TDK KONFERENCIA

Program és összefoglalók

2026/2027. I. félév



Debreceni Egyetem, Informatikai Kar

2026. november 5.

Meghívó

Szeretettel meghívunk mindenkit a **Debreceni Egyetem Informatikai Kar Tudományos Diákköri Bizottsága** által a 2026/2027. tanév I. félévében megrendezendő **Tudományos Diákköri Konferenciára**.

Időpont: 2026. november 5., 16:00

Helyszín: Debreceni Egyetem, Informatikai Kar,
F0, F01, 201 nagyelőadók, 202 tanterem

A rendezvényt támogatta:



Tudományos Diákköri Bizottság

Elnök és OTDT képviselő: Dr. Harangi Balázs, mb. tanszékvezető-helyettes, tudományos és pályázati dékánhelyettes

Titkár: Dr. Biró Piroska, egyetemi docens

Információk:

www.ik.unideb.hu/tdk

Ügyintézés:

I227

Kedd: 11:00–12:00

Szerda: 11:00–12:00

Felelős szerkesztők:

Dr. Biró Piroska, egyetemi docens

Dr. Kádek Tamás, adjunktus

Borítót és logót tervezte:

Biró Zsuzsanna, grafikus

Tartalomjegyzék

KÖSZÖNTŐ ÉS TUDNIVALÓK **4**

A TUDOMÁNYOS DIÁKKÖR **5**

PROGRAM **6**
- MEGNYITÓ 6
- SZEKCIÓK 6
- ÜNNEPÉLYES EREDMÉNYHIRDETÉS 6
- INTELLIGENS SZOFTVERRENDSZEREK ÉS SZÁMÍTÁSI INFRASTRUKTÚRÁK 7
- ADATTUDOMÁNY ÉS INTELLIGENS ADATFELDOLGOZÁS 8
- BIZTONSÁG, HÁLÓZATOK ÉS FORMÁLIS MÓDSZEREK 9
- ALKALMAZOTT MESTERSÉGES INTELLIGENCIA, KÉPFELDOLGOZÁS ÉS AUTONÓM RENDSZEREK 10

ÖSSZEFOGLALÓK **12**
- INTELLIGENS SZOFTVERRENDSZEREK ÉS SZÁMÍTÁSI INFRASTRUKTÚRÁK 12
- ADATTUDOMÁNY ÉS INTELLIGENS ADATFELDOLGOZÁS 19
- BIZTONSÁG, HÁLÓZATOK ÉS FORMÁLIS MÓDSZEREK 27
- ALKALMAZOTT MESTERSÉGES INTELLIGENCIA, KÉPFELDOLGOZÁS ÉS AUTONÓM RENDSZEREK 34

A TDK ELŐADÁSOK ÉRTÉKELÉSI SZEMPONTJAI **40**

A RÉSZTVEVŐK NÉVSORA **42**
- HALLGATÓK 42
- TÉMAVEZETŐK 44

Köszöntő és tudnivalók

Köszöntjük a 2026/2027. tanév I. félévi Tudományos Diákköri Konferencia előadóit, társszerzőit, a munkájukat irányító témavezetőket, a bíráló bizottságok tagjait, valamint minden kedves érdeklődőt. Bízunk abban, hogy a megrendezésre kerülő tudományos diákköri konferencia mindenki számára hasznos, új tapasztalatokkal szolgál majd.

Az előadások hossza legfeljebb 15 perc, melyet szintén legfeljebb 5 perces vita követ. Kérjük a résztvevőket az időkeretek pontos betartására. Mindenkit szeretettel várunk november 5-én!

A szervezők

A Tudományos Diákkör

A tudományos és művészeti diákkör a kötelező tananyaggal kapcsolatos ismeretek elmélyítését, a képzési követelményeket, a tantervi tananyagot meghaladó ismeretek elsajátítását, a hallgatói kutatómunkát, illetve a művészeti alkotótevékenységet elősegítő, ennek nyilvánosságot is biztosító önképzőköri forma. A tudományos és művészeti diákköri tevékenység az egyetemi, főiskolai tanulmányok kezdeti időszakában induló vagy az alsóbb évfolyamokon kezdődő, folyamatos tutoriális (mentor) jellegű hallgató-tanár műhelymunka, szakmai kapcsolat, a minőségi értelmiségi képzés fontos területe, a tehetséggondozás legfontosabb, legjelentősebb formája a hazai felsőoktatásban. A diáktudományos és művészeti tevékenység a tudományos és művészeti pályára való felkészítés, felkészülés legmagasabb szintje a doktori iskolát megelőző képzési szakaszban, s mint ilyen, a doktori képzés (PhD-, illetve DLA-képzés) egyik legjobb előiskolája.

A TDK keretei között folytatott tudományos és művészeti tevékenység kitartó, következetes munkán, folyamatos tanuláson és igazi megmérettetésen alapul. Megtanít érvelni, vitatkozni, mások igazát megismerni, elfogadni, néha még a „felnőtt” tudós nemzedéknek is példát mutatva örülni más sikereinek, elért eredményeinek. A szakmai, tudományos sikerek elérése mellett, vagy inkább mindezek előtt igényességre, a gondolkodás meg nem alkuvó becsületességére, a kutatói életforma nagyszerűségére, a felfedezés örömére, az új melletti kiállásra, de együttműködésre és toleranciára is nevel. A TDK-munka vállalása személyes döntés, amely a tudományos munka iránti alázattal, szorgos, kitartó munkával jár. A kölcsönös együttműködésen alapuló műhelymunka tanárnak, diáknak egyformán nagy lehetőség.

Olyan szellemi fellendülést eredményez, amely kedvező hatással van az egyetemi, de továbbtekintve hazánk tudományos és művészeti életének egészére is.

(Forrás: az OTDK kézikönyve)

Program

Felhívjuk a figyelmet, hogy az előadások kezdési időpontjai tájékoztató jellegű adatok, néhány perces eltérések előfordulhatnak.

Megnyitó

A konferencia elnöke: DR. HARANGI BALÁZS, mb. tanszékvezető-helyettes, tudományos és pályázati dékánhelyettes

Helyszín: Debreceni Egyetem, Informatikai Kar, földszint, F0

16:00 – 16:05 A konferencia megnyitása
PROF. DR. HAJDU ANDRÁS, egyetemi tanár, dékán

16:05 – 16:10 Résztvevők köszöntése, általános információk
DR. HARANGI BALÁZS, mb. tanszékvezető-helyettes, tudományos és pályázati dékánhelyettes

Szekciók

16:20 – 18:30, IK-F01
Intelligens szoftverrendszerek és számítási infrastruktúrák

16:10 – 18:40, IK-F0
Adattudomány és intelligens adatfeldolgozás

16:20 – 18:30, IK-202
Biztonság, hálózatok és formális módszerek

16:20 – 18:30, IK-201
Alkalmazott mesterséges intelligencia, képfeldolgozás és autonóm rendszerek

Ünnepélyes eredményhirdetés

19:00 F07
DR. HARANGI BALÁZS, mb. tanszékvezető-helyettes, tudományos és pályázati dékánhelyettes

Minden résztvevő megjelenésére feltétlenül számítunk az ünnepélyes eredményhirdetésen.

Intelligens szoftverrendszerek és számítási infrastruktúrák

Bíráló bizottság:

Prof. Dr. Baran Sándor, egyetemi tanár (elnök)
Dr. Kocsis Gergely, egyetemi docens
Dr. Piros Attila, EPAM

16:20 - 16:40 BIRKAS GERGŐ
Automatikus szoftvertesztelés generalista AI ágensekkel
Témavezető: Dr. Lakatos Róbert

16:40 - 17:00 KUKI CSABA SÁNDOR
Digitális döntéstámogatás a betegbiztonságért: A sürgősségi szakemberek igényei egy új gyógyszerdózis-kalkulátor mobilalkalmazással szemben
Témavezető: Dr. Biró Piroska

17:00 - 17:20 FINTOR GELLÉRT LÁSZLÓ
Élettani stresszválaszok és kiváltó események szinkronizált vizsgálata eseménynaplózó és szenzoros adatgyűjtő keretrendszerrel
Témavezetők: Dr. Gilányi Attila László és Dr. Kósa Karolina

17:20 - 17:40 ABDULLAEV BAKHTIER RAVSHANOVICH ÉS THE HOANG TUNG NGUYEN
Prompt Injection against Detector-Grounded LLM Log Analysts: Grounding, Explanation Changes and Deference
Témavezető: Dr. Herendi Tamás

17:40 - 17:50 Szünet

17:50 - 18:10 KOVALOV GEORGIY
Sparse Holographic Radiance Cascades.
Témavezető: Dr. Adamkó Attila Tamás

18:10 - 18:30 CHEBILI ANAS
Analyzing Pass-Ordering Effects in Multi-Level Compilers for ML Models
Témavezető: Dr. Herendi Tamás

Adattudomány és intelligens adatfeldolgozás

Bíráló bizottság:

Prof. Dr. Ispány Márton, egyetemi tanár (elnök)

