

BIOLOGIK VA TIBBIY JARAYON KETMA-KETLIGI

¹*Sh.Sh.Esanov*

²*A.SH.Ulug'berdiyev*

³*S.A.Abduraxmonov*

^{1,2,3}Assistant "Biotibbiyot muhandisligi, informatika va biofizika" kafedrası

Toshkent tibbiyot akademiyasi

samandarabdurahmonov00@gmail.com

+998889007996

Anotatsiya: Biotexnologiya va axborot texnologiyalari sohasidagi jadal taraqqiyot biologik ketma-ketlikka oid katta hajmdagi ma'lumotlar davrini boshlab bermogda. Sun'iy intellekt texnologiyalari, xususan, katta ma'lumotlarni tahlil qilish va tabiiy tilni qayta ishlash (NLP) usullari biologik ketma-ketliklarni chuqur o'rganishda muhim vosita sifatida namoyon bo'lmoqda. AlphaFold va ESM kabi ilg'or modellar oqsillarning tuzilishini oldindan belgilash va funksiyalarini izohlashda katta muvaffaqiyatlarga erishgan bo'lib, bu asosiy tadqiqotlar va dori vositalarini ishlab chiqarishda innovatsion yechimlar taqdim etmoqda. Mavjud yondashuvlar naqsh va statistikaga asoslangan, katta ma'lumotlar va NLP texnologiyalariga bo'lingan. AlphaFold 3 va ColabFold kabi modellar oqsillar va biomolekula komplekslarini yuqori aniqlikda bashorat qilishda yangi mezonlar yaratmoqda. Genomik va RNK til modellari (masalan, RNAErnie, ESM2) semantik tahlil orqali yangi bilimlarni ochib bermogda. Shu bilan birga, ma'lumotlarning to'liqligi, uzoq masofali bog'liqlik, multi-omics integratsiyasi va modellar mustahkamligi kabi muammolar dolzarb bo'lib qolmoqda. Sintetik ma'lumotlar, yarim nazoratli o'rganish va tushuntiriladigan AI texnikalari (SHAP, GNNExplainer) bu muammolarni yumshatishda istiqbolli. Ushbu yutuqlar kasalliklarni aniqlash, biomarkerlarni topish va tibbiyotda innovatsiyalarni jadallashtirishga xizmat qilmoqda, shu bilan biologik ketma-ketlik tahlilining kelajakdagi universal modellari shakllantirmogda.

Kalit so'zlar: Biologik ketma-ketlik; Sun'iy intellekt; Katta ma'lumotlar tahlili; Tabiiy tilni qayta ishlash (NLP); AlphaFold; ESM modeli; Oqsil tuzilishi; Biomolekulyar til modellari; Multi-omics integratsiyasi; Tushuntiriladigan AI; Biotibbiyot innovatsiyalari; Kasallik diagnostikasi; Dori vositalari ishlab chiqarish

