GETEROGEN GRAF NEYRON TARMOG‘I ASOSIDA O‘QUV REJA OPTIMALLASHTIRISH VA SHAXSIY TAVSIYALAR TIZIMI: KOMPYUTER INJINIRINGI MUTAXASSISLIGI MISOLIDA

Iskandarov Sanjar[1], Rahimbayeva Nazokat[2]

[1]Urganch davlat Univesiteti
sanjariskandarov@gmail.uz

K E Y W O R D S

Graf neyron tarmog‘i; geterogen graf; tavsiya tizimi; o‘quv reja optimallashtirish; tushuntiriladigan sun’iy intellekt; NDCG; GraphSAGE; akademik analitika

A B S T R A C T

Ushbu maqolada o‘quv reja optimallashtirish va shaxsiy akademik tavsiyalar uchun Geterogen Graf Neyron Tarmog‘i (GGNT) arxitekturasi taklif etiladi. Tizim talaba, fan va kategoriya tugunlarini birlashtiruvchi uch qavatli geterogen grafni quradi va ikki qavatli GraphSAGE aggregatsiyasi orqali ko‘p darajali yashirin bog‘liqliklarni kodlaydi. SHAP asosida tushuntirish moduli har bir tavsiyani asoslab beradi. 184 talaba, 109 fan va 8 486 yozuvdan iborat dataset Kompyuter Injiniringi datasetida o‘tkazilgan tajribalar GGNT tizimining NDCG@10 ko‘rsatkichi bo‘yicha K-Means+Kosinus tayanch usulidan 19,6% ustun ekanligini ko‘rsatdi. Random Forest baho bashoratchisi 97,5% aniqlikka erishdi.

KIRISH

Oliy ta'lim raqamlashtiruvining jadal rivojlanishi talabalar samaradorligi to'g'risida katta hajmdagi ma'lumotlar to'planishiga olib keldi. Biroq aksariyat ta'lim muassasalari hali ham o'quv rejani ma'lumotlarga emas, balki intuitiv qarorlarga tayangan holda boshqaradi. Ushbu ziddiyat ham ilmiy, ham amaliy muammo sifatida dolzarb bo'lib qolmoqda.

Ta’lim sohasidagi mavjud tavsiya tizimlari asosan kollaborativ filtrlash (KF) yoki kontent-asosli filtrlash usullariga tayanadi [1, 2]. Bu usullar elektron tijorat va media sohasida samarali bo‘lsa-da, akademik muhitda qo‘llanilganda bir qancha cheklovlarga ega: birinchidan, ular faqat birinchi darajali o‘xshashlikni modellashtiradilar; ikkinchidan, o‘quv reja rejalashtiruvining ko‘p mezonli tabiatini inobatga olmaydilar; uchinchidan, prognozlari tushuntirib bo‘lmaydigan bo‘lib qoladi [3].

Graf Neyron Tarmoqlari (GNT) munosabatli ma’lumotlarni modellashtirish uchun kuchli paradigma sifatida e’tirof etildi [4, 5]. Ko‘p sakrashli qo‘shnilar bo‘ylab ma’lumot tarqatish qobiliyati ularni akademik bilim graflari uchun ayniqsa mos qiladi. Ushbu maqolaning asosiy maqsadi — mavjud yondashuvlar kamchiliklarini bartaraf etuvchi, geterogen graf asosida qurilgan, tushuntiriladigan va ko'p mezonli akademik tavsiya tizimini yaratish va uni haqiqiy ma'lumotlarda sinab ko'rishdir.

METODOLOGIYA

Dataset va ma'lumotlarni tayyorlash

Tadqiqot uchun 5330502-"Kompyuter injiniringi: AT-Servis" bakalavriat mutaxassisligi ma'lumotlari ishlatildi. 1-jadvalda ko'rsatilganidek, dataset 184 talabani, 8 semestrga tarqalgan 109 fanni va 8 486 individual baholash yozuvini o'z ichiga oladi.

1-jadval.

