

## ОБЗОР ЭКСПЕРТНЫХ СИСТЕМ ПО ОПРЕДЕЛЕНИЮ ПАТОЛОГИЙ БЕРЕМЕННОСТИ С ИСПОЛЬЗОВАНИЕМ ТЕХНОЛОГИЙ МАШИННОГО ОБУЧЕНИЯ

Рипка Д. С.<sup>1</sup>, аспирант, [dar.stkr@gmail.com](mailto:dar.stkr@gmail.com)  
Рыбин С. В.<sup>1</sup>, канд. физ.-мат. наук, доцент, ✉ [rsvvm2leti@gmail.com](mailto:rsvvm2leti@gmail.com),  
[orcid.org/0000-0002-9095-3168](https://orcid.org/0000-0002-9095-3168)

<sup>1</sup>Санкт-Петербургский государственный электротехнический университет «ЛЭТИ» им. В. И. Ульянова (Ленина), ул. Профессора Попова, 5, корп. 3, 197022, Санкт-Петербург, Россия

### Аннотация

В области медицинской диагностики произошел значительный прогресс благодаря внедрению методов машинного обучения. В данной статье представлен комплексный обзор исследований по разработке экспертных систем для диагностики патологий беременности с использованием различных методов машинного обучения. Подчеркивая решающую роль данных для обучения и их предварительной обработки, в рассмотренных в статье публикациях сравнивается точность различных алгоритмов машинного обучения в этой области. Анализируемые исследования, в первую очередь, были сосредоточены на построении экспертных систем для диагностики патологий и осложнений, связанных с беременностью, с целью улучшения результатов лечения будущих матерей и их будущих детей. Используя широкий спектр подходов к машинному обучению, включая деревья решений, машины опорных векторов, случайный лес, искусственные нейронные сети и другие, была исследована эффективность каждого рассмотренного в статье алгоритма в точном прогнозировании проблем, связанных с беременностью.

Одним из ключевых моментов этого обзора является акцент на качестве и разнообразии обучающих данных. Особое внимание при анализе публикаций уделялось надежности и полноте наборов данных, позволяющих алгоритмам машинного обучения достигать более высокой диагностической точности. Для обзора были отобраны исследования, в которых для обучения были использованы параметры, признанные медицинским сообществом индикаторами различных патологий.

Важным критерием для выбора публикаций обзора было наличие в них предварительной обработки данных для коррекции пропущенных значений, шума и дисбаланса классов. Решение этих задач играет существенную роль в повышении производительности экспертных систем. В статье приводится оценка методологии и алгоритмов, используемых в рассмотренных публикациях, что дает ценную информацию для будущих исследований и разработок в этой области.

**Ключевые слова:** диагностика, патологии беременности, экспертные системы, машинное обучение.

**Цитирование:** Рипка Д. С., Рыбин С. В. Обзор экспертных систем по определению патологий беременности с использованием технологий машинного обучения // Компьютерные инструменты в образовании. 2023. № 4. С. 50–64. doi:10.32603/2071-2340-2023-4-50-64

## 1. ВВЕДЕНИЕ

Достижения в области машинного обучения и искусственного интеллекта произвели революцию в области медицинской диагностики, предложив новые возможности для раннего выявления и точного лечения различных заболеваний. В последние годы исследователи сосредоточили свое внимание на разработке экспертных систем, специально предназначенных для диагностики патологий беременности с целью улучшения результатов охраны здоровья матери и плода. Сочетание медицинского опыта с передовыми методами машинного обучения дает огромный потенциал для повышения точности и эффективности диагностики патологий беременности.

Статья состоит из трех частей: первая часть посвящена рассмотрению перспективных экспертных систем, во второй части рассмотрены последние аналитические обзоры методов машинного обучения (Machine Learning, ML) для прогнозирования патологий беременности, заключительная часть посвящена оценке трудностей и рисков применения машинного обучения в диагностике перинатальных осложнений.

## 2. ЭКСПЕРТНЫЕ СИСТЕМЫ

В исследовании D. S. Maylawati и соавторов предлагается экспертная система, использующая глубокую нейронную сеть (Deep neural network, DNN) с применением алгоритма обратного распространения ошибки (Back propagation algorithm, BPA) [1], которая направлена на раннее прогнозирование беременности с нарушениями [2]. Среди нарушений были выделены такие, как гиперемезис<sup>1</sup>, преэклампсия<sup>2</sup> и эклампсия<sup>3</sup>, пузырный занос<sup>4</sup> и внематочная беременность<sup>5</sup>.

Для исследования были взяты данные 172 медицинских карт с 17 входными параметрами и 5 выходными классами, один из которых — нормальное течение беременности. DNN содержит входной уровень, скрытый уровень и выходной уровень. А также обратное распространение обновляет вес в обратном процессе.

Обучение и тестирование проводилось по 6 сценариям: от использования 100 % данных для обучения и тестирования до использования 50 % данных для обучения и 50 % данных для тестирования с шагом в 10 %. Самая высокая точность наблюдалась, когда для обучения использовалось 80 % данных. В этом случае она составила 85,71 %. Для других случаев точность была значительно ниже, а ошибка составила 0,01.

Группа исследователей под руководством Н. К. V. S. Raghav изучала вопрос выявления женщин с высоким риском преждевременных родов, что важно для планирования беременности [3].

В исследовании приняли участие 170 пациентов, на вход нейронной сети подавалось 9 параметров. В исследовании использовались методы мягких вычислений, такие как регрессия Softmax с использованием нейронной сети (Deep neural network, DNN), и оптимизатор градиентного спуска для прогнозирования вероятности преждевременных родов. DNN в данной работе имитирует структуру и функции биологических нейронных сетей. Это нелинейная статистическая модель данных, используемая для определения

<sup>1</sup> Гиперемезис (hyperemesis gravidarum) — патологическое состояние в первой половине беременности.

<sup>2</sup> Преэклампсия — повышение артериального давления, сопровождается чрезмерным наличием белка в моче.

<sup>3</sup> Эклампсия — судороги у женщин с преэклампсией при отсутствии другой причины.

<sup>4</sup> Пузырный занос — патологические изменения внешнего слоя оболочки зародыша.

<sup>5</sup> Внематочная беременность — состояние, при котором плодное яйцо прикрепляется вне полости матки.

сложных взаимосвязей между входными и выходными переменными. Глубокая нейронная сеть в этом исследовании была обучена с использованием данного набора данных для прогнозирования вероятности преждевременных родов на основе входных переменных, таких как рост матери, паритет и тяжесть.

