|  |  |  |  |
| --- | --- | --- | --- |
| Szempontok | Python | R | AutoML |
| vizualizációs lehetőségek | * szkriptből nincs lehetőség vizualizáció létrehozására Serviceben * limitációk * nagyobb testreszabhatóságot eredményez * külső dependencia * beépített lehetőségek (pl. cross referencing) hiánya | * szkriptből nincs lehetőség vizualizáció létrehozására Serviceben * limitációk * nagyobb testreszabhatóságot eredményez * külső dependencia * beépített lehetőségek (pl. cross referencing) hiánya | * dataflowból serviceben nem lehet létrehozni reportot * forecasting desktopon – limitációkkal * desktopon kell elkészíteni (kivéve a training reportot) * bonyolult folyamat |
| developer experience | * serviceben nincs lehetőség a scriptelésre, csak kerülőutak (adatforrásban script) * egyszerű beépített lehetőségek * egyszerű telepítés * szegényes editor, fejlesztői környezet * rossz dokumentáció – nem kiszámítható * publikus adatforrásokat kell használni * serviceben ha frissül az adatot, újra kell tölteni az oldalt * sok fórumolvasást igényel * adatforrásba építve: bármilyen script editor használható, rugalmas * power query editorban: nem igényel + erőforrást * sok különböző rendszer – sok különböző licensz, bejelentkezés | * serviceben nincs lehetőség a scriptelésre, csak kerülőutak (adatforrásban script) * egyszerű beépített lehetőségek * egyszerű telepítés * szegényes editor, fejlesztői környezet * rossz dokumentáció – nem kiszámítható * publikus adatforrásokat kell használni * serviceben ha frissül az adat, újra kell tölteni az oldalt * sok fórumolvasást igényel * adatforrásba építve: bármilyen script editor használható, rugalmas * power query editorban: nem igényel + erőforrást * R nyelv korlátai * sok különböző rendszer – sok különböző licensz, bejelentkezés | * kódmentes * cross-platform (service miatt) * modellezés + prediction + vizualizációhoz szükséges service és desktop is * nem jól dokumentált * debugolás nehézkes |
| modell létrehozás (training, tesztelés, menthetőség) | * service: nem frissíthető, hozzáférhető (csak ha adatforrásban van a script) * komoly háttértudást igényel (programozás + statisztika + ML) * beépített run script editor nem erre készült, hanem adattranszformációra * nem menthető a modell (csak külső scriptnél) * kül. lehetőségek (több helyen tudjuk futtatni) * modell minden alkalommal frissül, mikor a dataset frissül * nem túl könnyen fenntartható/karbantartható * kis scriptek futtatására alkalmas * modul dependencia problémák | * service: nem frissíthető, hozzáférhető (csak ha adatforrásban van a script) * komoly háttértudást igényel (programozás + statisztika + ML) * beépített run script editor nem erre készült, hanem adattranszformációra * nem menthető a modell (csak külső scriptnél) * kül. lehetőségek (több helyen tudjuk futtatni) * modell minden alkalommal frissül, mikor a dataset frissül * nem túl könnyen fenntartható/karbantartható * kis scriptek futtatására alkalmas * multi-modulos programozás lehetőségeinek korlátai R-ben problémásak | * regresszióból nehézkes a forecasting előállítása * csak serviceben lehet modellt létrehozni * kivétel: forecasting egyenes, korlátolt lehetőségek * kevés modellezési lehetőség, nem flexibilis, nem testreszabható * training reportban a modell statisztikái láthatóak rögtön |
| adatelőkészítés / adatelérés (pl. szerverről, külső forrásból, adattisztítás, stb.) | * service: workflowt létrehozni csak premium/pro accounttal lehet, premium workspaceben, utána is datasetté kell alakítani, le kell tölteni * desktopon sokféle beépített lehetőség import + adattisztítás szempontjából * adatkonverziós problémák * sok lehetőség – több problémalehetőség * ha servicebe feltöltünk datasetet és szeretnénk frissíteni, personal gateway kell * applied stepek hibára futhatnak adatfrissítésnél * rossz dokumentáció | * service: workflowt létrehozni csak premium/pro accounttal lehet, premium workspaceben, utána is datasetté kell alakítani, le kell tölteni * desktopon sokféle lehetőség import + adattisztítás szempontjából * adatkonverziós problémák * sok lehetőség – több problémalehetőség * ha servicebe feltöltünk datasetet és szeretnénk frissíteni, personal gateway kell * applied stepek hibára futhatnak adatfrissítésnél * rossz dokumentáció | * adatkonverziós problémák (lokalizáció) * több datasourcenál kombinálási lehetőség nehézkes * serviceben történik (dataflow), predikció előkészítése a desktopon * limitált data source lehetőségek serviceben |
| Service – desktop integráció (desktop -> service, scheduled refresh lehetősége, modell menthet -> desktop) | * scheduled refresh lehetősége gateway használhatával * egyirányú kapcsolat – desktop -> service (publish + refresh gomb) * data flow létrehozása lehetséges serviceben, de le kell tölteni és újra feltölteni * sokféle külső integráció, néha bonyolult * gateway constraintek: állandó futási igény, 4 egymást követő hiba után letiltja a frissítést, 2 hónap meg nem tekintés után leáll * modelleken, dataseteken serviceben nem tudunk módosítani * onedriveos source esetén ha változik a struktúra, elérhetetlenné válnak az adatok | * scheduled refresh lehetősége gateway használhatával * egyirányú kapcsolat – desktop -> service (publish + refresh gomb) * data flow létrehozása lehetséges serviceben, de le kell tölteni és újra feltölteni * sokféle külső integráció, néha bonyolult * gateway constraintek: állandó futási igény, 4 egymást követő hiba után letiltja a frissítést, 2 hónap meg nem tekintés után leáll * modelleken, dataseteken serviceben nem tudunk módosítani * onedriveos source esetén ha változik a struktúra, elérhetetlenné válnak az adatok | * data flow: le kell tölteni és újra feltölteni * automata refresh: trainelt tábla változásánál lefut újra a trainelés + applyolás, ha a valós predikciós táblát változtatjuk, lefut az apply model, a report is refreshelődik (desktopon) |
| licenszek (pl. Pro, Premium változatban) | * premium: serviceben több funkció elérhető (pl. dataflow) * nem befolyásolja a licenszverzió: a scriptelés desktopon freeben, proban és premiumban is elérhető, megosztáshoz pro kell vagy premium | * premium: serviceben több funkció elérhető (pl. dataflow) * nem befolyásolja a licenszverzió: a scriptelés desktopon freeben, proban és premiumban is elérhető, megosztáshoz pro kell vagy premium | * autoML csak premiummal * külső azure ML csak premiummal érhető el – csak serviceben * forecast lehetőség desktopon reportinghoz free |