8D06114 – «Медицина саласындағы жасанды интеллект» білім беру бағдарламасы бойынша философия докторы (PhD) дәрежесін алу үшін ұсынылған Сулейменова Мадина Утегеновнаның «Жүрек-қан тамырлары ауруларының даму қаупін болжаудағы машиналық оқыту» тақырыбындағы диссертациясына

АҢДАТПА

Зерттеу тақырыбының өзектілігі. ЖҚА Қазақстанда және бүкіл әлемде өлім мен мүгедектіктің негізгі себебі болып табылады және өмір сүру ұзақтығына әсер етеді. ДДҰ деректеріне сәйкес, Қазақстан жүрек-қан тамырлары ауруларының (ЖҚА) даму қаупі жоғары өңірлерге жатады. Қазақстанда 2021 жылы сырқаттанушылық пен өлім – жітімнің өсуіне байланысты өмір сүру ұзақтығы 70,23 жасқа дейін төмендеді (2020 жылы-71,37). Ал соңғы екі жылда орташа өмір сүру ұзақтығы 2,95 жылға қысқарды, бұл 2013 жылғы деңгейге сәйкес келеді. Бұл көрсеткіштер жүрек-қан тамырларының ерте қартаюын болжауда және теріс нәтижелердің алдын алуда шешуші рөл атқаратын маркерлерді анықтау қажеттілігін көрсетеді. Машиналық оқыту және модельдеу сияқты жасанды интеллект (AI) әдістерін қолдану жеке тәуекелдерді анықтауға және ЖҚА даму тәуекелдерін болжауға мүмкіндік береді, бұл қазіргі заманғы медицинаның маңызды бөлігіне айналады, өйткені ол талдауды, болжауды және басқа процестерді жеделдетеді.

Машиналық оқыту болжаудың икемді негізін ұсынады, бірақ клиникалық енгізу қатаң валидация мен түсіндіруді қажет етеді. Біздің зерттеуде модель сапасы тек ROC-AUC арқылы ғана емес, сонымен қатар деректердің ағуынсыз орындалған ішкі валидация, калибровканы бағалау, Decision Curve Analysis негізіндегі клиникалық пайдалық және коэффициенттер, permutation importance, SHAP арқылы интерпретациялау тәсілдерімен кешенді түрде қарастырылды. Елеулі әдістемелік мәселе development-іріктеу аз болған кезде туындайды: дәл тексерусіз сапа жоғары болып көрінуі мүмкін, белгілерді таңдау тұрақсыз болуы мүмкін және модельдің төзімділігі белгісіз болады.

ЖҚА даму қаупін бағалау мультиморбидтілікпен, метаболикалық профильдердің гетерогенділігімен және популяцияға тән базалық тәуекелмен қиындайды, бұл кемсітушілікке де, ең бастысы ықтималдық болжамын калибрлеуге де әсер етуі мүмкін. Көптеген дәстүрлі тәуекел шкалалары жас аралас когорттарда жасалған және егде жастағы топтарға немесе жаңа жағдайларда қолданылған кезде шектеулі төзімділікті көрсете алады.

Осылайша, әдеттегі тәжірибеде болжамды құралдың пайдалылығы оның пациенттерді тәуекел дәрежесі бойынша бағалау қабілетімен ғана емес, сонымен қатар болжамды ықтималдықтардың шекті шешімдер қабылдау үшін қаншалықты калибрленгенімен және модельдің мінез-құлқын клиникалық тұрғыдан қаншалықты түсіндіруге болатынымен анықталады.

Жоғарыда айтылғандарға сүйене отырып, осы зерттеудің мақсаты мен міндеттері тұжырымдалған.

Диссертациялық жұмыстың мақсаты. Қайталанатын end-to-end процесі мен машиналық оқыту модельдерін қолдана отырып, 65 жастан асқан адамдарда

жүрек-қан тамырлары ауруларының даму қаупін болжаудың интеллектуалды жүйесін құру.