Корреляция между ростом матери и преждевременными родами: исследование показало четкую корреляцию между ростом матери и возможностью преждевременных родов. Было обнаружено, что женщины с низким ростом имеют более высокие шансы преждевременных родов по сравнению с женщинами с высоким ростом. Точность разработанного алгоритма составила 89,99 %.

Высокая точность прогностической модели может быть объяснена использованием регрессии Softmax в последнем слое классификатора на основе нейронной сети и Gradient Descent Optimizer [4] для оптимизации весов и смещений модели. Функция Softmax<sup>6</sup> преобразует необработанные выходные оценки модели в вероятности, гарантируя, что сумма предсказанных вероятностей равна 1. Это позволяет модели обеспечить вероятность преждевременных родов для заданного набора входных переменных.

Важно отметить, что исследование имеет ограничения из-за относительно небольшого набора данных. Тем не менее, это служит доказательством концепции и предполагает, что точность модели потенциально может быть улучшена с более крупным и разнообразным набором данных. Ученые из Индонезии исследовали возможность обнаружения преэклампсии с помощью нейронной сети, а также исследовали важность атрибута «Предыдущий случай ПЭ» для результатов классификации [5].

Данные для исследования состояли из 17 параметров, которые, как считается медицинским сообществом, влияют на риск возникновения преэклампсии. Эффективность разработанного алгоритма сравнивается с другими алгоритмами, такими как наивный Байес,  $k$ -ближайших соседей, линейной регрессией, логистической регрессией и методом опорных векторов, и было установлено, что разработанная глубокая нейронная сеть превосходит эти методы.

Для оценки эффективности DNN по сравнению с другими алгоритмами использовались несколько методов кросс-валидации.

1. Метод удержания (Hold-Out Validation): набор данных был разделен на обучающий набор (70 % данных) и тестовый набор (30 % данных). Модель была обучена на обучающем наборе, а затем оценена на тестовом наборе для измерения ее точности.
2.  $k$ -блочная кросс-валидация ( $k$ -Fold Cross Validation) для  $k = 10$ : набор данных был разделен на 10 подмножеств, а модель обучалась и тестировалась 10 раз, каждый раз используя другое подмножество в качестве тестового набора, а остальные — в качестве обучающего набора. Окончательные показатели производительности усреднялись по 10 итерациям.
3. Поэлементная кросс-валидация (Leave-One-Out, LOO): каждая запись данных использовалась в качестве тестового набора, а остальные данные служили обучающим набором. Модель обучалась и тестировалась для каждой записи в наборе данных, а производительность оценивалась на основе того, насколько хорошо модель предсказала каждую отдельную запись.

Эти эксперименты показали, что алгоритм нейронной сети достиг наилучшей точности с тремя проверочными тестами: разделение данных (92,46 %), 10-кратная перекрестная проверка (94,23 %) и проверка LOO (96,66 %). Те же данные применяются к нейронной сети после исключения информации о предыдущем случае ПЭ для процесса обучения. Результатом является правильная классификация до 96,66 % случаев преэклампсии по

<sup>6</sup> Функция Softmax — это обобщение логистической функции для многомерного случая.

всем параметрам в тестовой выборке. Прогнозируемые случаи преэклампсии по сумме результатов неизвестного проверочного теста составляют 90 %. И если не использовать информацию о предыдущем случае ПЭ, результат значительно уменьшится.

Одно из самых крупных исследований, направленных на изучение оценки риска возникновения преэклампсии на ранней стадии беременности, было проведено в 2009 году [6]. Набор данных включал 6838 случаев беременных женщин, из которых только 116 случаев (1,7 %) имели преэклампсию. Исследование было направлено на выявление факторов и параметров, которые можно было бы использовать для прогнозирования возникновения преэклампсии. В конечном итоге были выбраны 15 параметров (из имеющихся 26), наиболее влияющих на характеристику риска возникновения преэклампсии. Эти параметры включали такие характеристики, как среднее артериальное давление (САД), индекс пульсации матки (ИПМ), сывороточный маркер PAPP-A, этническая принадлежность, вес, рост, курение, употребление алкоголя, предыдущий случай преэклампсии, метод зачатия, состояние здоровья, принимаемые лекарства, срок беременности, длина макушки-крестца и семейный анамнез преэклампсии.

В исследовании использовались глубокие нейронные сети в качестве инструмента прогнозирования. В частности, рассматривались различные нейронные структуры с прямой связью, в том числе стандартные многослойные и многослойные структуры с различными размерами и функциями активации. После систематического поиска была предложена многослойная нейронная структура с четырьмя плитами в качестве окончательной прогностической модели.

Нейронная сеть добилась многообещающих результатов для прогнозирования преэклампсии. DNN правильно классифицировала 83,6 % случаев преэклампсии в обучающей выборке. В тестовом наборе нейронная сеть добилась еще лучших результатов, правильно классифицировав 93,8 % случаев преэклампсии. Для совершенно неизвестного проверочного набора нейронная сеть правильно предсказала 100 % случаев преэклампсии.

Напротив, методы множественной линейной регрессии и множественной нелинейной регрессии также исследовались в качестве возможных предикторов, однако результаты их прогнозирования существенно уступали тем, которые давала нейронная сеть.

Ученые пришли к выводу, что структура нейронной сети, использованная в исследовании, оказалась эффективным и надежным предиктором преэклампсии на основе использованного набора данных, и признали, что дальнейшее тестирование на более крупном и разнообразном наборе данных будет полезно для повышения точности и возможностей обобщения прогностической модели. Кроме того, была подчеркнута важность рассмотрения реальных, а не смоделированных данных для будущих исследований.

Эта же ученая группа занялась исследованием влияния курения и употребления алкоголя на развитие преэклампсии [7]. Исследование строилось на тех же данных, что и предыдущее, но в качестве входных параметров были добавлены вышеуказанные.

В исследовании изучалось влияние информации о курении и употреблении алкоголя на модель прогнозирования. Было обнаружено, что включение обеих частей информации значительно улучшило точность прогноза. Когда модели не предоставлялась информация о курении и алкоголе, вероятность правильного прогноза падала до 40 %.

Группа под руководством E. Purwanti предложила автоматизированную систему обнаружения преэклампсии на основе искусственных нейронных сетей [8]. На вход системы подавалось 11 параметров в виде факторов риска преэклампсии.

Набор данных включал 100 историй болезней пациентов, включая 70 медицинских карт пациентов с преэклампсией и 30 медицинских карт пациентов без преэклампсии. Каждая запись содержала информацию о различных факторах риска, таких как возраст

матери, индекс массы тела до и во время беременности, беременность в анамнезе, преэклампсия/эклампсия в анамнезе, аборт в анамнезе, АГ в анамнезе, анамнез сахарного диабета, семейный анамнез гипертензии, семейный анамнез сахарного диабета, беременность в паре.

