

АННОТАЦИЯ
диссертационной работы докторанта PhD
специальности 6D070300 – «Информационные системы»
Мукашевой Асель Коптлеувны
на тему «Исследование и разработка информационной
системы диагностики сахарного диабета на базе инструментов BigData
технологий»

Актуальность темы. Распространенность сахарного диабета в Казахстане носит характер эпидемии и приводит к большим финансовым затратам. Проведённый анализ современного состояния исследований по прогнозированию роста пациентов показал недостаточную проработанность в данной предметной области. Для усовершенствования данного вопроса было предложено использовать реальные статистические данные, с помощью которых можно выявить наиболее эффективный метод для прогнозирования роста пациентов сахарным диабетом на фоне пассивного выявления для дальнейшего определения количества закупаемого инсулина по регионам. Объем спрогнозированного количества инсулина в свою очередь позволит запланировать финансовые затраты во избежание лишних затрат.

Актуальность работы подчёркивается высоким уровнем распространенности болезни и отсутствием проблемно-ориентированных информационных технологий диагностики и выработки рекомендаций по лечению сахарного диабета. В связи с этим потребность в исследовании и разработке информационной системы поддержки врача-эндокринолога по диагностике сахарного диабета на базе инструментов BigData технологий высока, так как разработанная система позволит врачам использовать неструктурированные данные, на основе которых можно принимать оптимальное решение. Применение современных информационных технологий в системе здравоохранения позволит улучшить качество оказываемых медицинских услуг. Поэтому исследование этих вопросов определяет актуальность темы данной диссертационной работы.

Цель диссертационной работы. Исследование методов прогнозирования роста пациентов СД по Республике Казахстан, а также разработка информационной системы поддержки врача-эндокринолога по диагностике сахарного диабета на базе инструментов BigData технологии.

Поставленная цель определила основные задачи диссертации.

1. Проведение анализа современного состояния применения методов прогнозирования и использования технологий BigData в сфере медицинского обслуживания.

2. Разработка моделей и методов прогнозирования роста пациентов с помощью методов регрессионного анализа и с применением библиотеки scikit-learn для определения количества закупаемого инсулина.

3. Разработка системы на основе математического и программного обеспечения по диагностике сахарного диабета.

4. Разработка платформы информационной системы информационных ресурсов по сахарному диабету в среде BigData, а также сбор исходной информации.

5. Проведение исследовательского анализа, обработки больших данных с помощью сводной статистики и графических представлений данных.

6. Выполнение экспериментальных исследований, проведение анализа полученных результатов и формулировка выводов по работе в целом.

Таким образом, сформулирована цель и задачи исследования по прогнозированию возможного количества пациентов, которая позволит заранее запланировать количество закупаемого инсулина и других средств контроля диабета. В дальнейшем, планируется запустить данную информационную систему в открытом доступе для практического обучения методам диагностирования СД на основе полученных теоретических знаний студентами медицинских университетов.

Объектами исследования являются процессы прогнозирования больных СД на последующие года с целью подготовки процедуры закупа необходимого количества инсулина и информационная система диагностики пациентов в медицинских учреждениях на основе введенных данных с помощью инструментов технологии BigData.

Научная новизна диссертационной работы заключается в следующем:

– На основе множества допустимых решений статистических задач, полученные результаты позволили спрогнозировать количество пациентов.

– Распределённые вычисления, обеспечивающие эффективное выполнение разработанных алгоритмов;

– Разработана концептуальная модель процессов поиска, извлечения, обработки и анализа данных в больших массивах данных;

– Разработанная и апробированная на основе системного подхода информационная система для диагностики сахарного диабета на базе инструментов BigData технологий.

Методы исследования. Теоретическая основа диссертационной работы базируется на методах математической статистики, машинного обучения, математического анализа. Практическая реализация алгоритмов осуществлялась на основе методов объектно-ориентированного программирования и объектно-ориентированных баз данных.

Практическая значимость работы. Применение в системе методов для прогнозирования роста пациентов СД на последующие года позволит заранее запланировать количество закупаемого инсулина и других средств контроля диабета. Разработанная информационная система по диагностике СД позволила

на основе алгоритмического и программного обеспечения повысить качество диагностирования болезни.

Результаты исследований диссертационной работы позволят облегчить работу врачей-эндокринологов, т.к. ими могут быть использованы не только собственные знания, но и мировые ресурсы базы знаний о сахарном диабете. Также пользу от использования данной системы получают высшие учебные заведения, осуществляющие подготовку и обучение врачей-эндокринологов. Практическая ценность диссертационного исследования подтверждается полученным актом о внедрении результатов работы.

Апробация работы. Основные положения и результаты диссертационного исследования докладывались и обсуждались на научных семинарах кафедры «Кибербезопасность, обработка и хранение информации» КазНИТУ имени К.И.Сатпаева, а также на научных конференциях различного уровня.

Публикации. По теме диссертации опубликовано 14 публикаций, в том числе 5 статьи – в научных изданиях, рекомендованных Комитетом по контролю в сфере образования и науки МОН РК; 1 статья – в зарубежном журнале, которая входит в базу данных ISI Web of Knowledge (IF=2,827) и Scopus (76-перцентиль), 8 работ – в материалах международных конференций.

Структура и объем диссертации. Диссертация состоит из введения, 4 глав, заключения, списка использованных источников и приложения. Она изложена на 117 страницах машинописного текста, содержит 88 рисунков, 6 таблиц, список использованных источников из 139 наименований и приложений.

Во введении раскрыты актуальность, конкретизированы проблемы, связанные с исследуемой темой. Приведены цель и задачи исследования, научная новизна и практическая ценность работы, методы исследования.

