

TABIIY TILDA MATNLARNI QAYTA ISHLASH.

Abdurahmonova Umida Rustamovna,

Alisher Navoiy nomidagi Toshkent davlat o'zbek tili va adabiyoti universiteti “Kompyuter lingvistikasi va raqamli texnologiyalar” kafedrasida o'qituvchisi.

Qayumov Surxon O'lmas o'g'li,

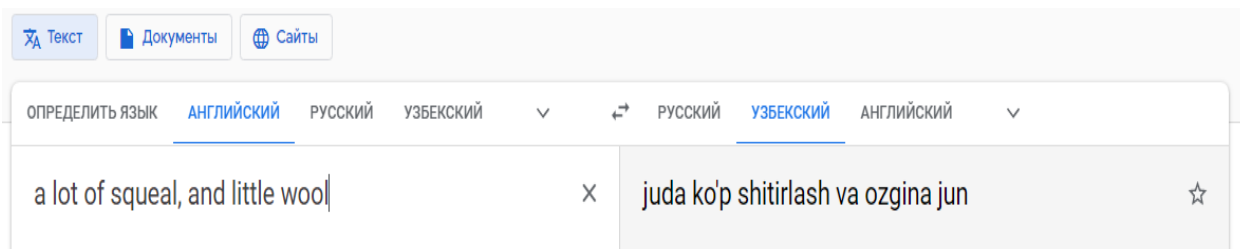
Alisher Navoiy nomidagi Toshkent davlat o'zbek tili va adabiyoti universiteti,
Kompyuter lingvistikasi mutaxassisligi 2-bosqich magistranti.

Annotatsiya. Ijobiy natijalarga ega bo'lgan tadqiqotlar mavjud bo'lsa-da, ilmiy sharhlarga ko'ra, bunday tadqiqotlarning barchasida uslubiy kamchiliklar mavjud. Hammasi bo'lib, NLP tarafdorlari da'vo qilgan natijalarni takrorlash mumkin bo'lmagan yuqori sifatli tadqiqotlar uch baravar ko'p bo'lgan. Bu tabiiy tilda (ya'ni odamlar gapiradigan va yozadigan tilda) matnlarni qayta ishlash bo'yicha keng ko'lamli vazifalardir. Klassik NLP muammolari to'plami mavjud bo'lib, ularning echimi amaliy ahamiyatga ega.

Kalit so'zlar: *til (lingvistik), maxsus texnika, zamonaviy neuro lingvistika, matn, korpus.*

Bu tabiiy tilda (ya'ni odamlar gapiradigan va yozadigan tilda) matnlarni qayta ishlash bo'yicha keng ko'lamli vazifalardir. Klassik NLP muammolari to'plami mavjud bo'lib, ularning echimi amaliy ahamiyatga ega.

Birinchi va eng muhim tarixiy vazifa - bu mashina tarjimasidir. Bu juda uzoq vaqtdan beri mavjud va katta taraqqiyot bor. Ammo yuqori sifatli to'liq avtomatik tarjimani (FAHQMT) olish vazifasi hali ham hal qilinmagan. Bu qaysidir ma'noda NLP ning dvigateli, bajarilishi mumkin bo'lgan eng katta vazifalardan biridir.

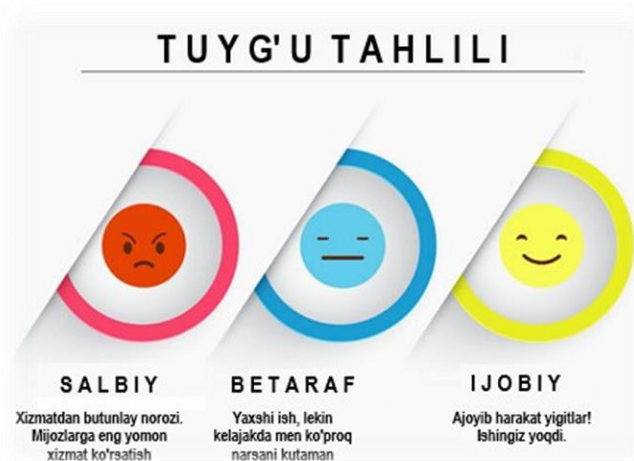


Ikkinchi vazifa - matnlarni tasniflash. Matnlar to'plami berilgan va vazifa bu matnlarni toifalarga ajratishdir. Nima? Bu korpus uchun savol.

