

## Нейронечеткая модель диагностики сахарного диабета

*А.С. Катасёв<sup>1,2</sup>, Д.З. Валиева<sup>1,3</sup>, Д.В. Катасёва<sup>1</sup>*

<sup>1</sup>*Казанский национальный исследовательский технический университет  
им. А.Н. Туполева-КАИ, Казань*

<sup>2</sup>*Казанский государственный энергетический университет, Казань*

<sup>3</sup>*ГБУ «Безопасность дорожного движения», Казань*

**Аннотация:** В статье решается задача диагностики сахарного диабета. Для ее решения предлагается построение нейронечеткой модели. Для обучения модели произведены поиск и подготовка исходных данных к анализу. Данные получены из общедоступного источника Kaggle. Подготовка данных к анализу произведена на базе аналитической платформы Deductor. Из подготовленного набора данных сформированы обучающая и тестовая выборки, использованные для построения модели. Сравнение полученных результатов с известными результатами других авторов позволило сделать вывод об адекватности модели и принципиальной возможности ее практического использования.

**Ключевые слова:** нейронечеткая модель, нечеткая нейронная сеть, сахарный диабет, моделирование, диагностика, машинное обучение.

### Введение

В настоящее время сахарный диабет является одной из наиболее актуальных медико-социальных проблем. Отмечается устойчивая тенденция к увеличению распространенности этого хронического заболевания во всем мире, что оказывает существенное воздействие на систему здравоохранения и экономику в целом. Сахарный диабет характеризуется высокой распространенностью и значительными затратами на лечение и профилактику осложнений. Данное заболевание ухудшает качество жизни пациентов, ограничивая их повседневную активность и функционирование.

Например, в 2023 году в России зарегистрировано 62,2 тысячи случаев временной нетрудоспособности населения, связанных с диагнозом сахарного диабета. Это составляет около 75% от общего числа граждан, временно утративших трудоспособность по причине проблем со здоровьем. Таким образом, сахарный диабет является одним из ключевых факторов, приводящих к временной утрате трудоспособности и, как следствие, к увеличению уровня безработицы [1].

Медицина сталкивается с проблемой позднего выявления сахарного диабета [2]. Своевременная диагностика данного заболевания играет ключевую роль в эффективном лечении и профилактике осложнений. Разработка эффективных методов диагностики сахарного диабета является важной задачей, требующей использования современных информационных технологий [3]. В работе для решения этой задачи предложено использование нейронечеткой модели [4], формирующей базу знаний, состоящей из набора нечетко-продукционных правил. Такой подход, базируясь на нечеткой логике и алгоритме нечеткого логического вывода, позволяет получать интерпретируемый результат [5], обеспечивая гибкость и адаптивность при анализе медицинских данных.

Для построения нейронечеткой модели диагностики сахарного диабета потребовалось произвести сбор и предварительную обработку исходных данных. Рассмотрим этапы и результат решения этой задачи.

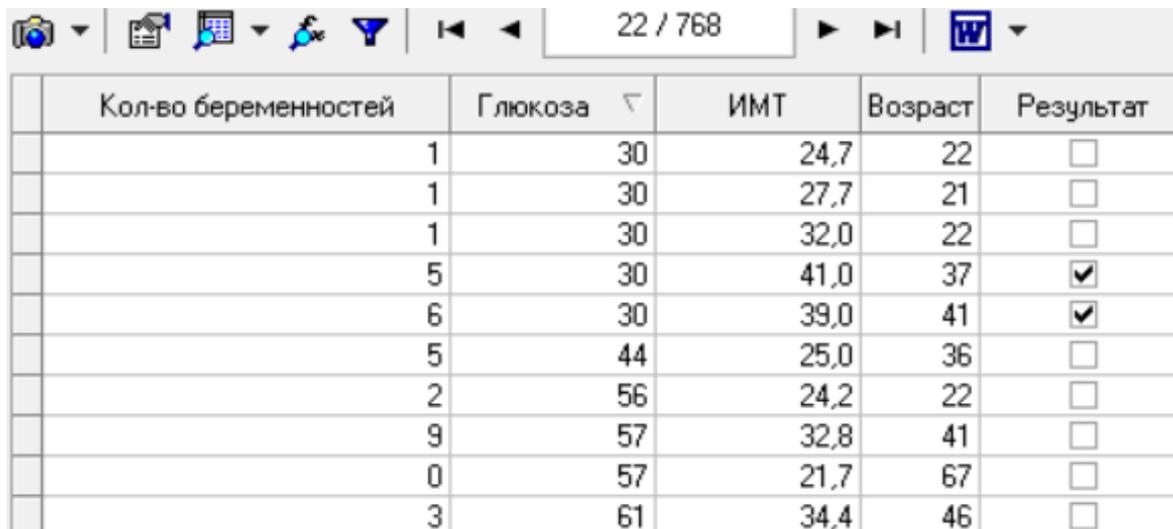
### **Получение и подготовка данных к анализу**

