

Л.Ш. Черикбаева¹, Б.Е. Түркістан^{1*}

¹Әл-Фараби атындағы Қазақ ұлттық университеті, Алматы қ., Қазақстан
*e-mail: Baha@mail.ru

МЕДИЦИНАЛЫҚ МӘЛІМЕТТЕРДІ ӨНДЕУДЕ МАШИНАЛЫҚ ОҚЫТУДЫҢ ТИІМДІ АЛГОРИТМДЕРІН ҚОЛДАНУ ЖӘНЕ ЗЕРТТЕУ

Аңдатпа

Денсаулық сақтаудағы машиналық оқыту пациенттердің нәтижелерін жақсарту үшін Интернет заттары ұсынатын медициналық деректердің өсіп келе жатқан көлемін пайдаланады. Бұл әдістер перспективалы қолданбаларды, сондай-ақ күрделі проблемаларды қамтамасыз етеді. Машиналық оқыту үш негізгі салада қолданылады: медициналық бейнелеу, медициналық құжаттарды табиғи тілде өңдеу және генетикалық ақпарат. Бұл салалардың көпшілігі диагностикаға, анықтауға және болжауға бағытталған. Медициналық құрылғылардың үлкен инфрақұрылымы қазіргі уақытта деректерді жасайды, бірақ мұндай деректерді тиімді пайдалану үшін қолдау көрсететін инфрақұрылым жиі болмайды. Денсаулық туралы ақпараттың көптеген әртүрлі нысандарының болуы деректерді пішімдеу кезінде кейбір мәселелерді тудырады және шуды арттыруы мүмкін. Машиналық оқытудың қысқаша тарихын, денсаулық сақтаудағы осы технологияның әдістері мен қазіргі жағдайы туралы кейбір негізгі білімдер қарастырылады. Мақалада жүрек ауруларының негізінде LogisticRegression, RandomForestClassifier, DecisionTreeClassifier, KNeighborsClassifier, GradientBoostingClassifier алгоритмдер қолданылып салыстырылды.

Түйін сөздер: машиналық оқыту, медицина, алгоритмдер, BigData, жүрек аурулары.

Аннотация

Л.Ш. Черикбаева¹, Б.Е. Түркістан¹

¹Казахский национальный университет им. аль-Фараби, г. Алматы, Казахстан

ПРИМЕНЕНИЕ И ИССЛЕДОВАНИЕ ЭФФЕКТИВНЫХ АЛГОРИТМОВ МАШИННОГО ОБУЧЕНИЯ ПРИ ОБРАБОТКЕ МЕДИЦИНСКИХ ДАННЫХ

Машинное обучение в здравоохранении использует все больше медицинских данных, предоставляемых Интернетом вещей, для улучшения результатов лечения пациентов. Эти методы обеспечивают многообещающие приложения, а также сложные проблемы. Машинное обучение используется в трех основных областях: медицинская визуализация, обработка медицинских документов на естественном языке и генетическая информация. Многие из этих областей сосредоточены на диагностике, обнаружении и прогнозировании. Большая инфраструктура медицинских устройств в настоящее время генерирует данные, но часто отсутствует инфраструктура для поддержки эффективного использования таких данных. Наличие множества различных форм медицинской информации вызывает некоторые проблемы при форматировании данных и может увеличить шум. Рассмотрена краткая история машинного обучения, некоторые базовые знания о методах и текущем состоянии этой технологии в здравоохранении. В статье сравниваются алгоритмы LogisticRegression, RandomForestClassifier, DecisionTreeClassifier, KNeighborsClassifier, GradientBoostingClassifier на основе болезней сердца.

Ключевые слова: машинное обучение, медицина, алгоритмы, BigData, болезни сердца.

Abstract

APPLICATION AND RESEARCH OF EFFECTIVE MACHINE LEARNING ALGORITHMS IN MEDICAL DATA PROCESSING

Cherikbayeva L.Sh.¹, Turkistan B.Y.¹

¹Al-Farabi Kazakh National University, Almaty, Kazakhstan

Machine learning in healthcare uses an increasing amount of medical data provided by Internet of Things to improve patient outcomes. These methods provide promising applications, as well as complex problems. Machine learning is used in three main areas: medical imaging, processing medical documents in natural language, and genetic information. Many of these areas focus on diagnosis, detection and prognosis. The large infrastructure of medical devices currently generates data, but there is often no infrastructure to support the efficient use of such data. The presence of many different forms of health information causes some problems when formatting data and can increase noise. A brief history of machine learning, some basic knowledge about the methods and current state of this technology in health care are considered.

The article compares LogisticRegression, RandomForestClassifier, DecisionTreeClassifier, KNeighborsClassifier, GradientBoostingClassifier algorithms based on heart disease.

Keywords: machine-learning, medicine, algorithms, BigData, heart disease.

