

2017-2018 GÜZ DÖNEMİ
ALGORİTMA ANALİZİ PROJESİ

**Locality Sensitive Hashing Yöntemi ile
Benzer Dokümanların Belirlenmesi**

SİNAN CAN KESKİN

11011055

YÖNTEM

Projede bizden belirli sayıdaki dokümanın LSH ile benzerlik oranlarının bulunması istenmiştir. Bunu yaparken LSH'in temel adımlarından ilk ikisi olan Shingling ve Minhashing adımları gerçekleşmiştir. Minhashing'in bize sağlayacağı faydalardan bahsetmek gerekirse, dokümanları kümeler halinde temsil etmemize ve benzerlik karşılaştırması yaparken bu kümeler üzerinden hareket etmemize olanak sağlayan bir algoritmadır. Dokümanları kümeler olarak temsil etmemize yarayacak işlem ise shingling'dir. Dışardan okunan bir K değerine göre dokümanların ayrı ayrı K-Shingle listeleri çıkartılmıştır. Bu projede harf shingle kullanılmıştır. Shingle yapısından bahsetmek gerekirse, örneğin $K = 4$ değeri için, dokumanın içeriği karakterler baştan sona kadar dörderli gruptara ayrılmış, gruptama işlemi yapılrken birer karakter ilerlenmiştir. Küme mantığıyla hareket ettiğimiz için her shingle her doküman içerisinde sadece bir kez kullanılmak zorundadır. Ayrıca alfabetik karakterler ve boşluk karakteri haricindeki karakterler shingle listeleri oluşturulurken dikkate alınmamıştır.

Örnek vermek gerekirse, "the idea behind" parçasındaki $K = 4$ shingleleri şu şekilde dir:

"the_ ", "he_i", "e_id ", "_ide", "idea", "dea_ ", "ea_b", "a_be", "_beh", "behi", "ehin", "hind"

Shinge listeleri struct içerisinde tutulmuştur. Bu structun içerisinde karakter matrisi ve her dokümanın unique shingle sayısını tutan değerler mevcuttur.

Geliştirdiğim algoritmanın ve programın ana adımları şu şekildedir;

1. Kullanıcıdan K sayısı ve karşılaştırma yapılacak olan doküman sayısı alınır.
2. Alınan dokuman sayısına göre shingle listeleri için bellek ayrırlır ve dosyalardan veri okunmaya başlanır(read_data fonksiyonu).
3. Veri, geçici değişkenlere kelime kelime okunur ve her kelimeden sonra bir boşluk eklenir. Dolayısıyla birden fazla boşluk olma durumları filtre edilmiş olur. Daha sonra buffera alınan veri alfabetik olmayan karakterlerden arındırılır ve büyük harfler küçük harfe çevrilerek işleme hazır hale getirilir.
4. find_shingle(K , doc_number) fonksiyonu, K sayısına ve dokuman sayısına göre shingle listelerini hazırlar. Her doküman için unique shingle listeleri ve her dokümanın unique shingle sayıları elde edilmiş olur.
5. Jaccard_distance(doc_number) fonksiyonu, dokümanların shingle listelerini karşılaştırarak jaccard benzerlik matrisini elde eder. Matris ekrana bastırılır.
6. Kullanıcı görmek isterse shingle listelerini ekrana yazdırabilir.
7. Jaccard matrisi bulunduktan sonra minhashing için tüm dokümanların shingle listeleri merge edilir, bu listedeki shinglelar da küme mantığı gereğince unique olmak zorundadır.
8. Genel shingle tablosu elde edildikten sonra, hangi shingle'in hangi dosyada bulunduğu gösteren map matrisi doldurulur. Map matrisinin boyutu $M \times 20$ 'dir. M , toplam unique shingle sayısıdır.
9. M sayısı find_closest_prime fonksiyonuna gönderilir ve fonksiyon dönüş değeri olarak M' ye en yakın ve M' den küçük asal sayıyı döndürür. Bu sayı minhashing işleminde kullanılacaktır.

10. M sayısı `find_signature_matrix` fonksiyonuna gönderilir. Signature matrix, 100×20 boyutundadır ve her dokümanın imza matrisidir. Buradaki 100 sayısı, kullandığımız hash fonksiyonu sayısından gelmektedir. Hash fonksiyonu sayısı arttıkça benzerlik hassasiyeti de artmaktadır.
11. Signature matrixin gözleri başlangıç olarak 9999 ileilklen dirilmiştir. Üretilen random hash değerlerine göre hash matrisi doldurulur. Hash matrisinin boyutu $M \times 100$ 'dür. Daha sonra map matristeki 1 olan gözlere bakılarak, hash matrisindeki en küçük değerler bulunur ve bu değer signature matrisin ilgili gözüne yazılır.
12. Bulunan imza matrisi imza benzerliğinin hesaplanması için `jaccard_by_signature` fonksiyonuna gönderilir. İmza benzerliklerinin bulunması, bu matrisin sütunlarının karşılaştırılması ile yapılır.
13. Elde edilen imza benzerlik matrisi ekrana yazdırılır.
14. `print_document_pairs` fonksiyonu, eşik değeri olarak 0.7, 0.8 ve 0.9'u alarak, shingle benzerliği ve imza benzerliği bu değerlerden fazla olan doküman çiftlerini ekrana yazdırır.
15. Kullanıcı isterse yeni bir K değeri girerek test etmeye devam edebilir.
16. K değeri '0' girilirse programdan çıkarılır.
17. $K = 4, 5, 8$ için; dokümanların unique shingle sayıları, toplam unique shingle sayıları, jaccard benzerlik tablosu, imza benzerlik tablosu ve threshold değerlerine(0.7, 0.8, 0.9) göre test sonuçları bölümünde detaylı olarak görülebilmektedir.

SONUÇ

1. K değeri 4, 5, 8 olarak artarken; dokümanların benzerlik oranlarının azaldığını gördüm. 4 ve 5 arasında bariz farklar olmasa da 4 ve 8, 5 ve 8 arasında %8-%9'lara varan benzerlik farklılıklarını gördüm.
2. Bence $K = 4$ için en doğru sonuçlar alındı. Çünkü çoğu doküman neredeyse birbirinin aynısı. K değeri büyündükçe elde ettiğimiz shingle kümelerinin çözünürlüğünün azaldığı kanısındayım. Ancak K değerini 4'ten aşağıya çekmek kelimelerin genel yapısı için uyumsuz olabilir, kelimeler genel olarak 4 harften büyiktür.
3. K-Shingle benzerliği, çoğu durumda imza matrisi benzerliğinden daha başarılı duruyor. Çünkü imza benzerliği hesaplarken rastgele erişimler yaptığımız için veri kaybı olması mümkün. Bu da benzerlik oranlarının değişmesine yol açar.

```
C:\Users\sinan.keskin\Desktop\algo\algo_proje\cmake-build-debug\algo_proje.exe
Isleme devam etmek icin K sayisini giriniz, cikmak icin '0' giriniz
```

```
K====>4
4
Dokuman sayisini giriniz=>
20
20

=====
K = 4 icin dokumanlarin unique shingle sayiları ==>

Dokuman => 1, unique shingle =>420
Dokuman => 2, unique shingle =>421
Dokuman => 3, unique shingle =>418
Dokuman => 4, unique shingle =>423
Dokuman => 5, unique shingle =>413
Dokuman => 6, unique shingle =>425
Dokuman => 7, unique shingle =>423
Dokuman => 8, unique shingle =>420
Dokuman => 9, unique shingle =>421
Dokuman => 10, unique shingle =>422
Dokuman => 11, unique shingle =>420
Dokuman => 12, unique shingle =>422
Dokuman => 13, unique shingle =>427
Dokuman => 14, unique shingle =>428
Dokuman => 15, unique shingle =>422
Dokuman => 16, unique shingle =>425
Dokuman => 17, unique shingle =>429
Dokuman => 18, unique shingle =>420
Dokuman => 19, unique shingle =>418
Dokuman => 20, unique shingle =>391
Shingle tablolari yazilsin mi ? e/E/h/Hh
```

Jaccard Benzerligi, K = 4

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
1	1.00	0.96	0.89	0.87	0.84	0.96	0.94	0.91	0.89	0.88	1.00	0.91	0.87	0.84	0.81	0.88	0.79	0.76	0.95	0.89
2	0.96	1.00	0.93	0.91	0.88	0.91	0.90	0.87	0.85	0.84	0.96	0.87	0.83	0.81	0.78	0.77	0.76	0.74	0.91	0.85
3	0.89	0.93	1.00	0.96	0.94	0.85	0.84	0.81	0.79	0.78	0.89	0.81	0.77	0.75	0.72	0.71	0.71	0.68	0.85	0.78
4	0.87	0.91	0.96	1.00	0.97	0.84	0.82	0.80	0.78	0.76	0.87	0.79	0.77	0.74	0.72	0.71	0.70	0.68	0.82	0.76
5	0.84	0.88	0.94	0.97	1.00	0.81	0.80	0.77	0.76	0.74	0.84	0.77	0.74	0.72	0.69	0.68	0.68	0.66	0.80	0.74
6	0.96	0.91	0.85	0.84	0.81	1.00	0.99	0.96	0.94	0.92	0.96	0.87	0.83	0.81	0.78	0.77	0.76	0.73	0.91	0.85
7	0.94	0.90	0.84	0.82	0.80	0.99	1.00	0.97	0.95	0.93	0.94	0.86	0.82	0.80	0.77	0.76	0.75	0.72	0.90	0.83
8	0.91	0.87	0.81	0.80	0.77	0.96	0.97	1.00	0.98	0.96	0.91	0.83	0.79	0.77	0.75	0.74	0.73	0.71	0.87	0.81
9	0.89	0.85	0.79	0.78	0.76	0.94	0.95	0.98	1.00	0.98	0.89	0.82	0.78	0.75	0.73	0.72	0.71	0.70	0.85	0.79
10	0.88	0.84	0.78	0.76	0.74	0.92	0.93	0.96	0.98	1.00	0.88	0.80	0.76	0.74	0.72	0.70	0.70	0.68	0.83	0.77
11	1.00	0.96	0.89	0.87	0.84	0.96	0.94	0.91	0.89	0.88	1.00	0.91	0.87	0.84	0.81	0.80	0.79	0.76	0.95	0.89
12	0.91	0.87	0.81	0.79	0.77	0.87	0.86	0.83	0.82	0.80	0.91	1.00	0.95	0.92	0.89	0.88	0.87	0.84	0.90	0.84
13	0.87	0.83	0.77	0.77	0.74	0.83	0.82	0.79	0.76	0.87	0.95	1.00	0.97	0.94	0.92	0.91	0.88	0.86	0.79	
14	0.84	0.81	0.75	0.74	0.72	0.81	0.80	0.77	0.75	0.74	0.84	0.92	0.97	1.00	0.97	0.95	0.94	0.91	0.83	0.77
15	0.81	0.78	0.72	0.72	0.69	0.78	0.77	0.75	0.73	0.72	0.81	0.89	0.94	0.97	1.00	0.98	0.97	0.94	0.80	0.78
16	0.80	0.77	0.71	0.71	0.68	0.77	0.76	0.74	0.74	0.70	0.80	0.88	0.92	0.95	0.98	1.00	0.99	0.96	0.79	0.77
17	0.79	0.76	0.71	0.70	0.68	0.76	0.75	0.73	0.71	0.70	0.79	0.87	0.91	0.94	0.97	0.99	1.00	0.97	0.78	0.76
18	0.76	0.74	0.68	0.68	0.66	0.73	0.72	0.71	0.70	0.68	0.76	0.84	0.88	0.91	0.94	0.96	0.97	1.00	0.75	0.73
19	0.95	0.91	0.85	0.82	0.80	0.91	0.90	0.87	0.85	0.83	0.95	0.90	0.86	0.83	0.80	0.79	0.78	0.75	1.00	0.93
20	0.89	0.85	0.78	0.76	0.74	0.85	0.83	0.81	0.79	0.77	0.89	0.84	0.79	0.77	0.78	0.77	0.76	0.73	0.93	1.00

=====TOPLAM UNIQUE SHINGLE SAYISI=>540

Merge edilmiş unique shingle tablosu yazılın mı? e/E/h/Hh

İmza Benzerligi, K = 4

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
1	1.00	0.96	0.93	0.89	0.82	0.98	0.98	0.97	0.95	0.95	1.00	0.97	0.97	0.95	0.93	0.92	0.92	0.90	0.99	0.94
2	0.96	1.00	0.96	0.92	0.85	0.94	0.94	0.93	0.91	0.91	0.96	0.94	0.94	0.92	0.90	0.89	0.89	0.87	0.95	0.90
3	0.93	0.96	1.00	0.99	0.86	0.91	0.91	0.90	0.88	0.88	0.93	0.91	0.91	0.89	0.87	0.86	0.86	0.84	0.92	0.87
4	0.89	0.92	0.93	1.00	0.93	0.89	0.89	0.88	0.86	0.86	0.89	0.87	0.87	0.85	0.83	0.82	0.82	0.80	0.88	0.83

5| 0.82|0.85|0.86|0.93|1.00|0.82|0.82|0.81|0.81|0.81|0.82|0.80|0.80|0.78|0.76|0.75|0.75|0.73|0.81|0.76|
6| 0.98|0.94|0.91|0.89|0.82|1.00|1.00|0.99|0.97|0.97|0.98|0.95|0.95|0.93|0.91|0.90|0.90|0.88|0.97|0.92|
7| 0.98|0.94|0.91|0.89|0.82|1.00|1.00|0.99|0.97|0.97|0.98|0.95|0.95|0.93|0.91|0.90|0.90|0.88|0.97|0.92|
8| 0.97|0.93|0.90|0.88|0.81|0.99|0.99|1.00|0.98|0.98|0.97|0.94|0.94|0.92|0.90|0.89|0.89|0.87|0.96|0.91|
9| 0.95|0.91|0.88|0.86|0.81|0.97|0.97|0.98|1.00|1.00|0.95|0.92|0.92|0.90|0.88|0.87|0.87|0.85|0.94|0.89|
10| 0.95|0.91|0.88|0.86|0.81|0.97|0.97|0.98|1.00|1.00|0.95|0.92|0.92|0.90|0.88|0.87|0.87|0.85|0.94|0.89|
11| 1.00|0.96|0.93|0.89|0.82|0.98|0.98|0.97|0.95|0.95|1.00|0.97|0.97|0.95|0.93|0.92|0.92|0.90|0.99|0.94|
12| 0.97|0.94|0.91|0.87|0.80|0.95|0.95|0.94|0.92|0.92|0.97|1.00|1.00|0.98|0.96|0.95|0.95|0.93|0.96|0.92|
13| 0.97|0.94|0.91|0.87|0.80|0.95|0.95|0.94|0.92|0.92|0.97|1.00|1.00|0.98|0.96|0.95|0.95|0.93|0.96|0.92|
14| 0.95|0.92|0.89|0.85|0.78|0.93|0.93|0.92|0.90|0.90|0.95|0.98|0.98|1.00|0.98|0.97|0.97|0.95|0.94|0.90|
15| 0.93|0.90|0.87|0.83|0.76|0.91|0.91|0.90|0.88|0.88|0.93|0.96|0.96|0.98|1.00|0.99|0.99|0.97|0.92|0.89|
16| 0.92|0.89|0.86|0.82|0.75|0.90|0.90|0.89|0.87|0.87|0.92|0.95|0.95|0.97|0.99|1.00|1.00|0.98|0.91|0.88|
17| 0.92|0.89|0.86|0.82|0.75|0.90|0.90|0.89|0.87|0.87|0.92|0.95|0.95|0.97|0.99|1.00|1.00|0.98|0.91|0.88|
18| 0.90|0.87|0.84|0.80|0.73|0.88|0.88|0.87|0.85|0.85|0.90|0.93|0.93|0.95|0.97|0.98|0.98|1.00|0.89|0.86|
19| 0.99|0.95|0.92|0.88|0.81|0.97|0.96|0.94|0.94|0.99|0.96|0.96|0.96|0.94|0.92|0.91|0.91|0.89|1.00|0.95|
20| 0.94|0.90|0.87|0.83|0.76|0.92|0.92|0.91|0.89|0.89|0.94|0.92|0.92|0.90|0.89|0.88|0.88|0.86|0.95|1.00|