I.KIRISH

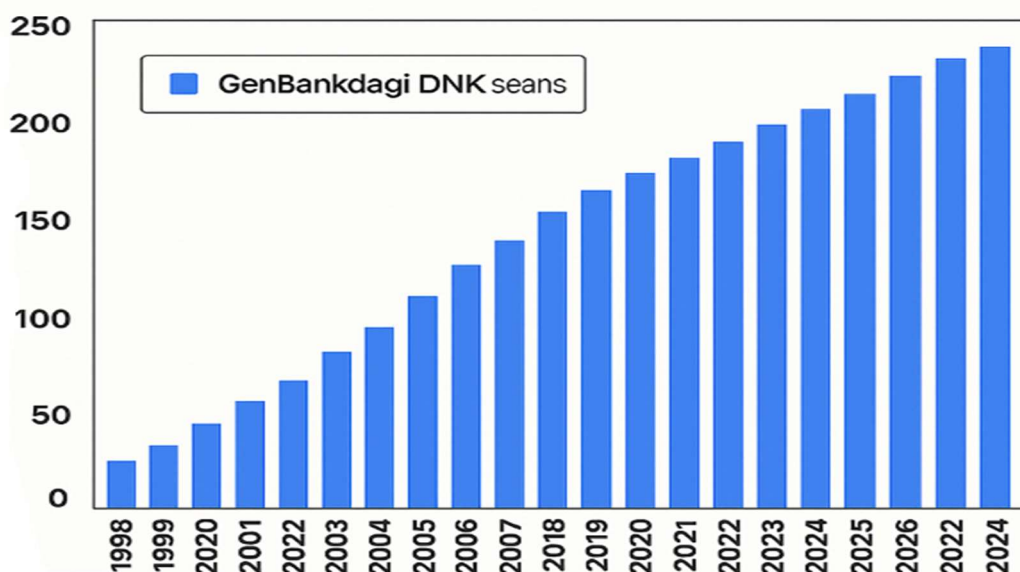
Biotexnologiya va IT sohasidagi jadal taraqqiyot tufayli biz biologik ketma-ketlikka oid katta hajmdagi ma'lumotlar davriga qadam qo'yamogdamiz. Organizm faoliyatini tushunish uchun ushbu murakkab va keng qamrovli ma'lumotlardan foydali xulosalar chiqarish jiddiy muammolardan biri bo'lib qolmoqda. Aynan shu yerda sun'iy intellekt texnologiyalari, xususan, katta ma'lumotlarni tahlil qilish hamda tabiiy tilni qayta ishlash usullari biologik ketma-ketlikni chuqur o'rganishda muhim vosita sifatida maydonga chiqmogda. Zamonaviy texnologiyalar biologik ma'lumotlardagi naqshlarni aniqlashni soddalashtirib, murakkab hayotiy jarayonlarga yangi ilmiy yondashuvlarni ochib bermogda. Ayniqsa, AlphaFold va ESM kabi ilg'or modellar oqsillarning tuzilishini oldindan belgilash hamda ularning funksiyalarini izohlashda

katta muvaffaqiyatlarga erishgan. Bu natijalar nafaqat asosiy biologik tadqiqotlarni jadallashtiradi, balki kasalliklarni aniqlash va dori vositalarini ishlab chiqishda innovatsion yechimlar va strategiyalarni taqdim etadi. Shu asosda, biz biologik ketma-ketlikni tahlil qilish sohasidagi ilgʻor yutuqlarni koʻrib chiqamiz hamda ularning tibbiyotda keng qoʻllanilishiga alohida eʼtibor qaratamiz. Shuningdek, biotibbiyot sohasidagi salohiyatni toʻliq ochib berish uchun ketma-ketlik tahlilida uzluksiz innovatsiyalar zarurligini taʼkidlaymiz, mavjud muammolarni yoritib, istiqbolli yoʻnalishlarni taklif etamiz.

II.ASOSIY QISM

Biotexnologiya hamda axborot texnologiyalari sohasidagi ilgʻor tadqiqotlar biologik ketma-ketlikka oid katta hajmdagi maʼlumotlar davrining shakllanishiga zamin yaratdi. Ushbu maʼlumotlar tarkibida biomolekulyar ketma-ketliklar, ularning fizik-kimyoviy xususiyatlari, tizimli tuzilmalari va funksional tafsilotlari mujassam. Ochiq manbalarda mavjud boʻlgan biomolekulyar ketma-ketliklar soni sezilarli darajada ortib bormoqda, bu esa ularning tuzilmasi va funksiyasiga oid izohlarning kengayishiga olib kelmoqda. Biologik ketma-ketlikka oid katta maʼlumotlarni chuqur tahlil qilish orqali biz biologik jarayonlar va mexanizmlar haqida yanada aniqroq tasavvurga ega boʻlamiz. Bu esa patogenezni oʻrganishda muhim ilmiy natijalarga erishish imkonini beradi.

