# Lista 5

Aluno: Yan Gustavo Pegyn Silva

**1 -**

**W(t, d) = (IF TF(t,d) > 0, 1 + log 10 TF(t,d), 0)**

**IDF(termo)= log10 N/DF(t)**

\*\* Considerei N = 10

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | IDF | Doc1 | Doc2 | Doc3 | Doc4 |
| alheio | 0,7 | 0 | 15 | 0 | 220 |
| apressado | 0,4 | 145 | 3 | 200 | 54 |
| bom | 0,52 | 30 | 200 | 1231 | 0 |
| errado | 0,4 | 12 | 83 | 120 | 400 |
| gente | 0,7 | 338 | 155 | 0 | 0 |

Calcule os ranqueamentos usando o modelo vetorial para as consultas abaixo:

W(t, d) = (IF TF(t,d) > 0, 1 + log 10 TF(t,d), 0)

TF.IDF = W \* IDF

Mdoc = sqrt(ponto1^2 + ponto2 ^2 + ... + pontoX^2)

**A – alheio, gente**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | IDF | TF.IDF | | | |
|  |  | Doc1 | Doc2 | Doc3 | Doc4 |
| alheio | 0,7 | W = 0  TF.IDF = 0 | W = 1 + 1,18 = 2,18  TF.IDF = 2,18 \* 0,7 = 1,53 | W = 0  TF.IDF = 0 | W = 1 + 2,34 = 3,34  TF.IDF = 3,34 \* 0,7 = 2,34 |
| gente | 0,7 | W = 1 + 2,53 = 3,53  TF.IDF = 3,53 \* 0,7 = 2,47 | W = 1 + 2,19 = 3,19  TF.IDF = 3,19 \* 0,7 = 2,23 | W = 0  TF.IDF = 0 | W = 0  TF.IDF = 0 |

Vdoc1 = (0; 2,47)

Vdoc2 = (1,53; 2,23)

Vdoc3 = (0; 0)

Vdoc4 = (2,34; 0)

Vcons = (0,7; 0,7)

Mdoc1 = (0 + 6,1) ^ 0,5 = 2,47

Mdoc2 = (2,34 + 4,97) ^ 0,5 = 2,7

Mdoc3 = (0 + 0) ^ 0,5 = 0

Mdoc4 = (5,48 + 0) ^ 0,5 = 2,34

Mcons = (0,49 + 0,49) ^ 0,5 = 0,99

Vdoc1n = (0; 1)

Vdoc2n = (0,57; 0,83)

Vdoc3n = (0; 0)

Vdoc4n = (1; 0)

Vconsn = (0,7; 0,7)

Cos(Vdoc1n, Vconsn) = (0 + 0,7) = 0,7

Cos(Vdoc2n, Vconsn) = (0,4 + 0,58) = 0,98

Cos(Vdoc3n, Vconsn) = (0 + 0) = 0

Cos(Vdoc4n, Vconsn) = (0,7 + 0) = 0,7

1º = doc2

2º = doc1 e doc4

4º = doc3

**B – apressado, errado**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | IDF | TF.IDF | | | |
|  |  | Doc1 | Doc2 | Doc3 | Doc4 |
| apressado | 0,4 | W = 1 + 2,16 = 3,16  TF.IDF = 3,16 \* 0,4 = 1,3 | W = 1 + 0,48 = 1,48  TF.IDF = 1,48 \* 0,4 = 0,59 | W = 1 + 2,3 = 3,3  TF.IDF = 3,3 \* 0,4 = 1,32 | W = 1 + 1,73 = 2,73  TF.IDF = 2,73 \* 0,4 = 1,1 |
| errado | 0,4 | W = 1 + 1,08 = 2,08  TF.IDF = 2,08 \* 0,4 = 0,83 | W = 1 + 1,92 = 2,92  TF.IDF = 2,92 \* 0,4 = 1,17 | W = 1 + 2,08 = 3,08  TF.IDF = 3,08 \* 0,4 = 1,23 | W = 1 + 2,6 = 3,6  TF.IDF = 3,6 \* 0,4 = 1,44 |

Vdoc1 = (1,3; 0,83)