K-SHINGLE BENZERLİKLERİNE GORE DOKUMAN ÇİFTLERİ

Threshold 0.70 icin Benzerlik orani esik degerinden fazla olan dokuman ciftleri==>

dokuman(1, 2) benzerlik orani=> 0.96	dokuman(1, 3) benzerlik orani=> 0.89	dokuman(1, 4) benzerlik orani=> 0.87
dokuman(1, 5) benzerlik orani=> 0.84	dokuman(1, 6) benzerlik orani=> 0.96	dokuman(1, 7) benzerlik orani=> 0.94
dokuman(1, 8) benzerlik orani=> 0.91	dokuman(1, 9) benzerlik orani=> 0.89	dokuman(1, 10) benzerlik orani=> 0.88
dokuman(1, 11) benzerlik orani=> 1.00	dokuman(1, 12) benzerlik orani=> 0.91	dokuman(1, 13) benzerlik orani=> 0.87
dokuman(1, 14) benzerlik orani=> 0.84	dokuman(1, 15) benzerlik orani=> 0.81	dokuman(1, 16) benzerlik orani=> 0.80
dokuman(1, 17) benzerlik orani=> 0.79	dokuman(1, 18) benzerlik orani=> 0.76	dokuman(1, 19) benzerlik orani=> 0.95
dokuman(1, 20) benzerlik orani=> 0.89	dokuman(2, 3) benzerlik orani=> 0.93	dokuman(2, 4) benzerlik orani=> 0.91
dokuman(2, 5) benzerlik orani=> 0.88	dokuman(2, 6) benzerlik orani=> 0.91	dokuman(2, 7) benzerlik orani=> 0.90
dokuman(2, 8) benzerlik orani=> 0.87	dokuman(2, 9) benzerlik orani=> 0.85	dokuman(2, 10) benzerlik orani=> 0.84
dokuman(2, 11) benzerlik orani=> 0.96	dokuman(2, 12) benzerlik orani=> 0.87	dokuman(2, 13) benzerlik orani=> 0.83
dokuman(2, 14) benzerlik orani=> 0.81	dokuman(2, 15) benzerlik orani=> 0.78	dokuman(2, 16) benzerlik orani=> 0.77
dokuman(2, 17) benzerlik orani=> 0.76	dokuman(2, 18) benzerlik orani=> 0.74	dokuman(2, 19) benzerlik orani=> 0.91
dokuman(2, 20) benzerlik orani=> 0.85	dokuman(3, 4) benzerlik orani=> 0.96	dokuman(3, 5) benzerlik orani=> 0.94
dokuman(3, 6) benzerlik orani=> 0.85	dokuman(3, 7) benzerlik orani=> 0.84	dokuman(3, 8) benzerlik orani=> 0.81
dokuman(3, 9) benzerlik orani=> 0.79	dokuman(3, 10) benzerlik orani=> 0.78	dokuman(3, 11) benzerlik orani=> 0.89
dokuman(3, 12) benzerlik orani=> 0.81	dokuman(3, 13) benzerlik orani=> 0.77	dokuman(3, 14) benzerlik orani=> 0.75
dokuman(3, 15) benzerlik orani=> 0.72	dokuman(3, 16) benzerlik orani=> 0.71	dokuman(3, 17) benzerlik orani=> 0.71
dokuman(3, 19) benzerlik orani=> 0.85	dokuman(3, 20) benzerlik orani=> 0.78	dokuman(4, 5) benzerlik orani=> 0.97
dokuman(4, 6) benzerlik orani=> 0.84	dokuman(4, 7) benzerlik orani=> 0.82	dokuman(4, 8) benzerlik orani=> 0.80
dokuman(4, 9) benzerlik orani=> 0.78	dokuman(4, 10) benzerlik orani=> 0.76	dokuman(4, 11) benzerlik orani=> 0.87
dokuman(4, 12) benzerlik orani=> 0.79	dokuman(4, 13) benzerlik orani=> 0.77	dokuman(4, 14) benzerlik orani=> 0.74
dokuman(4, 15) benzerlik orani=> 0.72	dokuman(4, 16) benzerlik orani=> 0.71	dokuman(4, 17) benzerlik orani=> 0.70
dokuman(4, 19) benzerlik orani=> 0.82	dokuman(4, 20) benzerlik orani=> 0.76	dokuman(5, 6) benzerlik orani=> 0.81
dokuman(5, 7) benzerlik orani=> 0.80	dokuman(5, 8) benzerlik orani=> 0.77	dokuman(5, 9) benzerlik orani=> 0.76
dokuman(5, 10) benzerlik orani=> 0.74	dokuman(5, 11) benzerlik orani=> 0.84	dokuman(5, 12) benzerlik orani=> 0.77
dokuman(5, 13) benzerlik orani=> 0.74	dokuman(5, 14) benzerlik orani=> 0.72	dokuman(5, 15) benzerlik orani=> 0.80
dokuman(5, 20) benzerlik orani=> 0.74	dokuman(6, 7) benzerlik orani=> 0.99	dokuman(6, 8) benzerlik orani=> 0.96
dokuman(6, 9) benzerlik orani=> 0.94	dokuman(6, 10) benzerlik orani=> 0.92	dokuman(6, 11) benzerlik orani=> 0.96
dokuman(6, 12) benzerlik orani=> 0.87	dokuman(6, 13) benzerlik orani=> 0.83	dokuman(6, 14) benzerlik orani=> 0.81
dokuman(6, 15) benzerlik orani=> 0.78	dokuman(6, 16) benzerlik orani=> 0.77	dokuman(6, 17) benzerlik orani=> 0.76
dokuman(6, 18) benzerlik orani=> 0.73	dokuman(6, 19) benzerlik orani=> 0.91	dokuman(6, 20) benzerlik orani=> 0.85
dokuman(7, 8) benzerlik orani=> 0.97	dokuman(7, 9) benzerlik orani=> 0.95	dokuman(7, 10) benzerlik orani=> 0.93
dokuman(7, 11) benzerlik orani=> 0.94	dokuman(7, 12) benzerlik orani=> 0.86	dokuman(7, 13) benzerlik orani=> 0.82
dokuman(7, 14) benzerlik orani=> 0.80	dokuman(7, 15) benzerlik orani=> 0.77	dokuman(7, 16) benzerlik orani=> 0.76
dokuman(7, 17) benzerlik orani=> 0.75	dokuman(7, 18) benzerlik orani=> 0.72	dokuman(7, 19) benzerlik orani=> 0.90
dokuman(7, 20) benzerlik orani=> 0.83	dokuman(8, 9) benzerlik orani=> 0.98	dokuman(8, 10) benzerlik orani=> 0.96
dokuman(8, 11) benzerlik orani=> 0.91	dokuman(8, 12) benzerlik orani=> 0.83	dokuman(8, 13) benzerlik orani=> 0.79
dokuman(8, 14) benzerlik orani=> 0.77	dokuman(8, 15) benzerlik orani=> 0.75	dokuman(8, 16) benzerlik orani=> 0.74
dokuman(8, 17) benzerlik orani=> 0.73	dokuman(8, 18) benzerlik orani=> 0.71	dokuman(8, 19) benzerlik orani=> 0.87
dokuman(8, 20) benzerlik orani=> 0.81	dokuman(9, 10) benzerlik orani=> 0.98	dokuman(9, 11) benzerlik orani=> 0.89
dokuman(9, 12) benzerlik orani=> 0.82	dokuman(9, 13) benzerlik orani=> 0.78	dokuman(9, 14) benzerlik orani=> 0.75
dokuman(9, 15) benzerlik orani=> 0.73	dokuman(9, 16) benzerlik orani=> 0.72	dokuman(9, 17) benzerlik orani=> 0.71
dokuman(9, 19) benzerlik orani=> 0.85	dokuman(9, 20) benzerlik orani=> 0.79	dokuman(10, 11) benzerlik orani=> 0.88
dokuman(10, 12) benzerlik orani=> 0.80	dokuman(10, 13) benzerlik orani=> 0.76	dokuman(10, 14) benzerlik orani=> 0.74
dokuman(10, 15) benzerlik orani=> 0.72	dokuman(10, 16) benzerlik orani=> 0.70	dokuman(10, 19) benzerlik orani=> 0.83
dokuman(10, 20) benzerlik orani=> 0.77	dokuman(11, 12) benzerlik orani=> 0.91	dokuman(11, 13) benzerlik orani=> 0.87
dokuman(11, 14) benzerlik orani=> 0.84	dokuman(11, 15) benzerlik orani=> 0.81	dokuman(11, 16) benzerlik orani=> 0.80
dokuman(11, 17) benzerlik orani=> 0.79	dokuman(11, 18) benzerlik orani=> 0.76	dokuman(11, 19) benzerlik orani=> 0.95
dokuman(11, 20) benzerlik orani=> 0.89	dokuman(12, 13) benzerlik orani=> 0.95	dokuman(12, 14) benzerlik orani=> 0.92
dokuman(12, 15) benzerlik orani=> 0.89	dokuman(12, 16) benzerlik orani=> 0.88	dokuman(12, 17) benzerlik orani=> 0.87
dokuman(12, 18) benzerlik orani=> 0.84	dokuman(12, 19) benzerlik orani=> 0.90	dokuman(12, 20) benzerlik orani=> 0.84
dokuman(13, 14) benzerlik orani=> 0.97	dokuman(13, 15) benzerlik orani=> 0.94	dokuman(13, 16) benzerlik orani=> 0.92
dokuman(13, 17) benzerlik orani=> 0.91	dokuman(13, 18) benzerlik orani=> 0.88	dokuman(13, 19) benzerlik orani=> 0.86
dokuman(13, 20) benzerlik orani=> 0.79	dokuman(14, 15) benzerlik orani=> 0.97	dokuman(14, 16) benzerlik orani=> 0.95
dokuman(14, 17) benzerlik orani=> 0.94	dokuman(14, 18) benzerlik orani=> 0.91	dokuman(14, 19) benzerlik orani=> 0.83
dokuman(14, 20) benzerlik orani=> 0.77	dokuman(15, 16) benzerlik orani=> 0.98	dokuman(15, 17) benzerlik orani=> 0.97
dokuman(15, 18) benzerlik orani=> 0.94	dokuman(15, 19) benzerlik orani=> 0.80	dokuman(15, 20) benzerlik orani=> 0.78
dokuman(16, 17) benzerlik orani=> 0.99	dokuman(16, 18) benzerlik orani=> 0.96	dokuman(16, 19) benzerlik orani=> 0.79
dokuman(16, 20) benzerlik orani=> 0.77	dokuman(17, 18) benzerlik orani=> 0.97	dokuman(17, 19) benzerlik orani=> 0.78
dokuman(17, 20) benzerlik orani=> 0.76	dokuman(18, 19) benzerlik orani=> 0.75	dokuman(18, 20) benzerlik orani=> 0.73

Threshold 0.80 icin Benzerlik orani esik degerinden fazla olan dokuman ciftleri==>

dokuman(1, 2) benzerlik orani=> 0.96
dokuman(1, 5) benzerlik orani=> 0.84
dokuman(1, 8) benzerlik orani=> 0.91
dokuman(1, 11) benzerlik orani=> 1.00
dokuman(1, 14) benzerlik orani=> 0.84
dokuman(1, 19) benzerlik orani=> 0.95
dokuman(2, 4) benzerlik orani=> 0.91
dokuman(2, 7) benzerlik orani=> 0.90
dokuman(2, 10) benzerlik orani=> 0.84
dokuman(2, 13) benzerlik orani=> 0.83
dokuman(2, 20) benzerlik orani=> 0.85
dokuman(3, 6) benzerlik orani=> 0.85
dokuman(3, 11) benzerlik orani=> 0.89
dokuman(4, 5) benzerlik orani=> 0.97
dokuman(4, 11) benzerlik orani=> 0.87
dokuman(5, 11) benzerlik orani=> 0.84
dokuman(6, 9) benzerlik orani=> 0.94
dokuman(6, 12) benzerlik orani=> 0.87
dokuman(6, 19) benzerlik orani=> 0.91
dokuman(7, 9) benzerlik orani=> 0.95
dokuman(7, 12) benzerlik orani=> 0.86
dokuman(7, 20) benzerlik orani=> 0.83
dokuman(8, 11) benzerlik orani=> 0.91
dokuman(8, 20) benzerlik orani=> 0.81
dokuman(9, 12) benzerlik orani=> 0.82
dokuman(10, 19) benzerlik orani=> 0.83
dokuman(11, 14) benzerlik orani=> 0.84
dokuman(11, 19) benzerlik orani=> 0.95
dokuman(12, 14) benzerlik orani=> 0.92
dokuman(12, 17) benzerlik orani=> 0.87
dokuman(12, 20) benzerlik orani=> 0.84
dokuman(13, 16) benzerlik orani=> 0.92
dokuman(13, 19) benzerlik orani=> 0.86
dokuman(14, 17) benzerlik orani=> 0.94
dokuman(15, 16) benzerlik orani=> 0.98
dokuman(15, 19) benzerlik orani=> 0.88
dokuman(17, 18) benzerlik orani=> 0.97
dokuman(1, 3) benzerlik orani=> 0.89
dokuman(1, 6) benzerlik orani=> 0.96
dokuman(1, 9) benzerlik orani=> 0.89
dokuman(1, 12) benzerlik orani=> 0.91
dokuman(1, 15) benzerlik orani=> 0.81
dokuman(1, 20) benzerlik orani=> 0.89
dokuman(2, 5) benzerlik orani=> 0.88
dokuman(2, 8) benzerlik orani=> 0.87
dokuman(2, 11) benzerlik orani=> 0.96
dokuman(2, 14) benzerlik orani=> 0.81
dokuman(3, 4) benzerlik orani=> 0.96
dokuman(3, 7) benzerlik orani=> 0.84
dokuman(3, 12) benzerlik orani=> 0.81
dokuman(4, 6) benzerlik orani=> 0.84
dokuman(4, 19) benzerlik orani=> 0.82
dokuman(6, 7) benzerlik orani=> 0.99
dokuman(6, 18) benzerlik orani=> 0.92
dokuman(6, 13) benzerlik orani=> 0.83
dokuman(6, 20) benzerlik orani=> 0.85
dokuman(7, 10) benzerlik orani=> 0.93
dokuman(7, 13) benzerlik orani=> 0.82
dokuman(8, 9) benzerlik orani=> 0.98
dokuman(8, 12) benzerlik orani=> 0.83
dokuman(9, 10) benzerlik orani=> 0.98
dokuman(9, 19) benzerlik orani=> 0.85
dokuman(11, 12) benzerlik orani=> 0.91
dokuman(11, 15) benzerlik orani=> 0.81
dokuman(11, 20) benzerlik orani=> 0.89
dokuman(12, 15) benzerlik orani=> 0.89
dokuman(12, 18) benzerlik orani=> 0.84
dokuman(13, 14) benzerlik orani=> 0.97
dokuman(13, 17) benzerlik orani=> 0.91
dokuman(14, 15) benzerlik orani=> 0.97
dokuman(14, 18) benzerlik orani=> 0.91
dokuman(15, 17) benzerlik orani=> 0.97
dokuman(16, 17) benzerlik orani=> 0.97
dokuman(19, 20) benzerlik orani=> 0.93
dokuman(1, 4) benzerlik orani=> 0.87
dokuman(1, 7) benzerlik orani=> 0.94
dokuman(1, 10) benzerlik orani=> 0.88
dokuman(1, 13) benzerlik orani=> 0.87
dokuman(1, 16) benzerlik orani=> 0.80
dokuman(2, 3) benzerlik orani=> 0.93
dokuman(2, 6) benzerlik orani=> 0.91
dokuman(2, 9) benzerlik orani=> 0.85
dokuman(2, 12) benzerlik orani=> 0.87
dokuman(2, 19) benzerlik orani=> 0.91
dokuman(3, 5) benzerlik orani=> 0.94
dokuman(3, 8) benzerlik orani=> 0.81
dokuman(3, 19) benzerlik orani=> 0.85
dokuman(4, 7) benzerlik orani=> 0.82
dokuman(5, 6) benzerlik orani=> 0.81
dokuman(6, 8) benzerlik orani=> 0.96
dokuman(6, 11) benzerlik orani=> 0.96
dokuman(6, 14) benzerlik orani=> 0.81
dokuman(7, 8) benzerlik orani=> 0.97
dokuman(7, 11) benzerlik orani=> 0.94
dokuman(7, 19) benzerlik orani=> 0.90
dokuman(8, 10) benzerlik orani=> 0.96
dokuman(8, 19) benzerlik orani=> 0.87
dokuman(9, 11) benzerlik orani=> 0.89
dokuman(10, 11) benzerlik orani=> 0.88
dokuman(11, 13) benzerlik orani=> 0.87
dokuman(11, 16) benzerlik orani=> 0.80
dokuman(12, 13) benzerlik orani=> 0.95
dokuman(12, 16) benzerlik orani=> 0.88
dokuman(12, 19) benzerlik orani=> 0.90
dokuman(13, 15) benzerlik orani=> 0.94
dokuman(13, 18) benzerlik orani=> 0.88
dokuman(14, 16) benzerlik orani=> 0.95
dokuman(14, 19) benzerlik orani=> 0.83
dokuman(15, 18) benzerlik orani=> 0.94
dokuman(16, 18) benzerlik orani=> 0.96

