SO‘ROV MATNIGA MUVOFIQ AXBOROTLARNI IZLASHDA ZAMONAVIY ALGORITMIK YONDASHUVLAR

Bozorov Rustambek Bozar o'g'li
Uzun tumani 56-maktab Informatika fani o'qituvchisi
E-mail: rustambek.bozorov.96@gmail.com

Qabul qilindi: 16.01.2026
Ko‘rib chiqildi: 17.01.2026
Tasdiqlandi: 19.01.2026

Kalit so‘zlar:
axborotlarni izlash, so‘rov matni, algoritmik yondashuvlar, tabiiy tilni qayta ishlash, semantik tahlil, mashinaviy o‘rganish, chuqur o‘rganish, axborot qidiruv tizimlari.

Annotatsiya. Mazkur tezisda so‘rov matniga muvofiq axborotlarni izlash masalasi va ushbu jarayonda qo‘llaniladigan zamonaviy algoritmik yondashuvlar tahlil qilinadi. Tadqiqot davomida an’anaviy axborot qidiruv modellarining imkoniyatlari hamda semantik va chuqur o‘rganishga asoslangan usullarning afzalliklari solishtirildi. So‘rov matnining mazmunini aniqlash, kontekstni hisobga olish va qidiruv natijalarining relevatligini oshirishda intellektual algoritmlarning o‘rni yoritib berildi. Olingan natijalar zamonaviy axborot qidiruv tizimlarini loyihalash va takomillashtirishda gibrid algoritmik yondashuvlardan foydalanish maqsadga muvofiq ekanini ko‘rsatadi. Tadqiqot xulosalari axborot qidiruv tizimlarining samaradorligini oshirishga qaratilgan ilmiy va amaliy ishlar uchun muhim ahamiyatga ega.

Kirish. Axborot texnologiyalarining jadal rivojlanishi va raqamli ma’lumotlar hajmining keskin ortishi zamonaviy jamiyatda axborotni samarali izlash va qayta ishlash muammosini dolzarb masalalardan biriga aylantirdi. Bugungi kunda foydalanuvchilar turli manbalarda – ilmiy maqolalar, elektron kutubxonalar, ijtimoiy tarmoqlar, veb-saytlar va ma’lumotlar bazalarida saqlanayotgan ulkan hajmdagi axborotlar orasidan aynan o‘z so‘roviga mos, aniq va ishonchli ma’lumotlarni tezkor topishga intilmoqdalar. Shu sababli so‘rov matniga muvofiq axborotlarni izlash masalasi zamonaviy axborot qidiruv tizimlarining asosiy vazifalaridan biri hisoblanadi. An’anaviy axborot qidiruv usullari asosan kalit so‘zlarga asoslangan bo‘lib, ular matnning faqat yuzaki mosligini aniqlashga xizmat qiladi. Bunday yondashuvlar ko‘p hollarda so‘rov matnining mazmuniy (semantik) mohiyatini to‘liq ochib bera olmaydi va natijada foydalanuvchiga kerakli axborotni aniqlik bilan taqdim etishda cheklovlarga duch keladi. Ayniqsa, tabiiy tillarning murakkabligi, sinonimlar, ko‘p ma’nolilik va kontekstga bog‘liqlik kabi omillar ushbu muammoni yanada murakkablashtiradi.

So‘nggi yillarda axborotlarni izlash sohasida zamonaviy algoritmik yondashuvlar keng qo‘llanila boshlandi. Ular orasida mashinaviy o‘rganish, chuqur o‘rganish (deep learning), tabiiy tilni qayta ishlash (Natural Language Processing – NLP), vektorli modellar, semantik tarmoqlar va neyron tarmoqlarga asoslangan usullar alohida o‘rin egallaydi. Ushbu yondashuvlar so‘rov matnini faqat so‘zlar to‘plami sifatida emas, balki mazmunli birlik sifatida tahlil qilish imkonini beradi hamda foydalanuvchi niyatini aniqlashda yuqori samaradorlikni ta’minlaydi. Zamonaviy algoritmlar yordamida so‘rov va hujjatlar o‘rtasidagi semantik o‘xshashlikni aniqlash, kontekstni hisobga olish va natijalarni relevatlik darajasiga ko‘ra saralash mumkin bo‘lmoqda. Bu esa axborot qidiruv tizimlarining aniqligi va ishonchliligini sezilarli darajada oshiradi. Shu nuqtayi nazardan, so‘rov matniga muvofiq axborotlarni izlashda zamonaviy algoritmik yondashuvlarni o‘rganish va tahlil qilish ilmiy va amaliy jihatdan muhim ahamiyat kasb etadi.

