

KOMPYUTER LINGVISTIKASI NUQTAYI NAZARIDAN O‘ZBEK TILIDA NEGATSIYA VOSITALARINI AVTOMATIK ANIQLASH MASALALAR

Toshxonov Lochinbek Tursunboyevich

Andijon qishloq xo‘jaligi va agrotexnologiyalar instituti,
Xorijiy tillar kafedrası katta o‘qituvchisi, f.f.f.d. (PhD)

Annotatsiya. Maqolada o‘zbek tilida negatsiya vositalarining (morfologik, sintaktik, leksik, pragmatik) avtomatik aniqlanishi kompyuter lingvistikasi nuqtayi nazaridan tahlil qilinadi. Ingliz tilidagi NLP tajribalari bilan qiyosiy o‘rganilib, o‘zbek tiliga mos algoritmik yechimlar taklif etiladi. Natijada maxsus “O‘zbek tilining Negatsiya Korpusi”ni yaratish zarurligi asoslab beriladi.

Kalit so‘zlar: kompyuter lingvistikasi, negatsiya, tabiiy tilni qayta ishlash (NLP), affiksial inkor, sintaktik inkor, leksik inkor, pragmatik inkor, korpus lingvistikasi, mashinaviy o‘qitish, algoritmik tahlil

ВОПРОСЫ АВТОМАТИЧЕСКОГО ВЫЯВЛЕНИЯ СРЕДСТВ НЕГАЦИИ В УЗБЕКСКОМ ЯЗЫКЕ С ТОЧКИ ЗРЕНИЯ КОМПЬЮТЕРНОЙ ЛИНГВИСТИКИ

Тошхонов Лочинбек Турсунбоевич

Андижанский институт сельского хозяйства и агротехнологий,
Кафедра иностранных языков, старший преподаватель, к.ф.н.

Аннотация. В статье рассматривается автоматическое выявление средств негации (морфологических, синтаксических, лексических, прагматических) в узбекском языке с точки зрения компьютерной лингвистики. Проведено сопоставительное исследование с практиками NLP в английском языке, предложены алгоритмические решения, адаптированные к узбекскому языку. В

результате обоснована необходимость создания специализированного «Корпуса негации узбекского языка».

Ключевые слова: компьютерная лингвистика, негация, обработка естественного языка (NLP), аффиксальная негация, синтаксическая негация, лексическая негация, прагматическая негация, корпусная лингвистика, машинное обучение, алгоритмический анализ.

COMPUTATIONAL LINGUISTICS PERSPECTIVE ON THE AUTOMATIC IDENTIFICATION OF NEGATION MEANS IN THE UZBEK LANGUAGE

Toshkhonov Lochinbek Tursunboyevich

Andijan Institute of Agriculture and Agrotechnologies,
Department of Foreign Languages, Senior Lecturer, PhD

Abstract. The article analyses the automatic identification of negation means (morphological, syntactic, lexical, pragmatic) in the Uzbek language from the perspective of computational linguistics. Comparative study with NLP practices in English is conducted, and algorithmic solutions adapted to the Uzbek language are proposed. As a result, the necessity of creating a specialised “Negation Corpus of the Uzbek Language” is substantiated.

Keywords: computational linguistics, negation, natural language processing (NLP), affixal negation, syntactic negation, lexical negation, pragmatic negation, corpus linguistics, machine learning, algorithmic analysis.