Threshold 0.90 icin Benzerlik orani esik degerinden fazla olan dokuman ciftleri==>

dokuman(1, 2) benzerlik orani=> 0.96
dokuman(1, 8) benzerlik orani=> 0.91
dokuman(1, 19) benzerlik orani=> 0.95
dokuman(2, 6) benzerlik orani=> 0.91
dokuman(2, 19) benzerlik orani=> 0.91
dokuman(4, 5) benzerlik orani=> 0.97
dokuman(6, 9) benzerlik orani=> 0.94
dokuman(6, 19) benzerlik orani=> 0.91
dokuman(7, 10) benzerlik orani=> 0.93
dokuman(8, 10) benzerlik orani=> 0.96
dokuman(11, 12) benzerlik orani=> 0.91
dokuman(12, 14) benzerlik orani=> 0.92
dokuman(13, 15) benzerlik orani=> 0.94
dokuman(14, 15) benzerlik orani=> 0.97
dokuman(14, 18) benzerlik orani=> 0.91
dokuman(15, 18) benzerlik orani=> 0.94
dokuman(17, 18) benzerlik orani=> 0.97
dokuman(1, 6) benzerlik orani=> 0.96
dokuman(1, 11) benzerlik orani=> 1.00
dokuman(2, 3) benzerlik orani=> 0.93
dokuman(2, 7) benzerlik orani=> 0.90
dokuman(3, 4) benzerlik orani=> 0.96
dokuman(6, 7) benzerlik orani=> 0.99
dokuman(6, 10) benzerlik orani=> 0.92
dokuman(7, 8) benzerlik orani=> 0.97
dokuman(7, 11) benzerlik orani=> 0.94
dokuman(8, 11) benzerlik orani=> 0.91
dokuman(11, 19) benzerlik orani=> 0.95
dokuman(12, 19) benzerlik orani=> 0.90
dokuman(13, 16) benzerlik orani=> 0.92
dokuman(14, 16) benzerlik orani=> 0.95
dokuman(15, 16) benzerlik orani=> 0.98
dokuman(16, 17) benzerlik orani=> 0.99
dokuman(19, 20) benzerlik orani=> 0.93
dokuman(1, 7) benzerlik orani=> 0.94
dokuman(1, 12) benzerlik orani=> 0.91
dokuman(2, 4) benzerlik orani=> 0.91
dokuman(2, 11) benzerlik orani=> 0.96
dokuman(3, 5) benzerlik orani=> 0.94
dokuman(6, 8) benzerlik orani=> 0.96
dokuman(6, 11) benzerlik orani=> 0.96
dokuman(7, 9) benzerlik orani=> 0.95
dokuman(8, 9) benzerlik orani=> 0.98
dokuman(9, 10) benzerlik orani=> 0.98
dokuman(12, 13) benzerlik orani=> 0.95
dokuman(13, 14) benzerlik orani=> 0.97
dokuman(13, 17) benzerlik orani=> 0.91
dokuman(14, 17) benzerlik orani=> 0.94
dokuman(15, 17) benzerlik orani=> 0.97
dokuman(16, 18) benzerlik orani=> 0.96

TMZA BENZERI TKI ERTNE GORE DOKUMAN CTETI ERT

Threshold 0.70 icin Benzerlik orani esik degerinden fazla olan dokuman ciftleri==>

dokuman(1, 2)imza benzerligi=> 0.96
dokuman(1, 5)imza benzerligi=> 0.82
dokuman(1, 8)imza benzerligi=> 0.97
dokuman(1, 11)imza benzerligi=> 1.00
dokuman(1, 14)imza benzerligi=> 0.95
dokuman(1, 17)imza benzerligi=> 0.92
dokuman(1, 20)imza benzerligi=> 0.94
dokuman(2, 5)imza benzerligi=> 0.85
dokuman(2, 8)imza benzerligi=> 0.93
dokuman(2, 11)imza benzerligi=> 0.96
dokuman(2, 14)imza benzerligi=> 0.92
dokuman(2, 17)imza benzerligi=> 0.89
dokuman(2, 20)imza benzerligi=> 0.90
dokuman(3, 6)imza benzerligi=> 0.91
dokuman(3, 9)imza benzerligi=> 0.88
dokuman(3, 12)imza benzerligi=> 0.91
dokuman(3, 15)imza benzerligi=> 0.87
dokuman(3, 18)imza benzerligi=> 0.84
dokuman(4, 5)imza benzerligi=> 0.93
dokuman(4, 8)imza benzerligi=> 0.88
dokuman(4, 11)imza benzerligi=> 0.89
dokuman(4, 14)imza benzerligi=> 0.85
dokuman(4, 17)imza benzerligi=> 0.82
dokuman(4, 20)imza benzerligi=> 0.83
dokuman(5, 8)imza benzerligi=> 0.81
dokuman(5, 11)imza benzerligi=> 0.82
dokuman(5, 14)imza benzerligi=> 0.78
dokuman(5, 17)imza benzerligi=> 0.75
dokuman(5, 20)imza benzerligi=> 0.76
dokuman(6, 9)imza benzerligi=> 0.97
dokuman(6, 12)imza benzerligi=> 0.95
dokuman(6, 15)imza benzerligi=> 0.91
dokuman(6, 18)imza benzerligi=> 0.88
dokuman(7, 8)imza benzerligi=> 0.99
dokuman(7, 11)imza benzerligi=> 0.98
dokuman(7, 14)imza benzerligi=> 0.93
dokuman(7, 17)imza benzerligi=> 0.90
dokuman(1, 3)imza benzerligi=> 0.93
dokuman(1, 6)imza benzerligi=> 0.98
dokuman(1, 9)imza benzerligi=> 0.95
dokuman(1, 12)imza benzerligi=> 0.97
dokuman(1, 15)imza benzerligi=> 0.93
dokuman(1, 18)imza benzerligi=> 0.90
dokuman(2, 3)imza benzerligi=> 0.96
dokuman(2, 6)imza benzerligi=> 0.94
dokuman(2, 9)imza benzerligi=> 0.91
dokuman(2, 12)imza benzerligi=> 0.94
dokuman(2, 15)imza benzerligi=> 0.90
dokuman(2, 18)imza benzerligi=> 0.87
dokuman(3, 4)imza benzerligi=> 0.93
dokuman(3, 7)imza benzerligi=> 0.91
dokuman(3, 10)imza benzerligi=> 0.88
dokuman(3, 13)imza benzerligi=> 0.91
dokuman(3, 16)imza benzerligi=> 0.86
dokuman(3, 19)imza benzerligi=> 0.92
dokuman(4, 6)imza benzerligi=> 0.89
dokuman(4, 9)imza benzerligi=> 0.86
dokuman(4, 12)imza benzerligi=> 0.87
dokuman(4, 15)imza benzerligi=> 0.83
dokuman(4, 18)imza benzerligi=> 0.80
dokuman(5, 6)imza benzerligi=> 0.82
dokuman(5, 9)imza benzerligi=> 0.81
dokuman(5, 12)imza benzerligi=> 0.80
dokuman(5, 15)imza benzerligi=> 0.76
dokuman(5, 18)imza benzerligi=> 0.73
dokuman(6, 7)imza benzerligi=> 1.00
dokuman(6, 10)imza benzerligi=> 0.97
dokuman(6, 13)imza benzerligi=> 0.95
dokuman(6, 16)imza benzerligi=> 0.90
dokuman(6, 19)imza benzerligi=> 0.97
dokuman(7, 9)imza benzerligi=> 0.97
dokuman(7, 12)imza benzerligi=> 0.95
dokuman(7, 15)imza benzerligi=> 0.91
dokuman(7, 18)imza benzerligi=> 0.88
dokuman(1, 4)imza benzerligi=> 0.89
dokuman(1, 7)imza benzerligi=> 0.98
dokuman(1, 10)imza benzerligi=> 0.95
dokuman(1, 13)imza benzerligi=> 0.97
dokuman(1, 16)imza benzerligi=> 0.92
dokuman(1, 19)imza benzerligi=> 0.99
dokuman(2, 4)imza benzerligi=> 0.92
dokuman(2, 7)imza benzerligi=> 0.94
dokuman(2, 10)imza benzerligi=> 0.91
dokuman(2, 13)imza benzerligi=> 0.94
dokuman(2, 16)imza benzerligi=> 0.89
dokuman(2, 19)imza benzerligi=> 0.95
dokuman(3, 5)imza benzerligi=> 0.86
dokuman(3, 8)imza benzerligi=> 0.90
dokuman(3, 11)imza benzerligi=> 0.93
dokuman(3, 14)imza benzerligi=> 0.89
dokuman(3, 17)imza benzerligi=> 0.86
dokuman(3, 20)imza benzerligi=> 0.87
dokuman(4, 7)imza benzerligi=> 0.89
dokuman(4, 10)imza benzerligi=> 0.86
dokuman(4, 13)imza benzerligi=> 0.87
dokuman(4, 16)imza benzerligi=> 0.82
dokuman(4, 19)imza benzerligi=> 0.88
dokuman(5, 7)imza benzerligi=> 0.82
dokuman(5, 10)imza benzerligi=> 0.81
dokuman(5, 13)imza benzerligi=> 0.80
dokuman(5, 16)imza benzerligi=> 0.75
dokuman(5, 19)imza benzerligi=> 0.81
dokuman(6, 8)imza benzerligi=> 0.99
dokuman(6, 11)imza benzerligi=> 0.98
dokuman(6, 14)imza benzerligi=> 0.93
dokuman(6, 17)imza benzerligi=> 0.90
dokuman(6, 20)imza benzerligi=> 0.92
dokuman(7, 10)imza benzerligi=> 0.97
dokuman(7, 13)imza benzerligi=> 0.95
dokuman(7, 16)imza benzerligi=> 0.90
dokuman(7, 19)imza benzerligi=> 0.97

dokuman(7, 20)imza benzerligi=> 0.92
dokuman(8, 11)imza benzerligi=> 0.97
dokuman(8, 14)imza benzerligi=> 0.92
dokuman(8, 17)imza benzerligi=> 0.89
dokuman(8, 20)imza benzerligi=> 0.91
dokuman(9, 12)imza benzerligi=> 0.92
dokuman(9, 15)imza benzerligi=> 0.88
dokuman(9, 18)imza benzerligi=> 0.85
dokuman(10, 11)imza benzerligi=> 0.95
dokuman(10, 14)imza benzerligi=> 0.90
dokuman(10, 17)imza benzerligi=> 0.87
dokuman(10, 20)imza benzerligi=> 0.89
dokuman(11, 14)imza benzerligi=> 0.95
dokuman(11, 17)imza benzerligi=> 0.92
dokuman(11, 20)imza benzerligi=> 0.94
dokuman(12, 15)imza benzerligi=> 0.96
dokuman(12, 18)imza benzerligi=> 0.93
dokuman(13, 14)imza benzerligi=> 0.98
dokuman(13, 17)imza benzerligi=> 0.95
dokuman(13, 20)imza benzerligi=> 0.92
dokuman(14, 17)imza benzerligi=> 0.97
dokuman(14, 20)imza benzerligi=> 0.90
dokuman(15, 18)imza benzerligi=> 0.97
dokuman(16, 17)imza benzerligi=> 1.00
dokuman(16, 20)imza benzerligi=> 0.88
dokuman(17, 20)imza benzerligi=> 0.88
dokuman(19, 20)imza benzerligi=> 0.95
dokuman(8, 9)imza benzerligi=> 0.98
dokuman(8, 12)imza benzerligi=> 0.94
dokuman(8, 15)imza benzerligi=> 0.90
dokuman(8, 18)imza benzerligi=> 0.87
dokuman(9, 10)imza benzerligi=> 1.00
dokuman(9, 13)imza benzerligi=> 0.92
dokuman(9, 16)imza benzerligi=> 0.87
dokuman(9, 19)imza benzerligi=> 0.94
dokuman(10, 12)imza benzerligi=> 0.92
dokuman(10, 15)imza benzerligi=> 0.88
dokuman(10, 18)imza benzerligi=> 0.85
dokuman(11, 12)imza benzerligi=> 0.97
dokuman(11, 15)imza benzerligi=> 0.93
dokuman(11, 18)imza benzerligi=> 0.90
dokuman(12, 13)imza benzerligi=> 1.00
dokuman(12, 16)imza benzerligi=> 0.95
dokuman(12, 19)imza benzerligi=> 0.96
dokuman(13, 15)imza benzerligi=> 0.96
dokuman(13, 18)imza benzerligi=> 0.93
dokuman(14, 15)imza benzerligi=> 0.98
dokuman(14, 18)imza benzerligi=> 0.95
dokuman(15, 16)imza benzerligi=> 0.99
dokuman(15, 19)imza benzerligi=> 0.92
dokuman(16, 18)imza benzerligi=> 0.98
dokuman(17, 18)imza benzerligi=> 0.98
dokuman(18, 19)imza benzerligi=> 0.89
dokuman(8, 10)imza benzerligi=> 0.98
dokuman(8, 13)imza benzerligi=> 0.94
dokuman(8, 16)imza benzerligi=> 0.89
dokuman(8, 19)imza benzerligi=> 0.96
dokuman(9, 11)imza benzerligi=> 0.95
dokuman(9, 14)imza benzerligi=> 0.90
dokuman(9, 17)imza benzerligi=> 0.87
dokuman(9, 20)imza benzerligi=> 0.89
dokuman(10, 13)imza benzerligi=> 0.92
dokuman(10, 16)imza benzerligi=> 0.87
dokuman(10, 19)imza benzerligi=> 0.94
dokuman(11, 13)imza benzerligi=> 0.97
dokuman(11, 16)imza benzerligi=> 0.92
dokuman(11, 19)imza benzerligi=> 0.99
dokuman(12, 14)imza benzerligi=> 0.98
dokuman(12, 17)imza benzerligi=> 0.95
dokuman(12, 20)imza benzerligi=> 0.92
dokuman(13, 16)imza benzerligi=> 0.95
dokuman(13, 19)imza benzerligi=> 0.96
dokuman(14, 16)imza benzerligi=> 0.97
dokuman(14, 19)imza benzerligi=> 0.94
dokuman(15, 17)imza benzerligi=> 0.99
dokuman(15, 20)imza benzerligi=> 0.89
dokuman(16, 19)imza benzerligi=> 0.91
dokuman(17, 19)imza benzerligi=> 0.91
dokuman(18, 20)imza benzerligi=> 0.86

