

The background is a futuristic, blue-toned digital environment. It features a central glowing sphere composed of a network of lines, surrounded by various data screens and panels displaying charts, graphs, and abstract patterns. The overall aesthetic is high-tech and data-driven.

**Dr. Vécsey Richárd Ádám**

# **A hálózati infrastruktúra és az MI modellek**

# Tartalom

- Betanító adatok és a rendszerterhelés
- Betanítási módszerek hálózatterhelése
- MI modellek erőforrásigénye
- Optimalizálási stratégiák
- Skálázhatósági szintek
- A jövő lehetséges irányai
- Igényfelmérés

# Rendszerterhelés



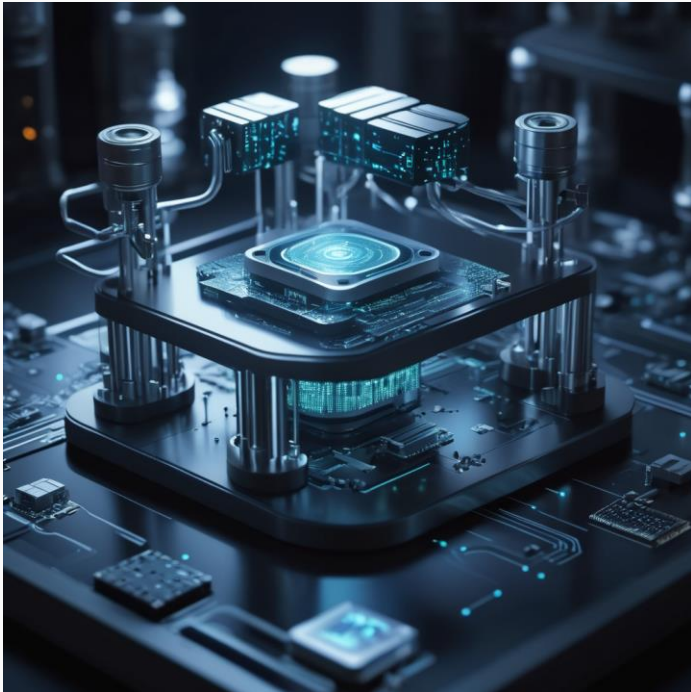
- Tárhely
- Sáv szélesség
- Számítási kapacitás
- Belső adatfolyam

# Rendszerterhelés



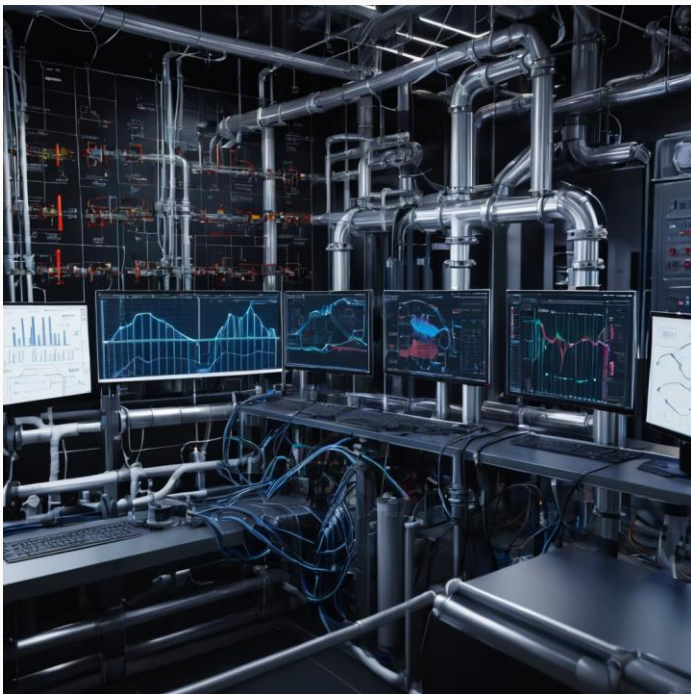
- Tárhely
- Sáv szélesség
- Számítási kapacitás
- Belső adatfolyam
  
