

## MASHINALI O'RGATISHGA ASOSLANGAN KASALLIKLAR DIAGNOSTIKASI

**Shukurillayev Kozimjon Shuxrat o'g'li**

ToshDTU 2-kurs magistranti

[kshukurillaev@mail.ru](mailto:kshukurillaev@mail.ru)

### ANNOTATSIYA

*Sun'iy intellekt (AI) sohasi bo'lgan Machine Learning (ML) tadqiqotchilar, shifokorlar va bemorlarga ushbu muammolarning bir qismini hal qilish imkonini beradi. Ko'pgina tadqiqotchilar va amaliyotchilar Mashinali o'rgatishga asoslangan kasallik diagnostikasi (MLBDD) va'dasini tasvirlamoqda, bu arzon va tezdir. An'anaviy diagnostika jarayonlari qimmat, ko'p vaqt talab qiladi va ko'pincha inson aralashuvini talab qiladi. An'anaviy diagnostika usullari inson imkoniyatlari bilan cheklangan bo'lsa-da, Mashinali o'rgatishga asoslangan tizimlar bunday cheklovlarga ega emas va mashinalar odamlar kabi charchamaydi. Natijada, sog'liqni saqlash tizimida ko'p sonli bemorlarning kutilmagan mavjudligi bilan kasallikni tashxislash usuli ishlab chiqilishi mumkin.*

Mashinali o'rgatish (ML) - bu matematik va statistik yondashuvlar yordamida asosiy xulosalar chiqarish uchun ma'lumotlar namunalarini tahlil qiladigan yondashuv bo'lib, mashinalarga dasturlashtirilmasdan o'rgatish imkonini beradi. Artur Samuel 1959 yilda asosiy yutuq birinchi marta e'tirof etilganida, tajribadan o'rgatish uchun o'yinlarda mashina o'rgatish va naqshni aniqlash algoritmlarini joriy qildi. Mashinali o'rgatishning asosiy printsiplari - bu vazifaga qarab bashorat qilish yoki qaror qabul qilish uchun ma'lumotlardan o'rgatishdir.<sup>[1]</sup> Mashinali o'rgatish (ML) texnologiyasi tufayli ko'p vaqt talab qiladigan vazifalar endi tez va minimal kuch bilan bajarilishi mumkin. Kompyuter quvvati va ma'lumotlar sig'iminining eksponentsial o'sishi bilan natijalarni deyarli mukammal aniqlik bilan bashorat qilish

uchun ma'lumotlarga asoslangan Mashinali o'rgatish modellarini o'rgatish osonlashmoqda. Bir nechta maqolalar Mashinali o'rgatishga turli yondashuvlarni taklif qiladi [2][3].

Mashinali o'rgatish algoritmlari odatda uchta toifaga bo'linadi: nazorat ostida, nazoratsiz va yarim nazorat ostida.<sup>[2]</sup> Biroq, 1-rasmda ko'rsatilganidek, Mashinali o'rgatish algoritmlarini turli xil o'rgatish yondashuvlari asosida bir nechta kichik guruhlariga bo'lish mumkin. Mashinali o'rgatishning mashhur algoritmlaridan ba'zilar qatoriga chiziqli regressiya, logistik regressiya, vektorli mashinalarni qo'llab-quvvatlash (SVM), tasodifiy o'rmon (RF) va Naive Bayes (NB) kiradi.<sup>[2]</sup>



**1-rasm.** Mashinali o'rgatish algoritmlarining har xil turlari.

## **Turli kasalliklarni tashxislash uchun Mashinali o'rgatish usullari**

Ko'pgina olimlar va amaliyotchilar kasallik tashxisida Mashinali o'rgatish (ML) yondashuvlaridan foydalanganlar. Ushbu bo'limda Mashinali o'rgatishga asoslangan kasalliklar diagnostikasi (MLBDD) ko'p turlari tasvirlangan, ularning ahamiyati va jiddiyligi tufayli katta e'tibor beriladi. Masalan, COVID-19 ning global ahamiyati tufayli bir nechta tadqiqotlar 2020 yildan hozirgi kungacha Mashinali o'rgatishdan foydalangan holda COVID-19 kasalligini aniqlashga qaratilgan bo'lib, u ham tadqiqotda ustuvor ahamiyatga ega. Yurak kasalliklari, buyrak kasalliklari, ko'krak saratoni, diabet, Parkinson kasalligi, Altsgeymer kasalligi va COVID-19 kabi og'ir kasalliklar qisqacha muhokama qilinadi, boshqa kasalliklar esa "Boshqa kasalliklar" bo'limida qisqacha muhokama qilinadi.