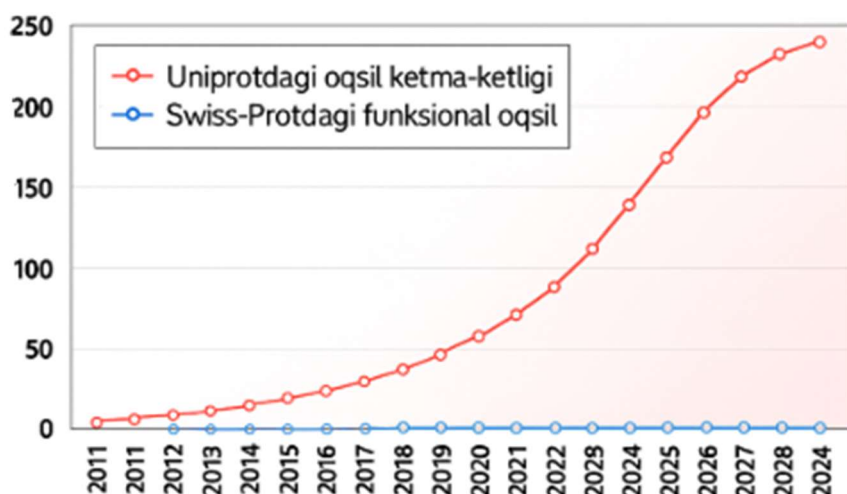
DNK seanslari o'sib borayotgan tendtsniyasi



Biologik ketma-ketlikni tahlil qilish tibbiy sohadagi koʻplab amaliyotlarda muhim axborot manbai sifatida xizmat qiladi. Ushbu yoʻnalishning rivojlanishi hayotiy jarayonlarni chuqurroq anglash imkonini yaratib, biotibbiyot tadqiqotlariga sezilarli hissa qoʻshmoqda. Katta hajmdagi maʼlumotlarni tahlil qilish va tabiiy tilni qayta ishlash (NLP) texnologiyalari biomolekulalarning funksional xaritalarini tuzishda muhim rol oʻynab, ilmiy izlanishlarda yoʻl koʻrsatuvchi vosita sifatida namoyon boʻlmoqda. AlphaFold modeli esa chuqur oʻrganish algoritmlaridan foydalangan holda

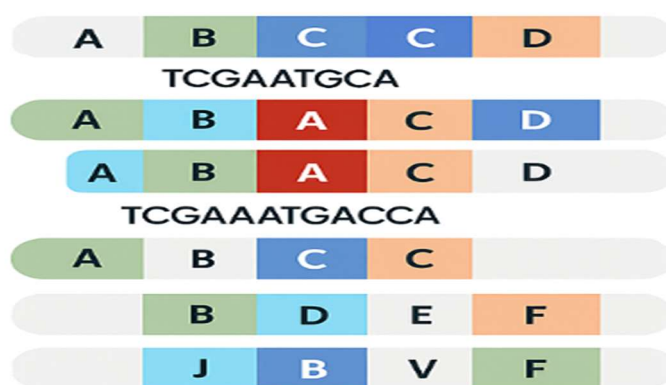
biologik ketma-ketlik asosida oqsil tuzilmalarini oldindan aniqlashga yordam beradi. Bu yondashuv ketma-ketlik va tuzilma o'rtasidagi tafovutni kamaytirib, oqsillar olamining keng qamrovli tasvirini taqdim etadi.

Oqsil malumotlarining o'sish tendentsiyasi



AlphaFold 3 modeli diffuziyaga asoslangan ilg'or arxitekturani qo'llab, turli biomolekulalardan tashkil topgan komplekslarning umumiy tuzilishini nihoyatda aniq tarzda oldindan belgilash imkonini beradi. Ushbu texnologik yutuq biologik ketma-ketlikni tahlil qilishda yangi mezon yaratib, sohaga tub burilish olib keldi. Biologik ketma-ketlikni chuqur tahlil qilish esa irsiy o'zgarishlarni aniqlash, biomarkerlarni identifikatsiyalash va yangi dori vositalarini topish kabi biotibbiyot sohasidagi tadqiqotlarda tobora keng qo'llanilmoqda. Bu esa kasalliklarni aniqlash va davolash usullarini sezilarli darajada takomillashtirishga xizmat qilmoqda.

DNK Mutatsiya Tahlili



Kasallikka Aloqador Molekula Identifikatsiyasi

Biologik ketma-ketlik tahlilidagi yutuqlar.