Amaliy nuqtai nazardan qo'llashning birinchi va eng muhim usullaridan biri bu harflarni spam va “xam”ga (spam emas) tasniflashdir.

Yana bir klassik variant - yangiliklarni toifalarga ya'ni ko'p toifali tasniflash - tashqi siyosat, sport, katta harid va boshqalar. Yoki, aytaylik, sizga xat keldi va siz onlayn-do'kondagi buyurtmalarni aviachiptalar va mehmonxonalarni bron qilishdan ajratmoqchisiz.

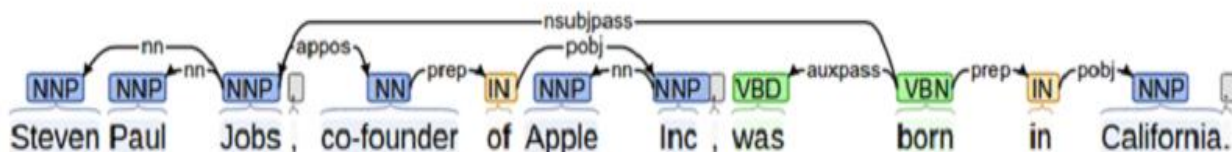
Matnni tasniflash muammosining uchinchi klassik qo'llanilishi - bu hissiyotlarni tahlil qilish. Masalan, sharhlarni ijobiy, salbiy va neytral deb tasniflash.



Matnlarni bo'lish mumkin bo'lgan ko'plab toifalar mavjud bo'lganligi sababli, matnni tasniflash NLPdagi eng mashhur amaliy muammolardan biri bo'lib qolgan.

Uchinchi vazifa NER deb nomlanadi. Biz matnda oldindan tanlangan ob'ektlar to'plamiga mos keladigan bo'limlarni tanlaymiz, masalan, matndagi barcha joylarni, shaxslarni va tashkilotlarni topishimiz kerak. "Ostap Bender – Qishloq qurilish loyiha direktori" matnida siz Ostap Bender shaxs, Qishloq va qurilish esa tashkilot ekanligini tushunishingiz kerak. Nima uchun bu vazifa amalda kerak va uni qanday hal qilish kerak. Uchinchi vazifa bilan bog'liq bo'lgan to'rtinchi vazifa - faktlar va munosabatlarni ajratib olish (munosabatlarni chiqarish). Masalan, ish (Kasb) munosabati mavjud. "Ostap Bender - Qishloq qurilish loyiha direktori" matnidano bizning qahramonimiz Qishloq va qurilish bilan professional

munosabatlar bilan bog‘liqligi aniq. Xuddi shu narsani boshqa ko‘plab yo‘llar bilan aytish mumkin: " Qishloq qurilish loyiha idorasini Ostap Bender boshqaradi" yoki "Ostap Bender leytenant Shmidtning oddiy o‘g‘lidan Qishloq qurilish loyiha idorasi boshlig‘igacha o‘tdi". Bu jumlar nafaqat predikat, balki tuzilishi jihatidan ham farqlanadi. Boshqa tez-tez aniqlangan munosabatlarga misollar - sotib olish / sotish (sotib olish va sotish), mulkchilik (egalik), atributlar bilan tug‘ilish fakti - sana, joy va boshqalar (tug‘ilgan) va boshqalar. Vazifaning aniq amaliy qo‘llanilishi yo‘qdek tuyuladi, lekin shunga qaramay, u tuzilmagan ma'lumotni strukturalashda qo‘llaniladi. Bundan tashqari, bu savol-javob va dialog tizimlarida, qidiruv tizimlarida - har doim savolni tahlil qilish va uning qaysi turga tegishli ekanligini, shuningdek, javobda qanday cheklovlar mavjudligini tushunish kerak bo‘lganda muhimdir.