Кіріспе

Денсаулық сақтау саласындағы цифрлық технологиялардың пайда болуы қолдануда да, практикалық жағынан да үздіксіз қиындықтармен сипатталады. Өртүрлі денсаулық сақтау жүйелерін біріктіру баяу болды және әлемнің көптеген бөліктерінде толық интеграцияланған денсаулық сақтау жүйесін қабылдау орындалмады [1]. Адам биологиясының тән табиғаты мен күрделілігі, сондай-ақ жекелеген пациенттер арасындағы әртүрлілік ауруларды диагностикалау мен емдеуде адам элементінің маңыздылығын дәйекті түрде көрсетті. Дегенмен, цифрлық технологиялардағы жетістіктер денсаулық сақтау мамандары үшін пациенттерге ең жақсы көмек көрсетуде таптырмас құралға айналатыны сөзсіз.

Сақтау көлемін, есептеу қуатын және деректерді беру жылдамдығын қоса алғанда, деректер технологияларын жақсарту көптеген салаларда, соның ішінде денсаулық сақтауда машиналық оқытуды кеңінен қолдануға мүмкіндік берді [2]. Жеке адамға сапалы медициналық көмек көрсетудің көп нұсқалы сипатына байланысты медицинадағы соңғы тенденциялар денсаулық сақтауда жекелендірілген медицина немесе «дәлдік медицина» тәсілінің қажеттілігін атап өтті. Жекелендірілген медицинаның мақсаты диагностикалық шешімдерді табу, болжау және талдау үшін денсаулық сақтау деректерінің үлкен көлемін пайдалану болып табылады, бұл дәрігерлер өз кезегінде әрбір жеке пациент үшін жүзеге асыра алады. Мұндай деректер генетикалық немесе отбасылық ақпаратты, медициналық бейнелеу деректерін, дәрілік комбинацияларды, жалпы пациенттердің денсаулығының нәтижелерін және қолданыстағы медициналық құжаттаманың табиғи тілде өңделуін қамтиды, бірақ олармен шектелмейді [3].

Ең алдымен медициналық және биомедициналық салалардағы машиналық оқытудың (ML) ең үлкен үш қосымшасына назар аударылады. Жылдам дамып келе жатқан сала ретінде денсаулық сақтау саласында машиналық оқытудың әлеуетті қосымшаларының кең ауқымы бар, олар персоналды басқару, сақтандыру саясаты, реттеу істері және т.б. сияқты саланың қосалқы аспектілерін қамтуы мүмкін [4]. Осылайша, осы тарауда қарастырылатын тақырыптар машиналық оқытудың үш жалпы қолданбасына дейін тарылды.

Біріншісі – магниттік-резонанстық томография (МРТ), компьютерленген осьтік томография (САТ) сканерлеуі, ультрадыбыстық (АҚШ) және позитронды-эмиссиялық томография (ПЭТ) сканерлері сияқты медициналық кескіндерде машиналық оқытуды пайдалану [5]. Бұл кескіндеу әдістерінің нәтижесі әдетте радиологты түсіндіру және диагноз қоюды талап ететін кескіндер жиынтығы немесе сериясы болып табылады. ML әдістері аурудың күйін немесе күрделі мәселені көрсете алатын кескіндерді болжау және табу үшін жылдам дамып келеді.

Екіншісі – медициналық құжаттарды табиғи тілде өңдеу. Көптеген елдерде электронды медициналық жазбаларға (EMR) итермелей отырып, көптеген денсаулық сақтау мамандарының келісімі бұл процесс баяу, жалықтырады және көптеген жағдайларда толығымен бұзылады. Бұл кейде пациенттердің жалпы денсаулығының нашарлауына әкелуі мүмкін [6]. Негізгі қиындықтардың бірі - көптеген ауруханалар мен емханаларда бар физикалық медициналық жазбалар мен құжаттаманың көлемі. Түрлі пішімдеу, қолмен жазылған жазбалар және толық емес немесе орталықтандырылмаған ақпараттың көптігі электронды медициналық жазбаларды қабылдауға көшуді тиімдірек етті.

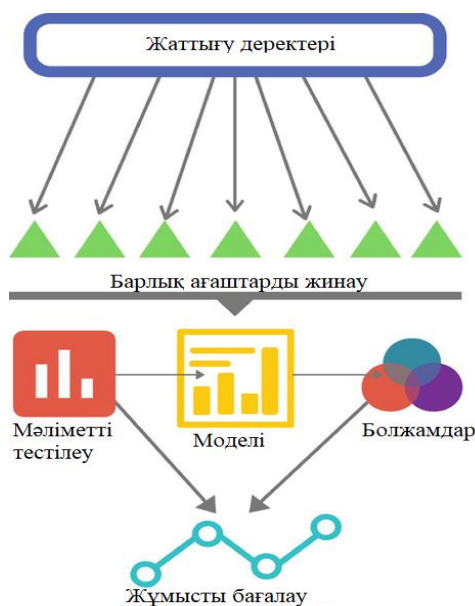
Үшінші машиналық оқыту қолданбасы ауруды болжау және аурудың себептерін табу үшін адам генетикасын пайдалануды қамтиды. Келесі ұрпақ секвенциясы (NGS) әдістерінің пайда болуымен және генетикалық деректердің жарылуымен, соның ішінде жалпы популяциялық генетикалық ақпараттың үлкен дерекқорларымен, генетиканың адам денсаулығына қалай әсер ететіні туралы маңызды ақпаратты анықтау әрекеті қазір көптеген зерттеулердің алдыңғы қатарында. Күрделі аурулардың қалай көрінетінін және генетиканың жеке адамның тәуекелін қалай арттыруы немесе азайтуы мүмкін екенін түсіну денсаулық сақтаудың алдын алуға көмектеседі [7]. Бұл дәрігерлерге күрделі ауруларды жұқтыру қаупін азайту үшін нақты пациенттерге күтім жасау жоспарын қалай бейімдеу керектігі туралы көбірек ақпарат бере алады.

