



## MATNLI MA'LUMOTLARNI INTELLEKTUAL QAYTA ISHLASH VA TASNIFLASH

---

*Choryorqulov G'iyos Husan o'g'li - assistent*

*O'zbekiston Milliy universitetining Jizzax filiali, O'zbekiston*

*[choryorqulov@jbnuu.uz](mailto:choryorqulov@jbnuu.uz)*

*Sa'dullayeva Sabina Rizamat qizi - talaba*

*O'zbekiston Milliy universitetining Jizzax filiali, O'zbekiston*

*[sadullayeva9030@gmail.com](mailto:sadullayeva9030@gmail.com)*

**Annotatsiya.** Matnli ma'lumotlarni qayta ishlash, matnning tuzilishini va uslubini o'zgartirish, matnning grammatik va ortografiyasini tekshirish yoki boshqa matn tahrirlash jarayonlarini tushunishni anglatadi. Matnli ma'lumotlarni qayta ishlashning bir necha turli usullari mavjud. Ular o'zgartirishlar, tuzatishlar, tarjima va redaktirovka kabi jarayonlarni o'z ichiga oladi. O'zgartirishlar ko'pincha, matnning ma'nolarini va uslubini o'zgartirishni anglatadi. Bunda, matn o'zgartirilishi zarur bo'lgan ma'lumotlar qo'shilishi mumkin yoki matnning o'ziga o'ziga qarab o'zgartirishlar kiritilishi mumkin.

**Kalit so'zlar:** matnli ma'lumotlar, axborot texnologiyalari (AT), sun'iy tafakkur tizimlari (STT), ma'lumotlarni intellektual tahlili (MIT).

Matnli ma'lumotlarni qayta ishlash, matn tahrirlashda juda muhim bo'lib, matndagi xatolar va noaniq ma'lumotlarni bartaraf qilish yoki matning ko'rsatilgan ma'noga to'g'ri kelishini ta'minlash yordam beradi.

Ma'lumotlar intellektual tahlili murakkab va ko'p bosqichli jarayon bo'lib, o'zida turli bosqichlarda ma'lumotlarni yig'ish, dastlabki ishlov berish va tahlil qilish jarayonlarini mujassamlashtiradi. Ma'lumotga dastlabki ishlov berish fazasida turli formatdagi ma'lumotlarni tartiblash, bir formatga keltirish uchun bir qator algoritmlarni qo'llashni talab qiladi. Ishlov berish fazasi ma'lumotlar integratsiyasi, ularni tozalash hamda tavsiflovchi belgilarni ajratishni qamrab oladi [1]. Ba'zi hollarda ma'lumotlar tavsifidagi xarajatni qisqartirish uchun belgilar fazosini samarali tanlashdan foydalanish mumkin. Ma'lumotlar qulay formatga keltirilgach, ularga ishlov berishda turli ko'rinishdagi algoritmlardan foydalanish mumkin bo'ladi.

Ma'lumotlarni intellektual tahlilash muammolari hal etilishida guruh masalalarda: timsollarni aniqlash, klasterizatsiya, tasniflash va chetlashishlar tahlili



kabi xilma-xilliklar uchun ssenariylarni takroriy qo‘llash kuzatiladi. Ma’lumotlarni intellektual tahlilining muayyan masalalarida yakuniy yechim ko‘rinishi uchun turli bloklarni birlashtirish yoki yangilangan algoritmlardan foydalanish mahorati talab etiladi [2].

Ma’lumotlarga ishlov berish jarayonini to‘g‘ri tashkil etilishi muhim ahamiyat kasb etadi. Buning uchun ham ma’lumotlarni to‘g‘ri tayyorlab olish, saralash ishlov berishdagi yutuq kafolati hisoblanadi, chunki analitik algoritmlar kirish ma’lumotlarining sifatiga o‘ta sezuvchan bo‘ladi. Ma’lumotlarni intellektual tahlili jarayoni analitik algoritmlarni bevosita qo‘llash uchun yaroqsiz bo‘lgan shakldagi turli xil manbalardan ishlov berilmagan ma’lumotlarni ajratib olish talab etiladi. Shuning uchun birlamchi ma’lumotlardan muhim belgilarni ajratib olish uchun ko‘plab usullar taklif qilingan. Olingan ma’lumotlarda sezilarli darajada xatolik, beqarorlik, ortiqchaliklar va yo‘qotilgan qiymatlar uchrashi mumkin. Yo‘qotilgan qiymatlarni kiritish yoki ma’lumotlardagi noaniqliklarni bartaraf etish uchun turli analitik usullar va ma’lumotlarni yig‘ish vositalari mavjud [3].