- On-the-fly adatfeldolgozás

# Rendszerterhelés



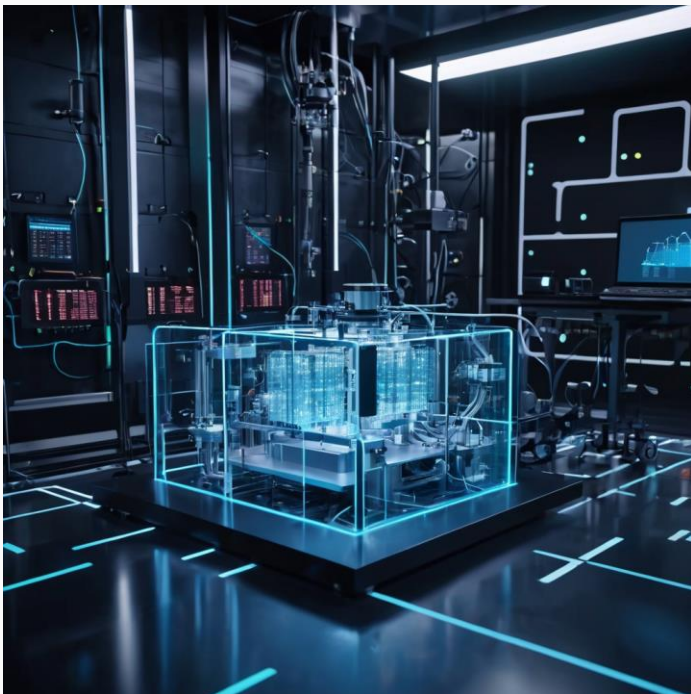
- Tárhely
- Sávszélesség
- Számítási kapacitás
- Belső adatfolyam
  
- Előfeldolgozott adatok

# Rendszerterhelés



- Tárhely
- Sáv szélesség
- Számítási kapacitás
- Belső adatfolyam
  
- Streaming adatok

# Rendszerterhelés



- Tárhely
- Sáv szélesség
- Számítási kapacitás
- Belső adatfolyam
  
- Tömörített adatok

# Rendszerterhelés



- Tárhely
- Sáv szélesség
- Számítási kapacitás
- Belső adatfolyam
  
- Tömbösített adatfeldolgozás



























# Rendszerterhelés



- Tárhely
- Sáv szélesség
- Számítási kapacitás
- Belső adatfolyam
  
- Szintetikus adatok

# Rendszerterhelés

	On-the-fly adatok	Előfeldolgozott adatok	Streaming adatok	Tömörített adatok	Tömbösített adatok	Szintetikus adatok
Tárhely						
Sávszélesség						
Számítási kapacitás						
Belső adatfolyam						

