

Кизимова Т.А., Риксен В.С., Шпак В.А., Максимович К.Ю., Галимов Р.Р.

Использование методов машинного обучения для прогнозирования нитратного азота в почве

Электронный научно-производственный журнал
«АгроЭкоИнфо»

УДК 631.416.1:519.2

Использование методов машинного обучения для прогнозирования нитратного азота в почве

Кизимова Т.А., Риксен В.С., Шпак В.А., Максимович К.Ю., Галимов Р.Р.

Сибирский федеральный научный центр агробиотехнологий РАН

Аннотация

В данной работе представлены модели мультиномиальной логистической регрессии и нейронной сети, которые позволяют прогнозировать и тем самым оперативно определять содержание нитратного азота в 0–40 см слое почвы перед посевом. Для обучения моделей использованы данные длительного многофакторного полевого опыта СибНИИЗиХ СФНЦА РАН за 2009-2018гг, заложенного в 1981 г. на черноземе выщелоченном в ОС «Элитная» - филиале СФНЦА РАН. Учитывая особенности статистической выборки (данных наблюдений и анализов), определены основные предикторы моделей, влияющие на содержание нитратного азота в почве (целевой показатель), они представлены качественными (предшественник, обработка почвы) и количественными (погодные условия и содержание продуктивной влаги перед посевом в слое 0–100 см) факторами с соответствующими градациями. Модели показали достаточно высокую достоверность при верификации на эмпирических данных и могут быть использованы в качестве инструмента для прогноза. Качество разработанной модели мультиномиальной логистической регрессии оценивали с помощью коэффициента детерминации, который по мере Нэйджелкера был равен 78 %, а по мере Коксаи Снелла 72%. Для определения предсказательной способности нейронной сети был проведён ROC-анализ, который показал, что площадь под ROC - кривой для каждой категории целевого показателя была близкая к 1, что говорит о высокой предсказательной силе данного метода. Проведена сравнительная оценка прогностических возможностей обученных моделей. Общая доля правильных прогнозов для мультиномиальной логистической регрессии составляет 80,6%, в модели нейронной сети 89,5%.

Ключевые слова: МУЛЬТИНОМИАЛЬНАЯ ЛОГИСТИЧЕСКАЯ РЕГРЕССИЯ, НЕЙРОННАЯ СЕТЬ, НИТРАТНЫЙ АЗОТ, ПОЧВА

Введение

Величина и качество урожая зерновых культур зависит от комплекса природных и агротехнических факторов, ведущее место среди которых занимает обеспеченность растений элементами питания, и прежде всего азотом. Урожайность зерновых культур снижается как при недостатке азота, так и при его избытке. Поэтому, определение оптимальной дозы удобрений приобретает важное значение. Однако оптимальная доза внесения азотных удобрений сильно варьируется, поскольку содержание азота не является постоянной величиной и сильно колеблется в зависимости от многих факторов: свойства почвы, предшественника, срока и вида основной обработки почвы, погодных условий и т.д. [1-5]. В России основным и более точным методом почвенной диагностики азотного питания полевых культур является традиционный метод. Он осуществляется на основе отбора образцов почвы в поле и лабораторных анализов их элементного состава. При этом данный метод имеет недостатки, такие как большая трудоемкость отбора образцов почвы для анализа, особенно в подпахотных горизонтах, и недостаточная изученность размера использования растениями минерального азота из нижних слоев почвы, что ограничивает применение этого подхода [6, 7]. Так же химические анализы требуют много времени, средств и сложны в эксплуатации. Поэтому создание альтернативных методов быстрого определения содержания азота в почве перед посевом, имеет важное значение для сельскохозяйственных угодий [8].

Сегодня мировое сельское хозяйство перешло на новый этап технологического развития (Agriculture 4.0), который заключается в комплексном внедрении цифровых технологий: интернета вещей, геоинформационных систем (ГИС), беспилотных летательных аппаратов, робототехнических устройств, мобильных приложений, платформенных технологий сбора, обработки данных и машинного обучения [9-12].

