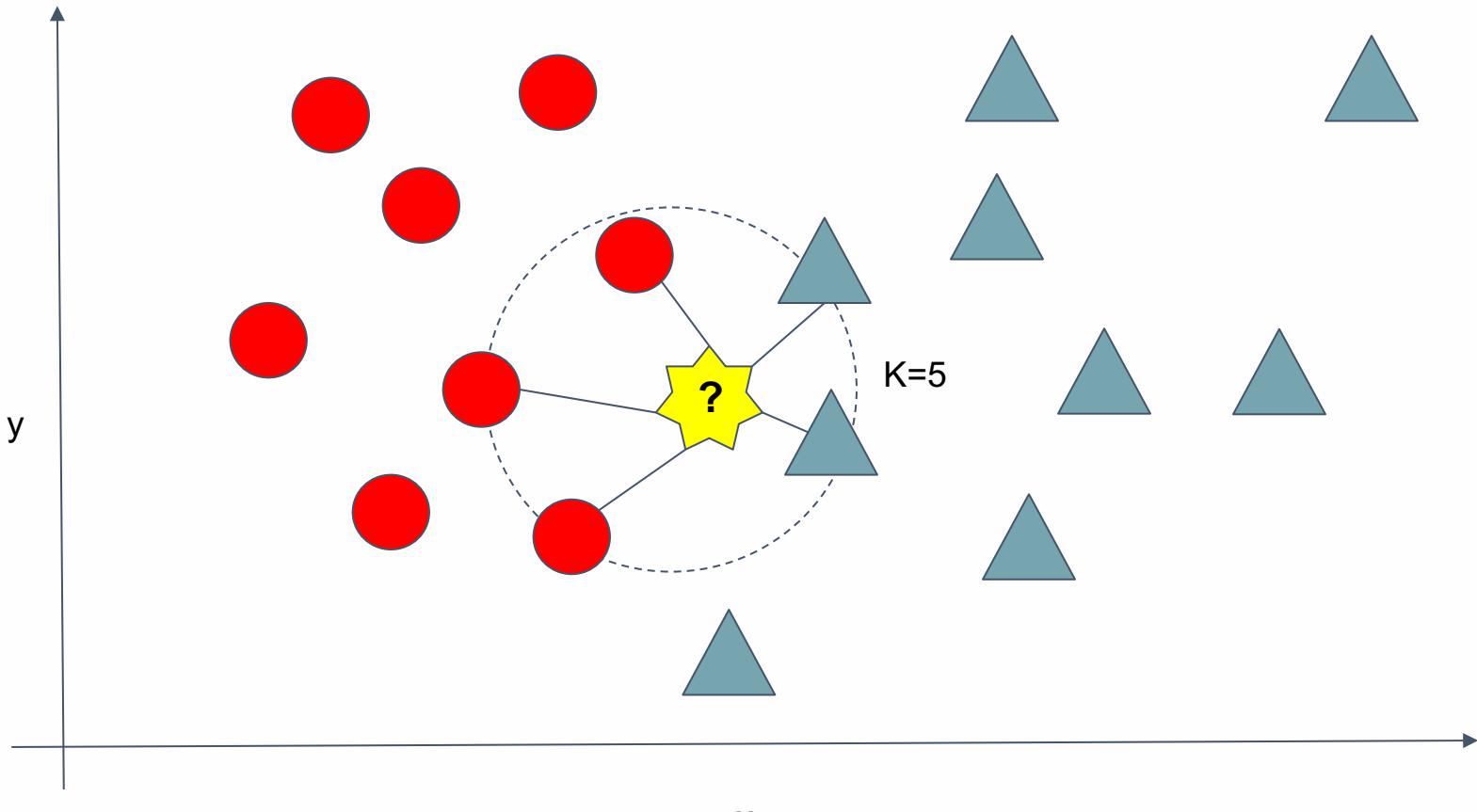
O algoritmo K-NN

- O K-NN (k -nearest neighbors) é um algoritmo de classificação baseado no conceito de “análise de vizinhança” ou seja, tem como fundamento analisar características dos seus vizinhos mais próximos
- A letra “K” indica o número de vizinhos em que ele irá analisar
- Proposto por T. Cover e P. Hart, no artigo "*Nearest neighbor pattern classification,*" em IEEE Transactions on Information Theory, vol. 13, no. 1, pp. 21-27, January 1967.
- Considerado um algoritmo “preguiçoso”, pois não realiza treinamento
- Dado uma nova entrada, o algoritmo tenta predizer o resultado com base nos vizinhos mais próximos desta nova entrada

