

Adattudomány és mesterséges intelligencia főspecializáció

Antal Péter, Bolgár Bence,
Gönczy László, Gyires-Tóth Bálint,
Szűcs Gábor, Toka László



MI? Miért? Hol?

“The History of every major Galactic Civilization tends to pass through three distinct and recognizable phases, ... known as the How, Why, and Where phases.

'How can we eat?

'Why do we eat?'

'Where shall we have lunch?'

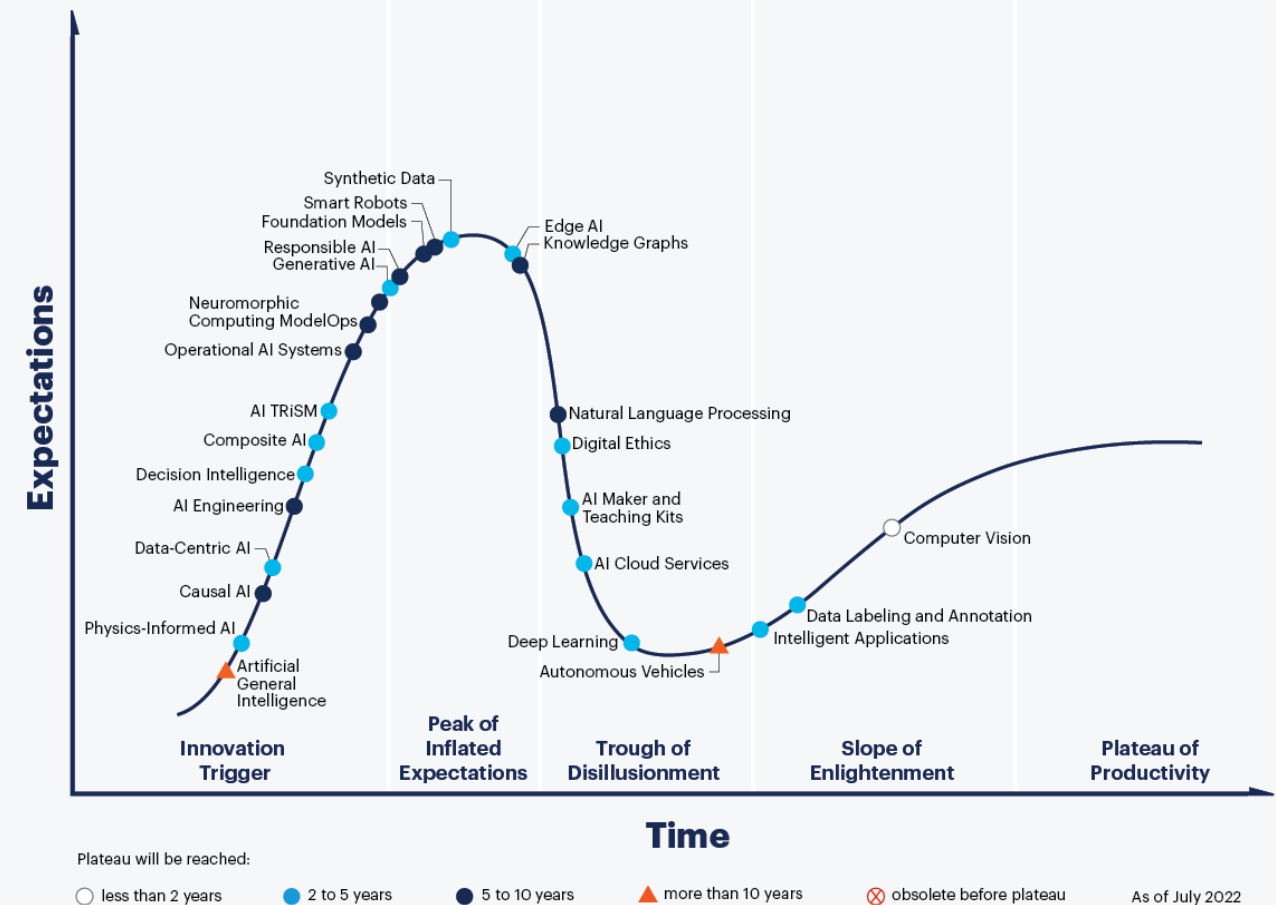
Douglas Adams: The Restaurant at the End of the Universe

-1950: Hogyan definiáljuk az intelligenciát?

