АҢДАТПА

Тасмурзаев Нұрдәулеттің диссертациялық жұмысының тақырыбы: «Медициналық заттар интернеті (IoMT) және машиналық оқыту әдістері негізінде жүректің ишемиялық ауруын анықтауға арналған интеллектуалды жүйені әзірлеу», философия докторы (PhD) дәрежесін алу үшін ұсынылған «8D07116 – Интеллектуалды басқару жүйелері» білім беру бағдарламасы бойынша

Зерттеудің өзектілігі. Жүректің ишемиялық ауруы (ЖИА) Қазақстан Республикасын қоса алғанда, бүкіл әлемде өлім мен мүгедектіктің жетекші себебі болып қала береді. Дүниежүзілік денсаулық сақтау ұйымының деректері бойынша, жүрек-қан тамырлары аурулары жыл сайын 17 миллионнан астам адамның өмірін қиып отыр. ЖИА-ны ерте әрі нақты диагностикалау тиімді емдеу мен асқынулардың алдын алудың аса маңызды алғышарты болып табылады. Сонымен қатар клиникалық практика тұрақты жүйелік шектеулерге тап болады: білікті мамандардың жетіспеушілігі, диагностикалық рәсімдердің жоғары құны және медициналық көмекке қолжетімділіктің теңсіздігі, әсіресе шалғай және ресурстары шектеулі өңірлерде. Осы тұрғыда клиникалық шешімдерді қабылдауды қолдау үшін машиналық оқытуға негізделген интеллектуалды жүйелерді әзірлеудің маңызы ерекше.

Логистикалық регрессияны және XGBoost, CatBoost, RandomForest сияқты ансамбльдік алгоритмдерді қоса алғанда, машиналық оқытудың дәстүрлі әдістері клиникалық деректер негізінде ЖИА-ны жіктеу үшін кеңінен қолданылады. Алайда бұл әдістердің айтарлықтай шектеулері бар. Біріншіден, клиникалық деректер жиындары әдетте кластардың айтарлықтай теңгерімсіздігімен сипатталады. Екіншіден, дәстүрлі модельдер белгілерді тәуелсіз кіріс айнымалылары ретінде қарастырады, бұл демографиялық көрсеткіштер, зертханалық деректер, ЭКГ нәтижелері және эхокардиография деректері арасындағы күрделі өзара әрекеттестікті лайықты модельдеуге кедергі келтіреді. Үшіншіден, қолжетімді клиникалық когорттардың көлемінің шектеулі болуы жоғары параметрлі архитектураларды қолданғанда шамадан тыс үйрену (overfitting) қаупін арттырады.

Өзіне-өзі назар аудару (self-attention) механизміне негізделген трансформерлік архитектуралар құрылымдалған медициналық деректерді талдау үшін түбегейлі жаңа мүмкіндіктер ашып, модельге белгілер арасындағы өзара байланыстарды айқын ескеруге мүмкіндік береді. Дегенмен деректер жиындарының көлемі шағын болғанда стандартты трансформер архитектуралары шамадан тыс үйренуге бейім, ал кіріс деректерінің клиникалық құрылымын елемеу клиникалық маңызы бар ақпараттың жоғалуына әкеледі.

Медициналық деректердің құпиялылығы мен қауіпсіздігіне қатысты жеке мәселе бар. Клиникалық жазбаларды бұлттық есептеу ортасына жіберу

пациенттердің құпиялылығын бұзу қаупін туғызады. Демек, интеллектуалды медициналық жүйелерді практикада енгізу деректерді құрылғыда (on-device) толықтай жергілікті өңдеуді талап етеді, ал бұл өз кезегінде периферийлік есептеулердің (edge computing) ендірілген платформаларының ресурстық шектеулеріне бейімделген ықшам модельдерді қажет етеді.

