# Mesterséges intelligencia és gépi tanulás tanulmányi verseny

# “MI ZH eredmények prediktálása”

A verseny során a 2018-as őszi félév Mesterséges intelligencia tárgy zárthelyijeinek (ZH) az eredményeit kell megjósolni rendre az egyes diákok esetében. Egy szakértői becsléshez hasonlóan ebben felhasználhatóak opcionálisan a diákok általános demográfiai jellemzői, a tanulmányi előzményeik, a házi feladat megoldási stratégiájuk általános jellemzői és konkrétan az egyes témákban elért házi feladat eredményeik, a tantárgy előadásainak látogatási gyakorisága és az ott felmutatott kompetenciaprofil a tantárgyi kvíz alapján. Az szakértői szakértelmet a versenyben természetes nyelvű leírások és statisztikai adatok biztosítják, amelyek kombinációja (fúziója) remélhetőleg az emberi teljesítményt meghaladó versenymegoldásokat eredményez!

**Bemeneti információk**

* A tárgy természetes nyelvű tematikája.
* A diákok 5 darab általános anonim jellemzője, mint például nem, kor és tanulmányi helyzet.
* A 2018-as házi feladatok eredeti természetes nyelvű leírásai.
* Az egyes diákok házi feladat (HF) megoldási profilja: a beadás folyamatának leírói és a szerzett pontszámok.
* A 2018-as órán elhangzott 52 darab kvízkérdés eredeti természetes nyelvű leírása.
* Az egyes diákok kvízpontszámai.
* A 2018-as ZH-k eredeti természetes nyelvű leírásai.
* Az egyes diákok ZH feladatpontszámai, összesített pontszámai és ZH teljesítése.
* Ezek az adatok teljes körűen 438 egyenletes eloszlás szerint kiválasztott diák esetében érhetőek el.
* A verseny 3. órájában válik elérhetővé egy előzőekkel diszjunkt 100 diákot tartalmazó VALIDATION adathalmaz, amely az előbb felsoroltakból nem tartalmazza a ZH eredményeket.

**Az elvárt kimenet**

Az elvárt kimeneti predikciók a ZH eredmények a VALIDATION adathalmazon, nevezetesen a következők:

* a bináris ZHTeljesitve,
* a skalár ZHOsszpontszam,
* a skalár ZHFeladatok pontszámai (iMSc feladatpontszámok predikcióját nem kérjük!).

**Az elvárt kimeneti formátum**

A predikciókat szintaktikailag egy RESULTS\_versenyzőiNEPTUNkód.CSV file-ban várjuk, amely első oszlopban tartalmazza a diákok VirtualisNeptun azonosítóit, majd rendre a prediktált ZHTeljesitve bináris (0/1) értéket, a ZHOsszpontszam értéket és a 2x7 darab ZHFeladat pontszámnak a prediktált értékeit. A file-t a MIT tanszék HF kezelői rendszerén keresztül (<https://hf.mit.bme.hu/> ) kérjük feltölteni “MIGT2019 tanulmányi verseny” címszó alatt 21 óráig. **A 21 óra után feltöltött feladatokat érvénytelennek tekintjük!**

**Teljesítménymetrikák (a RESULTS\*.CSV alapján)**

* Bináris félreosztályozási hiba a ZHTeljesitve esetén.
* Négyzetes hiba a ZHOsszpontszam esetén.
* Négyzetes hibák átlaga a ZHFeladatok esetén.

**Összesített helyezés**

A versenyen beadott megoldásokat a három teljesítménymetrika alapján külön-külön sorrendezzük, döntetlent megengedve, majd az egyes megoldások összesített helyezését az elért három sorrendjének az átlaga határozza meg.

**Ellenőrzés**

A MIT tanszék HF kezelői rendszerén (<https://hf.mit.bme.hu/> ) kérjük még feltölteni a következőket:

* A megoldást adó forráskódot tömörítve egy MIGT\_ versenyzőiNEPTUNkód gz/zip file-ban.
* A megoldás elméleti hátterének összefoglalóját 50-500 karakterben szóközökkel együtt.
* A megoldás lefuttatásához szükséges technikai feljegyzéseket 50-500 karakterben.

**A bemeneti adatok változóinak a részletes leírásai**

A diákok adatait egy vessző elválasztotta (CSV) file tartalmazza, amelynek első sora tartalmazza a változók alfanumerikus karakterekből álló neveit. Ezek közül kiemelendők a következők:

* A VirtualisNeptun egy egyedi azonosító, amelynek a megoldás beadásában van kritikus szerepe, mivel ez köti össze a VALIDATION file-ban szereplő input értékeket a RESULT file-ban szereplő prediktált értékekkel.
* Nem, SzuletesEv, SzuletesHely, TargyFelvetelHanyszor, Tankor: értelemszerűen
* kviz\*: az elért bináris (0/1) kvízpontszámok, ahol a \* azonosító feloldása a kvízkérdések\_2018.docx file-ban található meg.
* HF\*Darab: a feltöltések száma a \* sorszámú HF esetében
* HF\*ElsoBeadas/ HF\*UtolsoBeadas: az első/utolsó feltöltés ideje \*HF-nek
* HF\*Pont: az utolsó beadáshoz tartozó, így végleges pontszáma a \*HF-nek
* HF\*PontSzoras: a \*HF beadásainak szórása
* HF\*UtolsoBeadasDelta, HF\*ElsoBeadasDelta, HF\*BeadasDeltaAtlag, HF\*BeadasDeltaSzoras: a beadási határidőhöz képesti napokban mért értékek a \*HF esetében
* ZH1[1-7]: az első ZH feladataira kapott pontszámok
* ZH1i: első ZH iMSc feladatára kapott pontszám
* ZH1Total: az első ZH-ban elért összpontszám
* ZH2\*: a második ZH esetén analóg módon mint elsőnél
* ZHOsszpontszam: a ZH-kon elért összpontszám
* ZHTeljesitve: a tárgykövetelményekben meghatározott minimum elérése

Jó versenyzést kívánunk!

**ComBine.lab (Computational Biomedicine workgroup)**

**Artificial Intelligence group**

**Department of Measurement and Information Systems**