Adabiyotlar tahlili. Axborotlarni izlash (Information Retrieval – IR) sohasida olib borilgan ilmiy tadqiqotlar bir necha o‘n yilliklarni qamrab olib, ushbu yo‘nalish axborot texnologiyalari, sun’iy intellekt va kompyuter lingvistikasi bilan chambarchas bog‘liq holda rivojlanib kelmoqda. Dastlabki axborot qidiruv tizimlari hujjat va so‘rovlar o‘rtasidagi moslikni aniqlashda oddiy matematik va statistik modellarga tayangan. Ushbu yondashuvlar orasida eng keng tarqalganlari Boolean modeli, vektorli fazo modeli va ehtimollik (probabilistik) modellardir. Boolean modeli asosida ishlovchi tizimlarda hujjatlar va so‘rovlar kalit so‘zlar orqali “VA”, “YOKI”, “EMAS” mantiqiy operatorlari yordamida ifodalanadi. Mazkur modelning soddaligi va hisoblash nuqtayi nazaridan qulayligi uning asosiy afzalligi bo‘lsa-da, u natijalarni relevatlik darajasiga ko‘ra tartiblash imkonini bermaydi. Shu sababli foydalanuvchiga taqdim etilgan natijalar orasida haqiqatan muhim bo‘lgan axborotni aniqlash qiyinlashadi.

Vektorli fazo modeli esa hujjatlar va so‘rovlarni ko‘p o‘lchamli fazodagi vektorlar sifatida tasvirlaydi. Salton tomonidan taklif etilgan ushbu model TF-IDF (Term Frequency – Inverse Document Frequency) vaznlash mexanizmi orqali so‘zlarning ahamiyat darajasini aniqlashga imkon beradi. Tadqiqotchilar ushbu model yordamida hujjat va so‘rov o‘rtasidagi kosinus o‘xshashlikni hisoblash orqali nisbatan aniqroq qidiruv natijalariga erishilganini ta’kidlashadi. Biroq, bu yondashuv ham so‘zlarning semantik ma’nosini to‘liq hisobga olmaydi. Ehtimollik modellari doirasida Robertson va Sparck Jones tomonidan ishlab chiqilgan BM25 algoritmi axborot qidiruv tizimlarida uzoq vaqt davomida samarali yechim sifatida qo‘llanib kelmoqda. BM25 hujjat uzunligi, so‘z chastotasi va teskari hujjat chastotasini hisobga olgan holda relevatlikni baholaydi. Ko‘plab tadqiqotlarda ushbu algoritm an’anaviy statistik usullar orasida eng samaralilaridan biri sifatida e’tirof etilgan. So‘nggi yillarda tabiiy tilni qayta ishlashga asoslangan semantik yondashuvlar axborotlarni izlash

sohasida muhim burilish yasadi. NLP usullari yordamida matnlarni morfologik, sintaktik va semantik tahlil qilish, so‘zlar o‘rtasidagi bog‘lanishlarni aniqlash va kontekstni hisobga olish imkoniyati paydo bo‘ldi. LSA (Latent Semantic Analysis) va LDA (Latent Dirichlet Allocation) kabi yashirin semantik modellar hujjatlar mazmunini mavzular asosida tahlil qilishga xizmat qiladi.

