

Oriental Journal of Philology**ORIENTAL JOURNAL OF PHILOLOGY**

journal homepage:

<http://www.supportscience.uz/index.php/ojp/about>**COMPUTATIONAL LINGUISTICS: PROBLEMS AND SOLUTIONS***Aziza Vokhidillayevna Yormatova**Lecturer**UzSUWL**Tashkent, Uzbekistan***ABOUT ARTICLE**

Key words: Computational linguistics, natural language processing, artificial intelligence, speech, gestures, NLU, NLG, machine language, domain, context.

Received: 15.05.24

Accepted: 17.05.24

Published: 19.05.24

Abstract: In this article, some information about computational linguistics, some current problems in this computational linguistics and their solutions are discussed. Information about the working principle of computer linguistics and its special aspects is also provided. Also known as natural language processing (NLP), linguistics, computer science, and artificial intelligence (AI) are covered. Also discussed are language understanding, language generation, text processing and analysis, machine translation, speech recognition and synthesis Natural Language Understanding (NLU) and Natural Language Generation (NLG), dialog systems and conversational agents. This article also covers statistical and symbolic approaches, deep learning and neural models, language generation and understanding, real-world applications of multilingual and cross-linguistic NLP.

KOMPYUTER LINGVISTIKASI: MUAMMO VA YECHIMLAR*Aziza Voxidillayevna Yormatova**o'qituvchi**O'zDJTU**Toshkent, O'zbekiston***MAQOLA HAQIDA**

Kalit so'zlar: Kompyuter lingvistikasi, tabiiy tilni qayta ishlash, sun'iy intellekt, nutq, imo-ishoralar, NLU, NLG, mashina tili, domen, kontekst.

Annotatsiya: Ushbu maqolada, kompyuter lingvistikasi haqida ba'zi ma'lumotlar, hozirda ushbu kompyuter lingvistikasida uchrayotgan dolzarb muammolar va ularning yechimlari to'g'risida so'z yuritilgan. Kompyuter lingvistikasini

ishlash prinsipi va uning alohida jihatlari to'g'risida ham ma'lumotlar berib o'tilgan. Shu bilan birga tabiiy tilni qayta ishlash (NLP) sifatida ham tanilgan, tilshunoslik, kompyuter fanlari va sun'iy intellekt (AI) haqidagi ma'lumotlarga alohida to'xtalib o'tilgan. Shuningdek, tilni tushunish, til avlodi, matnga ishlov berish va tahlil qilish, mashina tarjimasini, nutqni aniqlash va sintez qilish Natural Language Understanding (NLU) va Natural Language Generation (NLG), dialog tizimlari va suhbat agentlari kabi tushunchalarga alohida to'xtalib o'tilgan. Ushbu maqolada shu bilan birga statistik va ramziy yondashuvlar, chuqur o'rganish va neyron modellari, tilni yaratish va tushunish, ko'p tilli va tillararo NLP Real dunyo ilovalari haqida ko'plab ma'lumotlar keltirilgan.

КОМПЬЮТЕРНАЯ ЛИНГВИСТИКА: ПРОБЛЕМЫ И РЕШЕНИЯ

Азиза Вохидиллаевна Ёрматова

Преподаватель

УзГУМЯ

Ташкент, Узбекистан

О СТАТЬЕ

Ключевые слова: компьютерная лингвистика, обработка естественного языка, искусственный интеллект, речь, жесты, NLU, NLG, машинный язык, предметная область, контекст.

Аннотация: В этой статье обсуждаются некоторая информация о компьютерной лингвистике, некоторые текущие проблемы в этой компьютерной лингвистике и их решения. Также представлена информация о принципе работы компьютерной лингвистики и ее особенностях. Охватываются также известные как обработка естественного языка (НЛП), лингвистика, информатика и искусственный интеллект (ИИ). Также обсуждаются понимание языка, генерация языка, обработка и анализ текста, машинный перевод, распознавание и синтез речи, понимание естественного языка (NLU) и генерация естественного языка (NLG), диалоговые системы и диалоговые агенты. В этой статье также рассматриваются статистические и символические подходы, глубокое обучение и нейронные модели, генерация и понимание языка, реальные применения многоязычного и межъязыкового НЛП.