Осы үш тақырыптың барлығында кездесетін ортақ мәселе - заттар интернетінен алынған денсаулық деректерін пациенттер мен дәрігер үшін түсінікті, пайдалы, сенімді ақпаратқа қалай аудару керек.

Арвиндер Диллон «Денсаулық сақтау деректерін талдаудағы машиналық оқыту» мақаласында деректердің осы алуан түрін талдау үшін болжамды жақсарту үшін бақыланатын, бақыланбайтын және күшейтілген алгоритмдер сияқты әртүрлі машиналық оқыту алгоритмдері пайдаланылады, оларды дәлдік, сезімталдық, ерекшелік, дәлдік, F1 ұпайы және қисық астындағы аумақ сияқты әртүрлі өнімділік параметрлері арқылы талдауға болатынын көрсеткен [8]. Бұл мақалада машиналық оқыту алгоритмдері анықталған және клиникалық, омикалық және сенсорлық деректер сияқты денсаулық сақтау деректерінің әртүрлі түрлерін талдау үшін машиналық оқыту алгоритмдерін пайдалану орындалған. Сауалнамадан денсаулық сақтау саласындағы деректердің әртүрлі түрлерін талдау үшін әртүрлі авторлар онкологиялық науқастардың өмір сүруін болжау үшін әртүрлі машиналық оқыту алгоритмдері мен мүмкіндіктерді алу әдістерін ұсынғаны туралы қорытындыға келді [9].

Джусси Тохка, Маркван Гилсб «Денсаулық пен сауықтыру қосымшалары үшін машиналық оқыту алгоритмдерін бағалау» мақаласында деректерге негізделген тәсілдерді пайдалана отырып, зерттеулер мен әзірлемелердегі күш-жігердің артуы оң және теріс әсер етеді. Оң жағы, жаңа қосымшалар мен шешімдер он жыл бұрын әлі де шешуге қиын деп саналатын мәселелерге жеткізіледі. Мысалы, кескінді тану, сөйлеуді тану, жалпы табиғи тілді өңдеу және медициналық технологиядағы, әсіресе кескінді талдау және көмекші диагностика саласындағы жоғары табысты жетістіктер туралы ойланыңыз. Жағымсыз жағы – AI және машиналық оқыту сияқты әдістерге қатысты шындыққа жанаспайтын күтулер тудыратын жағдай бар [10]. Жоғары үміттерді соңғы пайдаланушылар да, тұтынушылар да, сондай-ақ ғылыми қауымдастықтың өзі де қояды. Бұл нәтижені мұқият объективті бағалау қысымға ұшырайтын міндет болатын жағдайға әкеледі. Көріп отырғанымыздай, өнімділікті дұрыс бағалау тривиальды емес (мәселе мен деректерді түсінуді талап етеді) және көбінесе қымбат (деректерді сақтау қажет) және уақытты қажет етеді. Бұл әзірлеу/жаттығу жұмысының өзіне қарағанда жиі қызықты емес және объективті тексеруден кейінгі өнімділік бағалауы әзірлеудің бастапқы кезеңінде алынған «жоғары перспективалы» нәтижелерден аз болатын жағымсыз қасиеті бар [11].

Бақыланатын оқыту әдетте белгіленген деректер деп аталатын оқу деректерін пайдаланады. Жаттығу деректерінде бір немесе бірнеше кіріс бар және «белгіленген» шығысы бар. Модельдер жаңа деректерді болжауды жақсарту мақсатында (яғни, сынақ деректерінің жинағы) жаттығу кезінде өздерін бағалау үшін осы белгіленген нәтижелерді пайдаланады. Әдетте, бақыланатын оқыту үлгілері жіктеу және регрессия алгоритмдеріне бағытталған [12]. Медицинада жіктеу мәселелері өте жиі кездеседі. Көптеген клиникалық жағдайларда науқасты диагностикалау дәрігердің белгілі бір белгілер жиынтығын ескере отырып, ауруды жіктеуді қамтиды. Регрессия проблемалары өмірлік маңызды белгілер, ауру тарихы және салмақ сияқты белгілі бір деректер жиынтығын ескере отырып, ауруханада болжалды болу ұзақтығы сияқты сандық нәтижелерді болжауға бейім (Сурет 1).