1950 – 2010: Miért van szükségünk MI-re?

2010–: Hol lehet megrendelni?

Hype Cycle for Artificial Intelligence, 2022



[gartner.com](https://www.gartner.com)

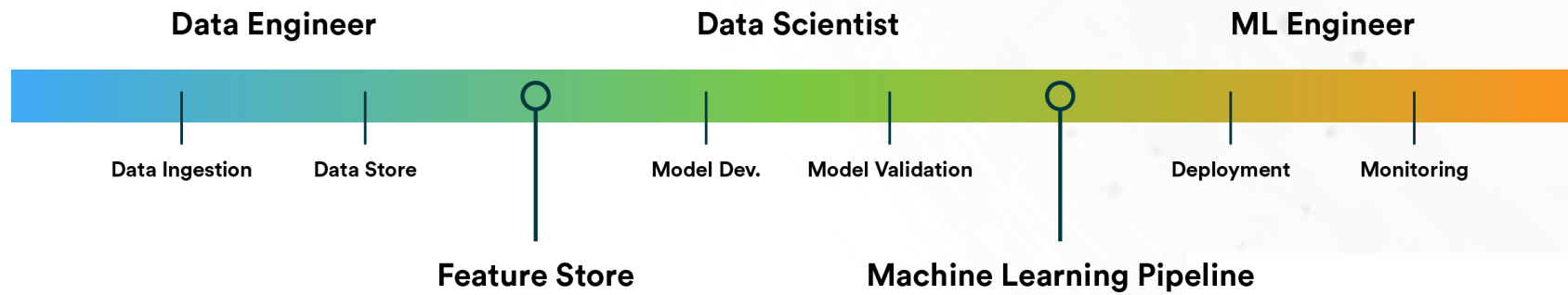
Source: Gartner
© 2022 Gartner, Inc. and/or its affiliates. All rights reserved. Gartner and Hype Cycle are registered trademarks of Gartner, Inc. and its affiliates in the U.S. 1957302

Gartner

Adat, MI, társadalom

Elméleti és gyakorlati MI tudás

a kutatás, fejlesztés, innováció, ipar és társadalom igényei szerint!



Egyéni adatelemzés/MI modellezés

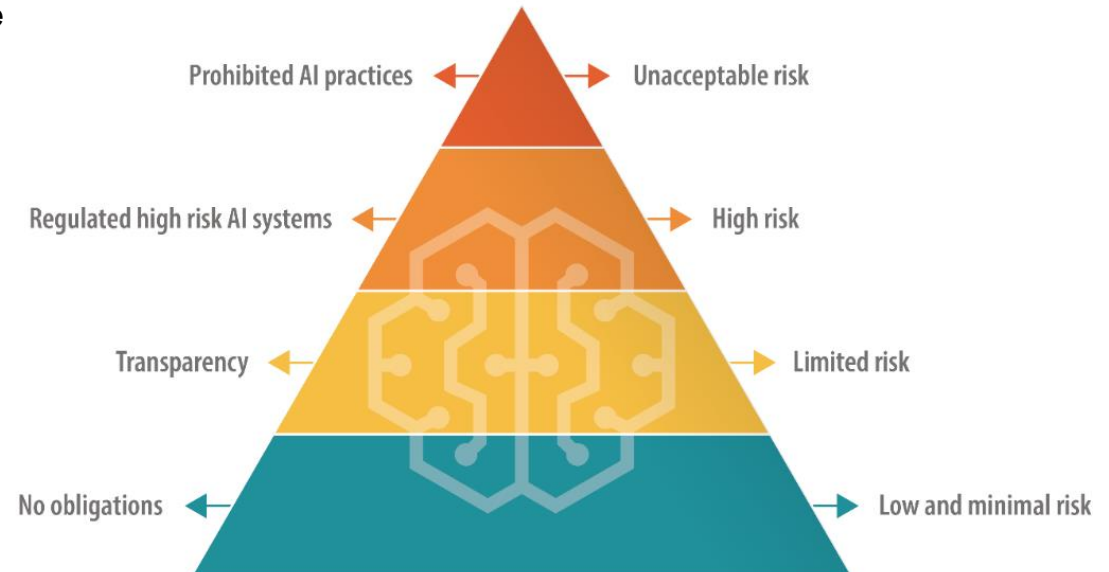
- CRISP-DM,...