So‘nggi yillarda kompyuter lingvistikasi va sun‘iy intellekt texnologiyalari jadal rivojlanib bormoqda. Xususan, tabiiy tilni qayta ishlash (Natural Language Processing — NLP) yo‘nalishida mashina tarjimasi, matnni avtomatik tahlil qilish, sentiment tahlili va nutqni tanish kabi masalalar dolzarb ahamiyat kasb etmoqda. Ushbu jarayonlarda tilning asosiy grammatik kategoriyalaridan biri bo‘lgan inkor alohida o‘rin tutadi, chunki negatsiya gapning umumiy semantikasini o‘zgartiradi va

matn mazmunini aniqlashda muhim rol o'ynaydi. O'zbek tilida negatsiya ko'plab shakllarda ifodalanadi: -ma/-me qo'shimchasi, emas, yo'q, hech kim, hech qachon kabi leksik birliklar, shuningdek, pragmatik darajadagi inkor shakllari. Ingliz tilida esa negatsiya asosan not va yordamchi fe'llar orqali amalga oshiriladi. Ushbu tafovutlar mashina tarjimai va NLP tizimlari uchun muhim muammo tug'diradi. Shu sababli tadqiqotning asosiy maqsadi o'zbek tilida negatsiya vositalarini avtomatik aniqlash usullarini ilmiy-nazariy jihatdan asoslashdan iboratdir. Bu borada bir qator vazifalar belgilangan: negatsiya tushunchasini NLP kontekstida aniqlash, o'zbek tilidagi negatsiya vositalarining tasnifini ishlab chiqish, ingliz tilidagi mavjud tajribalarni tahlil qilish, o'zbek tiliga mos algoritmik yechimlar taklif etish va negatsiya tahlilining amaliy qo'llanish sohalarini ko'rsatish. Tadqiqot ob'ekti sifatida o'zbek tilidagi negatsiya birliklari tanlandi, predmet sifatida esa negatsiyaning avtomatik tahlil mexanizmlari va algoritmlariga e'tibor qaratildi. O'rganish jarayonida korpus lingvistikasi yondashuvi asosida o'zbek tilidagi matnlar korpusidan keng ko'lamli misollar keltirilib, ular negatsiya vositalari nuqtayi nazaridan tahlil qilindi. Morfologik tahlil yordamida fe'llarda uchraydigan negatsiya qo'shimchasi -ma/-me ning shakllanish va qo'llanish xususiyatlari aniqlashtirildi. Qiyosiy tahlil orqali ingliz tilidagi NLP tizimlarida negatsiyani aniqlash bo'yicha mavjud tajribalar o'rganilib, ularning o'zbek tiliga mosligi baholandi. Algoritmik yondashuv asosida esa negatsiyani avtomatik belgilash uchun qoidaviy va mashinaviy o'qitishga asoslangan modellar ishlab chiqish taklif etildi. Tadqiqot manbalari sifatida Google Translate, DeepL, HuggingFace modellaridagi amaliy tajribalar, shuningdek, O'zbek til korpusi va ingliz tilidagi Brown Corpusdan foydalanildi. O'zbek tilida negatsiya vositalarining turli darajalarda ifodalanishi tasnif qilinib, morfologik inkor fe'llarga qo'shiladigan -ma/-me qo'shimchasi orqali (bormadi, qilmaydi), sintaktik inkor emas va yo'q birliklari yordamida (Bu to'g'ri emas), leksik inkor hech kim, hech qachon, hech narsa kabi birliklar orqali, pragmatik inkor esa kinoya va ironiyada yashirin shaklda amalga oshishi ko'rsatib berildi. Ingliz tilida esa negatsiya ko'proq not va yordamchi fe'llar orqali belgilanadi,