В исследовании для классификации использовалась искусственная нейронная сеть с обратным распространением. В конструкции системы использовалось 11 нейронов во входном слое, представляющих 11 факторов риска, и 1 нейрон в выходном слое, указывающий на риск преэклампсии или отсутствие риска.

В исследовании оценивалась производительность искусственной нейронной сети обратного распространения с различными параметрами, такими как количество нейронов в скрытом слое, максимальная эпоха и скорость обучения. Результаты были получены путем перекрестной проверки, чтобы уменьшить переоснащение и определить наилучшую комбинацию параметров.

Лучшая модель достигла уровня точности обучения 100 % и смогла отличить пациентов с риском преэклампсии от пациентов без него. Точность тестирования также достигла 100 % в лучшем сценарии, что указывает на то, что модель обратного распространения хорошо показала себя при распознавании новых шаблонов.

Кроме того, приемочный тест пользователя для раннего выявления преэклампсии показал положительный отклик, а аспекты разработки программного обеспечения, функциональные аспекты и аспекты визуальной коммуникации получили «хорошие» отзывы.

В целом исследование показало, что разработанная автоматизированная система выявления преэклампсии на основе искусственных нейронных сетей обладает высокой точностью в определении риска преэклампсии у беременных, что может дать важную информацию для профилактических мероприятий.

R. R. Janghel и другие разработали систему поддержки принятия решения врача для прогнозирования родов с использованием параметров, которые являются значимыми при определении патологий беременности, для того чтобы принять решение о методе родовспоможения [9]. В исследовании используются глубокие нейронные сети в качестве метода мягких вычислений.

В исследовании используется набор данных, состоящий из параметров, зарегистрированных во время родов. Этот набор содержит случаи как нормальных, так и хирургических родов. Три модели DNN обучаются с использованием разных алгоритмов: алгоритм обратного распространения ошибки (Back propagation algorithm, BPA), радиальная базисная функциональная сеть (RBFN)<sup>7</sup>, обучающая сеть векторного квантования (LVQN) [10].

Параметры из набора данных используются для обучения и тестирования каждой модели DNN. Модели оптимизируются путем корректировки их конструктивных факторов для достижения наивысшего балла распознавания.

В исследовании сообщается о следующих показателях точности оптимизированных моделей глубоких нейронных сетей: BPA — точность 93,75 %, RBFN — точность 99,00 %, LVQN — точность 87,5 %. На основании результатов исследования RBFN определена как лучшая сеть для задачи прогнозирования родов плода.

Группа ученых из Таиланда разработала прогностическую модель для оценки риска гипертонии, вызванной беременностью, осложнения, которое может иметь серьезные последствия для беременных женщин и плода [11].

---

<sup>7</sup> RBFN (Radial Basis Function Network, сеть RBF) — двухслойная сеть без обратных связей, содержащая скрытый слой радиально симметричных нейронов.

Исследователи использовали общедоступный набор данных от Logan (2020), который содержит информацию из исследования случай-контроль о 83 случаях преэклампсии и 5 случаях эклампсии среди 352 беременных женщин. Для анализа были выбраны 17 параметров, включая возраст матери, историю беременности, семейный анамнез диабета и гипертонии, родовые посещения, метод родоразрешения, беременность, многоплодную беременность, паритет, провинцию, место жительства, употребление алкоголя и табака.

Набор данных показал дисбаланс между здоровыми беременными женщинами (75 %) и женщинами с преэклампсией и эклампсией (25 %). Чтобы решить эту проблему, для балансировки данных был использован алгоритм Synthetic Minority Over-sampling Technique (SMOTE) [12].

Оценка риска гипертонии, вызванной беременностью, проводилась с использованием семи алгоритмов машинного обучения: логистической регрессии,  $k$ -ближайших соседей, дерева решений, случайного леса, многослойной нейронной сети, метода опорных векторов, наивного байесовского алгоритма.

Среди протестированных алгоритмов машинного обучения случайный лес достиг наивысшей точности 89,62 % для прогнозирования риска гипертонии, вызванной беременностью. Эта точность превзошла другие протестированные алгоритмы. В ходе исследования была успешно разработана прогностическая модель для оценки риска гипертонии, вызванной беременностью, с использованием машинного обучения.

Исследователи R. Chinnaiyan и S. Alex изучили в целом аномалии плода и алгоритмы машинного обучения для диагностики и прогнозирования [13]. Предлагаемое исследование направлено на разработку модели прогнозирующего классификатора здоровья плода с использованием алгоритмов машинного обучения.

Пробел в исследованиях указывает на необходимость более простых протоколов сканирования для выявления аномалий из-за нехватки квалифицированных специалистов по УЗИ в развивающихся странах, таких как Индия. Подходы машинного обучения считаются важными при разработке системы поддержки принятия медицинских решений для раннего прогнозирования аномалий плода. В настоящее время ученые считают, что самую высокую точность можно получить, используя методы глубокого обучения.

Исследователи S. Anbu и V. Sarmah описали модель машинного обучения, которая участвовала в конкурсе Microsoft Women Health Risk Assessment [14].

Подход включает в себя очистку данных, исследовательский анализ данных (exploratory data analysis, EDA<sup>8</sup>), проектирование признаков и построение модели. Модель решает задачу многоклассовой классификации женщин: по различным сегментам и подгруппам риска для здоровья на основе различных атрибутов, связанных с сексуальным здоровьем, образом жизни, культурой и другими факторами.

Авторы публикации использовали алгоритм XGBoost [15] (популярный алгоритм машинного обучения для повышения градиента) и добились точности 87,002 %.

Исследование D. K. Sreekantha и коллег было направлено на разработку экспертной системы с использованием алгоритмов машинного обучения для прогнозирования трудностей при интубации и создания системы распределения, которая назначает опытного анестезиолога для сложных случаев [16]. Они собрали репрезентативный набор данных из больницы и смоделировали большие наборы данных с помощью кода Python под руководством анестезиологов и медицинских преподавателей.

<sup>8</sup> Анализ основных свойств dataset, нахождение в нем общих закономерностей, распределений и аномалий.

Набор данных включал 19 характеристик пациента, таких как возраст, пол, индекс массы тела, балл Маллампати<sup>9</sup>, межрезцовый промежуток, движения головы и шеи, тироментальное расстояние и тест прикуса верхней губы (ULBT). Данные были разделены на легкие, сложные и невозможные случаи интубации на основе экспертной оценки анестезиолога.