Применение машинного обучения в сельском хозяйстве в настоящее время сопровождается массовым интересом со стороны мирового научного сообщества. В машинном обучении используются различные модели: линейной, логистической, полиномиальной регрессии, байесовских и нейронных сетей, метод опорных векторов, деревья принятия решений, случайные лес, метод k-ближайших соседей и др., выполняющие задачи предиктивной аналитики. Центральной сущностью предиктивной аналитики является задача определения предиктора или нескольких предикторов

(параметров или сущностей, которые влияют на прогнозируемое событие). Без прогнозирования трансформации условий, объектов и процессов, происходящих в сельском хозяйстве, трудно или почти невозможно принять правильное решение по его управлению. За рубежом эти методы активно используются для решения проблематики сельского хозяйства, в частности и для прогнозирования содержания азота в почве [13-18].

Поэтому, по нашему мнению, для более оперативного определения содержания азота в почве перед посевом как альтернатива традиционному методу, вполне возможно использовать различные методы искусственного интеллекта с элементами машинного обучения.

Цель исследований – построить с помощью методов машинного обучения модели способные прогнозировать содержание нитратного азота в почве, оценить точность прогнозных моделей.

Материалы и методы

Моделирование проводили на данных многофакторного полевого опыта СиНИ-ИЗиХ СФНЦА РАН (заложенного в 1981 г.), расположенном на территории ОС «Элитная» - филиала СФНЦА РАН Новосибирской области (центрально-лесостепная подзона). Данные включают результаты исследований четырёхпольного зернопарового севооборота (пар-пшеница-пшеница-пшеница) с разными вариантами основной обработки почвы с 2009 по 2018 год: вспашка (под 1-ю и 3-ю культуру на 20-22 см, под 2-ю на 25-27 см); безотвальная обработка (безотвальное рыхление стойками СибИМЭ под 1-ю и 3-ю культуры на глубину 20-22 см и под 2-ю на 25-27 см); минимальная обработка (плоскорезная обработка на глубину 10-12 см под все культуры ежегодно); нулевая обработка (без зяблевой обработки). Почва стационара – чернозём выщелоченный среднесуглинистого гранулометрического состава. Для моделирования использовали данные по содержанию нитратного азота в 0–40 см слое почвы перед посевом.

Построение мультиномиальной регрессии и нейронной сети осуществляли с помощью программного пакета модулей SPSS версия «26». При моделировании 80% данных исходной выборки использовались для обучения моделей, 20% в качестве тестирования. Размерность уровня содержания нитратного азота в почве устанавливали в соответствии со шкалой А.Е. Кочергина.

Результаты исследований

Построение мультиномиальной логистической регрессии (МЛР)

Для прогноза целевого показателя - содержание нитратного азота (N-NO₃ в кг/га почвы), исследовали связь этого показателя (зависимая переменная) от независимых переменных (факторов). При этом учитывались как качественные факторы: предшественник, способ обработки почвы, так и количественные, которые характеризуют погодные условия и запасы продуктивной влаги почвы перед посевом.

Качественные (категориальные) факторы имеют следующие градации:

- способ обработки почвы с 3 градациями: 0-нулевая обработка, 1- безотвальная обработка, 2-вспашка;

- предшественник имеет 4 градации: 0-пар, 1- первая пшеница по пару, 2- вторая пшеница по пару, 3- третья пшеница по пару.

К количественным факторам относятся погодные условия, а именно сумма активных температур воздуха $>0^{\circ}$ (больше нуля градусов) и осадков за периоды: сентябрь-ноябрь, март-апрель, содержание продуктивной влаги в метровом слое почвы перед посевом с/х культур.

Зависимую переменную можно классифицировать в качестве категориальной, при этом значения содержания нитратного азота в 0-40 см слое описывается тремя градациями: 0-очень низкое от 0-25 кг/га, 1-низкое 25-50 кг/га, 2-выше 50 кг/га (данная градация объединена - включает в себя среднее 50-75 кг/га и высокое 70-110 кг/га содержание нитратного азота, так как при построении моделей был недостаток исходных данных этих градаций).

Поскольку зависимая переменная N-NO₃ является категориальной и принимает значения перечисленных 3 категорий, то требуемую зависимость нельзя получить с использованием обычных регрессионных подходов. В этом случае требуемую зависимость можно получить с использованием модели МЛР или нейронных сетей (НС) [19-22].

