

PhD – 4. félévi beszámoló

Kozmológiai nagyskálás szerkezet vizsgálata szimulációkkal és gépi tanulással támogatott adatfeldolgozással

Pál Balázs*

ELTE, Komplex Rendszerek Fizikája Tanszék,
Csillagászat és Űrfizika doktori program

Témavezető: Dr. Csabai István

May 29, 2023

1. Az aktuális félévben elvégzett kutatások ismertetése

A tavaszi félév során – már Magyarországról – folytattam az előző félévben elkezdett kutatásomon történő munkát. Ebben a Subaru teleszkóp Prime Focus Spectrograph (PFS) műszeréhez készítettem egy olyan szoftveres, AI-alapú pipeline-t, mely a jövőbeli mérések során képes lesz a mért csillagspektrumok elemzésére. Az elsődleges cél az olyan autoencoder-alapú architektúrák fejlesztése és tesztelése volt, amik képesek egy zajos csillagspektrumról levonni a zajt, a spektrumban található információk elvesztése nélkül.

Az itthon töltött elmúlt 3 hónap során elvégeztem az eddigi modellek pontosságának tesztelését a teljes paraméterterre. Megvizsgáltam azok stabilitását és rekonstrukciós, valamint zajszűrő képességét a különböző fizikai és megfigyelési paraméterek függvényében.

A projekttel ezután tovább haladtam és elvégeztem a konvolúciós rétegeket tartalmazó, különféle autoencoder architektúrák vizsgálatát. Ezekről a modellektől azt várjuk, hogy invariánsan viselkednek a csillagspektrumok, vöröseltolódásból fakadó translációjára. A konvolúciós autoencoderrel kapott eredményeket összehasonlítottam a korábban vizsgált, teljesen összekötött rétegeket tartalmazó modellekkel kapottakkal, de eddig csupán kisebb pontosságbeli javulásokat tudtam velük elérni.

A témáról készített publikáció írása már folyamatban van, aminek első változatával a nyári hónapokban szeretnék elkészülni.

2. Elvégzett tárgyak

A félév során, az ELTE Fizika Doktori Iskola két tárgyát végeztem el, mindkettőt angol nyelven. Ezek az alábbiak voltak:

- **Lineáris- és nemlineáris MHD hullámok (FIZ/5/054):** A tárgy során J. P. Hans Goedbloed és Stefaan Poedts: *Principles of Magnetohydrodynamics* című könyvét dolgoztuk fel. A tárgy célja a magnetohidrodinamikába (MHD) történő betekintés volt,

*pal.balazs@ttk.elte.hu

mely során megismerkedtünk az MHD folyamatok alapvető fizikájával és matematikai leírásával, valamint a hozzá kapcsolódó fontos fizikai jelenségekkel. Tárgyaltunk az MHD hullámok leírásáról és tulajdonságairól, azok stabilitásáról és az őket befolyásoló tényezők (elsősorban gravitáció) hatásairól. A félév során minden óra előtt elolvastuk a könyv egy-egy fejezetét, majd azt az órán részletesen átbeszéltük egy feleltetéshez hasonló formában. Az órák alatt be kellett bizonyítanunk, hogy megértettük és átlátjuk az adott fejezetben olvasottakat, és ezt mind az MHD-n, mind általánosságban a fizikán belül, megfelelően el tudjuk helyezni. A félév végén az egyes órákon nyújtott szóbeli teljesítményünk és extra szorgalmi munkáink (pl. a könyvben szereplő bizonyos ábrák reprodukálása) alapján kaptunk megajánlott jegyet. Órai teljesítményem és a három MHD hullám Friedrichs-diagramjainak részletes elkészítése alapján a tárgyat jeles osztályzattal zártam.

