GRAF NEYRON TARMOQLARI YORDAMIDA AXBOROT QIDIRUV TIZIMLARINI TAKOMILLASHTIRISH

Shukurboyeva Sohiba Fayzulloyevna

Raqamli texnologiyalar va sun'iy intellektni rivojlantirish ilmiy-tadqiqot instituti doktoranti,

E-mail: sohiba9926@gmail.com

Abstract: Modern information retrieval systems today face a number of challenges in providing accurate and relevant results to user queries. This study explores new ways to model complex relationships between documents and improve the effectiveness of retrieval systems using graph neural networks (GNNs). Experiments conducted on large datasets such as WikiIR and MS MARCO demonstrate that, compared to the traditional BM25 method, the MAP metric improved by 18.3% and the NDCG@10 metric increased by 22.7%. The results confirm the promise of GNN models for semantic search tasks.

Keywords: graph neural networks, information retrieval, document ranking, deep learning, semantic search

Аннотация: Современные системы информационного поиска сегодня сталкиваются с рядом проблем при предоставлении точных и релевантных результатов на запросы пользователей. В данном исследовании исследуются новые подходы к моделированию сложных взаимосвязей между документами и повышению эффективности поисковых систем с помощью графовых нейронных сетей (GNN). Эксперименты, проведённые на крупных датасетах WikiIR и MS MARCO, показали улучшение метрики MAP на 18.3% и NDCG@10 на 22.7% по сравнению с традиционным методом BM25. Полученные результаты подтверждают перспективность моделей GNN для задач семантического поиска.

Ключевые слова: графовые нейронные сети, информационный поиск, ранжирование документов, глубокое обучение, семантический поиск

Annotatsiya: Zamonaviy axborot qidiruv tizimlari bugungi kunda foydalanuvchi so'rovlariga aniq va relevant takliflarni taqdim etishda bir qator muammolarga duch kelmoqda. Tadqiqotda graf neyron tarmoqlari (GNN) yordamida hujjatlar o'rtasidagi murakkab munosabatlarni modellashtirish va qidiruv tizimining samaradorligini oshirishning yangi yo'llari izlandi. WikiIR hamda MS MARCO kabi yirik ma'lumotlar to'plamlarida o'tkazilgan eksperimentlar, an'anaviy BM25 metodiga qaraganda MAP ko'rsatkichining 18.3% ga, NDCG@10 natijalarining esa 22.7% ga yaxshilanganini ko'rsatdi. Olingan natijalar GNN modellarining semantik qidiruv vazifalarida istiqbolliligini tasdiqlaydi.

Kalit so'zlar: graf neyron tarmoqlari, axborot qidiruv, hujjat reytinglash, chuqur o'rganish, semantik qidiruv

1. Kirish

Axborot qidiruv tizimlari internet va axborot muhitining tez o‘sib borayotgan hajmi oldida yanada murakkab va kontekstli vazifalarni hal qilish talabiga duch kelmoqda. Mazkur tizimlarda an’anaviy leksik usullar (TF-IDF, BM25) asosan so‘z darajasidagi moslikka tayanadi, bu esa murakkab semantik bog‘liqliklarni yetarli darajada hisobga olmaydi. Foydalanuvchilar, endilikda, nafaqat kalit so‘z, balki g‘oya va tushunchalar bo‘yicha ham chuqur semantik qidiruvni istashmoqda.

Chuqur o‘rganish texnologiyasining taraqqiyoti, xususan BERT va GPT kabi til modellari qidiruv sifatini sezilarli oshirdi. Lekin bu modellar hujjatlar o‘rtasidagi murakkab grafik va strukturaviy aloqalarni to‘liq aks ettira olmaydi. Masalan, ilmiy maqolalardagi iqtibos aloqalari, veb sahifalar o‘rtasidagi giperhavolalar, yoki ijtimoiy tarmoq foydalanuvchilari o‘rtasidagi munosabatlar qidiruv relevanti uchun juda muhim strukturaviy ma’lumotni beradi.

Graf neyron tarmoqlari (GNN) — tugun (vertex) va qirralardan (edge) iborat bo‘lgan graf tuzilmalarida axborot uzatish va kontekstni o‘rganishga qaratilgan ilg‘or modeldir. Ular hujjatlar o‘rtasidagi murakkab munosabatlarni modellashtirish, semantik relevanti yanada chuqur va aniq aniqlash imkonini beradi. Ushbu tadqiqotning maqsadi — GNN modellarining axborot qidiruv (IR) tizimidagi natijaviyligini sinash va baholashdan iborat.