Dr. Lakatos Róbert, adjunktus

Dr. Nagy-Lakatos Mária, adjunktus

16:10 - 16:30 SÓGOR GERGELY
Aktivitásfelismerés javítása a szenzorcsatornák megbízhatóságának figyelembevételével
Témavezető: Dr. Tóth János

16:30 - 16:50 PÓSZ PATRICIA
Felhősödés előrejelzések diszkrét utófeldolgozása osztályozó módszerekkel
Témavezető: Prof. Dr. Baran Sándor

16:50 - 17:10 CZAKÓ ORSOLYA
Fixpontok és stabilitás lineáris diszkriminanciaanalízisen alapuló öncímkéző modellekben
Témavezető: Dr. Bogacsovics Gergő

17:10 - 17:30 MAGURA ÁRON
Légifelvételeken és OpenStreetMap-adatokon alapuló szegmentálás a városi kerékpárút-hálózat felismerésére
Témavezető: Dr. Tóth Róbert

17:30 - 17:40 Szünet

17:40 - 18:00 ALY AHMED SAYED AHMED SOLIMAN
Design and Evaluation of a Reproducible Financial Market Data Pipeline with Data Quality Controls
Témavezető: Dr. Tóth Róbert

18:00 - 18:20 AFROZ SAMANTA
Safety-Aware Pedestrian Routing: A Held-Out Evaluation of Crime-Report-Based Route Recommendation
Témavezető: Dr. Vágner Anikó Szilvia

18:20 - 18:40 JUMAEV SHAKHRIYORBEK
Short-Term and Long-Term Forecasting of Corporate Financial Performance with LSTM Neural Networks: Evidence from the U.S. Consumer Staples Sector
Témavezető: Dr. Gáll József Mihály

Biztonság, hálózatok és formális módszerek

Bíráló bizottság:

Prof. Dr. Vaszil György, egyetemi tanár (elnök)
Dr. Pecsora Sándor, adjunktus
Dr. Tóth János, adjunktus

16:20 - 16:40 PANKOTAI IMRE
Decentralizált, NFC-alapú hitelesítési protokoll szimmetrikus kriptográfiával és elosztott kulcskezeléssel
Témavezető: Dr. Oláh Norbert

16:40 - 17:00 MAKÁR PÉTER ÁKOS
Nagypontosságú helymeghatározás terepi mérések automatizálásához
Témavezető: Dr. Kocsis Gergely

17:00 - 17:20 MIERLUT JONATHÁN ISTVÁN
Geolokáció-alapú eseményajánló rendszer fejlesztése és összehasonlító vizsgálata szabályalapú és Transformer-alapú megközelítésekkel
Témavezető: Dr. Biró Piroska

17:20 - 17:40 VAJDA MÁTÉ PÁL
Blokklánc alapú elektronikus szavazási rendszer adminisztrátor nélküli szavazatszámlálással
Témavezető: Dr. Pintér-Huszti Andrea

17:40 - 17:50 Szünet

17:50 - 18:10 TÓTH BENCE ÉS SZÁVA GRÉTA
Fogalomalapú adatátviteli mechanizmus zajos fizikai csatornán: szemantikus adaptív szinonimakódolás mesterséges intelligenciával
Témavezető: Prof. Dr. Gál Zoltán

18:10 - 18:30 TALAL BIN TAYYAB MUHAMMAD ÉS SUMAIYA ASHFAQUE
TRAIL: Formal Conversation Control and Evidence Retrieval in Long LLM Histories
Témavezetők: Dr. Bogacsovics Gergő és Dr. Hannush Carolin

Alkalmazott mesterséges intelligencia, képfeldolgozás és autonóm rendszerek

Bíráló bizottság:

Prof. Dr. Gál Zoltán, egyetemi tanár (elnök)
Dr. Baran Ágnes Éva, egyetemi docens
Dr. Tóth Róbert, adjunktus

16:20 - 16:40 ANGI DÁVID
Dinamikus objektumelkerülés megerősítéses tanulással a CARLA szimulátorban
Témavezető: Mészáros László

16:40 - 17:00 SZENTMIKLÓSSY BÁLINT
Pajzsmirigy szövettani digitális képek elemzése modern mélytanulási technikákkal
Témavezető: Dr. Tomán Henrietta

17:00 - 17:20 TÓMÁN PÉTER
Önvezető járművek sávkövetési pontosságának becslése
Témavezető: Dr. Sütő József

17:20 - 17:30 Szünet

17:30 - 17:50 ABDULAZIZ SALEM ALI SALEM
Investigation of the Impact of Ambient Lighting Conditions on the Navigation Accuracy of Autonomous Vehicles
Témavezető: Dr. Sütő József

17:50 - 18:10 YAR ASFAND
Representation Dynamics of Layer-Normalization Test-Time Adaptation on Camelyon17
Témavezető: Dr. Bogacsovics Gergő

18:10 - 18:30 KUMAR DANISH
Lightweight diffusion model for synthetic minority-class generation
Témavezető: Dr. Bogacsovics Gergő

Összefoglalók

Intelligens szoftverrendszerek és számítási infrastruktúrák

Automatikus szoftvertesztelés generalista AI ágensekkel

BIRKAS GERGÖ

Egy alkalmazás belső működése lehet, hogy hibátlan, viszont a grafikus felhasználói felülete (GUI) mégis akadályozhatja a felhasználót a szoftver célszerű használatában. A megfelelő szoftverminőség biztosításához ezért elengedhetetlen a GUI tesztelés, a tényleges gombok, szövegmezők és egyéb felületi elemek használatával. Ahhoz viszont, hogy ezeket a GUI teszteket automatikus módon el tudjuk végezni, továbbra is rengeteg fejlesztői tapasztalatra és időbefektetésre van szükség.

Ezt könnyíti meg a dolgozatban bemutatott Automated GUI Analysis and Testing Assistant (AGATA) PowerShell-alapú keretrendszer. Meglévő Windows-könyvtárakra építve strukturált, szöveges leírást biztosít a grafikus felület elemeiről, valamint lehetővé teszi azok programozott kezelését. Így az AGATA elősegíti a Nagy Nyelvi Modellekre épülő ágenseket a tesztelendő szoftverek feltérképezésében, és a tesztscriptek létrehozásában.

Az AGATA előnyeinek felmérése tíz Office- és Windows-alkalmazás GUI tesztelésével történt, ahol 74 külön munkameneten keresztül lett összehasonlítva az ágensek natív képessége az AGATA használatával. Átlagos tesztelkészítési idő 53%-al, 40,2 percről 18,9 percre csökkent, ami előny négy modellgeneráción át, GPT 5.5-től GPT 6.1-ig, továbbra is fennmaradt, ahol a legújabb modellek 66,5%-os csökkenést mutattak natív verzióikkal szemben.

A kísérletek során az ágensek több hibamódba is belefutottak, ahol időnként nem létező problémákat feltételeztek, majd ezek alapján tévesen lehetetlennek minősítették a feladatokat. Ugyanekkor előfordult, hogy az ágensek a GUI felületet megkerülve hozták létre a tesztfájlokat, így hamis teszteredményeket állítottak elő. Az AGATA használata viszont ezeket a hibamódokat minimalizálhatja a gondos tervezéssel, a szoftverkörnyezet feltérképezésével, és az ágensek által egyszerűen értelmezhető szöveges leírások biztosításával, lehetővé téve a megbízható GUI-tesztelés gyorsabb és hatékonyabb automatizálását.

Témavezető:

Dr. Lakatos Róbert, adjunktus
Adattudomány és Vizualizáció Tanszék

Digitális döntéstámogatás a betegbiztonságért: A sürgősségi szakemberek igényei egy új gyógyszerdózis-kalkulátor mobilalkalmazással szemben

KUKI CSABA SÁNDOR

A sürgősségi betegellátás során az egyik legfontosabb momentum a gyógyszerdózisok pontos meghatározása. E folyamatok során számos nehezítő tényező (pl. időnyomás, kognitív terhelés, stressz) érheti az ellátókat. A negatív hatásokból fakadó hibákat csökkenthetik a célzottan kifejlesztett, validált döntéstámogató eszközök, hozzájárulva a betegbiztonság növeléséhez. Mivel nem áll rendelkezésre a hazai sürgősségi ellátás specifikumaira épülő mobilalkalmazás, ennek megvalósítása hiánypótló. A fejlesztés előtt elengedhetetlen a célfelhasználók feltérképezése.

Kutatásunkban anonim, strukturált kérdőíves felmérést végeztünk a sürgősségi és prehospitális ellátók körében. A felmérés öt tematikus egységben vizsgálta a demográfiai hátteret, a dózisszámítási nehézségeket, a technológia-elfogadási modellek (TAM/UTAUT) dimenzióit, valamint a felületi (UI/UX) és funkcionális elvárásokat.