Vdoc2 = (0,59; 1,17)

Vdoc3 = (1,32; 1,23)

Vdoc4 = (1,1; 1,44)

Vcons = (0,4; 0,4)

Mdoc1 = (1,69 + 0,69) ^ 0,5 = 1,54

Mdoc2 = (0,35 + 1,37) ^ 0,5 = 1,31

Mdoc3 = (1,74 + 1,51) ^ 0,5 = 1,8

Mdoc4 = (1,21 + 2,07) ^ 0,5 = 1,81

Mcons = (0,16 + 0,16) ^ 0,5 = 0,57

Vdoc1n = (0,84; 0,54)

Vdoc2n = (0,45; 0,89)

Vdoc3n = (0,73; 0,68)

Vdoc4n = (0,61; 0,8)

Vconsn = (0,7; 0,7)

Cos(Vdoc1n, Vconsn) = (0,84\*0,7 + 0,54\*0,7) = (0,59 + 0,38) = 0,97

Cos(Vdoc2n, Vconsn) = (0,45\*0,7 + 0,89\*0,7) = (0,32 + 0,62) = 0,94

Cos(Vdoc3n, Vconsn) = (0,73\*0,7 + 0,68\*0,7) = (0,51 + 0,48) = 0,99

Cos(Vdoc4n, Vconsn) = (0,61\*0,7 + 0,8\*0,7) = (0,43 + 0,56) = 0,99

1º = doc3 e doc4

3º = doc1

4º = doc2

**C – alheio, bom, gente**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | IDF | TF.IDF | | | |
|  |  | Doc1 | Doc2 | Doc3 | Doc4 |
| alheio | 0,7 | W = 0  TF.IDF = 0 | W = 1 + 1,18 = 2,18  TF.IDF = 2,18 \* 0,7 = 1,53 | W = 0  TF.IDF = 0 | W = 1 + 2,34 = 3,34  TF.IDF = 3,34 \* 0,7 = 2,34 |
| bom | 0,52 | W = 1 + 1,48 = 2,48  TF.IDF = 2,48 \* 0,52 = 1,29 | W = 1 + 2,3 = 2,3  TF.IDF = 2,3 \* 0,52 = 1,2 | W = 1 + 3,09 = 4,09  TF.IDF = 4,09 \* 0,52 = 2,13 | W = 0  TF.IDF = 0 |
| gente | 0,7 | W = 1 + 2,53 = 3,53  TF.IDF = 3,53 \* 0,7 = 2,47 | W = 1 + 2,19 = 3,19  TF.IDF = 3,19 \* 0,7 = 2,23 | W = 0  TF.IDF = 0 | W = 0  TF.IDF = 0 |

Vdoc1 = (0; 1,29; 2,47)

Vdoc2 = (1,53; 1,2; 2,23)

Vdoc3 = (0; 2,13; 0)

Vdoc4 = (2,4; 0; 0)

Vcons = (0,7; 0,52; 0,7)

Mdoc1 = (0 ^ 2 + 1,29 ^ 2 + 2,47 ^ 2) ^ 0,5 = 2,79

Mdoc2 = (1,53 ^ 2 + 1,2 ^ 2 + 2,23 ^ 2) ^ 0,5 = 2,96

Mdoc3 = (0 ^ 2 + 2,13 ^ 2 + 0 ^ 2) ^ 0,5 = 2,13

Mdoc4 = (2,4 ^ 2 + 0 ^ 2 + 0 ^ 2) ^ 0,5 = 2,4

Mcons = (0,7 ^ 2 + 0,52 ^ 2 + 0,7 ^ 2) ^ 0,5 = 1,12

Vdoc1n = (0; 0,46; 0,89)

Vdoc2n = (0,52; 0,41; 0,75)

Vdoc3n = (0; 1; 0)

Vdoc4n = (1; 0; 0)

Vconsn = (0,63; 0,46; 0,63)

