



MATNNI USLUBIYATINI ANIQLASH

Toshniyozov F.O', Jiyanov O.P, Djurakulov D

Muhammad al-Xorazmiy nomidagi TATU Samarqand filiali assistenti

farruxtoshniyozov00@gmail.com

Muhammad al-Xorazmiy nomidagi TATU Samarqand filiali katta o'qituvchisi

01051987oybek@gmail.com

Muhammad al-Xorazmiy nomidagi TATU Samarqand filiali talabasi

Annotatsiya: *Bu maqolada lingvistik, statistik ma'lumotlarga va til vositalarining kombinatorikasiga asoslangan matematik, katta hajmdagi matnlarni tahlil qilishga mo'ljallangan usullar va matnlarni uslubiyatini aniqlash haqida ma'lumotlar keltirilgan.*

Kalit so'zlar: *stilometriya, mas, dedi, so'zlar, deydi, men, meni, mening, va, qaysi, bo'lsa, deb, keyin, lekin, qayerda,*

Matn nima haqida ekanligini aniqlashni **stilometriya** fani o'rganadi. Bunda matndagi gap bo'laklaridan foydalaniladi.

Biz ilmiy va badiiy asarlarni mazmunan bir-biridan qanday farq qilishini yaxshi bilamiz va ilmiy kitobni badiiy kitobdan bema'lol ajrata olamiz. Xo'sh, bunday asarlar til jihatidan farq qiladimi? Agar biz rus tilidagi ilmiy va badiiy matnlarda ismlar, fe'llar va olmoshlar sonini alohida hisoblasak, ilmiy matnlarda badiiy adabiyotga qaraganda taxminan ikki baravar kam fe'l borligini va taxminan bir yarim baravar kam olmoshlar borligini ko'ramiz mumkin. Lekin ushbu asarlarda taxminan 20% otlar ko'pligini uchratish mumkin. Gazeta matnlari va teatrlashtirilgan spektakllarning matnlari taxminan bir xil tarzda farqlanadi. Bularning barchasi oddiy bo'lib, turli janrdagi asarlarda gap bo'laklarining har xil nisbati qo'llaniladi: qayerda hodisa ko'p bo'lsa, fe'l ham ko'p bo'ladi; tavsiflar ko'p bo'lgan joyda otlar ko'proq bo'ladi; odamlar o'rtasida ko'proq dialog bo'lganda ko'proq olmoshlardan foydalaniladi.

Stilometriya (*stylometry*) ana shunday tamoyillarga asoslanadi. Turli yozuvchilar tomonidan turli janrlarda yozilgan matnlarning uslublarini Kvantitativ o'rganish aynan stilometriyaga asoslangan bo'ladi. Albatta, stilometriyada faqat gap qismlari hisobga olinmaydi. Tadqiqotchilar tilning hamma uchun bir xil bo'lgan umumiy xususiyatlaridan farqli o'laroq, statistik ahamiyatga ega bo'lishi mumkin bo'lgan nutq xususiyatlarini topishga harakat qilmoqdalar .



Shu jihatdan turli xil usullar mavjud bo'lib, ular: lingvistik, statistik ma'lumotlarga va til vositalarining kombinatorikasiga asoslangan matematik, katta hajmdagi matnlarni tahlil qilishga mo'ljallangan usullardir.

Stilometriya matnlarni tasniflash, janrlarni aniqlash, anonim va taxallusli asarlarning muallifligini aniqlash uchun ishlatilishi mumkin.

Stoni Bruk universiteti (AQSh) tadqiqotchilari stilometriya va mashinani o'rganish algoritmlaridan foydalanib, yangi kitoblar va hattoki kino skriptlarining qanchalik mashhurligini bashorat qilishdi. Shu yerda ular buni qanday qilishlarini ko'rib chiqaylik. Avvaliga olimlar 40 000 ga yaqin kitob saqlanadigan ingliz tilidagi *Project Gutenberg* (www.gutenberg.org) elektron kutubxonasi saytida turli janrdagi kitoblarni yuklab olish statistikasini ko'rib chiqdilar. So'ngra ular sarguzasht, sir, fantastika, tarixiy romanlar, romantika, she'riyat, ilmiy fantastika va qissa janrlari bo'yicha kitob va ssenariylarning eng ommabop 50 ta matni va mashhur bo'lmagan 50 ta matni tanlab oldilar. Ular nazorat ostidagi o'rganish algoritmi uchun o'quv namunasini yaratishdilar.

Kvantitativ tadqiqot uchun lingvistik parametrlar sifatida quyidagilar tanlandi:

- a) lug'at: unigramlar va bigrammalar, ya'ni alohida so'zlar va ikki so'z birikmalari;
- b) gap bo'laklari: matnlardagi so'zlarni gap bo'laklari bo'yicha taqsimlash;
- v) oddiy grammatik belgilar: eng oddiy sintaktik tuzilmalarning asarlarida tarqalishi;
- d) murakkab grammatik belgilar: ayrim murakkab sintaktik tuzilmalarning asarlarda tarqalishi;
- e) ohang va ma'nolar: his-tuyg'ularni bildiruvchi so'zlar va qo'shimcha hissiy yoki baholovchi ma'noga ega bo'lgan so'zlar.

