

УДК 004.021

BIG DATA КАК ОСНОВА ПОСТРОЕНИЯ ЭФФЕКТИВНОЙ СИСТЕМЫ МОНИТОРИНГА ПАНДЕМИЙ НА ПРИМЕРЕ COVID-19



Д.В. Орлов
Студент 4 курса,
кафедра ЭВМ, БГУИР
dim4work@gmail.com



С.Н. Нестеренков
Кандидат технических
наук, доцент,
декан факультета
компьютерных систем и
сетей
s.nesterenkov@bsuir.by



Д. А. Жалейко
инженер-программист
кафедра ЭВМ, БГУИР
a.n.markov@bsuir.by

Д.В. Орлов

Студент 4 курса специальности “Вычислительные машины, системы и сети” БГУИР.

С.Н. Нестеренков

Кандидат технических наук, доцент, декан факультета компьютерных систем и сетей Белорусского государственного университета информатики и радиоэлектроники, доцент кафедры программного обеспечения информационных технологий. Автор публикаций на тему машинного обучения, алгоритмов принятия решений, искусственных нейронных сетей и автоматизации.

Д.А. Жалейко

Окончил БГУИР в 2021 году по специальности “Вычислительные машины системы и сети”

Аннотация. После вспышки коронавируса (SARS-CoV-2) в 2019 году он заразил миллионы людей и унес жизни десятки тысяч человек. Во время пандемии коронавируса, применение Big Data стало одним из немногих мощных средств для борьбы с вирусом. В настоящее время многие страны и исследовательские учреждения используют Big Data для отслеживания и контроля распространения инфекционного заболевания. В будущем люди смогут использовать возможности Big Data для борьбы с такими эпидемиями. К примеру сравнительное геномное исследование вариантов вируса, ускоренное с помощью Big Data, может дать важную информацию о мутациях вируса. Эта статья поведает о применении Big Data во время пандемии COVID-19 и как бороться с подобными эпидемиями в будущем. Много приложений и программных решений было создано с использованием этих данных, отслеживающих и предсказывающих поведение инфекции. Следовательно такие приложения и методы обработки помогут человечеству в борьбе с похожими случаями в будущем.

Ключевые слова: Big Data, COVID-19, эпидемия, общество.

Введение.

Пандемия COVID-19 стала беспрецедентной ситуацией в области общественного здравоохранения. Эта болезнь стала серьезным вызовом эпидемиологическому планированию и системе здравоохранения.

Этот инцидент в области общественного здравоохранения, охвативший весь мир, не является первым в истории человечества; подобные случаи происходили несколько раз за последние сотни лет. Черная смерть, которая бушевала с 1346 по 1353 годам, унесла жизни около 200 миллионов людей. Пандемия холеры с 1852 по 1860 годы унесла около 1 миллиона жизней. Тем не менее с развитием медицина и института здравоохранения появление такой эпидемии как Коронавирус, предупреждает мир, что он не готов к таким случаям.

По данным Всемирной Организации Здравоохранения, по состоянию на 30 августа 2022 года, 216.303.376 случаев COVID-19 было подтверждено во всем мире, и число погибших достигло 4.498.451. Среди них США имеет наибольшее количество подтвержденных случаев, Индия идет следующей. Хотя более 4 миллиардов вакцин было выпущено, это никак не изменило ситуацию из-за нового варианта вируса Delta.

В настоящее время глобальные цели по борьбе с вирусом сосредоточены на сведении к минимуму его распространения и смертности, а также минимизировать ущерб обществу. Поэтому многие научно-исследовательские институты и государственные учреждения работают сообща и используют передовые технологии: анализ Big Data и искусственный интеллект для борьбы с такими пандемиями. Например США и Шанхай используют преимущества анализа Big Data для борьбы с коронавирусом. Ключевая информация, такая как физическое состояние и история путешествий более 100.000 человек была записана, чтобы справиться с пандемией. Эта статья направлена на применяемые технологии Big Data в условиях коронавируса. Узнайте, как использовать эти инструменты, чтобы позволить человечеству лучше реагировать на такие вспышки в будущем [1].

