

АНДАТПА

Жұмыстың жалпы сипаттамасы

Диссертация медициналық зертханаларды кешенді автоматтандыру арқылы зертханалық зерттеулердің нәтижелерін интерпретациялау үшін жасанды интеллект моделін жасауға арналған. Диссертация машиналық оқыту әдістерін, реляциялық алгебраны және OpenAI қызметтерін пайдалануды қарастырады.

Зерттеудің өзектілігі

Диссертациялық жұмыстың тақырыбы 2018-2022 жылдар аралығында жүзеге асырылған «Цифрлық Қазақстан» мемлекеттік цифрландыру бағдарламасымен тығыз байланысты, ол алдыңғы қатарлы технологияларды пайдалану арқылы ел халқының өмір сүру деңгейін арттыруға бағытталған стратегиялық кешенді бағдарлама болып табылады. сандық технологиялар. Бағдарлама аясында клиникалық-диагностикалық зертханаларды автоматтандыру маңызды рөл атқаратын Қазақстан Республикасы халқының электрондық денсаулық паспортын кең ауқымда енгізуге ерекше көңіл бөлінеді. Негізінен зертханалар гематологиялық, биохимиялық, иммунологиялық және т.б. автоматты және жартылай автоматты анализаторлармен жабдықталған. Автоматты анализаторлар сағатына 2000-нан астам талдау жүргізе алатынын атап өткен жөн, дегенмен негізгі көрсеткіштер жабдықтың үлгісі мен сипаттамаларына байланысты. Іс жүзінде барлық медициналық ұйымдарда автоматтандырылған жүйелер мен орталықтандырылған деректерді сақтау қоймалары жоқ, сондықтан әрбір зертханалық сынақ нәтижесі белгілі бір жабдықтың жергілікті деректер базасында қалады. Осылайша, сынақ нәтижелері стандартты анализатор пішіндерінде (көбінесе ағылшын тілінде) беріледі немесе қағазға жіберіледі, бұл адам қателеріне, көрсеткіштердегі сәйкессіздіктерге және деректердің жоғалуына ықпал етуі мүмкін. «Цифрлық Қазақстан» мемлекеттік бағдарламасында халықтың электронды денсаулық паспортына көңіл бөлініп, зертханалық ақпараттық жүйені (ЛАЖ) енгізе отырып, медициналық зертханалардағы күнделікті процестерді кешенді автоматтандыру қажеттілігі алға қойылған.

Пациентке бағдарлану цифрландыруды енгізу қағидасының негізгі негізі болып табылады, ол электрондық денсаулық паспортын қалыптастыруды жүзеге асыруда көрініс табады. mGov мобильді қосымшасында үзіндіні, төлқұжат деректерін, қай дәрігерге тағайындалғаныңызды және жақында қандай медициналық қызметтер алғаныңызды көре аласыз. Бұл ақпараттың барлығы ақпараттық қауіпсіздік заңнамасына сәйкес жеткілікті түрде қорғалған.

Электрондық денсаулық паспортында еліміздің медициналық мекемелерінде науқасты тексеру мен емдеудің барлық нәтижелері туралы мәліметтер, бұрын ауырған созылмалы аурулары туралы мәліметтер, науқаста аллергия тудыратын дәрілердің тізімі, пациент барған мамандардың барлық ұсыныстары, және т.б.

2020 жылдың 4 наурызында Қазақстан Республикасының Президенті Қасым-Жомарт Тоқаев «Цифрлық Қазақстан» мемлекеттік бағдарламасын жүзеге асыру барысы туралы есеп беру кездесуін өткізіп, цифрландыру бір саланы емес, бүкіл экономиканы дамытуға бағытталғанын атап өтті. мемлекет және тұтастай қоғамды өзгерту. Мемлекет басшысы осы уақытқа дейін технологияларды енгізу бойынша мемлекеттік органдардың алдына қойылған міндетке қол жеткізілмегенін, бағдарламаны сәтті жүзеге асыру үшін органдардың бірлігі мен үйлестіруі қажет екенін атап өтті, сондай-ақ Big Data технологияларын, жасанды интеллектті пайдаланудың маңыздылығын атап өтті. алгоритмдер сенімділіктің ең жоғары деңгейімен жүйелік деректер қорын қалыптастыруға ықпал етеді, қате немесе ескірген ақпаратты алып тастайды, деректерді жүйелейді және объективті тұтас суретті береді. Мемлекет басшысы Үкіметке Индустрия 4.0 жобасын кеңейту бойынша реттеуші және ынталандырушы шаралар кешенін әзірлеуді тапсырды.

Медициналық зертханаларды цифрландыру электрондық денсаулық сақтау саласының бір бөлігі болып табылатын перспективалы бағыт болып табылады. Медициналық зертханалар зертханалық зерттеулердің нәтижелері бойынша ақпараттың үлкен көлемін жинақтайды, бірақ ақпараттың әртүрлілігі сонша, ол «патчворк цифрландыру» элементтерін қамтиды.

McKinsey сарапшыларының пікірінше, бес жыл бұрын, тіпті медициналық технологиялардың серпінді дамуы, денсаулық сақтау саласы «Үлкен деректерді» қолданбалы қолдану бойынша басқа салалардан айтарлықтай артта қалды, бұл дәрігерлердің ұсынылатын қызметтерге сенімсіздігімен байланысты. ақпараттық технология құралдарын пайдалана отырып жасалған қорытындылар. Деректер ғылымындағы инновациялар (ағылшын тілінен аудармасы – data Science) емделушілерге көмектесіп қана қоймай, дәрігерлерді күнделікті жұмысының бір бөлігін босатады. Артықшылықтарды клиникалық тәжірибедегі кейбір процестерді оңтайландыратын чат боттары көрсетеді.