Eredményeink igazolták, hogy az ellátók többsége tapasztalt már számítási nehézséget, és egyöntetű nyitottság mutatkozik a hazai eljárásrendekre épülő kalkulátor használatára. A statisztikai elemzések alapján a jövőbeli használati szándékot nem a demográfiai változók (mint az életkor), hanem az észlelt szakmai hasznosság, a digitális rutin és a korábbi hibatapasztalatok határozzák meg. A felületi preferenciák közül a kiszámítható, hatóanyag alapú elrendezés társult a legmagasabb bizalommal, míg a funkcionális elvárások közül a dóziskorlát-figyelmeztetések, a kontraindikációk jelzése és a színkódolás jelentik a legfőbb pilléreket. Az eredmények rávilágítanak, hogy az intuitív elrendezés és a biztonsági funkciók prioritása garantálja a hatékony klinikai döntéstámogatást. A célfelhasználók körében mind az iOS, mind pedig az Android elterjedtnek bizonyult, ezért egy platformfüggetlen mobilalkalmazás fejlesztése lenne célszerű.

Témavezető:

Dr. Biró Piroska, egyetemi docens
Információ Technológia Tanszék

Élettani stresszválaszok és kiváltó események szinkronizált vizsgálata eseménynaplózó és szenzoros adatgyűjtő keretrendszerrel

FINTOR GELLÉRT LÁSZLÓ

A klinikai diagnosztika során a pácienseket érő akut stressz jelentősen torzíthatja a fiziológiai mérések eredményeit és orvosi értelmezését. Az élettani jelek és a környezeti események precíz, milliszekundum-pontosságú szinkronizációja elengedhetetlen, ám a jelenlegi klinikai rendszerekből ez az eseményalapú kontextus hiányzik. Jelen kutatás egy korábbi TDK-munka folytatása, amely rávilágított, hogy a szabad szöveges eseménynaplózás (átlagosan 19,74 másodperces késleltetéssel) módszertanilag jelentősen korlátozza a gyors vegetatív válaszreakciók követését és a mérések standardizálását. A kutatás célja egy új, integrált, protokollvezérelt eseménynaplózó és többérzékelős adatgyűjtő keretrendszer megtervezése és implementálása volt. A kialakított ötrétegű, moduláris szoftverarchitektúra szigorúan elválasztja a tudományos mérés életciklusát a technikai adatgyűjtéstől és a Bluetooth Low Energy alapú szenzorkommunikációtól. A szabad szöveges adatbevitelt egy előre definiált, zárt eseménykészletre épülő, egykattintásos állapotgép-logika váltotta fel. A keretrendszer a Movesense EKG-szenzorból érkező nyers adatokat érintetlenül, bináris formátumban őrzi meg, míg a szívfrekvencia-variabilitási mutatók kinyerését utólagos, transzparens jelfeldolgozási lánc végzi. A verifikációs mérések igazolták, hogy a protokollvezérelt kialakítás sikeresen minimalizálta a rögzítési késleltetést, a rendszer pedig minimális csomagvesztés mellett garantálja az EKG-adatfolyam és az eseménynapló mikroszekundum-szintű szinkronizációját. A nyers adatok megőrzése és a lokális jelfeldolgozás keresztvalidációja rámutatott, hogy a zárt, gyári szenzor-algoritmusok torzíthatják a stresszmutatókat, így a tudományos reprodukálhatóság csak a bemutatott „fehér dobozos” módszertannal biztosítható. A megalkotott keretrendszer mérsékeli a megfigyelői torzítást, ezáltal megbízható alapot teremt a jövőbeli pszichofiziológiai elemzésekhez és a mesterséges intelligencia modellek tanítóadatainak előállításához.

Témavezetők:

Dr. Gilányi Attila László, egyetemi docens
Információ Technológia Tanszék

Dr. Kósa Karolina, egyetemi tanár
Magatartástudományi Intézet

Prompt Injection against Detector-Grounded LLM Log Analysts: Grounding, Explanation Changes and Deference

ABDULLAEV BAKHTIER RAVSHANOVICH ÉS THE HOANG TUNG NGUYEN

Security operations centres increasingly ground LLM analysts in the verdict, probability and SHAP explanation of a machine-learning detector. Attackers can write text into client-controlled log fields such as the User-Agent, and in such a hybrid pipeline that text reaches the LLM twice: directly in the log line, and indirectly through the detector evidence computed from the same field. We call the second path explanation-mediated injection and measure it.

We built an isolated nginx testbed (700 labelled events) and a Random Forest detector with SHAP explanations, and injected 30 payloads from six families into 48 attack events (1,440 cases), with no-payload and benign controls. The LLM saw the raw log (A), the log with the detector block (B), or B with structural defences (C). A counterfactual arm kept the injected log but showed the clean event's detector block; further arms varied the parts of the block or showed a wrong detector. Eight models answered about 95,000 calls under pre-registered hypotheses and event-cluster bootstrap inference.

Grounding reduced gemma4:e4b's payload-induced flip rate from 25.8% to 11.6%; for Claude Haiku 4.5 the reduction (15.9% to 13.1%) was not significant. In the testbed the payloads lowered the detector's probability but never changed its verdict, and showing the clean event's detector block instead changed the flip rate by at most 1.0 percentage point in every model. The larger risk was deference: shown a wrong detector, the five models tested called 8-42% of clean attacks benign. Condition C lowered flips to 6.0% and 5.1% without false positives, but the whole gain came from the input-side defences, and the detector-anchored severity floor inherited the attacker's influence on the detector.

Detector grounding is therefore not a free defence: its benefit has to be measured for each model, and detector-based safeguards should rely on signals the attacker cannot write.

Témavezető:

Dr. Herendi Tamás, egyetemi docens
Számítógéptudományi Tanszék

Sparse Holographic Radiance Cascades.

KOVALOV GEORGIY

Radiance Cascades (RC) approximate global illumination by hierarchically merging radiance intervals, and Holographic Radiance Cascades (HRC) do so noiselessly, in a single shot, and in real time for 2D scenes — with higher light quality than RC: their anisotropic layout thins probes only along each frustum's axis, preserving full probe density perpendicular to it, so every cascade matches the base cascade's spatial accuracy and resolves sharp shadows and distant penumbras that RC's coarser higher cascades blur. In 2D this quality is nearly free, as the doubled angular resolution cancels the halved probe density and both methods scale as O($X^2 log X$); in 3D, the perpendicular plane is two-dimensional: RC thins probes $8\times$ per cascade while quadrupling directions, so each cascade halves in cost and an $X \times X \times X$ scene needs $\Theta(X^3)$ values in total, whereas HRC thins probes only $2\times$ against the same quadrupling, so cascade n holds $X^3 \cdot 2^n$ values and the total is $\Theta(X^4)$, dominated by the highest cascade. We present Sparse Holographic Radiance Cascades (SHRC), which make HRC content-proportional in computation and memory: occupancy masks over emission and transmittance let tracing, interval merging, and shading skip empty regions entirely, while cascade storage is allocated only where light transport occurs. SHRC retains HRC's noise-free, scene-agnostic quality on dynamic scenes with volumetric media and multibounce lighting, in 2D and 3D, with cost that scales with scene content rather than scene size.

Témavezető:

Dr. Adamkó Attila Tamás, egyetemi docens
Információ Technológia Tanszék

Analyzing Pass-Ordering Effects in Multi-Level Compilers for ML Models

CHEBILI ANAS

Motivation: Optimization passes in multi-level compilers do not act independently. An early transformation can change where later compiler stages execute a calculation, even when each pass is correct on its own. I investigate whether the order of MLIR passes changes activation placement. I also test whether this change survives lowering to native code and affects execution time.

Method: I evaluated all 120 orders of five Tensor/Linalg passes on a Dense-GELU-Dense computation using six model-derived shapes. I followed the resulting programs through Tensor MLIR, loop lowering, LLVM, and native code. I checked numerical correctness against independent reference programs, counted the executed activation calls, and measured representative states in randomized paired test blocks. I then repeated the order analysis on two SwiGLU shapes and examined ReLU6 and residual variants.

My results: The main finding is that pass order determines whether the activation is stored once or recomputed inside the following contraction. Exactly 40 of the 120 pipelines cause recomputation, while the remaining 80 preserve the intermediate activation result. This changes the number of activation evaluations from B x H to B x H x O. Here, B is the batch size, H is the hidden width, and O is the output width. One relative-order condition correctly classified every tested Dense-GELU-Dense and SwiGLU result. In the equal-shape Dense-GELU-Dense tests, recomputation was 84.6% slower. The two SwiGLU shapes showed slowdowns of 84.2% and 75.7%.

Conclusion: The results show that pass ordering can control a native-code cost. Fusion can remove an intermediate tensor, but it can also repeat an expensive calculation across an output dimension. A compiler should therefore consider both effects before it applies fusion. The identified order condition explains when this problem occurs and provides a practical basis for a compiler guard.

Témavezető:

Dr. Herendi Tamás, egyetemi docens
Számítógéptudományi Tanszék

Adattudomány és intelligens adatfeldolgozás

Aktivitásfelismerés javítása a szenzorcsatornák megbízhatóságának figyelembevételével

SÓGOR GERGELY

A viselhető szenzorok többcsatornás idősoros adatain alapuló aktivitásfelismerés gyakorlati alkalmazását egyaránt nehezítik a mérési hibák és az egyének közötti eltérések. A dolgozat azt vizsgálja, hogy az adatminőség milyen típusú és mértékű romlása esetén javítja az osztályozás robusztusságát egy külön tanított, csatornaszintű megbízhatósági kapu. A cél egyrészt a módszerrel elérhető javulás számszerűsítése, másrészt annak feltárása, hogy a hibás csatornák felismerése mely esetekben vezet ténylegesen pontosabb aktivitásfelismeréshez.