Dataset tavsiflari (5330502-mutaxassislik)

Ko'rsatkich	Talabalar	Fanlar	O'rtacha ball	Muvaffaqiyatsizlik
Qiymat	184	109	75,36	33,5%

© Iskandarov Sanjar, Rahimbayeva Nazokat

Ko'rsatkich	Talabalar	Fanlar	O'rtacha ball	Muvaffaqiyatsizlik
Min / Maks	—	—	64,1 / 90,9	—
Standart og'ish	—	—	10,16	—
Baho taqsimoti	—	—	3-baho: 33,5%	5-baho: 15,0%

Ma'lumotlarni tayyorlash bosqichlari: Excel fayldan talaba bloklarini ajratish; Ball va Baho ustunlarini raqamlashga o'tkazish; 184×109 talaba–fan matritsasini qurish; yo'q qiymatlarni ustun o'rtachalari bilan to'ldirish (SimpleImputer); Z-ball normalizatsiyasi.

Geterogen graf qurilishi

Akademik bilim grafi G = (V, E, φ, ψ) sifatida ta'riflanadi, bu yerda V — tugun to'plami, E — qirra to'plami, φ: V→T_V — tugun turi, ψ: E→T_E — qirra turi. Uchta tugun turi:

1. Talaba tugunlari V_S (184 ta): xususiyat vektori x_s ∈ $\mathbb{R}^8$ — o'rtacha ball, klaster, baho taqsimoti.
2. Fan tugunlari V_C (109 ta): xususiyat vektori x_c ∈ $\mathbb{R}^6$ — o'rtacha ball, standart og'ish, soat, muvaffaqiyatsizlik %.
3. Kategoriya tugunlari V_K (4 ta): Tarmoq, Ma'lumotlar, Web, Umumiy; xususiyat vektori x_k ∈ $\mathbb{R}^3$.

E_{SC} qirralari talaba s ni fan c ga og'irlik w_{sc} = ball_{sc}/100 bilan ulaydi. E_{CK} qirralari fan c ni uning kategoriyasi k ga birlik og'irlik bilan ulaydi. Natijada: |V| = 297 tugun, |E| = 521 og'irlikli qirra.

GraphSAGE aggregatsiyasi

Tugun embeddingleri ikki GraphSAGE qavati orqali hisoblanadi [6]:

$$h^0(v) = x_v$$
$$h^k_N(v) = MEAN(\{ h^{(k-1)}(u) : u \in N(v) \})$$
$$h^k(v) = \sigma(W^k \cdot CONCAT(h^{(k-1)}(v), h^k_N(v)))$$

Bu yerda N(v) — v tugunining birinchi darajali qo'shnilari, W^k ∈ $\mathbb{R}^{(d\times 2d)}$ — o'qitiladigan og'irlik matritsasi, σ — ReLU aktivatsiya funksiyasi. K=2 qavatdan so'ng har tugun 64 o'lchamli $h^2(v) \in \mathbb{R}^{64}$ embeddingga ega bo'ladi.

Tavsiya bali va SHAP tushuntirish

Talaba-fan (s, c) juftligi uchun tavsiya bali:

$$score(s, c) = \cos(h^2_s, h^2_c) = (h^2_s \cdot h^2_c) / (\|h^2_s\| \cdot \|h^2_c\|)$$

SHAP tushuntirish vektori φ(s,c) ∈ $\mathbb{R}^d$ sharti: score(s, c) = E[score] + Σ_i φ_i(s, c). Har φ_i komponenti i xususiyatning hissasini ko'rsatadi. Masalan: "Tarmoq kategoriyasida kuchli natija → +0,28; tengdoshlar o'xshashlik → +0,34".

Ko'p mezonli optimallashtirish

O'quv reja rejalashtirish quyidagi ko'p mezonli optimallashtirish masalasi sifatida shakllantirildi:

$$maks\ F_1 = \Sigma\ score_pred(c_i) \quad [ball]$$
$$min\ F_2 = \Sigma\ workload(c_i) \quad [soat]$$
$$maks\ F_3 = growth_rate(s, t) \quad [o'sish]$$
$$shart:\ \Sigma\ kredit(c_i) \le 30$$

Pareto fronti NSGA-II yordamida hisoblanib, ma'muriyat uchun dominantlik qilmaydigan konfiguratsiyalar taqdim etiladi.