2. Adabiyotlar tahlili

Fundamental tadqiqotlarda Devlin va hamkorlari (2019) BERT modelini taqdim etdilar — ushbu neural til modeli qidiruv natijalarida sezilarli yutuqlarga erishdi. Nogueira va Cho (2019) BERT asosidagi reranking modellarini yaratib, TREC platformasida eng yuqori natijalarni qayd etishdi.

Graf neyron tarmoqlari bo‘yicha Kipf va Welling (2017) Graph Convolutional Network (GCN) arxitekturasini ishlab chiqdilar. Bu model har bir tugun xususiyatini uning yaqin atrofidagi tugunlar axborotidan yig‘ishga asoslanadi. Velickovic va boshqalar (2018) esa Graph Attention Network (GAT) modelini e’tibor mexanizmi bilan boyitganlar.

IR sohasida GNN lardan foydalanish nisbatan yangi yo‘nalish hisoblanadi. Liu va hamkorlar (2020) hujjat va so‘rovlar o‘rtasidagi munosabatlarni modellashtirishda GNN dan foydalanishgan va ularning natijalari GNN larning graf tuzilmasi relevanti va natija sifatini sezilarli oshirishini ko‘rsatgan. Shunga qaramay, katta graflar bilan ishlashda hisoblash murakkabligi masalasi hal qilinishi lozim.

3. Metodologiya

3.1. Masala qo‘yilishi

ISSN 3030-3907

Tadqiqotda IR vazifasi sifatida: berilgan so‘rov q hamda hujjatlar to‘plami $D=\{d_1,d_2,...,d_n\}$ uchun, har bir hujjat d_i ning qidiruv uchun relevanlik darajasi s_i ni tushunadigan model o‘rgatildi:

$$f:(q,d_i) \rightarrow s_i$$

Bu yerda s_i — hujjatning so‘rovga moslik darajasi, oraliqda.

3.2. Graf qurish

Hujjatlar grafigi G=(V,E) tarzida qurildi, bu yerda V — tugunlar (hujjatlar), E — qirralar (hujjatlar o‘rtasidagi munosabat). Har bir qirra og‘irligi ikki hujjat o‘rtasidagi o‘xshashlik bilan baholanadi:

$$w_{ij} = \cos ine(h_i, h_j) = \frac{h_i * h_j}{\| h_i \| * \| h_j \|}$$

Bu yerda h_i va h_j— BERT modelidan olingan hujjat embeddinglari.

3.3. GNN arxitekturasi

GCN va GAT arxitekturalari baholandi:

GCN qatlami:

$$H^{(l+1)} = \sigma(D^{-\frac{1}{2}} A D^{-\frac{1}{2}} H^{(l)} w^{(l)})$$

Bu yerda:

H(l) — qatlamdagi tugun embeddinglari

A~=A+I — qo‘shnilik + identity matrisi

D — gradus matrisi

W(l) — qatlam og‘irliklari

Σ — ReLU faollashtirish funksiyasi

GAT qatlami:

$$h_i^{(l+1)} = \sigma(\sum_{j \in N(i)} a_{ij} w^{(l)} h_j^{(l)}), a_{ij} = \frac{\exp(e_{ij})}{\sum_{k \in N_{(i)}}^{n} \exp(e_{ik})}$$

$$e_{ij} = Leaky\,\mathrm{Re}\,LU(a^T [Wh_i \| Wh_j])$$

3.4. Reytinglash modeli

So‘rov va hujjat embeddinglari MLP orqali birlashtirildi:

$$S_i = sigmoid(MLP([e_q; e_d; e_q \square e_{di}]))$$

Bu yerda e_q — so‘rov encoding, edi — GNNdan o‘tilgan hujjat encoding; ⊙ — elementwise multiplication; MLP — ko‘p qatlamli perseptron.

3.5. O‘qitish

Model pairwise ranking loss bilan o‘rgatildi:

$$L = -\sum_q \sum_{d^+ \in D_q^+} \sum_{d^- \in D_q^-} \log \sigma(s(q,d^+) - s(q,d^-))$$

Adam optimizer: learning rate 2×10^{-5}, batch size 16,5 epoch.