Приложения использующие Big Data для исследования COVID-19

Некоторые исследователи и правительственные учреждения используют Big Data для отслеживания случаев заражения в реальном времени. Анализ эпидемии включает в себя все данные такие как положительные случаи заражения, смерти, людей выздоровевших от болезни, истории поездок, плотность населения, контакты с инфицированными. С применением искусственного интеллекта и машинного обучения, данные могут быть обработаны для построения модели болезни, которая может предсказать уровень заражения (высокий или низкий) и его влияние [2]. AarogyaSetu, мобильное приложение для слежения, запущенное правительством Индии, является отличным примером того, как Big Data можно использовать для борьбы с болезнью. Рабочий процесс этого приложения представлен на рисунке.



Рисунок 1. Рабочий процесс AarogyaSetu

Собирая данные о местоположении через GPS или Wi-Fi, а затем сравнивая его с существующей базой данных, приложение может выявить людей, которые могли контактировать с людьми, у которых подтвержденный случай заболевания, что обеспечивает раннее обнаружение и отслеживание цепочки заражения.

Кроме того Big Data также применяется для диагностики и прогноза COVID-19. Поскольку радиологическое проявление инфекции ковид и вирусной пневмонии очень

похожи, различие их в клинических условиях становится очень важным заданием. Таким образом КТ грудной клетки пациентов инфицированных COVID-19 и больных вирусной пневмонией используются в качестве входных данных системы глубокого обучения. В итоге финальная точность различия этих болезней составила 86%, а уровень точности в различии пациентов с ковид и здоровых составила 94%. Машинное обучение также приняло участие при разработке алгоритма прогнозирования.

Алгоритм используется для прогнозирования риска смертности инфицированных людей. Это помогает выявить пациентов с потенциальной потребностью в дальнейшей медицинской помощи, что позволяет медицинскому персоналу рационально распределять ресурсы больницы.

Будущее применение Big Data при пандемиях

Несмотря на некоторый прогресс в разработке вакцин и терапевтических средств, COVID-19 остается серьезной угрозой для здоровья населения, и продолжается исследование и внедрение новых методов борьбы с этой инфекцией. Исследователи разработали инструменты и методы из прошлого опыта, который может быть полезен в борьбе как с COVID-19 так и другими подобными заболеваниями в будущем. В этом разделе будут описаны четыре основные области применения, включая отслеживание, предсказание, диагностику и лечение.

