|  |  |
| --- | --- |
| Miskolci Egyetem  Gépészmérnöki és Informatikai Kar  Általános Informatikai Intézeti  *3515 Miskolc-Egyetemváros* |  |

**SZAKDOLGOZAT**

*Feladat címe:*

Mesterséges intelligencia alapú oktatást támogató alkalmazás tervezése

*Készítette:*

Szilvási Péter

*MSc szintű mérnökinformatika szakos*

*Alkalmazásfejlesztő szakirányos hallgató*

*Témavezető:*

Dr. Kovács László

*tanszékvezető, egyetemi tanár*

**Miskolc, 2020.**

# Bevezetés

## Téma választás ismertetése

placeholder

## Miért kell támogatni az oktatást?

placeholder

## Oktatás fontossága

placeholder

## Mesterséges intelligencia és az oktatás

placeholder

# Problémakör

## Az adat az új üzemanyag

placeholder

## A gépi tanulás folyamata

Machine learning pipeline

Adat pipeline

(Többi pipeline)

## Az adatforrás a kezdetek kezdete

Mik lehetnek az adat

Honnan szerezhető

## Saját adatforrásom ismertetése

Adat alapvetően: Ált leírás: null érték, zajok

Adathalmaz ismertetése

Átvezetés:

*Többet mond egy kép*

*mint száz bekezdés*

# **Reprezentációs módszerek**

## Miért hasznos az adatok megjelenítése?

Az adattudományban az egyik legfontosabb készség az adatok eloszlásának és összetettségének vizualizálása és megértése. Manapság azonban az online tanfolyamok nagy része elsősorban a gépi tanulásra és az algoritmus működésére összpontosít.

Ez nem azt jelenti, hogy a gépi tanulás ismerete elhanyagolandó. Nyilvánvalóan a legjobb adattudós könnyedén eligazodik a gépi tanulási algoritmusok nagy részével. Mindenesetre az adattudomány nem korlátozódik le a gépi tanulásra, inkább az egy képesség. A készség alatt azt kell érteni, hogy mennyire érted a probléma területet (domain), mennyire tudsz elmélyülni az adatokban. Elkapni a rejtett összefüggéseket és megtalálni az adatokban rejlő üzenetet. Az adat beszél magáról. A legjobb mód arra, hogy szóra bírjuk az **adatmegjelenítés**.

***Mi az adatmegjelenítés?***

Az adatmegjelenítés az információ (adatok) felvételének és vizuális kontextusba, például térképbe vagy grafikonba helyezésének módszere. A fő cél a nagy adatkészletek vizuális grafikába történő szűrése, hogy lehetővé tegye az adatokon belüli összetett kapcsolatok könnyű megértését.

Tehát a vizualizáció választ ad azokra a kérdésekre, amelyeket nem tudunk.

***Miért fontos az adatmegjelenítés?***

Az adatmegjelenítést egyre inkább minden sikeres adatközpontú elemzési stratégia alapvető utolsó lépésének tekintik. Amikor az adattudósok egy összetett projekt közepén vannak, szükségük van egy módra az összegyűjtött adatok megértésére. Figyelemmel kell kísérni és módosítani folyamatot annak biztosítása érdekében, hogy megfelelően funkcionáljon. Az adatmegjelenítés megkönnyíti az adatcsoportok mintáinak, előítéleteinek és kiugró értékeinek észlelését is.

A célközönség vezetése arra, hogy az üzleti felismerésekre összpontosítson, hogy felfedezzék a figyelmet igénylő területeket. Korábban észrevétlen kulcsfontosságú tények feltárása az adatforrásokkal kapcsolatban, hogy a döntéshozók segítsenek adatelemzési jelentések elkészítésében. Segít az érdekelt feleknek és a csapat többi tagjának minőségi információkkal szolgálni azáltal, hogy hatalmas mennyiségű adatot könnyen érthető képpé és grafikává alakít.

***Az adatmegjelenítés előnyei***

Figyelembe véve, hogy az adatok milyen hatást gyakorolnak a vállalkozás növekedésére, íme néhány előny

* Segít felismerni a legújabb fejlesztéseket a termék fejlesztése és az üzleti nyereség növelése érdekében.
* Az adatmegjelenítések megkönnyítik a kis és nagy adatok megértését az emberi agy számára, ami jobb elemzéshez vezet.
* Segít megérteni a történetet - Az emberi agy nem képes egyszerre nagy mennyiségű számot vagy szöveget megérteni, sőt csak elképzelni. Szüksége van egy vizuális ábrázolásra, hogy értelmezze őket, és ennek következtében a nyers adatokat kézzelfogható fogalommá alakítsa.

## Ábrázolási formák

***bevezetés***: történet, gépek előtt, hogy ábrázoltunk, mit ábrázoltunk (barlangokra strigulát stb.). Az ember maximum három dimenziót képes értelmezni. Az adatelemzésnél a dimenziót úgy kell felfogni, mint a tér leírására használtat. Szélesség, hosszúság és magasság helyett adatok tulajdonsága mentén ábrázolunk. Egy csillag esetében például: hőmérséklet, tömeg és fényesség. Az adatok természetéből kifolyólag számtalan dimenzió megjelenítés lehetséges.

***különböző tulajdonságok***at ugyanazon a koordinátán való ábrázolása

***2d***

***3d***

***Hogyan tudunk háromnál több dimenziót ábrázolni?***

***Saját adathalmazon ábrázolása***: Milyen megjelenítést használtam, több adathalmaz bemutatása, észrevételek, következtetések stb.

***Az ábrázolás az érme egyik oldala***: ábrázolás + statisztika. Az ábrázolás lehet megtévesztő. Mikor lehet megtévesztő és miért? Átvezetés: Nem szabad hinni mindig a szemünkben. Az ember képes beleesni az optikai illúziók csapdáiban. Ennek érdekében nyújthat segítséget az adatok statisztikai elemzése.

# Adatelemzés lépései és adathalmazok ismertetése