Threshold 0.80 icin Benzerlik orani esik degerinden fazla olan dokuman ciftleri==>

Threshold 0.90 icin Benzerlik orani esik degerinden fazla olan dokuman ciftleri==>

dokuman(1, 2)imza benzerligi=> 0.96	dokuman(1, 3)imza benzerligi=> 0.93	dokuman(1, 6)imza benzerligi=> 0.98
dokuman(1, 7)imza benzerligi=> 0.98	dokuman(1, 8)imza benzerligi=> 0.97	dokuman(1, 9)imza benzerligi=> 0.95
dokuman(1, 10)imza benzerligi=> 0.95	dokuman(1, 11)imza benzerligi=> 1.00	dokuman(1, 12)imza benzerligi=> 0.97
dokuman(1, 13)imza benzerligi=> 0.97	dokuman(1, 14)imza benzerligi=> 0.95	dokuman(1, 15)imza benzerligi=> 0.93
dokuman(1, 16)imza benzerligi=> 0.92	dokuman(1, 17)imza benzerligi=> 0.92	dokuman(1, 18)imza benzerligi=> 0.90
dokuman(1, 19)imza benzerligi=> 0.99	dokuman(1, 20)imza benzerligi=> 0.94	dokuman(2, 3)imza benzerligi=> 0.96

```

dokuman(2, 4)imza benzerligi=> 0.92
dokuman(2, 8)imza benzerligi=> 0.93
dokuman(2, 11)imza benzerligi=> 0.96
dokuman(2, 14)imza benzerligi=> 0.92
dokuman(2, 20)imza benzerligi=> 0.90
dokuman(3, 7)imza benzerligi=> 0.91
dokuman(3, 12)imza benzerligi=> 0.91
dokuman(4, 5)imza benzerligi=> 0.93
dokuman(6, 9)imza benzerligi=> 0.97
dokuman(6, 12)imza benzerligi=> 0.95
dokuman(6, 15)imza benzerligi=> 0.91
dokuman(6, 19)imza benzerligi=> 0.97
dokuman(7, 9)imza benzerligi=> 0.97
dokuman(7, 12)imza benzerligi=> 0.95
dokuman(7, 15)imza benzerligi=> 0.91
dokuman(7, 19)imza benzerligi=> 0.97
dokuman(8, 10)imza benzerligi=> 0.98
dokuman(8, 13)imza benzerligi=> 0.94
dokuman(8, 19)imza benzerligi=> 0.96
dokuman(9, 11)imza benzerligi=> 0.95
dokuman(9, 14)imza benzerligi=> 0.90
dokuman(10, 12)imza benzerligi=> 0.92
dokuman(10, 19)imza benzerligi=> 0.94
dokuman(11, 14)imza benzerligi=> 0.95
dokuman(11, 17)imza benzerligi=> 0.92
dokuman(11, 20)imza benzerligi=> 0.94
dokuman(12, 15)imza benzerligi=> 0.96
dokuman(12, 18)imza benzerligi=> 0.93
dokuman(13, 14)imza benzerligi=> 0.98
dokuman(13, 17)imza benzerligi=> 0.95
dokuman(13, 20)imza benzerligi=> 0.92
dokuman(14, 17)imza benzerligi=> 0.97
dokuman(14, 20)imza benzerligi=> 0.90
dokuman(15, 18)imza benzerligi=> 0.97
dokuman(16, 18)imza benzerligi=> 0.98
dokuman(17, 19)imza benzerligi=> 0.91
dokuman(2, 6)imza benzerligi=> 0.94
dokuman(2, 9)imza benzerligi=> 0.91
dokuman(2, 12)imza benzerligi=> 0.94
dokuman(2, 15)imza benzerligi=> 0.90
dokuman(3, 4)imza benzerligi=> 0.93
dokuman(3, 8)imza benzerligi=> 0.90
dokuman(3, 13)imza benzerligi=> 0.91
dokuman(6, 7)imza benzerligi=> 1.00
dokuman(6, 10)imza benzerligi=> 0.97
dokuman(6, 13)imza benzerligi=> 0.95
dokuman(6, 16)imza benzerligi=> 0.90
dokuman(6, 20)imza benzerligi=> 0.92
dokuman(7, 10)imza benzerligi=> 0.97
dokuman(7, 13)imza benzerligi=> 0.95
dokuman(7, 16)imza benzerligi=> 0.90
dokuman(7, 20)imza benzerligi=> 0.92
dokuman(8, 11)imza benzerligi=> 0.97
dokuman(8, 14)imza benzerligi=> 0.92
dokuman(8, 20)imza benzerligi=> 0.91
dokuman(9, 12)imza benzerligi=> 0.92
dokuman(9, 19)imza benzerligi=> 0.94
dokuman(10, 13)imza benzerligi=> 0.92
dokuman(11, 12)imza benzerligi=> 0.97
dokuman(11, 15)imza benzerligi=> 0.93
dokuman(11, 18)imza benzerligi=> 0.90
dokuman(12, 13)imza benzerligi=> 1.00
dokuman(12, 16)imza benzerligi=> 0.95
dokuman(12, 19)imza benzerligi=> 0.96
dokuman(13, 15)imza benzerligi=> 0.96
dokuman(13, 18)imza benzerligi=> 0.93
dokuman(14, 15)imza benzerligi=> 0.98
dokuman(14, 18)imza benzerligi=> 0.95
dokuman(15, 16)imza benzerligi=> 0.99
dokuman(15, 19)imza benzerligi=> 0.92
dokuman(16, 19)imza benzerligi=> 0.91
dokuman(19, 20)imza benzerligi=> 0.95
dokuman(2, 7)imza benzerligi=> 0.94
dokuman(2, 10)imza benzerligi=> 0.91
dokuman(2, 13)imza benzerligi=> 0.94
dokuman(2, 19)imza benzerligi=> 0.95
dokuman(3, 6)imza benzerligi=> 0.91
dokuman(3, 11)imza benzerligi=> 0.93
dokuman(3, 19)imza benzerligi=> 0.92
dokuman(6, 8)imza benzerligi=> 0.99
dokuman(6, 11)imza benzerligi=> 0.98
dokuman(6, 14)imza benzerligi=> 0.93
dokuman(6, 17)imza benzerligi=> 0.90
dokuman(7, 8)imza benzerligi=> 0.99
dokuman(7, 11)imza benzerligi=> 0.98
dokuman(7, 14)imza benzerligi=> 0.93
dokuman(7, 17)imza benzerligi=> 0.90
dokuman(8, 9)imza benzerligi=> 0.98
dokuman(8, 12)imza benzerligi=> 0.94
dokuman(8, 15)imza benzerligi=> 0.90
dokuman(9, 18)imza benzerligi=> 1.00
dokuman(9, 13)imza benzerligi=> 0.92
dokuman(10, 11)imza benzerligi=> 0.95
dokuman(10, 14)imza benzerligi=> 0.90
dokuman(11, 13)imza benzerligi=> 0.97
dokuman(11, 16)imza benzerligi=> 0.92
dokuman(11, 19)imza benzerligi=> 0.99
dokuman(12, 14)imza benzerligi=> 0.98
dokuman(12, 17)imza benzerligi=> 0.95
dokuman(12, 20)imza benzerligi=> 0.92
dokuman(13, 16)imza benzerligi=> 0.95
dokuman(13, 19)imza benzerligi=> 0.96
dokuman(14, 16)imza benzerligi=> 0.97
dokuman(14, 19)imza benzerligi=> 0.94
dokuman(15, 17)imza benzerligi=> 0.99
dokuman(16, 17)imza benzerligi=> 1.00
dokuman(17, 18)imza benzerligi=> 0.98

```

K====>5
5

=====
K = 5 için dokumanların unique shingle sayıları ==>

```

Dokuman => 1, unique shingle =>478
Dokuman => 2, unique shingle =>479
Dokuman => 3, unique shingle =>472
Dokuman => 4, unique shingle =>476
Dokuman => 5, unique shingle =>465
Dokuman => 6, unique shingle =>483
Dokuman => 7, unique shingle =>480
Dokuman => 8, unique shingle =>477
Dokuman => 9, unique shingle =>478
Dokuman => 10, unique shingle =>478
Dokuman => 11, unique shingle =>478
Dokuman => 12, unique shingle =>480
Dokuman => 13, unique shingle =>484
Dokuman => 14, unique shingle =>484
Dokuman => 15, unique shingle =>477
Dokuman => 16, unique shingle =>481
Dokuman => 17, unique shingle =>485
Dokuman => 18, unique shingle =>472
Dokuman => 19, unique shingle =>475
Dokuman => 20, unique shingle =>447
Shingle tablolari yazilsin mi ? e/E/h/Hh
h

```

Jaccard Benzerligi, K = 5

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
<hr/>																			
1 1.00 0.95 0.87 0.85 0.82 0.95 0.94 0.91 0.89 0.86 1.00 0.90 0.85 0.82 0.79 0.77 0.76 0.73 0.95 0.89																			
2 0.95 1.00 0.92 0.89 0.87 0.90 0.89 0.86 0.84 0.82 0.95 0.85 0.80 0.78 0.75 0.73 0.72 0.69 0.90 0.84																			
3 0.87 0.92 1.00 0.95 0.92 0.83 0.82 0.79 0.77 0.75 0.87 0.79 0.74 0.71 0.69 0.67 0.66 0.63 0.83 0.77																			
4 0.85 0.89 0.95 1.00 0.97 0.81 0.80 0.77 0.75 0.73 0.85 0.76 0.73 0.70 0.67 0.66 0.65 0.62 0.81 0.75																			
5 0.82 0.87 0.92 0.97 1.00 0.79 0.77 0.75 0.73 0.71 0.82 0.74 0.70 0.68 0.65 0.64 0.63 0.60 0.78 0.72																			
6 0.95 0.90 0.83 0.81 0.79 1.00 0.98 0.95 0.93 0.90 0.95 0.86 0.81 0.78 0.75 0.74 0.73 0.69 0.91 0.85																			
7 0.94 0.89 0.82 0.80 0.77 0.98 1.00 0.97 0.94 0.92 0.94 0.85 0.80 0.77 0.74 0.73 0.71 0.68 0.89 0.83																			
8 0.91 0.86 0.79 0.77 0.75 0.95 0.97 1.00 0.97 0.95 0.91 0.82 0.77 0.74 0.72 0.70 0.69 0.67 0.87 0.80																			
9 0.89 0.84 0.77 0.75 0.73 0.93 0.94 0.97 1.00 0.98 0.89 0.80 0.75 0.72 0.70 0.68 0.67 0.66 0.84 0.78																			
10 0.86 0.82 0.75 0.73 0.71 0.90 0.92 0.95 0.98 1.00 0.86 0.78 0.73 0.71 0.68 0.66 0.65 0.64 0.82 0.76																			
11 1.00 0.95 0.87 0.85 0.82 0.95 0.94 0.91 0.89 0.86 1.00 0.90 0.85 0.82 0.79 0.77 0.76 0.73 0.95 0.89																			
12 0.90 0.85 0.79 0.76 0.74 0.86 0.85 0.82 0.80 0.78 0.90 1.00 0.94 0.91 0.88 0.86 0.85 0.81 0.89 0.82																			
13 0.85 0.80 0.74 0.73 0.70 0.81 0.80 0.77 0.75 0.73 0.85 0.94 1.00 0.97 0.93 0.91 0.90 0.86 0.83 0.77																			
14 0.82 0.78 0.71 0.70 0.68 0.78 0.77 0.74 0.72 0.71 0.82 0.91 0.97 1.00 0.97 0.95 0.93 0.89 0.81 0.75																			
15 0.79 0.75 0.69 0.67 0.65 0.75 0.74 0.72 0.70 0.68 0.79 0.88 0.93 0.97 1.00 0.98 0.96 0.92 0.78 0.76																			

16| 0.77|0.73|0.67|0.66|0.64|0.74|0.73|0.70|0.68|0.66|0.77|0.86|0.91|0.95|0.98|1.00|0.98|0.94|0.76|0.74|
 17| 0.76|0.72|0.66|0.65|0.63|0.73|0.71|0.69|0.67|0.65|0.76|0.85|0.90|0.93|0.96|0.98|1.00|0.96|0.75|0.73|
 18| 0.73|0.69|0.63|0.62|0.60|0.69|0.68|0.67|0.66|0.64|0.73|0.81|0.86|0.89|0.92|0.94|0.96|1.00|0.72|0.70|
 19| 0.95|0.90|0.83|0.81|0.78|0.91|0.89|0.87|0.84|0.82|0.95|0.89|0.83|0.81|0.78|0.76|0.75|0.72|1.00|0.93|
 20| 0.89|0.84|0.77|0.75|0.72|0.85|0.83|0.80|0.78|0.76|0.89|0.82|0.77|0.75|0.76|0.74|0.73|0.70|0.93|1.00|

