

Vitor Amadeu Souza

Aprendizado de

Máquina

utilizando

Agglomerative Clustering

Programado em Python

© 2024 by Cerne Tecnologia e Treinamento Ltda.

© 2024 by Vitor Amadeu Souza

Nenhuma parte desta publicação poderá ser reproduzida sem autorização prévia e escrita de **Cerne Tecnologia e Treinamento Ltda.** Este livro publica nomes comerciais e marcas registradas de produtos pertencentes a diversas companhias. O editor utiliza as marcas somente para fins editoriais e em benefício dos proprietários das marcas, sem nenhuma intenção de atingir seus direitos.

Setembro de 2024

Direitos reservados por:

Cerne Tecnologia e Treinamento Ltda

Produção: Cerne Tecnologia e Treinamento

E-mail da Empresa: cerne@cerne-tec.com.br

Home Page: www.cerne-tec.com.br.com.br

Atendimento ao Consumidor: sac@cerne-tec.com.br

Contato com o Autor: vitor@cerne-tec.com.br



FEITO NO BRASIL

***“Venha também sobre mim a tua benignidade, ó Senhor, e a tua
salvação, segundo a tua palavra.”***

Sl 119:41

Cerne Tecnologia

A Cerne Tecnologia tem uma equipe preparada para desenvolvimento de projetos eletrônicos em diversas áreas: Médica, Entretenimento, Industrial, Robótica, Científica, Automobilística, Aeronáutica, etc. Trabalhamos com tecnologia microcontrolada usando o PIC, ARM, AVR, 8051, dsPIC, PIC24, PIC32 além do Arduino, Raspberry, Beaglebone etc. Desenvolvemos o projeto desde sua concepção até a entrega do produto final, passando pelas etapas de esquema elétrico, protótipo e desenvolvimento de circuito impresso.

Desenvolvemos aplicativos para smartphones/tablets Android, iOS, Blackberry, Windows Phone e no desenvolvimento de softwares a nível PC para plataforma Windows, usando ferramentas como o Visual Basic, C# e C++.

Atuamos na parte de montagem de placas, onde podemos fornecer ambos os serviços de desenvolvimento de projetos e produção ou apenas um destes.

Desenvolvemos esquemas elétricos e layout de PCI, tanto em tecnologia convencional como SMD.

Temos a flexibilidade de customizar um de nossos produtos, de modo a atender a uma necessidade específica do cliente, tornando o custo de desenvolvimento menor se comparado a construção de um projeto desde a sua fase inicial.

Desenvolvemos e fornecemos kits didáticos para diversos microcontroladores além de apostilas, livros e e-books.

Na hora de desenvolver um projeto ou equipar seu laboratório não hesite em nos contatar. Entre em contato conosco através do endereço cerne-tec.com.br para obter mais informações.



Sumário

Capítulo I – Metodologia de desenvolvimento	7
1. Introdução	7
Capítulo II – Programação em Python	8
1. Introdução	8
2. Operadores aritméticos	10
3. Operadores lógicos	12
4. Operadores de bits (bitwise operators)	13
5. Funções de conversão	13
6. Comentários	14
7. Variáveis	15
8. Operadores Relacionais	17
9. Trabalhando com strings	19
10. O comando If	23
11. O comando while	24
12. O comando for	26
13. Usando listas	28
14. Conhecendo as Tuplas	31
15. Dicionários	31
16. Conjuntos	32
17. Criando scripts	33
18. Pi e número de Euler	36
19. Funções matemáticas	37
20. Obtendo a data e hora	37
21. Calculando o tempo para executar uma rotina	38
22. Emitindo som	39
23. Calendar	39
24. Números complexos	40

25. Funções Pré-Definidas	40
26. Comando type	46
27. Entrada de dados	46
28. Comando break	48
29. Tratamento de erro	49
30. Impressão formatada	51
31. Função bool	51
32. Operador in	52
33. Operador randômico	53
34. Obtendo ajuda	55
35. Operador de formatação	56
36. Criando funções de usuário	57
37. Números perfeitos	60
38. Acesso a arquivos externos	60
Capítulo III – Plotando gráficos	63
Capítulo IV – Função sigmoide	86
Capítulo V – Perceptron	88
Capítulo VI – Agglomerative Clustering	98
1. Introdução	98
2. Conjunto de dados (dataset) Iris	99
3. Objetivo	101
4. Código-fonte	102

Capítulo I

Metodologia de desenvolvimento

1. Introdução

Este livro visa introduzir e explorar o conceito de agrupamento hierárquico utilizando o algoritmo *Agglomerative Clustering*, aplicado a um cenário prático usando Python. O foco está em ensinar como utilizar essa técnica de aprendizado não supervisionado para identificar padrões e relações inerentes aos dados, organizando-os em grupos baseados em similaridade. Em particular, abordamos um exemplo prático de agrupamento hierárquico de flores utilizando o famoso conjunto de dados Iris, demonstrando como o *Agglomerative Clustering* pode ser usado para agrupar as características morfológicas das plantas em diferentes clusters, facilitando a visualização e a interpretação das relações entre as classes. Utilizando o método de ligação Ward e a métrica de distância Euclidiana, exploramos como a organização hierárquica dos dados em uma árvore de cluster (dendrograma) pode revelar insights sobre a estrutura dos dados e a proximidade entre as diferentes espécies de flores.