### **Yurak kasalliklari**

Ko'pgina tadqiqotchilar va amaliyotchilar yurak kasalliklarini aniqlash uchun Mashinali o'rgatish (ML) yondashuvlaridan foydalanadilar.<sup>41[5]</sup> Masalan, Ansari va boshqalar (2011) neyro-loyqa integratsiyalashgan tizimlarga asoslangan yurak ishemik kasalliklarini tashxislash uchun avtomatlashtirilgan tizimni 89% aniqlik bilan taklif qilishdi.<sup>41</sup>. Tadqiqotning muhim kamchiliklaridan biri bu ko'p sinfli tasniflash, katta ma'lumotlarni tahlil qilish va sinflarni muvozanatsiz taqsimlash kabi turli stsenariylarda taklif qilingan metodologiyaning qanday ishlashini aniq tushuntirishning yo'qligi. Bundan tashqari, yaqinda tibbiyot sohalarida, ayniqsa, tibbiy bo'lmagan foydalanuvchilarga yondashuvni tushunishga yordam berish uchun katta rag'batlantirilgan modelning to'g'riligi uchun hech qanday izoh yo'q.

Rubin va boshqalar (2017) tartibsiz yurak tovushlarini aniqlash uchun chuqur konvolyutsion neyron tarmoq yondashuvlaridan foydalanadilar. Ushbu tadqiqot mualliflari o'quv ma'lumotlar to'plamining sezgirliги va o'ziga xosligini yaxshilash uchun yo'qotish funksiyasini moslashtirdilar. Ular taklif qilgan model 2016-yilda

PhysioNet Computing Competition tanlovida sinovdan o'tkazildi. Ular tanlovda 0,95 o'ziga xoslik va 0,73 sezgirlik yakuniy bashorati bilan ikkinchi o'rinni egalladilar.

Bundan tashqari, yaqinda chuqur o'rgatishga (DL) asoslangan algoritmlar yurak kasalliklarini aniqlashda e'tiborni tortdi. Miao va Miao va boshqalar (2018), masalan, ko'p sinfli morfologik naqsh asosida homilaning kardiokografik salomatligini tashxislash uchun DL-ga asoslangan usulni taklif qilishdi. Yaratilgan model homiladorlik asoratlari bo'lgan shaxslarning morfologik rasmini farqlash va tasniflash uchun ishlatiladi. Ularning dastlabki hisob-kitob natijalariga 88,02% aniqlik, 85,01% aniqlik va 0,85 F balli kiradi. Ushbu tadqiqot davomida ular haddan tashqari moslashish bilan bog'liq muammolarni hal qilish uchun bir nechta tashlab ketish strategiyalaridan foydalanganlar, bu esa pirovardida mashg'ulot vaqtini ko'paytirdi va ular yuqori aniqlik uchun savdo sifatida tan oldilar.

Mashinali o'rgatish dasturlari yurak kasalliklarini tashxislash uchun keng qo'llanilsa-da, ko'p sinfli tasniflashda nomutanosib ma'lumotlar bilan bog'liq muammolarni hal qiladigan tadqiqotlar o'tkazilmagan. Bundan tashqari, ko'p hollarda, yakuniy prognozda modelni tushuntirish mumkin emas.

### **Buyrak kasalliklari**

Ko'pincha buyrak kasalligi deb ataladigan buyrak kasalligi, nefropatiya yoki buyraklarning shikastlanishiga ishora qiladi. Buyrak kasalligi bilan og'riqan bemorlarda buyraklarning funktsional faolligi pasayadi, bu esa o'z vaqtida davolanmasa, buyrak etishmovchiligiga olib kelishi mumkin. Milliy Buyrak Jamg'armasining ma'lumotlariga ko'ra, dunyo aholisining 10% dan aziyat chekadi surunkali buyrak kasalligi (CKD) va har yili millionlab odamlar noto'g'ri davolanish tufayli vafot etadi. ML va DL asosida buyrak kasalliklarini tashxislash bo'yicha so'nggi yutuqlar buyrak kasalligi diagnostikasi bilan bog'liq testlarni engishga qodir bo'lmagan mamlakatlar uchun imkoniyat yaratishi mumkin.<sup>[8]</sup> Masalan, Charleonnan va boshqalar. (2016) to'rt xil Mashinali o'rgatish algoritmlarini baholash uchun ommaviy ma'lumotlar to'plamidan foydalangan: K-eng yaqin qo'shnilar (KNN),