В модели МЛР одна из категорий зависимой переменной объявляется опорной (эталонной), а все другие категории сравниваются с ней. Независимые переменные могут быть категориальными или количественными. Уравнение МЛР прогнозирует вероятность принадлежности зависимой переменной к категориям по значениям независимых

переменных. Окончательный выбор предсказываемой категории для зависимой переменной производится по правилу наибольшей вероятности принадлежности.

Для получения параметров (коэффициентов) МЛР, была использована выборка данных, сформированных в виде таблицы из 96 наблюдений (строк) и 8 факторов (столбцов), включая зависимую переменную. Число наблюдений распределено по показателю содержания нитратного азота в 0–40 см слое почвы следующим образом: очень низкое - 30, низкое - 47, выше 50 кг /га - 19. В качестве опорной категории выбрана вторая категория с содержанием нитратного азота выше 50 кг/га (можно выбрать любую другую).

Опишем кратко модель мультиномиальной логистической регрессии. Предполагается, что имеется серия из N наблюдений. Каждое наблюдение состоит из набора из m независимых переменных $X_i, i=1, \dots, m$ (также называемых предикторами) и соответствующего категориального значения зависимой переменной $Y_j, j=1, \dots, K$, которая может принимать одно из K возможных значений (категорий). Для каждой категории зависимой переменной (за исключением опорной) строится уравнение бинарной логистической регрессии, которое определяет отношение p_j - вероятности отнесения рассматриваемого наблюдения к этой категории к вероятности для опорной категории - p_1 :

$$\ln \left(\frac{p_j}{p_1} \right) = \alpha_j + \sum_i \beta_j^i X_i, \quad j = 2, \dots, K \quad (1)$$

где α_j – свободный член, β_j^i - коэффициент регрессии, связанный с i -й независимой переменной для категории j .

Неизвестные коэффициенты $\alpha_j, \beta_j^i, j=2, \dots, K, i=1, \dots, m$ совместно оцениваются для всей выборки наблюдений по максимальной апостериорной оценке (MAP), которая является расширением максимального правдоподобия с использованием регуляризирующих весов. Решение находится итеративно путем минимизации пересмотренных наименьших квадратов (IRLS) [20-22].

Используя формулу (1) и тот факт, что сумма всех K вероятностей должна составлять единицу, получим искомые вероятности принадлежности к категориям:

$$p_1 = 1 / (1 + \sum_j \exp(\alpha_j + \sum_i \beta_j^i X_i)) \quad (2)$$

и

$$p_j = \exp(\alpha_j + \sum_i \beta_j^i X_i) / (1 + \sum_j \exp(\alpha_j + \sum_i \beta_j^i X_i)), \text{ для } j=2, \dots, K \quad (3)$$

В таблице 1 приведены значения коэффициентов МЛР, их экспоненциальные значения и оценки значимости факторов, полученные с применением статистики Вальда.

Из формулы (3) для оценки вероятности принадлежности к категориям следует, что предикторы со значительными отрицательными коэффициентами уменьшают вероятность этой категории по отношению к эталонной категории, а предикторы с положительными коэффициентами, наоборот, увеличивают вероятность этой категории. О степени влияния предикторов на расчетные вероятности принадлежности к категориям говорят экспоненциальные значения соответствующих коэффициентов, приводимые в последнем столбце таблицы.

Таблица 1. Оценка параметров (коэффициентов) МЛР

NO ₃ по категориям ^a	Предикторы	Коэффициенты переменных (В)	Значимость коэффициента (Р)	Exp-Экспонента (В)
0	Сумма осадков, мм (сентябрь-ноябрь)	0,270	0,157	1,310
	Сумма осадков, мм (декабрь-февраль)	-1,069	0,054	0,343
	Сумма осадков, мм (март-апрель)	3,425	0,048	30,729
	Сумма температур >0° (сентябрь-ноябрь)	-0,017	0,067	0,983
	Сумма температур >0° (март-апрель)	0,176	0,000	1,192
	Запасы продуктивной влаги перед посевом в слое почвы (0–100 см), мм	-0,055	0,594	0,946
	Предшественник 1-я пшеница по пару	-14,819	0,029	3,7363E-07
	Предшественник 2-я пшеница по пару	-11,530	0,122	1,01301E-05
	Предшественник 3-я пшеница по пару	-13,818	0,017	1,01563E-06
	Предшественник пар	-24,202	0,011	3,09082E-11
	Обработка почвы безотвальная	1,151	0,504	3,161

