

TABIIY TILNI QAYTA ISHLASH (NLP) VA SENTIMENT TAHLILI

Qosimova Mehriniso Muzaffarxon qizi

FarDU Axborot tizimlari va texnologiyalari yo'nalishi 3-kurs talabasi

qosimovamehriniso50@gmail.com

Sobirjonov Behzod Qahramonovich

FarDU Axborot texnologiyalari kafedrası Katta o'qituvchisi

bekzodbekqahromonovich@gmail.com

<https://doi.org/10.5281/zenodo.20325187>

Annotatsiya: Mazkur maqolada tabiiy tilni qayta ishlash (Natural Language Processing – NLP) yo'nalishining nazariy asoslari, asosiy tushunchalari hamda amaliy qo'llanilish sohalari yoritiladi.[1:12] NLP sun'iy intellekt va kompyuter lingvistikasining integrativ yo'nalishi sifatida matn va nutqni avtomatik tahlil qilish, tushunish va generatsiya qilish masalalarini hal etishga qaratilgan.[7:15] Maqolada NLPning tipik vazifalari, jumladan tokenizatsiya, lemmatizatsiya, morfologik va sintaktik tahlil, nomlangan obyektlarni aniqlash, sentiment tahlili, mashina tarjimasi, savol-javob tizimlari hamda matnni qisqartirish kabi yo'nalishlar umumlashtiriladi. Shuningdek, klassik yondashuvlardan zamonaviy chuqur o'rganish (deep learning) va transformer arxitekturalarigacha bo'lgan rivojlanish tendensiyalari hamda ta'lim, tibbiyot, bank-moliya, media monitoring, call-markazlar va davlat xizmatlarida NLP yechimlarining ahamiyati tahlil qilinadi. [5:60] Tadqiqot natijasida NLP texnologiyalarini joriy etishda ma'lumotlar sifati, til resurslari, etik masalalar va xavfsizlik talablarini hisobga olish zarurligi asoslab beriladi.

Kalit so'zlar: NLP,tabiiy tilni qayta ishlash,kompyuter lingvistikasi,mashina o'rganishi,transformer,mashina tarjimasi,sentment tahlili.

NATURAL LANGUAGE PROCESSING (NLP) AND SENTIMENT ANALYSIS

Qosimova Mehriniso Muzaffarxon qizi

3rd-year student, Information Systems and Technologies program, FarDU

qosimovamehriniso50@gmail.com

Sobirjonov Behzod Qahramonovich

Senior Lecturer, Department of Information Technologies, FarDU

bekzodbekqahromonovich@gmail.com

Abstract:This article discusses the theoretical foundations, key concepts, and practical application areas of Natural Language Processing (NLP). As an integrative field of artificial intelligence and computational linguistics, NLP is aimed at solving problems related to the automatic analysis, understanding, and generation of text and speech.[1:12] The article summarizes the typical tasks of NLP, including tokenization, lemmatization, morphological and syntactic analysis, named entity recognition, sentiment analysis, machine translation, question-answering systems, and text summarization. [7:15]In addition, development trends ranging from classical approaches to modern deep learning and transformer architectures, as well as the significance of NLP solutions in education, healthcare, banking and finance, media monitoring, call centers, and public services are analyzed. The study substantiates the necessity of considering data quality, language resources, ethical issues, and security requirements in the implementation of NLP technologies.

Keywords: NLP, natural language processing, computational linguistics, machine learning, transformer, machine translation, sentiment analysis.[5:60]

ОБРАБОТКА ЕСТЕСТВЕННОГО ЯЗЫКА (NLP) И АНАЛИЗ ТОНАЛЬНОСТИ

Косимова Мехринисо Музаффархон кизи

Студентка 3 курса направления

«Информационные системы и технологии», ФарГУ

qosimovamehriniso50@gmail.com

Собиржонов Бехзод Кахрамонович

Старший преподаватель кафедры информационных технологий ФарГУ

bekzodbekqahromonovich@gmail.com

Аннотация: В данной статье рассматриваются теоретические основы, основные понятия и практические области применения обработки естественного языка (Natural Language Processing – NLP). NLP, как интегративное направление искусственного интеллекта и компьютерной лингвистики, направлено на решение задач автоматического анализа, понимания и генерации текста и речи. В статье обобщаются типичные задачи NLP, включая токенизацию, лемматизацию, морфологический и синтаксический анализ, распознавание именованных сущностей, анализ тональности, машинный перевод, системы вопросов и ответов, а также автоматическое реферирование текста. Кроме того, анализируются тенденции развития от классических подходов до современных методов глубокого обучения (deep learning) и архитектур transformer, а также значение NLP-решений в образовании, медицине, банковско-финансовой сфере, медиамониторинге, call-центрах и государственных услугах. В результате исследования обосновывается необходимость учета качества данных, языковых ресурсов, этических вопросов и требований безопасности при внедрении технологий NLP.