Әлемдік нарықта өз елінің зертханасында операциялық және өңдеу процестерін автоматтандыруға қызмет ететін көптеген зертханалық ақпараттық жүйелер бар. Әрбір зертханалық ақпараттық жүйе өз мемлекетінің заңнамалық нормаларына бейімделген және көбінесе басқа елдерге масштабтау өте күрделі процесс болып көрінеді, сондықтан жүйені бейімдеу бір жылдан астам уақытты алуы мүмкін.

Қазақстан Республикасының нарығына зертханалық ақпараттық жүйелердің отандық әзірлеушілері де, «K-Lab» LIS («Медитек» ЖШС, Ресей-Қазақстан), «Ариадна» (ООО «Брегис») сияқты көршілес елдердің шешімдері де кіреді. Ресей), «Сирока» ТЖК («SIROCA TECHNOLOGY» ЖШС, Қырғызстан), «Info Lab» ТЖК (ТОО «Inform Medical», Қазақстан), «CS-Soft» ТЖК («Mobile Service Group» ЖШС, Қырғызстан), LIS «TerraLab» (Terralab SOFT LLC, Украина), «Даму» ТЖК (КИИС) («DAMU» ақпараттық технологиялар орталығы» ЖШС, Қазақстан), «SmartLab» ЖТҚ (SmartLab Kazakhstan ЖШС, Қазақстан). Қазақстандық әзірлеушілердің артықшылығы - тиімділік пен икемділік, ал әзірлеушілер елде орналасқан, бұл оларға

медициналық зертханалардың қажеттіліктерін қанағаттандыру және жергілікті заңнамаға бейімделу үшін әзірлемелерді жылдам жүргізуге мүмкіндік береді.

Жұмыс аясында Қазақстанда зертханалық зерттеулердің бекітілген нормалары мен стандарттары жоқ екені анықталды. Қолданыстағы бұйрықтар мен стандарттар зертханалық зерттеулердің нәтижелеріне емес, зертхананың жұмысына бағытталған. Әдетте, көптеген зертханалар жабдықпен бірге жеткізілетін реагенттерге арналған нұсқаулықта көрсетілген анықтамалық мәндерді пайдаланады және бұл олардың жұмысында қателікке әкеледі. Әрбір реагент үшін реагентке арналған нұсқаулықта көрсетілген стандарттарға қарамастан, зертхана аймаққа және популяцияға байланысты оның стандарттарын анықтауы керек деген көрсеткіш болғандықтан.

Ақпараттық жүйелерді дамытудың және жасанды интеллект технологияларын заманауи медициналық тәжірибеге енгізудің маңыздылығын асыра бағалау мүмкін емес. Бұл технологиялар диагностикалық және емдеу процесінің тиімділігін арттыруда, сайып келгенде, науқастарды уақтылы және дәл емдеуді қамтамасыз етуде шешуші рөл атқарады. Медициналық шешімдерді қабылдауға маңызды қолдау көрсете отырып, AI ақпараттық жүйелері мен алгоритмдері медициналық мамандар үшін деректерді жылдам және дәл түсіндіруге, қателер қаупін азайтуға және пациенттерді басқару процестерін оңтайландыруға мүмкіндік беретін құнды құралға айналуға мүмкіндік береді. Бұл тәсіл диагностика процесін жылдамдатып қана қоймай, дәлірек және уақтылы емдеуді жеңілдетеді, сол арқылы айтарлықтай әлеуметтік және экономикалық пайда әкеледі.

Зерттеу мақсаты - Патологияны анықтауды автоматтандыру үшін денсаулық сақтау саласындағы зертханалық зерттеулердің нәтижелерін интерпретациялауға бағытталған жасанды интеллект моделін әзірлеу.

Осы мақсатқа жету үшін келесі **міндеттер** қойылды және шешілді:

- медициналық зертханалық диагностиканың әлемдік тәжірибесіндегі аналогтарын зерттеу;
- биохимиялық, иммунологиялық және гематологиялық мәліметтерді одан әрі талдау үшін зертханалық сынақ нәтижелерінің BIG DATA генерациясы;
- Қазақстан Республикасының аймақтық сипаттамалары мен зертханалық жабдықтарды өндірушілердің зертханалық сынақтар үшін эталондық мәндердің бірыңғай деректер базасын қалыптастыру;
- үлкен деректер (BIG DATA) негізінде зертханалық зерттеу деректерін интерпретациялау саласында кешенді автоматтандыру әдістемесін қалыптастыру;
- кешенді автоматтандыру негізінде зертханалық зерттеу нәтижелерінде патологияны анықтау үшін жасанды интеллект моделін жасау.
- «SmartLab» зертханалық ақпараттық жүйесі шеңберінде зертханалық сынақ нәтижелерін интерпретациялау модульдерін әзірлеу.

Зерттеу нысаны: денсаулық сақтаудағы зертханалық диагностика, тиісті медициналық құжаттары бар медициналық зертханалық жабдықтар, зертханадағы жасанды интеллект.

Зерттеу пәні: жасанды интеллект алгоритмдері және зертханалық зерттеулер нәтижелерінде патологияны анықтау алгоритмдері.

Диссертациялық талаптар шеңберіндегі ғылыми нәтижелер

Кіріспеде тақырыптың өзектілігі, зерттелу деңгейі мен ғылыми өңделуі негізделеді. Диссертациялық зерттеудің объектісі, пәні, мақсаты, міндеттері және әдістемесі көрсетілген. Ғылыми жаңалығы, қорғауға ұсынылған ғылыми ұстанымдары және зерттеу нәтижелерінің практикалық маңыздылығы көрсетілген. Зерттеу нәтижелерін сынақтан өткізу және жариялау туралы деректер келтіріліп, автордың ғылыми зерттеулерге қосқан жеке үлесі де көрсетіледі.