Зерттеудің міндеттері. Қойылған мақсатқа қол жеткізу үшін келесі міндеттерді шешу қажет:

1. Медицинада аурулардың даму қаупін болжауды анықтау үшін машиналық оқыту модельдерін қолдану мүмкіндіктерін зерттеу.
2. Клиникалық және биохимиялық көрсеткіштері бар 65 жастан асқан тұлғалардың development-іріктемесін қалыптастыру және деректердің "ағып" кетуінсіз, қайталанатын pipeline түрінде алдын ала өңдеу жүргізу.
3. Nested cross-validation әдісін пайдалана отырып, машиналық оқыту модельдерін салыстырмалы талдау, оқыту және out-of-fold (OOF) ықтималдықтарын алу.
4. Модельдердің сапасын дискриминациялық көрсеткіштер мен ықтималдық болжамдар бойынша бағалау, ықтималдықтардың калибровкасын тексеру, болжамдарды интерпретациялау және белгілер маңыздылығының тұрақтылығын permutation importance әдісімен анықтау.
5. Decision Curve Analysis (DCA) негізінде шешім қабылдау шектерін және тәуекел санаттарының шекараларын негіздеу, сондай-ақ development-іріктемеде Low / Intermediate / High тәуекел шкаласын қалыптастыру.
6. Әзірленген шкаланы тәуелсіз когортада сыртқы валидациядан өткізу, модельдің тасымалданғыштығын бағалау және ықтималдықтық рекалибровка жүргізу.

Зерттеу нысаны: Алматы қаласындағы 65 жастан асқан тұлғалардың клиникалық және биохимиялық кестелік мультимодальды деректері.

Зерттеу пәні. Клиникалық және биохимиялық деректер негізінде жүрек-қантамыр ауруларының даму қаупін болжауға арналған машиналық оқыту модельдері.

Зерттеу әдістері. Мультимодальды деректер жиналып, алдын ала өңделді, корреляциялық талдау жүргізілді және болжамдық модельдер (Elastic Net Logistic Regression, Random Forest) құрылды. Гиперпараметрлерді оптимизациялау ішкі және сыртқы валидация (n=695) орындалды. Модель сапасы ROC-AUC, PR-AUC, Brier, LogLoss және калибровка арқылы бағаланды, шешім қабылдау шектері DCA негізінде анықталды, ал факторлар OR, permutation importance және SHAP әдістерімен интерпретацияланды.

Алынған нәтижелердің ғылыми жаңалығы:

Кестелік мультимодальды деректерде ықтимал түсіндірілетін тәуекел шкаласын құру үшін қайталанатын end-to-end процесі ұсынылды, ол мыналарды біріктіреді: деректерді бұзбай стандартталған алдын ала өңдеу, болжау қабілетін шынайы бағалау, ықтималдық болжамдарының дұрыстығын тексеру.

Алғаш рет пациенттерді практикалық стратификациялау және шешім қабылдауды қолдау үшін клиникалық және биохимиялық маркерлеріне негізделген 65+ тұлғалар үшін ЖҚА тәуекелінің ықтималдық шкаласы ұсынылды. Құрал пациентті бағыттау үшін р тәуекел ықтималдығын және санатты (төмен/аралық/Жоғары) береді.

Қорғауға шығарылатын негізгі тұжырым:

1. Қайталанатын end-to-end машиналық оқыту процессі шағын

іріктемеде ЖҚА қаупін 93% дәлдікпен болжады.

2. Модельдің болжамдық қабілеті және ықтималдық болжам сапасы тәуелсіз когортада (n=695) сыртқы валидациямен расталды (ROC-AUC≈0.835). CARDIO RISK KZ интерпретацияланатын ықтималдық шкаласы 65+ тұлғаларда ЖҚА қаупін бағалауға қолдануға жарамды.

Жұмыстың теориялық және практикалық маңызы. Машиналық оқыту үлгілерін пайдалана отырып әзірленген түсіндірілетін ықтималдық шкаласы клиникалық және биохимиялық деректері бар Алматы қаласының тұрғындарының 65 және одан жоғары жастағы пациенттер топтары үшін жүрек-қан тамырлары ауруларының ықтимал даму тәуекелдерін бағалауға мүмкіндік береді. Бұл шкаланы қолдану жеке тәсілді ескере отырып, тиімді профилактикалық және емдік шешімдер қабылдауға мүмкіндік береді.