Csapatmunka

- MLOps

Jórávaló és megbízható MI

- GDPR, AI Act,...



Data source: [European Commission](https://europeancommission.eu).

A főspecializáció tárgyai, kapcsolódó képzések

ADATTUDOMÁNY ÉS MESTERSÉGES INTELLIGENCIA

Gépi tanulás

Mélytanulás

Intelligens
adatelemzés és
döntéstámogatás

Gépi tanulási
esettanulmányok

Haladó
adatelemzés

Kapcsolódó nemzetközi képzések

Human-Centred AI Masters MSc (HCAIM) opcionális kredit-alapú képzés, 2022-

EIT Digital Data Science MSc, átlapolt képzés 2. évben. 2024-

Kapcsolódó C tárgyak

Adatbiztonság és adatvédelem a gépi tanulásban

Megbízható mesterséges intelligencia és adatelemzés

Kapcsolódó választható tárgyak

Mesterséges intelligencia etikája

Mesterséges intelligencia és a jog

Neurális hálózatok

... és még több!

Erős ipari kapcsolatok

NVIDIA Deep Learning Institute, Continental, Siemens Healthineers, Richter, Magyar Nemzeti Bank, Ericsson

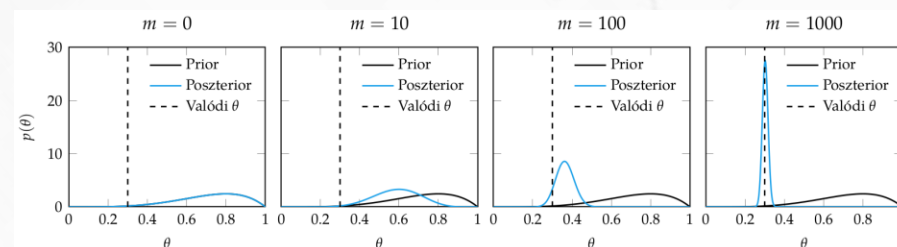
Erős kutatói háttér

Kutatási projektek, publikációk, szabadalmak

Gépi tanulás

- „Üres táblától a kódig”: filozófia, matematika, programozás,
- Átfogó kép a modern gépi tanulásról,
- Hatékony, gyakorlatban is használható implementációk,
- Interaktív kísérletek,
- Tippek, trükkök, sokféle gyakorlati alkalmazás.

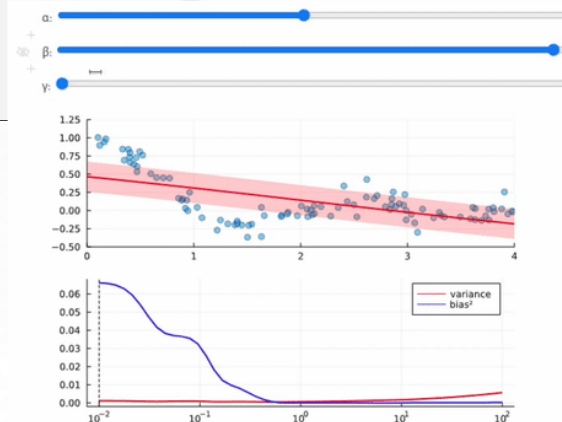
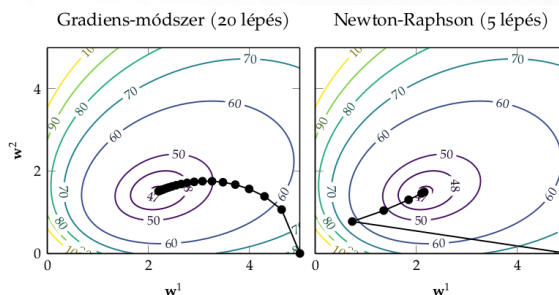
<https://github.com/bolgarbe/gtiegyzet>