Сурет 1. Жаттығу деректері

Нәтижесі

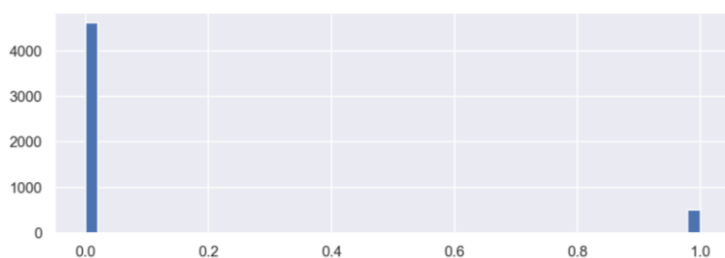
1. Деректер жиынын импорттау және зерттеу үшін жалпы деректер қоры құрылды. Деректер қорында 5000-ға жуық мәліметтер жиналды. Мәліметтер келесі айнымалаылар бойынша құрылды (2-сурет):

1. id: бірегей идентификатор.
2. жынысы: «ер», «әйел» немесе «басқа».
3. жасы: науқастың жасы.
4. гипертония: науқаста гипертония болмаса 0, гипертония болса 1.
5. жүрек_ауруы: егер науқаста жүрек ауруы болмаса 0, науқаста жүрек ауруы болса 1.
6. ever_married: «Жоқ» немесе «Иә».
7. жұмыс_түрі: "балалар", "Үкімет", "Ешқашан_жұмыс істемеген", "Жеке" немесе "Өзін-өзі жұмыспен қамтығандар".
8. тұрғын үй түрі: «Ауылдық» немесе «Қалалық».
9. орташа_глюкоза_деңгейі: қандағы глюкозаның орташа деңгейі.
10. BMI: дене салмағының индексі.
11. smoking_status: "бұрын темекі шегетін", "ешқашан темекі шекпеген", "темекі шегетін" немесе "белгісіз"*.
12. инсульт: егер науқаста инсульт болған болса 1 немесе жоқ болса 0.

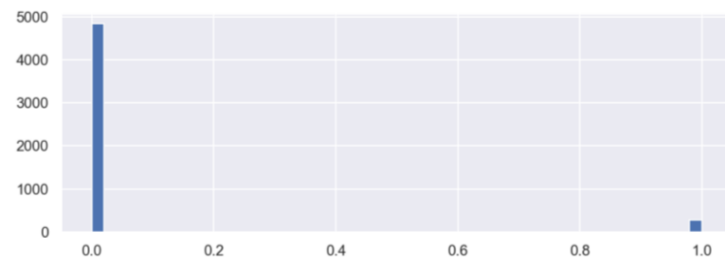
	id	gender	age	hypertension	heart_disease	ever_married	work_type	Residence_type	avg_glucose_level	bmi	smoking_status	stroke	
	2461	35446	Male	73.00	0	0	Yes	Govt_job	Rural	208.69	30.0	Unknown	0
	449	6171	Male	6.00	0	0	No	children	Urban	90.60	16.6	Unknown	0
	1861	26830	Female	47.00	0	0	Yes	Self-employed	Rural	68.37	29.4	smokes	0
	1113	15418	Female	80.00	0	0	Yes	Self-employed	Rural	90.43	34.2	never smoked	0
	5015	71590	Female	5.00	0	0	No	children	Rural	102.04	18.5	Unknown	0
	1698	24256	Male	35.00	0	0	Yes	Private	Rural	108.08	30.6	formerly smoked	0
	3088	44676	Male	1.64	0	0	No	children	Urban	115.12	21.1	Unknown	0
	1101	15266	Female	32.00	0	0	Yes	Private	Rural	77.67	32.3	smokes	0
	1843	26389	Female	2.00	0	0	No	children	Urban	120.85	16.2	Unknown	0
	689	9404	Female	44.00	0	0	Yes	Private	Rural	107.41	47.3	never smoked	0

Сурет 2. Деректер қоры

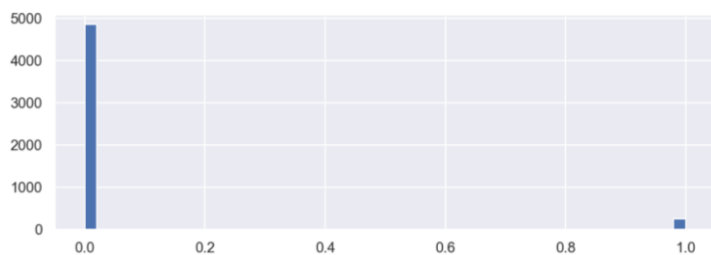
Медициналық мәліметтерді өңдеуде гипертонияның онымен таралуы, жүрек ауруларының таралуы, инсульт аурулары келесі суреттерде көрсетілген (Сурет 3, 4, 5):