Cos(Vdoc1n, Vconsn) = (0 + 0,46\*0,46 + 0,89\*0,63) = 0,77

Cos(Vdoc2n, Vconsn) = (0,52\*0,63 + 0,41\*0,46 + 0,75\*0,63) = 0,99

Cos(Vdoc3n, Vconsn) = (0 + 1\*0,46 + 0) = 0,46

Cos(Vdoc4n, Vconsn) = (1\*0,63 + 0 + 0) = 0,63

1º = doc2

2º = doc1

3º = doc4

4º = doc3

**2 -**

**Dados os TFs.**

**W(t, d) = (IF TF(t,d) > 0, 1 + log 10 TF(t,d), 0)**

**IDF(termo)= log10 N/DF(t) N=10**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | IDF | Doc1 | Doc2 | Doc3 | Doc4 |
| alho | 0,52 | 0 | 12 | 344 | 4544 |
| cebola | 0,52 | 23 | 0 | 32 | 332 |
| feijao | 0,4 | 323 | 230 | 30 | 223 |
| pepino | 0,52 | 223 | 333 | 5545 | 0 |
| rabanete | 0,7 | 323 | 230 | 0 | 0 |

Calcule os ranqueamentos usando o modelo vetorial para as

consultas:

W(t, d) = (IF TF(t,d) > 0, 1 + log 10 TF(t,d), 0)

TF.IDF = W \* IDF

Mdoc = sqrt(ponto1^2 + ponto2 ^2 + ... + pontoX^2)

**alho, cebola**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | IDF | TF.IDF | | | |
|  |  | Doc1 | Doc2 | Doc3 | Doc4 |
| alho | 0,52 | W = 0  TF.IDF = 0 | W = 1 + 1,08 = 2,08  TF.IDF = 2,08 \* 0,52 = 1,08 | W = 1 + 2,54 = 3,54  TF.IDF = 3,54 \* 0,52 = 1,84 | W = 1 + 3,66 = 4,66  TF.IDF = 4,66 \* 0,52 = 2,42 |
| cebola | 0,52 | W = 1 + 1,36 = 2,36  TF.IDF = 2,36 \* 0,52 = 1,23 | W = 0  TF.IDF = 0 | W = 1 + 1,51 = 2,51  TF.IDF = 2,51 \* 0,52 = 1,31 | W = 1 + 2,52 = 3,52  TF.IDF = 3,52 \* 0,52 = 1,83 |

Vdoc1 = (0; 1,23)

Vdoc2 = (1,08; 0)

Vdoc3 = (1,84; 1,31)

Vdoc4 = (2,42; 1,83)

Vcons = (0,52; 0,52)

Mdoc1 = (0 + 1,51) ^ 0,5 = 1,23

Mdoc2 = (1,17 + 0) ^ 0,5 = 1,08

Mdoc3 = (3,39 + 1,72) ^ 0,5 = 2,26

Mdoc4 = (5,86 + 3,35) ^ 0,5 = 3,03

Mcons = (0,27 + 0,27) ^ 0,5 = 0,74

Vdoc1n = (0; 1)

Vdoc2n = (1; 0)

Vdoc3n = (0,81; 0,58)

Vdoc4n = (0,8; 0,6)

Vconsn = (0,7; 0,7)

Cos(Vdoc1n, Vconsn) = (0 + 0,7) = 0,7

Cos(Vdoc2n, Vconsn) = (0,7 + 0) = 0,7

Cos(Vdoc3n, Vconsn) = (0,57 + 0,4) = 0,97

Cos(Vdoc4n, Vconsn) = (0,56 + 0,42) = 0,98

1º = doc4

2º = doc3

3º = doc3 e doc4

**pepino, rabanete**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | IDF | TF.IDF | | | |
|  |  | Doc1 | Doc2 | Doc3 | Doc4 |
| pepino | 0,52 | W = 1 + 2,35 = 3,35  TF.IDF = 3,35 \* 0,52 = 1,74 | W = 1 + 2,52 = 3,52  TF.IDF = 3,52 \* 0,52 = 1,83 | W = 1 + 3,74 = 4,74  TF.IDF = 4,74 \* 0,52 = 2,46 | W = 0  TF.IDF = 0 |
| rabanete | 0,7 | W = 1 + 2,51 = 3,51  TF.IDF = 3,51 \* 0,7 = 2,46 | W = 1 + 2,36 = 3,36  TF.IDF = 3,36 \* 0,7 = 2,35 | W = 0  TF.IDF = 0 | W = 0  TF.IDF = 0 |