$$\underbrace{p(\theta | \mathbf{y})}_{\text{poszterior}} \propto \underbrace{p(\mathbf{y} | \theta)}_{\text{likelihood}} \underbrace{p(\theta)}_{\text{prior}}$$

```
struct LinearRegression
    w # modellsúlyok
    λ # regularizáció

    LinearRegression(D; λ=0) = new(zeros(D), λ)
end
function train!(m::LinearRegression, φ, y)
    m.w .= (φ'φ + m.λ*I) \ (φ'y)
end
function predict(m::LinearRegression, φ_test)
    return φ_test * m.w
end
```



Gépi tanulási esettanulmányok labor

- Közös gondolkodás
- Valós adatok
- Tényleges üzleti probléma megoldása
- Python (DataFrame, Pandas)
- Jupyter Notebook
- Haladó megközelítés - Egy egyszerű folyamatból indulunk ki
- Gazdálkodás az idővel

Mélytanulás

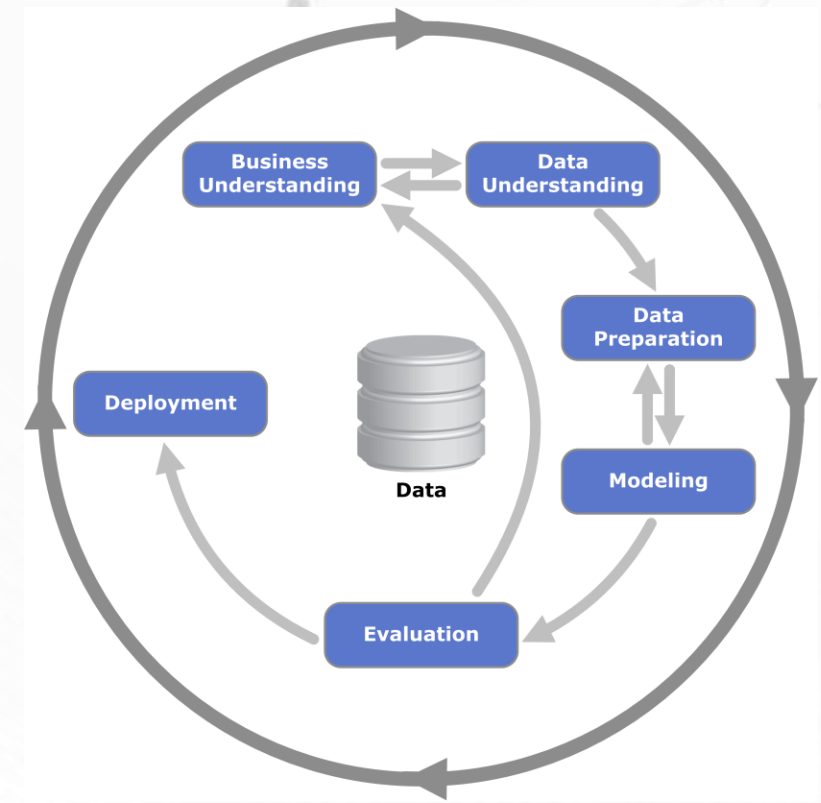
- Deep learning: Software 2.0
- Kezdő szintről középhaladóig
- Gyakorlati tudás megfelelő elméleti alapozással
- Szoftveres eszközök: PyTorch, TensorFlow, TensorBoard, Docker, NVIDIA driverek, Wandb.ai, stb.
- Házi feladatok, projekt alapú tanulás