Бірінші бөлімде денсаулық сақтаудағы зертханалық диагностиканың қазіргі жағдайын талдауды, денсаулық сақтаудағы жасанды интеллектті дамытудың талдауын және перспективаларын талдауды қоса алғанда, денсаулық сақтаудағы зертханалық диагностиканың қолданыстағы жасанды интеллект үлгілері мен жүйелеріне шолу және талдау, сондай-ақ шолу. зертханалық диагностикада жасанды интеллект моделін жасау бойынша жұмыс, сондай-ақ Қазақстан Республикасының нарығында ұсынылған зертханалық ақпараттық жүйелерге шолу жасау.

Екінші бөлімде зертханалық зерттеу нәтижелерін талдау процестерін модельдеу талқыланады. Жүйелік талдауды зертханалық диагностикада қолдануды, зертханалық операцияларды автоматтандыруды және патологияларды диагностикалау және болжау үшін зертханалық зерттеулерді оңтайлы қолдану әдістерін сипаттайды. Зертханалық нәтижелерді талдау үшін білім базасын құру және олардың белгіленген нормалардан ауытқуы негізінде стандартты мәндерді интерпретациялауды модельдеу де талқыланады.

Үшінші бөлімде денсаулық сақтаудағы зертханалық диагностика үшін жасанды интеллект моделін әзірлеу нәтижелері, зертханалық зерттеулер нәтижелерінің BIG DATA көмегімен деректерді талдау үшін машиналық оқыту әдістерін пайдалану, зертханалық диагностика үшін жасанды интеллект моделін әзірлеу нәтижелері, логикалық модель нәтижелерді интерпретациялауға арналған жасанды интеллект, сонымен қатар нәтижелерді интерпретациялау үшін жасанды интеллекттің математикалық моделін негіздейді.

Төртінші бөлім практикалық енгізуге және ақпараттық жүйенің архитектурасын енгізу нәтижелеріне арналған, ЛИС енгізу технологиялары, деректер қорының архитектурасы, ақпараттық жүйенің модульдері және олардың өзара әрекеттесуі, сондай-ақ Қазақстан Республикасының аумағындағы зертханаларда енгізу нәтижелері берілген.

Ғылыми жаңалығы:

1. Патологияларды анықтауды автоматтандыру үшін денсаулық сақтау саласындағы зертханалық зерттеулердің нәтижелерін интерпретациялау үшін жасанды интеллект моделі әзірленді.

2. Зертханалық жабдықты өндірушілер және Қазақстан Республикасының өңірлік сипаттамалары үшін зертханалық зерттеулердің эталондық мәндерінің бірыңғай деректер базасы қалыптастырылды, бұл

зертханалық зерттеулер нәтижелерінің сенімділігін арттыруға мүмкіндік береді.

3. Зертханалық жабдықтармен біріктіру үшін әмбебап платформа әзірленді.

4. Нейрондық желіні жинақтауда және градиентті күшейтуде машиналық оқыту әдістерін қолдану алгоритмі әзірленді.

Қорғауға ұсынылады:

1. BIG DATA (биохимиялық, иммунологиялық және гематологиялық сынақтар) көмегімен зертханалық зерттеулер нәтижелерін интерпретациялау үшін кешенді автоматтандыру әдістемесі әзірленді.

2. Зертханалық жабдықтарды өндірушілердің зертханалық зерттеулерге арналған эталондық мәндердің бірыңғай дерекқорын қалыптастыру әдістемесі және Қазақстан Республикасының өңірлік сипаттамалары.

3. Машиналық оқыту және жасанды интеллект әдістерін қолдану арқылы автоматтандыруға негізделген зертханалық зерттеулер нәтижелерінде патологияны анықтау үшін математикалық модель әзірленді.

4. «SmartLab» зертханалық ақпараттық жүйесінде медициналық зертханалық анализаторлармен біріктірілген жасанды интеллект үлгілерінің жиынтығы негізінде зертханалық зерттеулер нәтижелерінде патологияны автоматты түрде анықтауға арналған модульдер әзірленді.

Практикалық құндылығы және зерттеу нәтижелерін іске асыру

1. Жүргізілген зерттеулер нәтижесінде зертханалық жабдықтарды өндірушілер мен Қазақстан Республикасының аймақтық сипаттамалары үшін зертханалық зерттеулердің эталондық мәндерінің бірыңғай деректер базасы қалыптастырылды, ол денсаулық сақтау саласындағы мамандардың сараптамалық бағасын алды және енгізілді.

2. Медициналық зертханалық анализаторлармен біріктірілген жасанды интеллект үлгілерінің жиынтығы негізінде зертханалық зерттеулер нәтижелерінде патологияны автоматты түрде анықтау үшін алгоритмдер мен бағдарламалық модульдер әзірленді.

3. Лабораториялық анализаторлармен интеграциялау үшін платформа әзірленді, оны басқа салаларда (қаржы секторы, IoT технологиялары, қол жеткізуді басқару жүйелері) қолдануға болады.

4. Жасанды интеллектті басқа жасанды интеллектпен оқытуға мүмкіндік беретін OpenAI қолдану үшін қызметтер әзірленді.