Осылайша, клиникалық белгілер құрылымын ескеретін иерархиялық трансформер архитектурасын, теңгерімсіз және шағын деректер жиындарында сенімді оқытуды, сондай-ақ шешімдердің ашықтығын қамтамасыз ете отырып, модельді тиімді сығу және периферияда енгізу стратегиясын біріктіретін интеллектуалды жүйені әзірлеу теориялық та, қолданбалы да маңызы бар өзекті ғылыми міндет болып табылады.

Диссертациялық жұмыстың мақсаты жүректің ишемиялық ауруын анықтаудың дәлдігі мен тиімділігін арттыру үшін гибридті трансформерлік архитектураға және Edge-құрылғыларға арналған модельге негізделген интеллектуалды жүйені әзірлеу болып табылады.

Осы мақсатқа жету үшін жұмыста мынадай міндеттер шешіледі:

1) ЖИА-ны анықтау дәлдігін арттыру үшін теңгерімсіздікті жоюға арналған ADASYN алгоритмін және бейімделгіш өткірлікті азайту әдісімен (ASAM) оқытуды оңтайландыруды қамтитын гибридті трансформерлік архитектураны (HG-Former) әзірлеу;

2) Білімді дистилляциялауды және SHAP-ты қамтитын, модельді Edge-құрылғыларға көшіруге және тиімді енгізуге арналған модельді әзірлеу.

Зерттеу әдістері: кестелік деректерді жіктеуге арналған машиналық оқыту және терең оқыту әдістері, периферияда енгізу әдістері.

Зерттеу объектісі машиналық оқыту, түсіндірілетін жасанды интеллект және Edge-есептеулер әдістерін пайдалана отырып, жүректің ишемиялық ауруын анықтау үшін клиникалық медициналық деректерді интеллектуалды талдау процесі болып табылады.

Зерттеу пәні ЖИА-ны анықтауға арналған машиналық оқыту модельдері, әдістері мен алгоритмдері, соның ішінде HG-Former гибридті трансформерлік архитектурасы (ASAM оңтайландырғышымен) және Edge-құрылғыларда енгізуге арналған модельдерді сығу алгоритмдері болып табылады.

Алынған нәтижелердің ғылыми жаңалығы мынадан тұрады:

− иерархиялық трансформерлік архитектураны және Adaptive Sharpness-Aware Minimization әдісін біріктіретін жаңа гибридті архитектураға негізделген жүректің ишемиялық ауруын анықтау алгоритмі ұсынылды, бұл деректердің шектеулі жиыны жағдайында модельді тиімді оқытуды қамтамасыз етті;

− білімді дистилляциялауды және SHAP бағдарланған белгілерді іріктеуді бірлесіп қолдануға негізделген, Edge-құрылғыларға арналған жүректің ишемиялық ауруын анықтау моделі ұсынылды. Бұл жіктеудің жоғары дәлдігін сақтай отырып, пайдаланылатын белгілер санын едәуір қысқартуға мүмкіндік берді.

Қорғауға шығарылатын тұжырымдар:

− Z-Alizadeh Sani деректер жиынында оқытылған, HG-Former гибридті трансформерлік архитектурасына және Adaptive Sharpness-Aware Minimization әдісіне негізделген жүректің ишемиялық ауруын анықтау алгоритмі 96,72% дәлдікпен ЖИА-ны жіктеудің жоғары тиімділігін қамтамасыз етеді, бұл қолданыстағы баламалы әдістерден орта есеппен 10%-ға артық;

− білімді дистилляциялау және SHAP бағдарланған белгілерді іріктеу әдістерін қолдана отырып әзірленген, Edge-құрылғыларда жүректің ишемиялық ауруын анықтауға арналған модель 91,80% жіктеу дәлдігіне жетеді, бұл базалық ықшам модельдермен салыстырғанда орта есеппен 5-8%-ға жоғары. Бұл ретте пайдаланылатын белгілер саны 56-дан 16-ға дейін қысқартылды.