A kísérleteket a MHEALTH és a PAMAP2 adathalmazon végeztük, egydimenziós konvolúciós neurális hálózat (CNN) és időbeli konvolúciós hálózat (TCN) alkalmazásával. Az elsődleges kiértékelés személyenkénti kihagyásos (leave-one-subject-out) keresztvalidációval történt, rendre tíz, illetve nyolc alany bevonásával. A négy véletlen inicializálásból származó eredményeket alanyonként átlagoltuk, a bizonytalanságot pedig a tesztalanyok közötti variabilitás alapján jellemeztük.

A tiszta adatokon tanított osztályozókat különböző erősségű Gauss-zajjal, időszakos adatkimaradással, illetve teljes csatornakieséssel vizsgáltuk. A bemenet elé helyezett kapu egy kisméretű neurális hálózat, amely a jelek időbeli statisztikái alapján becsülte meg az egyes csatornák megbízhatóságát, majd ennek megfelelően súlyozta az osztályozó bemenetét. A kaput szintetikusan előállított minőségi célértékek felhasználásával tanítottuk. Összehasonlítási alapként a kapu nélküli modellek, valamint egy, a tanítóadatokon kalibrált, rögzített minőségi szabály szolgáltak.

Az elvégzett vizsgálatok erős zaj esetén jelentős, az alanyok és a két osztályozóarchitektúra mentén is következetesen megjelenő javulást igazoltak a kapu alkalmazásával. Ugyanakkor azt is sikerült megmutatni, hogy a csatornaminőség pontos becslése önmagában nem feltétlenül elegendő az aktivitásfelismerés teljesítményének javításához.

Témavezető:

Dr. Tóth János, adjunktus
Adattudomány és Vizualizáció Tanszék

Felhősödés előrejelzések diszkrét utófeldolgozása osztályozó módszerekkel

PÓSZ PATRICIA

Napjainkban minden fontosabb meteorológiai szolgálat ensemble előrejelzésekkel dolgozik, amiket egy numerikus időjárás előrejelző modell különböző paraméterezéssel történő futtatásaival állítanak elő. Ezen előrejelzések gyakran torzítottak, kalibrálatlanok, ezért utófeldolgozást igényelnek, melynek célja, hogy megbecsülje a vizsgált időjárási mennyiség jövőbeli értékének eloszlását. A TDK-dolgozat a gépi tanuláson alapuló utófeldolgozás hasznosságára és szükségességére kíván rávilágítani a felhősödés ensemble előrejelzések tekintetében.

Ezen mennyiség esetén az utófeldolgozás egy osztályozási feladat, mivel a felhőborítottság - a Meteorológiai Világszervezet szabványa szerint - 9 diszkrét értéket vehet fel attól függően, az égbolt hány nyolcadát borítják felhők. A felhősödés ensemble előrejelzések folytonos értéket vesznek fel a [0, 1] intervallumon, a diszkrét célváltozót pedig a különböző rendelkezésre álló adatokból származtatott bemenő paraméterekkel dolgozó gépi tanulási modellekkel próbáljuk előrejelezni, megadva az egyes osztályok valószínűségeit leíró előrejelző eloszlásokat.

Vizsgálatainkhoz az EUPPBench szabadon elérhető adatbázisát használjuk; speciálisan az ECMWF 2017-2018-as évre vonatkozó Nyugat-Európa egy részét lefedő felhősödés, hőmérséklet és csapadék előrejelzéseit, valamint a kapcsolódó felhősödés megfigyeléseket.

A felhasznált utófeldolgozó modellek teljesítményét a CRPS és Logarithmic Score metrikákkal értékeltük ki, megadva a nyers előrejelzésekből számolt mutatókhoz képest vett relatív javulást is. Ezen javulás szignifikáns voltát konfidencia intervallumok segítségével vizsgáltuk, az előrejelző eloszlások kalibráltságát pedig grafikusan, a PIT hisztogramokkal ellenőriztük.

Eredményképp az utófeldolgozó modellek szinte mindegyike statisztikailag szignifikáns javulást hozott az ensemble előrejelzésekhez képest, számottevően növelve azok kalibráltságát, a kapcsolódó időjárási változók bevonása pedig tovább javította a pontosságot.

Témavezető:

Prof. Dr. Baran Sándor, egyetemi tanár
Alkalmazott Matematika és Valószínűségszámítás Tanszék

Fixpontok és stabilitás lineáris diszkriminanciaanalízisen alapuló öncímkéző modellekben

CZAKÓ ORSOLYA

Az iteratív öncímkézés (iterative pseudo-labeling) egy olyan folyamat, melynek során a gépi tanuló modell predikcióit használjuk fel a tanítóhalmaz bővítésére, ami különösen hasznos olyan helyzetekben, ahol az adatok felcímkézése erőforrásigényes vagy kevés címkézett adat áll rendelkezésre. A módszer kockázata, hogy bizonyos esetekben a modell predikciói hibásak lehetnek, melyek pszeudocímkeként visszakerülhetnek a tanítóhalmazba, így a modell saját hibás döntéseit a következő tanítási körökben tovább erősítheti (confirmation bias). Ennek mérséklésére gyakori megoldás, hogy minden tanítási körben a tanítóhalmaz részét képezi valamennyi valódi címkével rendelkező adat, amit a továbbiakban horgonynak nevezek. Ennek kapcsán azt vizsgáltam, milyen feltételek mellett képes a horgony ellensúlyozni a hibás visszacsatolást, mikor maradhat fenn ennek ellenére is hibás állapot, valamint mekkora kritikus horgonysúly szükséges annak megszüntetéséhez.

Az öncímkézés vizsgálatához egy egydimenziós, kétosztályos, szimmetrikus, lineáris diszkriminanciaanalízisen alapuló modellt választottam, amely amellett, hogy az iteratív folyamat lényegi működését leírja, analitikusan is vizsgálható. Ebben az iteráció egy döntési küszöb változásával jellemezhető, ami lehetővé teszi a fixpontok, stabilitásuk és vonzási tartományaik elemzését, valamint annak meghatározását, mikor válhat egy hibás állapot önfenntartóvá és milyen beavatkozással szüntethető meg.

A további vizsgálatokban azt elemeztem, milyen mechanizmusok tartják fenn vagy szüntetik meg a hibás öncímkézési állapotokat, és hogyan változik ezek szerepe különböző osztályarányok, osztályszámok és modellbeállítások mellett. A jelenségeket az egyszerűsített modellben analitikusan, valós képi adatok rögzített reprezentációin, tanuló neurális hálóval, valamint többosztályos és eloszláseltolódásos beállításokban empirikusan vizsgáltam.

Témavezető:

Dr. Bogacsovics Gergő, adjunktus
Adattudomány és Vizualizáció Tanszék

Légifelvételeken és OpenStreetMap-adatokon alapuló szegmentálás a városi kerékpárút-hálózat felismerésére

MAGURA ÁRON

A városi kerékpáros infrastruktúra dinamikus bővülése új lehetőségeket teremt az automatizált nyilvántartási és felmérési eljárások alkalmazására, kiegészítve a hagyományos helyszíni bejárásokat. A légifelvételek mélytanulási elemzése ígéretes megoldás, azonban a tanításhoz szükséges referenciaadatok kézi előállítása rendkívül időigényes.

Kutatásom célja egy olyan szemantikus szegmentációs eljárás vizsgálata, amely az OpenStreetMap (OSM) nyílt, de hiányosságokkal és pozícióhibákkal rendelkező adataira támaszkodva ismer fel kerékpárutakat nagy felbontású légifelvételeken. Vizsgálati területként Valencia városát választottam kiterjedt kerékpárút-hálózata és a területről nyilvánosan elérhető 15 cm/pixel felbontású ortofotók miatt. A szegmentációhoz egy előtanított ResNet-34 enkóderrel ellátott U-Net-architektúrát alkalmaztam, a térbeli adatszivárgást és téves detektálásokat hálózatsűrűséggel rétegzett térbeli blokkokkal, valamint negatív mintavételezéssel mérsékeltem. A modellt szigorú és térbeli toleranciasávval ellátott relaxált metrikákkal értékeltem ki a teszthalmazon. Annak érdekében, hogy elkülönítsem a valós modellhibákat a referenciaadatok címkezajától, egy rétegzett véletlen mintán háromoldalú összevetést végeztem a modell előrejelzései, a térképi réteg és a kézzel annotált tényleges infrastruktúra között.

A pozitív tesztcsempék felén a modell szigorúan (IoU $\geq$ 0,6), további negyedén 1,05 m-es tűréssel egyezett az OSM-maszkkal. A csempék 22%-án jelentős hiány mutatkozott, amit a vizsgált példák alapján a lombkoronák és árnyékok kitakarása, valamint a peremhatások okozhattak, viszont a negatív kontrollcsempék 89%-án a modell egyáltalán nem jelzett tévesen. A rétegzett mintán (n = 15) szigorú átfedésben a modell kissé jobban illeszkedett a kézi referenciához, 1,05 m-es tűréssel azonban az OSM bizonyult jobbnak a modell szakaszos hiányai miatt. Az eredmények jelzik, hogy az eljárás kiindulópontot adhat az önkéntes térképi adatok geometriai pontosításához.