Сурет 3. Гипертония деректері

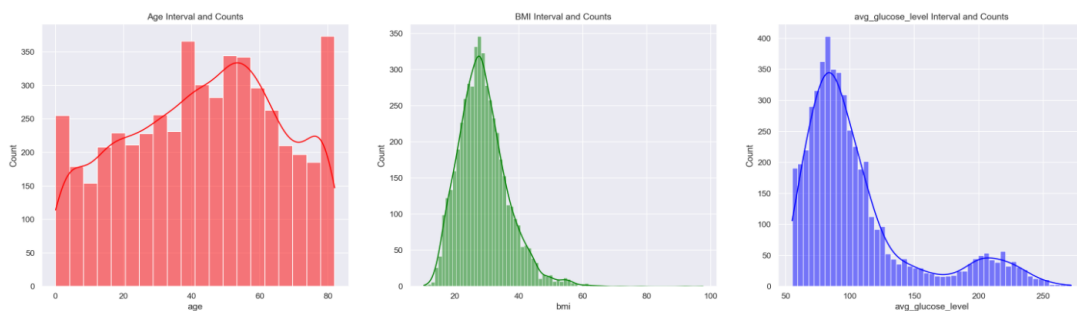


Сурет 4. Жүрек ауруларының деректері



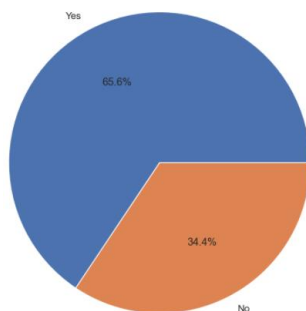
Сурет 5. Инсульт ауруларының деректері

6-суретте жас аралығын, BMI аралығы және avg_glucose_level аралық арқылы санау диаграмма түрінде көрсетілген.



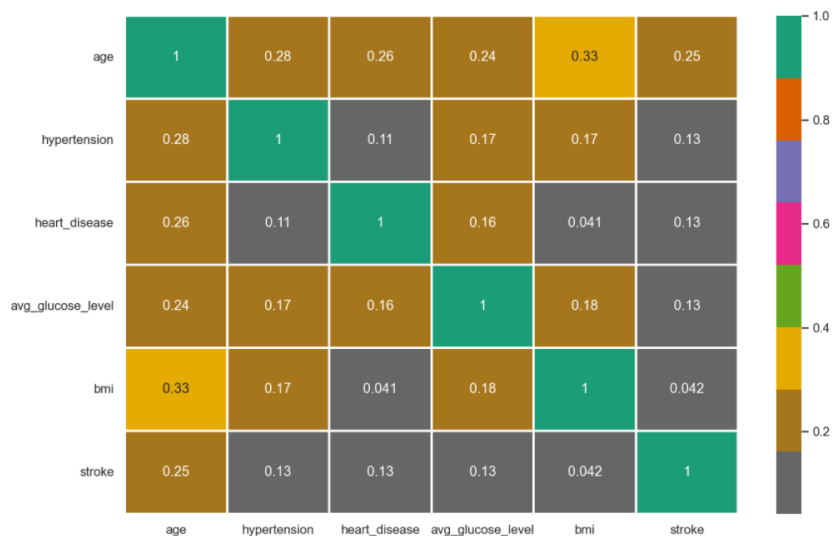
Сурет 6. Сандар

Жыныс бойынша инсульттің орташа көрсеткіші. Ер адамдарда инсульт жиілігі жоғары. Инсульт бойынша тұрмыс құрғандар мен құрмағандардың айырмашылығы 7-суретте көрсетілген.



Сурет 7. Инсульт деректерін бөлу

Корреляция бір немесе бірнеше айнымалылардың бір-бірімен байланысын түсіндіреді. Бұл айнымалылар мақсатты айнымалыны болжау үшін пайдаланылған кіріс деректер мүмкіндіктері болуы мүмкін (Сурет 8).

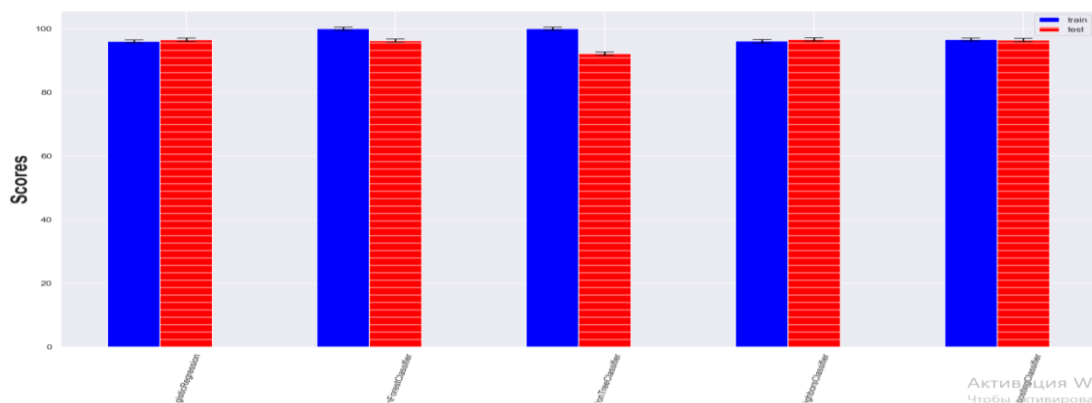


Сурет 8. Корреляция мәндері

Модельдерді өңдеу кезінде пайдаланылған алгоритмдер:

1. LogisticRegression;
2. RandomForestClassifier;
3. DecisionTreeClassifier;
4. KNeighborsClassifier
5. GradientBoostingClassifier.