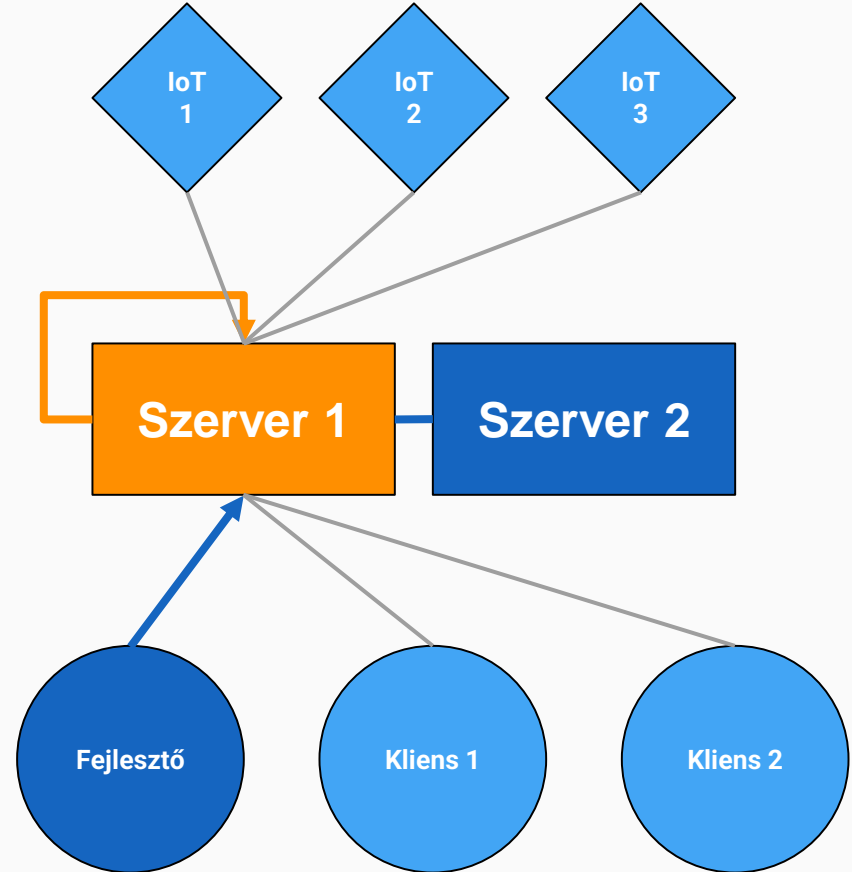
# Hálózatterhelés

- Központosított betanítás
- Elosztott betanítás
- Végponti betanítás
- Egyesített betanítás
- Online betanítás
- Hibrid betanítás



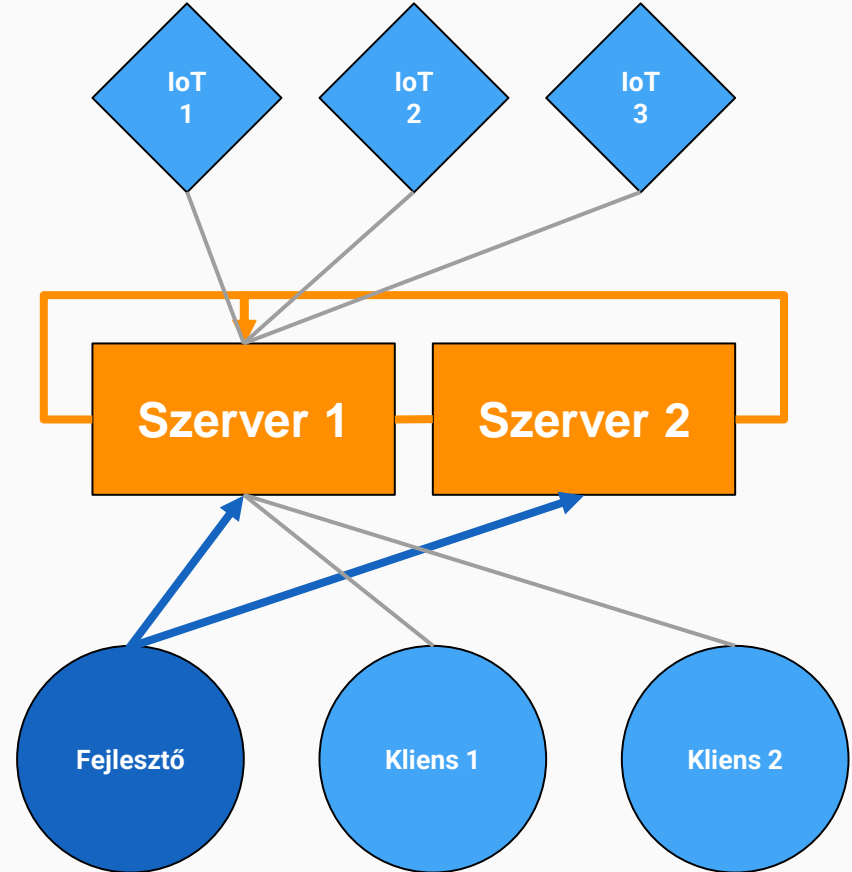
# Hálózatterhelés

- Központosított betanítás
- Elosztott betanítás
- Végponti betanítás
- Egyesített betanítás
- Online betanítás
- Hibrid betanítás



