



UNOESTE
SISTEMAS DE INFORMAÇÃO

FELIPE AUGUSTO GOBBO DOS SANTOS: RA: 262319020
LUIGI SARDELARI SCALIANTE: RA: 262318687
JOÃO CLÁUDIO MARTINEZ DA COSTA: RA:262319640
BHRENO GABRIEL MORAES FERREIRA: RA: 262318792

APLICAÇÃO COM ÁRVORE DE DECISÃO

Sumário

• A BASE DE DADOS:	3
• Introdução	3
• SOBRE A BASE DE DADOS:.....	3
• TIPOS DE ANIMAIS	4
• Mamíferos Terrestres	4
• Aves	4
• Répteis.....	4
• Peixes	4
• Anfíbios	4
• Insetos	4
• Moluscos e Artrópodes Aquáticos.....	4
• ATRIBUTOS	5
• ÁRVORE DE DECISÃO E MATRIZ DE CONFUSÃO	6
• ÁRVORE DE DECISÃO	6
• PREDIÇÕES.....	7
• ANÁLISE DE RESULTADOS:.....	8
• AUC (Area Under The Curve):.....	8
• CA (Classification Accuracy)	8
• Precision.....	8
• Recall	8
• F1	8
• Execução do Programa	9

A BASE DE DADOS:

Introdução

Explicação breve sobre o assunto e o tipo de linguagem usada:

O tema escolhido para a criação da árvore e desenvolvimento do código foi animais do zoológico. Na qual consiste em chegar ao resultado correto (espécie e nome do animal) através das características de cada um. Sendo assim com base na estrutura de árvore (algoritmo pronto resultado pelo Orange) inicializamos o algoritmo com linguagem C.

.

SOBRE A BASE DE DADOS:

A base de dados contém perguntas relacionadas a características de animais, gerais de animais, sejam físicas, tipo de pelagem, tamanho, domesticidade... E com base nos resultados, podemos chegar à sua classificação específica.

Link para a base de dados:

<https://archive.ics.uci.edu/dataset/111/zoo>

TIPOS DE ANIMAIS

Mamíferos Terrestres

Porco-da-terra, Antílope, Urso, Javali, Búfalo, Bezerro, Preá, Guepardo, Veado, Elefante, Morcego-da-fruta, Girafa, Cabra, Gorila, Hamster, Lebre, Leopardo, Leão, Lince, Vison, Toupeira, Mangustoo, Gambá, Órix, Ornitorrinco, Doninha, Pônei, Toninha, Puma, Gato, Guaxinim, Rena, Foca, Leão Marinho, Esquilo, Morcego-Vampiro, Ratazana, Canguru, Lobo.

Aves

Galinha, Corvo, Pomba, Pato, Flamingo, Gaivota, Falcão, Kiwi, Alaudidae, Avestruz, Periquito, Pinguim, Faisão, Réia, Talha-Mar, Skua, Pardal, Cisne, Abutre, Carriça.

Répteis

Víbora, Cobra d'água, Verme-lento, Tartaruga, Tuatara.

Peixes

Perca, Carpa, Peixe-Gato, Chub, Tubarão, Arinca, Herring, Lúcio, Piranha, Cavalo-Marinho, Solha, Arraia, Atum.

Anfíbios

Sapo, Salamandra, Rã, Perereca.

Insetos

Pulga, Mosquito, Abelha, Mosca, Joaninha, Mariposa, Cupim, Vespa.

Moluscos e Artrópodes Aquáticos

Molusco, Carangueijo, Lagostim, Lagosta, Pouvo, Escorpião, Vespa-do-mar, Lesma, Estrela-do-mar, Minhoca.

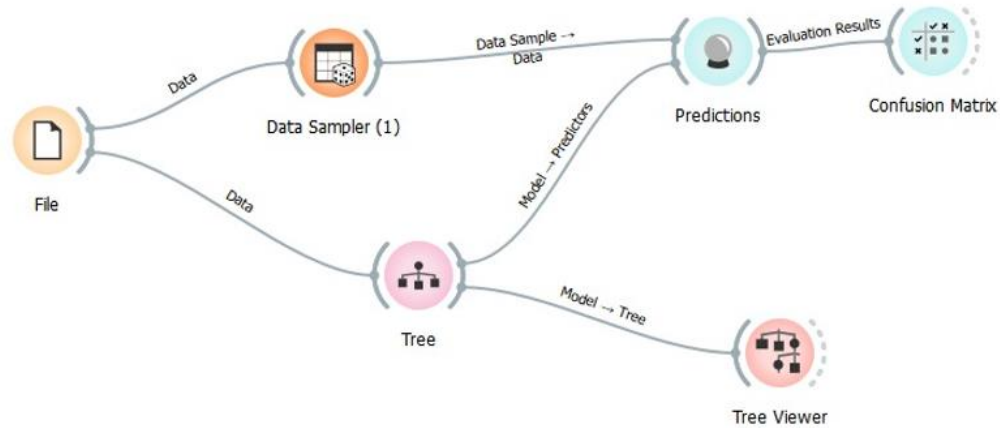
ATRIBUTOS

(Nome do atributo e tipo de domínio de valor).