=====TOPLAM UNIQUE SHINGLE SAYISI=>635

Merge edilmiş unique shingle tablosu yazılışın mı? e/E/h/Hh
h

İmza Benzerliği, K = 5

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
1 1.00 0.98 0.95 0.93 0.96 0.95 0.94 0.94 0.92 1.00 0.94 0.93 0.91 0.90 0.89 0.89 0.87 0.99 0.97																			
2 0.98 1.00 0.97 0.95 0.94 0.93 0.92 0.92 0.92 0.98 0.94 0.93 0.91 0.90 0.91 0.91 0.89 0.97 0.95																			
3 0.95 0.97 1.00 0.98 0.97 0.91 0.90 0.89 0.89 0.89 0.95 0.92 0.91 0.89 0.88 0.89 0.89 0.87 0.94 0.92																			
4 0.93 0.95 0.98 1.00 0.99 0.90 0.89 0.88 0.88 0.88 0.93 0.90 0.89 0.89 0.88 0.89 0.89 0.87 0.92 0.90																			
5 0.93 0.95 0.97 0.99 1.00 0.90 0.89 0.88 0.88 0.88 0.93 0.90 0.89 0.89 0.88 0.89 0.89 0.87 0.92 0.90																			
6 0.96 0.94 0.91 0.90 0.90 1.00 0.99 0.98 0.98 0.98 0.96 0.96 0.96 0.90 0.89 0.88 0.87 0.86 0.86 0.86 0.95 0.93																			
7 0.95 0.93 0.90 0.89 0.89 0.99 1.00 0.99 0.99 0.98 0.96 0.95 0.89 0.88 0.87 0.86 0.86 0.86 0.94 0.92																			
8 0.94 0.92 0.89 0.88 0.88 0.98 0.99 1.00 0.99 0.97 0.94 0.89 0.88 0.88 0.87 0.86 0.86 0.86 0.86 0.93 0.91																			
9 0.94 0.92 0.89 0.88 0.88 0.98 0.98 0.99 1.00 0.98 0.94 0.89 0.88 0.88 0.87 0.86 0.86 0.86 0.86 0.93 0.91																			
10 0.92 0.92 0.89 0.88 0.88 0.96 0.96 0.97 0.98 1.00 0.92 0.87 0.86 0.86 0.85 0.86 0.86 0.86 0.91 0.89																			
11 1.00 0.98 0.95 0.93 0.93 0.96 0.95 0.94 0.94 0.92 1.00 0.94 0.93 0.91 0.90 0.89 0.89 0.87 0.99 0.97																			
12 0.94 0.94 0.92 0.90 0.90 0.90 0.90 0.89 0.89 0.89 0.87 0.94 1.00 0.99 0.97 0.96 0.95 0.95 0.93 0.91																			
13 0.93 0.93 0.91 0.89 0.89 0.89 0.89 0.88 0.88 0.88 0.86 0.93 0.99 1.00 0.98 0.97 0.96 0.95 0.93 0.91																			
14 0.91 0.91 0.89 0.89 0.89 0.88 0.88 0.88 0.88 0.88 0.86 0.91 0.97 0.98 1.00 0.99 0.98 0.97 0.95 0.90 0.88																			
15 0.90 0.90 0.88 0.88 0.88 0.87 0.87 0.87 0.87 0.87 0.85 0.90 0.96 0.97 0.99 1.00 0.99 0.98 0.96 0.89 0.89																			
16 0.89 0.91 0.89 0.89 0.89 0.86 0.86 0.86 0.86 0.86 0.89 0.95 0.96 0.98 0.99 1.00 0.99 0.97 0.88 0.88																			
17 0.89 0.91 0.89 0.89 0.89 0.86 0.86 0.86 0.86 0.86 0.89 0.95 0.95 0.97 0.98 0.99 1.00 0.98 0.88 0.88 0.88																			
18 0.87 0.89 0.87 0.87 0.87 0.86 0.86 0.86 0.86 0.86 0.86 0.87 0.93 0.93 0.95 0.96 0.97 0.98 1.00 0.86 0.86																			
19 0.99 0.97 0.94 0.92 0.92 0.95 0.94 0.93 0.93 0.91 0.99 0.93 0.92 0.90 0.89 0.88 0.88 0.86 1.00 0.98																			
20 0.97 0.95 0.92 0.90 0.90 0.93 0.92 0.91 0.91 0.89 0.97 0.91 0.90 0.88 0.89 0.88 0.88 0.86 0.98 1.00																			

K-SHINGLE BENZERLİKLERINE GORE DOKUMAN ÇİFTLERİ

Threshold 0.70 için Benzerlik oranı esik değerinden fazla olan dokuman çiftleri=>

dokuman(1, 2) benzerlik oranı=> 0.95	dokuman(1, 3) benzerlik oranı=> 0.87	dokuman(1, 4) benzerlik oranı=> 0.85
dokuman(1, 5) benzerlik oranı=> 0.82	dokuman(1, 6) benzerlik oranı=> 0.95	dokuman(1, 7) benzerlik oranı=> 0.94
dokuman(1, 8) benzerlik oranı=> 0.91	dokuman(1, 9) benzerlik oranı=> 0.89	dokuman(1, 10) benzerlik oranı=> 0.86
dokuman(1, 11) benzerlik oranı=> 1.00	dokuman(1, 12) benzerlik oranı=> 0.90	dokuman(1, 13) benzerlik oranı=> 0.85
dokuman(1, 14) benzerlik oranı=> 0.82	dokuman(1, 15) benzerlik oranı=> 0.79	dokuman(1, 16) benzerlik oranı=> 0.77
dokuman(1, 17) benzerlik oranı=> 0.76	dokuman(1, 18) benzerlik oranı=> 0.73	dokuman(1, 19) benzerlik oranı=> 0.95
dokuman(1, 20) benzerlik oranı=> 0.89	dokuman(2, 3) benzerlik oranı=> 0.92	dokuman(2, 4) benzerlik oranı=> 0.89
dokuman(2, 5) benzerlik oranı=> 0.87	dokuman(2, 6) benzerlik oranı=> 0.90	dokuman(2, 7) benzerlik oranı=> 0.89
dokuman(2, 8) benzerlik oranı=> 0.86	dokuman(2, 9) benzerlik oranı=> 0.84	dokuman(2, 10) benzerlik oranı=> 0.82
dokuman(2, 11) benzerlik oranı=> 0.95	dokuman(2, 12) benzerlik oranı=> 0.85	dokuman(2, 13) benzerlik oranı=> 0.80
dokuman(2, 14) benzerlik oranı=> 0.78	dokuman(2, 15) benzerlik oranı=> 0.75	dokuman(2, 16) benzerlik oranı=> 0.73
dokuman(2, 17) benzerlik oranı=> 0.72	dokuman(2, 19) benzerlik oranı=> 0.90	dokuman(2, 20) benzerlik oranı=> 0.84
dokuman(3, 4) benzerlik oranı=> 0.95	dokuman(3, 5) benzerlik oranı=> 0.92	dokuman(3, 6) benzerlik oranı=> 0.83
dokuman(3, 7) benzerlik oranı=> 0.82	dokuman(3, 8) benzerlik oranı=> 0.79	dokuman(3, 9) benzerlik oranı=> 0.77
dokuman(3, 10) benzerlik oranı=> 0.75	dokuman(3, 11) benzerlik oranı=> 0.87	dokuman(3, 12) benzerlik oranı=> 0.79
dokuman(3, 13) benzerlik oranı=> 0.74	dokuman(3, 14) benzerlik oranı=> 0.71	dokuman(3, 19) benzerlik oranı=> 0.83
dokuman(3, 20) benzerlik oranı=> 0.77	dokuman(4, 5) benzerlik oranı=> 0.97	dokuman(4, 6) benzerlik oranı=> 0.81
dokuman(4, 7) benzerlik oranı=> 0.80	dokuman(4, 8) benzerlik oranı=> 0.77	dokuman(4, 9) benzerlik oranı=> 0.75
dokuman(4, 10) benzerlik oranı=> 0.73	dokuman(4, 11) benzerlik oranı=> 0.85	dokuman(4, 12) benzerlik oranı=> 0.76
dokuman(4, 13) benzerlik oranı=> 0.73	dokuman(4, 14) benzerlik oranı=> 0.70	dokuman(4, 19) benzerlik oranı=> 0.81
dokuman(4, 20) benzerlik oranı=> 0.75	dokuman(5, 6) benzerlik oranı=> 0.79	dokuman(5, 7) benzerlik oranı=> 0.77
dokuman(5, 8) benzerlik oranı=> 0.75	dokuman(5, 9) benzerlik oranı=> 0.73	dokuman(5, 10) benzerlik oranı=> 0.71
dokuman(5, 11) benzerlik oranı=> 0.82	dokuman(5, 12) benzerlik oranı=> 0.74	dokuman(5, 13) benzerlik oranı=> 0.70
dokuman(5, 19) benzerlik oranı=> 0.78	dokuman(5, 20) benzerlik oranı=> 0.72	dokuman(6, 7) benzerlik oranı=> 0.98
dokuman(6, 8) benzerlik oranı=> 0.95	dokuman(6, 9) benzerlik oranı=> 0.93	dokuman(6, 10) benzerlik oranı=> 0.90
dokuman(6, 11) benzerlik oranı=> 0.95	dokuman(6, 12) benzerlik oranı=> 0.86	dokuman(6, 13) benzerlik oranı=> 0.81
dokuman(6, 14) benzerlik oranı=> 0.78	dokuman(6, 15) benzerlik oranı=> 0.75	dokuman(6, 16) benzerlik oranı=> 0.74
dokuman(6, 17) benzerlik oranı=> 0.73	dokuman(6, 19) benzerlik oranı=> 0.91	dokuman(6, 20) benzerlik oranı=> 0.85
dokuman(7, 8) benzerlik oranı=> 0.97	dokuman(7, 9) benzerlik oranı=> 0.94	dokuman(7, 10) benzerlik oranı=> 0.92
dokuman(7, 11) benzerlik oranı=> 0.94	dokuman(7, 12) benzerlik oranı=> 0.85	dokuman(7, 13) benzerlik oranı=> 0.80
dokuman(7, 14) benzerlik oranı=> 0.77	dokuman(7, 15) benzerlik oranı=> 0.74	dokuman(7, 16) benzerlik oranı=> 0.73
dokuman(7, 17) benzerlik oranı=> 0.71	dokuman(7, 19) benzerlik oranı=> 0.89	dokuman(7, 20) benzerlik oranı=> 0.83

dokuman(3, 18)imza benzerligi=> 0.87
dokuman(4, 5)imza benzerligi=> 0.99
dokuman(4, 8)imza benzerligi=> 0.88
dokuman(4, 11)imza benzerligi=> 0.93
dokuman(4, 14)imza benzerligi=> 0.89
dokuman(4, 17)imza benzerligi=> 0.89
dokuman(4, 20)imza benzerligi=> 0.90
dokuman(5, 8)imza benzerligi=> 0.88
dokuman(5, 11)imza benzerligi=> 0.93
dokuman(5, 14)imza benzerligi=> 0.89
dokuman(5, 17)imza benzerligi=> 0.89
dokuman(5, 20)imza benzerligi=> 0.90
dokuman(6, 9)imza benzerligi=> 0.98
dokuman(6, 12)imza benzerligi=> 0.90
dokuman(6, 15)imza benzerligi=> 0.87
dokuman(6, 18)imza benzerligi=> 0.86
dokuman(7, 8)imza benzerligi=> 0.99
dokuman(7, 11)imza benzerligi=> 0.95
dokuman(7, 14)imza benzerligi=> 0.88
dokuman(7, 17)imza benzerligi=> 0.86
dokuman(7, 20)imza benzerligi=> 0.92
dokuman(8, 11)imza benzerligi=> 0.94
dokuman(8, 14)imza benzerligi=> 0.88
dokuman(8, 17)imza benzerligi=> 0.86
dokuman(8, 20)imza benzerligi=> 0.91
dokuman(9, 12)imza benzerligi=> 0.89
dokuman(9, 15)imza benzerligi=> 0.87
dokuman(9, 18)imza benzerligi=> 0.86
dokuman(10, 11)imza benzerligi=> 0.92
dokuman(10, 14)imza benzerligi=> 0.86
dokuman(10, 17)imza benzerligi=> 0.86
dokuman(10, 20)imza benzerligi=> 0.89
dokuman(11, 14)imza benzerligi=> 0.91
dokuman(11, 17)imza benzerligi=> 0.89
dokuman(11, 20)imza benzerligi=> 0.97
dokuman(12, 15)imza benzerligi=> 0.96
dokuman(12, 18)imza benzerligi=> 0.93
dokuman(13, 14)imza benzerligi=> 0.98
dokuman(13, 17)imza benzerligi=> 0.95
dokuman(13, 20)imza benzerligi=> 0.90
dokuman(14, 17)imza benzerligi=> 0.97
dokuman(14, 20)imza benzerligi=> 0.88
dokuman(15, 18)imza benzerligi=> 0.96
dokuman(16, 17)imza benzerligi=> 0.99
dokuman(16, 20)imza benzerligi=> 0.88
dokuman(17, 20)imza benzerligi=> 0.88
dokuman(19, 20)imza benzerligi=> 0.98
dokuman(3, 19)imza benzerligi=> 0.94
dokuman(4, 6)imza benzerligi=> 0.90
dokuman(4, 9)imza benzerligi=> 0.88
dokuman(4, 12)imza benzerligi=> 0.90
dokuman(4, 15)imza benzerligi=> 0.88
dokuman(4, 18)imza benzerligi=> 0.87
dokuman(5, 6)imza benzerligi=> 0.90
dokuman(5, 9)imza benzerligi=> 0.88
dokuman(5, 12)imza benzerligi=> 0.90
dokuman(5, 15)imza benzerligi=> 0.88
dokuman(5, 18)imza benzerligi=> 0.87
dokuman(6, 7)imza benzerligi=> 0.99
dokuman(6, 10)imza benzerligi=> 0.96
dokuman(6, 13)imza benzerligi=> 0.89
dokuman(6, 16)imza benzerligi=> 0.86
dokuman(6, 19)imza benzerligi=> 0.95
dokuman(7, 9)imza benzerligi=> 0.98
dokuman(7, 12)imza benzerligi=> 0.89
dokuman(7, 15)imza benzerligi=> 0.87
dokuman(7, 18)imza benzerligi=> 0.86
dokuman(8, 9)imza benzerligi=> 0.99
dokuman(8, 12)imza benzerligi=> 0.89
dokuman(8, 15)imza benzerligi=> 0.87
dokuman(8, 18)imza benzerligi=> 0.86
dokuman(9, 10)imza benzerligi=> 0.98
dokuman(9, 13)imza benzerligi=> 0.88
dokuman(9, 16)imza benzerligi=> 0.86
dokuman(9, 19)imza benzerligi=> 0.93
dokuman(10, 12)imza benzerligi=> 0.87
dokuman(10, 15)imza benzerligi=> 0.85
dokuman(10, 18)imza benzerligi=> 0.86
dokuman(11, 12)imza benzerligi=> 0.94
dokuman(11, 15)imza benzerligi=> 0.90
dokuman(11, 18)imza benzerligi=> 0.87
dokuman(12, 13)imza benzerligi=> 0.99
dokuman(12, 16)imza benzerligi=> 0.95
dokuman(12, 19)imza benzerligi=> 0.93
dokuman(13, 15)imza benzerligi=> 0.97
dokuman(13, 18)imza benzerligi=> 0.93
dokuman(14, 15)imza benzerligi=> 0.99
dokuman(14, 18)imza benzerligi=> 0.95
dokuman(15, 16)imza benzerligi=> 0.99
dokuman(15, 19)imza benzerligi=> 0.89
dokuman(16, 18)imza benzerligi=> 0.97
dokuman(17, 18)imza benzerligi=> 0.98
dokuman(18, 19)imza benzerligi=> 0.86
dokuman(3, 20)imza benzerligi=> 0.92
dokuman(4, 7)imza benzerligi=> 0.89
dokuman(4, 10)imza benzerligi=> 0.88
dokuman(4, 13)imza benzerligi=> 0.89
dokuman(4, 16)imza benzerligi=> 0.89
dokuman(4, 19)imza benzerligi=> 0.92
dokuman(5, 7)imza benzerligi=> 0.89
dokuman(5, 10)imza benzerligi=> 0.88
dokuman(5, 13)imza benzerligi=> 0.89
dokuman(5, 16)imza benzerligi=> 0.89
dokuman(5, 19)imza benzerligi=> 0.92
dokuman(6, 8)imza benzerligi=> 0.98
dokuman(6, 11)imza benzerligi=> 0.96
dokuman(6, 14)imza benzerligi=> 0.88
dokuman(6, 17)imza benzerligi=> 0.86
dokuman(6, 20)imza benzerligi=> 0.93
dokuman(7, 10)imza benzerligi=> 0.96
dokuman(7, 13)imza benzerligi=> 0.88
dokuman(7, 16)imza benzerligi=> 0.86
dokuman(7, 19)imza benzerligi=> 0.94
dokuman(8, 10)imza benzerligi=> 0.97
dokuman(8, 13)imza benzerligi=> 0.88
dokuman(8, 16)imza benzerligi=> 0.86
dokuman(8, 19)imza benzerligi=> 0.93
dokuman(9, 11)imza benzerligi=> 0.94
dokuman(9, 14)imza benzerligi=> 0.88
dokuman(9, 17)imza benzerligi=> 0.86
dokuman(9, 20)imza benzerligi=> 0.91
dokuman(10, 13)imza benzerligi=> 0.86
dokuman(10, 16)imza benzerligi=> 0.86
dokuman(10, 19)imza benzerligi=> 0.91
dokuman(11, 13)imza benzerligi=> 0.93
dokuman(11, 16)imza benzerligi=> 0.89
dokuman(11, 19)imza benzerligi=> 0.99
dokuman(12, 14)imza benzerligi=> 0.97
dokuman(12, 17)imza benzerligi=> 0.95
dokuman(12, 20)imza benzerligi=> 0.91
dokuman(13, 16)imza benzerligi=> 0.96
dokuman(13, 19)imza benzerligi=> 0.92
dokuman(14, 16)imza benzerligi=> 0.98
dokuman(14, 19)imza benzerligi=> 0.90
dokuman(15, 17)imza benzerligi=> 0.98
dokuman(15, 20)imza benzerligi=> 0.89
dokuman(16, 19)imza benzerligi=> 0.88
dokuman(17, 19)imza benzerligi=> 0.88
dokuman(18, 20)imza benzerligi=> 0.86