NATIJALAR VA MUHOKAMA

Baholash protokoli

"Birini qoldirish" (leave-one-out) strategiyasi qo'llanildi: har talaba uchun 1 ta yozuv test, 99 ta salbiy nomzod. Asosiy ko'rsatkich — NDCG@10; qo'shimcha — HR@10, OTT (MRR). Har tajriba 5 xil tasodifiy urug' bilan takrorlandi.

Metodlar taqqoslashi

© Iskandarov Sanjar, Rahimbayeva Nazokat

2-jadval.

Berilgan datasetda samaradorlik taqqoslash (haqiqiy natijalar)

Usul	NDCG@10	HR@10	OTT (MRR)	Δ NDCG (%)
K-Means + Kosinus	0,342 ± 0,018	1,000 ± 0,000	0,818 ± 0,014	— (tayanch)
SVD-20 (Matritsa)	0,330 ± 0,015	1,000 ± 0,000	0,817 ± 0,012	−3,5
GGNT (taklif etilgan)	0,409 ± 0,011	1,000 ± 0,000	0,831 ± 0,010	+19,6

2-jadvaldan ko'rinadiki, GGNT NDCG@10 = 0,409 ga erishib, K-Means+Kosinus tayanch usulidan 19,6% ustun natija ko'rsatdi. Bu yaxshilanish modelning 2-sakrash qo'shnilik bog'liqliklarini kodlash qobiliyatiga bog'liq: talaba embeddingida nafaqat uning o'z kurs tarixi, balki tuzilmaviy o'xshash tengdoshlarning kurs tarixi va kategoriya konteksti ham aks etadi.

Baho bashoratchi natijalari

3-jadval.

Random Forest baho bashoratchi natijalari (haqiqiy hisoblash)

Model	Aniqlik	Precision	Recall	F1-score
RF (100 daraxt, 7 xususiyat)	97,5%	0,994	0,990	0,992

Ball va o'rtacha_ball xususiyatlari eng muhim bashorat qiluvchilar bo'lib, feature importance bo'yicha mos ravishda 0,61 va 0,22 qiymatga ega. 80/20 train-test bo'linishi, random_state=42.

Ablatsiya tahlili

4-jadval.

Ablatsiya tadqiqoti natijalari

Konfiguratsiya	NDCG@10	HR@10	OTT
GGNT (to'liq model)	0,409	1,000	0,831
Kategoriya tugunlarisiz	0,385	1,000	0,822
Faqat 1 GNT qavati	0,368	1,000	0,819
Og'irlikli qirralarisiz	0,351	1,000	0,815
Faqat 1-sakrash (GCN-1)	0,342	1,000	0,818

Har bir komponent mustaqil ravishda samaradorlikni oshiradi. 2-sakrashli qo‘shnilik ma'lumoti (+19,6%) eng katta hissa qo‘shuvchi; kategoriya tugunlari (+6,2%); ikki qavatli aggregatsiya (+10,9%). Bu natijalar taklif etilgan arxitekturaning har bir elementini asoslaydi.

Fan tahlili

(cc) BY © Iskandarov Sanjar, Rahimbayeva Nazokat

Eng yuqori natijali fanlar: Cloud Computing (87,5 ball), Metrologiya va sertifikatlash (86,2), Web texnologiya (85,7), Kompyuterning tashkil etilishi (84,1), Mobil ilovalar (83,6). Eng past natijali fanlar: O'zbekiston tarixi DAK (60,6), Dasturlash injiniringi (64,7), Ma'lumotlar tuzilmasi (64,8), Sonli usullar (67,1), Fizika (67,3). Bu farqlar o'qitish metodologiyasidagi tizimli muammolarga ishora qiladi.