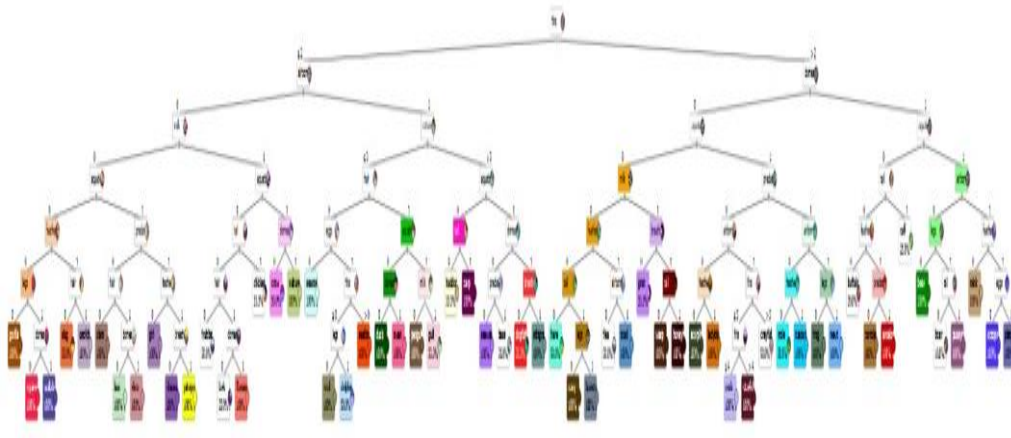
1. nome do animal:	Exclusivo para cada valor
2. pelo	Boolean
3. penas	Boolean
4. ovos	Boolean
5. leite	Boolean
6. aéreo	Boolean
7. aquático	Boolean
8. predador	Boolean
9. dentes	Boolean
10. coluna vertebral	Boolean
11. respiração	Boolean
12. venenoso	Boolean
13. barbatanas	Boolean
14. pernas	Numérico (conjunto dos valores: {0,2,4,5,6,8})
15. cauda	Boolean
16. doméstico	Boolean
17. tamanho do gato	Boolean
18. tipo	Numérico (valores inteiros no intervalo [1,7])

ÁRVORE DE DECISÃO E MATRIZ DE CONFUSÃO

Para gerarmos a Árvore de Decisão e a Matriz de Confusão, montamos a seguinte estrutura no programa Orange:



ÁRVORE DE DECISÃO



(Devido ao tamanho, a árvore ficou distorcida no word)

PREDIÇÕES

A árvore foi gerada com um conjunto de dados de 70% para treino e 30% para teste. Aqui estão os dados:

30%

Model	AUC	CA	F1	Prec	Recall	MCC
Tree	0.997	0.700	0.689	0.683	0.700	0.701

70%

Model	AUC	CA	F1	Prec	Recall	MCC
Tree	0.987	0.571	0.510	0.490	0.571	0.574

ANÁLISE DE RESULTADOS:

AUC (Area Under The Curve): indica o quão exato o modelo é em sua classificação. Ou seja, o quão preciso é ao indicar a classe de um registro.

Quanto mais próximo de 1, melhor: nosso modelo alcançou uma pontuação de 0,990, indicando que é bastante preciso ao indicar o nível de obesidade de um registro.

CA (Classification Accuracy): a acurácia do modelo é dada pela fórmula $CA =$

$$\frac{\text{número de predições corretas}}{\text{número total de predições}} .$$

Nosso modelo apresentou um CA de 0,590, o que não é ideal, pois o valor máximo e Melhor é 1.

Precision: Semelhante a CA, a Precision mede quantas vezes o modelo acertou em relação ao total de vezes em que ele tenta acertar.

Recall: Indica a quantidade de vezes em que o modelo acertou a classe em relação À quantidade de vezes em que deveria ter acertado.

F1: Indica o quão Precision e Recall estão condizentes entre si. Ou seja, quanto maior for F1, melhor, pois Precision e Recall estão em níveis parecidos (não Necessariamente alto ou baixo, mas, sim, próximos).