Threshold 0.80 icin Benzerlik orani esik degerinden fazla olan dokuman ciftleri==>

```

dokuman(11, 20)imza benzerligi=> 0.97 dokuman(12, 13)imza benzerligi=> 0.99 dokuman(12, 14)imza benzerligi=> 0.97
dokuman(12, 15)imza benzerligi=> 0.96 dokuman(12, 16)imza benzerligi=> 0.95 dokuman(12, 17)imza benzerligi=> 0.95
dokuman(12, 18)imza benzerligi=> 0.93 dokuman(12, 19)imza benzerligi=> 0.93 dokuman(12, 20)imza benzerligi=> 0.91
dokuman(13, 14)imza benzerligi=> 0.98 dokuman(13, 15)imza benzerligi=> 0.97 dokuman(13, 16)imza benzerligi=> 0.96
dokuman(13, 17)imza benzerligi=> 0.95 dokuman(13, 18)imza benzerligi=> 0.93 dokuman(13, 19)imza benzerligi=> 0.92
dokuman(13, 19)imza benzerligi=> 0.95 dokuman(14, 15)imza benzerligi=> 0.99 dokuman(14, 16)imza benzerligi=> 0.98
dokuman(14, 17)imza benzerligi=> 0.97 dokuman(14, 18)imza benzerligi=> 0.95 dokuman(14, 19)imza benzerligi=> 0.90
dokuman(14, 20)imza benzerligi=> 0.88 dokuman(15, 16)imza benzerligi=> 0.99 dokuman(15, 17)imza benzerligi=> 0.98
dokuman(15, 18)imza benzerligi=> 0.96 dokuman(15, 19)imza benzerligi=> 0.89 dokuman(15, 20)imza benzerligi=> 0.89
dokuman(16, 17)imza benzerligi=> 0.99 dokuman(16, 18)imza benzerligi=> 0.97 dokuman(16, 19)imza benzerligi=> 0.88
dokuman(16, 20)imza benzerligi=> 0.88 dokuman(17, 18)imza benzerligi=> 0.98 dokuman(17, 19)imza benzerligi=> 0.88
dokuman(17, 20)imza benzerligi=> 0.88 dokuman(18, 19)imza benzerligi=> 0.86 dokuman(18, 20)imza benzerligi=> 0.86
dokuman(19, 20)imza benzerligi=> 0.98
-----
```

Threshold 0.90 için Benzerlik orani esik değerinden fazla olan dokuman çiftleri==>

```

dokuman(1, 2)imza benzerligi=> 0.98 dokuman(1, 3)imza benzerligi=> 0.95 dokuman(1, 4)imza benzerligi=> 0.93
dokuman(1, 5)imza benzerligi=> 0.93 dokuman(1, 6)imza benzerligi=> 0.96 dokuman(1, 7)imza benzerligi=> 0.95
dokuman(1, 8)imza benzerligi=> 0.94 dokuman(1, 9)imza benzerligi=> 0.94 dokuman(1, 10)imza benzerligi=> 0.92
dokuman(1, 11)imza benzerligi=> 1.00 dokuman(1, 12)imza benzerligi=> 0.94 dokuman(1, 13)imza benzerligi=> 0.93
dokuman(1, 14)imza benzerligi=> 0.91 dokuman(1, 15)imza benzerligi=> 0.90 dokuman(1, 19)imza benzerligi=> 0.99
dokuman(1, 20)imza benzerligi=> 0.97 dokuman(2, 3)imza benzerligi=> 0.97 dokuman(2, 4)imza benzerligi=> 0.95
dokuman(2, 5)imza benzerligi=> 0.95 dokuman(2, 6)imza benzerligi=> 0.94 dokuman(2, 7)imza benzerligi=> 0.93
dokuman(2, 8)imza benzerligi=> 0.92 dokuman(2, 9)imza benzerligi=> 0.92 dokuman(2, 10)imza benzerligi=> 0.92
dokuman(2, 11)imza benzerligi=> 0.98 dokuman(2, 12)imza benzerligi=> 0.94 dokuman(2, 13)imza benzerligi=> 0.93
dokuman(2, 14)imza benzerligi=> 0.91 dokuman(2, 15)imza benzerligi=> 0.90 dokuman(2, 16)imza benzerligi=> 0.91
dokuman(2, 17)imza benzerligi=> 0.91 dokuman(2, 19)imza benzerligi=> 0.97 dokuman(2, 20)imza benzerligi=> 0.95
dokuman(3, 4)imza benzerligi=> 0.98 dokuman(3, 5)imza benzerligi=> 0.97 dokuman(3, 6)imza benzerligi=> 0.91
dokuman(3, 7)imza benzerligi=> 0.90 dokuman(3, 11)imza benzerligi=> 0.95 dokuman(3, 12)imza benzerligi=> 0.92
dokuman(3, 13)imza benzerligi=> 0.91 dokuman(3, 19)imza benzerligi=> 0.94 dokuman(3, 20)imza benzerligi=> 0.92
dokuman(4, 5)imza benzerligi=> 0.99 dokuman(4, 6)imza benzerligi=> 0.90 dokuman(4, 11)imza benzerligi=> 0.93
dokuman(4, 12)imza benzerligi=> 0.90 dokuman(4, 19)imza benzerligi=> 0.92 dokuman(4, 20)imza benzerligi=> 0.90
dokuman(5, 6)imza benzerligi=> 0.90 dokuman(5, 11)imza benzerligi=> 0.93 dokuman(5, 12)imza benzerligi=> 0.90
dokuman(5, 19)imza benzerligi=> 0.92 dokuman(5, 20)imza benzerligi=> 0.90 dokuman(6, 7)imza benzerligi=> 0.99
dokuman(6, 8)imza benzerligi=> 0.98 dokuman(6, 9)imza benzerligi=> 0.98 dokuman(6, 10)imza benzerligi=> 0.96
dokuman(6, 11)imza benzerligi=> 0.96 dokuman(6, 12)imza benzerligi=> 0.90 dokuman(6, 19)imza benzerligi=> 0.95
dokuman(6, 20)imza benzerligi=> 0.93 dokuman(7, 8)imza benzerligi=> 0.99 dokuman(7, 9)imza benzerligi=> 0.98
dokuman(7, 10)imza benzerligi=> 0.96 dokuman(7, 11)imza benzerligi=> 0.95 dokuman(7, 19)imza benzerligi=> 0.94
dokuman(7, 20)imza benzerligi=> 0.92 dokuman(8, 9)imza benzerligi=> 0.99 dokuman(8, 10)imza benzerligi=> 0.97
dokuman(8, 11)imza benzerligi=> 0.94 dokuman(8, 19)imza benzerligi=> 0.93 dokuman(8, 20)imza benzerligi=> 0.91
dokuman(9, 10)imza benzerligi=> 0.98 dokuman(9, 11)imza benzerligi=> 0.94 dokuman(9, 19)imza benzerligi=> 0.93
dokuman(9, 20)imza benzerligi=> 0.91 dokuman(10, 11)imza benzerligi=> 0.92 dokuman(10, 19)imza benzerligi=> 0.91
dokuman(11, 12)imza benzerligi=> 0.94 dokuman(11, 13)imza benzerligi=> 0.93 dokuman(11, 14)imza benzerligi=> 0.91
dokuman(11, 15)imza benzerligi=> 0.90 dokuman(11, 19)imza benzerligi=> 0.99 dokuman(11, 20)imza benzerligi=> 0.97
dokuman(12, 13)imza benzerligi=> 0.99 dokuman(12, 14)imza benzerligi=> 0.97 dokuman(12, 15)imza benzerligi=> 0.96
dokuman(12, 16)imza benzerligi=> 0.95 dokuman(12, 17)imza benzerligi=> 0.95 dokuman(12, 18)imza benzerligi=> 0.93
dokuman(12, 19)imza benzerligi=> 0.93 dokuman(12, 20)imza benzerligi=> 0.91 dokuman(13, 14)imza benzerligi=> 0.98
dokuman(13, 15)imza benzerligi=> 0.97 dokuman(13, 16)imza benzerligi=> 0.96 dokuman(13, 17)imza benzerligi=> 0.95
dokuman(13, 18)imza benzerligi=> 0.93 dokuman(13, 19)imza benzerligi=> 0.92 dokuman(13, 20)imza benzerligi=> 0.90
dokuman(14, 15)imza benzerligi=> 0.99 dokuman(14, 16)imza benzerligi=> 0.98 dokuman(14, 17)imza benzerligi=> 0.97
dokuman(14, 18)imza benzerligi=> 0.95 dokuman(14, 19)imza benzerligi=> 0.90 dokuman(15, 16)imza benzerligi=> 0.99
dokuman(15, 17)imza benzerligi=> 0.98 dokuman(15, 18)imza benzerligi=> 0.96 dokuman(15, 19)imza benzerligi=> 0.99
dokuman(16, 18)imza benzerligi=> 0.97 dokuman(17, 18)imza benzerligi=> 0.98 dokuman(19, 20)imza benzerligi=> 0.98

```

K==>8
8

=====
K = 8 için dokumanların unique shingle sayıları ==>

```

Dokuman => 1, unique shingle =>558
Dokuman => 2, unique shingle =>556
Dokuman => 3, unique shingle =>547
Dokuman => 4, unique shingle =>549
Dokuman => 5, unique shingle =>538
Dokuman => 6, unique shingle =>564
Dokuman => 7, unique shingle =>558
Dokuman => 8, unique shingle =>555
Dokuman => 9, unique shingle =>553
Dokuman => 10, unique shingle =>552
Dokuman => 11, unique shingle =>559
Dokuman => 12, unique shingle =>561
Dokuman => 13, unique shingle =>562
Dokuman => 14, unique shingle =>561
Dokuman => 15, unique shingle =>554
Dokuman => 16, unique shingle =>559
Dokuman => 17, unique shingle =>560
Dokuman => 18, unique shingle =>539
Dokuman => 19, unique shingle =>553
Dokuman => 20, unique shingle =>526
Shingle tablolari yazilsin mi ? e/E/h/Hh
h

```

Jaccard Benzerligi, K = 8

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
1 1.00 0.93 0.84 0.81 0.78 0.95 0.90 0.85 0.82 1.00 0.87 0.81 0.78 0.75 0.74 0.71 0.67 0.64 0.88	2 0.93 1.00 0.91 0.87 0.84 0.88 0.87 0.83 0.79 0.76 0.93 0.81 0.76 0.73 0.70 0.68 0.66 0.62 0.88 0.81	3 0.84 0.91 1.00 0.94 0.90 0.80 0.79 0.75 0.72 0.69 0.84 0.74 0.68 0.66 0.63 0.62 0.60 0.55 0.79 0.73	4 0.81 0.87 0.94 1.00 0.96 0.77 0.75 0.73 0.69 0.66 0.81 0.71 0.66 0.63 0.60 0.59 0.57 0.76 0.70	5 0.78 0.84 0.90 0.96 1.00 0.74 0.72 0.69 0.66 0.63 0.78 0.68 0.63 0.60 0.58 0.56 0.55 0.50 0.73 0.67	6 0.95 0.88 0.80 0.77 0.74 1.00 0.97 0.93 0.89 0.85 0.95 0.83 0.77 0.75 0.71 0.70 0.68 0.63 0.90 0.83														

7| 0.93|0.87|0.79|0.75|0.72|0.97|1.00|0.96|0.92|0.88|0.93|0.82|0.76|0.73|0.70|0.69|0.67|0.62|0.88|0.82|
 8| 0.90|0.83|0.75|0.73|0.69|0.93|0.96|1.00|0.95|0.92|0.90|0.79|0.73|0.70|0.67|0.66|0.64|0.62|0.85|0.78|
 9| 0.85|0.79|0.72|0.69|0.66|0.89|0.92|0.95|1.00|0.96|0.85|0.75|0.69|0.67|0.64|0.63|0.61|0.59|0.80|0.74|
 10| 0.82|0.76|0.69|0.66|0.63|0.85|0.88|0.92|0.96|1.00|0.82|0.72|0.67|0.64|0.61|0.60|0.58|0.56|0.77|0.71|
 11| 1.00|0.93|0.84|0.81|0.78|0.95|0.93|0.90|0.85|0.82|1.00|0.88|0.81|0.79|0.75|0.74|0.72|0.67|0.94|0.88|
 12| 0.87|0.81|0.74|0.71|0.68|0.83|0.82|0.79|0.75|0.72|0.88|1.00|0.93|0.90|0.86|0.85|0.82|0.77|0.86|0.80|
 13| 0.81|0.76|0.68|0.66|0.63|0.77|0.76|0.73|0.69|0.67|0.81|0.93|1.00|0.96|0.92|0.90|0.88|0.82|0.80|0.74|
 14| 0.78|0.73|0.66|0.63|0.60|0.75|0.73|0.70|0.67|0.64|0.79|0.90|0.96|1.00|0.96|0.94|0.91|0.86|0.77|0.71|
 15| 0.75|0.70|0.63|0.60|0.58|0.71|0.70|0.67|0.64|0.61|0.75|0.86|0.92|0.96|1.00|0.98|0.95|0.89|0.74|0.72|
 16| 0.74|0.68|0.62|0.59|0.56|0.70|0.69|0.66|0.63|0.60|0.74|0.85|0.90|0.94|0.98|1.00|0.97|0.91|0.72|0.71|
 17| 0.71|0.66|0.60|0.57|0.55|0.68|0.67|0.64|0.61|0.58|0.72|0.82|0.88|0.91|0.95|0.97|1.00|0.94|0.70|0.68|
 18| 0.67|0.62|0.55|0.53|0.50|0.63|0.62|0.62|0.59|0.56|0.67|0.77|0.82|0.86|0.89|0.91|0.94|1.00|0.65|0.64|
 19| 0.94|0.88|0.79|0.76|0.73|0.90|0.88|0.85|0.80|0.77|0.94|0.86|0.80|0.77|0.74|0.72|0.70|0.65|1.00|0.93|
 20| 0.88|0.81|0.73|0.70|0.67|0.83|0.82|0.78|0.74|0.71|0.88|0.80|0.74|0.71|0.72|0.71|0.68|0.64|0.93|1.00|