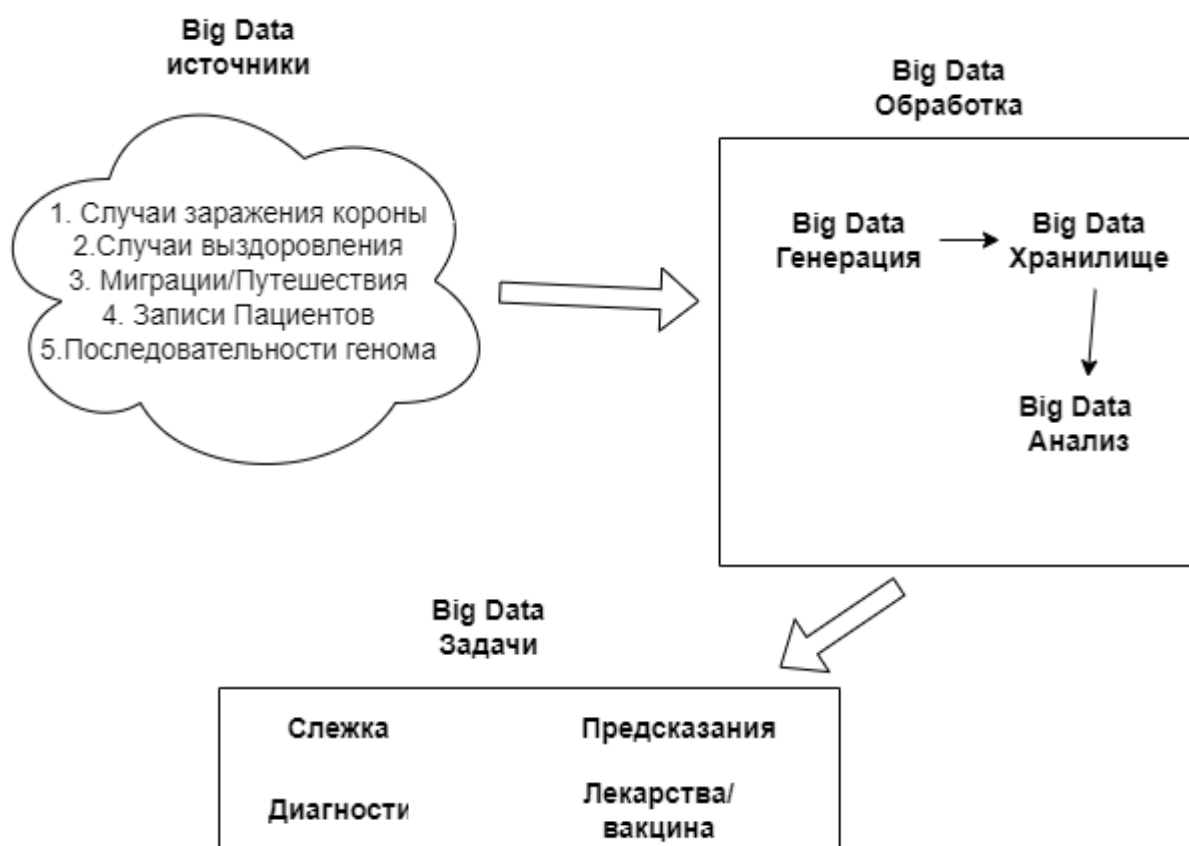


Рисунок 2. Борьба с инфекциями используя Big Data

Одна из особенностей COVID-19 является то, что больные могут передавать вирус до того, как у них появятся какие-либо симптомы. Следовательно вирус становится очень заразным, это критично для правительства, чтобы отслеживать и сдерживать инфекцию, а также принимать своевременные меры для борьбы.

Помимо отслеживания и прогнозирования распространения вируса, Big Data также могут применяться для диагностики и лечения. Надежная, быстрая и удобная диагностика необходима для борьбы с вирусом.

В условиях пандемии это может спасти большое количество жизней, но и ограничить его распространение; большое количество медицинских данных могут выступать в качестве обучающего набора для машинного обучения. Прежде всего, применение Big Data может значительно сократить время обнаружения болезни. Так как для определения точного исследования нужно обладать большим количеством оборудования, денег и времени. Следовательно, у многих стран нету возможностей, чтобы эффективно противостоять вирусу.

Мобильное приложение основанное на Big Data: AI4COVID-19, которое можно использовать для постановки диагноза на основании кашля.

Точность этого приложения при диагностике достигает 90%, что делает его удобным, дешевым и эффективным средством для борьбы с вирусом, которое может позволить любое медицинское учреждение.

Big Data широко используется в исследованиях лекарственных средств. Вакцины, таблетки и т.п., разработанные на основе этих данных, могут сыграть решающую роль в реагировании на подобные эпидемии. Google DeepMind - компания известная известная благодаря программе для игры в го AlphaGo, теперь может использовать искусственный интеллект для предсказания мембраны SARS-CoV-2 белков, которые могут быть полезны для разработки новых лекарств [3-4].

Заключение

Были рассмотрены и представлены механизмы контроля и профилактические меры, которые люди использовали до сих пор, включающие в себя Big Data и инструменты работы над ними, в условиях глобального кризиса вызванного COVID-19 пандемией. Также исследуются будущее развитие этого направления и инструменты для чрезвычайного реагирования на подобные эпидемии.

Во-первых, Big Data можно использовать для разработки приложений отслеживания и прогнозирования болезни. Затем более быстрый и безопасный клинический диагноз может быть поставлен с помощью искусственного интеллекта на основе большого кластера данных.

Во-вторых, данные Big Data могут быть задействованы при изучении вируса, и поисков лекарства от него. Хотя уже сейчас видны положительные результаты применения этих технологий для борьбы с коронавирусом, методы обработки данных находятся на относительно ранней стадии. Следовательно, в данном случае трудно судить о том, какой алгоритм больше подходит для обнаружения вирусов. Другой вопрос - конфиденциальность и безопасность персональных данных людей. Во время пандемии COVID-19 правительства просили граждан делиться личной информацией такой как: местоположение, история путешествий, отчеты тестов и т.д.