Сенімділік дәрежесі және апробациялау нәтижелері: Диссертация тақырыбы бойынша зерттеу нәтижелері Scopus дерекқорындағы журналдарда 3 мақала**,** Қазақстан Республикасы Ғылым және жоғары білім министрлігінің Ғылым және жоғары білім саласында сапаны қамтамасыз ету комитеті ұсынатын ғылыми басылымдарда 2 мақала жарияланды, сондай-ақ авторлық құқықпен қорғалатын объектілерге құқықтарды мемлекеттік тізілімге мәліметтер енгізу туралы 2 куәлік алынды.

Scopus базасындағы журналдық мақалалар:

Suleimenova, M., Abzaliyev, K., Abzaliyeva, S., & Nassyrova, N. (2026). An Interpretable and Externally Validated Model for Cardiovascular Disease Risk Assessment in Older Adults. Applied Sciences, 16(8), 3903. https://doi.org/10.3390/app16083903

Abzaliyev, K., Suleimenova, M., Chen, S., Mansurova, M., Abzaliyeva, S., Manapova, A., Kurmanova, A., Bugibayeva, A., Sundetova, D., Bitemirova, R., Baizhigitova, N., Abdykassymova, M., & Sagalbayeva, U. *Predicting Cardiovascular Aging Risk Based on Clinical Data Through the Integration of Mathematical Modeling and Machine Learning.* Applied Sciences**, 2025,** 15(9)**,** 5077. https://doi.org/10.3390/app15095077 **(Scopus: Q2)**

Abzaliyev, K., Suleimenova, M., Abzaliyeva, S., Mansurova, M., Shomanov, A., Bugibayeva, A., Tolemisova, A., Kurmanova, A., & Nassyrova, N. *Multidimensional Visualization and AI-Driven Prediction Using Clinical and Biochemical Biomarkers in Premature Cardiovascular Aging.* **Biomedicines**, 2025, **13(10)**, 2482. https://doi.org/10.3390/biomedicines13102482 **(Scopus: Q1)**

Ғылыми еңбектің негізгі нәтижелерін жариялау үшін Қазақстан Республикасы Ғылым және жоғары білім министрлігінің Ғылым және жоғары білім саласында сапаны қамтамасыз ету комитеті ұсынатын ғылыми басылымдар тізбесіндегі журналдардағы мақалалар:

Suleimenova, M., Chen Siming, Shomanov, A., Abzaliyev, K., Kurmanova, A., Abzaliyeva, S., Abdykassymova, M., Sagalbayeva, U., Bitemirova, R., Sundetova, D., Bugibayeva, A. *AI in Identifying Aging Markers.* Фтизиопульмонология**,** №01(43), 2024, 130–139 б.

https://www.doi.org/10.26212/2227-1937.2024.50.73.018

Сулейменова М.У., Мансурова М.Е. *Искусственный интеллект как инструмент исследования признаков старения.* Вестник НИА РК**,** №2, 2024.

Авторлық құқықпен қорғалатын объектілерге құқықтарды мемлекеттік тізілімге мәліметтер енгізу туралы куәліктер:

«CARDIO RISK KZ: Интеллектуальная система для расчёта индивидуального риска преждевременного сердечно-сосудистого старения» атауымен ЭЕМ-ге арналған бағдарламаға авторлық құқықпен қорғалатын объектілерге құқықтарды мемлекеттік тізілімге мәліметтер енгізу туралы куәлік. № 64683, 2025.

«Программа для вероятностной оценки риска сердечно-сосудистых заболеваний у лиц 65 и старше» атауымен ЭЕМ-ге арналған бағдарламаға авторлық құқықпен қорғалатын объектілерге құқықтарды мемлекеттік тізілімге мәліметтер енгізу туралы куәлік. № 68445, 2026.

Зерттеушінің жеке үлесі. Диссертациялық жұмыстың қойылған мақсаттарына сәйкес міндеттер орындалды. 65+ жастағы тұлғалардың клиникалық және биохимиялық деректері негізінде қайта өндіріле алатын pipeline әзірленіп, деректердің ағып кетуінсіз Elastic Net логистикалық регрессия және Random Forest модельдері құрылды. Модельдер nested cross-validation арқылы бағаланып, ROC-AUC, PR-AUC, Brier, LogLoss және калибровка көрсеткіштері есептелді. Decision Curve Analysis негізінде тәуекел шектері (t_low=0.23, t_high=0.40) анықталып, Low/Intermediate/High шкаласы қалыптастырылды. Әзірленген шкала тәуелсіз сыртқы когортада (n=695) тексеріліп, нәтижелер бағдарламалық калькулятор түрінде іске асырылды.