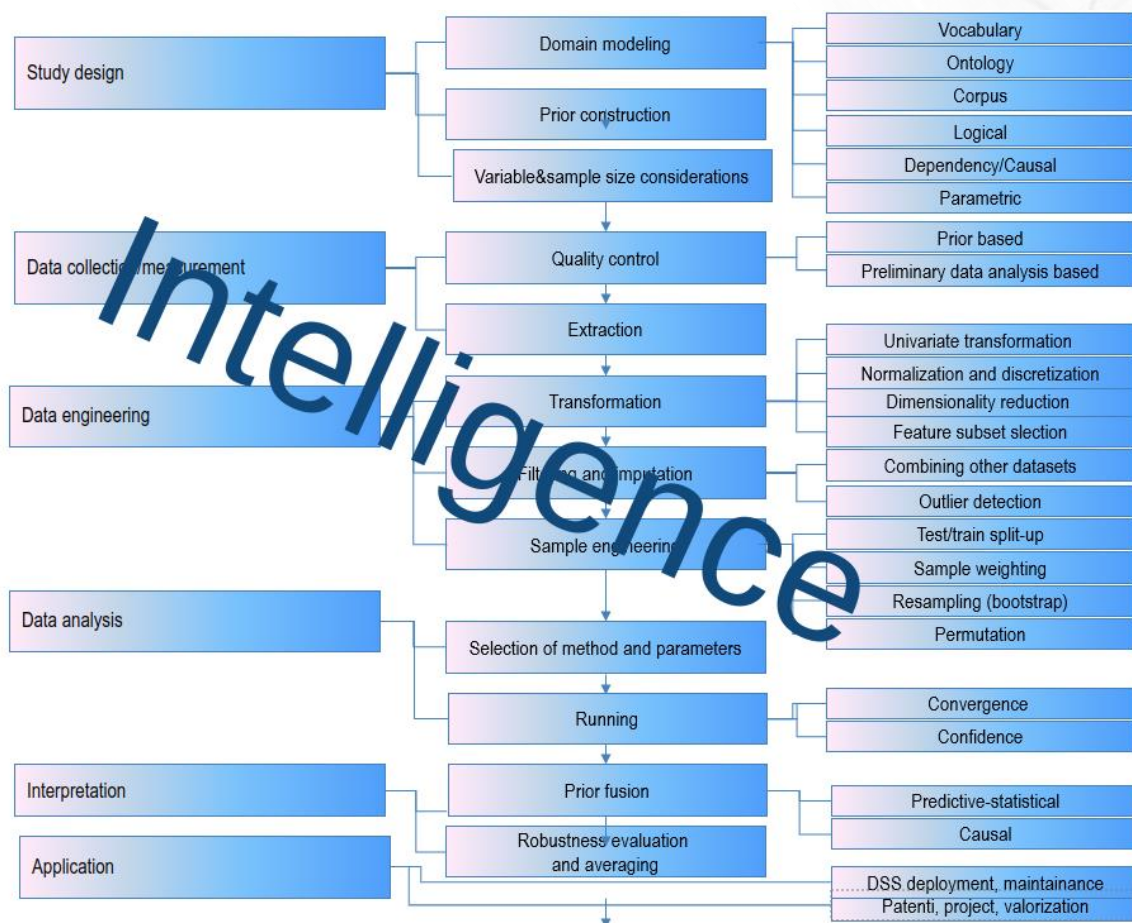
DEEP
LEARNING
INSTITUTE

Haladó adatelemzési módszerek labor

- Egy teljes adatbányászati munkafolyamat gyakorlása, a valós adatszetekben és feladatokban előforduló buktatók megismerésére és leküzdésére
- A féléves adatbányászati feladatot és a hozzá tartozó adatszettet több opció közül lehet választani
- A választott adathoz tartozó mentor kíséri végig a projektmunkát, ami opcionálisan akár kis csapatban is végezhető
- Fokozott hangsúly az adatelőkészítésen (adatbázis és adatformátum kiválasztása, adattisztítás, stb.) ami a gyakorlatban is hosszadalmas fázis
- Gyors és hatékony módszerek az adatvizualizációra és előzetes adatelemzésre
- Gépi tanulási modellek kiválasztása, előállítása, összehasonlítása (hibrid, mélytanulás, stb.)
- A modellek kiértékelése a feladathoz hangolva (metrika-választás, bootstrapping, az eredmények javítása, hiperparaméter-hangolással, boosting alkalmazásával, stb.)
- Az előállított adatbányászati folyamat gyakorlati alkalmazása (felhőbe történő telepítés, etikai kérdések, adatvédelem)