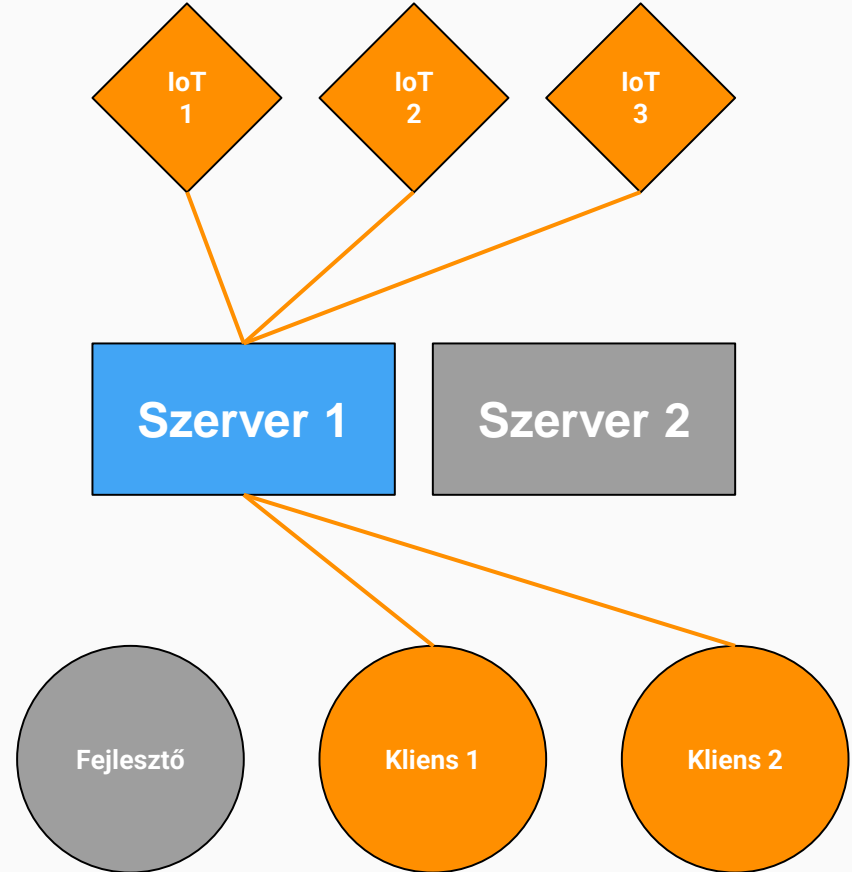
# Hálózatterhelés

- Központosított betanítás
- Elosztott betanítás
- Végponti betanítás
- Egyesített betanítás
- Online betanítás
- Hibrid betanítás