Первичным этапом построения моделей машинного обучения, к которым относится нейронечеткая модель, является получение исходных данных. В работе поиск данных для построения модели диагностики сахарного диабета осуществлялся по следующим открытым репозиториям и поисковым системам [6, 7]: Kaggle, UCI Machine Learning Repository, Amazon Web Services Registry of Open Data, Google Dataset Search, Harvard Dataverse и Microsoft Research Open Data. Набор данных о диабете найден на платформе Kaggle [8]. Набор содержит 768 строк и 9 столбцов: 8 входных («число беременностей», «концентрация глюкозы в плазме крови», «уровень инсулина», «артериальное давление», «индекс массы тела», «толщина кожно-жировой складки», «функция родословной диабета», «возраст») и 1 выходной, указывающий на наличие (1) или отсутствие (0) заболевания.

---

В программе Deductor произведены оценка качества и подготовка исходных данных к анализу. После импорта данных с помощью встроенных инструментов Deductor проведена предварительная оценка качества данных. Выявлено, что необходима дальнейшая очистка и подготовка данных, в частности, обработка выбросов и экстремальных значений. Следующим шагом стала проверка данных на наличие дубликатов и противоречий. Далее проведен корреляционный анализ, который определил степень влияния каждого входного столбца на выходной. В результате анализа столбцы «толщина кожи» (корреляция 0,065), «артериальное давление» (0,114), «инсулин» (0,125) и «функция родословной диабета» (0,184) были исключены из выборки из-за их низкой коррелированности с выходом.

В результате предобработки получена выборка с 4 входными параметрами и 1 выходным. Фрагмент выборки представлен на рисунке 1.



Кол-во беременностей	Глюкоза	ИМТ	Возраст	Результат
1	30	24,7	22	<input type="checkbox"/>
1	30	27,7	21	<input type="checkbox"/>
1	30	32,0	22	<input type="checkbox"/>
5	30	41,0	37	<input checked="" type="checkbox"/>
6	30	39,0	41	<input checked="" type="checkbox"/>
5	44	25,0	36	<input type="checkbox"/>
2	56	24,2	22	<input type="checkbox"/>
9	57	32,8	41	<input type="checkbox"/>
0	57	21,7	67	<input type="checkbox"/>
3	61	34,4	46	<input type="checkbox"/>

Рисунок 1. – Фрагмент итоговой выборки данных для анализа

На основе подготовленных данных построена нейронечеткая модель. Построение производилось в программном комплексе «Нейронечеткая система формирования нечетких моделей оценки дискретного состояния объектов» [9]. Рассмотрим этот этап моделирования более подробно.

## Построение нейронечеткой модели

Построение нейронечеткой модели начинается с загрузки данных для анализа. На рисунке 2 представлено окно загрузки с возможностью выбора числа нечетких градаций для каждого входного параметра.

■ Выберите столбцы с градацией

Выберите поля с градациями

Beremennosti    Glukoza    IMT    Vozrast

Выберите число градаций для каждого поля

5   4   4   5

	Beremennosti	Glukoza	IMT	Vozrast	Rezultat
▶ 1	6	148	33,6	50	1
2	1	85	26,6	31	0
3	8	183	23,3	32	1
4	1	89	28,1	21	0
5	0	137	43,1	33	1
6	5	116	25,6	30	0
7	3	78	31	26	1
8	10	115	35,3	29	0
9	2	197	30,5	53	1
10	8	125	12,2	54	1
11	4	110	37,6	30	0
12	10	160	30	34	1

Далее  Показать таблицу

Рисунок 2. – Окно загрузки данных для анализа

Оптимальное число градаций выбирается экспертным путем или подбирается экспериментально.