Зерттеудің әдістемелік негізі және метрологиялық қамтамасыз етілуі

Диссертациялық жұмыста орындаудың әрбір кезеңінде келесі әдістердің қатары қолданылды: Жүйелік талдау әдісі – медициналық зертхананың операциялық қызметінің техникалық, зертханалық мәліметтерін, іске асыру алгоритмдерін қалыптастыру; Декомпозиция әдісі – бағдарламалық кодты, модульдерді және ақпараттық жүйенің барлық функционалдығын әзірлеу; Машиналық оқыту әдістері – деректердің үлкен массивін зерттеу кезінде

үлгілерді анықтау; Реляциялық алгебра әдістері - зертханалық зерттеулер нәтижелерінде патологияны анықтауға арналған функционалдылықты қоса алғанда, деректер бойынша барлық операциялар. Өртүрлі үлгідегі заманауи зертханалық анализаторларда метрологиялық (өлшеу) жұмыстары жүргізілді, олар күнделікті орындалатын зерттеулердің сапасын бақылаудан, сондай-ақ стандарттау және метрология органдарымен жыл сайынғы тексеруден өтеді.

Алынған нәтижелердің сенімділігі мен негізділігі, қорытындылар және диссертацияда баяндалған ұсыныстар клиникалық диагностикалық зертханаларда енгізу актілерімен, сондай-ақ машиналық оқыту үлгілерін болжау дұрыстығын бағалау көрсеткіштерімен расталады.

Өтініш берушінің басылымдары және жеке үлесі. 13 ғылыми жұмыс жарияланды, оның ішінде 3 жұмыс Қазақстан Республикасы Білім және ғылым министрлігі Білім және ғылым саласындағы бақылау комитеті ұсынған басылымдарда, 1 мақала халықаралық Scopus деректер базасына енгізілген ғылыми жарияланым, басқа басылымдарда 4 мақала, халықаралық ғылыми-практикалық конференцияларда 5 мақала. Компьютерлік бағдарламаны мемлекеттік тіркеу туралы 2 куәлік, «SmartLab» тауар белгісінің 1 сертификаты алынды.

Диссертациялық жұмыстың құрылымы мен көлемі. Ұсынылған диссертациялық жұмыс кіріспеден, төрт бөлімнен, қорытындыдан, 101 атаудан пайдаланылған дереккөздер тізімінен, компьютерлік мәтіннің 131 бетінде берілген, 63 суреттен, 18 кестеден және 6 қосымшадан тұрады.

АННОТАЦИЯ

на диссертацию Кадиркулова Куаныша Кайсаровича на тему: «Разработка модели искусственного интеллекта по лабораторной диагностике в здравоохранении» на соискание степени доктора философии PhD по специальности 8D06101 - Аналитика больших данных.

Общая характеристика работы

Диссертационная работа посвящена разработке модели искусственного интеллекта по интерпретации результатов лабораторных исследований за счет комплексной автоматизации медицинских лабораторий. В диссертации рассматриваются методы машинного обучения, реляционной алгебры, а также использование сервисов OpenAI.

Актуальность исследования

Тема диссертационной работы имеет тесную связь с Государственной программой по цифровизации «Цифровой Казахстан» реализованной в период с 2018 по 2022 гг., которая является стратегической комплексной программой, нацеленной на повышение уровня жизни населения страны за счет использования передовых цифровых технологий. В рамках программы уделяется особое внимание на масштабную реализацию электронного паспорта здоровья населения Республики Казахстан, где автоматизация клиническо – диагностических лабораторий занимает немаловажную роль. В основном лаборатории оснащаются автоматическими и полуавтоматическими анализаторами, такими как гематологические, биохимические, иммунологические и другие. Имеет место отметить, что автоматические анализаторы могут проводить более 2000 анализов в час, хотя основные показатели зависят от модели и характеристик оборудования. На практике, не все медицинские организации имеют автоматизированные системы и централизованные хранилища данных, следовательно, каждый результат лабораторных тестов остается на локальных базах данных определенного оборудования. Таким образом, результаты анализов предоставляются на стандартных бланках анализатора (зачастую на английском языке) либо переносятся на бумажный носитель, что может способствовать ошибкам, обусловленным человеческим фактором, несоответствиям показателей и утрате данных. В государственной программе «Цифровой Казахстан» уделяют внимание на электронный паспорт здоровья населения, выдвигается необходимость комплексной автоматизации рутинных процессов медицинских лабораторий с внедрением лабораторной информационной системы (ЛИС).

Пациента-ориентированность является главной основой принципа внедрения цифровизации, что отражена в реализации формирования электронного паспорта здоровья. В мобильном приложении mGov можно посмотреть выписку, паспортные данные, к какому врачу прикреплены, какие медицинские услуги в последнее время получали. Вся эта информация достаточно защищена в соответствии с законодательством об информационной безопасности.

Электронный паспорт здоровья содержит данные обо всех результатах обследования и лечения пациента в медицинских учреждениях страны, информацию о перенесенных хронических заболеваниях, список лекарств, вызывающих аллергию у пациента, все рекомендации специалистов которых посещал пациент и другое.

Президент Республики Казахстан Касым-Жомарт Токаев 4 марта 2020 года провел отчетное заседание по реализации Государственной программы «Цифровой Казахстан», и отметил, что цифровизация направлена не для развития одного сектора, а всей экономики государства и преобразовании общества в целом. Президентом подчеркнуто, что пока поставленная перед государственными органами задача по внедрению технологий не достигнута, что для успешной реализации программы необходима согласованность и скоординированность органов, а также было отмечено важность применения технологий «Big Data», алгоритмов «искусственного интеллекта», что способствует к формированию систем базы данных с наивысшим уровнем достоверности, отбрасывать некорректную или устаревшую информацию, систематизировать данные и выдавать объективную целую картину. Глава государства поручил Правительству разработать комплекс регуляторных и стимулирующих мер по расширению проекта Индустрия 4.0.