Execução do Programa

Para melhor entender a dinâmica do programa mostraremos o início:

```
1  #include<stdio.h>
2  #include<stdlib.h>
3  #include<ctype.h>
4  #include<string.h>
5  #include<conio2.h>
6  #include <windows.h>
7
8  void LimpaMensagem(void)
9  {
10     int x,y;
11     for(x=20;x<78;x++){
12         gotoxy(x,23);
13         printf(" ");
14     }
15 }
16
17 void LimpaTela(void)
18 {
19     int x,y;
20     //LIMPA QUADRADO
21     for(x=30;x<=78;x++){
22         for(y=6;y<=20;y++){
23             gotoxy(x,y);
24             printf(" ");
25         }
26     }
27     //LIMPA MENSAGEM
28     for(x=20;x<78;x++){
29         gotoxy(x,23);
30         printf(" ");
31     }
32 }
33 }
```

Acima na imagem é mostrado o começo do código, importamos bibliotecas e criamos funções para desenvolver.

Logo após os créditos e as criações, começamos as perguntas:

```

127 void menu()
128 {
129     gotoxy(40,7);
130     printf("#*** Arvore zoologico ***#");
131     gotoxy(47,9);
132     printf("FIPP UNOESTE");
133     /*gotoxy(30,9);
134     backgroundcolor(15);
135     printf("Numero de animais possiveis: ");
136     gotoxy(30,10);
137     backgroundcolor(15);
138     printf("Fonte dos dados: ");
139     gotoxy*/
140
141     gotoxy(20,23);
142     printf("Aperte qualquer tecla para continuar");
143
144     getch();
145 }
146
147 void Erro(void)
148 {
149     gotoxy(20,23);
150     printf("Valor invalido!, por favor apenas 'S' ou 'N'");
151     getch();
152     LimpaMensagem();
153 }
154
155 int verifica(int pos)
156 {
157     int i;
158     for(i=0; i<100 && pos != 'S' && pos!= 'N' && pos!=27; i++)
159     {
160         Erro();
161         pos = toupper(getch());
162     }
163     return pos;
164 }
165
166 void Animais(void)
167 {
168     char x,op='S',p;
169     int i=7;
170
171     menu();
172     LimpaTela();
173     while(op == 'S')

```

Finalizando a primeira parte do código que inicializa o programa, começamos as decisões (if else):

```

67 {
68     char x,op='S',p;
69     int i=7;
70
71     menu();
72     LimpaTela();
73     while(op == 'S')
74     {
75         gotoxy(30,7);
76         printf("O animal tem barbatana? (S/N)");
77         gotoxy(20,23);
78         printf("Ou Pressione 2 [Esc] - Sair!");
79         x = toupper(getch());
80         x = verifica(x);
81         if(x == 'S' && x != 27) //1
82         {
83             gotoxy(30,8);
84             printf("Aerotransportada? (S/N)");
85             gotoxy(20,23);
86             printf("Ou Pressione 2 [Esc] - Sair!");
87             x = toupper(getch());
88             x = verifica(x);
89             if(x == 'S' && x != 27) //2
90             {
91                 gotoxy(30,9);
92                 printf("O animal e mamifero? (S/N)");
93                 gotoxy(20,23);
94                 printf("Ou Pressione 2 [Esc] - Sair!");
95                 x = toupper(getch());
96                 x = verifica(x);
97                 LimpaMensagem();
98                 if(x == 'S' && x != 27) //3
99                 {
100                     gotoxy(30,10);
101                     printf("O animal e aquatico? (S/N)");
102                     x = toupper(getch());
103                     x = verifica(x);
104                     if(x == 'S' && x != 27) //4
105                     {
106                         gotoxy(30,11);
107                         printf("Possui classe de animal? (S/N)");
108                         x = toupper(getch());
109                         x = verifica(x);
110                         if(x == 'S' && x != 27) //5
111                         {
112                             gotoxy(30,12);

```

Assim foi se desenvolvendo o programa na temática "Zoo".

E ao fim do programa temos as devidas decisões tomadas de forma correta conforme as características do animal:

```

65 x = toupper(getch());
66 x = verifica(x);
67 if(x == 'S')
68 {
69     gotoxy(20,23);
70     printf("O animal escolhido foi: 'Boar'");
71 }
72 else
73 {
74     gotoxy(20,23);
75     printf("O animal escolhido foi: 'Pussycat'");
76 }
77 }
78 }
79 else
80 {
81     gotoxy(30,11);
82     printf("Possui classe animal? (S/N)");
83     x = toupper(getch());
84     x = verifica(x);
85     if(x == 'S')
86     {
87         gotoxy(20,23);
88         printf("O animal escolhido foi: 'Octopus'");
89     }
90     else
91     {
92         gotoxy(30,12);
93         printf("Possui penas? (S/N)");
94         x = toupper(getch());
95         x = verifica(x);
96         if(x == 'S')
97         {
98             gotoxy(20,23);
99             printf("O animal escolhido foi: 'Mink'");
100         }
101         else
102         {
103             gotoxy(20,23);
104             printf("O animal escolhido foi: 'Platplus'");
105         }
106     }

```