Yechilishi lozim bo‘lgan yana bir muhim masalalardan biri ma’lumotlarning xilma-xilligi hisoblanadi, ularda analitikasi turli xil atributlar uchraydi. Shu sababli ma’lumotlarni intellektual tahlili algoritmlarining bevosita qo‘llanilishi murakkab kechadi. Bunday holatlarda ma’lumotlar turining portativligi muhim bo‘lib, unda ba’zi bir atributlar to‘plami oldindan belgilangan formatga keltiriladi. Ko‘p o‘lchovli format odatda oddiyligi tufayli afzal ko‘riladi. Aslida har qanday ma’lumot turini ko‘p o‘lchovli ko‘rinishga keltirib olish mumkin.

Sun‘iy tafakkur tizimlari (STT) axborot texnologiyalarida mutahassis bo‘lmagan foydalanuvchiga intellektual (qiyin shakllanuvchi, evristik) masalalarni kompyuter bilan cheklangan tabiiy tilda muloqot qilishi oqibatida qo‘yish va yechish imkoniyatini yaratadi [4].

Qaralayotgan masalalar sinfiga tegishlilarning asosiy belgilari – ularning qiyin yoki umuman shakllantirishning mumkin emasligida, bu holatda determinanlashgan algoritmlar usullari bilan bir qatorda evristik usullar, shu bilan birga ekspert bilimlari va malakalari asosida ushbu masalalarni yechish uchun qo‘llanilishi bilan ajralib turadi. Shuning uchun bunday “odam-mashina” tizimlari bilimlarga asoslangan tizimlar deb yuritiladi. Bunday amaliy “odam-mashina” ko‘rinishidagi keng tarqalgan suni’y tafakkuriy tizimlari vakili “Ekspert tizimlari” hisoblanadi [5].

Axborot texnologiyalarini (AT)ni intellektuallashtirishning maqsadi ularning hususiyatlarini EHMda ifodalash va formal bo‘lmagan bilimlarni qayta ishlashdan



iborat. Formal bo‘lmagan tushunchasi ostida qat’iy matematik vositalar bilan ifodalana olmaydiganlarini tushunamiz [6].

**ATni intellektuallashtirish** ularni xususiyatlarga ajratib chiqish hisoblanadi:

- O‘qitish jarayonida to‘plangan bilimlar bilan bog‘liq masalani yechishda qo‘llanilishi mumkinligi;
- Xulosalash qoidalarini generasiya qilishi imkoniyati;
- Axborotni umumlashtirish imkoniyati;
- Qabul qilingan vaqt, aniqlik va narh qiymatlari bilan bog‘liq kuchsiz tizimlashtirilgan masalalarni yechishda **ST vositalaridan** foydalanish imkonining mavjudligi [7].

ATni intellektuallashtirish STning asosiy vositalarini qo‘llashni nazarda tutadi, ularning asosiylari quyidagilar:

- Fikr yuritish mexanizmlari;
- Hatti-harakat modellarri;
- Evolyusiyasi modellari;
- parallel i assosiativ hisoblar modellari;
- samaradorlik modellari;
- tabiiy-biologik tanlash, muvofiqlashtirish va moslashtirish prosedurasini modellari [8].

ATni intellektuallashtirishning asosiy paradigmalari quyidagilar hisoblanadi:

- Evolyusion modellashtirish i dasturlash.
- Neyron to‘rli modellashtirish i hisoblash.
- Noravshan mantiq va noravshan to‘plamlar.
- Haqiqatga yaqin fikrlash va mantiqiy xulosalash.
- Xaos nazariyasi.

Evolyusion modellashtirish va dasturlash (Evolution Modelling and Programming), neyron to‘rli modellashtirish i hisoblash (Neural Network Modelling and Computing), genetik algoritmlar (Genetic algorithms), shu bilan birga DNK – hisoblash (DNA Computing), hujayra avtomatlari (Cellular Automata) va chumoli algoritmlari (Ant Colony Algorithms) Tabiiy hisoblash (Natural Computing) nomini olgan sinfni hosil qilishadi [9, 10].