Цифровизация медицинских лабораторий является перспективным направлением, являющимся частью электронного здравоохранения. Медицинские лаборатории аккумулируют огромные объемы информации по результатам лабораторных исследований, однако информация настолько разнообразная что несет в себе элементы «лоскутной цифровизации».

Согласно утверждениям аналитиков McKinsey, пять лет назад, даже динамичное развитие медицинских технологий, сфера здравоохранения с позиции прикладного применения «Big Data» существенно отставала от других отраслей, это связано с отсутствием доверия врачей к предоставляемым выводам сделанных с использованием инструментов информационных технологий. Инновации в data science (перевод с англ. – наука о данных) могут не только помочь больным, но и избавить врачей от части рутинной работы. Плюсы демонстрируют чат боты, которые оптимизируют некоторые процессы в клинической практике.

На мировом рынке существует множество лабораторных информационных систем, которые служат для автоматизации операционных, лечебных процессов в лаборатории своей страны. Каждая лабораторная информационная система адаптирована под законодательные нормы своего государства, и зачастую масштабирование в другие страны представляются крайне сложным процессом так, так адаптация системы может занимать в сроки более одного года.

На рынке Республики Казахстан представлены как отечественные разработчики лабораторных информационных систем, так и решения ближнего зарубежье, такие как ЛИС «K-Lab» (ТОО «Meditec», Россия-Казахстан), ЛИС «Ариадна» (ООО «Брегис», Россия), ЛИС «Siroca» (ТОО «SIROCA TECHNOLOGY», Кыргызстан), ЛИС «Info Lab» (ТОО «Inform

Medical», Казахстан), ЛИС «CS-Soft» (ОсОО «Мобайл Сервис Групп», Кыргызстан), ЛИС «TerraLab» (ООО «Терралаб СОФТ», Украина), ЛИС «Даму» (КМИС) (ТОО «Центр Информационных Технологий «ДАМУ», Казахстан), ЛИС «SmartLab» (ТОО «SmartLab Kazakhstan», Казахстан). Преимуществом Казахских разработчиков является оперативность и гибкость, и разработчики находятся на территории страны, что позволяет быстро осуществлять разработки под потребности медицинских лабораторий и адаптироваться под местные законодательные акты.

В рамках работы было выявлено, отсутствие в Казахстане утвержденных норм и стандартов для лабораторных исследований. Существующие приказы и стандарты фокусируются на работе самой лаборатории, а не на результаты лабораторных исследований. Как правило, большинство лабораторий используют референсные значения, указанные в инструкциях для реактивов, поставляемых с оборудованием, и это приводит к ошибке в их работе. Так как для каждого реактива имеется указание, что независимо от стандартов, указанных в инструкциях к реактиву, лаборатория должна определить свои нормы, в зависимости от региона и популяции.

Важность разработки информационных систем и внедрения технологий искусственного интеллекта в современной медицинской практике трудно переоценить. Эти технологии играют ключевую роль в усилении эффективности процесса диагностики и лечения, обеспечивая, в итоге, своевременное и точное лечение пациентов. Обеспечивая необходимую поддержку при принятии медицинских решений, созданные информационные системы и алгоритмы ИИ становятся ценным инструментом для медицинских профессионалов, позволяя им оперативно и точно интерпретировать данные, минимизировать риск ошибок и оптимизировать процессы управления пациентами. Такой подход не только ускоряет процесс постановки диагноза, но и способствует более точному и своевременному назначению лечения, предоставляя, таким образом, значительные социальные и экономические выгоды.

Цель исследования - разработка и реализация модели искусственного интеллекта, способной автоматически интерпретировать результаты лабораторных исследований в области здравоохранения.

Для достижения поставленной цели были поставлены и решены следующие **задачи**:

- Изучение существующих методов и технологий в области лабораторной диагностики в здравоохранении: проанализировать современные методы и подходы к лабораторной диагностике, выявив их преимущества и недостатки.
- Формирование методики комплексной автоматизации в области интерпретации данных лабораторных исследований на основе больших данных (big data), которая позволит автоматически интерпретировать результаты лабораторных исследований для дальнейшего анализа биохимических, иммунологических и гематологических данных.

- Формирование единой базы референсных значений лабораторных исследований: создать базу данных, содержащую референсные значения лабораторных параметров, учитывая производителей лабораторного оборудования и региональные особенности Республики Казахстан.
- Разработка и обучение модели искусственного интеллекта для выявления патологии в результатах лабораторных исследований с последующей интеграцией разработанных модулей интерпретации результатов в лабораторную информационную систему «SmartLab».

Объект исследования: является лабораторная диагностика в здравоохранении, медицинское лабораторное оборудование с надлежащей медицинской документацией, искусственный интеллект в лаборатории.

Предмет исследования: модели и алгоритмы искусственного интеллекта по выявлению патологии в результатах лабораторных исследований.

Научные результаты в рамках требований к диссертации

Во введении обосновывается релевантность темы, уровень изученности и научной проработанности. Представлены объект, предмет, цель, задачи и методология диссертационного исследования. Выделяются научная новизна, научные позиции, представленные к защите, и практическая значимость результатов исследования. Приводятся данные об апробации и публикациях результатов исследования, а также указывается персональный вклад автора в научные исследования.