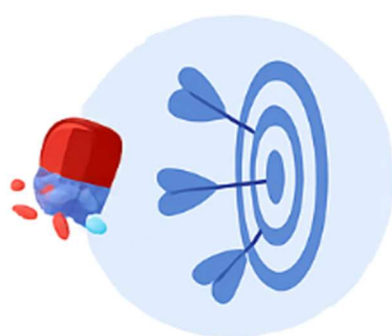
Biologik ketma-ketlikni tahlil qilishning mavjud yondashuvlari uchta asosiy toifaga bo'linadi:

1. Naqsh va statistikaga asoslangan metodlar,

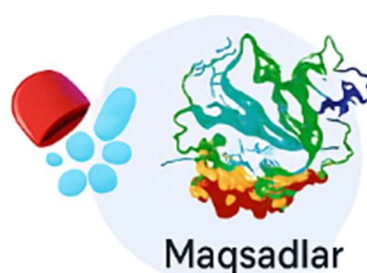
2. Katta hajmdagi biologik ma'lumotlarga tayanuvchi usullar,

3. Tabiiy tilni qayta ishlash (NLP) texnologiyalariga asoslangan yondashuvlar.

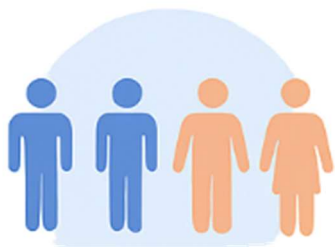
Naqsh va statistikaga asoslangan usullar, avvalo, ketma-ketliklarning statistik xususiyatlari hamda ilgari to'plangan bilimlarga tayanadi. Bu yondashuvlar odatda ketma-ketlikdagi muhim belgilarni aniqlash uchun maxsus qoidalar yoki statistik modellarni qo'llaydi. Biologik ketma-ketlik xususiyatlarini aniqlashda turli statistik va fizik-kimyoviy yondashuvlar qo'llaniladi. Masalan, nukleotidlar yoki aminokislotalarning ma'lum pozitsiyalardagi uchrashish chastotasini tahlil qilish orqali vaznli matritsalar tuzish mumkin. Bu matritsalar ketma-ketlikdagi muhim naqshlarni tavsiflashda xizmat qiladi. Aminokislota indeksleri esa ularning fizik-kimyoviy va biokimyoviy xususiyatlarini kodlash uchun hisoblanadi. Nuklein kislotalar ketma-ketligini ifodalashda esa mahalliy k-mer naqshlar, fizik-kimyoviy parametrlar va nukleotidlar o'rtasidagi korrelyatsiyalar kabi global xususiyatlar birgalikda qo'llaniladi. DPCfam modeli aminokislotalarning pozitsion taqsimotiga asoslanib, ketma-ketlikni moslashtirish orqali protein oilalarini avtomatik tarzda aniqlaydi. DeepSoluE esa oqsillarning eruvchanligini oldindan baholash uchun fizik-kimyoviy naqshlar va aminokislotalarning taqsimlangan tasvirlarini birlashtirgan gibrid xususiyatlar asosida ishlaydi.



**Patogen-Xost
Özaro ta'sir**



**Patogen-Xost
Oksilni Aniqlash**



Katta ma'lumotlarning biologik ketma-ketligiga asoslangan usullar esa biologik ketma-ketlikka asoslangan katta ma'lumotlar yondashuvlari odatda ketma-ketliklarning asosiy xususiyatlarini aniqlash va ularni miqdoriy, barqaror ko'rinishda ifodalash uchun keng ko'lamdagi ma'lumotlar to'plamidagi murakkab bog'liqliklarga tayanadi. AlphaFold 2 modeli aminokislotalar ketma-ketligini, bir nechta ketma-

