

1. feladat: Döntsd el, hogy a lentebbi pontok osztályozáshoz vagy clusterezéshez tartoznak! Lehetnek olyanok, amelyek nem egyértelműek, azokat jelöld.

1. Azonosítsd a betegek különböző típusú rákos megbetegedéseit a laboratóriumi eredmények alapján. = **osztályozás (címke; rákos megbetegedés)**
2. Csoportosítsd a vásárlókat vásárlási szokásaik alapján anélkül, hogy előzetes címkéket használnál. = **klaszterezés**
3. Osztályozd a leveleket spam és nem spam kategóriákba azok tartalma alapján. = **osztályozás (címke; spam vagy nem-spam)**
4. Találd meg a közös jellemzőket birtokló termékcsoportokat egy webáruház kínálatában, címkék nélkül. = **klaszterezés**
5. Azonosítsd és csoportosítsd a különböző nyelveken íródott szövegeket. = **osztályozás (címke; nyelv)**
6. Találd ki algoritmust, amely képes azonosítani a különböző autómárkákat a járművek jellemzői alapján. = **osztályozás (címke; autómárka)**
7. Osztályozd a filmeket műfajuk alapján azok leírása alapján. = **osztályozás (címke; film műfaja)**
8. Csoportosítsd az ügyfeleket koruk és vásárlási gyakoriságuk alapján. = **klaszterezés**
9. Határozd meg, hogy egy adott e-mail üzenet fontos-e vagy sem. = **osztályozás (címke; fontos vagy nem fontos)**
10. Azonosítsd a városokat lakosságszámuk és földrajzi helyzetük alapján. = **mindkettő lehet (Ha osztályozás, akkor a címke a város. Klaszterezni akkor lehet pl., ha a lakosok pontokként jelennek meg a térképen.)**
11. Osztályozd a diákokat teljesítményük alapján különböző szintű osztályokba. = **osztályozás (címke; teljesítményi szintek)**
12. Hozz létre csoportokat az állatok között azok élőhelye alapján. = **mindkettő lehet (Ha több különböző állat van, akkor lehet osztályozás, amiben az állat neve a címke. Egyféle állat esetén klaszterezés.)**
13. Azonosítsd a különböző éghajlati zónákat a világon meteorológiai adatok alapján. = **mindkettő lehet (Ha osztályozás, akkor a címke a zóna neve.)**
14. Csoportosítsd a növényeket azok virágzásának ideje szerint. = **osztályozás (címke; a növényfaj neve)**
15. Határozd meg a sportolók teljesítményét különböző sportágakban. = **mindkettő lehet (Ha osztályozás, akkor a címke valamilyen teljesítményszint.)**
16. Azonosítsd a csillagképeket az éjszakai égbolton. = **osztályozás (címke; csillagkép neve)**
17. Osztályozd a mobilalkalmazásokat kategóriák szerint, mint például játék, üzleti, oktatás. = **osztályozás (címke; alkalmazáskategória)**
18. Hozz létre csoportokat a számítógépes vírusok között a terjedési módjuk alapján. = **klaszterezés**
19. Azonosítsd a különböző pénzügyi tranzakciótípusokat, mint például vásárlás, átutalás, befizetés. = **osztályozás (címke; tranzakció típusa)**
20. Csoportosítsd a zenei stílusokat a hangzásuk alapján. = **osztályozás (címke; zenei stílus)**
21. Azonosítsd a különböző tengeri élőlényeket azok fizikai jellemzői alapján. = **osztályozás (címke; tengeri élőlény fajtája)**
22. Osztályozd az ételeket azok fő összetevője alapján, például hús, zöldség, tészta. = **kérdés értelmezésétől függően mindkettő lehet (Ha a "hús, zöldség, tészta" fő összetevő és nincs célkategória; klaszterezés. Ha a "hús, zöldség, tészta" a címke, akkor osztályozás.)**
23. Csoportosítsd a hoteleket azok szolgáltatásai és árkategóriája alapján. = **mindkettő lehet (Ha osztályozás, akkor a címke lehet pl. a hotel csillagozása.)**
24. Határozd meg, hogy egy adott növény mérgező-e vagy sem. = **osztályozás (címke; mérgező vagy nem mérgező)**

25. Azonosítsd a különböző gépjárműveket azok üzemanyag-fogyasztása alapján. = **osztályozás (címke; gépjármű-típus)**
26. Osztályozd a könyveket a kiadásuk éve alapján. = **mindkettő lehet (Osztályozás esetén lehet a címke például; könyv műfaja, terjedelme, író neve, stb...)**
27. Csoportosítsd a festményeket azok stílusa és a festő korának alapján. = **klaszterezés**
28. Azonosítsd a különböző országokat azok GDP-je alapján. = **osztályozás (címke; ország neve)**
29. Osztályozd a diákokat az általuk választott szakkörök alapján. = **osztályozás (címke; szakkör neve)**
30. Hozz létre csoportokat a filmrendezők között a filmjeik általános témája alapján. = **klaszterezés**

2. feladat: Döntsd el, hogy a következő képességek és feladatok Data Scientisthez, ML Engineerhez vagy Data Engineerhez tartozik inkább! Lehetnek olyanok, amelyeket többhöz is sorolhatsz.