masalan, sentiment tahlilida not good kabi ifodalar semantik inversiya orqali bad tarzida talqin qilinadi, ikki karra inkor (double negation) esa NLP tizimlari uchun murakkab hodisa sanaladi. O'zbek tilida negatsiyani avtomatik aniqlashda esa bir qator amaliy muammolar mavjud: affiksial negatsiya (bor → bormadi)ni avtomatik ajratish qiyin, chunki negatsiya morfemasi fe'ning boshqa grammatik qo'shimchalari bilan qo'shib ketadi; leksik negatsiya (hech kim, hech qachon) gapning boshqa qismlari bilan murakkab bog'lanib, semantik noaniqlik keltirib chiqaradi; pragmatik negatsiya (ironik yoki evfemistik shakllar) esa avtomatik tahlilda deyarli aniqlanmaydi. Shu bois, negatsiyani aniqlash uchun bir nechta algoritmik yechimlar ishlab chiqilishi zarur: qoidaviy yondashuv morfologik negatsiyani aniqlashda maxsus affiks lug'atini yaratishni talab etadi; mashina o'rganishi sintaktik kontekst asosida negatsiyani belgilashga yordam beradi (masalan, emas bilan keluvchi predikativ konstruktsiyalarni aniqlash); semantik modellashtirish leksik negatsiya birliklarini turli vaziyatlarda avtomatik belgilash imkonini beradi; korpusga asoslangan yondashuv esa o'zbek tilidagi matnlar asosida negatsiya belgilarini qamrab oluvchi maxsus korpus yaratishni taqozo etadi. Ayni paytda o'zbek tilida negatsiyani avtomatik aniqlash masalasi hali to'liq ishlab chiqilmagan. Ingliz tilidagi sentiment tahlili va mashina tarjima tizimlari o'zbek tiliga moslashtirilganda ko'plab xatoliklar kuzatiladi. Masalan, Hech kim kelmadi gapini Nobody came tarzida to'g'ri tarjima qilish mumkin, biroq Nobody didn't come kabi xatoliklar ham uchraydi. Bundan tashqari, o'zbek tilidagi affiksial negatsiya NLP tizimlari uchun alohida murakkablik tug'diradi, chunki bormadim so'zida shaxs, zamon va negatsiya birgalikda ifodalanadi, ingliz tilida esa ular alohida so'zlarga bo'linadi: I did not go. Shu bois, mazkur muammolarni hal qilish uchun maxsus "O'zbek tilining Negatsiya Korpusi"ni yaratish va uni mashinaviy o'qitish modellarida qo'llash zarur bo'ladi.

Xulosa o'rnida shuni ta'kidlash joizki, o'zbek tilida negatsiya morfologik, sintaktik, leksik va pragmatik darajalarda ifodalanadi. Ingliz tilidagi NLP tajribalari negatsiyani aniqlashda samarali natija bergan bo'lsa-da, ular o'zbek tilining tuzilishiga to'liq mos kelmaydi. Xususan, affiksial negatsiya (*-ma/-me*)ni avtomatik

tahlil qilish murakkab bo'lib, uni amalga oshirish uchun maxsus algoritmlar ishlab chiqilishi zarur. O'zbek tilida negatsiya vositalarining to'liq tasnifi va maxsus korpusi yaratilishi esa mashina tarjimasi, sentiment tahlili va nutqni avtomatik anglash sohalarida sezilarli yutuqlarga erishish imkonini beradi. Ushbu tadqiqot natijalari kompyuter lingvistikasi va tabiiy tilni qayta ishlash yo'nalishida yangi ilmiy izlanishlarni shakllantirish uchun nazariy hamda amaliy asos bo'lib xizmat qiladi.

Foydalanilgan adabiyotlar.

1. Cameron-Faulkner, T., Lieven, E., & Tomasello, M. (2007). A construction-based analysis of child negation in English. *Journal of Child Language*, 34(2), 251–282.
2. Horn, L. R. (2001). *A Natural History of Negation*. Stanford: CSLI Publications.
3. Dahl, Ö. (1979). Typology of sentence negation. *Linguistics*, 17(1-2), 79–106.
4. Morante, R., & Blanco, E. (2012). SemEval-2012 Task 7: Automatic detection of negation scope and focus. *Proceedings of the First Joint Conference on Lexical and Computational Semantics*. Montréal.
5. Toshxonov L.T. Формирование отрицания с префиксами. – Rossiya, «Question of Science and Education», №15(99), 2020. – Б.58–64.
6. Toshxonov L.T. Badiiy matnda inkorni ifodalashning lingvistik vositalari. – Rossiya, «Экономика и социум», №10(89), 2021. – Б.5.