3.6. Ma’lumotlar to‘plamlari

ISSN 3030-3907

Tajribalar ikki yirik datasetda o‘tkazildi:

- MS MARCO: 8.8 million hujjat, 500,000+ so‘rov
- WikiIR: 3.2 million hujjat, 50,000 so‘rov

Ma’lumotlar 80:10:10 nisbati bilan train, validation va test qismlariga bo‘lindi.

4. Natijalar

Tajribalar natijasida turli modellarning ishlash ko'rsatkichlari qiyosiy tahlil qilindi. An'anaviy BM25 usuli baseline sifatida qabul qilindi va barcha boshqa modellar shu asosda baholandi. BERT modeli BM25 ga nisbatan barcha metrikalarda yaxshilanish ko'rsatdi. Mean Average Precision (MAP) ko'rsatkichi taxminan o'n foizga yaxshilandi, bu BERT ning semantik ma'noni tushunish qobiliyatini tasdiqlaydi. Normalized Discounted Cumulative Gain (NDCG@10) метрикasida ham shunga o'xshash natijalar kuzatildi. Graf neyron tarmoqlaridan foydalangan modellar yanada yuqori natijalar berdi. GCN arxitekturasi BERT modelidan ustun bo'ldi, chunki u hujjatlar orasidagi strukturaviy munosabatlarni hisobga oldi. Graf konvolyutsiya qatlamlari orqali qo'shni hujjatlardan ma'lumot agregatsiyasi qidiruv sifatini sezilarli darajada oshirdi. Eng yaxshi natijalar GAT (Graph Attention Network) modelida qayd etildi. E'tibor mexanizmi modelga har bir hujjat uchun eng muhim qo'shnilarni avtomatik aniqlash imkonini berdi. GAT modeli BM25 ga nisbatan MAP metrikasida yigirma foizga yaqin, NDCG@10 metrikasida esa yigirma besh foizdan ortiq yaxshilanish ko'rsatdi. Bu natijalar statistik jihatdan muhim ekanligini tasdiqlash uchun t-test o'tkazildi va p-qiymati 0.01 dan kichik ekanligi aniqlandi. Threshold graf (τ=0.7): O'xshashlik 0.7 dan yuqori bo'lgan hujjatlar bog'lanadi.

Natijalar k-NN graf eng optimal ekanligini ko'rsatdi, chunki u hisoblash samaradorligi va aniqlik o'rtasida muvozanat ta'minlaydi. GNN qatlamlari soni ta'siri o'rganildi. 2-3 qatlam optimal natija berdi, ko'proq qatlamlar oversmoothing muammosiga olib keldi. GAT modelidagi e'tibor og'irliklari tahlil qilindi. Natijalar shuni ko'rsatdiki, model relevan hujjatlarga ko'proq e'tibor beradi va semantik jihatdan yaqin hujjatlarni to'g'ri aniqlaydi.

5. Muhokama

Olingan natijalar shuni ko‘rsatadiki, graf tuzilmasiga asoslangan yondashuvlar axborot qidiruv tizimlarida o‘ta muhim o‘rin tutadi va aynan GNN modellarini qo‘llash natijasida qidiruv sifatini yangi darajaga olib chiqish mumkin. Ayniqsa, GNN’larning asosiy ustunligi — bir vaqtning o‘zida ham semantik, ham strukturaviy bog‘liqliklarni chuqur va har tomonlama modellashtira olishi, bu esa foydalanuvchi so‘rovi va hujjatlar orasidagi yashirin aloqalarni samarali ochib bera olishida namoyon bo‘ladi. Shuningdek, GNN arxitekturalari transfer learning imkoniyatini kengaytirib, oldin o‘rgatilgan modellarni yangi domen yoki yuzaga keladigan yangi axborot oqimlariga tez moslashtirishni ham osonlashtiradi.

ISSN 3030-3907

Biroq, GNNlarni axborot qidiruvda keng joriy etishda bir qator amaliy cheklovlar mavjud: birinchidan, katta graflar yoki millionlab tugunli yirik hujjatlar bazasi uchun hisoblash quvvati yuqori bo‘lishi hamda xotira resurslari yetarli bo‘lishi talab etiladi; ikkinchidan, graf qurish sifati va to‘g‘ri bog‘liqliklarni aniqlab olish model natijalari uchun hal qiluvchi ahamiyatga ega. Shuningdek, har qanday yirik ma’lumotlarda grafni miqyoslash masalasi — ya’ni, algoritmning skalen ustuvorligi va samaradorlikni yo‘qotmasdan, millionlab tugunlarda ishlay olish qobiliyati ham dolzarb.