Зерттеудің теориялық және практикалық маңызы. Жоғары дәлдікті HG-Former гибридті трансформерлік архитектурасын, шағын және теңгерімсіз медициналық деректерде орнықты оқытуды және модельді Edge-құрылғыларда жергілікті енгізу мүмкіндігін біріктіретін жүректің ишемиялық ауруын анықтауға арналған интеллектуалды жүйе әзірленді. Ұсынылған тәсіл ЖИА-ны жіктеудің дәлдігі мен сенімділігін арттырады, шамадан тыс үйрену қаупін азайтады, SHAP-талдау мен attention-механизмдері есебінен шешім қабылдаудың ашықтығын қамтамасыз етеді, сондай-ақ медициналық деректерді бұлтқа жібермей өңдеуге мүмкіндік береді, бұл интеллектуалды медициналық жүйелердің құпиялылығын, қолжетімділігін және практикалық қолданылуын арттырады. Әзірленген жүйенің практикалық қолданылуы оның «Кардиология және ішкі аурулар ғылыми-зерттеу институты» АҚ-дағы ZhurekAI интеллектуалды диагностикалық платформасына енгізілуімен расталады.

Ғылыми нәтижелердің сенімділігі терең оқытудың озық әдістерін, соның ішінде трансформерлердің гибридті архитектураларын (HG-Former) және перифериялық есептеулерді (Edge ML) қолданумен қамтамасыз етіледі. Нәтижелер нақты клиникалық деректердегі есептеу эксперименттерімен, MIMIC-IV деректер қорынан алынған тәуелсіз когортта жүргізілген сыртқы валидациямен, білімді дистилляциялау және SHAP талдау әдістерін сәтті қолданумен, сондай-ақ ұсынылған модельдерді жүректің ишемиялық ауруын жіктеудің қолданыстағы әдістерімен сәтті салыстырумен расталады.

Тақырыптың ғылыми-зерттеу бағдарламаларының жоспарларымен байланысы. Диссертациялық жұмыс Қазақстан Республикасы Ғылым және жоғары білім министрлігінің ғылыми-зерттеу жобалары шеңберінде орындалды (қаржыландыру көзі: Қазақстан Республикасы Ғылым және жоғары білім министрлігінің Ғылым комитеті):

− AP26103523 «Блокчейн және IoMT (Internet of Medical Things) технологиясын пайдалана отырып, ұйқы бұзылыстарын мониторингілеу және диагностикалау интеллектуалды жүйесін әзірлеу» гранттық қаржыландыруы (ГҚ), 2025-2027 жж.

Диссертациялық жұмыстың құрылымы мен мазмұны. Диссертация кіріспеден, төрт негізгі тараудан, қорытындыдан және пайдаланылған дереккөздер тізімінен тұрады.

Кіріспеде тақырыптың өзектілігі негізделген, зерттеудің мақсаты мен міндеттері нақты айқындалған, зерттеу объектісі мен пәні анықталған, жұмыстың ғылыми жаңалығы мен практикалық маңызы ашылған, сондай-ақ қорғауға шығарылатын негізгі тұжырымдар баяндалған.

Бірінші тарауда ЖИА диагностикасына арналған машиналық оқытудың классикалық әдістерін, кардиологиядағы жасанды интеллекттің заманауи әдістерін, трансформер архитектураларының математикалық негіздерін, модельдерді оңтайландыру және сығу алгоритмдерін, сондай-ақ қолданыстағы гибридті және аралас тәсілдерге сыни талдауды қамтитын әдебиетке шолу ұсынылған.

Екінші тарауда ұсынылған жүйенің математикалық модельдері баяндалған, деректер жиындары мен оларды алдын ала өңдеу әдістері, HG-Former архитектурасының негізгі техникалық компоненттері, оқыту және оңтайландыру рәсімдері сипатталған, сондай-ақ архитектуралық компоненттерді дәйекті түрде алып тастау (ablation) жолымен жүргізілген алдын ала салыстырмалы эксперименттердің нәтижелері берілген.

