

TUGAS BESAR KECERDASAN BUATAN

- TOPIK** : KMEANS
- KASUS** : Peringkasan dokumen
- DESKRIPSI** : Sebuah dokumen yang terlalu panjang seringkali membuat pembaca menjadi enggan untuk mempelajarinya. Peringkasan dokumen adalah salah satu cara untuk memudahkan pembaca untuk memperoleh informasi yang penting dari suatu dokumen. Dalam system peringkasan otomatis harus mengolah dokumen yang berupa data teks menjadi data numerik sehingga dapat diproses oleh algoritma atau biasa disebut proses ekstraksi fitur. Salah satu cara ekstraksi fitur adalah Tf-Idf. Nilai hasil Tf-Idf tersebut digunakan sebagai masukan algoritma KMEANS.

Pada tabel berikut ini adalah hasil ekstraksi fitur Tf-Idf dari suatu dokumen yang akan diringkaskan

Dokumen 1		Dokumen 2		Dokumen 3	
KALIMAT	TF-IDF	KALIMAT	TF-IDF	KALIMAT	TF-IDF
K1	10.91	K1	1.91	K1	15.91
K2	5.316	K2	4.16	K2	2.35
K3	9.155	K3	19.55	K3	8.155
K4	9.145	K4	7.14	K4	9.54
K5	8.678	K5	8.90	K5	8.87
K6	15.653	K6	15.67	K6	17.53
K7	10.257	K7	13.47		
		K8	7.44		

Dokumen 4		Dokumen 5		Dokumen 6	
KALIMAT	TF-IDF	KALIMAT	TF-IDF	KALIMAT	TF-IDF
K1	10.953	K1	1.91	K1	1.91
K2	6.316	K2	4.16	K2	4.35
K3	7.135	K3	14.55	K3	18.53
K4	18.786	K4	17.14	K4	20.54
K5	10.63	K5	9.90	K5	9.77
K6	11.357	K6	8.69	K6	13.53
		K7	12.87	K7	8.66
		K8	5.33		

Dokumen 7		Dokumen 8		Dokumen 9	
KALIMAT	TF-IDF	KALIMAT	TF-IDF	KALIMAT	TF-IDF
K1	12.19	K1	11.91	K1	15.91
K2	3.36	K2	6.26	K2	2.35
K3	5.55	K3	9.54	K3	8.155
K4	8.14	K4	17.14	K4	9.54
K5	18.68	K5	9.98	K5	8.87
K6	11.53	K6	4.79	K6	17.53
K7	11.70	K7	12.37		
		K8	7.44		

Tentukan anggota cluster yang dihasilkan oleh KMEANS untuk masing-masing dokumen.

TUGAS BESAR KECERDASAN BUATAN

- TOPIK

:

KMEANS
- KASUS

:

Pengelompokan warna dalam citra
- DESKRIPSI

:

Sebuah citra yang terdiri dari beberapa warna akan ditentukan kelompok-kelompok warna. Citra tersebut memiliki format RGB. Setiap piksel dalam citra tersebut memiliki 3 nilai yaitu nilai warna merah (R), hijau (G), dan biru (B).

Berikut ini nilai RGB dari citra tersebut.

Piksel	R	G	B	Piksel	R	G	B
1	232	176	55	41	130	5	210
2	27	209	69	42	64	26	168
3	125	30	188	43	230	115	182
4	92	12	150	44	250	148	223
5	162	61	157	45	253	134	130
6	77	91	151	46	140	202	254
7	178	46	213	47	129	226	77
8	102	42	150	48	238	218	6
9	96	54	182	49	243	132	40
10	155	76	246	50	43	197	167
11	105	208	35	51	14	24	242
12	228	210	147	52	235	84	19
13	40	221	185	53	93	76	59
14	186	166	88	54	48	120	232
15	154	254	185	55	86	231	31
16	204	181	9	56	31	71	147
17	249	197	141	57	191	133	187
18	209	22	210	58	130	75	223
19	155	148	230	59	254	99	34
20	155	245	100	60	38	190	118
21	185	172	111	61	56	12	125
22	22	165	109	62	47	184	136
23	225	40	214	63	11	49	163
24	87	150	90	64	126	173	43
25	141	31	10	65	13	238	139
26	73	39	8	66	109	32	216
27	232	174	6	67	179	164	93
28	253	20	231	68	207	153	227
29	160	143	178	69	141	227	21
30	14	6	227	70	29	233	170
31	162	9	146	71	96	29	228
32	61	24	124	72	71	159	173
33	201	6	213	73	60	142	66
34	209	249	173	74	217	182	229
35	197	98	239	75	200	45	137
36	175	211	198	76	207	11	10
37	153	217	40	77	183	194	52
38	106	0	53	78	182	254	169
39	20	27	41	79	128	42	190
40	107	3	26	80	32	250	1

Tentukan anggota masing-masing cluster dengan menggunakan KMEANS