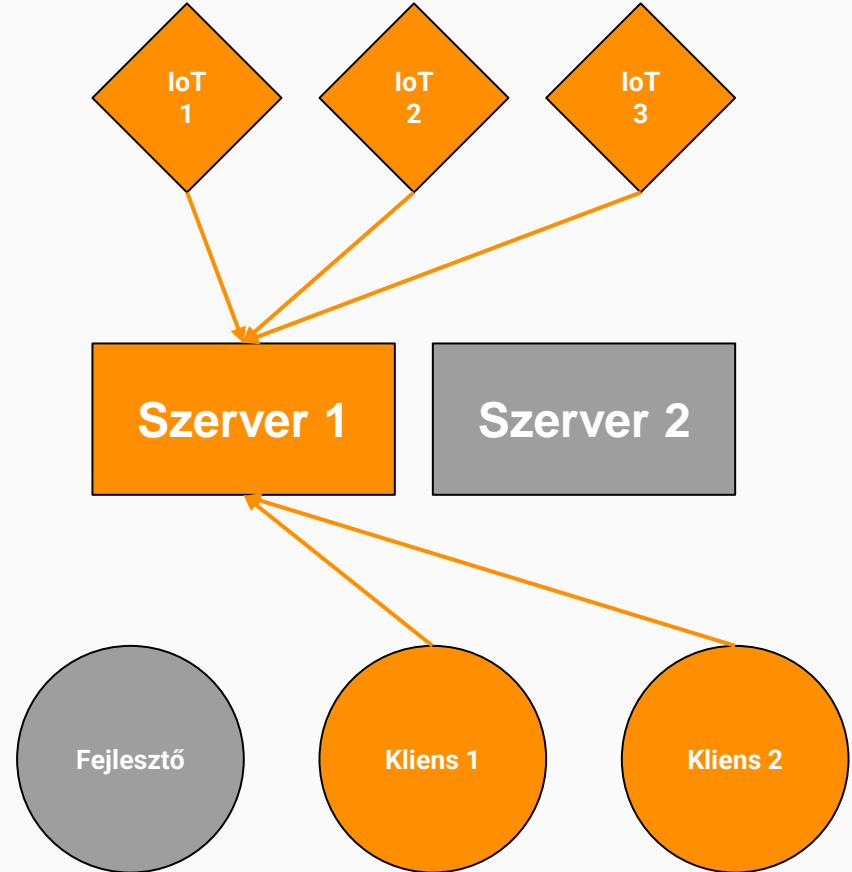
# Hálózatterhelés

- Központosított betanítás
- Elosztott betanítás
- **Végponti betanítás**
- Egyesített betanítás
- Online betanítás
- Hibrid betanítás



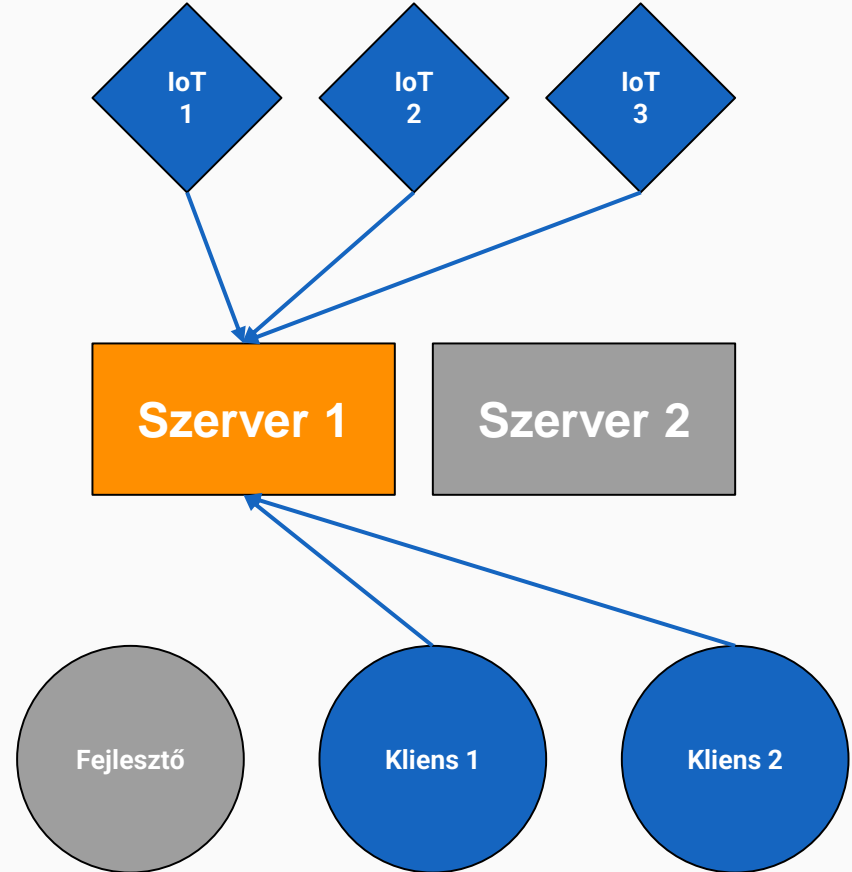
# Hálózatterhelés

- Központosított betanítás
- Elosztott betanítás
- Végponti betanítás
- Egyesített betanítás
- Online betanítás
- Hibrid betanítás



