



Multimédia tartalmak intelligens feldolgozása - Házi feladat megoldási terv

Távközlési és Médiainformatikai Tanszék

Készítette:	Benda Krisztián
Neptun-kód:	J1CEI3
Ágazat:	Adat- és Médiainformatika
E-mail cím:	krisztianbenda@gmail.com
Konzulens(ek):	Dr. Szűcs Gábor
E-mail címe(ik):	szucs@tmit.bme.hu

Téma címe: Videókon arcok érzelmi elemzése

Feladat

A 2021-as évben a szorgalmi házi feladat **videókon arcok érzelmi elemzése** témakörhöz kapcsolódik. Ehhez az osztályozási feladathoz rendelkezésre áll egy tanulói állomány (felcímkezett érzelem információkkal: boldogság, düh, szomorúság, undor, meglepetés és félelem); a félév során el kell végezni a prediktív becslést a tesztszállományra és ki kell értékelni az eredményt.

A megoldáshoz bármilyen segédsoftver használható; cél a tesztadatsor célváltozójának minél jobb predikciója (az 1. szakaszban 1 beadási lehetőség van, a 2. szakaszban pedig plusz egy beadási lehetőség) és minél több saját kontribúció elérése. A házi feladat elfogadásához szükséges tennivalók és határidejük:

- Február 28-ig el kell készíteni egy dokumentumot (1 - 3 oldal) a feladat rövid értelmezésével és megoldási ötletekkel, tervvel (szucs@tmit.bme.hu címre).
- Április 20-án: A megoldásnak azt a részét, ami már készen van, be kell mutatni szóban (Teamsen) egy kiselőadás keretében. Ha van már egy kezdetleges, egyszerű megoldással előállított predikció a célváltozóhoz, akkor azt kérem elküldeni e-mail-en április 18-ig (1. szakasz).
- Május 7-ig lehetőség van a tesztadatsor célváltozójának még egy predikciójára (2. szakasz). Ezen kívül május 7-ig kell elkészíteni a megoldást bemutató dokumentumot.
- Május 11-én kiselőadást kell tartani a végső megoldásról, valamint meg kell válaszolni a felmerülő kérdéseket.

2018/2019. 1. félév

Feladat rövid értelmezése

A feladat során videók tartalmát kell elemezni mesterséges intelligencia algoritmusokkal. Osztályozni kell a videókban látható vizualitást a megjelenő érzelmek tekintetében. A feladat elvégzéséhez rendelkezésre áll egy tanuló és egy teszt videóállomány, melyek az alábbi linken érhetők el:

https://drive.google.com/drive/folders/1kYYWZiPBjaxX99wI75kDbirgS_YFdioY?usp=sharing

A teszt adatállomány 270 videót tartalmaz. A tanuló adathalmazban 6 érzelem van elkülönítve: düh, undor, félelem, boldogság, szomorúság, meglepetés. Nagyjából 70 videóval tudunk számolni érzelmekként külön-külön.

A tanuló adathalmazon mesterséges intelligencia modell felépítése a feladat, amely képes a videókat feldolgozni és hozzájuk érzelmet társítani. A modell használhatóságát a teszt adathalmazra adott predikcióval tudjuk bemutatni. Ezt a predikciót kell beadni, melynek egyszeri javítására is lehetőség van.

A munka követésére két szóbeli és egy írásbeli beszámoló szolgál a félév során.

Megoldási ötletek

A feladat elvégzéséhez valamilyen mesterséges intelligencia alapú megoldásra gondolhatunk ugyanis tanító adatok alapján kell osztályozni ismeretlen videókat. Megoldási ötletként a konvolúciós neurális hálózatokat tudnám mondani, mint felügyelt gépi tanulási technikát, mert a mély neurális hálók már számtalanszor bizonyították képesztályozási rátermettségüket a tématerületen. De, ha a multimédia tartalomból sikeresen előállítottunk számadatokat a megfelelő struktúrában, akkor más predikciós osztályozási algoritmusok is szóba jöhetnek, mint például döntési fákra alapuló modellek.

Terv

1. Irodalom kutatás a témában, különböző megoldások vizsgálata, kipróbálása
 2. Órán tanult módszerek áttanulmányozása
 3. Videók letöltése, szoftveres beolvasása, kezelése
 4. Videók képekre bontása, felvágása
 5. Képekről jellemzők kinyerése, tulajdonságvektorok megállapítása
 6. Neurális hálózat és/vagy más tanuló megoldás elkészítése
 7. Tanítás, hálózat javítása
 8. Predikció elkészítése
 9. Az elkészült megoldás javítása, továbbfejlesztése
 10. Bemutató elkészítése és előadása
-