Несколько алгоритмов машинного обучения были реализованы и протестированы на наборе данных, включая логистическую регрессию, дерево решений, случайный лес, метод опорных векторов и  $k$ -ближайших соседей. Производительность каждого алгоритма оценивалась с точки зрения общей точности и точности перекрестной проверки. Результаты показали следующее: точность логистической регрессии — 90,35 %, точность при дереве решений — 95,67 %, точность случайного леса — 77,34 %, точность метода опорных векторов — 77,34 %, самая высокая точность по  $k$ -ближайшим соседям — 96,17 %. По результатам алгоритм дерева решений был признан наиболее выгодным, обеспечивающим наибольшую точность при меньших вычислительных затратах.

Исследование J. F. Carreño и P. Qiu было сосредоточено на прогнозировании преэклампсии на основе данных протеомики, полученных от беременных женщин [17]. Цель исследования состояла в том, чтобы изучить, как методы уменьшения размерности и методы наложения временных рядов могут быть полезны для прогнозирования преэклампсии с использованием данных протеомики.

В указанной статье использовались два независимых набора данных из разных групп пациентов. Один набор данных, состоящий из 166 пациентов, был из Детройта, а другой набор данных, состоящий из 36 пациентов, был из Стэнфорда. Каждый набор данных включал протеомную информацию для каждой пациентки в несколько моментов времени во время беременности. Временные точки, выбранные для каждого пациента, не были одинаковыми, у разных пациентов было разное количество временных точек.

В исследовании изучались два метода уменьшения размерности. Импералистический конкурентный алгоритм (ICA)<sup>10</sup> использовался в качестве метода выбора параметров для выбора соответствующих признаков для целей классификации. Алгоритм Sample Progression Discovery (SPD) [18] на основе комбинации модифицированного метода  $k$ -средних и минимального остового дерева применялся для кластеризации признаков, которая группирует белки с высокой корреляцией в кластеры.

В исследовании рассматривались два метода наложения временных рядов для протеомных данных. Простое общее среднее: несколько моментов времени для каждого пациента усреднялись, чтобы суммировать данные в один вектор, представляющий среднее значение экспрессии белка для этого пациента. Суммирование по трем точкам: отметки по времени для каждой пациентки были разделены на три подгруппы, соответствующие трем триместрам беременности. Было рассчитано итоговое значение экспрессии белка для каждого триместра, чтобы суммировать данные по 3 временным точкам.

Для прогнозирования преэклампсии использовались два алгоритма классификации: случайный лес и метод опорных векторов. Исследование показало, что все четыре пайплайна анализа (Independent component analysis (ICA) с общим средним значением, ICA с итоговым значением по трем точкам, SPD с общим средним значением и SPD с итоговым значением по трем точкам) достигли одинаковой точности прогнозирования в диапазоне от 85 % до 93 %. Это указывало на то, что как методы уменьшения размерности,

<sup>9</sup> Балл Маллампати — известный метод прогнозирования трудной интубации трахеи.

<sup>10</sup> Семейство вычислительных методов, используемых для решения задач оптимизации различных типов, не требует градиента функции в процессе оптимизации.

так и методы наложения временных рядов сохраняют полезную информацию для прогнозирования преэклампсии.

Кроме того, при анализе данных каждого триместра отдельно исследование показало, что первые два триместра предоставили более информативные данные для прогнозирования преэклампсии, чем последний триместр. Кроме того, метод суммирования, основанный на трехточечных средних значениях, превзошел классификационный анализ, основанный на отдельных триместрах, демонстрируя эффективность интеграции информации из нескольких триместров.

В целом исследование продемонстрировало потенциал использования протеомных данных и алгоритмов машинного обучения для прогнозирования преэклампсии и подчеркнуло важность учета данных из разных триместров для раннего прогнозирования.

Группой исследователей из Пакистана, Южной Кореи и Саудовской Аравии [19] разработана система прогнозирования рисков осложнений беременности. Предложена новая архитектура глубокой нейронной сети DT-BiLTCN на основе деревьев решений, двунаправленной сети долговременной памяти (Bidirectional Long short-term memory, BiLSTM) [20] и временной сверточной сети (Temporal Convolutional Network, TCN) [21]. Были проведены эксперименты на наборе данных из 1218 образцов, собранных в службах охраны материнства, больницах и поликлиниках. Дисбаланс классов был устранен с помощью увеличения числа примеров миноритарного класса (Synthetic Minority Over-sampling Technique, SMOTE).

Также следует упомянуть медицинский чат-бот Med-PaLM 2 [22]. Этот бот с искусственным интеллектом был разработан в 2023 г. компанией Google для медицинских консультаций, обобщения анализов, записей, рецептов и т. п. Нейросеть была обучена на корпусе различных медицинских данных. Однако Med-PaLM 2 пока имеет те же проблемы, что и другие чат-боты: иногда отвечает неточно, упускает важные детали. Согласно последнему исследованию Google, модель пока еще отвечает на вопросы хуже, чем обычные врачи. Тем не менее бот уже хорошо показывает себя в аргументации, понимании проблемы и принятии взвешенных решений.

### 3. АНАЛИТИЧЕСКИЕ ОБЗОРЫ

В работе Sr. R. Surendiran и др. [23] проведен обзор методов ML для прогнозирования преждевременных родов. Преждевременные роды — распространенное осложнение беременности — являются причиной 35 % из 3,1 млн. случаев смерти, связанных с беременностью, ежегодно страдает около 15 млн. детей.

В большинстве исследований этой проблемы до настоящего времени использовались данные электронных медицинских карт, электромиографии матки и электрогистерографии. В данном систематическом обзоре рассмотрены подходы на основе данных трансвагинального ультразвукового сканирования, а также с помощью ультразвуковых и магнитно-резонансных изображений.

Также задаче прогнозирования преждевременных родов с использованием алгоритмов машинного обучения посвящен систематический обзор литературы команды ученых Калифорнийского университета и университета Турку под руководством Z. Sharifi-Heris [24]. Был проведен комплексный поиск источников в 7 библиографических базах данных до 15 мая 2021 года. Оценивалось качество исследований, извлекалась и приводилась описательная информация, включая описательные характеристики данных, процессы ML-моделирования и их эффективность.



В работе [25] группа чилийских исследователей анализировала классические методы машинного обучения для прогнозирования перинатальных осложнений. Рассматривались в основном электронные медицинские карты (48 %), медицинские изображения (29 %) и биологические маркеры (19 %), а также другие типы признаков, такие как датчики и частота сердечных сокращений плода (4 %). В качестве основных перинатальных осложнений были выделены преэклампсия и недоношенность. В качестве метрики точности использовалась AUC. Для прогнозирования недоношенности наилучшие результаты показал метод опорных векторов (точность 95,7 %), для прогнозирования неонатальной смертности — метод XBoost (точность 99,7 %).