Ключевые слова: NLP, обработка естественного языка, компьютерная лингвистика, машинное обучение, transformer, машинный перевод, анализ тональности.

Kirish: Tabiiy tilni qayta ishlash (NLP) - bu kompyuterlar va inson tili o'rtasidagi o'zaro ta'sirga qaratilgan sun'iy intellektning (AI) dinamik kichik sohasi. Asosiysi, NLP odamlar bilan muloqot qilish va kompyuter tushunchasi o'rtasidagi bo'shliqni bartaraf etishga qaratilgan bo'lib, mashinalarga inson tilini tabiiy va intuitiv tarzda talqin qilish, tahlil qilish va yaratish imkonini beradi.[1:45]NLP katta hajmdagi til ma'lumotlarini qayta ishlash va tushunish uchun hisoblash lingvistikasi, mashina o'rganish va chuqur o'rganishning kuchli tomonlarini birlashtiradi. Ushbu ilg'or texnologiyalardan foydalangan holda, NLP kompyuterlarga og'zaki yoki yozma ravishda inson tilining nuanslari, konteksti va maqsadini tushunishga imkon beradi. Bu qobiliyat virtual yordamchilar va chatbotlardan tortib til tarjimai va hissiyotlarni tahlil qilishgacha bo'lgan keng ko'lamli ilovalar uchun zarurdir.

Tabiiy tilni qayta ishlash nlp ning asosiy maqsadi odamlar bilan muammosiz o'zaro aloqada bo'ladigan, texnologiyani yanada qulayroq va kundalik muloqot ehtiyojlarimizga javob beradigan tizimlarni yaratishdir. NLP rivojlanishda davom etar ekan, u bizning mashinalar bilan o'zaro munosabatimizni o'zgartiradi, samarali, samarali va insoniy muloqot uchun yangi imkoniyatlarni ochadi.

Bugungi kunda ma'lumotlarni intellektual qayta ishlash tabiiy tilni qayta ishlash (Natural Language Processing, NLP)ning ommabop vazifalaridan biri hisoblanadi. Bunda matematik tenglamalar, formulalar, paradigmalar, shablonlarni matn shaklida ifodalash va uni qayta ishlash uchun matn semantikasini (tarkibini) tushunish uchun maqsadida tasniflash va fragmentatsiya kabi amallar bajarilishi lozim. Dasturchilar kompyuterlar va odamlarga tabiiy til yordamida o'zaro ma'lumot almashish imkonini beruvchi mexanizmlarini ishlab chiqadilar.[4:120]

Kompyuterlar NLP tufayli inson tilini o'qishi, sharhlashi, tushunishi va javob qaytarishi mumkin. Qoidaga ko'ra, ishlov berish mashinaning aql darajasiga asoslanadi.NLP algoritmlari inson xabarlarini uning uchun mazmunli bo'lgan raqamli shakldagi ma'lumotlarga aylantiradi. NLP jarayonlari va ularning algoritmlari qanday ishlashini tushunish juda muhimdir. Zamon bilan hamnafas bo'lish va bu texnologiyalar imkoniyatlaridan samarali foydalanish lozim. Hozirda NLPda hal qilinadigan vazifalar orasida eng muhimlari quyidagilar (1-rasm) [1,2]:

mashina tarjimai (machine translation) – NLP texnologiyalarini ishlab chiquvchilarga yuklatilgan birinchi klassik vazifa; imloni tekshirish (grammar and spell checking) – o'zbek tilidagi matnlar imlosini avtomatik tekshirish va tuzatish; matn tasnifi (text classification) – matn semantikasini aniqlash; nomlanuvchi obyektlarni aniqlash (named-entity recognition, NER) – ma'noga ega obyektlarni aniqlash; umumlashtirish (summarization) – matnni soddalashtirilgan shaklga umumlashtirish; matn generatsiyasi (text generation) – AI tizimlarini yaratishda foydalaniladigan vazifalardan biri; tematik modellashtirish (topic modeling) – kata hajmdagi matnlardan yashirin mavzularni ajratib olish.[8:171]