Модельдерді оқыту және тестілеу кезеңі 9-суретте көрсетілген.



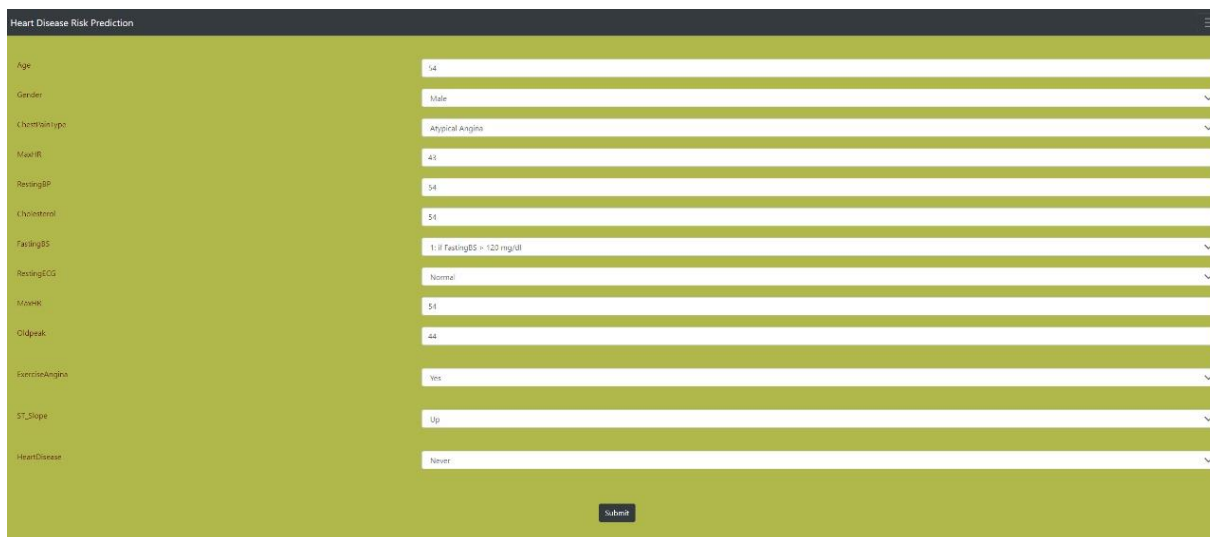
Сурет 9. Модельдерді оқыту және тестілеу

Модельдерді оқыту және тестілеу кезеңінде медициналық мәліметтерді өндеуде машинаны оқытудың тиімді алгоритмдерді пайдаланудың нәтижелері 1-кестеде көрсетілген:

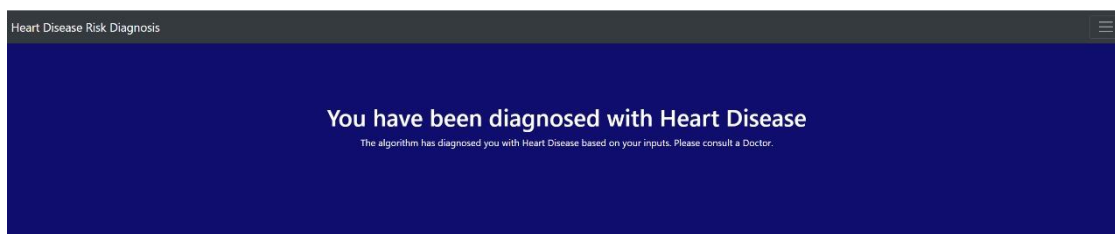
Кесте 1. Машиналық оқыту алгоритмдерінің нәтижесі

Алгоритмдер	Нәтижелері
LogisticRegression	96.52432969215492
RandomForestClassifier	96.22641509433963
DecisionTreeClassifier	92.15491559086395
KNeighborsClassifier	96.62363455809336
GradientBoostingClassifier	96.42502482621649

Осы нәтижелерге сүйене отырып KNeighborsClassifier алгоритмі салыстырмалы түрде жақсы нәтиже көрсетті. Машиналық оқытудың негізінде Flask фреймворкы арқылы электрондық веб-сайт әзірленді (Сурет 10, 11).