# Hálózatterhelés

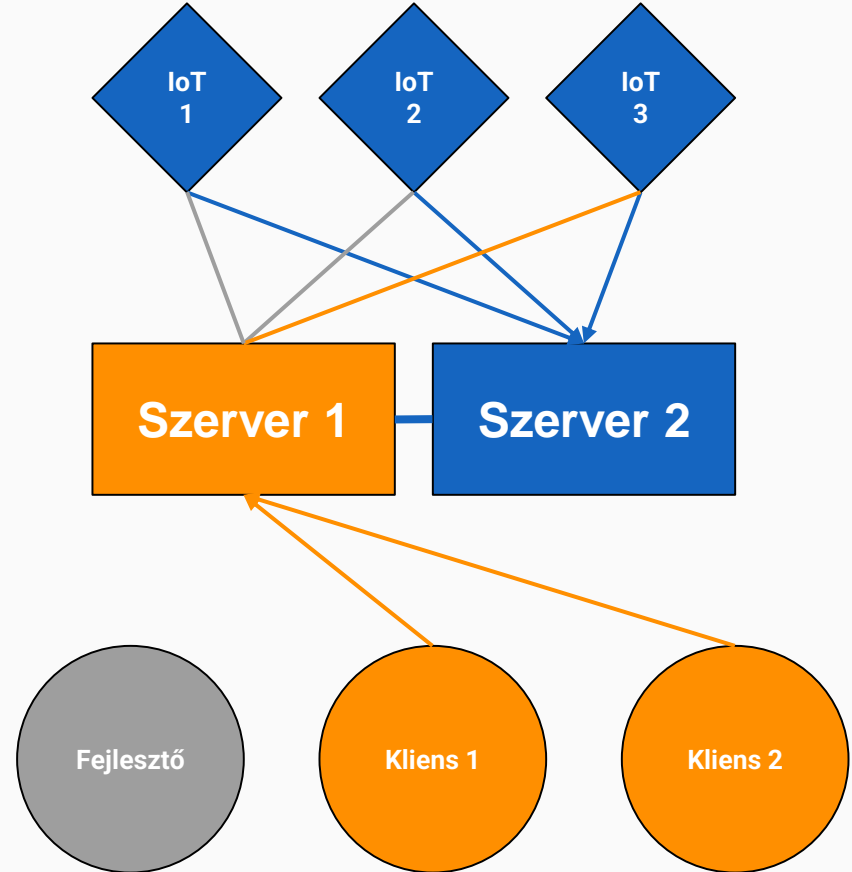
- Központosított betanítás
- Elosztott betanítás
- Végponti betanítás
- Egyesített betanítás
- Online betanítás
- Hibrid betanítás





# Hálózatterhelés

- Központosított betanítás
- Elosztott betanítás
- Végponti betanítás
- Egyesített betanítás
- Online betanítás
- Hibrid betanítás



# Erőforrásigény



- Gépi látás
- Hangfeldolgozás
- Szövegfeldolgozás
- Pénzügyi predikciók
- Játékok MI modelljei
- Ajánlórendszerek

# Erőforrásigény: gépi látás



- **Betanítás**  
150 GB–1 TB (1–10 MB)  
10–200 óra
- **Modell**  
100–600 MB
- **Inferencia**  
1–3 ms/kép  
100 KB–5 MB (224x224 px)

# Erőforrásigény: hangfeldolgozás



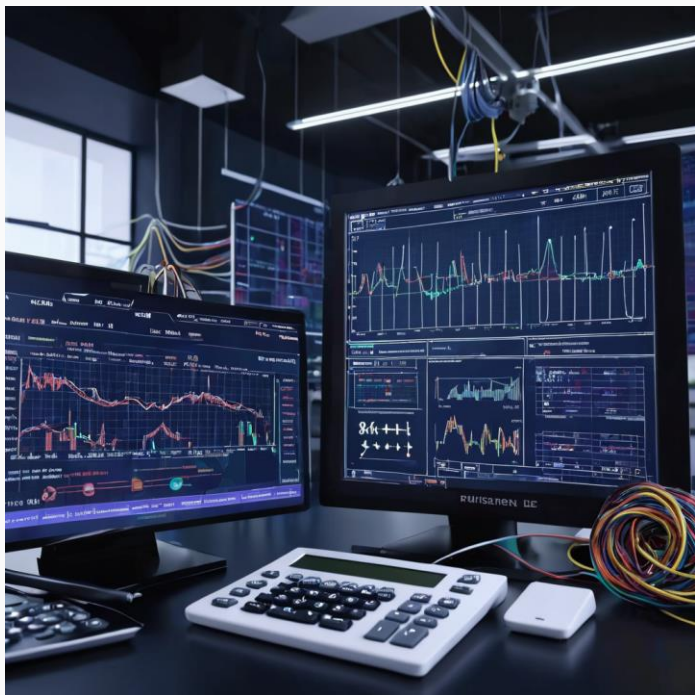
- **Betanítás**  
100 GB–200 GB (50–90 MB/óra)  
72–500 óra
- **Modell**  
50–500 MB
- **Inferencia**  
5–60 sec/perc  
500 KB–2 MB (5-10 sec)