vektorni qo'llab-quvvatlash mashinasi (SVM), logistik regressiya (LR) va qarorlar daraxti klassifikatorlari va 98,1%, 98,3% aniqlik, mos ravishda 96,55% va 94,8%. Aljaaf va boshqalar (2018) shunga o'xshash tadqiqot o'tkazdilar. Mualliflar RPART, SVM, LOGR va MLP kabi turli xil Mashinali o'rgatish algoritmlarini sinovdan o'tkazdilar, CKD tomonidan qo'llaniladigan taqqoslanadigan ma'lumotlar to'plamidan foydalangan holda va MLP surunkali buyrak kasalligini aniqlashda yaxshiroq (98,1%) ishlaganligini aniqladi. Surunkali buyrak kasalligini aniqlash uchun Ma va boshqalar. (2020) ko'plab manbalardan olingan ma'lumotlarni o'z ichiga olgan ma'lumotlar to'plamidan foydalanadi. Ular taklif qilgan geterogen modifikatsiyalangan sun'iy neyron tarmog'i (HMANN) modeli 87-99% aniq edi.

### **Sut bezlari saratoni**

Ko'pgina tibbiyot olimlari erta bosqichda tashxis qo'yish uchun potentsial yechim sifatida Mashinali o'rgatish (ML) asosida ko'krak bezi saratoni tahlilini taklif qilishdi. Masalan, Miranda va Felipe (2015) ko'krak saratoni tasnifi uchun loyqa mantiqqa asoslangan kompyuter diagnostika tizimlarini taklif qilishdi. Loyqa mantiqning boshqa klassik Mashinali o'rgatish usullaridan afzalligi shundaki, u radiolog mutaxassisining fikri va uslubiga taqlid qilish orqali hisoblash murakkabligini minimallashtirishi mumkin. Agar foydalanuvchi kontur, shakl va zichlik kabi parametrlarni kiritsa, algoritm afzal ko'rgan usuli asosida saraton tasnifini taklif qiladi. Miranda va Felipe (2015) tomonidan taklif qilingan model taxminan 83,34% aniqlikka ega edi. Mualliflar tajriba uchun tasvirlarning taxminan teng nisbatidan foydalanganlar, natijada aniqlik va adolat yaxshilangan. Biroq, tadqiqot ularning natijalarini talqin qilishni tushunarli tarzda ko'rib chiqmaganligi sababli, umumiy aniqlik yaxshi va yomon xulqli o'smalarni tasniflashning haqiqiy aniqligini ko'rsatadi degan xulosaga kelish qiyin bo'lishi mumkin. Bundan tashqari, har bir sinf uchun modellarning haqiqiy bashoratini namoyish qilish uchun chalkashlik matritsasi taqdim etilmaydi.

Zheng va boshqalar (2014) k-means klasterlash (KMC) va SVM yordamida ko'krak bezi saratoniga tashxis qo'yish uchun gibrid strategiyalarni taqdim etdi. Ularning taklif qilgan modeli o'lchovdagi qiyinchiliklarni sezilarli darajada kamaytirdi va Viskonsin diagnostik ko'krak bezi saratoni (WDBC) ma'lumotlar to'plamidan foydalangan holda 97,38% aniqlikka erishdi. Ma'lumotlar to'plami odatda taqsimlanadi va 10 toifaga bo'lingan 32 xususiyatga ega. Ularning taklif etayotgan modeli teng bo'lmagan sinf nisbati bo'lgan ma'lumotlar to'plamida yaxshiroq ishlaydi, degan xulosaga kelish qiyin, unda etishmayotgan qiymat ham bo'lishi mumkin.

Mashinali o'rgatishning eng yaxshi modellarini aniqlash uchun Asri va boshqalar. (2016) Viskonsindagi ko'krak saratoni (WBC) ma'lumotlar to'plamiga SVM, DT (C4.5), NB va KNN kabi turli ML yondashuvlarini qo'llagan. Ularning topilmalariga ko'ra, SVM 97,13% aniqlik bilan boshqa barcha ML algoritmlaridan ustun keldi [<sup>14</sup>]. Biroq, agar bir xil tajriba boshqa ma'lumotlar bazasida takrorlansa, natijalar farq qilishi mumkin. Bundan tashqari, asosiy haqiqiy qiymatlar bilan birga bo'lgan eksperimental natijalar qaysi ML modeli eng yaxshi yoki yo'qligini aniqlashda aniqroq baho berishi mumkin.