Intelligens adatelemzés és döntéstámogatás



$$p(Model | Data) \propto p(Data | Model)p(Model)$$

$$p(prediction | data) =$$

$$= \sum_i p(pred. | Model_i) p(Model_i | data)$$

$$a^* = \arg \max_i \sum_j U(o_j) p(o_j | a_i)$$

Probability theory

Logical inference of independencies

Inference of (conditional) marginals

Most probable explanation

Maximum a posteriori explanation

Causal inference

Logically inferred causation

Quantitative effect strength inference

Do-calculus

Decision theory

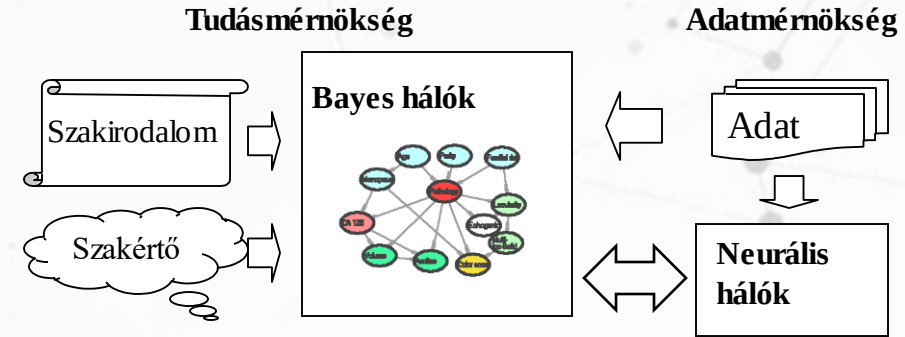
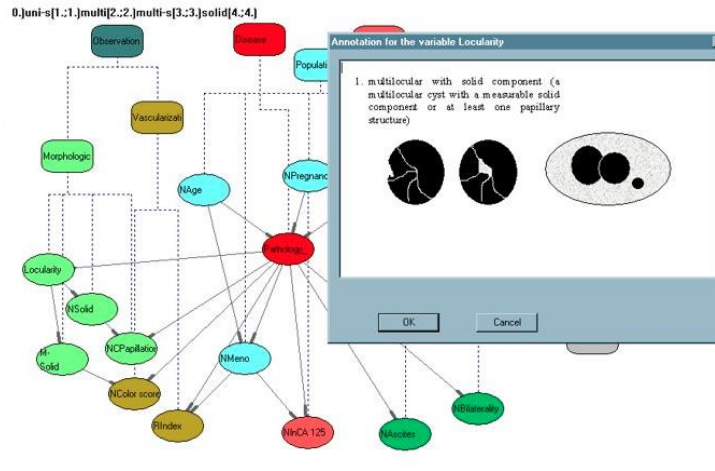
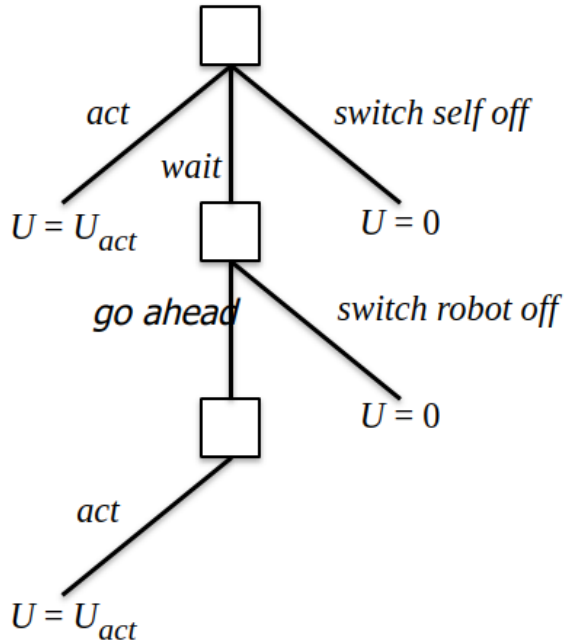
Expected utility/loss

Optimal decision: simple, sequential

Intelligent reasoning

Counterfactual reasoning

Általános intelligencia és adatelemzés



Russell, Stuart: New AI, AI25@BME VIK, 2018

Russell, Stuart. *Human compatible: Artificial intelligence and the problem of control*. Penguin, 2019.