На основе загруженных данных для последующего построения модели формируются обучающая и тестовая выборки. Для этого используется метод группового случайного сэмпинга с замещением, обеспечивающий однородность формируемых выборок [4].

На основе сформированных выборок произведено обучение нечеткой нейронной сети. График обучения представлен на рисунке 3.

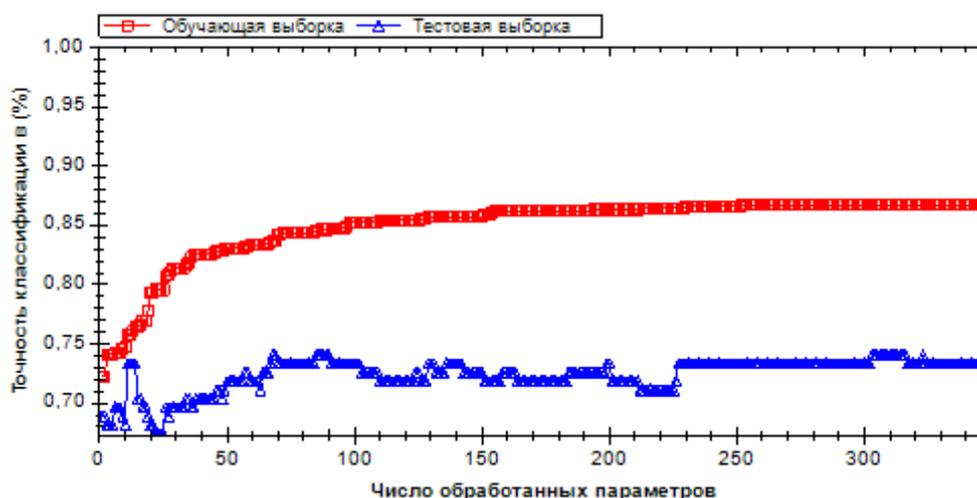


Рисунок 3. – График обучения нечеткой нейронной сети

Достигнутые результаты обучения представлены на рисунке 4.

Цикл обучения ННС	9
Параметр ФП / всего параметров	38 / 38
Эпоха ГА (настройка параметра)	11
Время обучения ННС (чч:мм:сс)	00:55:32
Точность классификации, % (train)	86,72
Точность классификации, % (test)	73,33

Рисунок 4. – Результаты обучения нечеткой нейронной сети

Нечеткая нейронная сеть прошла 9 циклов обучения, длившихся в общей сложности 55 минут 32 секунды. В каждом цикле производилась настройка 38 параметров функции принадлежности. Для обучающей выборки точность классификации составила 86,72%, для тестовой – 73,33%.

Следует отметить, что в различных научных работах исследуется применение методов машинного обучения для диагностики сахарного диабета [10]. Результаты этих исследований представлены в таблице 1.

Таблица № 1

Точность диагностики сахарного диабета в различных исследованиях

Наименование исследования	Метод машинного обучения	Достигнутая точность
Pima Indian Diabetes Database and Machine Learning Models for Diabetes Prediction	Логистическая регрессия	0,79
	Случайные леса	0,77
	Опорные вектора	0,75
Prediction of Pima Indians Diabetes with Machine Learning Algorithms	Логистическая регрессия	0,8
	Деревья решений	0,75
	Опорные вектора	0,78
A comparative study of diabetes detection using the Pima Indian diabetes database	Случайные леса	0,78

Сравнивая результаты обучения нечеткой нейронной сети с табличными данными можно сделать вывод, что точность нейронечеткой модели (на тестовых данных) немного уступает точности традиционных методов машинного обучения [11]. Однако принципиальная возможность использования нейронечеткой модели для диагностики сахарного диабета наряду с интерпретируемостью получаемых решений свидетельствуют о ее адекватности и возможности эффективного практического использования.

### Выводы

Проведенное исследование показало, что нейронечеткая модель является эффективным инструментом диагностики сахарного диабета. Она способна «объяснять» выходной результат за счет использования нечетких и лингвистических переменных. В дальнейшем модель может быть использована в составе интеллектуальной системы поддержки принятия решений, помогая специалистам решать поставленную задачу.