Témavezető:

Dr. Tóth Róbert, adjunktus
Információ Technológia Tanszék

Design and Evaluation of a Reproducible Financial Market Data Pipeline with Data Quality Controls

ALY AHMED SAYED AHMED SOLIMAN

Reliable financial analysis depends on the quality, consistency, and reproducibility of the underlying market data. Financial data pipelines face challenges such as inconsistent API responses, invalid observations, duplicate records, and the need to recover from processing failures. This research presents the design and implementation of a layered ETL pipeline for acquiring, validating, and preparing historical equity-market data for analytical applications.

The system uses Python and the Twelve Data API to collect daily open, high, low, close, and volume (OHLCV) observations for five US equities. API responses are preserved in their original JSON format to support traceability and independent reprocessing. A canonical data model standardizes the observations, while explicit validation rules detect invalid prices, inconsistent price ranges, and negative trading volumes. Valid records are stored in partitioned Apache Parquet files, and rejected records are isolated in a quarantine layer together with their validation failures.

The architecture separates raw data preservation, data-quality processing, analytical storage, and business transformations. It defines a PostgreSQL dimensional warehouse for financial observations and derived indicators, with explicit keys, constraints, and an idempotent loading strategy. Automated tests verify parsing, validation, storage behavior, error handling, and repeated processing without duplicate Silver records.

The main contribution is a reproducible, testable approach to financial data engineering that emphasizes data integrity and transparent processing decisions. The work provides a practical foundation for dependable financial analytics and demonstrates how layered architecture, validation, and idempotency can improve the reliability of data preparation.

Témavezető:

Dr. Tóth Róbert, adjunktus
Információ Technológia Tanszék

Safety-Aware Pedestrian Routing: A Held-Out Evaluation of Crime-Report-Based Route Recommendation

AFROZ SAMANTA

SafeWalk is a pedestrian navigation app that ranks Google Maps walking alternatives by adding a penalty for reported incidents near each route. I tested whether this reduces exposure to later incidents, which had not been evaluated.

Method: The real recommendation endpoint was run on 568 random origin-destination pairs (1-3 km) in Nottingham, Cambridge, Birmingham and York, using police.uk street-level crime data. Routes were chosen from 2025 incidents and judged on held-out incidents from January-June 2026 within 50 m of the route. I compared the recommender with the shortest route, a randomly chosen alternative and a count-only variant (every incident counts 1) and varied the buffer radius, the penalty scale and the severity weights (hand-chosen versus weights derived from the ONS Crime Severity Score). Differences were paired by trip with 95% bootstrap intervals.

Results: Safety-aware routing reduced exposure to held-out incidents by 21-29% compared with the shortest route, with a mean detour of only 2-4%, in all four cities. A random alternative gave no reduction (-0.3 to +6.5%) and a longer detour, so the benefit comes from the selection. Weighting incidents by severity did not outperform counting them, even for robbery alone, and the two weighting schemes chose the same route in 88-97% of trips. A 100 m buffer captured about 90% of all incidents. The penalty scale was the sensitive parameter: beyond about 5 m per severity point routes stopped changing, so the original default of 50 m per point ignores walking distance on dense data, while values of about 1-5 keep the safety gain with distance still counting.

Conclusion: Penalising routes by nearby reported incidents lowers exposure to future recorded incidents at a small detour. Results concern recorded crime in four English cities and do not measure victimisation or perceived safety.

Témavezető:

Dr. Vágner Anikó Szilvia, adjunktus
Információ Technológia Tanszék

Short-Term and Long-Term Forecasting of Corporate Financial Performance with LSTM Neural Networks: Evidence from the U.S. Consumer Staples Sector

JUMAEV SHAKHRIYORBEK

Most studies that use deep learning in finance try to predict stock prices. Far fewer try to forecast the figures in the financial statements themselves, even though analysts, lenders and company managers work with these forecasts all the time. In this paper I test how well a Long Short-Term Memory (LSTM) network can forecast the financial performance of U.S. Consumer Staples companies, and how the accuracy changes when the forecast horizon grows from one quarter to five years.

I collected quarterly financial statement data for 200 U.S.-listed Consumer Staples companies from SEC EDGAR XBRL, covering 2015 to 2024. The model has two stacked LSTM layers with 64 units each and a dropout of 0.2, and it uses the previous eight quarters as input. It predicts four indicators at the same time: revenue, net income, diluted earnings per share and operating margin. I trained it on 190 companies and kept the remaining ten completely separate for testing. For longer horizons the model forecasts recursively, one quarter at a time. At each step I also rebuild the other input features from the new predictions instead of keeping them fixed, which is the main methodological contribution of the work.

In the short term the forecasts were accurate. The one-quarter-ahead error for revenue was 3.2% MAPE, and for all four indicators the error stayed under 10% up to four quarters ahead. Net income was the hardest to predict (9.5% at Q+4), while operating margin was the most stable. The model got the direction of change right in 75–85% of cases. Over longer horizons the error kept growing and reached 18.3% for revenue after five years. My results suggest that LSTM models are useful for short-term forecasting of company fundamentals in a stable sector, but their long-term forecasts should be treated with care.

Témavezető:

Dr. Gáll József Mihály, egyetemi docens
Alkalmazott Matematika és Valószínűségszámítás Tanszék

Biztonság, hálózatok és formális módszerek

Decentralizált, NFC-alapú hitelesítési protokoll szimmetrikus kriptográfiával és elosztott kulcskezeléssel

PANKOTAI IMRE

A globális ellátási láncok digitalizációjával a fizikai termékek biztonságos hitelesítése egyre nagyobb kihívás, mivel a hagyományos áruvédelmi mechanizmusok nem nyújtanak megfelelő védelmet a kifinomult hamisításokkal szemben. Ezt jól szemlélteti a 2020-ban felderített olaszországi Sassicaia-ügy, melynek során az elkövetők az eredetivel vizuálisan megegyező csomagolás, illetve hologramos zárjegyek alkalmazásával láttak el több ezer palack kommersz bort, bizonyítva a jelenleg elterjedt statikus fizikai védelmi megoldások sebezhetőségét. Bár az NFC chipek hatékony védelmet és dinamikus azonosítást tesznek lehetővé, a szerveroldali infrastruktúra központosított kulcskezelése egyetlen meghibásodási pontot (SpoF) hoz létre. A dolgozat célja egy decentralizált, NFC-alapú termékhitelesítési séma tervezése és implementálása, amely figyelembe veszi az NFC chipek korlátozott erőforrását és kiküszöböli a központosított hitelesítési rendszerek biztonsági kockázatait. A javasolt megoldás a hitelesítéshez szükséges szimmetrikus AES kulcsot a Shamir-féle titokmegosztással titokrészekre bontja az elosztott hálózati csomópontok között. A termék eredetiségének vizsgálatát a csomópontok biztonságos többpárti számítással (SMPC) végzik el, mely magába foglalja a BGW-protokoll és Beaver-hármasok alkalmazását. Így a termék hitelesítés anélkül történik, hogy a teljes titkos kulcs a folyamat során bárkinél rekonstruálódna. A decentralizált ellenőrzések auditálhatóságát a Hyperledger Fabric engedélyköteles blokklánc technológia integrálása biztosítja. Ez manipulációmentes naplózást garantál, mely a „rosszindulatú, de óvatos” támadói modellel szemben is védelmet nyújt. A Python-alapú szoftveres prototípus értékelése alapján a megvalósított decentralizált AES-hitelesítés csekély hálózati forgalmat igényel, és a hálózat méretétől függően 2–6 másodperces válaszidőt eredményez. Ez az eredmény prémium ipari felhasználásnál elfogadható kompromisszumot kínál a biztonságosabb eredetiségvizsgálatért.

Témavezető:

Dr. Oláh Norbert, adjunktus
Adattudomány és Vizualizáció Tanszék

Nagypontosságú helymeghatározás terepi mérések automatizálásához

MAKÁR PÉTER ÁKOS

A nagypontosságú helymeghatározás kulcsfontosságú szerepet tölt be a geodéziai, geoinformatikai, környezetmonitorozási és precíziós mezőgazdasági alkalmazásokban. A kutatás célja egy költséghatékony, RTK-GNSS technológiára épülő autonóm rendszer fejlesztése, amely egy rover segítségével képes a terepi adatgyűjtési folyamatok támogatására és automatizálására, miközben nagy pontosságú pozícióinformációt biztosít.

A precíziós pozicionáló rendszer egy többfrekvenciás GNSS vevőegységre, RTK korrekciós szolgáltatásokra, valamint egy beágyazott számítógépes platformra épül. A fejlesztés részeként olyan szoftveres környezet készül, amely lehetővé teszi a mérési adatok gyűjtését, tárolását, feldolgozását és megjelenítését, továbbá támogatja a mérési pontok kezelését és a terepi munkafolyamatok nyomon követését. A kutatás kiemelt feladata olyan adatfeldolgozási és hibacsökkentő módszerek kidolgozása, amelyek ismételt mérések és minőségi mutatók figyelembevételével növelik a helymeghatározás pontosságát és megbízhatóságát.

