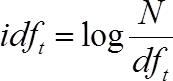
Aşağıdaki cümleleri içeren bir listeniz olduğunu varsayın. Text isimli değişkeni program kodunuzun başında kullanın.

**text=["duran duran sang wild boys in 1984","wild boys don't remain forever wild","who brought wild flowers","it was john krakauer who wrote in to the wild"]**

Bu listedeki benzersiz terimleri tespit eden ve aşağıdaki formüllere göre cümleleri ifade eden bir kod yazacaksınız.

* + *N*- toplam cümle sayısı
  + *dft* - t terimi geçen cümle sayısı (Örnek: **wild** isimli terim)

Bu formülde, logaritmanın tabanı da 10'dur. Benzersiz terimleri saptadıktan ve her terim için *idft* değerini hesapladıktan sonra, tüm cümleleri aşağıdaki *tfidf* değerlerine göre ifade edeceksiniz. Burada *tft,d*, ilgili cümlede t teriminin geçme sayısını temsil eder.



Son olarak, kodun çıktısı aşağıdaki gibi olacaktır:

['duran', 'sang', 'wild', 'boys', 'in', '1984', "don't", 'remain', 'forever', 'who', 'brought', 'flowers', 'it', 'was', 'john', 'krakauer', 'wrote', 'to', 'the']

[1.204, 0.602, 0.0, 0.301, 0.301, 0.602, 0.0, 0.0, 0.0, 0.0, 0.0, 0.0, 0.0, 0.0, 0.0, 0.0, 0.0, 0.0, 0.0]

[0.0, 0.0, 0.0, 0.301, 0.0, 0.0, 0.602, 0.602, 0.602, 0.0, 0.0, 0.0, 0.0, 0.0, 0.0, 0.0, 0.0, 0.0, 0.0]

[0.0, 0.0, 0.0, 0.0, 0.0, 0.0, 0.0, 0.0, 0.0, 0.301, 0.602, 0.602, 0.0, 0.0, 0.0, 0.0, 0.0, 0.0, 0.0]

[0.0, 0.0, 0.0, 0.0, 0.301, 0.0, 0.0, 0.0, 0.0, 0.301, 0.0, 0.0, 0.602, 0.602, 0.602, 0.602, 0.602, 0.602, 0.602]

Yukarıda görüldüğü gibi, benzersiz terimlerin listesi ekrana yazdırılır. Daha sonra her cümle, uzunluğu benzersiz terimlerin sayısına eşit olan bir matris ile ifade edilir. Karşılık gelen satırlardaki değerler, her cümledeki terimlerin **tfidf** değerleridir.