#### 4. ЗАКЛЮЧЕНИЕ

В настоящее время всё больше экспертных систем для оценки риска развития патологий беременности в основе своей имеют алгоритмы машинного обучения [26, 27].

К сожалению, в настоящий момент существует множество препятствий для массового внедрения нейронных сетей в области практического здравоохранения. Наиболее существенные — это ограниченность и большая вариативность (зашумленность) наборов данных, на которых производилось обучение моделей, что приводит к некоторой предвзятости их решений, недостаточная для реальной клинической практики точность работы моделей, а также проблемы конфиденциальности и этики.

Рассмотрим эти проблемы более подробно. Разметку медицинских данных Data Science специалистам производить очень сложно, потому что у подавляющего большинства отсутствует медицинское образование. Кроме того, на вариативность данных влияют:

- различное оборудование, используемое для получения данных, разные настройки и неисправности этого оборудования;
- отсутствие единых протоколов осмотра;
- ошибки в работе лаборантов.

Однако и при разметке врачами возможно появление шумов в разметке:

- каждый врач может разметить одну и ту же единицу данных по-разному — в зависимости от своего опыта и образования;
- из-за жестких требований к обезличенности данных врачи-разметчики чаще всего интерпретируют исследования пациентов, не зная их истории болезни, иногда даже возраста и пола;
- при вводе медицинских данных существует множество неформальных соглашений;
- врач может разметить одно и то же исследование по-разному, например, вследствие усталости или качества изображения монитора.

К сожалению, примеры сбора и подготовки обезличенных медицинских данных в России пока весьма немногочисленны, не отвечают потребностям разработчиков нейронных сетей и, как правило, содержат относительно небольшое количество записей в наборе — от десятков до несколько тысяч пациентов. Кроме того, использование методов анонимизации<sup>11</sup> и псевдоанонимизации требует внесения изменений в текущее законодательство [28].

---

<sup>11</sup> Анонимизация — способ обработки идентификационной информации так, чтобы по полученным данным нельзя было определить, к какому именно лицу они относятся.

Вторая острая проблема — этическая. При разработке и внедрении систем поддержки принятия решений возникают весьма острые вопросы, например:

- можно ли использовать обезличенные медицинские данные для разработок и исследований без согласия пациента?
- кто будет отвечать за неизбежные ошибки в работе алгоритмов машинного обучения — врач, пациент или разработчик?

Также необходимо принимать во внимание, что страховые компании могут использовать предсказательные системы для отказа в страховании пациентам с более высоким риском.

Вместе с тем, внедрение ИИ показало хорошие результаты при оценке профилактического здоровья пациента [29].

В Российской Федерации развитие методов медицинской диагностики на основе машинного обучения обусловлено указом Президента РФ от 10.10.2019 № 490 «О развитии искусственного интеллекта в Российской Федерации» (вместе с «Национальной стратегией развития искусственного интеллекта на период до 2030 года») [30]. Кроме того, за последние несколько лет в России в силу вступили стандарты, устанавливающие общие требования к системам искусственного интеллекта в медицине, например ГОСТ Р 59921.8-2022, ГОСТ Р 59921.9-2022.

### Список литературы

1. *Werbos P. J.* Beyond regression: New tools for prediction and analysis in the behavioral sciences / Ph. D. thesis, Harvard University. Cambridge, MA, 1974.
2. *Maylawati D. S., Ramdhani M. A., Zulfikar W. B., Taufik I., Darmalaksana W.* Expert system for predicting the early pregnancy with disorders using artificial neural network // Proc. of 2017 5th Int. Conf. on Cyber and IT Service Management (CITSM). 8–10 Aug. 2017. Denpasar, Indonesia, 2017. P. 1–6. doi:10.1109/CITSM.2017.8089243
3. *Raghav H. K. V. S., Devi S. P., Rengaraj N., Thanranikumar E.* Prediction of Preterm Pregnancies using Soft Computing Techniques Neural Networks and Gradient Descent Optimizer // Proc. 2018 Int. Conf. on Computer Communication and Informatics (ICCCI). 4–6 Jan. 2018. Coimbatore, India, 2018. P. 1–4. doi:10.1109/ICCCI.2018.8441432
4. *Ruder S.* An Overview of Gradient Descent Optimization Algorithms // arXiv:1609.04747v2[cs.LG], 2017.
5. *Tahir M., Badriyah T., Syarif I.* Neural Networks Algorithm to Inquire Previous Preeclampsia Factors in Women with Chronic Hypertension During Pregnancy in Childbirth Process // Proc. 2018 Int. Electronics Symposium on Knowledge Creation and Intelligent Computing (IES-KCIC). 23–30 Oct. 2018. Bali, Indonesia, 2018. P. 51–55. doi:10.1109/KCIC.2018.8628588
6. *Neocleous C. K., Anastasopoulos P., Nikolaidis K. H., Schizas C. N., Neokleous K. C.* Neural networks to estimate the risk for preeclampsia occurrence // Proc. 2009 Int. Joint Conf. on Neural Networks. 14–19 Jun. 2009. Atlanta, GA, USA, 2009. P. 2221–2225. doi: 10.1109/IJCNN.2009.5178820
7. *Neocleous C., Nikolaidis K., Neokleous K., Schizas C.* Neural networks to investigate the effects of smoking and alcohol abuse on the risk for preeclampsia // Proc. 2009 9th International Conference on Information Technology and Applications in Biomedicine. 5–7 Nov. 2009. P. 1–5. Larnaka, Cyprus, 2009. doi:10.1109/ITAB.2009.5394421
8. *Purwanti E., Preswari I. S., Ernawati E.* Early Risk Detection of Pre-eclampsia for Pregnant Women Using Artificial Neural Network // International Journal of Online and Biomedical Engineering (iJOE). 2019. Vol. 15, № 02. P. 71–71. doi:10.3991/ijoe.v15i02.9680
9. *Janghel R. R., Shukla A., Tiwari R., Tiwari P.* Clinical Decision Support System for Fetal Delivery Using Artificial Neural Network // Proc. of 2009 International Conference on New Trends in Information and Service Science. 30 Jun – 2 Jul 2009. Beijing, China, 2009. P. 1070–1075. doi:10.1109/niss.2009.66