Capítulo II

Programação em Python

1. Introdução

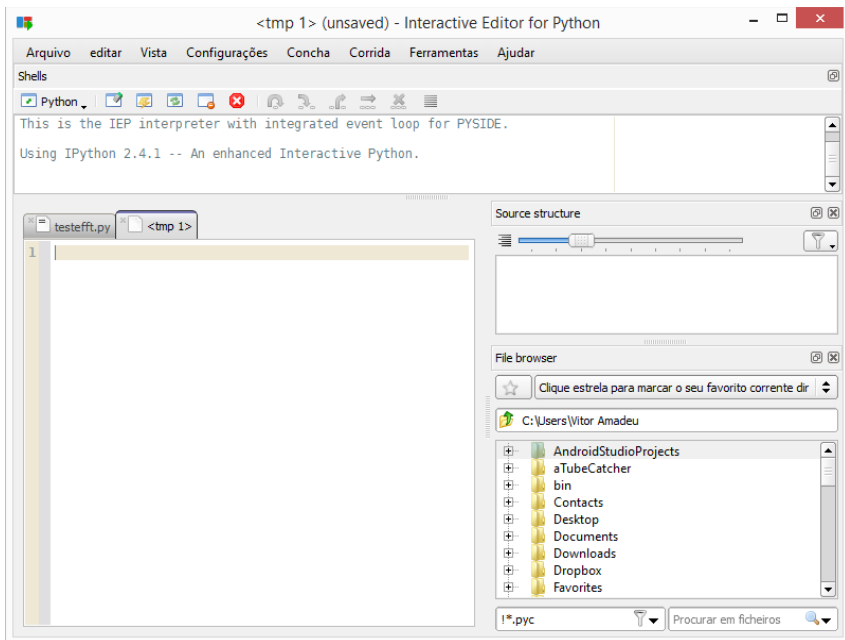
Nesta literatura a distribuição Pyzo foi utilizada, no qual a última versão pode ser baixada através do link abaixo.

<http://www.pyzo.org/downloads.html>

Baixe e instale também a última versão do Python disponível em:

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

A vantagem desta distribuição é que ela por padrão já vem com as bibliotecas a serem utilizadas ao longo desta obra. Após a instalação inicialize o programa, a tela a seguir será apresentada.



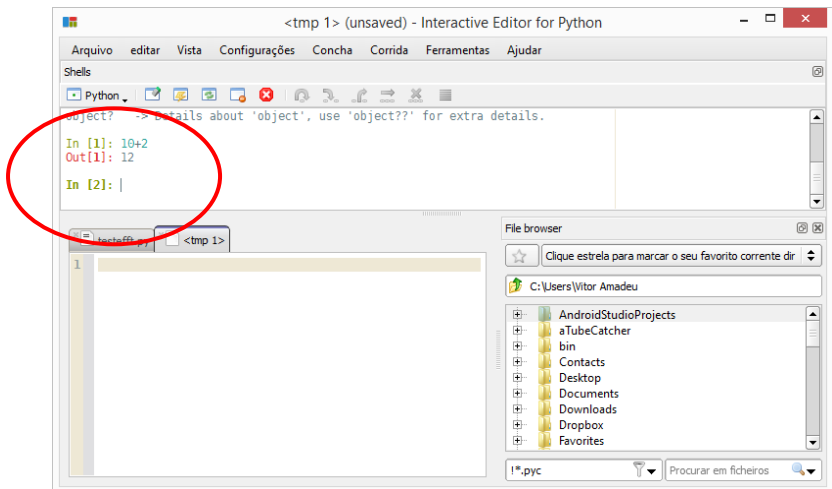
Outra possibilidade é usar a plataforma online chamada Google Colab, no qual é possível criar *notebooks* (cadernos) para executar e testar programas.

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

O Python é um software interpretado, ou seja, cada comando digitado no ambiente é logo executado após você pressionar o enter do teclado. Nos próximos tópicos, estaremos exercitando diversos exemplos no Python, de forma a entender como o mesmo funciona.

2. Operadores aritméticos

Podemos usar o Python como uma calculadora, bastando neste caso digitar diretamente a expressão matemática no mesmo, usando neste caso o prompt. Observe abaixo:



Note que ao digitar a expressão e pressionar o enter, o comando é imediatamente processado, tendo como resultado a soma da operação. Outra forma é escrever o programa como um script, salvá-lo e executá-lo em seguida, indo no menu Corrida-> Corrida arquivo como script.

O Python possui diversos operadores aritméticos, como os citados a seguir.