

Lili Horváth

Data Engineer gyakornok

✉lili.horvath@example.com

☎+36 30 708 31

📍Veszprém

MUNKATAPASZTALAT

Data Engineer gyakornok

01/2022 – Jelen

TanulóHíd, Veszprém

- Felépítettem a Python és Pandas alapú adatfolyamokat a tanulási események minősége kezelésére a TanulóHíd jelenlegi csapatában; a termék- és platformmérnökökkel 410 ezer tanulási esemény léptékben dolgoztam, így a határidőre elkészült futások aránya 92%-ról 99,6%-ra nőtt.
- Újraterveztem a SQL modelleket és particionálást a „Kurzusaktivitási pipeline” projektben éles üzemeltetési felelősséggel; 9 fős mérnöki csapattal együttműködve a medián lekérdezési idő 24 másodpercről 4,1 másodpercre csökkent.
- Mérhetővé tettem a séma-, frissesség- és lineage-ellenőrzéseket 38 éles táblán a tanulási események minősége környezetében; 14 elemzővel egyeztetve 63%-kal kevesebb adateltérést mértünk.
- Automatizáltam a backfill-folyamatokat, riasztásokat és runbookokat a Airflow futásaihoz Veszprém csapatában; 12 piacot támogattunk, és negyedévente 86 ügyeleti órát takarítottunk meg.

Data Platform Engineer

04/2019 – 12/2021

TanulóHíd Adatstúdió, Veszprém

- Felépítettem a Python és Pandas alapú adatfolyamokat a tanulási események minősége kezelésére a TanulóHíd korábbi csapatában; a termék- és platformmérnökökkel 410 ezer tanulási esemény léptékben dolgoztam, így a határidőre elkészült futások aránya 92%-ról 99,6%-ra nőtt.
- Újraterveztem a SQL modelleket és particionálást a „Kurzusaktivitási pipeline” projektben migrációs felelősséggel; 6 fős mérnöki csapattal együttműködve a medián lekérdezési idő 24 másodpercről 4,1 másodpercre csökkent.
- Mérhetővé tettem a séma-, frissesség- és lineage-ellenőrzéseket 38 éles táblán a tanulási események minősége környezetében; 5 alkalmazáscsapattal egyeztetve 63%-kal kevesebb adateltérést mértünk.
- Automatizáltam a backfill-folyamatokat, riasztásokat és runbookokat a Airflow futásaihoz Veszprém csapatában; 4 üzleti egységet támogattunk, és negyedévente 86 ügyeleti órát takarítottunk meg.

PROJEKTEK

Kurzusaktivitási pipeline

Dokumentált adatterméket építettem a tanulási események minősége igényeire, tesztelt transzformációkkal és elemzést támogató szállítási modellel.

Technológiák: Python, Pandas, SQL, Airflow

KÉSZSÉGEK

Technológiák: Python, Pandas, SQL, Airflow, Docker

Szállítás: Adatmodellezés, Adatminőség, CI/CD, Megfigyelhetőség, Dokumentáció

Együttműködés: Követelmény-workshop, Incidenskezelés, Mentorálás, Agilis szállítás

OKTATÁS

Pannon Egyetem

09/2015 – 06/2019

Gazdaságinformatikus BSc Adatmérnökség és szoftverrendszerek • Veszprém

TANÚSÍTVÁNYOK

AWS Certified Cloud Practitioner – Szakmai tanúsító szervezet

05/2024

Data Reliability Workshop – Data Engineering Guild

08/2023

NYELVEK

Magyar (Anyanyelvi)

Angol (Munkanyelv)