Mohammed va boshqalar (2020) deyarli bir xil tadqiqot o'tkazdilar. Mualliflar Mashinali o'rgatishning eng yaxshi usullarini topish uchun uchta Mashinali o'rgatish algoritmidan foydalanadilar: DT (J48), NB va Sequential Minimum Optimization (SMO) va tajriba ikkita mashhur ma'lumotlar to'plamida o'tkazildi: WBC va ko'krak saratoni ma'lumotlar to'plami. Ushbu tadqiqotning qiziqarli jihati shundaki, ular ma'lumotlar muvozanati muammolariga e'tibor qaratdilar va qayta namunalangan ma'lumotlarni etiketlash protseduralaridan foydalanish orqali nomutanosiblik muammosini minimallashtirdilar. Ularning natijalari shuni ko'rsatdiki, SMO algoritmlari boshqa ikkita tasniflagichdan ustun bo'lib, ikkala ma'lumot to'plami uchun 95% dan ortiq aniqlikka erishdi. Biroq, nomutanosiblik omilini kamaytirish uchun ular qayta-qayta namuna olish tartib-qoidalaridan foydalanganlar, bu esa ma'lumotlar xilma-xilligi ehtimolini kamaytiradi. Natijada, ushbu uchta Mashinali



o'rgatish usulining ishlashi odatda taqsimlanmagan yoki muvozanatsiz ma'lumotlar to'plamida pasayishi mumkin.

Assegui (2021) eng yaxshi k-eng yaqin qo'shni (KNN) sozlamalarini aniqlash uchun grid qidiruv usulidan foydalangan. Ularning tadqiqoti shuni ko'rsatdiki, parametrlarni sozlash modelning ishlashiga sezilarli ta'sir ko'rsatdi. Ular parametrlarni nozik sozlash orqali 94,35% aniqlikka erishish mumkinligini ko'rsatdi, KNN esa sukut bo'yicha 90% aniqlikka erishdi.

Ko'krak bezi saratonini aniqlash uchun Bhattacharjee va boshqalar (2020) orqaga tarqaladigan neyron tarmog'idan (BNN) foydalangan. Tajriba WBC ma'lumotlar to'plamida to'qqiz xususiyatga ega bo'lgan va ular 99,27% aniqlikka erishgan. Alshaieji va boshqalar (2021) ko'krak saratoni o'smalarini tasniflash uchun sayoz ANN modelini ishlab chiqish uchun WBCD va WDBI ma'lumotlar to'plamidan foydalangan. Mualliflar tavsiya etilgan model o'smalarni xarakteristikalar yoki sozlash algoritmlarini tanlamasdan 99,85% gacha to'g'ri tasniflashi mumkinligini ko'rsatdi.

Sultana va boshqalar (2021) WBC ma'lumotlar to'plamida boshqa ANN arxitekturasidan foydalangan holda ko'krak saratonini aniqlaydi. Ular turli xil NN arxitekturalaridan foydalanganlar, shu jumladan ko'p qatlamli perceptron neyron tarmog'i (MLP), Jordan/Elman NN, modulli neyron tarmog'i (MNN), to'g'ridan-to'g'ri uzatiladigan umumiy neyron tarmog'i (GFFNN), o'z-o'zini tashkil qiluvchi xususiyatlar xaritasi (SOFM), SVM neyron tarmog'i, ehtimollik neyron tarmog'i (PNN) va takroriy neyron tarmog'i (RNN). Ularning yakuniy hisoblash natijalari shuni ko'rsatadiki, PNN ushbu tadqiqotda qo'llanilgan boshqa NN modellaridan 98,24% aniqlik bilan ustundir. Biroq, bu tadqiqot boshqa ko'plab tadqiqotlar kabi izohlanmaydi, chunki u bashorat bosqichida qaysi xususiyatlar eng muhimligini ko'rsatmaydi.

**REFERENCES**

1. Samuel, A.L. Dama o'yinidan foydalangan holda Mashinali o'rgatish bo'yicha ba'zi tadqiqotlar. *IBM J. Res. Dev.* 1959, 3, 210–229.
2. Brownlee, J. Python bilan Mashinali o'rgatish mahorati. *Maks. O'qish. Mastery Pty Ltd.* 2016, 527, 100–120.
3. Husayn, E. H.; Emam, MM; Ali, AA; Sugantan, P.N. Tibbiy tasvirga asoslangan ko'krak bezi saratonini davolash uchun chuqur va Mashinali o'rgatish usullari: keng qamrovli sharh. *Ekspert tizimi. dek.* 2021, 167, 114161.
4. Ansari, Kaliforniya; Gupta, N.K. Neyro-loyqa integratsiyalashgan tizim yordamida yurak-qon tomir kasalliklarining avtomatlashtirilgan diagnostikasi. *Axborot va kommunikatsiya texnologiyalari bo'yicha Butunjahon kongressi materiallarida 2011 yil, Mumbai, Hindiston, 2011 yil 11-14 dekabr; 1379–1384-betlar.*