NO ₃ по категориям ^а	Предикторы	Коэффициенты переменных (B)	Значимость коэффициента (P)	Exp-Экспонента (B)
	Обработка почвы Вспашка	-0,455	0,051	0,634
	Обработка почвы нулевая	0 ^с	-	-
1	Сумма осадков, мм (сентябрь-ноябрь)	0,320	0,213	1,377
	Сумма осадков, мм (декабрь-февраль)	-1,027	0,047	0,358
	Сумма осадков, мм (март-апрель)	3,173	0,033	23,889
	Сумма температур >0° (сентябрь-ноябрь)	-0,013	0,998	0,987
	Сумма температур >0° (март-апрель)	0,179	0,000	1,196
	Запасы продуктивной влаги перед посевом в слое почвы (0–100 см), мм	-0,046	0,651	0,955
	Предшественник 1-я пшеница по пару	-14,548	0,0526	5,04348E-07
	Предшественник 2-я пшеница по пару	-12,049	0,046	6,14421E-06
	Предшественник 3-я пшеница по пару	-13,675	0,034	1,2405E-06
	Предшественник пар	-21,652	0,051	4,1614E-10
	Обработка почвы безотвальная	0,130	0,933	1,139
	Обработка почвы Вспашка	-1,013	0,019	0,363
	Обработка почвы нулевая	0 ^с	-	-

Примечание: а – опорная категория 2; с – избыточный параметр (присваивается значение ноль).

Большие значения по абсолютной величине коэффициентов предикторов свидетельствуют о значимости этих факторов. Если значимость коэффициента $P < 0,05$, то связь статически значима. Результат $P > 0,05$ говорит о том, что связь между переменными слабая или не обнаружена. Из таблицы видно, что самыми значимыми являются следующие факторы: предшественник, сумма осадков и температур (март-апрель), сумма осадков (декабрь-февраль), обработка почвы.

Адекватность модели оценена по критерию отношения правдоподобия. Получена

высокая статистическая значимость построенной модели по критерию Хи квадрат ($<0,001$), которая свидетельствует о ее высокой пригодности для решения практических задач.

Для оценки качества модели обычной линейной регрессии используется показатель R-квадрат, который описывает ту часть дисперсии, которую можно объяснить с помощью регрессии. В случае МЛР такую роль играет показатель Псевдо R-квадрат. Наиболее распространенными являются меры, предложенные Нэйджелкерком, Коксом и Снеллом. В нашем случае объясненная часть дисперсии составляет по мере Нэйджелкерка (как правило, наиболее применяемой) 78 %, а по мере Кокса и Снелла 72%, что говорит о высоких прогностических возможностях метода.

В качестве критерия оценки прогнозной модели определена величина отклонения (погрешность) фактического содержания нитратного азота от прогнозируемого. В таблице 2 представлены сравнительные прогностические способности данного метода, проверенные на исходной выборке.

Таблица 2. Классификационная таблица МЛР

Количество наблюдений	Предсказанные			% правильных прогнозов
	Очень низкое 0–25 кг/га	Низкое 25–50 кг/га	Выше 50 кг/га	
Очень низкое 0–25 кг/га (30)	21	9	0	69,7
Низкое 25–50 кг/га (47)	6	38	3	81,1
Выше 50 (19)	0	1	18	95,5
Общая доля				80,6

Для выборки наблюдений с очень низким содержанием нитратного азота (0–25 кг/га) доля правильных прогнозов составила 69,7%, из них верно спрогнозировано 21 наблюдение, неправильными были 9 прогнозов, которые попали в категорию низкое содержание (25–50 кг/га). Для категории низкое содержание нитратного азота правильно было спрогнозировано 38 наблюдений из 47, при этом процент правильных прогнозов составил 81,1%. Самый большой процент (соответствует 95,5) правильных прогнозов у градации выше 50–70 кг/га, из 19 наблюдений неверно было спрогнозировано 1

наблюдение относящиеся к категории с низким содержанием. Общая доля правильных прогнозов по всем категориям составляет 80,6%.

Построение нейронной сети (НС).

Модели нейронных сетей, используя подходящую архитектуру, также позволяют оценивать вероятности принадлежности зависимой переменной к возможным категориям (нейроны выходного слоя) по значениям используемых предикторов (нейроны входного слоя). Итоговый выбор предсказываемой категории производится также по правилу наибольшей вероятности принадлежности.