Үшінші тарауда HG-Former гибридті архитектурасы сипатталған, оның ішінде клиникалық белгілерді топтау және токенизация механизмі, трансформердің топішілік және топаралық энкодерлері, ASAM алгоритмін интеграциялау, Z-Alizadeh Sani деректер жиынындағы эксперименттердің нәтижелері, SHAP талдауы және назар аудару механизмін түсіндіру. ASAM оңтайландырғышымен бірлесіп оқытылған HG-Former толық конфигурациясы бөлек қалдырылған (hold-out) тест таңдамасында AUC 0,9755, дәлдік 0,9672, F1-өлшемі 0,9762, Macro Recall 0,9767 және PR-AUC 0,9919 көрсеткіштеріне қол жеткізіп, CatBoost, XGBoost, RandomForest және FT-Transformer сияқты машиналық оқытудың он классикалық базалық моделінен, сондай-ақ swarm (үйірлік) алгоритмдерімен оңтайландырылған бес стекинг-ансамбльден айтарлықтай асып түсті. MIMIC-IV деректер қорынан алынған 800 пациенттен тұратын тәуелсіз когортта жүргізілген сыртқы валидация AUC 0,9620 және дәлдік 0,9038 көрсетті.

Төртінші тарауда перифериялық құрылғыларда енгізудің оңтайландырылған алгоритмі ұсынылған, ол білімді ProtoNN ықшам студент-моделіне дистилляциялауды, SHAP бағдарланған белгілерді іріктеуді, кванттауды және прунингті, Raspberry Pi Zero 2 W құрылғысында TFLite енгізуді қамтиды. Білімді дистилляциялау мен SHAP негізінде белгілерді іріктеуді бірлесіп қолдану кіріс белгілері кеңістігін 56-дан клиникалық тұрғыдан ақпаратты 16 айнымалыға дейін қысқартуға, сонымен бірге модельді TFLite форматында 4,35 КБ өлшемге дейін сығуға мүмкіндік берді. Студент-модель 91,80% жіктеу дәлдігіне және 0,9651 AUC-ке қол жеткізіп, бір болжамға шығару

кідірісін 0,0740 мс-қа дейін төмендетті — бұл бастапқы мұғалім-моделден шамамен 400 есе жылдам.

Қорытындыда диссертацияның негізгі нәтижелері мен тұжырымдары баяндалған.

Жарияланымдар. Диссертациялық зерттеудің нәтижелері 6 жұмыста жарияланған. Олардың ішінде Web of Science және Scopus дерекқорларына енгізілген халықаралық ғылыми журналдарда 5 мақала, ҚР ҒЖБМ Ғылым және жоғары білім саласында сапаны қамтамасыз ету комитеті ұсынған журналдарда 1 мақала, сондай-ақ диссертациялық зерттеуді практикалық іске асыру нәтижесі бойынша 2 авторлық куәлік және 1 пайдалы модельге патент бар.

Ғылыми жарияланымдар:

1. Tasmurzayev, N.; Amangeldy, B.; Imankulov, T.; Imanbek, B.; Postolache, O.A.; Konysbekova, A. A Wearable IoT-Based Measurement System for Real-Time Cardiovascular Risk Prediction Using Heart Rate Variability. Eng 2025, 6, 259. https://doi.org/10.3390/eng6100259 (IF: 0.4, Scopus Percentile: 58%, Q2).

2. Tasmurzayev, N.; Amangeldy, B.; Imanbek, B.; Baigarayeva, Z.; Imankulov, T.; Dikhanbayeva, G.; Amangeldi, I.; Sharipova, S. Digital Cardiovascular Twins, AI Agents, and Sensor Data: A Narrative Review from System Architecture to Proactive Heart Health. Sensors 2025, 25, 5272. https://doi.org/10.3390/s25175272 (IF: 3.5, Scopus Percentile: 88%, Q1).