A fejlesztett autonóm rover a navigációs algoritmusok és a nagy pontosságú pozícióadatok felhasználásával képes előre meghatározott mérési pontok megközelítésére és mérési feladatok végrehajtására.

Témavezető:

Dr. Kocsis Gergely, egyetemi docens
Informatikai Rendszerek és Hálózatok Tanszék

Geolokáció-alapú eseményajánló rendszer fejlesztése és összehasonlító vizsgálata szabályalapú és Transformer-alapú megközelítésekkel

MIERLUT JONATHÁN ISTVÁN

A digitális eseményplatformok növekvő kínálata megnehezítheti a felhasználók számára a személyes érdeklődésüknek és aktuális helyzetüknek megfelelő programok megtalálását. Kutatásom célja egy geolokáció-alapú eseményajánló mobilalkalmazás fejlesztése, valamint két eltérő ajánlási megközelítés felhasználói vizsgálaton alapuló összehasonlítása.

Az elkészített Eventure alkalmazás a felhasználó által megadott érdeklődési kategóriák, a maximális utazási távolság és az aktuális földrajzi helyzet figyelembevételével ajánl eseményeket. A kutatás során egy explicit, szabályalapú és súlyozott ajánlási módszert hasonlítok össze egy előtanított Transformer-modellre épülő szemantikus ajánlási megközelítéssel. A résztvevők véletlenszerűen kerülnek a két rendszer egyikéhez, miközben az alkalmazás felülete és a tesztelés körülményei mindkét csoport számára azonosak. A felhasználók több ajánlást értékelnek azok megfelelősége és a részvételi szándék szempontjából, amelyeket egy kiegészítő kérdőíves felmérés egészít ki.

A vizsgálat célja annak feltárása, hogy egy egyszerűbb, jól értelmezhető szabályalapú rendszer, illetve egy szemantikus reprezentációkat alkalmazó gépi tanulási megközelítés milyen mértékben képes releváns eseményajánlásokat biztosítani. A kutatás így nemcsak a két módszer teljesítményének összehasonlításához járul hozzá, hanem az eseményajánló rendszerek gyakorlati tervezéséhez és a komplexebb gépi tanulási módszerek alkalmazhatóságának értékeléséhez is.

Témavezető:

Dr. Biró Piroska, egyetemi docens
Információ Technológia Tanszék

Blokklánc alapú elektronikus szavazási rendszer adminisztrátor nélküli szavazatszámlálással

VAJDA MÁTÉ PÁL

A közigazgatás gyors digitalizációja a biztonságos, skálázható és egyetemesen ellenőrizhető elektronikus szavazást a modern kriptográfiai kutatások kiemelten fontos témájává tette. Egy olyan anonimitást és adatvédelmet biztosító, ellenőrizhető, időalapú elektronikus szavazási protokollt javasolunk, amelyet a Hyperledger Fabric rendszerhez terveztünk.

A szavazólapokat úgynevezett „one-hot" vektorokként kódoljuk, majd komponensenként titkosítjuk homomorf elliptikus görbéjű ElGamal eljárással, ezen komponensek érvényességét 1-out-of-n típusú, nem interaktív nullaismeretű bizonyításokkal tesszük. A többszöri szavazást a kriptográfiai ellenőrzést megelőző, konstans idejű indexelt állapotlekérdezésekkel akadályozzuk meg, miközben a szavazási fázisok közötti átmeneteket a TimeFabric megbízható időbélyegei biztosítják. A megbízható időbélyegek lehetővé teszik a fázisok közötti automatikus váltást és az automatikusan indított szavazatszámlálást, anélkül, hogy a számlálási fázisban adminisztrátorra lenne szükség.

A rendszer biztonságát egy valószínűségi polinomiális idejű modellben formalizáljuk, amely védelmet nyújt a rosszindulatú hatósági szereplőkkel és szavazókkal szemben. Végül megvalósítási szempontokat is bemutatunk: részletezzük az $O(n^2)$ nagyságrendű jóváhagyási többletterhelést, valamint megmutatjuk, hogy tíz jelölt esetén a szavazat mindössze 4,18 KB helyet foglal. Ezzel bizonyítjuk a rendszer gyakorlati kivitelezhetőségét az engedélyköteles elosztott blokkláncon.

Témavezető:

Dr. Pintér-Huszti Andrea, egyetemi docens
Adattudomány és Vizualizáció Tanszék

Fogalomalapú adatátviteli mechanizmus zajos fizikai csatornán: szemantikus adaptív szinonimakódolás mesterséges intelligenciával

TÓTH BENCE ÉS SZÁVA GRÉTA

Egy olyan beágyazott, ESP32-mikrokontrollerekre épülő, fogalomalapú szemantikus kommunikációs rendszert fejlesztettünk ki, amely zajos környezetben robusztus adatátvitelt biztosít. A hagyományos, bitszintű pontosságra törekvő megoldásokkal ellentétben rendszerünk a természetes nyelvek adaptív redundanciáját veszi alapul és az információ nyelvi jelentéstartalmát védi. A kommunikáció során többszintű szemantikai kódolást és PAM4-modulációt használunk. A rendszer egy 258×5-ös szemantikus kódkönyvet definiál, amely fogalmanként 5 független szinonimát tartalmaz. Ha egy kódszó a csatornazaj miatt torzul, a rendszer újraküldés helyett egy másik szinonimára vált. A vevő oldal egy exponenciálisan súlyozott mozgóátlag (EWMA) segítségével folyamatosan figyeli a csatorna minőségét, és adaptív nyugták (ACK, NACK) valamint automatikus szintváltások révén igazítja az átviteli paramétereket a pillanatnyi viszonyokhoz. A szinonimák zajmentesítését egy pehelysúlyú neurális hálózat végzi. A gyorsabb működés érdekében, szigorú feltételek teljesülése mellett, ki lehet kerülni a mesterséges intelligencia használatát (AI bypass). Eredményeink azt igazolják, hogy a bitismétlés elhagyásával és fogalomalapú kódolással még a kiemelkedően zajos peremhálózati környezetben is fenntartható megbízható kommunikáció. Ez új és hatékony alternatívát jelent a modern B5G/6G rendszerek fejlesztésében.

Témavezető:

Prof. Dr. Gál Zoltán, egyetemi tanár
Informatikai Rendszerek és Hálózatok Tanszék

TRAIL: Formal Conversation Control and Evidence Retrieval in Long LLM Histories

TALAL BIN TAYYAB MUHAMMAD ÉS SUMAIYA ASHFAQUE

As conversations with large language models grow, repeatedly supplying full history increases context use and can introduce information unrelated to the current task. We aim to make memory access explicit and select relevant evidence for each request. Long conversational histories create separate problems of control, evidence selection and answer generation. TRAIL (Threaded Regular Agent Interaction Language) addresses the control boundary through a regular language whose fields specify thread identity, memory eligibility, operations and execution policy. The methodology separates syntax recognition, finite compatibility rules and state dependent execution. We applied Lean proofs to verify selected deterministic properties; runtime conformance is evaluated separately from semantic resolution and answer correctness. The retrieval architecture combines an authorized candidate pool with either a deterministic evidence selector or a SPLADE ranker. On a retrospective, oracle assisted scope stress panel, TRAIL+SPLADE retains source recall 0.9669 while reducing mean context from 8,044 to 4,620 tokens and increasing MRR from 0.9350 to 0.9735 compared with global SPLADE. A separate 500-question LongMemEval-S evaluation with a Qwen3-8B reader reports 33.0% official answer accuracy for oracle-scoped TRAIL+SPLADE and 28.4% for global SPLADE. An equivalent JSON scope gives the same 33.0%. Separate interface benchmarks show that TRAIL sigils require 19–31 bytes to represent control instructions, compared with 293–299 bytes for full-control JSON, excluding the message body. Median parsing time for TRAIL remains below 0.03 ms across tested message sizes; JSON is faster for short message bodies, while TRAIL is faster for 32 KB bodies. Finally, our main scientific contribution is a formal, executable interface that makes memory access explicit and combines scope control with evidence retrieval, supported by Lean proofs of selected properties and measured gains in context efficiency.

