

Programozás 2 – 11. gyakorlat

(2020. 11. 24.)

1. Olvassa be a TrumpTweets.csv nevű állományt, és végezze el rajta a következő előfeldolgozó feladatokat.
 - Az állományból nyerte ki a Twitter üzenetek szövegét
 - A szöveget alakítsa kisbetűssé
 - Távolítsa el a felesleges tag-eket (http, pic., #, @)
 - Távolítsa el a felesleges írásjeleket
2. Hozzon létre a Twitter üzenetek szavaiból egy Bag of Words listát, és számolja ki a különböző szavak arányát az összes szóhoz képest.
 - Távolítsa el a Bag of Words halmazból a kötőszavakat, és végezze el újra a szavak arányának a vizsgálatát.
3. Végezze el a következő vizualizációkat a Twitter üzenetek szavainak felhasználásával:
 - Plotolja ki egy függvény segítségével a leggyakrabban előforduló n szót
 - Plotolja ki a leggyakrabban előforduló n szó eloszlását az összes Twitter üzenetre vonatkozóan
 - Készítsen egy Pie Chartot, ami ábrázolja a szavak arányát annak megfelelően, hogy az pozitív, negatív vagy semleges érzelmű szó-e.