В первой главе диссертации представлены актуальные проблемы диагностики и прогнозирования сахарного диабета в современном мире. Проанализированы недостатки распространенности сахарного диабета в мире. Проанализированы достоинства и недостатки существующих на сегодняшний день современных медицинских информационных систем диагностики. Приведены цели и задачи исследования.

Во второй главе диссертации исследованы методы применения технологии BigData в сфере мирового здравоохранения. Рассмотрены и проанализированы информационные технологии, решающие медицинские задачи.

В третьей главе описаны модели прогнозирования на базе методов регрессионного анализа. В результате экспериментов для сравнительного анализа между проведенными вычислениями были дополнительно рассмотрены возможности библиотеки scikit-learn для решения задач машинного обучения, которые тоже продемонстрировали рост пациентов. Проведен

исследовательский анализ обработки больших данных с помощью сводной статистики и графических представлений данных.

В четвертой главе на основе программной реализации информационной системы рассмотрены результаты исследования диагностики сахарного диабета с помощью инструментов BigData технологии. Разработана информационная модель диагностирования СД с помощью БД, и на основе математического аппарата показан перечень стандартных медицинских анализов, с помощью которых можно диагностировать диабет.

В заключении отражены основные результаты и выводы диссертационной работы.

По теме диссертации опубликовано 14 публикаций:

1. А.К. Мукашева, Н.П. Сапарходжаев. «Қазақстан Республикасының аумағында BigData технологиясы негізінде науқастарға қызмет көрсету жүйесін әзірлеу». Вестник КазННТУ, №6 (124) ноябрь 2017г. – Алматы: КазННТУ имени Сатпаева, 2017, стр. 193-197.

2. Н. П. Сапарходжаев, Г. К. Балбаев, А.К. Мукашева. «Разработка информационной системы на основе технологий BigData для диагностики и лечения диабета». Вестник АУЭС, №4 (6) (43) 2018, Алматы

3. Н.П. Сапарходжаев, А.К. Мукашева. «Параметрлерді және ұсыныстарды бағалау үшін ақпараттық ресурстарды пайдалануға және қалыптастыруға арналған ақпарат жүйесін құрастыру: BigData құралдары негізінде MongoDB-мен жұмыс». Вестник КазННТУ, №6 (130) 2018г. – Алматы: КазННТУ имени Сатпаева, 2018, стр. 198-204.

4. Н.П. Сапарходжаев, А.К. Мукашева. «Прогнозирование распространенности диабета в республике Казахстан на основе методов регрессионного анализа». Вестник АУЭС, №5 2019, Алматы

5. А.К. Мукашева, Н.П. Сапарходжаев. «Прогнозирование распространенности диабета в республике Казахстан на основе методов регрессионного анализа». Вестник КазННТУ, №5 () 2019г. – Алматы: КазННТУ имени Сатпаева, 2019.

6. А.К. Mukasheva, N. Saparkhojayev. «The concept of monetization of IoT-based project: case of Medical System in Kazakhstan». The 15th International conference information technologies and management 2017 April 27-28, 2017, ISMA University, Riga, Latvia.

7. А.К. Мукашева, Н.П. Сапарходжаев. «Разработка системы обслуживания пациентов на основе технологии BigData на территории Республики Казахстан». МНПК «Математические методы и информационные технологии макроэкономического анализа и экономической политики» КазННТУ им. К.И.Сатпаева, стр. 182-187, 11-12 апреля 2017г. – с.182-187

8. Н.П. Сапарходжаев., А.К. Мукашева, «Анализ системы для диагностики сахарного диабета на основе технологии BigData» ISBN 978-601-323-111-2. Труды Международных Сатпаевских чтений «Инновационные решения

традиционных проблем: инженерия и технологии», 12 апреля 2018г. И66 – Алматы: КазНИТУ имени Сатпаева, 2018.-1254-1256.

9. N. Saparkhojayev, A.K. Mukasheva, «Introduction to BigData technology for diagnosis of diabetes». Information Technologies and Management, 2018 April 26 – 27, “Information Technologies, Management and Society”, ISMA University, Riga, Latvia.

10. N. Saparkhojayev, A.K. Mukasheva, «The development of information system of formation and use of information resources for evaluation of parameters and evaluation of recommendations based on BigData technology tools: work with MongoDB». International Conference on Cyber Security and Computer Science (ICONCS'18), Oct 18-20, 2018 Safranbolu, Turkey

11. A. Mukasheva, N. Saparkhojayev, Z. Akanov, A. Algazieva, «Prevalence of diabetes in the republic of Kazakhstan based on regression analysis methods». ISBER. International Conference on Research in E-Learning & Distance Education, Social Sciences, Economics and Management. July 13-14, 2019 / Turkey.

12. Mukasheva A., Saparkhojayev N., Akanov Z. et al. Forecasting the Prevalence of Diabetes Mellitus Using Econometric Models. Diabetes Therapy (2019). <https://doi.org/10.1007/s13300-019-00684-1>.

13. Мукашева А. К., Н. П. Сапарходжаев, И. А. Зимин. «Разработка информационной системы на основе технологии BigData». «Научное творчество: научно-исследовательская практика и приоритетные направления» / Материалы XXII республиканской научной конференции студентов и молодых ученых, посвященной 1150-летию Абу Насыр аль-Фараби: Том I. (23 апреля 2020 г), стр. 527-532, ISBN 978-601-285-096-3.

14. А. А. Куандыков, А.К. Мукашева, И. А. Зимин, Обработка журналов веб серверов для получения статистики сайтов с помощью инструмента Apache Spark. МНПК «Модели инновационных решений повышения конкурентоспособности отечественной науки» (4 июня 2020 г, г. Челябинск). В 2 ч. Ч. 2 / - Уфа: OMEGA SCIENCE, 2020. – стр. 79-84, ISBN 978-5-907347-21-2.