- **Rádiócsillagászat II. (FIZ/5/010):** Ezt a kurzust az előző féléves Rádiócsillagászat I. (FIZ/5/009) tárgy folytatásaként vettem fel. Míg az előző félévben a rádiócsillagászat elméleti alapjaira koncentráltunk, a mostani félév során a tényleges, rádiócsillagászati megfigyelésekkel kapcsolatos tudnivalókkal ismerkedtünk meg. A félév végén egy 15 perces prezentációval kellett egy szabadon választott rádiócsillagászati cikkből, vagy témából felkészülni a jegyszerzéshez. Ehelyett én a kozmikus mikrohullámú háttérsugárzás témáját választottam, amiből a kurzus egy extra tanóráját tartottam meg a félév során. A tárgyat végül jeles osztályzattal zártam.

3. Oktatott tárgyak

- **Korszerű számítástechnikai módszerek a fizikában 1. (korszam1f19va):** A félév során a kurzus egyik óraadójaként vettem részt, Dr. Papp Gábor, Szigeti Balázs, Dudás Bence és Dr. Bíró Gábor mellett. Előadásaimban a git és GitHub felépítését és használatát, valamint az adattudomány koncepcionális és gyakorlati alapjait ismerttettem. Az év során a hallgatóknak több beadandót, majd egy nagyobb projektet kellett elkészíteniük, melyekben a kurzus folyamán tanultakat kellett alkalmazniuk.
- **Tudományos modellezés számítógépes laboratórium (dsscimodf20lm):** Ezen az MSc-seknek meghirdetett kurzuson minden hallgató 1-1 projektet választ a félév elején, melyek mindegyike valamilyen tudományos szimuláció elvégzését tűzi ki célul. A hallgatóknak a félév során a projektjeiken történő haladásáról kéthetente kell egy prezentáció és egy jegyzőkönyv formájában beszámolniuk. Minden hallgatóhoz egy-egy oktató van rendelve, akik meghatározzák a projektjük haladási irányát, valamint segítik a hallgatót a felmerülő problémák megoldásában. Ebben a félévben egy gravitációs N-test szimulációkkal kapcsolatos projekt témavezetésében vettem részt, aminek tematikáját az eddigi félévekhez képest teljesen átdolgoztam és modernizáltam. A témát választó hallgatónak a GADGET-4 és GIZMO szimulációs szoftverekkel kellett megismerkednie, különféle példaszimulációkat készítenie, majd a GIZMO segítségével egy kozmológiai N-test szimulációt lefuttatnia, vizualizálnia és kiértékelnie.

4. Előadások és konferencia-részvételek

- 2023. április 26-án a Budapesti Gazdasági Egyetem (BGE) meghívására az adattudományról tartottam előadást a téma iránt érdeklődő hallgatók és oktatók számára. Az előadás során a hallgatók megismerkedhettek az adattudomány fogalmával, a tudományban ma már exponenciális ütemben egyre fontosabbá váló *data-driven* szemléletmóddal,

a gépi tanulással és neurális hálózatokkal, valamint mindezek gyakorlati felhasználási módjaival a tudomány és ipar területein.

- 2023. május 15-16-án a Wigner FK által szervezett XIII. GPU Day(s) konferencián vettem részt. Itt számtalan hazai és külföldi kutató között tarthattam meg én is előadásomat a Johns Hopkins Egyetemen végzett kutatásomról. Az előadásom címe *AI learn stellar spectroscopy* volt, melyben bemutattam a kutatás fizikai és technikai motivációt, annak céljait, ismertettem a problémához való megközelítést és az ahhoz használt módszerket, valamint az eddigi eredményeket is.

5. Publikációk

Február végén rövid várakozást követően megjelent *SARS-CoV-2 receptor-binding domain deep mutational AlphaFold2 structures* (Kilim et al., 2023) című cikkünk, melyben a SARS-CoV-2 vírus eddig azonosított 7 fő törzsének mutációt vizsgáltuk az AlphaFold2 szoftver segítségével. A kutatásban 26733 mutáció 3D struktúráját szintetizáltuk, melyeknek teljes adatbázisát publikusan elérhetővé tettük és azt a *Nature – Scientific Data* folyóiratában ismertettük.

Hivatkozások

- [1] Oz Kilim et al. “SARS-CoV-2 receptor-binding domain deep mutational AlphaFold2 structures”. In: *Scientific Data* 10.1 (2023), p. 134.