Сурет 10. Веб-сайттың көрінісі



Сурет 11. Жүрек ауруларын енгізгендегі шығатын нәтиже

Қорытынды

Денсаулық сақтау саласында машиналық оқыту сияқты цифрлық технологияларды қолдану қызықты дәуірге енуде. Информатика, биология, инженерия, химия және информатиканың соқтығысуы күрделі аурулардың пайда болуына ықпал ететін тұқым қуалайтын және қоршаған орта факторлары туралы білімімізді тездетеді. Қатерлі ісік диагнозын болжауда көшірме санының вариациясын пайдалану әлеуеті қызықты. Тұқым қуалайтын қатерлі ісік қаупіне ықпал ету үшін геномдық ландшафттың өзара байланысын түсінудің түсіндірілетін әдісін жасау үшін машиналық оқытуды пайдалану пациенттің денсаулығын жеке деңгейде жақсартуы мүмкін. Кескінді талдау көптеген диагностикалық әрекеттердің негізгі бөлігі болып табылады және радиологиялық диагностиканың дәлдігін жақсартуды жалғастырады. Қатерлі массаларды анықтау және бар диагнозды растау және тексеру қателерді азайта отырып, пациенттің нәтижелерін жақсартуға мүмкіндік береді. Жалпы жасалынған жұмыстар бойынша бұл бағдарлама толығымен жасалынды. Жұмыс аясында келесі міндеттер шешілді:

1. Медициналық деректердің мәселесі зерттелді, оны шешудің тәсілдері талданды.
2. Жүрек ауруларының негізінде мәліметтер қоры құрылды құрылды.
3. LogisticRegression, RandomForestClassifier, DecisionTreeClassifier, KNeighborsClassifier, GradientBoostingClassifier алгоритмдері қарастырылды және жүрек ауруларын анықтағанда шешімдері салыстырылды.
5. Әдістердің тиімділігіне тестілеу жүргізілді.

Бұл зерттеу ҚР Білім және ғылым министрлігі №BR10965172 ғылыми жобасы аясында орындалды.

Пайдаланылған әдебиеттер тізімі:

1 Аднан Н., Нордин Шахрина М.д., Рахман И., Нур А. Влияние передачи знаний на принятие фермерами решений в отношении устойчивых методов ведения сельского хозяйства. World J Sci Technol Sustain Dev. 2018.

- 2 Чен Т., Гестрин К., Xgboost: Масштабируемая система повышения дерева, в: Труды 22-й Международной конференции Acta Sigkdd по обнаружению знаний и интеллектуальному анализу данных, стр. 785–794.
- 3 Фукусима К. Когнитрон: Самоорганизующаяся многослойная нейронная сеть. Биологическая кибернетика. 1975;20(3-4):121-136
- 4 Датта, С., Баруа, Р., Дас, Дж. Применение искусственного интеллекта в современной системе здравоохранения. Недавнее использование этого природного полимера в альгинатах; IntechOpen: Риека, Хорватия, 2020 г.
- 5 Raheja K, Dubey A, Chawda R. Анализ данных и его важность в здравоохранении. *Int Computer Trends and Technology J.* 2018;48:176-180.
- 6 Коциантис С.Б. Контролируемое машинное обучение: обзор методов классификации. Информатика (Люблина). 2007;31(3):249-268
- 7 Сениор А.В., Эванс Р., Джампер Дж., Киркпатрик Дж., Сифре Л., Грин Т. и другие. Улучшенное предсказание структуры белка с использованием потенциалов глубокого обучения. *Природа.* 2020; 577 (7792): 706–10.
- 8 Сринивасан С., Джонсон Н.Т., Коркин Д. Гибридный подход с глубокой кластеризацией для надежного профилирования типов клеток с использованием данных секвенирования РНК одной клетки. *БиоРксив.* 2019:511626.
- 9 Букерш А., Ван Дж. Модели прогнозирования трафика на основе машинного обучения для интеллектуальных транспортных систем. *Компьютерная сеть.* 2020;181
- 10 Шувал, Р.; Фейн, Дж. А.; Савани, Б.; Мохти, М.; Наглер, А. Машинное обучение и искусственный интеллект в гематологии. бр. Дж. Гематол. 2021, 192, 239–250.
- 11 А.С. Лундерволд, А. Лундерволд Обзор глубокого обучения в медицинской визуализации с акцентом на МРТ *Z. Med. Phys.*, 29 (2019), стр. 102-127
- 12 Фридман Дж.Х. Аппроксимация жадной функции: машина повышения градиента. *Анналы статистики.* 2001;29(5):1189-1232
- 13 Панда Б., Маджи Б. Новое улучшенное предсказание структурного класса белка с использованием глубокой рекуррентной нейронной сети. *Эволюционный интеллект.* 2018;1-8.
- 14 Фудзиёси Х., Хиракава Т., Ямасита Т. Распознавание изображений на основе глубокого обучения для автономного вождения. *Рез. IATSS.* 2019;43(4):244–52.
- 15 К. Амбруаз, Г.Дж. McLachlan Предвзятость отбора при извлечении генов на основе данных экспрессии генов на микрочипах *Proc. Natl. akad. nauch. Blok. SHtaty Ameriki*, 99 (2002), стр. 6562-6566.