Vdoc1 = (1,74; 2,46)

Vdoc2 = (1,83; 2,35)

Vdoc3 = (2,46; 0)

Vdoc4 = (0; 0)

Vcons = (0,52; 0,7)

Mdoc1 = (3,02 + 6,05) ^ 0,5 = 3,01

Mdoc2= (3,35 + 5,52) ^ 0,5 = 2,97

Mdoc3 = (6,05 + 0) ^ 0,5 = 2,46

Mdoc4 = 0

Mcons = (0,27 + 0,49) ^ 0,5 = 0,87

Vdoc1n = (0,58; 0,82)

Vdoc2n = (0,62; 0,79)

Vdoc3n = (1; 0)

Vdoc4n = (0; 0)

Vconsn = (0,6; 0,81)

Cos(Vdoc1n, Vconsn) = (0,35 + 0,66) = 1,01

Cos(Vdoc2n, Vconsn) = (0,37 + 0,64) = 1,01

Cos(Vdoc3n, Vconsn) = (0,6 + 0) = 0,6

Cos(Vdoc4n, Vconsn) = (0 + 0) = 0

1º = doc1 e doc2

3º = doc3

4º = doc4

**alho, feijão, rabanete**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | IDF | TF.IDF | | | |
|  |  | Doc1 | Doc2 | Doc3 | Doc4 |
| alho | 0,52 | W = 0  TF.IDF = 0 | W = 1 + 1,08 = 2,08  TF.IDF = 2,08 \* 0,52 = 1,08 | W = 1 + 2,54 = 3,54  TF.IDF = 3,54 \* 0,52 = 1,84 | W = 1 + 3,66 = 4,66  TF.IDF = 4,66 \* 0,52 = 2,42 |
| feijao | 0,4 | W = 1 + 2,51 = 3,51  TF.IDF = 3,51 \* 0,4 = 1,4 | W = 1 + 2,36 = 3,36  TF.IDF = 3,36 \* 0,4 = 1,34 | W = 1 + 1,48 = 2,48  TF.IDF = 2,48 \* 0,4 = 0,99 | W = 1 + 2,35 = 3,35  TF.IDF = 3,35 \* 0,4 = 1,34 |
| rabanete | 0,7 | W = 1 + 2,51 = 3,51  TF.IDF = 3,51 \* 0,7 = 2,46 | W = 1 + 2,36 = 3,36  TF.IDF = 3,36 \* 0,7 = 2,35 | W = 0  TF.IDF = 0 | W = 0  TF.IDF = 0 |

Vdoc1 = (0; 1,4; 2,46)

Vdoc2 = (1,08; 1,34; 2,35)

Vdoc3 = (1,84; 0,99; 0)

Vdoc4 = (2,42; 1,34; 0)

Vcons = (0,52; 0,4; 0,7)

Mdoc1 = (0 + 1,96 + 6,05) ^ 0,5 = 2,83

Mdoc2 = (1,17 + 1,8 + 5,53) ^ 0,5 = 2,92

Mdoc3 = (3,39 + 0,98 + 0) ^ 0,5 = 2,09

Mdoc4 = (5,86 + 1,8 + 0) ^ 0,5 = 2,77

Mcons = (0,27 + 0,16 + 0,49) ^ 0,5 = 0,96

Vdoc1n = (0; 0,5; 0,87)

Vdoc2n = (0,37; 0,46; 0,8)

Vdoc3n = (0,88; 0,47; 0)

Vdoc4n = (0,87; 0,48; 0)

Vconsn = (0,54; 0,42; 0,73)

Cos(Vdoc1n, Vconsn) = (0 + 0,21 + 0,64) = 0,85

Cos(Vdoc2n, Vconsn) = (0,2 + 0,19 + 0,58) = 0,97

Cos(Vdoc3n, Vconsn) = (0,48 + 0,2 + 0) = 0,68

Cos(Vdoc4n, Vconsn) = (0,47 + 0,2 + 0) = 0,67

1º = doc2

2º = doc1

3º = doc3

4º = doc4