Нейронные сети являются гибким средством прогнозирования процессов, зависящих от разнообразных факторов. Их применение в настоящее время является приоритетным при решении подобных задач.

В нашем случае выбор сделан в пользу наиболее популярной модели нейронной сети - многослойный персептрон (MLP). Нейронная сеть описывается направленным графом с взвешенными связями, в котором нейроны являются вершинами, а синаптические связи (веса) – дугами. MLP - нейронная сеть прямого распространения сигнала (без обратных связей), в которой входной сигнал преобразуется в выходной, проходя последовательно через несколько слоев. Первый из таких слоев называют входным, последний - выходным. Кроме входного и выходного слоев, в многослойном персептроне есть один или несколько промежуточных слоев, которые называют скрытыми.

Сеть, изображенная на рис. 1, имеет 13 входов, на которые поступают сигналы в виде значений предикторов, идущие по синапсам на 3 нейрона, образующие выходной слой данной сети и выдающие три выходных сигнала для каждой категории N-NO₃. Имеется 2 скрытых слоя соответственно с 5 и 4 искусственными нейронами.

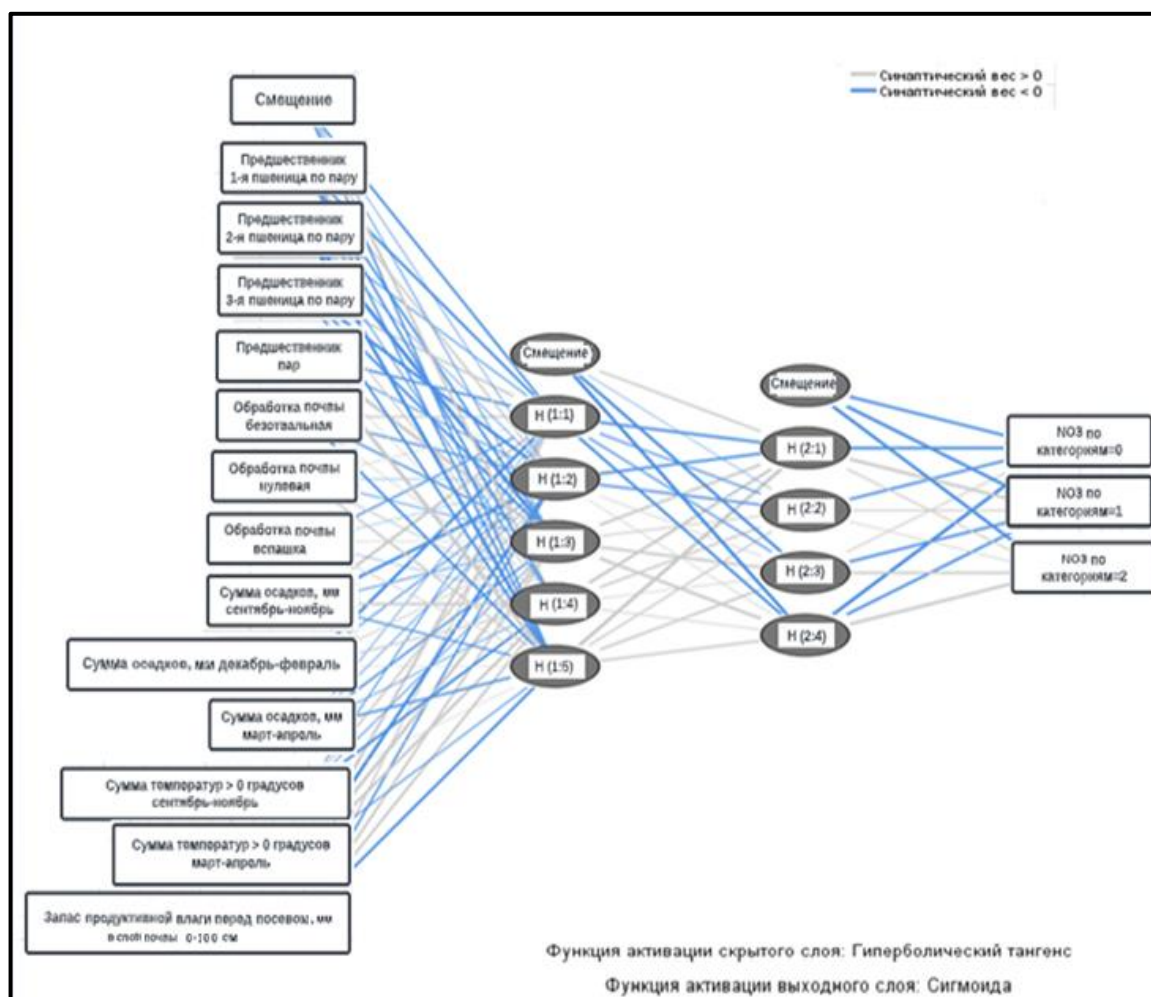


Рис. 1. Структура нейросети

Определение синаптических весов нейронной сети производится на базе деления исходной выборки на два множества: обучающее и тестовое. На обучающем множестве происходит обучение нейронной сети. На тестовом множестве осуществляется проверка построенной модели. Эти множества не пересекаются.