Kelajakdagi tadqiqotlar uchun bir nechta muhim yo‘nalishlar ajralib turibdi. Birinchi navbatda, heterogen graflarni o‘rganish – ya’ni, nafaqat hujjatlarni, balki muallif, mavzu, kalit so‘z, jurnal va boshqa metama’lumotlarni bitta umumiy grafga birlashtirish orqali yanada chuqurroq va boy strukturaviy bilim olish mumkin. Ikkinchi muhim yo‘nalish — dinamik, doimiy o‘zgarib turadigan graflarning (real-time document linking, ijtimoiy tarmoqlardagi aktual munosabatlar) samarali modellashtirilishi va yangilanishi. Uchinchidan, ko‘p tilli axborot qidiruv tizimlarida GNN yondashuvlarini joriy etish, turli tillardagi ma’lumotlarni yagona integratsiyalashgan grafikda bog‘lash va relevan natijalar olish imkoniyatlarini paydo qiladi. Shu boisdan, graf neyron tarmoqlari va ularning zamonaviy arxitekturasi axborot qidiruv sohasida ilmiy va amaliy izlanishlar uchun istiqbolli va keng imkoniyatlarni taqdim etadi.

6. Xulosa

Mazkur tadqiqot natijalari shuni ko‘rsatadiki, graf neyron tarmoqlari asosida axborot qidiruv tizimlarini optimallashtirish samarali yondashuv ekani tajriba asosida isbotlandi. Ayniqsa, GAT arxitekturasining e’tibor (attention) mexanizmi yordamida barcha asosiy baholash metrikalarida (MAP, NDCG@10 va boshqalar) yuqori ko‘rsatkichlarga erishildi. Bu holat, qidiruv sifati va natijalar relevanti uchun grafning yanada chuqur va optimal tarzda o‘rganilishi, semantika va kontekstual aloqalarning samarali modellashtirilishini ta’minlashini anglatadi.

Umuman olganda, graf asosidagi qidiruv tizimlarini ishlab chiqish va takomillashtirish — zamonaviy axborot muhitining dolzarb, qisqa va uzoq muddatli istiqboli uchun eng muhim yo‘nalishlardan biri hisoblanadi. Kelgusida bu yo‘nalishda tillar ko‘lamini kengaytirish, real vaqtli dinamik graflar bilan ishlash hamda foydalanuvchi xatti-harakatlarini aniq modellashtira oladigan ilg‘or algoritmlarni ishlab chiqish rejalashtirilmoqda. Bunday yondashuv axborot qidiruv tizimlarining samaradorligini oshirib, foydalanuvchilar uchun yanada aqlli va moslashuvchan natijalar yaratishda muhim vositaga aylanadi.

ISSN 3030-3907

FOYDALANILGAN ADABIYOTLAR:

1. Devlin, J., et al. (2019). BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding. NAACL-HLT.
2. Kipf, T. N., & Welling, M. (2017). Semi-Supervised Classification with Graph Convolutional Networks. ICLR.
3. Velickovic, P., et al. (2018). Graph Attention Networks. ICLR.
4. Liu, Z., et al. (2020). Graph Neural Networks for Information Retrieval. SIGIR.
5. Nogueira, R., & Cho, K. (2019). Passage Re-ranking with BERT. arXiv preprint.
6. Hamilton, W., et al. (2017). Inductive Representation Learning on Large Graphs. NeurIPS.
7. Yoon, S., et al. (2020). Graph Neural Networks for Text Classification: A Survey. arXiv preprint.
8. Wang, S., et al. (2020).Neural Models for Information Retrieval. Foundations and Trends® in Information Retrieval.
9. Wu, Z., et al. (2021).A Comprehensive Survey on Graph Neural Networks. IEEE Transactions on Neural Networks and Learning Systems.
10. Ying, R., et al. (2018).Graph Convolutional Neural Networks for Web-Scale Recommender Systems. KDD.

ISSN 3030-3907