Ma’lumotlarga ishlov berishda ma’lumotlar intellektual tahlilining yana bir apparati klaster tahlili hisoblanadi. Ma’lumotlarni tahlilashda klasterlash masalasini hal etish uchun keng turdagi algoritmlar ishlab chiqilgan, bulardan vakilli asoslash, iyerarxik, ehtimoliy, zichlikka asoslangan, grafikka asoslangan va matrisali asoslash



kabi bir qator usullarini keltirish o'tish o'rinli. Barcha usullar odatda algoritmi klasterlar soni, zichligi yoki matrisaviy darajasini asoslash kabi ayrim parametrlarni ajratishni talab qiladi [11,12]. Vakilga asoslangan hamda ehtimoliy usullar klasterlarning shakliga cheklov qo'ymaydi, biroq klasterlarning zichligi o'zgarishiga yaxshiroq moslashadi. Boshqa tomondan, aglomerativ va zichlikka asoslangan usullar klasterlar shakliga moslashuvchanroq bo'ladi, biroq klasterlarning o'zgaruvchan zichligiga moslashishi jarayoni qiyin kichadi [13, 14]. Grafikka asoslangan usullar turli shakldagi va zichlikdagi eng yaxshi moslashuvni ta'minlaydi, o'z navbatida amalga oshirishga resurs ko'proq talab qiladi. Klasterlash uchun tashqi va ichki tekshirish mezonlari mavjud bo'lsada, ko'pincha turli xil algoritmlar noto'g'ri yoki asosiy ma'lumotlardagi ichki klasterlarni aniq aks ettira olmaydi. Bunday qiymatlardan ehtiyotkorlik bilan foydalanish talab etiladi [15].

Ma'lumotlar intellektual tahlilining asosiy masalalaridan biri bu tasniflash masalasi hisoblanadi. Ushbu masala ma'lumotlarni tavsiflovchi  $X_i$  – belgilar bilan berilgan  $S_i$  – o'quv tanlanma obyektlari asosida kiruvchi obyektning belgilangan sinflardan biriga tegishlilikini aniqlashni talab etadi. Ma'lumotlar tasnifi uchun turli xil modellar taklif qilingan [16].

Ma'lumotlarni intellektual tahlili (MIT) o'z tarkibiga asosan olti turdagi masalalarni oladi – tasniflash (klassifikatsiya), klasterizatsiya, regressiya, biror vaqt oralig'ida bashorat qilish, assosiasiya, tartibni aniqlash hamda chetlashishlarni tahlil qilish hisoblanadi. MITning eng keng tarqalgan masalalaridan biri klassifikatsiya masalasi hisoblanib, uning yordamida predmet soha bo'yicha obyektning tavsiflovchi belgilari yordamida u yoki bu sinfga ta'luqliligini aniqlanadi. Klasterizatsiya avtomatik klassifikatsiya bo'lib, unda oldindan sinflar berilmagan bo'ladi, obyektlar klasterlarga avtomatik tarzda ajratiladi va ular ajratilgan klasteriga (sinfiga) tegishli deb olinadi. Klasterizatsiya usullari bilan aksariyat hollarda katta hajmli turli tipli ma'lumotlarni tadqiq qilish bilan bog'liq masalalarni yechish uchun xizmat qiladi [17]. Regression tahlil o'zgaruvchilar o'rtasidagi munosabatlar miqdoriy ifodalangan holda ularning ayrim kombinatsiyalari ko'rinishida ishlatiladi. Asosan standart statistik usullar qo'llaniladi. Bulardan biri chiziqli regressiyadir [18]. Vaqt ketma-ketliklarini bashorat qilish boshlang'ich ma'lumotlarni yoki yarim qayta ishlangan ma'lumotlarni tahlil qilish asosida bo'lg'usi bashorat o'zgaruvchilari qiymatini baholash imkonini beradi. Assotsiatsiya bir nechta hodisa va faktlarning bog'liqligini topish usuli hisoblanadi. Tartibni keltirib chiqarish tahlil qilinayotgan ma'lumotlarni vaqt tartibi zanjirini aniqlash hamda sochilib yotgan vaqt tartiblarini



guruhlashtirishni amalga oshiradi. Ya'ni qo'yilgan maqsad yoki oldindan kutilgan yechimdan og'ishni aniqlashtiradi [19].

