

# Szakedolgozat kivonat

## Toxikus szövegek detektálása

Homolya Panni

2021

Az interneten, különösen a közösségi médiában jelen van a gyűlöletbeszéd és egyéb toxikus tartalom, aminek felismerésével, osztályozásával és kiszűrésével az adattudomány is elkezdett foglalkozni. A szakdolgozat célja a toxikus szövegek detektálása, gépi tanulás módszereket felhasználva.

A dolgozat első felében a kapcsolódó tanulmányok és adathalmaz ismertetését követően, áttekintést adunk az alap adattudományi fogalmakról. Részletesen ismertetjük a felügyelt tanulás, a magyarázó és a magyarázott változó fogalmát, illetve a tanító, a validációs és teszhalmaz jelentését. Továbbá bemutatjuk a logisztikus regressziós modell működését, a gradiens módszert és a modell kiértékelésénél használt mérőszámokat, mint például a precision, recall, F-mérték értékeket.

A dolgozat második felében a szöveges adat előfeldolgozási lépéseivel foglalkozunk, mivel a gépi tanulás modellek szöveges adatokkal nem tudnak dolgozni, így vektorra alakítjuk azokat. A vektorra alakítást attribútumgenerálással hajtjuk végre a tokenizációt és a szótövesítést követően. A tanítás során felépítünk egy alapmodellt, ami 57.02%-os F-mértéket eredményez, aminek az eredményén további kísérletek során javítani próbálunk. A kísérletezés során további három modellt építünk fel, amik közül a gazetteer változót tartalmazó modell sikeresen javít az alapmodell F-mértékén.

Az utolsó részben az alapmodellen elvégzett hibaelemzést mutatjuk be, aminek következtében a három másik modellben szereplő változókat létrehoztuk. Továbbá a tanítás során kapott eredményeinket, illetve a tesztadaton kijövő eredményeket ismertetjük. A dolgozat folyamán az alapmodell sikeres javítása 63.84%-os F-mértéket eredményezett.