ketliklarni moslashtirish natijalarini hamda tizimli shablon ma'lumotlarini kirish sifatida qabul qiladi. U chuqur o'rganish algoritmlarida diqqat mexanizmidan foydalanib, oqsillarning strukturaviy xususiyatlarini yuqori aniqlikda tasvirlay oladi. Bu model deyarli barcha ma'lum oqsillarni qamrab olgan holda, ularning tuzilmasini aniq prognozlashga erishadi. AlphaFold 2 va keyinchalik AlphaFold 3 ning muvaffaqiyati shuni ko'rsatadiki, chuqur o'rganish texnologiyalari biologik ketma-ketlik ma'lumotlaridan murakkab xususiyatlarni avtomatik tarzda ajratib olishga qodir. Bu esa oqsil va biomolekulalar o'rtasidagi o'zaro ta'sir tuzilmalarini yuqori aniqlikda bashorat qilish imkonini beradi. ColabFold esa AlphaFold 2 modelini tez homologiyali qidiruv strategiyasi bilan birlashtirib, oqsil tuzilmalari va ularning komplekslarini prognozlash jarayonini sezilarli darajada tezlashtiradi. Ushbu kombinatsiya nafaqat hisoblash samaradorligini oshiradi, balki tuzilmani bashorat qilish imkoniyatlarini kengaytiradi. Natijada, biologik tadqiqotlar va farmatsevtika sohasida kuchli va samarali vositalar paydo bo'ladi. Tabiiy tilni qayta ishlashga asoslangan usullar bu biologik ketma-ketliklar va tabiiy til o'rtasidagi o'xshashliklar asosida yaratilgan NLP (tabiiy tilni qayta ishlash) metodlari biologik ma'lumotlardagi chuqur naqshlarni semantik darajada talqin qilish imkonini beradi. Bu yondashuvlar ketma-ketliklarning lingvistik xususiyatlarini o'rganib, ularni semantik fazolarga proyeksiya qilish orqali yanada aniqroq tasvirlaydi. Biomolekulyar til modellari molekulalarning kontekstual til xususiyatlarini tahlil qilib, ularning semantik ifodalarini hosil qiladi. Masalan, Genomik oldindan o'rgatilgan tarmoq DNK til modeli sifatida genomik ketma-ketliklar asosida nazoratsiz o'rganish orqali gen tuzilmalari, DNK motivlari va genom miqyosidagi variantlarning ta'sirini aniqlashga xizmat qiladi. RNAErnie esa RNK til modeli bo'lib, motivlarni oldindan o'rgatish va turga mos nozik sozlash bosqichlarini o'z ichiga oladi. ESM2 modeli esa to'g'ridan-to'g'ri ketma-ketlikdan atom darajasidagi tuzilmani tez va aniq prognozlash imkonini beradi hamda metagenomik hududlarda ilgari noma'lum bo'lgan bilimlarni ochib beradi. NLP texnologiyalarining biologik ketma-ketlik tahliliga qo'llanishi biomolekulyar til modellariga "hayot kitobi"dagi ma'nolarni anglash imkonini yaratadi.

Kasallik bilan bog'liq molekulalarni aniqlash esa yangi terapevtik maqsadlar va molekulyar biomarkerlarni topishda muhim ahamiyatga ega. Dastlabki hisoblash metodlari turli biologik sharoitlarda gen ekspressiya profillarini solishtirish orqali individual kasalliklarni tahlil qilishga qaratilgan edi. Bugungi kunda esa biologik katta ma'lumotlar tahlili bir vaqtning o'zida bir nechta kasalliklar bo'yicha tegishli molekulalarni aniqlash imkonini bermoqda. Molekulalarning yashirin statik va dinamik o'zaro ta'sir xususiyatlarini aniqlash uchun ketma-ketliklar, ekspressiya profillari, genetik variantlar va molekulyar tarmoqlar kabi turli manbalardagi oldingi bilimlar grafik modellar orqali birlashtiriladi. Yo'lni boyitish, dori sezgirlikni baholash va eksperimental tekshiruvlar kabi quyi oqim tahlillari esa molekulyar patologik mexanizmlarni tizimli o'rganishni yanada chuqurlashtiradi. Biologik ketma-ketlikni tahlil qilishdagi qiyinchiliklar biologik ketma-ketlikni tahlil qilishda erishilgan yutuqlarga qaramay, bir qator muhim muammolar hanuzgacha dolzarb bo'lib qolmoqda. Eng asosiy muammolardan biri ma'lumotlarning to'liqligi va