В первом разделе приведен обзор и анализ существующих модели и систем искусственного интеллекта по лабораторной диагностике в здравоохранении, в том числе анализ современного состояния лабораторной диагностики в здравоохранении, анализ и перспектива развития искусственного интеллекта в здравоохранении, приведен обзор работ по разработке модели искусственного интеллекта в лабораторной диагностике, а также сделан обзор лабораторным информационным системам представленные на рынке Республики Казахстан.

В разделе два рассматривается моделирование процессов анализа результатов лабораторных исследований. Описывается применение системного анализа к лабораторной диагностике, автоматизация лабораторных операций и методы оптимального применения лабораторных тестов для диагностики и прогноза патологий. Также рассматриваются создание базы знаний для анализа результатов лабораторных исследований и моделирование интерпретации стандартных значений на основе их отклонений от установленных норм.

В третьем разделе приводятся результаты разработки модели искусственного интеллекта по лабораторной диагностике в здравоохранении, применения методов машинного обучения для анализа данных по BIG DATA результатов лабораторных исследований, результаты разработки модели искусственного интеллекта по лабораторной диагностике, логическая модель искусственного интеллекта по интерпретации результатов, а также обоснована

математическая модель искусственного интеллекта по интерпретации результатов.

Четвертый раздел посвящен практической реализации и результатам внедрения информационной системы, приведены технологии реализации ЛИС, архитектура базы данных, модули информационной системы и их взаимодействие, а также результаты внедрения в лабораториях на территории Республики Казахстан.

Научная новизна:

1. Методика комплексной автоматизации в области интерпретации данных лабораторных исследований на основе больших данных (big data), позволяющая автоматически интерпретировать результаты лабораторных исследований для дальнейшего анализа биохимических, иммунологических и гематологических данных.
2. Методика формирования единой базы референсных значений лабораторных исследований с учетом производителей лабораторного оборудования и региональных особенностей Республики Казахстан.
3. Модель искусственного интеллекта для выявления патологии в результатах лабораторных исследований и интеграция разработанных модулей интерпретации результатов в лабораторную информационную систему «SmartLab».

Положения, выносимые на защиту:

1. Разработана методика комплексной автоматизации по интерпретации результатов лабораторных исследований по BIG DATA (биохимических, иммунологических и гематологических тестов).

2. Методика формирования единой базы референсных значений лабораторных исследований по производителям лабораторного оборудования и региональных особенностей Республики Казахстан.

3. Разработана математическая модель по выявлению патологии в результатах лабораторных исследований на основе автоматизации с применением методов машинного обучения.

4. Разработаны модули в лабораторной информационной системе «SmartLab» для автоматического выявления патологии в результатах лабораторных исследований на основе комплекса моделей искусственного интеллекта, интегрируемая с медицинскими лабораторными анализаторами.

Практическая ценность и реализация результатов исследований

1. В результате проведенных исследований сформирована Единая база референсных значений лабораторных исследований по производителям лабораторного оборудования и региональных особенностей Республики Казахстан, которая получила экспертную оценку специалистов в области здравоохранения и внедрена.

2. Разработаны алгоритмы и программные модули для автоматического выявления патологии в результатах лабораторных исследований на основе комплекса моделей искусственного интеллекта, интегрируемая с медицинскими лабораторными анализаторами.

3. Разработана платформа по интеграции с лабораторными анализаторами, которая может применяться в других сферах (финансовый сектор, IoT технологии, системы контроля доступа).

4. Разработаны сервисы по использованию OpenAI, которая позволяет обучать искусственный интеллект другим искусственным интеллектом.

Методологическая база и метрологическое обеспечение исследований

В диссертационной работе на каждом этапе реализации были использованы ряд следующих методов: Метод системного анализа – формирование технических, лабораторных данных, алгоритмов реализации по операционной деятельности медицинской лаборатории; Метод декомпозиции – разработка программного кода, модулей и всего функционала информационной системы; Методы машинного обучения – определение закономерностей при изучении большого массива данных; Методы реляционной алгебры – все операции над данными, включая функционал по выявлению патологии в результатах лабораторных исследований. Метрологические (измерительные) работы выполнялись на современных лабораторных анализаторах разных моделей, которые проходят ежедневный контроль качества на выполняемые исследования, а также ежегодную поверку органами стандартизации и метрологии.

Достоверность и обоснованность полученных результатов, выводов и рекомендаций, изложенных в диссертации, подтверждены актами внедрения в клинику – диагностические лаборатории, а также показателями оценки точности прогнозирования моделей машинного обучения.

Публикации и личный вклад соискателя. Опубликованы 13 научных трудов, в том числе 3 работы в изданиях, рекомендованных Комитетом по контролю в сфере образования и науки МВОН РК, 1 статья в международном научном издании, входящем в базу данных Scopus, 4 статьи в других изданиях, 5 статей в международных научно-практических конференциях. Получены 2 свидетельства о государственной регистрации программы для ЭВМ, 1 свидетельство о товарном знаке «SmartLab».

Структура и объем диссертационной работы. Представленная диссертационная работа состоит из введения, четырех разделов, заключения, списка использованных источников из 101 наименований, изложенных на 131 страницах компьютерного текста, включает 63 рисунков, 18 таблиц и 6 приложений.

8D06101 – Үлкен деректердің аналитикасы мамандығы бойынша философия докторы PhD дәрежесін алу үшін дайындалған **Кадиркулов Куаныш Кайсаровичтің «Денсаулық сақтаудағы зертханалық диагностика үшін жасанды интеллект моделін жасау»** диссертациясының тақырыбына жазылған

ABSTRACT

for the dissertation of Kuanysh Kadirkulov on the topic: “Development of an artificial intelligence model for laboratory diagnostics in healthcare” for the degree of Doctor of Philosophy PhD in specialty 8D06101 - Big Data Analytics.