=====TOPLAM UNIQUE SHINGLE SAYISI=>797

Merge edilmiş unique shingle tablosu yazılışın mı? e/E/h/Hh
h

İmza Benzerliği, K = 8

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
1 1.00 0.96 0.93 0.88 0.87 0.97 0.98 0.95 0.94 0.94 1.00 0.92 0.89 0.88 0.86 0.84 0.83 0.82 0.98 0.93																			
2 0.96 1.00 0.96 0.92 0.91 0.93 0.94 0.91 0.90 0.90 0.96 0.94 0.91 0.91 0.88 0.86 0.85 0.84 0.94 0.89																			
3 0.93 0.96 1.00 0.94 0.93 0.92 0.91 0.88 0.87 0.87 0.93 0.92 0.90 0.91 0.88 0.86 0.85 0.84 0.91 0.86																			
4 0.88 0.92 0.94 1.00 0.99 0.87 0.86 0.84 0.83 0.83 0.88 0.90 0.88 0.88 0.86 0.84 0.83 0.82 0.86 0.82																			
5 0.87 0.91 0.93 0.99 1.00 0.86 0.85 0.83 0.82 0.82 0.87 0.89 0.87 0.87 0.85 0.83 0.84 0.83 0.85 0.81																			
6 0.97 0.93 0.92 0.87 0.86 1.00 0.99 0.96 0.95 0.95 0.97 0.90 0.88 0.87 0.87 0.85 0.83 0.82 0.81 0.99 0.94																			
7 0.98 0.94 0.91 0.86 0.85 0.99 1.00 0.97 0.96 0.96 0.98 0.90 0.90 0.87 0.86 0.84 0.82 0.81 0.80 1.00 0.95																			
8 0.95 0.91 0.88 0.84 0.83 0.96 0.97 1.00 0.99 0.99 0.95 0.88 0.85 0.85 0.83 0.81 0.80 0.79 0.97 0.92																			
9 0.94 0.90 0.87 0.83 0.82 0.95 0.96 0.99 1.00 1.00 0.94 0.87 0.85 0.85 0.83 0.82 0.81 0.80 0.96 0.91																			
10 0.94 0.90 0.87 0.83 0.82 0.95 0.96 0.99 1.00 1.00 0.94 0.87 0.85 0.85 0.83 0.82 0.81 0.80 0.96 0.91																			
11 1.00 0.96 0.93 0.88 0.87 0.97 0.98 0.95 0.94 0.94 1.00 0.92 0.89 0.88 0.86 0.84 0.83 0.82 0.98 0.93																			
12 0.92 0.94 0.92 0.90 0.89 0.90 0.90 0.88 0.87 0.87 0.92 1.00 0.96 0.95 0.91 0.89 0.88 0.87 0.90 0.86																			
13 0.89 0.91 0.90 0.88 0.87 0.88 0.87 0.85 0.85 0.85 0.89 0.96 1.00 0.97 0.93 0.90 0.89 0.88 0.87 0.83																			
14 0.88 0.91 0.91 0.88 0.87 0.87 0.86 0.85 0.85 0.85 0.88 0.95 0.97 1.00 0.96 0.93 0.92 0.91 0.86 0.82																			
15 0.86 0.88 0.88 0.86 0.85 0.85 0.84 0.83 0.83 0.83 0.86 0.91 0.93 0.96 1.00 0.97 0.96 0.95 0.84 0.81																			
16 0.84 0.86 0.86 0.84 0.83 0.83 0.82 0.81 0.82 0.82 0.84 0.89 0.90 0.93 0.97 1.00 0.98 0.97 0.82 0.79																			
17 0.83 0.85 0.85 0.83 0.84 0.82 0.81 0.80 0.81 0.81 0.83 0.88 0.89 0.92 0.96 0.98 1.00 0.99 0.81 0.78																			
18 0.82 0.84 0.84 0.82 0.83 0.81 0.80 0.79 0.80 0.80 0.82 0.87 0.88 0.91 0.95 0.97 0.99 1.00 0.80 0.77																			
19 0.98 0.94 0.91 0.86 0.85 0.99 1.00 0.97 0.96 0.96 0.98 0.90 0.90 0.87 0.86 0.84 0.82 0.81 0.80 1.00 0.95																			
20 0.93 0.89 0.86 0.82 0.81 0.94 0.95 0.92 0.91 0.91 0.93 0.86 0.83 0.82 0.81 0.79 0.78 0.77 0.95 1.00																			

K-SHINGLE BENZERLİKLERINE GORE DOKUMAN CİFTLERİ

Threshold 0.70 icin Benzerlik orani esik degerinden fazla olan dokuman ciftleri=>

dokuman(1, 2) benzerlik orani=> 0.93	dokuman(1, 3) benzerlik orani=> 0.84	dokuman(1, 4) benzerlik orani=> 0.81
dokuman(1, 5) benzerlik orani=> 0.78	dokuman(1, 6) benzerlik orani=> 0.95	dokuman(1, 7) benzerlik orani=> 0.93
dokuman(1, 8) benzerlik orani=> 0.90	dokuman(1, 9) benzerlik orani=> 0.85	dokuman(1, 10) benzerlik orani=> 0.82
dokuman(1, 11) benzerlik orani=> 1.00	dokuman(1, 12) benzerlik orani=> 0.87	dokuman(1, 13) benzerlik orani=> 0.81
dokuman(1, 14) benzerlik orani=> 0.78	dokuman(1, 15) benzerlik orani=> 0.75	dokuman(1, 16) benzerlik orani=> 0.74
dokuman(1, 17) benzerlik orani=> 0.71	dokuman(1, 19) benzerlik orani=> 0.94	dokuman(1, 20) benzerlik orani=> 0.88
dokuman(2, 3) benzerlik orani=> 0.91	dokuman(2, 4) benzerlik orani=> 0.87	dokuman(2, 5) benzerlik orani=> 0.84
dokuman(2, 6) benzerlik orani=> 0.88	dokuman(2, 7) benzerlik orani=> 0.87	dokuman(2, 8) benzerlik orani=> 0.83
dokuman(2, 9) benzerlik orani=> 0.79	dokuman(2, 10) benzerlik orani=> 0.76	dokuman(2, 11) benzerlik orani=> 0.93
dokuman(2, 12) benzerlik orani=> 0.81	dokuman(2, 13) benzerlik orani=> 0.76	dokuman(2, 14) benzerlik orani=> 0.73
dokuman(2, 19) benzerlik orani=> 0.88	dokuman(2, 20) benzerlik orani=> 0.81	dokuman(3, 4) benzerlik orani=> 0.94
dokuman(3, 5) benzerlik orani=> 0.90	dokuman(3, 6) benzerlik orani=> 0.80	dokuman(3, 7) benzerlik orani=> 0.79
dokuman(3, 8) benzerlik orani=> 0.75	dokuman(3, 9) benzerlik orani=> 0.72	dokuman(3, 11) benzerlik orani=> 0.84
dokuman(3, 12) benzerlik orani=> 0.74	dokuman(3, 19) benzerlik orani=> 0.79	dokuman(3, 20) benzerlik orani=> 0.73

dokuman(4, 5) benzerlik orani=> 0.96
dokuman(4, 8) benzerlik orani=> 0.73
dokuman(4, 19) benzerlik orani=> 0.76
dokuman(5, 7) benzerlik orani=> 0.72
dokuman(6, 7) benzerlik orani=> 0.97
dokuman(6, 10) benzerlik orani=> 0.85
dokuman(6, 13) benzerlik orani=> 0.77
dokuman(6, 16) benzerlik orani=> 0.70
dokuman(7, 8) benzerlik orani=> 0.96
dokuman(7, 11) benzerlik orani=> 0.93
dokuman(7, 14) benzerlik orani=> 0.73
dokuman(7, 20) benzerlik orani=> 0.82
dokuman(8, 11) benzerlik orani=> 0.90
dokuman(8, 14) benzerlik orani=> 0.70
dokuman(9, 10) benzerlik orani=> 0.96
dokuman(9, 19) benzerlik orani=> 0.80
dokuman(10, 12) benzerlik orani=> 0.72
dokuman(11, 12) benzerlik orani=> 0.88
dokuman(11, 15) benzerlik orani=> 0.75
dokuman(11, 19) benzerlik orani=> 0.94
dokuman(12, 14) benzerlik orani=> 0.90
dokuman(12, 17) benzerlik orani=> 0.82
dokuman(12, 20) benzerlik orani=> 0.80
dokuman(13, 16) benzerlik orani=> 0.98
dokuman(13, 19) benzerlik orani=> 0.80
dokuman(14, 16) benzerlik orani=> 0.94
dokuman(14, 19) benzerlik orani=> 0.77
dokuman(15, 17) benzerlik orani=> 0.95
dokuman(15, 20) benzerlik orani=> 0.72
dokuman(16, 19) benzerlik orani=> 0.72
dokuman(17, 19) benzerlik orani=> 0.70
dokuman(4, 6) benzerlik orani=> 0.77
dokuman(4, 11) benzerlik orani=> 0.81
dokuman(4, 20) benzerlik orani=> 0.70
dokuman(5, 11) benzerlik orani=> 0.78
dokuman(6, 8) benzerlik orani=> 0.93
dokuman(6, 11) benzerlik orani=> 0.95
dokuman(6, 14) benzerlik orani=> 0.95
dokuman(6, 19) benzerlik orani=> 0.90
dokuman(7, 9) benzerlik orani=> 0.92
dokuman(7, 12) benzerlik orani=> 0.82
dokuman(7, 15) benzerlik orani=> 0.70
dokuman(8, 9) benzerlik orani=> 0.95
dokuman(8, 12) benzerlik orani=> 0.79
dokuman(8, 19) benzerlik orani=> 0.85
dokuman(9, 11) benzerlik orani=> 0.85
dokuman(9, 20) benzerlik orani=> 0.74
dokuman(10, 19) benzerlik orani=> 0.77
dokuman(11, 13) benzerlik orani=> 0.81
dokuman(11, 16) benzerlik orani=> 0.74
dokuman(11, 20) benzerlik orani=> 0.88
dokuman(12, 15) benzerlik orani=> 0.86
dokuman(12, 18) benzerlik orani=> 0.77
dokuman(13, 14) benzerlik orani=> 0.96
dokuman(13, 17) benzerlik orani=> 0.88
dokuman(13, 20) benzerlik orani=> 0.74
dokuman(14, 17) benzerlik orani=> 0.91
dokuman(14, 20) benzerlik orani=> 0.71
dokuman(15, 18) benzerlik orani=> 0.89
dokuman(16, 17) benzerlik orani=> 0.97
dokuman(16, 20) benzerlik orani=> 0.71
dokuman(19, 20) benzerlik orani=> 0.93
dokuman(4, 7) benzerlik orani=> 0.75
dokuman(4, 12) benzerlik orani=> 0.71
dokuman(5, 6) benzerlik orani=> 0.74
dokuman(5, 19) benzerlik orani=> 0.73
dokuman(6, 9) benzerlik orani=> 0.89
dokuman(6, 12) benzerlik orani=> 0.83
dokuman(6, 15) benzerlik orani=> 0.71
dokuman(6, 20) benzerlik orani=> 0.83
dokuman(7, 10) benzerlik orani=> 0.88
dokuman(7, 13) benzerlik orani=> 0.76
dokuman(7, 19) benzerlik orani=> 0.88
dokuman(8, 18) benzerlik orani=> 0.92
dokuman(8, 13) benzerlik orani=> 0.73
dokuman(8, 20) benzerlik orani=> 0.78
dokuman(9, 12) benzerlik orani=> 0.75
dokuman(10, 11) benzerlik orani=> 0.82
dokuman(10, 20) benzerlik orani=> 0.71
dokuman(11, 14) benzerlik orani=> 0.79
dokuman(11, 17) benzerlik orani=> 0.72
dokuman(12, 13) benzerlik orani=> 0.93
dokuman(12, 16) benzerlik orani=> 0.85
dokuman(12, 19) benzerlik orani=> 0.86
dokuman(13, 15) benzerlik orani=> 0.92
dokuman(13, 18) benzerlik orani=> 0.82
dokuman(14, 15) benzerlik orani=> 0.96
dokuman(14, 18) benzerlik orani=> 0.86
dokuman(15, 16) benzerlik orani=> 0.98
dokuman(15, 19) benzerlik orani=> 0.74
dokuman(16, 18) benzerlik orani=> 0.91
dokuman(17, 18) benzerlik orani=> 0.94