muvozanatining yetishmasligi. Garchi biz biologik ketma-ketlikka oid katta hajmdagi ma'lumotlar davriga kirgan bo'lsak-da, molekulyar funksiyalarni — masalan, qo'shilish variantlari, tartibga soluvchi elementlar va o'zaro ta'sir zonalarini — izohlash jarayoni katta xarajatlar va chuqur mutaxassislikni talab qiladi. Bu esa ko'plab ketma-ketliklarning yetarli darajada belgilab chiqilmaganligiga olib keladi va natijada tahlil modellarini yaratish jarayonini murakkablashtiradi. Kam uchraydigan hujayra turlari yoki noyob kasalliklar uchun mavjud bo'lgan namunalarning miqdori kamligi, ketma-ketlik chuqurligining cheklanganligi va toifalar o'rtasidagi sezilarli tafovutlar nomutanosib va yetarli bo'lmagan ma'lumotlar to'plamlarini yuzaga keltiradi. Bu holat biologik jarayonlar haqida mazmunli xulosalar chiqarishni sezilarli darajada qiyinlashtiradi. Ushbu muammolarni yumshatish uchun sintetik ma'lumotlar yaratish va ma'lumotlarni ko'paytirishning ilg'or usullarini joriy etish mumkin. Shuningdek, domenlararo bilimlarni uzatish, yarim nazorat ostida o'rganish va faol o'rganish kabi yondashuvlar izohlanmagan ma'lumotlardan samarali foydalanishga yordam beradi. Bu metodlar biologik ketma-ketlik tahlilining aniqligini oshirish va ilmiy natijalarni chuqurlashtirishda muhim rol o'ynaydi. *Uzoq masofaga bog'liqlik va murakkab o'zaro ta'sir*. Biologik ketma-ketliklar murakkab va o'zaro bog'liq naqshlarni namoyish etadi. Birinchidan, ketma-ketliklar ichidagi muhim ma'lumotlar uzoq pozitsiyalar bo'ylab taqsimlanishi mumkin, masalan, oqsil ketma-ketligida bir-biridan uzoqda joylashgan aminokislotalar qoldiqlari o'rtasidagi o'zaro ta'sirlar, ammo uning tuzilishi va funktsiyasiga ta'sir qiladi. Ikkinchidan, bir nechta mintaqalar va segmentlar o'rtasidagi chiziqli bo'lmagan bog'liqliklar va o'zaro ta'sirlar an'anaviy modellar uchun ushbu murakkab naqshlarni aniqlashni qiyinlashtiradi. Tabiiy tilni qayta ishlashda biologik ketma-ketlik katta ma'lumotlari va til modellariga asoslanib, biologik ketma-ketlikni tahlil qilish muammolari xususiyatlariga ko'ra yashirin semantikani ochish uchun samarali biologik til modellarini loyihalash arziydi. *Multi-omics ma'lumotlar integratsiyasi*. Turli xil omiks ma'lumotlarini birlashtirish molekulyar tartibga solish mexanizmlarini yanada to'liqroq tushunish imkonini beradi. Biroq, multi-omics ma'lumotlarini tizimli modellashtirish yuqori heterojenlik va o'lchovlilik, past signal-shovqin nisbati, partiya effektlari va omiks ma'lumotlar to'plamining boshqa noyob atributlari tufayli jiddiy qiyinchiliklarni keltirib chiqaradi. Turli omika qatlamlarini uyg'unlashtirish va bog'lash mazmunli tushunchalarni olish uchun juda muhimdir. Ushbu muammoni hal qilish uchun ma'lumotlarni to'g'rilash jarayonlarini yo'q qilish kerak va grafiklarni o'rganish, chuqur ko'p ko'rinishda o'rganish va kontrastli o'rganish kabi ba'zi chuqur o'rganish usullari naqshlarni ajratib olish uchun birlashtirilishi mumkin, bu omik ma'lumotlarning muvofiqligini ta'minlaydi va turli molekulyar darajalar o'rtasidagi korrelyatsiyalarni ushlaydi. Modelning mustahkamligi va amaliyligi: biologik tahlil uchun muhim omillar ko'plab mavjud ketma-ketlikni tahlil qilish modellari odatda ma'lum bir ma'lumotlar to'plami yoki vazifaga moslashtirilgan bo'ladi. Biroq, bu modellarning ko'pi real biologik muhitda barqaror ishlash va amaliy qo'llash nuqtai nazaridan yetarli darajada mustahkam emas. Ular ko'pincha haddan tashqari moslashishga moyil bo'lib, ma'lumotlardagi xatoliklar bilan kurashishda qiynaladi. Bu esa yangi yoki ilgari ko'rilmagan biologik ketma-ketliklarni tahlil