10. *Kohonen T.* Learning vector quantization // M. A. Arbib ed., *The Handbook of Brain Theory and Neural Networks*. Cambridge, MA, USA: MIT Press, 1995. P. 537–540.
11. *Wanriko S., Hnoohom N., Wongpatikaseree K., Jitpattanukul A., Musigavong O.* Risk Assessment of Pregnancy-induced Hypertension Using a Machine Learning Approach // Proc. of 2021 Joint International Conference on Digital Arts, Media and Technology with ECTI Northern Section Conference on Electrical, Electronics, Computer and Telecommunication Engineering. 03–06 Mar 2021. P. 233–237. Cha-am, Thailand, 2021; doi:10.1109/ECTIDAMTNCON51128.2021.9425764
12. *Chawla N. V., Bowyer K. W., Hall L. O., Kegelmeyer W. P.* SMOTE: Synthetic Minority Over-sampling Technique // *Journal of Artificial Intelligence Research*. 2002. Vol. 16. P. 321–357. doi:10.1613/jair.953.
13. *Chinnaiyan R., Alex S.* Machine Learning Approaches for Early Diagnosis and Prediction of Fetal Abnormalities // Proc. 2021 Int. Conf. on Computer Communication and Informatics (ICCCI). 27–29 Jan 2021. Coimbatore, India, 2021. P. 1–3. doi:10.1109/ICCCI50826.2021.9402317
14. *Anbu S., Sarmah B.* Machine learning approach for predicting womens health risk // Machine learning approach for predicting womens health risk // Proc. of 2017 4th International Conference on Advanced Computing and Communication Systems (ICACCS). 6–7 Jan 2017. Coimbatore, India, 2017. P. 1–4. doi:10.1109/ICACCS.2017.8014588
15. *Chen T., Guestrin C.* XGBoost: A Scalable Tree Boosting System // Proc. of the 22nd Int. Conf. on Knowledge Discovery and Data Mining. 13–17 Aug 2016. San Francisco California, USA, 2016. P. 785–794. doi:10.1145/2939672.2939785
16. *Sreekantha D. K. et al.* Prediction of difficulties in Intubation using an Expert system // Proc. 2019 IEEE Int. Conf. on Distributed Computing, VLSI, Electrical Circuits and Robotics (DISCOVER). 11–12 Aug 2019. Manipal, India, 2019. P. 1–7. doi:10.1109/DISCOVER47552.2019.9007952
17. *Carreño J. F., Qiu P.* Feature selection algorithms for predicting preeclampsia: A comparative approach // Proc. of 2020 IEEE International Conference on Bioinformatics and Biomedicine (BIBM). 16–19 Dec 2020. Seoul, Korea (South), 2020. P. 2626–2631. doi:10.1109/BIBM49941.2020.9313108
18. *Qiu P., Gentles A. J., Plevritis S. K.* Discovering Biological Progression Underlying Microarray Samples // *PLoS Computational Biology*. 2011. Vol. 7, № 4. P. e1001123. doi:10.1371/journal.pcbi.1001123
19. *Raza A., Siddiqui H. U. R., Munir K., Almutairi M., Rustam F., Ashraf I.* Ensemble learning-based feature engineering to analyze maternal health during pregnancy and health risk prediction // *PLOS ONE*. 2022. Vol. 17, № 11. P. e0276525. doi:10.1371/journal.pone.0276525
20. *He K., Zhang X., Ren S., Sun J.* Delving Deep into Rectifiers: Surpassing Human-Level Performance on ImageNet Classification // Proc. of 2015 IEEE International Conference on Computer Vision (ICCV). 7–13 Dec 2015. P. 1026–1034. Santiago, Chile, 2015. doi:10.1109/iccv.2015.123 1026–1034
21. *Lea C. et al.* Temporal convolutional networks: A unified approach to action segmentation // Proc. 14th European Conference Computer Vision — ECCV 2016. 11–14 Oct. 2016. Amsterdam, The Netherlands, 2016. P. 47–54. doi:10.1007/978-3-319-49409-8\_7
22. *Singhal K. et al.* Towards Expert-Level Medical Question Answering with Large Language Models // arXiv:2305.09617[cs.LG], 2023.
23. *Surendiran Dr. R., Aarathi R., Thangamani M., Sugavanam S., Sarumathy R.* A Systematic Review using Machine Learning Algorithms for Predicting Preterm Birth // *International Journal of Engineering Trends and Technology*. 2022. Vol. 70, № 5. P. 46–59. doi:10.14445/22315381/IJETT-V70I5P207
24. *Sharifi-Heris Z., Laitala J., Arola A., Rahmani A. M., Bender M.* Machine Learning Approach for Preterm Birth Prediction Using Health Records: Systematic Review // *JMIR Med Inform.* 2022. Vol. 10, № 4. P. e33875. doi:10.2196/33875
25. *Bertini A., Salas R., Chabert S., Sobrevia L., Pardo F.* Using Machine Learning to Predict Complications in Pregnancy: A Systematic Review // *Front. Bioeng. Biotechnol.* 2012. Vol. 9. P. 780389. doi:10.3389/fbioe.2021.780389
26. *Damaraji G. M., Permanasari A. E., Hidayah I.* A Review of Expert System for Identification Various Risk in Pregnancy // Proc. of 3rd International Conference on Information and Communications Technology (ICOIACT). 24–26 Nov 2020. Yogyakarta, Indonesia, 2020. P. 99–104. doi:10.1109/ICOIACT 50329.2020.9332003
27. *Ramla M., Sangeetha S., Nickolas S.* Machine Learning for High Risk Pregnancies Pre-Term Birth Prediction: A Retrospective // *International Journal of Engineering and Technology(UAE)*. 2018. Vol.

- 7, № 2.22. P. 5–8. doi:10.14419/ijet.v7i2.22.11799.
28. Столбов А. П. Обезличивание персональных данных в здравоохранении // Врач и информационные технологии. 2017. № 3. С. 25–36.
29. Курдюмов Д. А., Кашин А. В., Рябов Н. Ю., Новицкий Р. Э., Гусев А. В. Опыт применения технологий искусственного интеллекта для развития профилактического здравоохранения на примере Кировской области // Менеджер здравоохранения. 2023. № 6. С. 62–69. doi:10.21045/1811-0185-2023-6-62-69
30. Указ Президента Российской Федерации от 10.10.2019 № 490 «О развитии искусственного интеллекта в Российской Федерации» [Электронный ресурс] <http://publication.pravo.gov.ru/Document/View/0001201910110003> (дата обращения: 16.11.2023).

Поступила в редакцию 23.09.2023, окончательный вариант — 16.11.2023.