При обучении и тестировании, синаптические веса и смещения выбираются таким образом, чтобы минимизировать ошибку обучения, представляющую собой меру разности между фактическими и полученными на выходе сети значениями сигналов. В таблице 3 представлены оценки эффективности построенной нейронной сети, полученные на обучающейся и тестовой выборках. Для первой градации количество достоверно спрогнозированных наблюдений составляет 23, для второй 45 и для третьей 18.

Таблица 3. Классификационная таблица нейронной сети

Количество наблюдений	Предсказанные			% правильных прогнозов
	Очень низкое 0–25 кг/га	Низкое 25–50 кг/га	Выше 50 кг/га	
Очень низкое 0–25 кг/га (30)	23	7	0	76,6
Низкое 25–50 кг/га (47)	1	45	1	95,7
Выше 50 кг/га (19)	0	1	18	95,5
Общая доля				89,5

Видно, что общий процент правильно классифицируемых наблюдений достаточно высокий и составляет 89,5. Точность данного метода на 10% выше, чем в модели МЛР.

Еще одним критерием эффективности нейросети являются так называемые ROC-кривые, описывающие качество классификации каждой категории зависимой переменной (рис. 2).

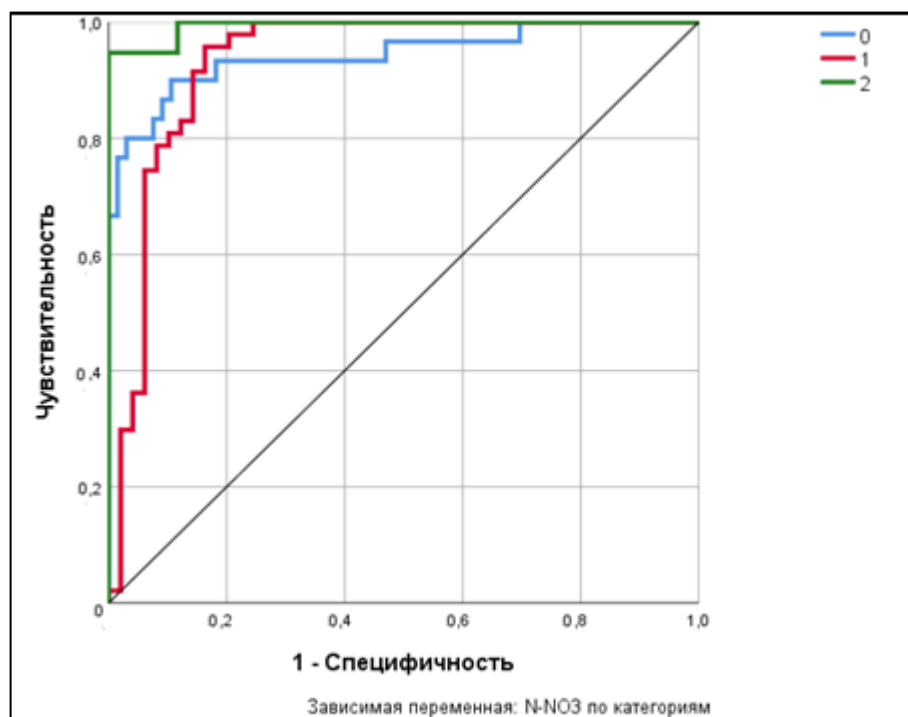


Рис. 2. График ROC кривой нейронной сети

На рис. 2 приведены ступенчатые ROC-кривые для трёх категорий содержания нитратного азота. Прямая линия по диагонали представляет ROC-кривую чисто