Témavezetők:

Dr. Bogacsovics Gergő, adjunktus
Adattudomány és Vizualizáció Tanszék

Dr. Hannush Carolin, adjunktus
Számítógéptudományi Tanszék

Alkalmazott mesterséges intelligencia, képfeldolgozás és autonóm rendszerek

Dinamikus objektumelkerülés megerősítéses tanulással a CARLA szimulátorban

ANGI DÁVID

Dolgozatomban egy megerősítéses tanulás alapú önvezető ágenst fejlesztettem, amely autópályán és városban egyaránt navigál, és a lassabb járműveket a KRESZ betartásával előzi meg: szaggatott vonal és szabad bal sáv esetén kiáll, majd visszasorol. A kísérleteket a CARLA szimulátor Town04-es pályáján végeztem. Az ágens a Stable-Baselines3 SAC algoritmusa folytonos kormány- és gázakcióval. Egy 160×80 pixeles RGB-kamera és 64 csatornás LiDAR nyers adatai helyett saját autoenkóderek látenseit kapja; ezekhez szenzoronként 40k mintát rögzítettem és három architektúrát hasonlítottam össze. Az ágenssel együtt tanuló jellemzőkinyerő a LiDAR 16×11×32-es jellemzőtérképét és a kamera látensét 256 elemre képezi, majd összeadással fuzionálja; ehhez egy MLP-ág 64 elemű nem szenzoros jellemzőt fűz, így 320 elemű vektor jut az ágenshez. A legösszetettebb rész a jutalomfüggvény volt: öt változatot értékeltem 10 rögzített, összesen 34 km-es, sűrű forgalmú útvonalon, párokban, egy-egy változtatás hatását mérve. Az első a sávtartást és a szabályos előzést jutalmazta. A második TTC-büntetése óvatosabbá tette az ágenst (37,7 → 34,4 km/h, sávközépen töltött idő 88 → 72%), az ütközések viszont nem csökkentek szignifikánsan (0,96 → 0,91 db/km). A harmadikban (TTC-küszöb 7 → 3 s, előzési alaptag 0,3 → 0,4, +5-ös bónusz +2/+3 részre bontva, büntetés lassabb jármű mögötti haladásra) a várttal ellentétben óvatosabb lett: simább kormányzás (0,036 → 0,029), de kisebb sebesség (34,4 → 31,0 km/h) és kevesebb előzés (121 → 104). A negyedik a kézzel formált tagokat egy haladás-alapú taggal váltotta ki (34,7 km/h, 128 előzés, 0,027-es kormányváltozás). Az ötödik potenciál-alapú formálással (előzés alatt 2-ről 5-re növő potenciál) 36,3 km/h-t, 97,6%-os útvonalteljesítést (96,1% helyett) és 149 előzést ért el, de ezek az eltérések nem szignifikánsak, összhangban azzal, hogy a formálás elméletileg csak a tanulást gyorsítja.

Témavezető:

Mészáros László, tanársegéd
Adattudomány és Vizualizáció Tanszék

Pajzsmirigy szövettani digitális képek elemzése modern mélytanulási technikákkal

SZENTMIKLÓSSY BÁLINT

A TCGA a rákkutatás történetének egyik meghatározó kutatási projektje. A dolgozatom során a TCGA-THCA adatbázisból származó, digitalizált szövettani felvételeket (Whole Slide Imaging – WSI) használtam fel. A gigapixeles felbontású metszetekből előfeldolgozás során kisebb képrészleteket (patch) képeztem. A kialakított adatbázisom megközelítőleg 360 ezer tanító képet tartalmaz, amelyet 500 beteg felvételeiből mintavételeztem. A modellezést egy háromlépcsős keretrendszerben valósítottam meg: kiinduló referenciaként (baseline) egy egyszerű, kilapított képi vektorokon (flattened features) alapuló osztályozót értékeltem ki, amelyet egy szövettani mintázatokra optimalizált, nulláról tanított egyedi konvolúciós neurális hálózat (CNN), majd egy ImageNet-en előtanított és a feladathoz finomhangolt ResNet-50 architektúra követett. A modellek teljesítményét standard metrikák (pontosság, szenzitivitás, specificitás, F1-pontszám, valamint ROC-AUC görbék) mentén, kiterjedt tévesztési mátrixok elemzésével hasonlítottam össze. A dolgozat eredményei rámutatnak a mély reprezentációk és az átviteli tanulás (transfer learning) hozzáadott értékére a szövettani diagnosztikában a hagyományos megközelítésekkel szemben.

Témavezető:

Dr. Tomán Henrietta, adjunktus
Adattudomány és Vizualizáció Tanszék

Önvezető járművek sávkövetési pontosságának becslése

TÓMÁN PÉTER

Napjaink önvezető járműveinek és vezetéstámogató rendszereinek egyik legdinamikusabban kutatott, biztonságkritikus feladata a robusztus sávtartás.

A dolgozat írása során elsősorban arra a kérdésre kerestem a választ, hogy a különböző sávkövetést megvalósító módszerek mennyire hatékonyan működnek a gyakorlatban. Az összehasonlításhoz többféle metrikát használtam. Bizonyos metrikák teszthalmazon számított pontosságot vizsgálnak, mások a valós jármű vezérlésének minőségét mérik.

Az általam választott kutatási feladatra, tesztjárműként, a népszerű Donkey car platformot használtam. A sávkövetés pontosságának mérését a Debreceni Egyetem Informatikai Karának laborjában kialakított tesztpályán végeztem. A jármű irányításához egy saját robotikai alkalmazást fejlesztettem ki ROS2 (Robot Operating System 2) alapokon. A ROS2 keretrendszer segítségével lehetőség nyílt egy olyan vezérlő alkalmazás létrehozására, amely a későbbiekben minimális szoftveres változtatással futtatható lesz más típusú tesztjárműveken is. Ezen az alkalmazáson belül, definiáltam az adatgyűjtéshez használt módszertant, amivel a tanításhoz szükséges adatokat rögzítettem. Az így összegyűjtött adatok alapján különböző mélytanuló modelleket hoztam létre és tanítottam be.

A mérések rávilágítottak, hogy nincs egyértelmű győztes: mindegyik modell alkalmazása valamilyen metrika szempontjából kompromisszumot követel.

A mérések jól mutatják a módszerek közti kompromisszumot. A geometriai modell finomabb kormányzást biztosít (az 1. és 2. derivált 0,0073 és 0,0100, szemben a CNN 0,0143 és 0,0215 értékeivel). A konvolúciós modell viszont stabilabban követi a sáv közepét, ráadásul a geometriai modell 4,53 milliszekundumos futási idejét kevesebb mint a felére, 2,14 milliszekundumra csökkenti.

Témavezető:

Dr. Sütő József, egyetemi docens
Informatikai Rendszerek és Hálózatok Tanszék

Investigation of the Impact of Ambient Lighting Conditions on the Navigation Accuracy of Autonomous Vehicles

ABDULAZIZ SALEM ALI SALEM

This study investigates the effect of ambient illumination on the navigation accuracy of a camera-based small-scale autonomous vehicle. A DonkeyCar platform equipped with a camera and a deep-learning model was tested on an indoor track under six illumination conditions ranging from normal laboratory lighting to extreme low-light conditions. Five autonomous runs were performed for each condition, resulting in 30 experimental runs.

Software was developed to evaluate navigation accuracy from recorded frames. Canny edge detection was applied, and detected boundary points were fitted using RANSAC to estimate the lane center. The difference between the lane center and the camera center was used to calculate the mean squared pixel error (MSE). Because increasingly dark conditions also caused the evaluator to reject more frames, detection coverage was reported separately and an adjusted MSE was additionally used to reduce selection bias in the raw error values.

The results show that large reductions in illumination did not produce proportionally large degradation in navigation performance. The vehicle successfully completed the runs under the first five conditions, while all five runs under the darkest condition resulted in navigation failure. Even under that condition, however, the vehicle navigated most of the track. Raw MSE decreased from 221.91 px^2 in the brightest condition to 110.71 px^2 in the darkest one, while the evaluator's detection failure rate increased from 1.69% to 66.82%. After accounting for detection coverage, the adjusted score increased to 336.23 px^2. These results also show that evaluation based only on successfully measurable frames can underestimate performance degradation under severe low-light conditions.

Témavezető:

Dr. Sütő József, egyetemi docens
Informatikai Rendszerek és Hálózatok Tanszék

Representation Dynamics of Layer-Normalization Test-Time Adaptation on Camelyon17

YAR ASFAND

Test-time adaptation promises robustness to hospital-to-hospital shift without target labels, retraining, or source data, and its cheapest form adjusts a small parameter subset at inference to sharpen the model's predictions. Almost all supporting evidence, however, comes from convolutional networks with batch normalization evaluated on synthetic corruptions, which describes neither the architecture nor the shift a histopathology model meets.

We evaluate the layer-normalization analog of this method on Camelyon17-WILDS, where the benchmark splits by hospital rather than by image, across six frozen backbones spanning supervised, self-distilled, masked-reconstruction and joint-embedding-predictive pretraining at two scales, with three seeds. In the configuration most commonly reported, one gradient step over the normalization affine parameters, accuracy moves by less than half a percent for every backbone.

Rather than stop at that null result, we instrument the adaptation and measure how far the penultimate representation travels, as a mean cosine distance between frozen and adapted features. It travels between 1.9e-6 and 5.9e-5, two to four orders of magnitude below a one percent change. The absence of an effect does not show that these representations resist adaptation. It shows they were barely touched.

Adapting harder does not rescue the method. Widening the adapted subset to the final transformer block and raising the step count to twenty grows displacement by roughly five orders of magnitude, but accuracy falls monotonically along both axes, to a mean loss of 11.6 percent at the largest budget, and the worst backbone drops to the chance floor.