Keyinchalik, ta'lim berish LibLinear mashinani o'rganish to'plamidan foydalangan holda o'quv namunasi matnlardan dastlabki 1000 ta jumla va xuddi shu elektron kutubxonadan test namunasi asosida baholash asosida o'tkazildi. Natija ancha yuqori bo'lib, asarlarning mashhurligini baholashning aniqligi 74% hamda sarguzasht janrida esa 84% ga yetgan.

Tadqiqotchilar asarlarning mashhurligiga ta'sir ko'rsatgan lingvistik parametrlarni tahlil qilishdi. Misol uchun, mashhur matnlardagi unigramlar turkumida shunday inglizcha so'zlar mavjud edi (so'zlar juda oddiy va joyni tejash uchun ularni tarjimasiz qoldiramiz): *emas, dedi, so'zlar, deydi, men, meni, mening, va, qaysi, bo'lsa, deb, keyin, lekin, qayerda, nima, kimdan, beri, qachon, yuqoriga, ichiga, tashqariga, keyin, ichida, ichidan, tanildi, esladi.*



Mashhur bo‘lmagan asarlarda quyidagi so‘zlar ustunlik qildi: *hech, xatar, yomon, yomonroq, qullar, qul, qattiq, o‘ldirilgan, ko‘kargan, og‘ir, qamoqxona, yuz, qo‘l, tana, terilar, xona, plyaj, bay, tepaliklar, prospekt, qayiq, eshik, xoilyayman, ketdi, oldi, vada, yig‘ladi, qichqirdi, sakradi, yaltiradi, turtki, juda, nafassiz, zarracha, mutlaqo, mukammal, istaklar, ishlar.*

Keyingi ishlarda olimlar o‘qitilgan algoritmdan Gutenberg loyihasining raqamli kutubxonasida bo‘lmagan kitoblarning mashhurligini tekshirish uchun foydalandilar. Kitobxonlar orasida muvaffaqiyat bilan qabul qilingan yozuvchilar, jumladan, nufuzli adabiy mukofotlarga sazovor bo‘lgan ayrim adiblar, M. Servantes, Ch.Dikkins, E.Xeminguey, J.Apdayk, R.Stivenson, F.Rotlar ishlaridan ham 10 ta kitob tanlab olindi.

Unigramlarda ishlaydigan algoritmi bu kitoblarning barchasi keng ommaga mashhur bo‘lishi kerakligini ko‘rsatdi. Murakkab sintaktik tuzilmalar ustida ishlaydigan algoritmi 10 ta kitobdan 7 tasi muvaffaqiyatli degan xulosaga keldi. Misol uchun, Xemingueyning “Chol va dengiz” qissasi uncha mashhur bo‘lmagan kitoblar toifasiga kirdi, chunki bu asarda muallif asosan sodda gaplardan foydalangan.

Tadqiqotchilar ssenariy janri bo‘yicha bir xil yuqori ko‘rsatkichlarga erishmagan bo‘lishsa-da, endilikda ularning algoritmi nafaqat filmlar muvaffaqiyatini oldindan aytish, balki ilmiy ishlar muvaffaqiyatini tahlil qilish uchun ham takomillashtirilishi mumkinligiga ishonchlari komil. Olimlarning ta’kidlashicha, ishlab chiqilgan metodika yordamida har qanday matnning “o‘qilishi”ni, ya’ni oddiy o‘quvchi nuqtai nazaridan qanchalik sodda va tushunarli yozilganligiga ham baho berish mumkin. Bu imkoniyat amaliy ahamiyatga ega bo‘lib, insho yoki kurs ishingizni qanday qilib o‘qituvchiga yanada qiziqarli bo‘lishi va uning bahosini qanday oshirishningiz mumkinligini ayta oladigan matn muharririni tasavvur qilishingizga yordam beradi.

Foydalanilgan adabiyotlar

1. Qarshiyev A.B., Tursunov M.S. “Matnlarni statistik tahlili uchun amaliy dasturlar majmuasi” //“Axborot va telekommunikatsiya texnologiyalari muammolari” ilmiy-texnik konferensiyaning ma’ruzalar to‘plami, II qism. –Toshkent, 2015. Sahifalar 35-37.
2. Mingboyev Ulug‘bek S. A., Djumayev Sindor O. N. E. TA’LIMDA FOYDALANILADIGAN MOBIL ILOVALARNING AFZALLIKLARI VA KAMCHILIKLARI //International Journal of Contemporary Scientific and Technical Research. – 2023. – C. 279-282.
3. Djumayev, S. N., D. A. Shodmonov, and N. E. Ortiqova. "Ta'limda mobil



illovalardan foydalanish imkoniyatlari: Possibilities of Using Mobile Applications in Education." INTERNATIONAL JOURNAL OF THEORETICAL AND APPLIED ISSUES OF DIGITAL TECHNOLOGIES 4.2 (2023): 90-95.