Threshold 0.80 icin Benzerlik orani esik degerinden fazla olan dokuman ciftleri==>

dokuman(1, 2) benzerlik orani=> 0.93
dokuman(1, 6) benzerlik orani=> 0.95
dokuman(1, 9) benzerlik orani=> 0.85
dokuman(1, 12) benzerlik orani=> 0.87
dokuman(1, 20) benzerlik orani=> 0.88
dokuman(2, 5) benzerlik orani=> 0.84
dokuman(2, 8) benzerlik orani=> 0.83
dokuman(2, 19) benzerlik orani=> 0.88
dokuman(3, 5) benzerlik orani=> 0.90
dokuman(4, 5) benzerlik orani=> 0.96
dokuman(6, 8) benzerlik orani=> 0.93
dokuman(6, 11) benzerlik orani=> 0.95
dokuman(6, 20) benzerlik orani=> 0.83
dokuman(7, 10) benzerlik orani=> 0.88
dokuman(7, 19) benzerlik orani=> 0.88
dokuman(8, 10) benzerlik orani=> 0.92
dokuman(9, 10) benzerlik orani=> 0.96
dokuman(10, 11) benzerlik orani=> 0.82
dokuman(11, 19) benzerlik orani=> 0.94
dokuman(12, 14) benzerlik orani=> 0.90
dokuman(12, 17) benzerlik orani=> 0.82
dokuman(13, 15) benzerlik orani=> 0.92
dokuman(13, 18) benzerlik orani=> 0.82
dokuman(14, 17) benzerlik orani=> 0.91
dokuman(15, 17) benzerlik orani=> 0.95
dokuman(16, 18) benzerlik orani=> 0.91
dokuman(1, 3) benzerlik orani=> 0.84
dokuman(1, 7) benzerlik orani=> 0.93
dokuman(1, 10) benzerlik orani=> 0.82
dokuman(1, 13) benzerlik orani=> 0.81
dokuman(2, 3) benzerlik orani=> 0.91
dokuman(2, 6) benzerlik orani=> 0.88
dokuman(2, 11) benzerlik orani=> 0.93
dokuman(2, 20) benzerlik orani=> 0.81
dokuman(3, 6) benzerlik orani=> 0.80
dokuman(4, 11) benzerlik orani=> 0.81
dokuman(6, 9) benzerlik orani=> 0.89
dokuman(6, 12) benzerlik orani=> 0.83
dokuman(7, 8) benzerlik orani=> 0.96
dokuman(7, 11) benzerlik orani=> 0.93
dokuman(7, 20) benzerlik orani=> 0.82
dokuman(8, 11) benzerlik orani=> 0.90
dokuman(9, 11) benzerlik orani=> 0.85
dokuman(11, 12) benzerlik orani=> 0.88
dokuman(11, 20) benzerlik orani=> 0.88
dokuman(12, 15) benzerlik orani=> 0.86
dokuman(12, 19) benzerlik orani=> 0.86
dokuman(13, 16) benzerlik orani=> 0.90
dokuman(14, 15) benzerlik orani=> 0.96
dokuman(14, 18) benzerlik orani=> 0.86
dokuman(15, 18) benzerlik orani=> 0.89
dokuman(17, 18) benzerlik orani=> 0.94
dokuman(1, 4) benzerlik orani=> 0.81
dokuman(1, 8) benzerlik orani=> 0.90
dokuman(1, 11) benzerlik orani=> 1.00
dokuman(1, 19) benzerlik orani=> 0.94
dokuman(2, 4) benzerlik orani=> 0.87
dokuman(2, 7) benzerlik orani=> 0.87
dokuman(2, 12) benzerlik orani=> 0.81
dokuman(3, 4) benzerlik orani=> 0.94
dokuman(3, 11) benzerlik orani=> 0.84
dokuman(6, 7) benzerlik orani=> 0.97
dokuman(6, 10) benzerlik orani=> 0.85
dokuman(6, 19) benzerlik orani=> 0.90
dokuman(7, 9) benzerlik orani=> 0.92
dokuman(7, 12) benzerlik orani=> 0.82
dokuman(8, 9) benzerlik orani=> 0.95
dokuman(8, 19) benzerlik orani=> 0.85
dokuman(9, 19) benzerlik orani=> 0.80
dokuman(11, 13) benzerlik orani=> 0.81
dokuman(12, 13) benzerlik orani=> 0.93
dokuman(12, 16) benzerlik orani=> 0.85
dokuman(13, 14) benzerlik orani=> 0.96
dokuman(13, 17) benzerlik orani=> 0.88
dokuman(14, 16) benzerlik orani=> 0.94
dokuman(15, 16) benzerlik orani=> 0.98
dokuman(16, 17) benzerlik orani=> 0.97
dokuman(19, 20) benzerlik orani=> 0.93

Threshold 0.90 icin Benzerlik orani esik degerinden fazla olan dokuman çiftleri=>

dokuman(1, 2) benzerlik orani=> 0.93
dokuman(1, 11) benzerlik orani=> 1.00
dokuman(2, 11) benzerlik orani=> 0.93
dokuman(4, 5) benzerlik orani=> 0.96
dokuman(6, 11) benzerlik orani=> 0.95
dokuman(7, 11) benzerlik orani=> 0.93
dokuman(9, 10) benzerlik orani=> 0.96
dokuman(13, 14) benzerlik orani=> 0.96
dokuman(14, 15) benzerlik orani=> 0.96
dokuman(15, 16) benzerlik orani=> 0.98
dokuman(16, 18) benzerlik orani=> 0.91
dokuman(1, 6) benzerlik orani=> 0.95
dokuman(1, 19) benzerlik orani=> 0.94
dokuman(3, 4) benzerlik orani=> 0.94
dokuman(6, 7) benzerlik orani=> 0.97
dokuman(7, 8) benzerlik orani=> 0.96
dokuman(8, 9) benzerlik orani=> 0.95
dokuman(11, 19) benzerlik orani=> 0.94
dokuman(13, 15) benzerlik orani=> 0.92
dokuman(14, 16) benzerlik orani=> 0.94
dokuman(15, 17) benzerlik orani=> 0.95
dokuman(17, 18) benzerlik orani=> 0.94
dokuman(1, 7) benzerlik orani=> 0.93
dokuman(2, 3) benzerlik orani=> 0.91
dokuman(3, 5) benzerlik orani=> 0.90
dokuman(6, 8) benzerlik orani=> 0.93
dokuman(7, 9) benzerlik orani=> 0.92
dokuman(8, 10) benzerlik orani=> 0.92
dokuman(12, 13) benzerlik orani=> 0.90
dokuman(13, 16) benzerlik orani=> 0.90
dokuman(14, 17) benzerlik orani=> 0.91
dokuman(16, 17) benzerlik orani=> 0.97
dokuman(19, 20) benzerlik orani=> 0.93

IMZA BENZERLİKLERİNE GÖRE DOKUMAN ÇİFTLERİ

Threshold 0.70 icin Benzerlik orani esik degerinden fazla olan dokuman ciftleri==>

dokuman(1, 2)imza benzerligi=> 0.96
dokuman(1, 5)imza benzerligi=> 0.87
dokuman(1, 8)imza benzerligi=> 0.95
dokuman(1, 11)imza benzerligi=> 1.00
dokuman(1, 14)imza benzerligi=> 0.88
dokuman(1, 17)imza benzerligi=> 0.83
dokuman(1, 20)imza benzerligi=> 0.93
dokuman(2, 5)imza benzerligi=> 0.91
dokuman(2, 8)imza benzerligi=> 0.91
dokuman(2, 11)imza benzerligi=> 0.96
dokuman(2, 14)imza benzerligi=> 0.91
dokuman(2, 17)imza benzerligi=> 0.85
dokuman(2, 20)imza benzerligi=> 0.89
dokuman(3, 6)imza benzerligi=> 0.92
dokuman(3, 9)imza benzerligi=> 0.87
dokuman(3, 12)imza benzerligi=> 0.92
dokuman(3, 15)imza benzerligi=> 0.88
dokuman(3, 18)imza benzerligi=> 0.84
dokuman(4, 5)imza benzerligi=> 0.99
dokuman(1, 3)imza benzerligi=> 0.93
dokuman(1, 6)imza benzerligi=> 0.97
dokuman(1, 9)imza benzerligi=> 0.94
dokuman(1, 12)imza benzerligi=> 0.92
dokuman(1, 15)imza benzerligi=> 0.86
dokuman(1, 18)imza benzerligi=> 0.82
dokuman(2, 3)imza benzerligi=> 0.96
dokuman(2, 6)imza benzerligi=> 0.93
dokuman(2, 9)imza benzerligi=> 0.90
dokuman(2, 12)imza benzerligi=> 0.94
dokuman(2, 15)imza benzerligi=> 0.88
dokuman(2, 18)imza benzerligi=> 0.84
dokuman(3, 4)imza benzerligi=> 0.94
dokuman(3, 7)imza benzerligi=> 0.91
dokuman(3, 10)imza benzerligi=> 0.87
dokuman(3, 13)imza benzerligi=> 0.90
dokuman(3, 16)imza benzerligi=> 0.86
dokuman(3, 19)imza benzerligi=> 0.91
dokuman(4, 6)imza benzerligi=> 0.87
dokuman(1, 4)imza benzerligi=> 0.88
dokuman(1, 7)imza benzerligi=> 0.98
dokuman(1, 10)imza benzerligi=> 0.94
dokuman(1, 13)imza benzerligi=> 0.89
dokuman(1, 16)imza benzerligi=> 0.84
dokuman(1, 19)imza benzerligi=> 0.98
dokuman(2, 4)imza benzerligi=> 0.92
dokuman(2, 7)imza benzerligi=> 0.94
dokuman(2, 10)imza benzerligi=> 0.90
dokuman(2, 13)imza benzerligi=> 0.91
dokuman(2, 16)imza benzerligi=> 0.86
dokuman(2, 19)imza benzerligi=> 0.94
dokuman(3, 5)imza benzerligi=> 0.93
dokuman(3, 8)imza benzerligi=> 0.88
dokuman(3, 11)imza benzerligi=> 0.93
dokuman(3, 14)imza benzerligi=> 0.91
dokuman(3, 17)imza benzerligi=> 0.85
dokuman(3, 20)imza benzerligi=> 0.86
dokuman(4, 7)imza benzerligi=> 0.86


```

dokuman(12, 19)imza benzerligi=> 0.90 dokuman(12, 20)imza benzerligi=> 0.86 dokuman(13, 14)imza benzerligi=> 0.97
dokuman(13, 15)imza benzerligi=> 0.93 dokuman(13, 16)imza benzerligi=> 0.90 dokuman(13, 17)imza benzerligi=> 0.89
dokuman(13, 18)imza benzerligi=> 0.88 dokuman(13, 19)imza benzerligi=> 0.87 dokuman(13, 20)imza benzerligi=> 0.83
dokuman(14, 15)imza benzerligi=> 0.96 dokuman(14, 16)imza benzerligi=> 0.93 dokuman(14, 17)imza benzerligi=> 0.92
dokuman(14, 18)imza benzerligi=> 0.91 dokuman(14, 19)imza benzerligi=> 0.86 dokuman(14, 20)imza benzerligi=> 0.82
dokuman(15, 16)imza benzerligi=> 0.97 dokuman(15, 17)imza benzerligi=> 0.96 dokuman(15, 18)imza benzerligi=> 0.95
dokuman(15, 19)imza benzerligi=> 0.84 dokuman(15, 20)imza benzerligi=> 0.81 dokuman(16, 17)imza benzerligi=> 0.98
dokuman(16, 18)imza benzerligi=> 0.97 dokuman(16, 19)imza benzerligi=> 0.82 dokuman(17, 18)imza benzerligi=> 0.99
dokuman(17, 19)imza benzerligi=> 0.81 dokuman(18, 19)imza benzerligi=> 0.80 dokuman(19, 20)imza benzerligi=> 0.95
-----
```

Threshold 0.90 icin Benzerlik orani esik degerinden fazla olan dokuman ciftleri==>

```

dokuman(1, 2)imza benzerligi=> 0.96 dokuman(1, 3)imza benzerligi=> 0.93 dokuman(1, 6)imza benzerligi=> 0.97
dokuman(1, 7)imza benzerligi=> 0.98 dokuman(1, 8)imza benzerligi=> 0.95 dokuman(1, 9)imza benzerligi=> 0.94
dokuman(1, 10)imza benzerligi=> 0.94 dokuman(1, 11)imza benzerligi=> 1.00 dokuman(1, 12)imza benzerligi=> 0.92
dokuman(1, 19)imza benzerligi=> 0.98 dokuman(1, 20)imza benzerligi=> 0.93 dokuman(2, 3)imza benzerligi=> 0.96
dokuman(2, 4)imza benzerligi=> 0.92 dokuman(2, 5)imza benzerligi=> 0.91 dokuman(2, 6)imza benzerligi=> 0.93
dokuman(2, 7)imza benzerligi=> 0.94 dokuman(2, 8)imza benzerligi=> 0.91 dokuman(2, 9)imza benzerligi=> 0.90
dokuman(2, 10)imza benzerligi=> 0.90 dokuman(2, 11)imza benzerligi=> 0.96 dokuman(2, 12)imza benzerligi=> 0.94
dokuman(2, 13)imza benzerligi=> 0.91 dokuman(2, 14)imza benzerligi=> 0.91 dokuman(2, 19)imza benzerligi=> 0.94
dokuman(3, 4)imza benzerligi=> 0.94 dokuman(3, 5)imza benzerligi=> 0.93 dokuman(3, 6)imza benzerligi=> 0.92
dokuman(3, 7)imza benzerligi=> 0.91 dokuman(3, 11)imza benzerligi=> 0.93 dokuman(3, 12)imza benzerligi=> 0.92
dokuman(3, 13)imza benzerligi=> 0.90 dokuman(3, 14)imza benzerligi=> 0.91 dokuman(3, 19)imza benzerligi=> 0.91
dokuman(4, 5)imza benzerligi=> 0.99 dokuman(4, 12)imza benzerligi=> 0.90 dokuman(6, 7)imza benzerligi=> 0.99
dokuman(6, 8)imza benzerligi=> 0.96 dokuman(6, 9)imza benzerligi=> 0.95 dokuman(6, 10)imza benzerligi=> 0.95
dokuman(6, 11)imza benzerligi=> 0.97 dokuman(6, 12)imza benzerligi=> 0.90 dokuman(6, 19)imza benzerligi=> 0.99
dokuman(6, 20)imza benzerligi=> 0.94 dokuman(7, 8)imza benzerligi=> 0.97 dokuman(7, 9)imza benzerligi=> 0.96
dokuman(7, 10)imza benzerligi=> 0.96 dokuman(7, 11)imza benzerligi=> 0.98 dokuman(7, 12)imza benzerligi=> 0.90
dokuman(7, 19)imza benzerligi=> 1.00 dokuman(7, 20)imza benzerligi=> 0.95 dokuman(8, 9)imza benzerligi=> 0.99
dokuman(8, 10)imza benzerligi=> 0.99 dokuman(8, 11)imza benzerligi=> 0.95 dokuman(8, 18)imza benzerligi=> 0.97
dokuman(8, 20)imza benzerligi=> 0.92 dokuman(9, 10)imza benzerligi=> 1.00 dokuman(9, 11)imza benzerligi=> 0.94
dokuman(9, 19)imza benzerligi=> 0.96 dokuman(9, 20)imza benzerligi=> 0.91 dokuman(10, 11)imza benzerligi=> 0.94
dokuman(10, 19)imza benzerligi=> 0.96 dokuman(10, 20)imza benzerligi=> 0.91 dokuman(11, 12)imza benzerligi=> 0.92
dokuman(11, 19)imza benzerligi=> 0.98 dokuman(11, 20)imza benzerligi=> 0.93 dokuman(12, 13)imza benzerligi=> 0.96
dokuman(12, 14)imza benzerligi=> 0.95 dokuman(12, 15)imza benzerligi=> 0.91 dokuman(12, 19)imza benzerligi=> 0.90
dokuman(13, 14)imza benzerligi=> 0.97 dokuman(13, 15)imza benzerligi=> 0.93 dokuman(13, 16)imza benzerligi=> 0.90
dokuman(14, 15)imza benzerligi=> 0.96 dokuman(14, 16)imza benzerligi=> 0.93 dokuman(14, 17)imza benzerligi=> 0.92
dokuman(14, 18)imza benzerligi=> 0.91 dokuman(15, 16)imza benzerligi=> 0.97 dokuman(15, 17)imza benzerligi=> 0.96
dokuman(15, 18)imza benzerligi=> 0.95 dokuman(16, 17)imza benzerligi=> 0.98 dokuman(16, 18)imza benzerligi=> 0.97
dokuman(17, 18)imza benzerligi=> 0.99 dokuman(19, 20)imza benzerligi=> 0.95
-----
```

K====>0
0

Process finished with exit code 0