## Литература

1. Драгомирова Л.Г., Хайрутдинова З.Л., Шубин Л.Л. Динамика заболеваемости сахарным диабетом в Российской Федерации с 2001-2021 год // Modern Science. 2022. № 5-4. С. 59-62.
2. Балахонова Е.А. Гестационный сахарный диабет // Мирская наука. 2023. № 2 (71). С. 158-162.
3. Жолдас Н., Мансурова М.Э. Информационные технологии для мониторинга юных пациентов с сахарным диабетом // Автоматика и программная инженерия. 2021. № 1 (35). С. 11-20.
4. Катасёва Д.В., Баринаева А.О. Формирование баз знаний интеллектуальных систем на примере нейронечеткого анализа медицинских данных // Вестник Технологического университета. 2022. Т. 25. № 2. С. 67-70.
5. Нурматова Е.В., Зайцев Е.И. Моделирование оценки состояния объектов на основе запросов с нечеткими условиями // Инженерный вестник Дона, 2024, № 1. URL: [ivdon.ru/ru/magazine/archive/n1y2024/8936](http://ivdon.ru/ru/magazine/archive/n1y2024/8936)
6. Prior F., Smith K., Almeida J., Kathiravelu P., Kurc T., Saltz J., Fitzgerald T.J. Open access image repositories: high-quality data to enable machine learning research // Clinical Radiology. 2020. Vol. 75. No. 1. P. 7-12.
7. Ямашкин С.А., Ямашкина Е.О., Ямашкин А.А. Репозиторий нейросетевых моделей для анализа пространственных данных // Инженерный вестник Дона, 2022, № 12. URL: [ivdon.ru/ru/magazine/archive/n12y2022/8082](http://ivdon.ru/ru/magazine/archive/n12y2022/8082).
8. Kuroki M. Integrating data science into an econometrics course with a kaggle competition // Journal of Economic Education. 2023. Vol. 54. No. 4. P. 364-378.
9. Катасёв А.С., Дагаева М.В., Хасбиуллин М.Ф. Нейронечеткая система формирования нечетких моделей оценки дискретного состояния объектов // Научно-технический вестник Поволжья. 2023. № 7. С. 60-63.



10. Кротова О.С., Пиянзин А.И., Хворова Л.А. Интеллектуальный анализ данных в диагностике сахарного диабета // Южно-Сибирский научный вестник. 2019. № 1 (25). С. 39-42.

11. Андриянов Н.А., Орлов Е.А. Разработка модели машинного обучения для оценки состояния глаз водителя // Инженерный вестник Дона, 2022, № 5 URL: ivdon.ru/ru/magazine/archive/n5y2022/7633.

### References

1. Dragomirova L.G., Hajrutdinova Z.L., SHubin L.L. Modern Science. 2022. № 5-4. pp. 59-62.

2. Balahonova E.A. Mirovaya nauka. 2023. № 2 (71). pp. 158-162.

3. ZHoldas N., Mansurova M.E. Avtomatika i programmaya inzheneriya. 2021. № 1 (35). pp. 11-20.

4. Kataseva D.V., Barinova A.O. Vestnik Tekhnologicheskogo universiteta. 2022. Т. 25. № 2. pp. 67-70.

5. Nurmatova E.V., Zajcev E.I. Inzhenernyj vestnik Dona, 2024, № 1. URL: ivdon.ru/ru/magazine/archive/n1y2024/8936

6. Prior F., Smith K., Almeida J., Kathiravelu P., Kurc T., Saltz J., Fitzgerald T.J. Clinical Radiology. 2020. Vol. 75. No. 1. P. 7-12.

7. YAmashkin S.A., YAmashkina E.O., YAmashkin A.A. Inzhenernyj vestnik Dona, 2022, № 12. URL: ivdon.ru/ru/magazine/archive/n12y2022/8082.

8. Kuroki M. Journal of Economic Education. 2023. Vol. 54. No. 4. P. 364-378.

9. Katasev A.S., Dagaeva M.V., Hasbiullin M.F. Nauchno-tehnicheskij vestnik Povolzh'ya. 2023. № 7. pp. 60-63.

10. Krotova O.S., Piyanzin A.I., Hvorova L.A. YUzhno-Sibirskij nauchnyj vestnik. 2019. № 1 (25). pp. 39-42.

11. Andriyanov N.A., Orlov E.A. Inzhenernyj vestnik Dona, 2022, № 5. URL: ivdon.ru/ru/magazine/archive/n5y2022/7633.

**Дата поступления: 19.09.2024**

**Дата публикации: 31.10.2024**

---