**Рипка Дарина Станиславовна**, аспирант кафедры алгоритмической математики СПбГЭТУ «ЛЭТИ» им. В. И. Ульянова (Ленина), [dar.stkr@gmail.com](mailto:dar.stkr@gmail.com)

**Рыбин Сергей Витальевич**, кандидат физико-математических наук, доцент кафедры алгоритмической математики СПбГЭТУ «ЛЭТИ» им. В. И. Ульянова (Ленина),  
✉ [rsvvm2leti@gmail.com](mailto:rsvvm2leti@gmail.com)

---

Computer tools in education, 2023

№ 4: 50–64

<http://cte.eltech.ru>

doi:10.32603/2071-2340-2023-4-50-64

## **A Review of Expert Systems For Detecting Pregnancy Abnormalities Using Machine Learning Technologies**

Ripka D. S.<sup>1</sup>, Postgraduate, [dar.stkr@gmail.com](mailto:dar.stkr@gmail.com)  
Rybin S. V.<sup>1</sup>, Cand. Sc., Associate Professor, ✉ [rsvvm2leti@gmail.com](mailto:rsvvm2leti@gmail.com),  
[orcid.org/0000-0002-9095-3168](https://orcid.org/0000-0002-9095-3168)

<sup>1</sup>Saint Petersburg Electrotechnical University,  
5, building 3, st. Professora Popova, 197022, Saint Petersburg, Russia

### **Abstract**

The field of medical diagnosis has made significant progress with the introduction of machine learning techniques. This article presents a comprehensive review of studies in which researchers have developed expert systems for diagnosing pregnancy abnormalities using various machine learning techniques. Emphasizing the crucial role of data for learning and its preprocessing, the publications reviewed in the article compare the accuracy of different machine learning algorithms in this area. The analyzed research has primarily focused on building expert systems to diagnose pregnancy-related pathologies and complications in order to improve outcomes for expectant mothers and their future children. Using a wide range of machine learning approaches, including decision trees, support vector machines, random forest, artificial neural networks, and others, the effectiveness of each algorithm reviewed in this article in accurately predicting pregnancy-related problems was investigated.

One of the key points of this review is the emphasis on the quality and variety of training data. Particular attention was paid to the reliability and completeness of the datasets in analyzing the publications to allow machine learning algorithms to achieve higher diagnostic accuracy. Studies were selected for the review that used parameters recognized by the medical community as indicators of various pathologies for training.

An important criterion for selecting the review publications was the availability of data preprocessing in them to correct missing values, noise and class imbalance. These tasks play an essential role in improving the performance of the developed expert systems. By critically evaluating the methodologies and algorithms used in the reviewed publications, this paper provides valuable information for future research and development in this area.

**Keywords:** *diagnostics, birth pathologies, expert systems, machine learning.*

**Citation:** D. S. Ripka and S. V. Rybin, "A Review of Expert Systems For Detecting Pregnancy Abnormalities Using Machine Learning Technologies," *Computer tools in education*, no. 4, pp. 50–64, 2023 (in Russian); doi:10.32603/2071-2340-2023-4-50-64

### Список литературы

1. P. J. Werbos, "Beyond regression: New tools for prediction and analysis in the behavioral sciences," Ph. D. thesis, Harvard University, Cambridge, MA, 1974.
2. D. S. Maylawati, M. A. Ramdhani, W. B. Zulfikar, I. Taufik, and W. Darmalaksana, "Expert system for predicting the early pregnancy with disorders using artificial neural network," in *Proc. of 2017 5th Int. Conf. on Cyber and IT Service Management (CITSM), Denpasar, Indonesia, 8–10 Aug. 2017*, pp. 1–6, 2017; doi:10.1109/CITSM.2017.8089243
3. H. K. V. S. Raghav, S. P. Devi, N. Rengaraj, and E. Thanranikumar, "Prediction of Preterm Pregnancies using Soft Computing Techniques Neural Networks and Gradient Descent Optimizer," in *Proc. 2018 Int. Conf. on Computer Communication and Informatics (ICCCI), Coimbatore, India, 4–6 Jan. 2018*, pp. 1–4, 2018; doi:10.1109/ICCCI.2018.8441432
4. S. Ruder, "An Overview of Gradient Descent Optimization Algorithms," *arXiv:1609.04747v2[cs.LG]*, 2017.
5. M. Tahir, T. Badriyah, and I. Syarif, "Neural Networks Algorithm to Inquire Previous Preeclampsia Factors in Women with Chronic Hypertension During Pregnancy in Childbirth Process," in *Proc. 2018 Int. Electronics Symposium on Knowledge Creation and Intelligent Computing (IES-KCIC), Bali, Indonesia, 23–30 Oct. 2018*, pp. 51–55, 2018; doi:10.1109/KCIC.2018.8628588
6. C. K. Neocleous, P. Anastopoulos, K. H. Nikolaidis, C. N. Schizas, and K. C. Neocleous, "Neural networks to estimate the risk for preeclampsia occurrence," in *Proc. 2009 Int. Joint Conf. on Neural Networks, Atlanta, GA, USA, 14–19 Jun. 2009*, pp. 2221–2225, 2009; doi: 10.1109/IJCNN.2009.5178820
7. C. Neocleous, K. Nikolaidis, K. Neocleous, and C. Schizas, "Neural networks to investigate the effects of smoking and alcohol abuse on the risk for preeclampsia," in *Proc. 2009 9th International Conference on Information Technology and Applications in Biomedicine, Larnaka, Cyprus, 5–7 Nov. 2009*, pp. 1–5, 2009; doi:10.1109/ITAB.2009.5394421
8. E. Purwanti, I. S. Preswari, and E. Ernawati, "Early Risk Detection of Pre-eclampsia for Pregnant Women Using Artificial Neural Network," *International Journal of Online and Biomedical Engineering (iJOE)*, vol. 15, no. 02, pp. 71–71, 2019; doi:10.3991/ijoe.v15i02.9680
9. R. R. Janghel, A. Shukla, R. Tiwari, and P. Tiwari, "Clinical Decision Support System for Fetal Delivery Using Artificial Neural Network," in *Proc. of 2009 International Conference on New Trends in Information and Service Science, Beijing, China, 30 Jun – 2 Jul 2009*, pp. 1070–1075, 2009; doi:10.1109/niss.2009.66
10. T. Kohonen, "Learning vector quantization," in *M. A. Arbib ed., The Handbook of Brain Theory and Neural Networks*, Cambridge, MA, USA: MIT Press, pp. 537–540, 1995.
11. S. Wanriko, N. Hnoohom, K. Wongpatikaseree, A. Jitpattanakul, and O. Musigavong, "Risk Assessment of Pregnancy-induced Hypertension Using a Machine Learning Approach," in *Proc. of 2021 Joint International Conference on Digital Arts, Media and Technology with ECTI Northern Section Conference*

