

Szakmai beszámoló

A 2019-es, 8 hetes nyári szakmai gyakorlat célja különböző kompetenciák megszerzése volt. Ezek közé tartozott az *SAP HANA* technológia megismerése, működésének megértése dokumentáció alapján, a képesség arra, hogy a munkánkat egyéb, releváns technológiák segítségével egészítsük ki, és csapatként együtt tudjunk működni.

Munkánkat az *SAP*-val való ismerkedéssel kezdtük. Eleinte témavezetőink hasznos linkekkel láttak el minket, ennek megfelelően az első héten javarészt az elméleti tudás megalapozása folyt. Tanáraink továbbá adtak nekünk egy vázlatot arról, hogy a későbbi, nagy projektünkben milyen szerepkörök lesznek elérhetők (adatbányászat, front-end fejlesztés stb.), és már ekkor felmérték a hallgatók érdeklődési körét (minden résztvevő megnevezett egy elsődleges és egy másodlagos területet, amelyben szívesen dolgozna).

Az elméleti alapozás után témavezetőink segítettek elindítani az első kisebb projektet – együtt létrehoztuk az első táblánkat, és megkaptuk az első feladatunkat. Ehhez közzétettek számunkra egy csv formátumú adatbázist, mely vállalatok különböző adatait tartalmazta. A mi feladatunk volt ezeket az adatokat beolvasni, szükség esetén javítani, és *odata* segítségével, valamilyen földrajzi vonatkozásban megjeleníteni. Szükség volt továbbá kalkulációs nézet létrehozására *SAP*-ban.

Ezeket az alfeladatokat csoportokra osztottuk szét; amikor az egyik csoport sikeresen megcsinált valamit, azt bemutatta a többieknek. Így végül szinte egy egységes csoportprojekt jött létre, melyet a végén mindenki a saját meglátása szerint tudott tovább csiszolni.

Ez a „miniprojekt” nemcsak az *SAP* mélyebb megértése szempontjából volt fontos, hanem azért is, hogy minden hallgató belekóstolhasson mindenféle területbe, amelyet a témavezetők korábban felvázoltak. El kellett döntenünk, melyik feladatkört választjuk.

Engem kezdetben az adatbányászat témakör érdekelt, és ehhez tartottam is magam. Akik velem együtt így döntöttek, azt az utasítást kapták, hogy kezdjenek el megismerkedni az *SAP HANA Predictive Analysis Library*-vel (*PAL*), amely az *SAP HANA* saját adatbányász-függvényeinek könyvtára.

Így eleinte ismételten dokumentáció-olvasással foglalkoztunk. A függvények meghívásához szükségünk volt egy script szerverre, melyet egyik tanárunk hamarosan létrehozott nekünk; így elkezdhattunk a gyakorlatban is kísérletezni az olvasottakkal.

Az utolsó három hétre kaptunk egy új adatbázist, amely szenzorok időjárás- és légszennyezés-méréseit tartalmazta adott időpontokban. Ezen kellett a *PAL* könyvtárról való tudásunkat kamatoztatni.

Itt az adatbányász csapat két csoportra szakadt, én magam Tihanyi Balázs hallgatótársammal dolgoztam. Az adatbázison futtattunk egy klaszterezési eljárást (*Affinity Propagation*, röviden *AP*), egy kiugró adat vizsgálatot (*Anomaly Detection*) és egy idősor analízist (*ARIMA*). Célunk elsősorban az volt, hogy a *PAL* három különböző területéről sajátítsunk el egy-egy eljárást, értsük meg működésüket mélyebben. Munkánk során volt, hogy egy hibás gondolkodás vezetett rá a jobb megértésre. Minderre bővebben a *pal_dokumentacio* című dokumentum tér ki külön.

Összefoglalva elmondható, hogy a gyakorlat során megtanultunk különböző technológiákat dokumentáció alapján elsajátítani, szükség szerint kombinálni, működésüket megérteni, és csapatban dolgozni. Fejlődésünknek ugyan határt szabott az időkorlát, de a munkát mindig lehet folytatni. Úgy vélem, a kezdeti célkitűzéseket sikerült teljesítenem, hasznos kompetenciákra tettem szert.