AF-SM-04\_05.2020.G

|  |  |
| --- | --- |
| **კურსის სახელწოდება:** | ხელოვნური ინტელექტი და მანქანური სწავლების საწყისები |
| **ლექტორი:** | დავით დათუაშვილი |
| **სტუდენტი:** |  |

1. მიჰყევით ინსტრუქციებს და გაუშვით თვითოეული კოდი მიყოლებით, ასევე უპასუხეთ ქვევით მოცემულ კითხვებს (10 ქულა)

* მოცემულია ქვევით სიის ტიპის მასივი, რომელიც შეიცავს ტექსტებს :

documents = ["This little kitty came to play when I was eating at a restaurant."**,** "Merley has the best squooshy kitten belly."**,** "Google Translate app is incredible."**,** "If you open 100 tab in google you get a smiley face."**,** "Best cat photo I've ever taken."**,** "Climbing ninja cat."**,** "Impressed with google map feedback."**,** "Key promoter extension for Google Chrome."]

მოცემული თვითოეული ტექსტი აღწერს გარკვეული შინაარსის მომენტს.(kitty არის კნუტი).განსაღვრეთ თუ მოცემულ დოკუმენტის ტექსტებში რამდენ განსხვავებულ მომენტზე არის საუბარი და KMeans ის კლასტერებში გაუწერეთ ის რაოდენობა .მანამდე მიჰყევით ქვევით მოცემულ ინსტრუქციას :

ტექსტი დავყავი 2 კლასტერად კატის და Google-ს.

* გამოიძახეთ შემდეგი ბიბლიოთეკები :

from sklearn.feature\_extraction.text import TfidfVectorizer  
from sklearn.cluster import KMeans  
from sklearn.metrics import adjusted\_rand\_score

* გადაიყვანეთ ტექსტები რიცხვებში შემდეგი ფუნქციის გამოყენებით(ითვლის სიტყვების სიხშირეების შესაბამის წონებს)

vectorizer = TfidfVectorizer(stop\_words='english')  
X = vectorizer.fit\_transform(documents)

* კლასტერინგისთვის გამოიყენეთ შემდეგი ბრძანება :

model = KMeans(n\_clusters=true\_k**,** init='k-means++'**,** max\_iter=**100,** n\_init=**1**)  
model.fit(X)

სადაც true\_k არის თქვენს მიერ შერჩეული კლასტერების რაოდენობა

True\_k-ს მნიშვნელობა იქნება 2 რადგან ტექსტი დავყავით ორ ნაწილად.

* მოახდინეთ პროგნოზირება შემდეგი თქვენს მიერ შერჩეული ტექსტის შემდეგი ბრანების გამოყენებით :

Y = vectorizer.transform([write your text in here])  
prediction = model.predict(Y)  
print(prediction)

კვადრატულ ფრჩხილებში უნდა ჩაწეროთ ტექსტი ბრწყალებში (მოიფიქრეთ ტექსტი ისე რომ ახლოს იყოს დატრენინგებულ ტექსტებთან)

Vectorizer.transform-ში თუ ჩავწერდით წინადადებას რომელიც შეეხება კატას მასინ ალგორითმი დააბრუნებს 0-ს ხოლო თუ გუგლს მაშინ 1-ს რადგან სულ გვაქვს ორი კლასტერი. შესაბამისად შესაბამის ტექსტზე დააბრუნა თავისი მნიშვნელობა.

A screenshot of a computer program

Description automatically generated with medium confidence A screenshot of a computer program

Description automatically generated with medium confidence