

9	A)																
10	termo	fi1	fi2	fi3	fi4	fi5	TFi1	TFi2	TFi3	TFi4	TFi5	IDFi	d1	d2	d3	d4	d5
11	logan	1	1	0	1	0	1	1	0	1	0	0.736	0.736	0.736	0	0.736	0
12	ororo	1	0	1	1	0	1	0	1	1	0	0.736	0.736	0	0.736	0.736	0
13	stark	0	1	1	0	0	0	1	1	0	0	1.321	0	1.321	1.321	0	0
14	parker	0	2	0	0	0	0	2	0	0	0	2.643	0	4.643	0	0	0
15	groot	0	0	1	3	2	0	0	1	2.584	2	0.736	0	0	0.736	1.905	1.473
16	rocket	0	0	1	0	2	0	0	1	0	2	1.321	0	0	1.321	0	2.643
17	x-men	1	0	0	1	0	1	0	0	1	0	1.321	1.321	0	0	1.321	0
18	vingadores	0	1	2	0	0	0	1	2	0	0	1.321	0	1.321	2.643	0	0
19	guardiões	0	0	2	0	1	0	0	2	0	1	1.321	0	0	2.643	0	1.321

20

21	norma												1.681	5.058	4.306	2.541	3.301
----	-------	--	--	--	--	--	--	--	--	--	--	--	-------	-------	-------	-------	-------

22

23 B)

24 $q1 = \{\text{logan ororo x-men}\}.$

25 $TF_q = (1, 1, 0, 0, 0, 0, 1, 0, 0)$

26 $IDF_q = (0.736, 0.736, 0, 0, 0, 0, 1.321, 0, 0)$

27 $TF-IDF_q = (0.736, 0.736, 0, 0, 0, 0, 1.321, 0, 0)$

28 $||q|| = 1.681$

29

30 $d1 = (0.736*0.736 + 0.736*0.736 + 1.321*1.321) / 3.362 = 0.841$

31 $d2 = (0.736*0.736 + 0.736*0 + 1.321*0) / 6.739 = 0.080$

32 $d3 = (0.736*0 + 0.736*0.736 + 1.321*0) / 5.987 = 0.090$

33 $d4 = (0.736*0.736 + 0.736*0.736 + 1.321*1.321) / 4.222 = 0.669$

34 $d5 = (0.736*0 + 0.736*0 + 1.321*0) / 4.982 = 0$

35

36 Como é possível notar, d1 e d4 apresentam o maior grau de similaridade por conter todos os termos.

37 Entretanto, por ser menor, o grau de d1 se torna maior que de d4.

38 D5, por outro lado, não possui nenhum dos termos da consulta, portanto, seu grau de similaridade é 0.

39

40 C) Como obtemos a representação vetorial de cada documento, conseguimos sim compara-los.

41 A similaridade entre o d1 e d5 seria igual a 0, visto que nenhum dos termos presentes em d1 estão

42 presentes em d2, e vice versa. Ou seja, não há semelhança alguma entre eles.

43 O calculo para esta similaridade é o seguinte:

44

45 $TF-IDF_d1 = (0.736, 0.736, 0, 0, 0, 0, 1.321, 0, 0)$

46 $TF-IDF_d5 = (0, 0, 0, 0, 1.473, 2.643, 0, 0, 1.321)$

47 $sim(d1,d5) = (0.736*0 + 0.736*0 + 0*1.473 + 0*2.643 + 1.321*0 + 0*1.321) / 4.982 = 0$