The quiet standard configuration is therefore not a safe version of this method but a sub-threshold one. Representation drift belongs alongside accuracy in any evaluation of test-time adaptation, since accuracy alone cannot distinguish an intervention that does nothing from one kept harmless only by its smallness.

Témavezető:

Dr. Bogacsovics Gergő, adjunktus
Adattudomány és Vizualizáció Tanszék

Lightweight diffusion model for synthetic minority-class generation

KUMAR DANISH

Class imbalance is a common problem in medical image classification, especially when the underrepresented class is the clinically important one. The ISBI 2025 PS3C Pap smear cell classification dataset is a good example: out of 81,632 labeled training images in three categories, only 2,366 are abnormal ("unhealthy") cells. Standard augmentation techniques such as flipping, rotating, color jitter, and random erasing improve things, but only up to a point. They shuffle around what the model has already seen without adding cell shapes or structures that were never in the training set.

This work explores whether a small conditional denoising diffusion probabilistic model (DDPM) can fill that gap by generating synthetic unhealthy-class images with morphological variation the original data lacks. The model uses a reduced U-Net backbone with three channel stages instead of four, self-attention at one resolution, and 500 diffusion steps rather than the standard thousand. It fits on a single 8 GB consumer GPU and trains overnight. To measure what these reductions cost, a full-sized DDPM with four channel stages, multi-resolution attention, and 1,000 steps is trained and evaluated under the same conditions.

The downstream classifier is a two-model EfficientNetV2-S ensemble. The first model stays the same across every experiment; only the second model's training data changes. The DDPM-generated images are compared against three alternatives: no extra data at all, images from a conditional GAN (DCGAN), and class-specific heavy augmentation, at 250 and 500 added images. A combined condition tests whether synthetic data and heavy augmentation work better together than either one alone. All conditions use the same data split and held-out test set, evaluated with weighted F1 and per-class recall. Synthetic image quality is measured separately using Fréchet Inception Distance and t-SNE plots of real versus generated samples.

Témavezető:

Dr. Bogacsovics Gergő, adjunktus
Adattudomány és Vizualizáció Tanszék

A TDK előadások értékelési szempontjai

1. Előadói stílus, gazdálkodás az idővel (0-10 pont)

***a)* Stílus**

0 pont – ha az előadás csapongó, hiányos;

2 pont – ha az előadás nehezen követhető, gondatlanul szerkesztett, nyelvtani hibával;

4 pont – ha az előadás csak kisebb hibákat tartalmaz, érthető;

6 pont – ha az előadás gyakorlatilag hibátlan, jól követhető.

***b)* Gazdálkodás az idővel**

0 pont – ha az előadást az elnöknek kell leállítani;

2 pont – ha az előadás részei aránytalanok, vagy az előadót figyelmeztetni kell;

4 pont – ha az előadás arányos, tartja az időt.

2. Szemléltető eszközök használata (0-5 pont)

***a)* A prezentált anyag minősége**

0 pont – rossz minőségű prezentációs anyag;

1 pont – megfelelő minőségű prezentációs anyag;

2 pont – nagyon jó.

***b)* A prezentált anyag bemutatásának minősége**

0 pont – csak felolvas;

1 pont – csak kevés többletet ad a kész prezentációhoz képest;

2 pont – magyarázza az ábrákat, értelmezi az ottani állításokat;

3 pont – kiváló előadó.

3. Eredmények bemutatása (0-10 pont)

***a)* Az eredmények mennyisége**

0 pont – nincs kiemelkedő eredmény, és a ráfordított munka mennyisége is megkérdőjelezhető;

2 pont – nincs kiemelkedő eredmény, de sok munka van benne;

4 pont – sok munka, sok eredménnyel.

***b)* Az eredmények bemutatási módja**

0 pont – gyakorlatilag nincsenek eredmények vagy nem mutatja be;

2 pont – az eredmények bemutatása nem hangsúlyos;

4 pont – ha az eredmények egyértelműen megállapíthatók, de nem lát módot a hasznosításra/közlésre;

6 pont – ha az eredmények egyértelműen megállapíthatók, van működő, tesztelt berendezés, eljárás, közlemény.

4. Vitakészség (0-5 pont)

0 pont – nem tud a kérdésekre meggyőzően válaszolni;

2 pont – bizonytalan egyes válaszokban;

4 pont – alapvetően jól érvel, de nem meggyőző;

5 pont – jól érvel, a kérdésekre lényegi választ ad, meggyőző.

A résztvevők névsora

Hallgatók

1. **Abdulaziz Salem Ali Salem**
   Computer Science Engineering BSc, 37. oldal

2. **Abdullaev Bakhtier Ravshanovich**
   Computer Science BSc, 15. oldal

3. **Afroz Samanta**
   Computer Science Engineering BSc, 24. oldal

4. **Aly Ahmed Sayed Ahmed Soliman**
   Computer Science BSc, 23. oldal

5. **Angi Dávid**
   Adattudomány MSc, 34. oldal

6. **Birkas Gergö**
   Adattudomány MSc, 12. oldal

7. **Chebili Anas**
   Computer Science Engineering BSc, 17. oldal

8. **Czakó Orsolya**
   Mesterséges intelligencia informatikus MSc, 21. oldal

9. **Fintor Gellért László**
   Mérnökinformatikus MSc, 14. oldal

10. **Jumaev Shakhriyorbek**
    Business Informatics BSc, 25. oldal

11. **Kovalov Georgiy**
    Computer Science BSc, 16. oldal

12. **Kuki Csaba Sándor**
    Programtervező informatikus BSc, 13. oldal

13. **Kumar Danish**
    Computer Science BSc, 39. oldal

14. **Magura Áron**
    Adattudomány MSc, 22. oldal

15. **Makár Péter Ákos**
    Mérnökinformatikus MSc, 28. oldal

16. **Mierlut Jonathán István**
    Programtervező informatikus BSc, 29. oldal

17. **Pankotai Imre**
    Adattudomány MSc, 27. oldal

18. **Pósz Patricia**
    Gazdaságinformatikus BSc, 20. oldal

19. **Sógor Gergely**
    Adattudomány MSc, 19. oldal

20. **Sumaiya Ashfaque**
    Computer Science BSc, 32. oldal

21. **Száva Gréta**
    Mérnökinformatikus BSc, 31. oldal

22. **Szentmiklóssy Bálint**
    Adattudomány MSc, 35. oldal

23. **Talal Bin Tayyab Muhammad**
    Computer Science BSc, 32. oldal

24. **The Hoang Tung Nguyen**
    Computer Science BSc, 15. oldal

25. **Tóman Péter**
    Adattudomány MSc, 36. oldal

26. **Tóth Bence**
    Mérnökinformatikus BSc, 31. oldal

27. **Vajda Máté Pál**
    Programtervező informatikus MSc, 30. oldal

28. **Yar Asfand**
    Computer Science BSc, 38. oldal

Témavezetők

1. **Dr. Adamkó Attila Tamás**
   egyetemi docens, Információ Technológia Tanszék

2. **Prof. Dr. Baran Sándor**
   egyetemi tanár, Alkalmazott Matematika és Valószínűségszámítás Tanszék

3. **Dr. Biró Piroska**
   egyetemi docens, Információ Technológia Tanszék

4. **Dr. Bogacsovics Gergő**
   adjunktus, Adattudomány és Vizualizáció Tanszék

5. **Dr. Gáll József Mihály**
   egyetemi docens, Alkalmazott Matematika és Valószínűségszámítás Tanszék

6. **Prof. Dr. Gál Zoltán**
   egyetemi tanár, Informatikai Rendszerek és Hálózatok Tanszék

7. **Dr. Gilányi Attila László**
   egyetemi docens, Információ Technológia Tanszék

8. **Dr. Hannush Carolin**
   adjunktus, Számítógéptudományi Tanszék

9. **Dr. Herendi Tamás**
   egyetemi docens, Számítógéptudományi Tanszék

10. **Dr. Kocsis Gergely**
    egyetemi docens, Informatikai Rendszerek és Hálózatok Tanszék

11. **Dr. Kósa Karolina**
    egyetemi tanár, Magatartástudományi Intézet

12. **Dr. Lakatos Róbert**
    adjunktus, Adattudomány és Vizualizáció Tanszék

13. **Mészáros László**
    tanársegéd, Adattudomány és Vizualizáció Tanszék

14. **Dr. Oláh Norbert**
    adjunktus, Adattudomány és Vizualizáció Tanszék

15. **Dr. Pintér-Huszti Andrea**
    egyetemi docens, Adattudomány és Vizualizáció Tanszék

16. **Dr. Sütő József**
    egyetemi docens, Informatikai Rendszerek és Hálózatok Tanszék

17. **Dr. Tomán Henrietta**
    adjunktus, Adattudomány és Vizualizáció Tanszék

18. **Dr. Tóth János**
    adjunktus, Adattudomány és Vizualizáció Tanszék

19. **Dr. Tóth Róbert**
    adjunktus, Információ Technológia Tanszék

20. **Dr. Vágner Anikó Szilvia**
    adjunktus, Információ Technológia Tanszék



Debrecen
2026