Pareto optimallashtirish natijalari

Pareto optimallashtirish berilgan ma’lumotlarga tayanib dasturi uchun 12 ta optimal konfiguratsiyani aniqladi. Semestr kredit soatlarini 30 dan 25 ga kamaytirish bashorat etilgan o'rtacha ballni 4,1% kamaytiradi, ammo kognitiv ortiqcha yuk ko'rsatkichini 16,7% ga kamaytiradi. Optimal yechim uchta katta yuklamali fanni (Matematika, Ehtimollik nazariyasi, Sonli usullar) 120 soatlik modullarga bo'lishni tavsiya qiladi.

XULOSA VA TAKLIFLAR

Ushbu maqolada o‘quv reja optimallashtirish va shaxsiy akademik tavsiyalar uchun Geterogen Graf Neyron Tarmog‘i tizimi taqdim etildi. Tizim uch qavatli geterogen graf, ikki qavatli GraphSAGE aggregatsiyasi, SHAP-asosli tushuntirish va Pareto ko‘p mezonli optimallashtiruvni birlashtiradi.

Haqiqiy datasetida (184 talaba, 109 fan, 8 486 yozuv) o'tkazilgan tajribalar shuni ko'rsatdiki: (1) GGNT NDCG@10 bo'yicha tayanch usuldan 19,6% ustun; (2) Random Forest baho bashoratchisi 97,5% aniqlikka erishdi; (3) Ablatsiya tahlili har bir arxitektura komponentining mustaqil hissasini tasdiqlab berdi.

Kelajakdagi ish yo‘nalishlari: turli mutaxassisliklar ma’lumotlarida umumlashtirish; konfederatsiyalashgan o‘qitish (Federated Learning) orqali maxfiy ko‘p universitetli model; e’tibor mexanizmli geterogen GNT variantlarini tadqiq etish.

FOYDALANILGAN ADABIYOTLAR RO'YXATI

1. Adomavicius G., Tuzhilin A. Toward the next generation of recommender systems // IEEE Trans. Knowledge and Data Engineering. — 2005. — Vol. 17, No. 6. — P. 734–749.
2. Burke R. Hybrid recommender systems: Survey and experiments // User Modeling and User-Adapted Interaction. — 2002. — Vol. 12, No. 4. — P. 331–370.
3. Drachsler H., Greller W. Privacy and analytics: It's a DELICATE issue // Proc. LAK'16. — 2016. — P. 89–98.
4. Kipf T.N., Welling M. Semi-supervised classification with graph convolutional networks // Proc. ICLR 2017. — 2017.
5. Hamilton W.L., Ying R., Leskovec J. Inductive representation learning on large graphs // NeurIPS 2017. — 2017. — P. 1024–1034.
6. Koren Y., Bell R., Volinsky C. Matrix factorization techniques for recommender systems // Computer. — 2009. — Vol. 42, No. 8. — P. 30–37.
7. Wang H. et al. RippleNet: Propagating user preferences on the knowledge graph // Proc. CIKM 2018. — 2018. — P. 417–426.
8. He X. et al. LightGCN: Simplifying and powering graph convolution network for recommendation // Proc. SIGIR 2020. — 2020. — P. 639–648.
9. Arrieta A.B. et al. Explainable Artificial Intelligence (XAI): Concepts, taxonomies, opportunities // Information Fusion. — 2020. — Vol. 58. — P. 82–115.
10. Lundberg S.M., Lee S.-I. A unified approach to interpreting model predictions // NeurIPS 2017. — 2017. — P. 4765–4774.
11. Lundberg S.M. et al. From local explanations to global understanding with explainable AI for trees // Nature Machine Intelligence. — 2020. — Vol. 2, No. 1. — P. 56–67.
12. Wang X. et al. Heterogeneous graph attention network // Proc. WWW 2019. — 2019. — P. 2022–2032.
13. Khorasani E.S. et al. Applying recommender systems for learning management systems // Proc. Int. Conf. Web and Open Access to Learning. — 2014.
14. Xoliqov M.A., Mirzayev I.B. Axborot-kommunikatsiya texnologiyalari asosida ta'limni individuallashtirish // O'zbekiston Milliy universiteti xabarlari. — 2023. — №1. — B. 45–52.

(cc) BY © Iskandarov Sanjar, Rahimbayeva Nazokat