

Avtomatik Savol Yaratishda sun'iy ong (Artificial Intelligence - AI) va tabiiy tilni ishlatishi(Natural Language Processing-NLP), taklif va tadbiri qilingan yechimlar

*Yuldashev Nodirbek Abdumannob o'g'li,
Andijon mashinasozlik institute,
"Texnologik jarayonlar va
ishlab chiqarishlarni avtomatlashtirish
boshqarish" ixtisosligi bo'yicha
mustaqil izlanuvchi, Andijan, Uzbekistan
+99897 164-03-96
nodirbek4405979@gmail.com*

Annotatsiya: Mazkur maqolada ta'lim boshqaruv tizimlarida, tashkilot hamda korxonalarining kadrlarni qayta tayyorlov hamda malakaviy imtixon jarayonlarida test tuzish jarayonlarini avtomatlashtirish va bu jarayonlarda sun'iy ong va tabiiy tillarini qo'llashning afzalliklari, taklif-mulohazalar hamda tadbiri qilingan yechimlar haqida fikr yuritiladi.

Kalit so'zlar: Avtomatik Savol Yaratish (AQG-Automatic Question Generation), tabiiy tilni (Natural Language Processing-NLP), sun'iy ong(AI), Python Packages, transformers' tools, techniques.

Avtomatik Savol Yaratish (AQG-Automatic Question Generation) - bu har bir boshqaruv tizimlari (LMS) uchun yangi va samarali yondashuv bo'lib, bu texnologiya sun'iy intellekt (AI) va tabiiy tildan foydalanganda holda ishlatadi(NLP). AQG orqali, test oluvchi va test topshiruvchilar uchun samarali, interaktiv va eng muhim bo'lgan test sinovlarini, tezkor hamda vaqtdan unumli foydalangan holda avtomatik savol yaratish imkoniyatini oshiradi. Ushbu maqola orqali AQG ning test olish va olingan bilimlarni qay darajada mukammal egallaganliklarini tekshirish tizimlari uchun qanday foydali bo'lishi, AI va tabiiy tilning (NLP) usulini qanday ishlatishini va bu texnologiya bilan avtomatik test tuzish tizimlarini qanday takomillashtirish haqida to'liqroq fikr-mulohazalar yuritiladi.

Birinchi bo'lib, AQG ning qanday yaratilishi va uning ahamiyati haqida fikr yuritsak. AQG – biz taklif qilgan yechimda birinchi o'rinda Open questions tipli savollarni yaratganda Python dasturlash tilining "*Haystack package*" dan foydalanildi. *Haystack* - ochiq manba Python kutubxonasi bo'lib, u suhbatdosh AI tizimlarini, xususan, savollarga javob berish va hujjatlarni qidirishni o'z ichiga olgan tizimlarni yaratish va ulardan foydalanish vositalari va usullarini taqdim etadi. U ko'pincha tabiiy tildagi so'rovlarni tushuna oladigan va ularga javob beradigan

chatbotlar, qidiruv tizimlari va savollarga javob berish tizimlarini yaratish uchun ishlatiladi.

Haystackning asosiy xususiyatlari va tarkibiy qismlaridan ba'zilari:

1. Hujjatlar do'koni(ombori): Haystack sizga hujjatlarni (masalan, matnli fayllar, PDF-fayllar, HTML-sahifalar) turli formatlarda saqlash va indekslash imkonini beradi, bu esa katta hajmdagi hujjatlardan ma'lumotlarni qidirish va olishni osonlashtiradi.
2. Oldindan ishlov berish va tozalash: u qidiruv va qidirish sifatini yaxshilash uchun zarur bo'lgan matn ma'lumotlarini oldindan qayta ishlash va tozalash vositalarini o'z ichiga oladi.
3. Savolga javob: Haystack o'rnatilgan savol-javob (QA) komponentini taqdim etadi, u tabiiy til savollariga asoslangan hujjatlardan javoblarni olish uchun mashinani o'rganish modellaridan foydalanadi.
4. Retrieverlar: Haystack turli xil retriever komponentlarini qo'llab-quvvatlaydi, jumladan, Elasticsearch, BM25 va Dense Passage Retrieval (DPR), ular so'rov bo'yicha tegishli hujjatlarni topishga yordam beradi.
5. Readers: Readers hujjatlardan ma'lumot olish uchun ishlatiladi. Haystack turli xil hujjat formatlari va tillari uchun turli xil o'quvchi modellarini qo'llab-quvvatlaydi.
6. Transformatorlar: QA vazifalari uchun transformatorga asoslangan modellarni (masalan, BERT, RoBERTa) birlashtira olasiz, bu modellarni Haystack doirasida nozik sozlash va ulardan foydalanishni osonlashtiradi.
7. REST API: Haystack REST API-ni taqdim etadi, bu HTTP so'rovlari orqali so'zlashuv AI tizimini o'rnatish va o'zaro ta'sir qilishni osonlashtiradi.

