

# Gépi tanulás a gyakorlatban - Projektmunka

Az értékelés szempontjai:

- órai munka és aktivitás a szemeszter során
  - o a végső jegyen nem változtat vagy javít
- a beadandó alfeladataiból mennyi teljesül
  - o a jegy túlnyomó részét adja
- a védés során a kérdésekre mennyire ad a hallgató kielégítő választ
  - o amennyiben bármilyen csalás derül ki<sup>1</sup>, vagy a hallgató egyáltalán nem érti az elméletet, a végső jegy elégtelen
  - o ha a hallgató precíz válaszokat ad, pontosan érti a projektjét, a végső jegyen egy osztályzatot javíthat.

A projektmunka értékelése:

A lentebb felsorolt szempontokból kell tetszőleges számút implementálni. A helyesen lefejlesztett alfeladatok összege adják a beadandó jegyét. A pontozás 0-ról indul és 5-ig halad.

A beadandó osztályozása:

0-2,5 pont: elégtelen

2,5-3,5 pont: elégséges

3,5-4 pont: közepes

4-4,5 pont: jó

4,5- pont: jeles

*Ha a pont kétesre áll, akkor a védés és az órai munka eredménye kerekítik fel/le a jegyet.*

Értékelési szempontok:

A beadandóban **neurális hálóval** kell **osztályzó** modellt készíteni. Amennyiben a modell más algoritmus, a jegy automatikusan elégtelen. A neurális háló lehet alap, de választható speciális architektúra is (például konvolúciós) is.

A neurális háléhoz használj Tensorflow-t vagy PyTorch-ot. A fejlesztést notebook környezetben készítsd el.

Az adathalmazt [Kaggle](#)-ről vagy [HuggingFace](#)-ről szerezd be.

---

<sup>1</sup> A projektmunka során nagy nyelvi modellek használhatók. A hallgatónak ettől függetlenül tisztában kell lennie a kód működésével és az elmélettel.

Szempont	Pontszám	Megjegyzés
A tanításhoz legalább 6 feature van használva.	0,5	Nem számít bele az output (y), valamint sorazonosító ID-k.
A projektben a hiányzó/kiugró értékek statisztikai módszerekkel szűrve vannak.	0,5	
A projektben a hiányzó/kiugró értékek valamilyen speciális módszerrel vannak kiszűrve.	0,75	Pl. klaszterezés.
A beadandó tartalmaz beágyazott, kategorikus változót.	0,25	
Az adathalmazhoz a hallgató hozzáad valamilyen új, származtatott feature-t, amely releváns a feladatmegoldáshoz.	0,25	Nem számít bele a kategorikus változók átalakítása.
A projekt tartalmaz normalizálást/standardizálást	0,15	
Az előfeldolgozás tartalmaz valamilyen vizualizációt	0,25	Pl.: eloszlásvizsgálat, outlierok felismerése stb.
A modell valamilyen speciális neurális háló architektúrát használ.	0,75	Pl. konvolúciós neurális háló
A neurális háló legalább három rejtett réteggel dolgozik.	0,25	Csak olyan rejtett rétegek számítanak, ahol a neuronok számát meg kell adni. <sup>2</sup>
A neurális háló legalább öt rejtett réteggel dolgozik.	0,5	Csak olyan rejtett rétegek számítanak, ahol a neuronok számát meg kell adni. <sup>3</sup>
A modell tartalmaz speciális rétegeket és ezeket a hallgató indokolni tudja.	0,25	Pl. regularizáció
A tanítás tartalmaz Early Stoppingot.	0,25	
A neurális háló 75% feletti pontossággal osztályoz validációs és teszt adathalmazon	0,25	
A neurális háló 90% feletti pontossággal osztályoz validációs és teszt adathalmazon	0,25	
A modell teljesítményét a hallgató nem csak pontossággal méri.	0,25	
A neurális háló teljesítményének megjelenítése valamilyen vizuális alapon	0,5	
A kód letisztult, átlátható	0,2	

<sup>2</sup> Például `model.add(Dropout(0.2))` nem számít bele.

<sup>3</sup> Itt sem számít bele például Dropout, aktivációs függvények.

A projekt Githubon elérhető, tartalmaz dokumentációt.	0,25	
---	------	--

**Megjegyzés:** A szempontok nagyobb összpontszámot adnak, mint 5. Nem szükséges az összes alfeladatot lefejleszteni.