Жұмыстың көлемі мен құрылымы. Диссертация қазақ тілінде жазылды. Диссертация кіріспеден, 4 тараудан және қорытындыдан тұрады. Диссертацияның жалпы көлемі 86 беттен, 24 суреттен, 10 кестеден, 53 қолданылған әдебиеттен тұрады.

Кіріспеде жұмыстың өзектілігі анықталып, 65 жастан асқан тұлғаларда жүрек-қантамыр ауруларының даму қаупін болжауға арналған клиникалық тұрғыдан қолдануға ыңғайлы және интерпретацияланатын машиналық оқыту модельдерін әзірлеудің қажеттілігі көрсетілді. Сонымен қатар, диссертациялық жұмыстың мақсаты, міндеттері, әдістері, алынған нәтижелердің ғылыми жаңалығы, жұмыстың теориялық және практикалық маңызы жазылып, жарияланымдар тізімі берілген.

Бірінші тарауда жүрек-қантамыр ауруларының, жүрек-қантамырлық қартаюдың және 65 жастан асқан тұлғалардағы қауіп факторларының қазіргі жағдайына әдеби шолу жасалды. Сондай-ақ жүрек-қантамырлық қауіпті бағалауда қолданылатын дәстүрлі шкалалар мен машиналық оқыту әдістеріне талдау жүргізіліп, олардың артықшылықтары мен шектеулері көрсетілді. Егде жастағы популяция үшін тәуекелді бағалаудың интерпретацияланатын және клиникалық тұрғыдан қолдануға ыңғайлы жаңа тәсілдерін әзірлеудің өзектілігі негізделді.

Екінші тарауда зерттеу материалдары мен әдістері сипатталды.

Development cohort (n=100) және external validation cohort (n=695) деректерін дайындау, алдын ала өңдеу, клиникалық және зертханалық көрсеткіштерді іріктеу, корреляциялық талдау және көпөлшемді визуализация әдістері қарастырылды. Сонымен қатар, машиналық оқыту модельдерін құру, nested cross-validation, out-of-fold болжамдарын алу, интерпретация әдістерін қолдану, Decision Curve Analysis негізінде тәуекел шектерін анықтау, сыртқы валидация жүргізу және цифрлық прототипті әзірлеу кезеңдері баяндалды.

Үшінші тарауда жүрек-қантамыр ауруларының қаупін болжауға арналған модельдерді құру және бағалау нәтижелері келтірілді. Логистикалық регрессия, Random Forest және басқа тәсілдер салыстырылып, клиникалық қолдану үшін интерпретацияланатын және тұрақты модель ретінде финалдық нұсқа таңдалды. Модельдердің сапасы ROC-AUC, PR-AUC, Brier score, LogLoss көрсеткіштері бойынша бағаланды, калибровка нәтижелері талданды, ал DCA арқылы t_low = 0.23 және t_high = 0.40 тәуекел шектері негізделді. Сонымен қатар, тәуелсіз сыртқы валидация нәтижелері көрсетіліп, 65 жастан асқан тұлғаларда жүрек-қантамыр аурулары қаупін бағалауға арналған CARDIO RISK KZ интеллектуалды жүйесі ұсынылды.

Төртінші тарауда зерттеу нәтижелеріне талқылау жасалды. Әзірленген жүрек-қантамыр аурулары қаупін болжау моделінің ғылыми және практикалық маңыздылығы негізделіп, клиникалық логистикалық регрессия моделінің неге финалдық модель ретінде таңдалғаны түсіндірілді. Негізгі предикторлар ретінде тұқымқуалаушылық, дене салмағының индексі, физикалық белсенділік және қант диабетінің рөлі көрсетіліп, әзірленген CARDIO RISK KZ шкаласының клиникалық тәжірибеде қолдану мүмкіндіктері, зерттеудің шектеулері мен болашақ даму бағыттары сипатталды.

Қорытындыда диссертацияның негізгі нәтижелері мен қорытындылары берілген.