

4. házi feladat

(kit?z?/javító: Pálovics Róbert)

Gyakori feladat, hogy két halmazról kell eldöntenünk, mennyire "http://wiki.math.bme.huhasonlóak"http://wiki.math.bme.hu. Például

- meg kell mondanunk, hogy két vásárló mennyire vett azonos termékeket
- két ember mennyire olvas hasonló könyvet, hallgat azonos zenét, ...

Az alapfeladat, hogy eldöntsük Sage-ben két listáról, hogy "http://wiki.math.bme.humennyire tartalmaznak azonos elemeket"http://wiki.math.bme.hu. (A listák egy-egy matematikai értelemben vett halmaznak felelnek meg.)

- Multiplicitás nélküli halmazok összehasonlításának egyik módja a Jaccard index kiszámítása. Ez egyszer?en két halmaz metszetének és uniójának hányadosa: Jaccard index
- Írj sage függvényt, mely
- bemenetként megkap két listát,
- kiszámolja a két lista Jaccard-indexét,
- nem dolgozik multiplicitással, tehát például a ['ferrari','mercedes','ferrari', 'mercedes'] és ['mercedes','williams','mercedes'] listák hasonlósága 1/3.
- A függvény neve legyen "http://wiki.math.bme.hu"jaccard_index"http://wiki.math.bme.hu.
- Számítsd ki a függvénnyel a ['U2','Isis', 'Slash','sleepmakeswaves', 'Coldplay', 'Maybeshewill'] és ['Isis','Slash','Amplifier', 'Led Zeppelin', 'U2'] listák hasonlóságát!
- Írj "http://wiki.math.bme.hu"jaccard_multi"http://wiki.math.bme.hu függvényt, mely a metszet és unió számítása esetén a halmazok elemeit multiplicitással veszi figyelembe. Tehát az el?z? példára visszatérve: ['ferrari','mercedes','ferrari', 'mercedes'] és ['mercedes','williams','mercedes'] listák hasonlósága most 2/5.
- Ellen?rizd le az 1/3 (jaccard_index) és 2/5 (jaccard_multi) eredményket az autós listákra.
- Segítség: a gyakorlat "http://wiki.math.bme.hu"szószámláló"http://wiki.math.bme.hu és "http://wiki.math.bme.hu"szótár összehasonlító"http://wiki.math.bme.hu függvényeinél látott ötletek segíthetnek. Azonban a feladat nagyon sokféleképp megoldható, és ennek megfelelően sokféle jó megoldást el tudunk fogadni.

A házit most a következőképp küldjétek: A sage worksheetet a szokásos konvenció alapján (T2_HF4_kovacspisti) nevezzétek el, majd osszátok meg a sage szerveren **palovics** felhasználóval. Ha bármi kérdés merül fel, jelezzetek, írjatok nyugodtan.

Update: Ha meg van nyitva egy worksheet-etek, akkor a megosztó gomb, jobb fent a kék Share gomb, utána írjátok a box-ba a nevet akivel meg akarjátok osztani, majd az Invite Collaborators gomb.