qilishda samaradorlikni pasaytiradi. Ko‘p vazifali tahlil va umumlashtirishdagi cheklovlar modellar bir nechta biologik vazifalarni tahlil qilishda izchil natijalar bera olmaydi, chunki ularning umumlashtirish qobiliyati va turli vaziyatlarga moslashuvchanligi cheklangan. Bu holat biologik tadqiqotlarda keng qo‘llaniladigan universal yondashuvlar yaratishni qiyinlashtiradi.

III.XULOSA

Hisoblash murakkabligi va biologlar uchun qiyinchiliklar bunday modellarning hisoblash jarayonlari murakkab bo‘lib, biologik tadqiqotchilar uchun texnik jihatdan tushunish va qo‘llashda muammolar tug‘diradi. Shu sababli, modelning haddan tashqari moslashishini oldini olish va umumlashtirish qobiliyatini oshirish uchun tartibga solish usullari (regularization) hamda ansambl yondashuvlari (ensemble frameworks) taklif etiladi. Kelajakdagi yechim: universal va qulay modellar biologik talqinning ahamiyati va muammolari garchi mavjud hisob-kitob modellari biologik ketma-ketliklarni bashorat qilishda samarali bo‘lsa-da, bu modellarning ichki ishlash mexanizmlari ko‘pincha murakkab va tushunarsiz bo‘lib, ular “qora quti” sifatida qaraladi. Bunday noaniqliklar esa model natijalarining ishonchlilikini pasaytiradi va ularni real biotibbiy tadqiqotlarda qo‘llashni cheklaydi. Izohlash texnikalari va ularning cheklovlari bu ba’zi tushuntirish usullari mavjud bo‘lsa-da, ular odatda faqat ma’lum vazifalarga mos keladi va keng biologik kontekstda izoh berishdan ko‘ra, modelning bashorat qilish qobiliyatini oshirishga qaratilgan. Bu esa biologlar uchun model natijalarini to‘liq tushunishni qiyinlashtiradi. SHAP qiymatlari va GNNExplainer kabi tushuntiriladigan sun’iy intellekt texnikalari model qarorlarini izohlashda istiqbolli vositalar hisoblanadi. Biroq, bu texnikalarning biologik kontekstga mos tushuntirishlar berishi uchun biologlar bilan yaqin hamkorlikda ishlash zarur. Bu yondashuv model natijalarining amaliy biotibbiyotda qo‘llanishini yanada ishonchli va tushunarli qiladi.

IV.FOYDALANILGAN ADABIYOTLAR

1. Jumper, J., Evans, R., Pritzel. Highly accurate protein structure prediction with AlphaFold. *Nature*, 596(7873), 583–589. (2021)
2. Abramson, J., Adler, J., Dunger. Accurate structure prediction of biomolecular interactions with AlphaFold 3. *Nature*, 630(8016), 493–500. (2024)
3. Davranov, K., Abdusamatov, S., Khujamshukurov, N. A. Advances in plant biotechnology and sequence analysis in Uzbekistan agriculture. *International Journal of Biotechnology and Molecular Biology Research*, 1–12. (2022).
4. Lin, Z., Akin, H.,. Evolutionary-scale prediction of atomic-level protein structure with a language model. *Science*, 379(6639), 1123–1130. (2023)
5. SA Abdurakhmonov, JT Abdurazzoqov, DB Elmurotova. Zamonaviy biosensornlarning tibbiyotdagi ahamiyati. *Innovations in Science and Technologies*. 2024

6. Turdikulova, S., Dalimov, Z., & Ashurov, A. Genome-wide association studies in Central Asian populations: Implications for biotechnology. *Journal of Human Genetics*, 411–420. 2019
7. Umer, M., Rout, R. K., & Khandelwal, P. (2024). *Computational Techniques for Biological Sequence Analysis*. Routledge. ISBN: 978-1032630267
8. Plonsey, R., & Barr, R. C. *Bioelectricity: A Quantitative Approach*. Springer 2007. ISBN: 978-0387488646