KIRISH

Hozirda mamlakatimizda kompyuter lingvistikasi doirasida turli chora-tadbirlar amalga oshirilmoqda, shuningdek O‘zbekiston Respublikasi Prezidentining “Raqamli iqtisodiyot va elektron hukumatni keng joriy etish chora-tadbirlari to‘g‘risida”gi hamda “Aholiga davlat ijtimoiy xizmatlari va yordam taqdim etish tartib-taomillarini avtomatlashtirish bo‘yicha qo‘shimcha chora-tadbirlar to‘g‘risida”gi qarorlari va boshqa normativ-huquqiy hujjatlar mamlakatimizda raqamlashtirishni jadallashtirish va ijtimoiy-iqtisodiy sohalarga zamonaviy texnologiyalarni joriy qilishga qaratilgan tadbirlar amalga oshirilmoqda.[1]

ASOSIY QISM

Kompyuter lingvistikasi, shuningdek, tabiiy tilni qayta ishlash (NLP) sifatida ham tanilgan, tilshunoslik, kompyuter fanlari va sun‘iy intellekt (AI) markazida joylashgan sohadir. U inson tilining hisoblash jihatlariga qaratilgan bo‘lib, kompyuterlarga tabiiy tilni insonlar uchun mazmunli va foydali tarzda tushunish, yaratish va boshqarish imkonini beradi. Kompyuter lingvistikasining ba’zi asosiy jihatlari:

1. Tilni tushunish: Kompyuter lingvistikasi kompyuterlarga inson tilini turli shakllarda, jumladan matn, nutq va imo-ishoralarda tushunish imkonini beruvchi algoritm va modellarni ishlab chiqishga intiladi. Bu lingvistik kirishdan ma’no chiqarish uchun tahlil qilish, semantik tahlil va nutqni qayta ishlash kabi vazifalarni o‘z ichiga oladi.

2. Til avlodi: Tilni tushunishdan tashqari, kompyuter lingvistikasi tabiiy til mahsulotini yaratish bilan ham shug‘ullanadi. Bunga matn yaratish, nutq sintezi va dialog yaratish kabi vazifalar kiradi, bunda kompyuterlar izchil va kontekstga mos keladigan lingvistik natijalarni ishlab chiqaradi.

3. Matnga ishlov berish va tahlil qilish: Kompyuter lingvistikasi katta hajmdagi matn ma’lumotlarini qayta ishlash va tahlil qilish usullarini o‘z ichiga oladi. Bunga qidiruv tizimlaridan tortib, ijtimoiy media tahlillarigacha bo‘lgan ilovalarda qo‘llaniladigan ma’lumotlarni qidirish, matn tasnifi, hissiyotlarni tahlil qilish va matnni umumlashtirish kabi vazifalar kiradi.

4. Mashina tarjimasi: Kompyuter lingvistikasining ko‘zga ko‘ringan qo‘llanmalaridan biri bu mashina tarjimasi bo‘lib, kompyuterlar matn yoki nutqni bir tildan ikkinchi tilga avtomatik tarzda tarjima qiladi. Mashina tarjimasi tizimlari aniq va ravon tarjimalarga erishish uchun statistik modellarga, qoidalarga asoslangan yondashuvlarga yoki yaqinda neyron tarmoq arxitekturalariga tayanadi.

5. Nutqni aniqlash va sintez qilish: Kompyuter lingvistikasi nutqni aniqlash (og‘zaki tilni matnga aylantirish) va nutq sintezi (matndan og‘zaki nutqni yaratish) tizimlarini ishlab chiqishda hal qiluvchi rol o‘ynaydi. Ushbu tizimlar virtual yordamchilarda, ovoz bilan boshqariladigan qurilmalarda va nogironlar uchun qulaylik vositalarida qo‘llaniladi.

6. Natural Language Understanding (NLU) and Natural Language Generation (NLG): NLU kompyuterlarga inson tilini tushunish imkonini yaratishga, NLG esa tabiiy til chiqishini yaratishga qaratilgan. Ikkala soha ham inson tili va mashina tushunchasi o'rtasidagi tafovutni bartaraf etish uchun semantik vakillik, sintaktik tahlil va pragmatik talqinni o'z ichiga oladi.

7. Dialog tizimlari va suhbat agentlari: Kompyuter lingvistikasi chatbotlar va virtual yordamchilar kabi dialog tizimlari va suhbat agentlarini ishlab chiqishda markaziy o'rin tutadi. Ushbu tizimlar foydalanuvchilar bilan tabiiy tilda suhbatlar olib boradi, bu murakkab tilni tushunish, yaratish va dialoglarni boshqarish imkoniyatlarini talab qiladi.