O algoritmo K-NN



O algoritmo K-NN

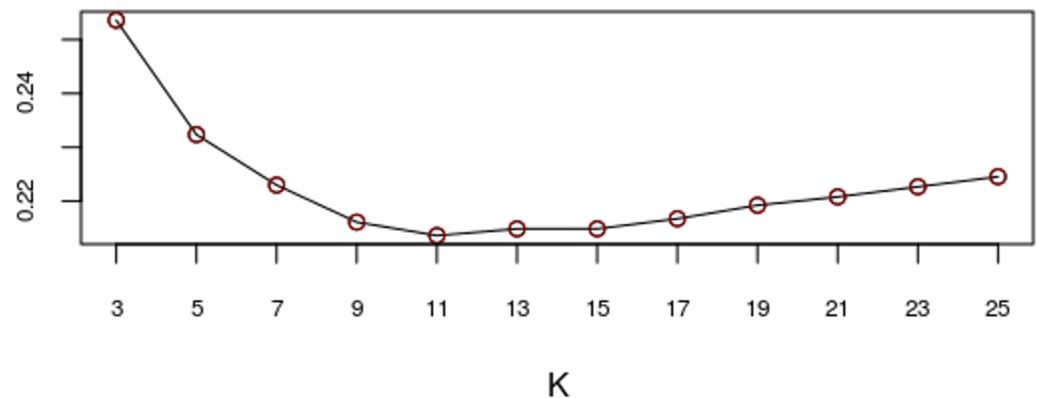


Resposta:

O algoritmo K-NN

Pergunta: Qual é o melhor valor de K?

- Não há uma regra
- Geralmente, o valor de K é dado de forma empírica
- Em alguns artigos, o valor de K é definido como 30% dos dados
- Há também a análise da curva de erro, selecionando o K de menor erro.



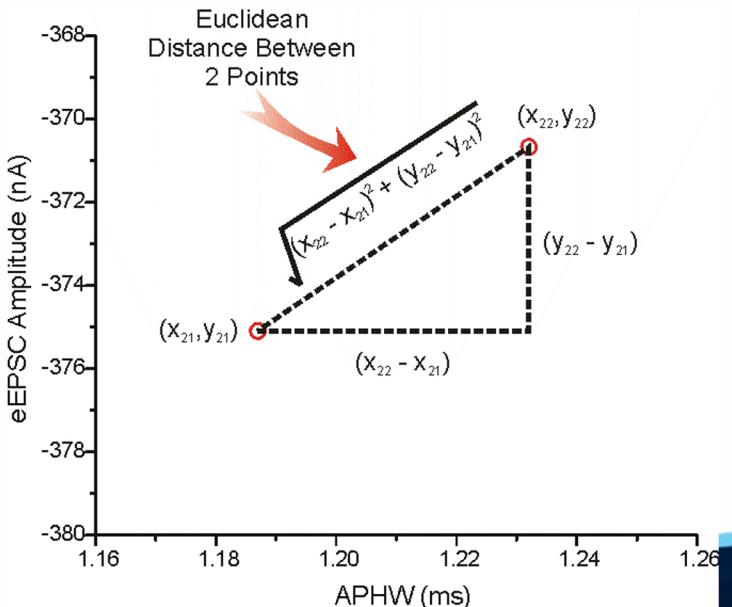
O algoritmo K-NN

Métricas para cálculo de distância entre os pontos

- Existem várias formas de calcular a distâncias entre os pontos vizinhos (manhattan, minkowski, hamming), porém a mais utilizada é a **distância euclidiana**, dada pela notação matemática abaixo:

$$d(x, y) = \sqrt{\sum_{i=1}^n (x_i - y_i)^2}$$

x = Característica do dado de exemplo
y = Característica do novo dado



O algoritmo K-NN

Funcionamento do algoritmo K-NN:

1. Recebe a nova entrada
1. Calcula a distância desta nova entrada com todos os exemplos
1. Ordena os resultados na ordem crescente
1. Seleciona os “top K” resultados
1. Conta os elementos de cada classes
1. Retorna a classe com mais elementos (maior similaridade) à nova entrada.

O algoritmo K-NN

Vantagens

- Simples de ser implementado, e fácil de compreender seu resultado
- Pode ser utilizado tanto e classificação como regressão

Desvantagens

- Alto custo computacional para realizar suas previsões
- Grandes quantidades de dados podem onerar consideravelmente no tempo para previsão

Exemplo prático

Exemplo prático

Sistema de concessão de crédito

- Seu objetivo é determinar a concessão ou negação de crédito com base em características entre bons e mau pagadores.
- Características que serão utilizadas:
 - Idade
 - Tempo de contratação na empresa
 - Restrição de crédito
 - Salário
 - Se costuma pagar suas contas em dia

Go to -> Código fonte!!!

Outros exemplos de aplicações do K-NN

Sistemas de recomendação

Vendas realizadas

Cliente	Produto	Produto	Produto	Produto
João	Notebook	Mouse	Impressora	Mochila
Antônio	Notebook	Mouse	-	-
José	Notebook	Mouse	Impressora	
Diego	Pendrive	Mouse	-	-

Novas vendas

Cliente	Carrinho	Sugestões de compra
Novo cliente A	Notebook, mouse	Impressora, mochila
Novo cliente B	Notebook, mouse, impressora	Mochila

Outros exemplos de aplicações do K-NN

Medicina

- Predizer o risco de um paciente hospitalizado por ataque cardíaco, ter o segundo ataque através da análise dados demográficos, histórico do paciente, hábito alimentar, etc...
- Classificar o fator de risco de câncer de próstata, através da análise de dados demográficos e variáveis clínicas.



Encerramento

Encerramento

Github do projeto: https://github.com/rodolphonascimento/Delphi_KNN

Para facilitar, pesquisar no Google: **delphi_knn**

Contatos:

 nascimento@gko.com.br

 <https://www.linkedin.com/in/rodolphonascimento/>



Perguntas?

OBRIGADO