Tasniflash masalasi turkumiga o'qitish orqali klassifikasiyalash, avtomatik klassifikasiya (taksonomiya – klasterizasiya), bashorat qilish kabi boshqaruvning asosiy elementlari kiradi. Mazkur predmet soha bo'yicha dunyoning ko'zga ko'ringan olimlari tadqiqot olib borgan. Tasniflash masalasining kelib chiqish xarakteri tadqiq qilinayotgan obyektga qarab turlicha bo'ladi [20]. Jumladan, XXI asrda yuz bergan axborot resurslaridan cheklanmagan miqdorda foydalanishga imkon bergan veb-texnologiyalarning rivojlanishi katta-katta ma'lumotlar jamlanmasiga ega turkum axborot resurslarini vujudga keltirdi hamda katta tezlik bilan boyitilishiga va rivojlantirilishiga olib keldi. Mazkur tendensiya foydalanuvchilarning keng qatlamida saralanmagan ma'lumotlardan foydalanish, katta miqdordagi vaqtlarini ma'lumotlar tahliliga sarflashlariga olib keldi. Shu tufayli yuqorida nomi zikr qilingan MIT yo'nalishini kelib chiqishiga sabab bo'ldi. Matnli hujjatlardan tashkil topgan ma'lumotlarni intellektual tahlillashni amalga oshirish uchun samarali tasniflash (matnni belgilari asosida toyifalashni amalga oshirish) mexanizmini ishlab chiqish dolzarb muammolar sirasiga kirdi [21]. Quyida matnni tasniflash borasida amalga oshirilgan tadqiqotlar tahlili keltirilgan bo'lib, navbatdagi bosqich tadqiqotlar uchun ustqurmani shakllantirib olishga erishiladi. Matnli ma'lumotlarni tasniflash bo'yicha ishlab chiqarish – sanoat industriyasida matnni tahlillash orqali masalalarni hal etish borasida Etto va boshqalar tomonidan matn tahlili dasturlarini qo'llash masalalarini ko'rib chiqilgan. Mualliflar tomonidan mazkur predmet soha bo'yicha uch yo'nalish: dastur konteksti, qo'llaniladigan usul va yondashuvlar hamda amalga oshirish tartibidan foydalanilgan. Ular real sanoat sektorida foydalaniladigan matnni tahlil qilish usullarini qo'llashda yuzaga keladigan muammo va cheklovlarni aniqlagan. Shuningdek vujudga keladigan muammolarni bartaraf etish va ulardan amaliy masalani yechishda foydalanish yondashuvlarini (takliflar ko'rinishda) shakllantirdi [22].

### **Foydalanilgan adabiyotlar:**

1. Nizomiddin N. et al. TA'LIMDA DASTURLASH JARAYONINI BAHOLASHGA ASOSLANGAN AVTOMATLASHTIRILGAN TIZIMNI TADBIQ ETISH //International Journal of Contemporary Scientific and Technical Research. – 2023. – C. 24-28.