Mashinaviy o‘rganishga asoslangan yondashuvlar axborot qidiruv tizimlarining moslashuvchanligini sezilarli darajada oshirdi. “Learning to Rank” algoritmlari foydalanuvchi xatti-harakatlari va tarixiy ma’lumotlar asosida qidiruv natijalarini tartiblashni o‘rganadi. Ushbu usullar relevatlikni an’anaviy qoidalarga emas, balki real foydalanuvchi baholariga tayangan holda aniqlaydi. Chuqur o‘rganish texnologiyalarining rivojlanishi bilan neyron tarmoqlarga asoslangan modellar, xususan Word2Vec, GloVe, FastText kabi so‘z vektorlashtirish usullari keng qo‘llanila boshlandi. Ushbu modellar so‘zlar o‘rtasidagi semantik o‘xshashlikni raqamli vektorlar orqali ifodalab, sinonim va kontekstga yaqin tushunchalarni aniqlash imkonini beradi. Keyingi bosqichda Transformer arxitekturasiga asoslangan BERT, GPT va shunga o‘xshash modellar so‘rov matnini chuqur kontekstual tahlil qilishda yuqori natijalarni ko‘rsatdi.

Tadqiqot muhokamasi. Mazkur tadqiqot doirasida so‘rov matniga muvofiq axborotlarni izlashda qo‘llaniladigan an’anaviy va zamonaviy algoritmik yondashuvlarning samaradorligi tahlil qilindi. O‘rganish jarayonida kalit so‘zlarga asoslangan modellarning imkoniyatlari bilan semantik va intellektual yondashuvlarning afzalliklari solishtirildi. Natijalar shuni ko‘rsatdiki, zamonaviy algoritmlar foydalanuvchi so‘rovining mazmunini chuqurroq anglash orqali axborot qidiruv jarayonini sezilarli darajada takomillashtiradi. Tadqiqot natijalariga ko‘ra, an’anaviy vektorli va ehtimollik modellar, jumladan TF-IDF va BM25 algoritmlari strukturalashgan va nisbatan kichik hajmdagi ma’lumotlar bazalarida barqaror natijalar beradi. Biroq so‘rov matni murakkablashgan, ko‘p ma’noli yoki kontekstga boy bo‘lgan holatlarda ushbu yondashuvlarning samaradorligi pasayishi kuzatildi. Bu esa kalit so‘zlarga qattiq bog‘liqlik va semantik ma’lumotlarning yetarli darajada hisobga olinmasligi bilan izohlanadi.

Semantik tahlilga asoslangan algoritmlar esa so‘rov va hujjatlar o‘rtasidagi mazmuniy bog‘liqlikni aniqlashda yuqori aniqlikni namoyon etdi. Xususan, so‘z vektorlashtirish usullari yordamida sinonim va ma’nodosh tushunchalarni aniqlash imkoniyati kengaydi. Tadqiqot davomida Word2Vec va FastText modellaridan foydalanilgan holatlarda qidiruv natijalarining relevatlik darajasi an’anaviy modellarga nisbatan yuqoriroq bo‘lgani aniqlandi. Bu holat foydalanuvchi so‘rovining aniq ifodalanmagan yoki umumiy shaklda berilgan vaziyatlarida ayniqsa yaqqol namoyon bo‘ldi. Chuqur o‘rganishga asoslangan kontekstual modellar, jumladan Transformer arxitekturasiga tayangan algoritmlar, so‘rov matnining kontekstini to‘liqroq tahlil qilish imkonini berdi. Tadqiqot natijalari BERT kabi modellar murakkab grammatik tuzilmalar,

bog‘lovchi so‘zlar va kontekstga bog‘liq ma’nolarni aniqlashda yuqori samaradorlikka ega ekanini ko‘rsatdi. Shu bilan birga, ushbu modellar qidiruv natijalarini foydalanuvchi niyatiga yaqinlashtirishda muhim rol o‘ynaydi.