8. Sanoat va tadqiqotlardagi ilovalar: Kompyuter lingvistikasi turli sohalarda, jumladan sog'liqni saqlash (klinik matn qidirish), moliya (iqtisodiyotni tahlil qilish), ta'lim (til bo'yicha repetitorlik) va ko'ngilochar (video o'yinlardagi dialog tizimlari) kabi ilovalarga ega. Bu, shuningdek, neyron til modellari, ko'p tilli ishlov berish va tushuntiriladigan AI kabi sohalarda davom etayotgan tadqiqotlar bilan jonli tadqiqot sohasidir.[2]

Umuman olganda, kompyuter lingvistikasi inson tilini tushunishimizni rivojlantirishda va tabiiy tilda odamlar va mashinalar o'rtasidagi aloqani osonlashtiradigan texnologiyalarni ishlab chiqishda hal qiluvchi rol o'ynaydi. U sun'iy intellekt, mashinani o'rganish va lingvistik nazariyalardagi yutuqlar tufayli tez rivojlanishda davom etmoqda.

Kompyuter lingvistikasi tabiiy tilning murakkabligidan tortib til texnologiyalaridan foydalanishda axloqiy mulohazalargacha bo'lgan bir qancha muammolarga duch keladi. Kompyuter lingvistikasidagi ba'zi asosiy muammolar va potentsial echimlar:

1. Noaniqlik va kontekst sezgirligi: Tabiiy til o'z-o'zidan noaniq va kontekstga sezgir bo'lib, kompyuterlar uchun inson tilini to'g'ri tushunish va yaratishni qiyinlashtiradi. Yechimlarga quyidagilar kiradi:

- lingvistik kontekstga asoslangan ma'nolarni ajratish uchun semantik tahlil va kontekstni modellashtirish uchun ilg'or algoritmlarni ishlab chiqish.

- Izohlangan ma'lumotlarning katta korpusidan o'rganish va muayyan sohalarda tilni tushunishni yaxshilash uchun mashinani o'rganish usullarini integratsiyalash.

2. Lingvistik manbalarning yetishmasligi: Ko'pgina tillarda til modellarini o'rgatish va baholash uchun zarur bo'lgan izohli korpuslar, leksikonlar va grammatik ramkalar kabi yetarli lingvistik resurslar yo'q. Yechimlar quyidagilarni o'z ichiga oladi:

- Resurssiz tillar uchun lingvistik ma'lumotlarni to'plash va izohlash uchun kraudsorsing harakatlari.

- Til resurslarini yaratish va almashish uchun tadqiqotchilar, hukumatlar va til jamoalari o'rtasidagi hamkorlik tashabbuslari.

3. Domenga xos muammolar: Tilni tushunish va yaratish vazifalari turli domenlar va ilovalarda farqlanadi, ular maxsus bilim va moslashtirilgan yechimlarni talab qiladi. Yechimlarga quyidagilar kiradi:

- Mavjud til modellarini sozlash yoki domenga xos ta'lim ma'lumotlari orqali muayyan domenlarga moslashtirish.

- Ixtisoslashgan sohalarda tilni tushunishni yaxshilash uchun domenga xos ontologiyalar va bilim grafiklarini ishlab chiqish.

4. Axloqiy muammolar: Til texnologiyalari o'quv ma'lumotlarida mavjud bo'lgan noto'g'ri qarashlarni davom ettirishi yoki algoritmik qarorlar orqali yangi noto'g'ri qarashlarni kiritishi mumkin, bu esa zararli natijalarga olib keladi. Yechimlar quyidagilarni o'z ichiga oladi:

- Til modellari va ilovalaridagi noxolisliklarni aniqlash va yumshatish uchun xolis audit va adolatni baholashni o'tkazish.

- ta'lim ma'lumotlarida xilma-xillik va inklyuzivlikni rag'batlantirish va til texnologiyalarini rivojlantirishda turli jamoalar vakillarini ta'minlash.