2. Choryorqulov G. H., & Qosimov NS (2023) //ELEKTRON JADVAL MODELINING TAVSIFLANISHI. PEDAGOGS Jurnal. – Т. 30. – №. 3. – С. 67-73.
3. Чорркулов Г., Норматов Н., Мамараимов А. Роль анализа текстовых связей в электронных документах в информационной безопасности //Информатика и инженерные технологии. – 2023. – Т. 1. – №. 1. – С. 67-71.
4. Норматов Н., Мамараимов А. Ta'lim tizimida baholash tizimini avtomatlashtirishni joriy etish jarayonlari va foydalanish metodlari //Информатика и инженерные технологии. – 2023. – Т. 1. – №. 2. – С. 356-359.
5. Мамараимов А., Чорёркулов Г., Норматов Н. Tanib olish modullarini dasturiy amalga oshirish //Информатика и инженерные технологии. – 2023. – Т. 1. – №. 2. – С. 38-44.
6. Sanoqul o'g'li Q. N. et al. ELEKTRON HUIJAT ALMASHINUVINI AVTOMATLASHTIRISH MODELINI ANALITIK TAHLILI //Лучшие интеллектуальные исследования. – 2023. – Т. 10. – №. 5. – С. 89-100.
7. Mamatkulovich B. B., Shuhrat o'g'li M. S., Jasurjonovich B. J. SPECIAL DEEP CNN DESIGN FOR FACIAL EXPRESSION CLASSIFICATION WITH A SMALL AMOUNT OF DATA //Open Access Repository. – 2023. – Т. 4. – №. 3. – С. 472-478.
8. Mamatkulovich B. B. et al. Simplified machine learning for image-based fruit quality assessment //Eurasian Journal of Research, Development and Innovation. – 2023. – Т. 19. – С. 8-12.
9. Mamatkulovich, B. B., Dilshod o'gli, Y. A., & Akmal o'g'li, A. A. (2023). Predicting daily energy production in a blockchain-based P2P energy trading system. Texas Journal of Engineering and Technology, 18, 7-11.
10. Javlon, K., & Erali, M. (2023). STRUCTURE AND PRINCIPLE OF OPERATION OF FULLY CONNECTED NEURAL NETWORKS. International Journal of Contemporary Scientific and Technical Research, 136-141.
11. Мустафоев, Е., & Холматов, Ж. (2023). Brayl matn tasviri sifatini oshirish usullari. Информатика и инженерные технологии, 1(2), 23-27.
12. Obid o'g' A. S. J. et al. Numpy Library Capabilities. Vectorized Calculation In Numpy Va Type Of Information //Eurasian Research Bulletin. – 2022. – Т. 15. – С. 132-137.
13. Javlon X. et al. Классификатор движения рук с использованием биомиметического распознавания образов с помощью сверточных нейронных сетей с методом динамического порога для извлечения движения с использованием датчиков EF //Journal of new century innovations. – 2022. – Т. 19. – №. 6. – С. 352-357.



14. Фитратович В. и др. МАТЕМАТИЧЕСКАЯ МОДЕЛЬ МНОГОФАЗНОЙ ФИЛЬТРАЦИИ В НЕФТЕГАЗОВОМ ПЛАСТЕ ПРИ ЕГО ЗАВОДНЕНИИ //INTERNATIONAL CONFERENCES ON LEARNING AND TEACHING. – 2022. – Т. 1. – №. 4. – С. 520-525.
15. Jamshid S. ENTROPY EVALUATION CRITERION IN DECISION TREE ALGORITHM EVALUATION //International Journal of Contemporary Scientific and Technical Research. – 2023. – С. 236-239.
16. Салимов Ж., Абулаева А. Классификации дерева в машинном обучении и гиперпараметрах //Информатика и инженерные технологии. – 2023. – Т. 1. – №. 1. – С. 71-79.
17. Obid o'g'li S. J., Nodir o'g'li X. A., Jasurjonovich B. J. SUPERVISED LEARNING REGRESSION ALGORITHM SIMPLE LINEAR REGRESSION //Academia Science Repository. – 2023. – Т. 4. – №. 04. – С. 69-76.
18. Ramazon, Mixliyev, and Babayarov Abdusattor. "MIKROSKOP YORDAMIDA HUYAYRALARDAGI QON VA OQ QON HUYAYRALARI SONI BO'YICHA BEMORLARNING SOG'LIG'INI ANIQLASH." International Journal of Contemporary Scientific and Technical Research (2023): 133-137.
19. Norqo'ziyev, Q. (2023). MOBIL ROBOTLAR UCHUN YO'LNI REJALASHTIRISH ALGORITMI. Research and Implementation. извлечено от <https://fer-teach.uz/index.php/rai/article/view/746>
20. Норкозиёв, К., & Тоджиев, А. (2023). Использование искусственных нейронных сетей при разработке алгоритма поиска оптимального пути мобильных роботов в динамических средах. Информатика и инженерные технологии, 1(1), 25–29. извлечено от <https://inlibrary.uz/index.php/computer-engineering/article/view/25025>
21. Тоджиев, А., & Норкузиёв, К. (2023). The role of artificial intelligence technology in individualized teaching . Информатика и инженерные технологии, 1(2), 153–156. извлечено от <https://inlibrary.uz/index.php/computer-engineering/article/view/25014>
22. Ramazon, Mixliyev, and Babayarov Abdusattor. "MIKROSKOP YORDAMIDA HUYAYRALARDAGI QON VA OQ QON HUYAYRALARI SONI BO'YICHA BEMORLARNING SOG'LIG'INI ANIQLASH." International Journal of Contemporary Scientific and Technical Research (2023): 133-137.