Endi bir nechta tanlov savollari ya'ni (Multiple Choice Questions MCQ) tipli savol tuzadigan jarayonga keladigan bo'lsak MCQ uchun question generation uchun train qilingan bir nechta machine learning modellar ishlatilgan. Ularni ishlatish uchun hozirgi kunda NLP da mashhur bo'lgan transformers toollari ishlatilgan. Python package lardan modelni ishlatish uchun pytorch, ma'lumotlarni tahlil qilish uchun numpy, pandas, datasetni tahlil qilish uchun *sklearn*, AI joyi uchun transformers ishlatilgan. MCQ ni generate qilish 3 bosqichga bo'linadi (code qismida ham alohida 3 ta folderga bo'lingan):

- 1 - qism textni olib uni tokenize qilish (tokenize degani textni raqamli korinihsga keltirib olish) va undan savol generate qilish uchun transformerni t5-small modeli ishlatilgan
- 2 - qism generate qilingan savolga to'g'ri javobni aniqlash, javob generate qilish uchunham transformerni t5-small modeli ishlatilgan

3 - qism savolga oid javobga yaqin bo'lgan 3 ta noaniq javobni generate qilish bular distractor deb ataladi, distractorlarni generate qilish uchun ham transformerni t5-small modeli ishlatilgan.

Agar distractorlarni bu model yetarlicha generate qila olmasa qo'shimchasida Sense2Vec vectorization texnikasi ishlatilgan. Bularni barchasini bajarish orqasida shu question generate qilish muammolari hal qilindi, open questions va MCQ lar generate bolyapti, bu orqali bir inson uzoq vaqt davomida otirib bir textni tahlil qilib unga question yaratgunicha, AI questionlarni answerlari bilan generate qiladi va bu bir necha barobar tezroq bo'ladi.

Xulosa: Mening bu ilmiy maqolani nashr etishdan asosiy maqsadim bunday tizim yoki modulda O'rta Osiyoda ilmiy tadqiqot olib borayotgan mutaxassislar kamchilikni tashkil qiladi, shuning uchun, maqolamdan boshqa tashkilot va shunga o'xshash ilmiy tadqiqot ustida ishlayotgan va ishlamoqchi bo'lgan yosh olimlar uchun, taklif qilinayotgan modulni test qilib ko'rish uchun, fikr yoki tajribalarini sinab ko'rish uchun ishlatishlari mumkin. Yangi va ilg'or texnologiyalarga asoslangan, yuqori saviyadagi bu kabi tizimlashtirish, avtomatik savollar yaratish, savollarni baholash va moslashuvchan amaliyot uchun ilovalar amalga oshirilishi va tizimga oson integratsiya qilinishi mumkinligini anglashlari muhim, bu esa bunday modullarni tashkilot va korxonalar uchun kadrlarni qayta tayyorlash va ularning malakarini tekshirish uchun avtomatik savol yaratuvchi tizimlarga tezroq va osonroq integratsiyasi uchun ijobiy tasirini o'tkazadi.

Adabiyotlar:

1. Williams, H. (2013) "Intellectual Property Rights and Innovation: Evidence from the Human Genome," *Journal of Political Economy*, 121(1): 1-27
2. R. J. de Ayala, *Theory and practice of item response theory*. Guilford Publications, 2013.
3. Leung, M.K.K., A. Delong, B. Alipanahi, and B.J. Frey (2016) "Machine Learning in Genomic Medicine: A Review of Computational Problems and Data Sets," *Proceedings of the IEEE*, 104(1): 176-197.
4. Yates, T. (2016). Automated generation of questions from factual, natural language sentences
5. J. Papoušek and R. Pelánek, "Impact of adaptive educational system behaviour on student motivation", in *Proceedings of Artificial Intelligence in Education (AIED)*, 2015.
6. R. Singh, A. Singhrova, and R. Bhatia. "Test Case Generation Tools-A Review". In: *International Journal of Electronics Engineering (ISSN: 0973-7383)* (2018).