Keyingi ikkita vazifa, ehtimol, eng shov-shuvdir. Bular savol-javob va dialog tizimlari (chat-botlar). Amazon Alexa, Alisa suhbat tizimlarining klassik namunalaridir. Ularning to‘g‘ri ishlashi uchun ko‘plab NLP muammolarini hal qilish kerak. Misol uchun, matn tasnifi biz maqsadli chatbot stsenariylaridan biriga tushib qolishimizni aniqlashga yordam beradi. Masalan, "valyuta kurslari masalasi". Skript shablonini to‘ldiruvchilarini aniqlash uchun munosabatlarni ajratib olish kerak va umumiy mavzular bo‘yicha dialog o‘tkazish vazifasi ("chat") biz biron bir stsenariyga tushmagan vaziyatda bizga yordam beradi. Savol-javob tizimlari ham tushunarli va foydali narsadir. Siz mashinaga savol berasiz, mashina javobni ma'lumotlar bazasidan yoki matnlar korpusidan qidiradi. Bunday tizimlarga IBM Watson yoki Wolfram Alpha misol bo‘la oladi. Klassik NLP muammosining yana bir misoli umumlashtirishdir. Muammoni shakllantirish oddiy - tizim kirish sifatida

katta matnni oladi va chiqish kattaroq matnning mazmunini qandaydir tarzda aks ettiruvchi kichikroq matndir. Masalan, mashina matnning parafrasasini, uning sarlavhasini yoki izohini yaratishi kerak.

Article

(cnn) to allay possible concerns, boston prosecutors released video friday of the shooting of a police officer last month that resulted in the killing of the gunman. the officer wounded, john moynihan, is white. angelo west, the gunman shot to death by officers, was black. after the shooting, community leaders in the predominantly african-american neighborhood of (...)

Human-written summary

boston police officer john moynihan is released from the hospital. video shows that the man later shot dead by police in boston opened fire first. moynihan was shot in the face during a traffic stop.

Generated summary (See et al., 2017)

boston prosecutors released video friday of the shooting of a police officer last month. the gunman shot to death by officers , was black . one said the officers were forced to return fire. he was placed in a medically induced coma at a boston hospital.

Yana bir mashhur vazifa - argumentatsiyani qazib olish, matnda asoslanishni izlash. Sizga fakt va matn berilgan, matndan bu faktning asosini topishingiz kerak. Bu, albatta, NLP vazifalarining to'liq ro'yxati emas. Ularning sanog'I ko'p. Umuman olganda, tabiiy til matni bilan bajarilishi mumkin bo'lgan barcha narsalarni NLP vazifalari sifatida tasniflash mumkin, faqat sanab o'tilgan mavzular eshitiladi va ular eng aniq amaliy qo'llanmalarga ega.

Foydalanilgan adabiyotlar:

1. Aripov, M., Sharipbay, A., Abdurakhmonova, N., Razakhova B.: Ontology of grammar rules as example of noun of Uzbek and Kazakh languages. In: Abstract of the VI International Conference “Modern Problems of Applied Mathematics and Information Technology - Al-Khorezmiiy 2018”, pp. 37–38, Tashkent, Uzbekistan (2018)
2. Abduraxmonova, N. Z. "Linguistic support of the program for translating English texts into Uzbek (on the example of simple sentences): Doctor of Philosophy (PhD) il dis. aforef." (2018).
3. Garri Alder - Nlp ning zamonaviy psixotexnologiyalari. 2003-2011 Nashriyotchi: Profit Style, Phoenix
4. S. V. Ilyinskiy. - Jamoat bilan aloqa. Reklama. Marketing. Neyro-lingvistik dasturlash. Operatsion ma'lumotnoma lug'ati AST: Sharq-G'arb, 2006. S. 204.