Natija va muhokama

Tabiiy tilni qayta ishlash nlp ning asosiy maqsadi odamlar bilan muammosiz o'zaro aloqada bo'ladigan, texnologiyani yanada qulayroq va kundalik muloqot ehtiyojlarimizga javob beradigan tizimlarni yaratishdir. NLP rivojlanishda davom etar ekan, u bizning mashinalar bilan o'zaro munosabatimizni o'zgartiradi, samarali, samarali va insoniy muloqot uchun yangi imkoniyatlarni ochadi. Natural Language Processing (NLP) – chatbotlardan tortib qidiruv tizimlari va Alexa va Siri kabi ovozli yordamchilargacha bo'lgan hamma narsani quvvatlovchi mashinalarga inson tilini tushunish, talqin qilish va qayta ishlashga yordam beruvchi ilg'or AI texnologiyasi. NLP kompyuter fanidan kelib chiqadi va uni rivojlantirish uchun hisoblash va analitik asoslarni ta'minlaydigan ma'lumotlar fanlari bilan chambarchas bog'liq. NLP-dan tabiiy tillarni dasturlash tillariga tarjima qilish, odamlar bilan muloqot qilish va kodni bog'lash uchun ham foydalanish mumkin.Tizimlar va kompyuterlar insoniy muloqotni aniq taqlid qila olishining asosiy sabablaridan biri audio, matnlar, ijtimoiy media kanallaridagi suhbat ma'lumotlari, videolar, elektron pochta xabarlar va boshqalar ko'rinishidagi ma'lumotlarning ko'pligidir. Biroq, NLP modellarini samarali o'rgatish uchun etarli ma'lumotlarga ega bo'lish juda muhim, chunki mashinani aniq o'rganish va nutqni aniqlash uchun katta etiketli ma'lumotlar to'plami zarur. Puxta sintaksislarning rivojlanishi modellarga odamlar muloqotidagi nuanslarni, jumladan kinoya, omonimlar, hazil va boshqalarni aniq tushunish imkonini berdi.[4:155]

Tabiiy tilni qayta ishlashga yondashuvlar.

Nazorat ostidagi NLP: E-pochtalarni tasniflash kabi aniq bashorat qilish uchun modellarni etiketlangan ma'lumotlarga o'rgatadi. Nazorat qilinmagan NLP: Mavzuni modellashtirish kabi vazifalar uchun foydali bo'lgan naqshlarni topish uchun etiketlanmagan ma'lumotlar bilan ishlaydi. Tabiiy tilni tushunish (NLU): Mashinalarga inson tilining ma'nosini talqin qilish va tushunishga yordam beradi. Tabiiy til avlodi (NLG): Xulosa yozish yoki chatbot javoblari kabi insonga o'xshash matn yaratadi. Tabiiy tilni qayta ishlash (NLP) bozori ajoyib va'da bermoqda va

156.80 yilga kelib taxminan 2030 milliard dollarga baholanishi kutilmoqda. Bu o'sish yillik CAGR 27.55% ni tashkil qiladi. Bundan tashqari, yirik tashkilotlarning 85% dan ortig'i 2025 yilgacha NLP ni o'zlashtirish ustida ishlamoqda. [5:602] NLP ning hayratlanarli o'sishiga turli sabablar sabab bo'ladi. NLP-ni bunday ommaviy qabul qilish va tarqatish ham qimmatga tushadi, bu erda McKinsey hisobotida NLP-ni avtomatlashtirish ishlarning 8 foizini eskirganligini aniqladi. Biroq, hisobotda aytilishicha, bu yangi ish o'rinlarining 9 foizini yaratish uchun javobgar bo'ladi. Natijalarning aniqligi haqida gap ketganda, zamonaviy NLP modellari GLUE benchmarkida 97% aniqlik haqida xabar berdi. Tabiiy tillarni qayta ishlash (NLP) tizimlari katta hajmdagi tuzilmagan ma'lumotlarni tahlil qilish va tegishli ma'lumotlarni olish uchun mashinani o'rganish algoritmlaridan foydalanadi. Algoritmlar naqshlarni aniqlash va shu naqshlar asosida xulosalar chiqarishga o'rgatilgan. Hujjatlarning samaradorligi va aniqligini oshirish NLP tomonidan yaratilgan hujjat odamlar avtomatik ravishda yarata olmaydigan har qanday asl matnni aniq umumlashtiradi. Shuningdek, u inson samaradorligini oshirish uchun katta hajmdagi ma'lumotlarni tahlil qilish kabi takrorlanadigan vazifalarni bajarishi mumkin. Katta va murakkab matnli kontentning qisqacha mazmunini avtomatik ravishda yaratish imkoniyati Tabiiy ishlov berish tili hujjatlardan faktlarni olish, his-tuyg'ularni tahlil qilish yoki nomli ob'ektlarni aniqlash kabi oddiy matnni qazib olish vazifalari uchun ishlatilishi mumkin. Tabiiy qayta ishlash inson xatti-harakatlari va his-tuyg'ularini tushunish kabi murakkabroq vazifalar uchun ham ishlatilishi mumkin. Sentiment tahlili (yoki his-tuyg'ular tahlili) — bu matndagi **ijobiy, salbiy yoki neytral** munosabatni avtomatik aniqlash uchun tabiiy tilni qayta ishlash (NLP) va mashinali o'rganish texnologiyasidir. U mijozlar fikrlarini (reviews), ijtimoiy tarmoq postlarini va so'rovlarni tahlil qilib, brend obro'sini boshqarish va qaror qabul qilishda yordam beradi. [1, 2, 3] NLP odamlar bilan matn yoki ovoz orqali muloqot qilish uchun sun'iy intellektdan foydalanadigan chatbotlar va kompyuter dasturlarida qo'llanilishi mumkin. Chatbot odam nima yozayotganini tushunish va tegishli javob berish uchun NLP dan foydalanadi. Ular, shuningdek, tashkilotga bir nechta kanallar bo'ylab 24/7 mijozlarni qo'llab-quvvatlash imkonini beradi. [3:310] Tabiiy tillar imlo xatolari, matn terish xatolari va uslubdagi nomuvofiqliklarga to'la. Masalan, “jarayon” so'zi “jarayon” yoki “qayta ishlash” deb yozilishi mumkin. Urg'u yoki lug'atingizda bo'lmagan boshqa belgilar qo'shsangiz, muammo yanada kuchayadi.