References:

- 1 Adnan N., Nordin SHahrina M.d., Rahman I., Nur A. (2018) Vliyanie peredachi znaniy na prinyatie fermerami reshenij v otnošenii ustojchivyh metodov vedeniya sel'skogo hozyajstva [Impact of knowledge transfer on farmers' decision-making on sustainable agricultural practices]. *World J Sci Technol Sustain Dev.* (In Russian)
- 2 T. CHen, K. Gestrin, Xgboost: Mashtabiruemaya sistema povysheniya dereva [Scalable Tree Enhancement System]. в: *Trudy 22-j Mezhdunarodnoj konferencii Acta Sigkdd po obnaruzheniyu znaniy i intellektual'nomu analizu dannyh.* 785–794. (In Russian)
- 3 Fukusima K. Kognitron: (1975) Samoorganizuyushchayasya mnogoslajnaya nejronnaya set'. *Biologicheskaya kibernetika* [Self-organizing multi-layered neural network. *Biological cybernetics*]. 20(3-4):121-136. (In Russian)
- 4 Datta, S., Barua, R., Das, Dzh. (2020) Primenenie iskusstvennogo intellekta v sovremennoj sisteme zdavoohraneniya. Nedavnee ispol'zovanie etogo prirodnogo polimera v al'ginatah [The use of artificial intelligence in the modern healthcare system. Recent use of this natural polymer in alginates]. *IntechOpen: Rieka, Horvatiya.* (In Russian)
- 5 Raheja K, Dubey A, Chawda R. (2018) Analiz dannyh i ego vazhnost' v zdavoohranenii [Data analysis and its importance in healthcare]. *Int Computer Trends and Technology J.* 48:176-180. (In Russian)
- 6 Kociantis S.B. (2007) Kontroliruemoe mashinnoe obuchenie: obzor metodov klassifikacii [Supervised machine learning: an overview of classification methods]. *Informatika (Lyublyana).* 31(3):249-268. (In Russian)
- 7 Senior A.V., Evans R., Dzhampier Dzh., Kirkpatrick Dzh., Sifre L., Grin T. i drugie. (2020) Uluchshennoe predskazanie struktury belka s ispol'zovaniem potencialov glubokogo obucheniya [Improved prediction of protein structure using deep learning potentials]. *Priroda.* 577 (7792): 706–10. (In Russian)
- 8 Srinivasan S., Dzhonson N.T., Korkin D. (2019) Gibridnyj podhod s glubokoj klasterizaciej dlya nadezhnogo profilirovaniya tipov kletok s ispol'zovaniem dannyh sekvenirovaniya RNK odnoj kletki [A hybrid approach with deep clustering for reliable cell type profiling using single cell RNA sequencing data] *BioRksiv.* 511626. (In Russian)
- 9 Bukersh A., Van Dzh. (2020) Modeli prognozirovaniya trafika na osnove mashinnogo obucheniya dlya intellektual'nyh transportnyh sistem [Traffic forecasting models based on machine learning for intelligent transport systems.]. *Komp'yuternaya set'.* 181. (In Russian)

10 SHuval, R.; Fejn, Dzh. A.; Savani, B.; Mohti, M. ; Nagler, A. (2021) *Mashinnoe obuchenie i iskusstvennyj intellekt v gematologii [Machine learning and artificial intelligence in hematology] br. Dzh. Gematol. 192, 239–250. (In Russian)*

11 A.S. Lundervold, A. Lundervol'd (2019) *Obzor glubokogo obucheniya v medicinskoj vizualizacii s akcentom na MRT [Overview of Deep Learning in Medical Imaging with an emphasis on MRI] Z. Med. Phys., 29, 102-127. (In Russian)*

12 Fridman Dzh.H. (2001) *Approksimaciya zhadnoj funkicii: mashina povysheniya gradienta [Approximation of a greedy function: a gradient boosting machine]. Annaly statistiki. 29(5):1189-1232. (In Russian)*

13 Panda B., Madzhi B. (2018) *Novoe uluchshennoe predskazanie strukturnogo klassa belka s ispol'zovaniem glubokoj rekurrentnoj nejronnoj seti [A new improved prediction of the structural class of a protein using a deep recurrent neural network]. Evolyucionnyj intellekt. 1-8. (In Russian)*

14 Fudziyosi H., Hirakava T., YAmasita T. (2019) *Raspoznavanie izobrazhenij na osnove glubokogo obucheniya dlya avtonomnogo vozheniya [Deep learning-based image recognition for autonomous driving] Rez. IATSS. 43(4):244–52. (In Russian)*

15 K. Ambrucz, G.Dzh. McLachlan (2002) *Predvzyatost' otbora pri izvlechenii genov na osnove dannyh ekspressii genov na mikrochipah [Selection bias in gene extraction based on gene expression data on microarrays] Proc. Natl. akad. nauch. Blok. SHtaty Ameriki, 99. 6562-6566. (In Russian)*