

Név: _____

Magyar MI Diákolimpia
Országos Válogató
1. Forduló

Versenyzői Azonosító: _____

A feladatsor megoldására pontosan **60 perc** áll rendelkezésedre. A forduló **28 kérdésből áll**, a teljes feladatsor maximálisan **200 pontot ér**.

Ez a feladatsor a **Magyar Mesterséges Intelligencia Diákolimpia** első fordulója, amelyet papíralapon, segédeszközök használata nélkül kell teljesíteni.

A kérdések négy fő témakörben vizsgálják a tudásotokat: **gépi tanulás** (Machine Learning, ML), **számítógépes látás** (Computer Vision, CV), **természetes nyelvfeldolgozás** (Natural Language Processing, NLP), valamint **megerősítéssel tanulás** (Reinforcement Learning, RL).

A diákolimpia során összesen **600 pont** szerezhető. Az első forduló eredménye beleszámít a végső pontszámba.

Sok sikert kívánunk!

Feleletválasztós kérdések (5 pont kérdésenként)

1. Tanítás és kiértékelés során hány független adathalmazra van szükség L1 regularizáció és korai megállás együttes alkalmazása esetén?
☐ Egy
☐ Kettő
☐ Három
☐ Négy
☐ Attól függ, milyen loss függvényt használunk
2. Az alábbiak közül melyik **nem** biológiai ihletésű?
☐ Hebb-i tanulás
☐ ADAM
☐ Önszervező térkép
☐ Konvolúciós háló
☐ Spike-time dependent plasticity
3. Az alábbiak közül válaszd ki a felügyelt tanulási algoritmusokat!
☐ K-means
☐ K-legközelebbi szomszéd (K-NN)
☐ Önszerveződő térkép
☐ Perceptron tanulás
☐ PCA
4. Az alábbi módszerek közül melyek használják **csak az elsőrendű deriváltat**?
☐ SGD Nesterov momentummal
☐ ADAM
☐ Newton-módszer
☐ Koordináta csökkentés módszere
☐ L-BFGS
5. Vizsgáld meg az alábbi két állítást:
 - A sigmoid igazából egy transzformált tangens hiperbolikus.
 - A softplus általánosságban nem teljesít rosszabbul a ReLU-nál.

- ☐ Mindkettő igaz
 - ☐ Az első igaz, a második hamis
 - ☐ Az első hamis, a második igaz
 - ☐ Mindkettő hamis
 - ☐ A második függ a konkrét feladattól
6. Mi a helyes teendő, ha a tanítás során NaN értékű gradiens keletkezik?
- ☐ Semmi, ez nem befolyásolja a tanítást
 - ☐ Frissítés után véletlenszerű kis lépés megtétele
 - ☐ Az aktuális lépést hagyjuk ki
 - ☐ Hagyjuk ki a lépést, helyette tegyünk egy kis véletlenszerű lépést
 - ☐ A teljes tanítást újra kell indítani
7. Melyik módszerhez kapcsolódik a Pointer Network?
- ☐ Seq2Seq architektúra
 - ☐ Bahdanau figyelem
 - ☐ Nesterov momentum
 - ☐ Másoló mechanizmus
 - ☐ Transformers
8. Melyik állítás **igaz** a dropout technikáról?
- ☐ Minden lépésben ugyanazokat a súlyokat nullázza
 - ☐ A háló súlyainak egy részét véglegesen nullára állítja
 - ☐ A kimenet bizonyos részeit véglegesen nullára állítja
 - ☐ A háló egyes részeit ideiglenesen kikapcsolja tanítás közben
 - ☐ A tanulási rátát csökkenti lokálisan
9. Mely állítások igazak az adathalmazokra vonatkozóan?
- ☐ A validációs és teszt adathalmaz mindig azonos
 - ☐ A validációs adathalmazt a tanítás során használjuk
 - ☐ A teszt adathalmaz csak a végső értékelésre szolgál
 - ☐ A teszt és tanítóhalmaz méreteinek azonosnak kell lennie
 - ☐ A validációs halmazra nincs mindig szükség
10. Az alábbiak közül melyek tartoznak a **hiperparaméterek** közé?
- ☐ Tanulási ráta (learning rate)
 - ☐ Batch mérete
 - ☐ Modell pontossága (accuracy)
 - ☐ Súlyok aktuális értékei
 - ☐ Rétegek száma a hálóban
11. Mely állítások igazak a veszteségfüggvényre (loss) mesterséges neurális hálók esetén?
- ☐ Értéke mindig 0 és 1 között van
 - ☐ A tanítás során nőnie kell
 - ☐ Mindig monoton csökken
 - ☐ Differenciálhatónak kell lennie a tanításhoz (pl. SGD esetén)
 - ☐ A logloss jobb, mint az MSE