Xulosa: Tabiiy tilni qayta ishlash (NLP) sun'iy intellektning eng muhim va tez rivojlanayotgan yo'nalishlaridan biri bo'lib, inson va kompyuter o'rtasidagi muloqotni yanada qulay, tezkor va samarali tashkil etishga xizmat qiladi. NLP texnologiyalari yordamida matn va nutqni avtomatik tahlil qilish, tushunish hamda qayta ishlash imkoniyati yaratilmoqda. Tokenizatsiya, matnni tasniflash, mashina tarjimasini, chatbotlar, ovozli yordamchilar va sentiment tahlili kabi yo'nalishlar bugungi kunda keng qo'llanilmoqda. Shuningdek, NLP mashinali o'rganish va chuqur o'rganish usullari bilan integratsiyalashgan holda katta hajmdagi tuzilmagan ma'lumotlarni qayta ishlash, foydali xulosalar chiqarish hamda inson hissiyotlarini aniqlash imkonini bermoqda. Ayniqsa, sentiment tahlili mijozlar fikrlarini o'rganish, ijtimoiy tarmoqlardagi munosabatlarni tahlil qilish va tashkilotlarning qaror qabul qilish jarayonlarini takomillashtirishda muhim ahamiyat kasb etadi. NLP texnologiyalarining rivojlanishi ta'lim, tibbiyot, bank-moliya, media monitoring, davlat xizmatlari va biznes jarayonlarida samaradorlikni oshirishga xizmat qilmoqda. Shu bilan birga, NLP tizimlarini yaratishda katta hajmdagi sifatli ma'lumotlar bazasi, til resurslari va etik hamda xavfsizlik masalalariga alohida e'tibor qaratish zarur. Kelajakda NLP inson va texnologiya o'rtasidagi o'zaro aloqani yanada tabiiylashtirib, yangi innovatsion imkoniyatlarni yaratishda muhim o'rin egallaydi. [7:210]

References:

1. Jurafsky D., Martin J. H. Speech and Language Processing: An Introduction to Natural Language Processing, Computational Linguistics, and Speech Recognition. – 3rd ed. – Draft version. – Stanford : Stanford University, 2023. – 716 p.[1:12]
2. Manning C. D., Schütze H. Foundations of Statistical Natural Language Processing. – Cambridge : MIT Press, 1999. – 680 p.[7:15]
3. Bird S., Klein E., Loper E. Natural Language Processing with Python. – Sebastopol : O’Reilly Media, 2009. – 504 p.
4. Goodfellow I., Bengio Y., Courville A. Deep Learning. – Cambridge : MIT Press, 2016. – 775 p.[5:60]
5. Vaswani A., Shazeer N., Parmar N., et al. Attention Is All You Need // Advances in Neural Information Processing Systems (NeurIPS). – 2017. – Vol. 30. – P. 5998–6008.[4:120]
6. Devlin J., Chang M. W., Lee K., Toutanova K. BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding // Proceedings of NAACL-HLT. – 2019. – P. 4171–4186. [1:78]
7. Goldberg Y. Neural Network Methods in Natural Language Processing. – San Rafael : Morgan & Claypool Publishers, 2017. – 309 p.[7:210]