- on *Electrical, Electronics, Computer and Telecommunication Engineering, Cha-am, Thailand, 03–06 Mar 2021*, pp. 233–237, 2021; doi:10.1109/ECTIDAMTNCON51128.2021.9425764
12. N. V. Chawla, K. W. Bowyer, L. O. Hall, and W. P. Kegelmeyer, “SMOTE: Synthetic Minority Over-sampling Technique,” *Journal of Artificial Intelligence Research*, vol. 16, pp. 321–357, 2002; doi:10.1613/jair.953
  13. R. Chinnaiyan and S. Alex, “Machine Learning Approaches for Early Diagnosis and Prediction of Fetal Abnormalities,” in *Proc. 2021 Int. Conf. on Computer Communication and Informatics (ICCCI), Coimbatore, India, 27–29 Jan 2021*, pp. 1–3, 2021; doi:10.1109/ICCCI50826.2021.9402317
  14. S. Anbu and B. Sarmah, “Machine learning approach for predicting womens health risk,” in *Proc. of 2017 4th International Conference on Advanced Computing and Communication Systems (ICACCS), Coimbatore, India, 6–7 Jan 2017*, pp. 1–4, 2017; doi: 10.1109/ICACCS.2017.8014588
  15. T. Chen and C. Guestrin, “XGBoost: A Scalable Tree Boosting System,” in *Proc. of the 22nd Int. Conf. on Knowledge Discovery and Data Mining, San Francisco California, USA, 13–17 Aug 2016*, pp. 785–794, 2016; doi:10.1145/2939672.2939785
  16. D. K. Sreekantha et al., “Prediction of difficulties in Intubation using an Expert system,” in *Proc. 2019 IEEE Int. Conf. on Distributed Computing, VLSI, Electrical Circuits and Robotics (DISCOVER), Manipal, India, 11–12 Aug 2019*, pp. 1–7, 2019; doi:10.1109/DISCOVER47552.2019.9007952
  17. J. F. Carreño and P. Qiu, “Feature selection algorithms for predicting preeclampsia: A comparative approach,” in *Proc. of 2020 IEEE International Conference on Bioinformatics and Biomedicine (BIBM), Seoul, Korea (South), 16–19 Dec 2020*, pp. 2626–2631, 2020; doi:10.1109/BIBM49941.2020.9313108
  18. P. Qiu, A. J. Gentles, and S. K. Plevritis, “Discovering Biological Progression Underlying Microarray Samples,” *PLoS Computational Biology*, vol. 7, no. 4, p. e1001123, 2011; doi:10.1371/journal.pcbi.1001123
  19. A. Raza, H. U. R. Siddiqui, K. Munir, M. Almutairi, F. Rustam, and I. Ashraf, “Ensemble learning-based feature engineering to analyze maternal health during pregnancy and health risk prediction,” *PLOS ONE*, vol. 17, no. 11, p. e0276525, 2022; doi:10.1371/journal.pone.0276525
  20. K. He, X. Zhang, S. Ren, and J. Sun, “Delving Deep into Rectifiers: Surpassing Human-Level Performance on ImageNet Classification,” in *Proc. of 2015 IEEE International Conference on Computer Vision (ICCV), Santiago, Chile, 7–13 Dec 2015*, pp. 1026–1034, 2015; doi:10.1109/iccv.2015.123 1026–1034
  21. C. Lea, et al. “Temporal convolutional networks: A unified approach to action segmentation,” in *Proc. 14th European Conference Computer Vision – ECCV 2016, Amsterdam, The Netherlands, 11–14 Oct. 2016*, pp 47–54, 2016; doi:10.1007/978-3-319-49409-8\_7
  22. K. Singhal, et al., “Towards Expert-Level Medical Question Answering with Large Language Models,” *arXiv:2305.09617[cs.LG]*, 2023.
  23. Dr. R. Surendiran, R. Aarthi, M. Thangamani, S. Sugavanam, and R. Sarumathy, “A Systematic Review using Machine Learning Algorithms for Predicting Preterm Birth,” *International Journal of Engineering Trends and Technology*, vol. 70, no. 5, pp. 46–59, 2022; doi:10.14445/22315381/IJETT-V70 I5P207
  24. Z. Sharifi-Heris, J. Laitala, A. Arola, A. M. Rahmani, and M. Bender, “Machine Learning Approach for Preterm Birth Prediction Using Health Records: Systematic Review,” *JMIR Med Inform*, vol. 10, no. 4, p. e33875, 2022; doi:10.2196/33875
  25. A. Bertini, R. Salas, S. Chabert, L. Sobrevia, and F. Pardo, “Using Machine Learning to Predict Complications in Pregnancy: A Systematic Review,” *Front. Bioeng. Biotechnol*, vol. 9, p. 780389, 2012; doi: 10.3389/fbioe.2021.780389
  26. G. M. Damaraji, A. E. Permanasari, and I. Hidayah, “A Review of Expert System for Identification Various Risk in Pregnancy,” in *Proc. of 3rd International Conference on Information and Communications Technology (ICOIACT), Yogyakarta, Indonesia, 24–26 Nov 2020*, pp. 99–104, 2020; doi:10.1109/ICOIACT 50329.2020.9332003
  27. M. Ramla, S. Sangeetha, and S. Nickolas, “Machine Learning for High Risk Pregnancies Pre-Term Birth Prediction: A Retrospective,” *International Journal of Engineering & Technology*, vol. 7, no. 2.22, pp. 5–8, 2018; doi:10.14419/ijet.v7i2.22.11799
  28. A. P. Stolbov, “De-identification of personal data in health care,” *Medical doctor and information technologies*, no. 3, pp. 25–36, 2017 (in Russian).

29. D. A. Kurdyumov et al., “Experience in the application of artificial intelligence technologies for the development of preventive health care on the example of the Kirov region,” *Manager Zdravookhrameniya*, pp. 62–69, 2023 (in Russian); doi:10.21045/1811-0185-2023-6-62-69
30. “Ukaz Prezidenta Rossiiskoi Federatsii ot 10.10.2019 № 490 O razvitii iskusstvennogo intellekta v Rossiiskoi Federatsii,” in *Oftsil'noe opublikovanie pravovykh aktov*, 2019 (in Russian). [Online]. Available: <http://publication.pravo.gov.ru/Document/View/0001201910110003>

*Received 23-09-2023, the final version — 16-11-2023.*

**Darina Ripka, Postgraduate, Department of Algorithmic Mathematics, Saint Petersburg Electrotechnical University, [dar.stkr@gmail.com](mailto:dar.stkr@gmail.com)**

**Sergey Rybin, Candidate of Sciences (Phys.-Math.), Associate Professor of the Department of Algorithmic Mathematics, Saint Petersburg Electrotechnical University, ✉ [rsvvm2leti@gmail.com](mailto:rsvvm2leti@gmail.com)**