Hadfield-Menell, D., Dragan, A., Abbeel, P. and Russell, S., 2017, March. The off-switch game. In *Workshops at the Thirty-First AAAI Conference on Artificial Intelligence*.

Antal, P., Mészáros, T., De Moor, B., & Dobrowiecki, T. (2001). Annotated Bayesian Networks: a tool to integrate textual and probabilistic medical knowledge. In Proceedings 14th IEEE Symposium on Computer-Based Medical Systems. CBMS 2001 (pp. 177-182).

Antal, Peter, Geert Fannes, Dirk Timmerman, Yves Moreau, and Bart De Moor. "Bayesian applications of belief networks and multilayer perceptrons for ovarian tumor classification with rejection." *Artificial intelligence in medicine* 29, no. 1-2 (2003): 39-60.

Projekt témák

- Bayesi számítások, adat- és tudásfúziós módszerek
- Oksági kutatás, hálózattudomány
- Egészségügyi, kemo- és bioinformatika: gyógyszerkutatás, genetika
- Elosztott mesterséges intelligencia: multiágens rendszerek, federált tanulás
- Megbízható, jóra való és magyarázható mesterséges intelligencia
- Okosotthon, okos viselhető elektronikai eszközök integrálása
- Vezetőtámogató technológiák kutatása és fejlesztése
- Mélytanulás alapú gépi látás, orvosi képfeldolgozás, sportanalitika
- Természetes nyelvfeldolgozás (NLP) a közigazgatásban és pénzügyekben
- Beszédtechnológiák, beszéd-szintézis és beszédfelismerés
- Gráf neurális hálózatok, 3D rácshálók modellezése, tudásgráfok
- Önfelügyelt technológiák
- Adatvédelmi és biztonsági problémák elemzése
- Felhő- és 5G rendszerek MI-alapú menedzsmentje
- ... és még sok más!

Ajánlott mellékspecializáció

Felhasználói élmény - UX és interakció mellékspecializáció
(erős MI fókusszal)

...részletek a terem előtt és
<https://spec.vik.bme.hu/> oldalon.



MESTERSÉGES INTELLIGENCIA ÉS GÉPI TANULÁS TANULMÁNYI VERSENY

HELYSZÍNE: online(!)

IDEJ: November 23. szerda 17:00-21:00

VERSENYRE JELENTKEZÉS: MIGT2022” Teams csoport, csatlakozási kód: ghk67gm

MINTA FELADAT

A verseny egy feladatból áll, amely egy prediktor létrehozását várja el...

KATEGÓRIÁK

Két kategória: (1) akik nem végezték el a Mesterséges intelligencia tantárgyat és (2) akik teljesítették.

HASZNÁLHATÓ SEGÉDESZKÖZÖK

Bármilyen..

(Emberi segítség igénybe vétele a versenyből történő azonnali kizárást vonja maga után.)

Kapcsolat, további információk

Dr. Antal Péter, antal.peter@vik.bme.hu

<https://spec.vik.bme.hu/>

Partnereink



DEEP
LEARNING
INSTITUTE



SIEMENS
Healthineers



ERICSSON



ARTIFICIAL INTELLIGENCE
National Laboratory

ADATTUDOMÁNY ÉS MESTERSÉGES INTELLIGENCIA MSc főspecializáció

*Elméleti és gyakorlati MI tudás az
ipar igényeihez igazítva*

- Erős ipari kapcsolatok ✓
- Erős kutatói háttér ✓
- Nemzetközi képzések ✓

Gépi tanulás
Gépi tanulás esettanulmányok
Mélytanulás
Haladó adatelemzés
Intelligens adatelemzés és
döntéstámogatás

Megbízható MI és adatelemzés
Adatbiztonság és adatvédelem a
gépi tanulásban

MI etikája
MI és a jog