5. Talqin va tushuntirish mumkinligi: Kompyuter lingvistikasidagi qora quti modellari shaffoflikka ega bo'lmasligi mumkin, bu ularning qarorlarini izohlashni va ularning asosiy mexanizmlarini tushunishni qiyinlashtiradi. Yechimlarga quyidagilar kiradi:

- Til modellari o'z bashoratlariga qanday erishishi haqida tushuncha beradigan izohlanadigan modellar va tushuntiriladigan AI usullarini ishlab chiqish.

- Shaffoflik va foydalanuvchi ishonchini oshirish uchun lingvistik bilimlar va inson tomonidan o'qiladigan tasvirlarni namunaviy natijalarga kiritish.

6. Ma'lumotlarning maxfiyligi va xavfsizligi: Til texnologiyalari maxfiy yoki shaxsiy ma'lumotlarni qayta ishlashni o'z ichiga olishi mumkin, bu esa ma'lumotlar maxfiyligi va xavfsizligiga oid xavotirlarni keltirib chiqarishi mumkin. Yechimlar quyidagilarni o'z ichiga oladi:

- Muhim lingvistik ma'lumotlarni himoya qilish uchun shifrlash, anonimlashtirish va kirishni boshqarish kabi ishonchli ma'lumotlarni himoya qilish choralarini qo'llash.

- Til texnologiyalarini ishlab chiqish va qo'llashda GDPR va mas'uliyatli AI tamoyillari kabi maxfiylik qoidalari va axloqiy ko'rsatmalarga rioya qilish.

7. Resurs intensiv trening: Keng miqyosli til modellarini o'rgatish atrof-muhit muammolariga hissa qo'shadigan katta hisoblash resurslari va energiya sarfini talab qiladi. Yechimlarga quyidagilar kiradi:

- Ish unumdorligini saqlab qolgan holda hisoblash xarajatlari va energiya sarfini kamaytirish uchun model arxitekturasini va o'qitish algoritmlarini optimallashtirish.

- Modelni o'qitishni bir nechta qurilmalarda tarqatish yoki model hajmini kamaytirish uchun federatsiyalangan o'rganish yoki modelni jiplash kabi muqobil yondashuvlarni o'rganish.[3]

Kompyuter lingvistikasidagi ushbu muammolarni hal qilish fanlararo hamkorlikni, axloqiy xabardorlikni va sohani mas'uliyatli va inklyuziv tarzda rivojlantiradigan innovatsion yechimlarni ishlab chiqish bo'yicha doimiy tadqiqot sa'y-harakatlarini talab qiladi.

Kompyuter lingvistikasidagi munozaralar sohaning fanlararo xususiyatini va uning nazariy tilshunoslik va tabiiy tilni qayta ishlash (NLP) va sun'iy intellekt (AI)dagi amaliy qo'llanilishiga aloqadorligini aks ettiruvchi keng mavzularni qamrab oladi. Kompyuter lingvistikasidagi ba'zi asosiy munozaralar va muhokamalar:

1. Tilni ifodalash: Munozaralar tilni hisoblash yo'li bilan qanday eng yaxshi ifodalash, jumladan, sintaktik tuzilmalar, semantik tasvirlar va nutq tuzilmalari atrofida bo'ladi. Tabiiy tilning murakkabligi va o'zgaruvchanligini qo'lga kiritish uchun eng mos shakllar va ramkalar haqida munozaralar mavjud.

2. Statistik va ramziy yondashuvlar: Kompyuter lingvistikasida mashina o'rganish va ehtimollik modellari kabi statistik usullarning, qoidaga asoslangan tizimlar va rasmiy grammatika kabi ramziy yondashuvlarga nisbatan nisbiy afzalliklari haqida munozaralar davom etmoqda. Tadqiqotchilar har bir yondashuvning kuchli va cheklovlarini muhokama qiladilar va ularni samarali integratsiya qilish yo'llarini o'rganadilar.

3. Chuqur o'rganish va neyron modellari: Chuqur o'rganishning kuchayishi bilan neyron tarmoq modellarini turli NLP vazifalariga, jumladan, tilni modellashtirish, mashina tarjimasini va hissiyotlarni tahlil qilishda qo'llash haqida munozaralar olib borilmoqda. Tadqiqotchilar neyron arxitekturaning imkoniyatlari va cheklovlarini, shuningdek, ularning mustahkamligi va talqin qilinishini yaxshilash usullarini o'rganadilar.