Однако люди часто не желают делиться своими личными данными, потому что они заботятся о ее конфиденциальности.

Более того Big Data могут содержать дезинформацию, что несомненно имело негативное влияние на борьбу с пандемией.

Несмотря на то что имеется ряд проблем, это не отменяет того факта что Big Data во время пандемии является мощным оружием для людей в борьбе с болезнью. В будущем нужно проводить исследования в этой области для решения проблем на этапе сбора данных и их обработки.

Список литературы

- [1] Big Data: Concepts, Technology and Architecture / В. Balusamy [и др.]. – Hoboken : John Wiley & Sons, Inc., 2021. – 368 с.
- [2] Big-Data Analytics and Cloud Computing: Theory, Algorithms and Applications / М. Trovati [и др.]. – Berlin : Springer International Publishing, 2015. – 185 с.
- [3] Нестеренков, С.Н. Применение больших данных в электронном образовании / С.Н. Нестеренков, М.И. Макаров, Н.В. Ющенко, А.Д. Радкевич // BIG DATA and Advanced Analytics = BIG DATA и анализ высокого уровня : сб. материалов V Междунар. науч.-практ. конф. (Республика Беларусь, Минск, 13-14 марта 2019 года). В 2 ч. Ч. 2 / редкол. : В. А. Богуш [и др.]. - Минск : БГУИР, 2019. - С. 242-245.
- [4] Беляк, А. А. Анализ производительности технологии Hadoop / А. А. Беляк, С. Н. Нестеренков // BIG DATA and Advanced Analytics = BI DATA и анализ высокого уровня: сб. научных статей VII Междунар. науч.-практ. конф. (Республика Беларусь, Минск, 19-20 мая 2021 года): / редкол. : В. А. Богуш [и др.]. – Минск : Бестпринт, 2021. – С. 343–346.
- [5] Security and Privacy for Big Data, Cloud Computing and Applications / W. Ren [и др.]. – London : The Institution of Engineering and Technology, 2019. – 328 с.
- [6] Калоша, А.Л. Система анализа качества текстовых коллекций / А.Л. Калоша, М.А. Медунецкий, М.П. Хоронко, А.А. Александров, А.И. Гридасов, С.Н. Нестеренков // BIG DATA and Advanced Analytics = BIG DATA и анализ высокого уровня : сб. материалов VI Междунар. науч.-практ. конф. (Республика Беларусь, Минск, 20-21 мая 2020 года): в 3 ч. Ч. 2 / редкол. : В. А. Богуш [и др.]. - Минск : Бестпринт, 2020. - С. 369-375.

BIG DATA AS A BASIS FOR BUILDING AN EFFECTIVE PANDEMIC MONITORING SYSTEM ON THE EXAMPLE OF COVID-19

D.V. ORLOV

*Pregraduated student of the
BSUIR,
department of electronic
computers*

S.N. NESTERENKOV

*PhD, Associate professor, Dean
of the Faculty of Computer
Systems and Networks*

D.A. ZHALEIKO

*Software engineer
BSUIR,
department of electronic
computers*

Center for Informatization and Development of the Belarusian University of State Informatics and Radioelectronics, Republic of Belarus.

E-mail: dimm4work@gmail.com, s.nesterenkov@bsuir.by, a.n.markov@bsuir.by

Abstract. Since the outbreak of the coronavirus (SARS-CoV-2) in 2019, it has infected millions and claimed the lives of tens of thousands. During the coronavirus pandemic, the use of Big Data has become one of the few powerful tools to fight the virus. Currently, many countries and research institutions use Big Data to track and control the spread of an infectious disease. In the future, people will be able to use the power of Big Data to fight such epidemics. For example, a comparative genomic study of virus variants, accelerated with the help of Big Data, can provide important information about virus mutations. This article will tell you about the use of Big Data during the COVID-19 pandemic and how to deal with similar epidemics in the future. Many applications and software solutions have been created using this data to track and predict infection behavior. Therefore, such applications and processing methods will help humanity in the fight against similar cases in the future.

Keywords. Big Data, COVID-19, epidemic, society