

ELTE Fizika Doktori Iskola I. féléves beszámoló

Fiam Regina Nóra FPYM01

December 2025

1. Bevezetés

A jelenlegi MI-modellek – mint a transzformerek vagy a konvolúciós neurális hálózatok (CNN) – sűrű matematikai műveletekre épülnek, ahol az adatok minden neuronon átáramlanak, ami rendkívüli számítási kapacitást igényel. Ezzel szemben az agy ihlette architektúrák, például a tüskékódolású neurális hálózatok (SNN), az idegrendszer eseményvezérelt működését másolják: csak akkor „tűzelnek”, ha az információ elér egy bizonyos küszöbértéket. Ez ritkább és energiahatékonyabb adatfeldolgozást tesz lehetővé. Míg a klasszikus modellek statikus súlyokkal dolgoznak, az agyszerű struktúrák gyakran alkalmaznak plaszticitást, ahol a kapcsolatok az időbeli dinamika alapján, lokálisan alakulnak. Ez képessé teszi a modelleket a valódi időbeliség kezelésére és a folyamatos tanulásra anélkül, hogy a hosszú kontextusokban elveszne a korábban megszerzett információ.

Kutatásaim két fő területre összpontosítanak: a biológiai jellegű problémákra és a rádiós telekommunikációra. Közös pontjuk olyan gépi tanulási architektúrák fejlesztése, amelyek valamilyen módon az agyi információfeldolgozás mintáit követik. A telekommunikációban az időbeli adatok hatalmas mennyisége miatt kulcsfontosságú az agyszerű modellek energiahatékonyasága, hiszen a hálózati eszközök száma és az adatforgalom exponenciálisan növekszik.

A biológiai adatoknál szintén látható, hogy a transzformerek figyelem-mechanizmusa – mivel minden adatpontot egyszerre vet össze – óriási memóriagigánt generál. Hatékonyabb megoldást jelentenek az állapotter-modellek (SSM), amelyek az emberi munkamemóriához hasonlóan egy belső „állapotot” frissítenek ahelyett, hogy minden lépésben hatalmas mátrixokat számolnának újra. Az SSM lineáris skálázódás mellett képes kezelni akár végtelenül hosszú adatsorokat is. A módszer előnye, hogy a rögzített matematikai rétegek helyett differenciálegyenletekkel leírható, folytonos rendszert használ, amely természetesebben kezeli az időbeli összefüggéseket.

2. Az aktuális félévben elvégzett kutatások ismertetése

Ebben a félévben génexpressziós adatok perturbációinak elemzésével foglalkoztam, és egy olyan modellt fejlesztettem, amely képes megjósolni a perturbációk sejtállapotra gyakorolt hatásait. A munkához az ARC Institute által meghirdetett *Virtual Cell Challenge* biztosította az adatokat: több mint 150 gén perturbációját vizsgáltuk összesen 220 ezer sejtben. A kutatás során bebizonyosodott, hogy a transzformer modellek nehezen alkalmazhatóak ezen a területen az attention-mechanizmus számítási igénye miatt. Ezért egy neurális állapotter-modellen (SSM) alapuló architektúra fejlesztésébe kezdtem.

Mivel valódi idősoros adataink nincsenek – csak a perturbáció előtti és utáni állapotot ismerjük –, a tanítás során a két állapot közötti átmenetet próbáltam rekonstruálni minden sejtben, így tanítva meg a rendszernek a sejtbiológiai dinamikát. A következő félévben a modell finomítását egy hasonló, de elhízással kapcsolatos génperturbációs kihívás adatain tervezem folytatni.

A KDP program keretében, az Ericssonnal partnerségben, csatornabecslési problémák megoldására is elkezdtem SSM-alapú modelleket fejleszteni. A rádiófrekvenciás környezet gyors változásai miatt a pontos csatornaállapot-információ kinyerése jelentős számítási kapacitást igényel, különösen a MIMO technológiáknál. Az SSM architektúra itt azért jelent komoly előrelépést a klasszikus RNN-ekhez képest, mert kiküszöböli a gradiens-eltűnés vagy -felrobbanás problémáját, így sokkal hosszabb időtávú összefüggéseket képes stabilan megtanulni.

Míg az RNN-ek számítási igénye a sorozathosszal lineárisan nő és nehezen párhuzamosíthatóak, a modern SSM-ek (mint az S4 vagy a Mamba) konvolúciós formában is reprezentálhatóak, ami nagyságrendekkel gyorsabb tanítást tesz lehetővé. A csatornabecslés során ez azt jelenti, hogy a modell precízebben követi a fading jelenségeket, ami jobb hibaarányt (BER) és alacsonyabb átlagos négyzetes hibát (MSE) eredményez. Az adatok szimulációjához az NVIDIA Sionna csomagját használtam, a számításokat pedig az Ericsson GPU-in végeztem el.

3. Tanulmányi tevékenység az aktuális félévben

- A gépi tanulás új eredményei szeminárium (FIZ/3/092) – 6 kredit
- Adatbányászat és gépi tanulás (FIZ/3/084) – 6 kredit
- Irányított kutatómunka (FIZ/S1-S4/K18) – 18 kredit

4. Konferenciák az aktuális félévben

- **3MT konferencia:** A KDP-s kutatásunkat ismertettem egy 3 perces előadásban a II. KDP konferencia keretein belül.
- **MABIT Bioinformatikai konferencia:** Előadás tartása
- **NTU-HUN-REN Machine Learning Workshop:** Részvétel

5. Szakmai közéleti tevékenység

- Doktorandusz társaimmal Dudás Bence ötlete alapján elindítottuk a „Kaggle Club”-ot machine learning versenyek közös megoldására. Elsőként a *Virtual Cell Challenge*-en dolgoztunk, ahol gének perturbációinak sejtállapotokra gyakorolt hatásait akartuk megjósolni. Ennek folytatásaként a Broad Institute elhízáskutatással kapcsolatos kihívására fókuszálunk a következő félévben, ahol az RNA-Seq adatok elemzése során szerzett tapasztalatainkat hasznosítjuk.
- Az ELTE KDP doktoranduszként részt vettem az Ericsson Kutatók Éjszakája programján, ahol tudománynépszerűsítő tevékenységet végeztem.

6. Elismerések

Kooperatív Doktori Program (KDP) ösztöndíj az Ericssonnal partnerségben.