3. Tasmurzayev, N.; Imanbek, B.; Boltaboyeva, A.; Dikhanbayeva, G.; Zhussupbekov, S.; Saparbayeva, Q.; Amirkhanova, G. Explainable AI for Coronary Artery Disease Stratification Using Routine Clinical Data. Algorithms 2025, 18, 693. https://doi.org/10.3390/a18110693 (IF: 2.3, Scopus Percentile: 78%, Q1).

4. Tasmurzayev, N.; Baigarayeva, Z.; Amangeldy, B.; Imanbek, B.; Kurmanbek, S.; Dikhanbayeva, G.; Amirkhanova, G. Interpretable Machine Learning for Coronary Artery Disease Risk Stratification: A SHAP-Based Analysis. Algorithms 2025, 18, 697. https://doi.org/10.3390/a18110697 (IF: 2.3, Scopus Percentile: 78%, Q1).

5. Tasmurzayev, N.; Amangeldy, B.; Baigarayeva, Z.; Boltaboyeva, A.; Imanbek, B.; Maeda-Nishino, N.; Zhussupbekov, S.; Baidauletova, A. Enhancing Cardiovascular Disease Classification with Routine Blood Tests Using an Explainable AI Approach. Algorithms 2025, 18, 708. https://doi.org/10.3390/a18110708 (IF: 2.3, Scopus Percentile: 78%, Q1).

6. Tasmurzayev, N.; Turmakhanbet, D.; Kakharov, A.; Aitkazin, M.; Baidauletova, A.; Kozhamberdiyeva, M. CVD Prediction from HRV Derived from Wearable PPG. JPCSIT, vol. 3, no. 4, Dec. 2025. https://doi.org/10.26577/jpcsit202542.

7. Авторлық куәлік: Турмаханбет Д. Н., Тасмурзаев Н. М., Мауленбеков М. А., Ибашова А. Б., Сапарбаева К. Ж., Иманбек Б. Т., Болтабоева А. К., Айтказин М. К., Искакова А. М., Кудьярова Ж. Б., Амангельды Б. С., Байгараева Ж. Е., Жолдыбаева Ж. К., Байдаулетова А. И. Жүрек ырғағының вариабельділігін және PPG пульстік толқынының амплитудасын талдау негізінде ментальдық стресс

деңгейін анықтау әдісі. Авторлық құқық объектісіне құқықтарды мемлекеттік тіркеу туралы куәлік № 62846, 2025 жылғы 9 қазан.

8. Авторлық куәлік: Турмаханбет Д. Н., Тасмурзаев Н. М., Мауленбеков М. А., Ибашова А. Б., Сапарбаева К. Ж., Иманбек Б. Т., Болтабоева А. К., Айтказин М. К., Искакова А. М., Кудьярова Ж. Б., Амангельды Б. С., Байгараева Ж. Е., Жолдыбаева Ж. К., Байдаулетова А. И. Фотоплетизмограмма талдауына негізделген тамыр жасы мен артериялардың қатаңдығын бағалауға арналған бағдарламалық кешен. Авторлық құқық объектісіне құқықтарды мемлекеттік тіркеу туралы куәлік № 63527, 2025 жылғы 29 қазан.

9. Пайдалы модельге патент: Иманбек Б. Т., Болтабоева А. К., Амангельды Б. С., Тасмурзаев Н. М., Нурланов Ю. Ю., Копенов А. Б., Ибрашева З. С., Омирбекова Ж. Ж., Мансурова М. Я., Байгараева Ж. Е., Амирханова Г. А. Жүрек-қан тамырлары жүйесінің параметрлерін мониторингілеу жүйесі. Пайдалы модельге патент № 10282, 2025 жылғы 7 қараша.