12. Mely állítások igazak a súlyinicializálásra?
- ☐ A súlyokat mindig nullára kell állítani
 - ☐ Általában a súlyokat véletlenszerűen inicializáljuk
 - ☐ A kezdőérték nem számít
 - ☐ A súlyokat újratanulásból is származtathatjuk (pl. transfer learning)
 - ☐ A bias értékek mindig 1-re vannak állítva
13. Mely állítások igazak a K-means klaszterező algoritmusra?
- ☐ A klaszterek száma (k) a módszer paramétere
 - ☐ Minden pont a hozzá legközelebbi középponthoz tartozik
 - ☐ A veszteségfüggvény erősen befolyásolja a konvergenciát
 - ☐ Nem igényel inicializációt
 - ☐ Csak szférikus klasztereket tud kezelni
14. Mely állítások igazak a normalizálásra és skálázásra?
- ☐ Ha nem alkalmazzuk, a nagy értékű változók dominálhatnak
 - ☐ A normalizálás során az átlagot nullára hozzuk
 - ☐ A normalizálás és skálázás ugyanaz
 - ☐ A skálázás mindig az átlaggal való osztást jelenti
 - ☐ A skálázás gyakran a szórással való osztást is magában foglalja
15. Mely metrikákat használják gyakran bináris osztályozási feladatok értékelésére?
- ☐ F1-score
 - ☐ ROC-AUC
 - ☐ Precision és Recall
 - ☐ Mean Squared Error
 - ☐ BLEU score
16. Mely változók befolyásolják a tanulást megerősítéses tanulásban?
- ☐ Tanulási ráta (α)
 - ☐ Diszkont faktor (γ)
 - ☐ Felfedezési arány (ϵ)
 - ☐ Hibatolerancia
 - ☐ Dropout arány
17. Mely jellemzők írják le a Q-learning algoritmust?
- ☐ Off-policy
 - ☐ Táblázatos értékfrissítést alkalmazhat
 - ☐ Nem igényel környezetmodell ismeretet
 - ☐ Szabály-alapú következtetésen alapul
 - ☐ A jövőbeli jutalmakat nem veszi figyelembe
18. Melyik állítás igaz az Attention mechanizmusra?
- ☐ Képes kontextusfüggő reprezentációkat létrehozni
 - ☐ Egy token hatását a többi tokenhez viszonyítva súlyozza
 - ☐ Kizárólag képek feldolgozására használjuk
 - ☐ A tanulási ráta optimalizálására szolgál

☐ Csak visszacsatolt hálókbán alkalmazható

19. Mely állítások igazak a konvolúciós rétegekre a számítógépes látásban?

- ☐ Képesek helyi mintázatok felismerésére
- ☐ Tömegével csökkentik a tanulandó paraméterek számát
- ☐ Ugyanazt a szűrőt alkalmazzák különböző pozíciókra
- ☐ Minden neuron külön súlyt kap minden képponthoz
- ☐ Csak fekete-fehér képeken használhatók

20. Mely tényezők vezethetnek overfittinghez egy gépi tanulási modellben?

- ☐ Túl bonyolult modell kevés adattal
- ☐ Hiányzó regularizáció
- ☐ Dropout használata
- ☐ Túl hosszú tanítás validáció nélkül
- ☐ Alacsony tanulási ráta

Kifejtős kérdések (20 pont kérdésenként)

1. **Milyen limitációkat von maga után a lineáris és konvolúciós réteg használata szekvenciális bemeneteken?**

Fogalmazd meg, hogy milyen problémák adódhatnak ezeknek a rétegeknek a közvetlen használatából, amikor a bemenet időbeli vagy szekvenciális természetű.

Térj ki arra is, hogyan lehet ezeket a korlátokat áthidalni modern architektúrák vagy tanulási stratégiák segítségével.

2. **Mit nevezünk batch optimalizációnak?**

Ismertesd a fogalmat, magyarázd el, hogyan működik a gyakorlatban, és hasonlítsd össze más optimalizálási stratégiákkal:

- Teljes halmazos (batch)
- Mini-batch
- Sztochasztikus (SGD)

Fejtsd ki, milyen hatással van a batch-méret növelése vagy csökkentése:

- A tanulás stabilitására
- A gradiens becslésére
- A konvergencia sebességére
- Az általánosítási képességre

3. **Ismertesd a korai megállás technikát! Milyen probléma megoldásában lehet ez hasznos?**

Írd le, hogyan működik az early stopping módszer, milyen típusú metrikákat figyelünk, és milyen feltétellel állítjuk le a tanítást.

Mutasd be, milyen jelenség kivédésére szolgál, és milyen körülmények között célszerű alkalmazni.

4. **Bináris osztályozás: FP és FN jelentősége**

Egy bináris osztályozási feladatban a válaszok lehetnek igaz pozitívak (TP), igaz negatívak (TN), hamis pozitívak (FP) vagy hamis negatívak (FN).

Adj példát olyan alkalmazásra, ahol:

- Fontosabb csökkenteni az FP értékét még az FN rovására is
- Fontosabb minimalizálni az FN értéket, még ha több FP keletkezik is

Indokold meg mindkét példát. Milyen típusú alkalmazásokban számítanak ezek a különbségek?

5. Hogyan befolyásolja a diszkont faktor (γ) értéke a tanulást megerősítéses tanulásban?

Írd le, mit jelent a diszkont faktor (γ), és hogyan jelenik meg a visszatekintő (return) vagy értékfüggvény számításában.

Térj ki a következő szempontokra:

- Mit jelent, ha γ közel van 0-hoz? Mit, ha közel van 1-hez?
- Milyen típusú környezeteknél célszerű alacsonyabb vagy magasabb γ -t választani?
- Milyen viselkedési különbségekhez vezethet egy túl kicsi vagy túl nagy diszkont faktor?