# Erőforrásigény: szövegfeldolgozás



- **Betanítás**  
5 TB–X PB (100 GB)  
100–5000 óra (2 óra)
- **Modell**  
2,5–400 GB
- **Inferencia**  
0,1 ms–2 sec  
100 KB–1 MB

# Erőforrásigény: pénzügyi predikció



- **Betanítás**  
1 GB–300 GB (1–10 MB)  
30–100 óra
- **Modell**  
10–500 MB
- **Inferencia**  
0,2–5 ms  
10 KB–2 MB

# Erőforrásigény: játék



- **Betanítás**  
15 GB–500 GB (50–100 MB)  
70–120 óra
- **Modell**  
100–500 MB
- **Inferencia**  
5 ms–3 sec  
1–5 MB

# Erőforrásigény: ajánlórendszer

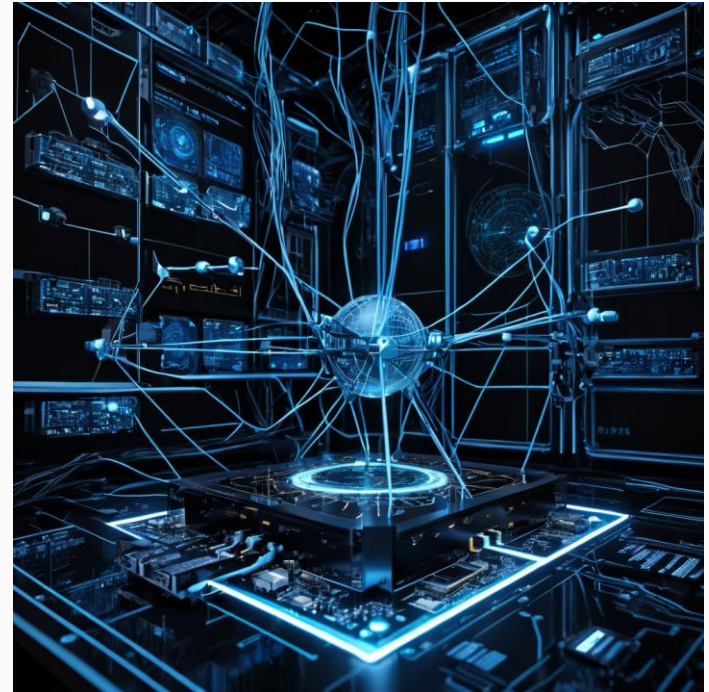


- **Betanítás**  
1 GB–100 TB (1–10 MB)  
10–500 óra
- **Modell**  
20 MB–1 GB
- **Inferencia**  
20ms–2 sec  
700 KB–5 MB



# Optimalizálási lehetőségek

- Gyorsítótárazás
- Késleltetett hozzáférés
- Dinamikus modellek
- Célhardver váltás
- Párhuzamos modellek
- Dinamikus erőforrás allokáció



# Metrikák

- Késleltetés
- Elérhető sávszélesség
- CPU/GPU kihasználtság
- Betöltési/működési idő
- Tárhely I/O
- Elveszett csomagok
- Feldolgozott adatok



# Skálázhatóság

- Modellméret
  - Speciális modellek
  - Nagyobb modellek
- Kérések száma
  - Párhuzamosítás
  - Dinamikus erőforrások
  - Kvantálás



# Jövő



- Egyedi eszközök
- 6G
- Új hálózati protokollok
- Adatvédelmi megoldások
- Dinamikus erőforrás-kezelés
- Hibatűrő algoritmusok

# Igényfelmérés



- Adatmennyiség a fejlesztés során
- Adatátvitel gyakorisága
- Adatbetöltés sebessége
- Betanítás helye
- Modellek ki- és betöltésének gyakorisága

# Igényfelmérés



- Hálózati forgalom időbeli jellege
- Inferencia adatcsomagok mérete
- Inferencia válaszidő igénye
- Speciális célhardver igénye
- Adatvédelmi igények



**Dr. Vécsey Richárd Ádám**