General characteristics of the work

The dissertation is devoted to the development of an artificial intelligence model for interpreting the results of laboratory tests through complex automation of medical laboratories. The dissertation examines methods of machine learning, relational algebra, and the use of OpenAI services.

Relevance of the study

The topic of the dissertation work has a close connection with the State Digitalization Program “Digital Kazakhstan” implemented in the period from 2018 to 2022, which is a strategic comprehensive program aimed at improving the standard of living of the country’s population through the use of advanced digital technologies. Within the framework of the program, special attention is paid to the large-scale implementation of the electronic health passport of the population of the Republic of Kazakhstan, where the automation of clinical and diagnostic laboratories plays an important role. Basically, laboratories are equipped with automatic and semi-automatic analyzers, such as hematological, biochemical, immunological and others. It is worth noting that automatic analyzers can conduct more than 2000 analyzes per hour, although the main indicators depend on the model and characteristics of the equipment. In practice, not all medical organizations have automated systems and centralized data storage facilities; therefore, each laboratory test result remains on the local databases of certain equipment. Thus, test results are provided on standard analyzer forms (often in English) or transferred to paper, which can contribute to human errors, inconsistencies in indicators and data loss. The state program “Digital Kazakhstan” pays attention to the electronic health passport of the population and puts forward the need for comprehensive automation of routine processes in medical laboratories with the introduction of a laboratory information system (LIS).

Patient-centeredness is the main basis of the principle of introducing digitalization, which is reflected in the implementation of the formation of an electronic health passport. In the mGov mobile application, you can view your extract, passport data, which doctor you are assigned to, and what medical services you have recently received. All this information is sufficiently protected in accordance with information security legislation.

The electronic health passport contains data on all the results of examination and treatment of the patient in the country’s medical institutions, information about previous chronic diseases, a list of medications that cause allergies in the patient, all recommendations of specialists whose visits the patient visited, and more.

On March 4, 2020, President of the Republic of Kazakhstan Kassym-Jomart Tokayev held a reporting meeting on the implementation of the State Program “Digital Kazakhstan”, and noted that digitalization is not aimed at developing one sector, but the entire economy of the state and transforming society as a whole. The

President emphasized that so far the task set for government agencies to implement technologies has not been achieved, that for the successful implementation of the program, consistency and coordination of bodies is necessary, and also noted the importance of using Big Data technologies, artificial intelligence algorithms, which contributes to the formation of systems databases with the highest level of reliability, discard incorrect or outdated information, systematize data and provide an objective whole picture. The Head of State instructed the Government to develop a set of regulatory and incentive measures to expand the Industry 4.0 project.

Digitalization of medical laboratories is a promising area that is part of e-health. Medical laboratories accumulate huge amounts of information based on the results of laboratory tests, but the information is so diverse that it carries elements of “patchwork digitalization.”

According to McKinsey analysts, five years ago, even the dynamic development of medical technologies, the healthcare sector in terms of the applied application of “Big Data” lagged significantly behind other industries, this is due to the lack of trust of doctors in the provided conclusions made using information technology tools. Innovations in data science (translation from English - data science) can not only help patients, but also relieve doctors from some of their routine work. The advantages are demonstrated by chat bots that optimize some processes in clinical practice.

There are many laboratory information systems on the world market that serve to automate operational and treatment processes in the laboratory of their country. Each laboratory information system is adapted to the legislative norms of its state, and often scaling to other countries seems to be an extremely complex process, so adapting the system can take more than one year.

The market of the Republic of Kazakhstan includes both domestic developers of laboratory information systems and solutions from neighboring countries, such as LIS "K-Lab" (LLP "Meditec", Russia-Kazakhstan), LIS "Ariadna" (LLC "Bregis", Russia), LIS "Siroca" (LLP "SIROCA TECHNOLOGY", Kyrgyzstan), LIS "Info Lab" (TOO "Inform Medical", Kazakhstan), LIS "CS-Soft" (LLC "Mobile Service Group", Kyrgyzstan), LIS "TerraLab" (Terralab SOFT LLC, Ukraine), LIS "Damu" (KIIS) (Information Technology Center "DAMU" LLP, Kazakhstan), LIS "SmartLab" (SmartLab Kazakhstan LLP, Kazakhstan). The advantage of Kazakhstani developers is efficiency and flexibility, and the developers are located in the country, which allows them to quickly carry out developments to meet the needs of medical laboratories and adapt to local legislation.

As part of the work, it was revealed that there are no approved norms and standards for laboratory research in Kazakhstan. Existing orders and standards focus on the work of the laboratory itself, and not on the results of laboratory tests. Typically, most laboratories use the reference values specified in the instructions for the reagents supplied with the equipment, and this leads to error in their work. Since for each reagent there is an indication that, regardless of the standards specified in the instructions for the reagent, the laboratory must determine its standards, depending on the region and population.

The importance of developing information systems and introducing artificial intelligence technologies in modern medical practice cannot be overestimated. These technologies play a key role in enhancing the efficiency of the diagnostic and treatment process, ensuring timely and accurate treatment for patients. By providing essential support for medical decision-making, AI information systems and algorithms are becoming a valuable tool for medical professionals, allowing them to interpret data, minimize the risk of errors and optimize patient management processes quickly and accurately. This approach not only speeds up the diagnosis process, but also facilitates more accurate and timely treatment, thereby providing significant social and economic benefits.

The aim of the study – development of an artificial intelligence model aimed at interpreting the results of laboratory tests in the healthcare sector to automate the detection of pathology.