Muhokama jarayonida zamonaviy algoritmik yondashuvlarning ayrim cheklovlari ham aniqlanib, ular tahlil qilindi. Xususan, chuqur o‘rganishga asoslangan modellarni o‘qitish uchun katta hajmdagi belgilangan (labelled) ma’lumotlar talab etiladi. Bundan tashqari, hisoblash resurslariga bo‘lgan yuqori ehtiyoj va tizimning murakkabligi ularni real vaqt rejimida qo‘llashni qiyinlashtirishi mumkin. Shu sababli bunday yondashuvlarni amaliyotga tatbiq etishda optimallashtirish va gibrid modellarni ishlab chiqish zarurati yuzaga keladi. Tadqiqot natijalari shuni ko‘rsatadiki, eng maqbul yechim an’anaviy va zamonaviy algoritmlarni integratsiyalash asosida qurilgan gibrid axborot qidiruv tizimlarini yaratishdan iborat. Masalan, dastlabki bosqichda BM25 kabi tezkor algoritmlar yordamida hujjatlar to‘plami qisqartirilib, keyingi bosqichda chuqur semantik modellar orqali aniq saralash amalga oshirilishi mumkin. Bunday yondashuv samaradorlik va hisoblash tezligi o‘rtasidagi muvozanatni ta’minlaydi. Umuman olganda, tadqiqot muhokamasi zamonaviy algoritmik yondashuvlar so‘rov matniga muvofiq axborotlarni izlashda yuqori aniqlik va moslashuvchanlikni ta’minlashini ko‘rsatdi. Ushbu yondashuvlar foydalanuvchi ehtiyojlariga mos, intellektual va samarali axborot qidiruv tizimlarini yaratish uchun mustahkam ilmiy asos bo‘lib xizmat qiladi. Ilmiy adabiyotlarda qayd etilishicha, zamonaviy algoritmik yondashuvlar axborotlarni izlashda aniqlik, to‘liqlik va tezkorlikni oshirishga xizmat qilmoqda. Shu bilan birga, hisoblash resurslariga bo‘lgan talabning ortishi, modelni o‘qitish murakkabligi va ma’lumotlar sifati bilan bog‘liq muammolar ham dolzarbligicha qolmoqda. Shunday bo‘lsa-da, olib borilgan tadqiqotlar shuni ko‘rsatadiki, semantik va intellektual yondashuvlar asosida qurilgan qidiruv tizimlari kelajakda axborotlarni izlash samaradorligini yanada oshirish uchun mustahkam ilmiy asos bo‘lib xizmat qiladi.

Xulosa. Mazkur tezisda so‘rov matniga muvofiq axborotlarni izlashda qo‘llaniladigan an’anaviy va zamonaviy algoritmik yondashuvlar tahlil qilindi. Tadqiqot natijalari shuni ko‘rsatdiki, kalit so‘zlarga asoslangan an’anaviy modellar ma’lum sharoitlarda samarali bo‘lsa-da, murakkab va kontekstga boy so‘rovlarni qayta ishlashda cheklangan imkoniyatlarga ega. Zamonaviy semantik va chuqur o‘rganishga asoslangan algoritmlar esa so‘rov matnining mazmunini chuqur tahlil qilish orqali axborot qidiruv jarayonining aniqligi va relevatligini sezilarli darajada oshiradi. Shu bilan birga, hisoblash resurslariga bo‘lgan yuqori talab va model murakkabligi ushbu yondashuvlarning asosiy muammolaridan biri sifatida namoyon bo‘ladi. Tadqiqot natijalari asosida gibrid algoritmik yondashuvlardan foydalanish axborot qidiruv tizimlarining samaradorligini oshirish uchun eng maqbul yechim ekanligi xulosaga kelindi.

Foydalanilgan adabiyotlar

1. Manning C.D., Raghavan P., Schütze H. Introduction to Information Retrieval. — Cambridge: Cambridge University Press, 2008.
2. Baeza-Yates R., Ribeiro-Neto B. Modern Information Retrieval: The Concepts and Technology behind Search. — 2nd ed. — Addison-Wesley, 2011.
3. Robertson S., Zaragoza H. The Probabilistic Relevance Framework: BM25 and Beyond. — Foundations and Trends in Information Retrieval, 2009.
4. Jurafsky D., Martin J.H. Speech and Language Processing. — 3rd ed. — zearson, 2023.
5. Mikolov T., Sutskever I., Chen K., Corrado G., Dean J. Distributed Representations of Words and Phrases and their Compositionality. — Advances in Neural Information Processing Systems, 2013.
6. Devlin J., Chang M.-W., Lee K., Toutanova K. BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding. — NAACL-HLT, 2019.