1. Adatok előkészítése és tisztítása. = **Data Engineer**
2. Statikus adatelemzés és hipotézisvizsgálat. = **Data Scientist**
3. Big Data technológiák, mint például Hadoop és Spark. = **Data Engineer**
4. Adattárházak és ETL folyamatok kialakítása. = **Data Engineer**
5. Deep learning modellek production-be helyezése. = **ML Engineer**
6. Adatvizualizáció és jelentéskészítés. = **Data Scientist és ML Engineer**
7. Idősoros elemzés és elprejelzés. = **Data Scientist**
8. Adatázis-kezelés és lekérdezések SQL-ben. = **Data Engineer**
9. Adatbiztonság és adatkezelési szabályozások. = **Data Engineer és ML Engineer**
10. Algoritmus optimalizálás és teljesítményelemzés. = **Data Scientist és ML Engineer**
11. Kísérleti kialakítás és A/B tesztelés. = **ML Engineer**
12. Felhőalapú szolgáltatások, mint az AWS vagy Azure. = **ML Engineer**
13. ML rendszerek pipeline-jainak kialakítása. = **Data Engineer**
14. Adatpiaci trendek és üzleti intelligencia. = **Data Scientist**
15. Generative AI. = **Data Scientist**
16. Adatbányászat és komplex adatelemzés. = **Data Scientist**
17. Adatintegráció és adatáramlások kezelése. = **Data Engineer**
18. Deployment technikák. = **ML Engineer**
19. Docker és Kubernetes konténerizációs technológiák. = **ML Engineer**
20. Prediktív modellezés és valószínűségi statisztikák. = **Data Scientist**

3. feladat: Írj legalább 10 példát regresszióra!

1. Műanyagpalackokban lágyító mennyiség meghatározása a palack mechanikai tulajdonságai alapján.
2. Elektronikai eszköz várható élettidejének előrebecslése a vásárlás helye és a vásárló adatai alapján.
3. Adott kép alapján megbecsülni, hogy a pixelek hány százaléka tartozik emberekhez.
4. Weboldal popularitásának előrebecslése a weboldal szöveges és vizuális tartalma alapján.
5. Termék várható fogyásának a becslése az adott hónapra az előző hónapban eladott mennyiségek alapján.
6. Kurzor x-y pozíciójának becslése 2 másodperc múlva az előzetes kurzortrajektória alapján.
7. Alvásminőségi pontszám előrejelzése az adott napon fogyasztott ételek tápanyagtartalma szerint.
8. Gazdasági újságok szövegei alapján egy ország következő évi egy főre jutó GDP-jének előrejelzése.

9. YouTube csatornára feliratkozók száma és videó hossza alapján a várható nézettség becslése.
10. Hőmérsékleti és csapadékmennyiségi trendek alapján méhek méztermelésének előrejelzése.

3. feladat: Írj legalább 10 példát reinforcement learningre!

1. Robotporszívó útvonaltervezése.
(Környezet állapottere: a szoba, különböző állapotokban megtisztítva. Akciótér: porszívó sebessége, gyorsulása, szívási sebessége. Jutalom: a szoba tisztasági mértéke.)
2. Sakkjáték vagy Go-játék során nyerő stratégia előállítás.
(Környezet állapottere: játéktábla különböző konfigurációi. Akciótér: lépés a bábukkal/kövekkel. Jutalom: a játék becsült véghelyezete.)
3. Önvezető autó vezetési stratégiái.
(Környezet állapottere: az autót körülvevő tárgyak, táblák, jelzések, más járművek. Akciótér: kormány mozdítása, gáz/fék/kuplung, kézifék, stb... Jutalom: a véghelyezetől való távolság, környezet és utasok épsége.)
4. Porszívóügynök eladásai statisztikáinak javítása.
(Környezet állapottere: a vevő hangulata, nonverbális jelzései, mondatai. Akciótér: elmondható természetes szövegek. Jutalom: a vevő hajlandósága az adás-vételre.)
5. Large language modellek esetén túsztárgyalási stratégia képesség javítása.
(Környezet állapottere: a túszt és a fogvatartó hangulata, nonverbális jelzései, mondatai. Akciótér: elmondható természetes szövegek. Jutalom: fogvatartó hajlandósága a túszt átadására.)
6. Részvénykereskedési modell kidolgozása.
(Környezet állapottere: múltbéli részvényértékek, esetleg gazdasági hírek. Akciótér: eladás/vásárlás. Jutalom: profit.)
7. Sudoku kitöltése.
(Környezet állapottere: sudoku tábla állapota. Akciótér: adott helyre adott karakter lerakása és/vagy eltávolítása. Jutalom: tábla validitása és kitöltöttségi mértéke.)
8. Weboldal látogatottságának növelése.
(Környezet állapottere: összes lehetséges weboldal. Akciótér: weboldal szerkesztése, digitális elemek lerakása és/vagy eltávolítása. Jutalom: látogatottság.)
9. Large language modellek esetén nyelvtanítás.
(Környezet állapottere: diákok nyelvtudása feladatokon keresztül felmérve. Akciótér: nyelvi feladatok összeállítása. Jutalom: diákok teljesítménye referenciateladatokon.)
10. Filmajánlórendszer kidolgozása.
(Környezet állapottere: felhasználói adatok, pl. megnézett filmek és sorozatok adatai, filmválogatás időpontja, adott film oldalán eltöltött idő mennyisége, stb... Akciótér: ajánlható filmek listája. Jutalom: ajánlott film lejátszása.)