To achieve this goal, the following **tasks** were set and solved:

- study of analogues in the world practice of medical laboratory diagnostics;
- generation of BIG DATA of laboratory test results for further analysis of biochemical, immunological and hematological data;
- formation of a unified database of reference values for laboratory tests by manufacturers of laboratory equipment and regional characteristics of the Republic of Kazakhstan;
- formation of a methodology for complex automation in the field of interpretation of laboratory research data based on big data (BIG DATA);
- development of an artificial intelligence model for identifying pathology in laboratory research results based on complex automation.
- Development of modules for interpreting laboratory test results within the laboratory information system “SmartLab”.

Object of the study: laboratory diagnostics in healthcare, medical laboratory equipment with proper medical documentation, artificial intelligence in the laboratory.

Subject of the study: artificial intelligence models and algorithms for identifying pathology in laboratory test results.

Scientific results within the framework of the requirements for the dissertation

The introduction substantiates the relevance of the topic, the level of study and scientific elaboration. The object, subject, purpose, objectives, and methodology of the dissertation research are presented. Scientific novelty, scientific positions presented for defense, and the practical significance of the research results are highlighted. Data on the testing and publication of research results are provided, and the author’s personal contribution to scientific research is also indicated.

The first section provides an overview and analysis of existing artificial intelligence models and systems for laboratory diagnostics in healthcare, including an analysis of the current state of laboratory diagnostics in healthcare, analysis and prospects for the development of artificial intelligence in healthcare, and an overview of work on the development of an artificial intelligence model in

laboratory diagnostics , as well as a review of laboratory information systems presented on the market of the Republic of Kazakhstan.

Section two discusses the modeling of laboratory research results analysis processes. Describes the application of systems analysis to laboratory diagnostics, automation of laboratory operations, and methods for optimal use of laboratory tests for the diagnosis and prognosis of pathologies. The creation of a knowledge base for analyzing laboratory results and modeling the interpretation of standard values based on their deviations from established norms are also discussed.

The third section presents the results of developing an artificial intelligence model for laboratory diagnostics in healthcare, the use of machine learning methods for analyzing data using BIG DATA of laboratory research results, the results of developing an artificial intelligence model for laboratory diagnostics, a logical model of artificial intelligence for interpreting results, and substantiates the mathematical artificial intelligence model for interpreting results.

The fourth section is devoted to the practical implementation and results of implementing the architecture of the information system; the technologies for implementing LIS, the architecture of the database, modules of the information system and their interaction are presented, as well as the results of implementation in laboratories on the territory of the Republic of Kazakhstan.

Scientific novelty:

1. An artificial intelligence model has been developed for interpreting the results of laboratory tests in healthcare to automate the detection of pathology.

2. A unified database of reference values of laboratory tests has been formed for manufacturers of laboratory equipment and regional characteristics of the Republic of Kazakhstan, which makes it possible to increase the reliability of laboratory research results.

3. A universal platform for integration with laboratory equipment has been developed.

4. An algorithm has been developed for applying machine learning methods in neural network stacking and gradient boosting.

Provisions to be defended:

1. A comprehensive automation technique has been developed for interpreting the results of laboratory tests using BIG DATA (biochemical, immunological, and hematological tests).

2. Methodology for forming a unified database of reference values for laboratory research by manufacturers of laboratory equipment and regional characteristics of the Republic of Kazakhstan.

3. A mathematical model has been developed to identify pathology in the results of laboratory tests based on automation using machine learning and artificial intelligence methods.

4. Modules have been developed in the laboratory information system “SmartLab” for automatic detection of pathology in the results of laboratory tests based on a set of artificial intelligence models, integrated with medical laboratory analyzers.

Practical value and implementation of research results

1. As a result of the research, a Unified database of reference values for laboratory tests was formed for manufacturers of laboratory equipment and regional characteristics of the Republic of Kazakhstan, which received expert assessment from specialists in the field of healthcare and was implemented.

2. Algorithms and software modules have been developed for automatic detection of pathology in the results of laboratory tests based on a set of artificial intelligence models, integrated with medical laboratory analyzers.

3. A platform has been developed for integration with laboratory analyzers, which can be used in other areas (financial sector, IoT technologies, access control systems).

4. Services have been developed for the use of OpenAI, which allows artificial intelligence to be trained by other artificial intelligence.

Methodological basis and metrological support for research

In the dissertation work, at each stage of implementation, a number of the following methods were used: System analysis method - the formation of technical, laboratory data, implementation algorithms for the operational activities of a medical laboratory; Decomposition method – development of program code, modules and all functionality of the information system; Machine learning methods – identifying patterns when studying a large array of data; Methods of relational algebra - all operations on data, including functionality for identifying pathology in the results of laboratory tests. Metrological (measuring) work was carried out on modern laboratory analyzers of various models, which undergo daily quality control for the studies performed, as well as annual verification by standardization and metrology bodies.

The reliability and validity of the results obtained, conclusions and recommendations set out in the dissertation are confirmed by acts of implementation in clinical diagnostic laboratories, as well as by indicators for assessing the prediction accuracy of machine learning models.

Publications and personal contribution of the applicant. 13 scientific papers were published, including 3 works in publications recommended by the Committee for Control in the Sphere of Education and Science of the Ministry of Education and Science of the Republic of Kazakhstan, 1 article in an international scientific publication included in the Scopus database, 4 articles in other publications, 5 articles in international scientific practical conferences. Received 2 certificates of state registration of a computer program, 1 certificate of the “SmartLab” trademark.

Structure and scope of the dissertation work. The presented dissertation work consists of an introduction, four sections, a conclusion, a list of sources used from 101 titles, presented on 131 pages of computer text, includes 63 figures, 18 tables and 6 appendices.