

SAVOL GENERATSIYASI UCHUN LINGVISTIK MODEL

Nasriddinova Iroda Sayfiddin qizi
Kompyuter lingvistikasi mutaxassisligi
1-bosqich magistranti
irodanasriddinova3@gmail.com
ToshDO'TAU

Annotatsiya. Mazkur tezisda savol generatsiyasi jarayoni lingvistik model asosida tahlil qilinadi. Tadqiqotda savol hosil qilish jarayoni matndan axborotni ajratish, uni lingvistik jihatdan qayta ishlash va savol shakliga o'tkazish bosqichlaridan iborat tizimli jarayon sifatida ko'rib chiqiladi. Taklif etilgan model morfologik, sintaktik va semantik tahlil darajalarini o'z ichiga olgan holda savol generatsiyasini bosqichma-bosqich tashkil etishga qaratilgan. Tezisda savol turlari, matn tuzilishining o'rni hamda lingvistik omillarning savol shakllanishiga ta'siri yoritilgan. Shuningdek, lingvistik model zamonaviy neyron yondashuvlar bilan solishtirilib, uning afzalliklari va cheklovlari ko'rsatib berilgan. Tadqiqot natijalari savol generatsiyasida lingvistik yondashuvlarning muhimligini asoslab, kelgusida gibrid modellardan foydalanish istiqbolli ekanini ko'rsatadi.

Kalit so'zlar: *savol generatsiyasi, lingvistik model, matn tuzilishi, sintaktik tahlil, morfologik tahlil, semantik tahlil, tabiiy tilni qayta ishlash, NLP, sun'iy intellect.*

Abstract. This paper analyzes the process of question generation based on a linguistic model. Question generation is considered as a structured process that involves identifying information from text, processing it linguistically, and transforming it into question form. The proposed model organizes this process into several stages, including morphological, syntactic, and semantic analysis. The study examines different types of questions, the role of text structure, and the influence of linguistic factors on question formation. In addition, the linguistic model is compared with modern neural approaches, highlighting its advantages and limitations. The findings demonstrate the importance of linguistic-based methods in

question generation and suggest that hybrid approaches combining linguistic and neural models may provide more effective results in future research.

Keywords: *question generation, linguistic model, text structure, syntactic analysis, morphological analysis, semantic analysis, natural language processing, NLP, artificial intelligence.*

Savol generatsiyasi (Question Generation, QG) matndan axborotni aniqlash va uni savol shaklida ifodalash bilan bog'liq murakkab lingvistik jarayon bo'lib, tabiiy tilni qayta ishlash (Natural Language Processing, NLP) sohasining muhim yo'nalishlaridan biri hisoblanadi [1; 2]. Ushbu yo'nalish nafaqat nazariy, balki amaliy jihatdan ham keng qo'llanilib, ayniqsa ta'lim tizimida avtomatik testlar yaratish, bilimlarni baholash, o'quv materiallarini qayta ishlashda muhim ahamiyat kasb etadi. Shu bilan birga, savol generatsiyasi sun'iy intellekt tizimlarida, jumladan chatbotlar, virtual assistentlar va interaktiv o'quv platformalarida ham keng qo'llanilmoqda.

Amaliy kuzatishlar shuni ko'rsatadiki, savol hosil qilish jarayonini oddiy transformatsiya sifatida talqin qilish yetarli emas. Ya'ni, gapdagi ma'lum bir so'zni savolga aylantirish orqali hosil qilingan savollar har doim ham mazmunli va tabiiy bo'lavermaydi. Buning sababi shundaki, savol matndan ajratib olingan alohida birlikka emas, balki butun matnning tuzilishiga va undagi bog'lanishlarga tayanadi. Shu sababli savol generatsiyasida matnning sintaktik va semantik tuzilishini chuqur tahlil qilish zarur hisoblanadi.

Savol generatsiyasi jarayonini to'g'ri tushunish uchun avvalo savol turlarini aniqlab olish muhimdir. Umuman olganda, savollar yopiq va ochiq savollarga ajratiladi. Yopiq savollar “ha/yo‘q” tipidagi javobni talab qilsa, ochiq savollar muayyan axborotni aniqlashga qaratilgan bo'lib, “kim”, “nima”, “qachon”, “qayerda”, “nima sababdan” kabi savol birliklari orqali ifodalanadi. Aynan ochiq

savollar savol generatsiyasi tizimlarida murakkabroq hisoblanadi, chunki ular matndagi aniq axborotga yo'naltirilgan bo'ladi.

Matndan savol hosil qilish jarayoni odatda bir necha bosqichda amalga oshiriladi. Birinchi bosqichda matndan asosiy axborot ajratib olinadi. Bu jarayon axborot markazini aniqlash bilan bog'liq bo'lib, gapdagi eng muhim komponentlarni topishni talab qiladi. Ikkinchi bosqichda ushbu komponentlar lingvistik jihatdan tahlil qilinadi. Uchinchi bosqichda esa aniqlangan axborot savol shakliga o'tkaziladi [3]. Ushbu bosqichlarning har biri savolning yakuniy sifati va tabiiyligiga bevosita ta'sir ko'rsatadi.