4. Tilni yaratish va tushunish: Munozaralar tabiiy tilni tushunish (NLU) va avlod (NLG) muammolariga, jumladan, noaniqlik, kontekst sezgirligi va umumiy fikrlash masalalariga qaratilgan. Tadqiqotchilar til modellarining aniqligi, izchilligi va turli lingvistik hodisalarni boshqarish qobiliyatini yaxshilash usullarini o'rganmoqdalar.

5. Ko'p tilli va tillararo NLP: Ko'p til va lahjalarda ishlay oladigan til texnologiyalariga bo'lgan ehtiyoj tufayli ko'p tilli va tillararo NLPga qiziqish ortib bormoqda. Munozaralar tilga agnostik modellarni yaratish, tillar o'rtasida bilimlarni uzatish va tilga xos muammolarni hal qilish usullariga qaratilgan.

6. Real dunyo ilovalari: Munozaralar axborot olish, savollarga javob berish, chatbotlar, hissiyotlarni tahlil qilish va sog'liqni saqlash kabi sohalarda kompyuter lingvistikasining amaliy qo'llanilishini o'rganadi. Tadqiqotchilar NLP tizimlarini real dunyo sharoitlarida joylashtirish

muammolarini muhokama qiladilar va ulardan foydalanish qulayligi, ishonchliligi va foydalanish imkoniyatini oshirish yo‘llarini o‘rganadilar.[4;5]

Umuman olganda, kompyuter lingvistikasidagi munozaralar ushbu sohaning dinamik va rivojlanayotgan tabiatini aks ettiradi, tadqiqotchilar inson tili haqidagi tushunchamizni oshirish va innovatsion til texnologiyalarini rivojlantirish uchun doimiy ravishda yangi g‘oyalar, metodologiyalar va ilovalarni o‘rganishadi.[6]

XULOSA

Xulosa qilib aytganda, kompyuter lingvistikasi bugungi hayotimizga olib kelayotgan muhim qulayliklarni tan olishimiz kerak. Sun‘iy intellekt, mashinani o‘rganish va tabiiy tillarni qayta ishlash sohasidagi yutuqlardan foydalangan holda, kompyuter lingvistikasi bizning texnologiya bilan o‘zaro munosabatlarimiz va ma‘lumotlarga kirishda inqilob qildi.

Xulosa qilib aytganda, kompyuter lingvistikasi bizning muloqot qilish, ma‘lumotlarga kirish va texnologiya bilan o‘zaro munosabatimizni o‘zgartirib, kundalik hayotimizni yaxshilaydigan ko‘plab qulayliklarni taqdim etdi. Soha taraqqiyotda davom etar ekan, biz hayotimizning turli jabhalarida qulaylik va foydalanish imkoniyatini yanada oshirib, til texnologiyalarida keyingi innovatsiyalar va takomillashuvlarni kutishimiz mumkin.

FOYDALANILGAN ADABIYOTLAR RO‘YXATI

1. Sun‘iy intellekt — to‘rtinchi sanoat inqilobining asosi. <https://xabar.uz/uz/texnologiya/suniy-intellekt-tortinchi-sanoat>
2. “Kompyuter lingvistikasi: muammo va yechimlar” mavzusidagi an’anaviy xalqaro ilmiy–amaliy onlayn konferensiya. <https://fledu.uz/uz/kompyuter-lingvistikasi-xalqaro-konferensiya-2022/>
3. Kompyuter lingvistikasi: muammolar, yechim, istiqbollar <https://azkurs.org/kompyuter-lingvistikasi-muammolar-yechim-istiqbollar-v2.html>
4. MATNNI QAYTA ISHLASH VA TAHLIL QILISH USULLARI Abduraxmonova Umida Rustamovna
5. Vol. 1 №. 01 (2021) <http://compling.navoiy-uni.uz/95>
6. “KOMPYUTER LINGVISTIKASI: MUAMMO VA YECHIMLAR” MAVZUSIDA AN’ANAVIY XALQARO ILMIY-AMALIY ONLAYN KONFERENSIYA. <https://fledu.uz/uz/kompyuter-lingvistikasi-muammo-va-yechimlar-mavzusida-ananaviy-xalqaro-ilmiy-amaliy-onlayn-konferensiyasi-o%ca%bbtkaziladi/>
7. The Oxford Handbook of Computational Linguistics (2nd edn) Ruslan Mitkov (ed.)
8. Computational linguistics. https://en.wikipedia.org/wiki/Computational_linguistics