Mazkur ishda savol generatsiyasi uchun lingvistik model taklif etiladi. Ushbu model matnni qayta ishlashni uch asosiy darajada – morfologik, sintaktik va semantik darajalarda tashkil etadi. Morfologik darajada so'zlarning grammatik shakllari, jumladan kelishik, egalik va boshqa qo'shimchalar aniqlanadi. Sintaktik darajada gap tuzilishi, ya'ni ega, kesim va boshqa bo'laklar o'rtasidagi munosabatlar tahlil qilinadi. Semantik darajada esa matndagi umumiy mazmun, sabab-oqibat va boshqa mantiqiy bog'lanishlar aniqlanadi [4].

Taklif etilayotgan lingvistik modelni quyidagi bosqichlar orqali ifodalash mumkin:

1. Matnni segmentatsiya qilish va gaplarga ajratish
2. Morfologik tahlil
3. Sintaktik tahlil
4. Semantik bog'lanishlarni aniqlash
5. Axborot markazini tanlash
6. Savol strukturasini shakllantirish

Ushbu model savol generatsiyasini bosqichma-bosqich va izchil tarzda tashkil etish imkonini beradi. Ayniqsa, har bir bosqich aniq ajratilgani sababli model tushunarli va nazorat qilinadigan tizim sifatida qaraladi.

Modelning ishlashini quyidagi oddiy misol orqali ko'rsatish mumkin:

Matn: “Toshkent O'zbekiston poytaxti.”

Tahlil: “Toshkent” – ega, “poytaxti” – kesimga bog'langan atribut

Savol: “O'zbekiston poytaxti qaysi shahar?”

Bu misolda savol generatsiyasi matn tuzilishiga asoslangan holda amalga oshirilgani ko'rinadi. Savol hosil qilish jarayonida ega va kesim o'rtasidagi munosabat asosiy rol o'ynaydi.

Zamonaviy savol generatsiyasi tizimlarida neyron tarmoqlarga asoslangan yondashuvlar keng qo'llanilmoqda. Masalan, GPT tipidagi modellar katta hajmdagi ma'lumotlar asosida savollarni avtomatik hosil qiladi [5]. Bunday yondashuvlar yuqori moslashuvchanlikka ega bo'lib, murakkab matnlar bilan ishlash imkonini beradi. Biroq ularning asosiy kamchiligi shundaki, savol qanday hosil qilingani aniq tushuntirib berilmaydi. Lingvistik modellarning ustun tomoni esa ularning izohlanadigan (interpretable) tizim ekanidadir. Har bir bosqich aniq ko'rsatiladi va savol shakllanishi mantiqan tushuntiriladi. Shu bilan birga, bunday modellarning kamchiligi – murakkab va kontekstga boy matnlarda moslashuvchanlikning pastligi hisoblanadi.

Savol generatsiyasi jarayonida bir qator muammolar ham mavjud. Eng avvalo, matn kontekstini to'liq tushunish murakkab vazifa hisoblanadi. Bundan tashqari, ko'p ma'noli so'zlar savol generatsiyasiga salbiy ta'sir ko'rsatishi mumkin. Shuningdek, hosil qilingan savollarning tabiiyligi muammosi ham dolzarb bo'lib, ba'zi hollarda savollar sun'iy ko'rinishi mumkin [6].

Xulosa qilib aytganda, lingvistik modelga asoslangan savol generatsiyasi matn tuzilishini chuqur tahlil qilishga tayanadi va bu jarayonda sintaktik, morfologik hamda semantik omillar muhim rol o'ynaydi. Taklif etilgan model savol generatsiyasini tizimli tashkil etish imkonini beradi hamda mazmunli va grammatik jihatdan to'g'ri savollar hosil qilishga xizmat qiladi. Kelgusida lingvistik va neyron



yondashuvlarni birlashtirgan gibrid modellarning ishlab chiqilishi ushbu yo‘nalishda yanada samarali natijalar berishi kutiladi.

Foydalanilgan adabiyotlar ro‘yxati

1. Heilman M., Smith N. A. *Automatic Factual Question Generation from Text*. Carnegie Mellon University, 2011.
2. Kurdi G., Leo J., Parsia B., Sattler U., Al-Emari S. *A Systematic Review of Automatic Question Generation for Educational Purposes*. International Journal of Artificial Intelligence in Education, 2020.
3. Du X., Shao J., Cardie C. *Learning to Ask: Neural Question Generation for Reading Comprehension*. ACL, 2017.
4. Flor M., Riordan B. *A Semantic Role-Based Approach to Automatic Question Generation*. 2018.
5. Brown T. et al. *Language Models are Few-Shot Learners*. NeurIPS, 2020.
6. Das B. et al. *Automatic Question Generation: A Survey*. 2021.