случайного классификатора. Эффективный классификатор отклоняется от этой линии, как можно выше, максимально прижимаясь к левому верхнему углу. Способом сравнения классификаторов являются значения площади под ROC- кривой. Наиболее совершенный классификатор имеет площадь под ROC - кривой близкую к 1, чисто случайный классификатор - площадь 0,5 (демонстрирует непригодность выбранного метода). Значения в таблице 4 для каждой категории есть вероятность правильной классификации наблюдения в сравнении с классификацией других категорий.

Таблица 4. Значения площади ROC кривой

N-NO ₃ по категориям	Площадь ROC кривой
0 (0–25 кг/га)	0,943
1 (25–50 кг/га)	0,931
2 (выше 50 кг/га)	0,994

Полученные оценки площадей говорят о высоком качестве классификаторов каждой категории целевого показателя и пригодности применения данной нейросети для решения прогностических задач.

Так же проведён анализ чувствительности влияния каждого предиктора на предсказанные значения нейросетевой модели (рис. 3).

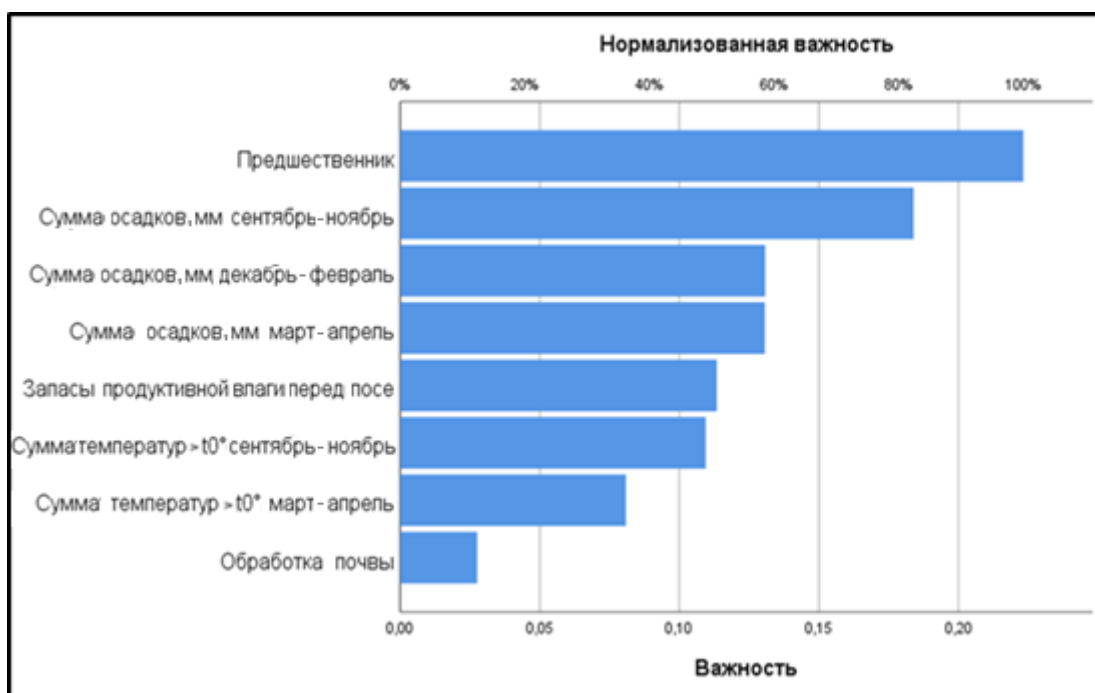


Рис. 3. Оценка влияния факторов на целевой показатель

При данном анализе исходная выборка разбивается (делится) на обучающую и тестовую, в результате чего создаётся диаграмма отображения важности каждого предиктора (представленная в нормализованном виде). Наиболее важный предиктор принимается за 100%, остальные предикторы сортируются в порядке убывания значений важности.

В данном случае анализ чувствительности показывает, что наибольшее влияние на целевую переменную оказывает такой фактор как предшественник, в меньшей степени погодные условия, запас продуктивной влаги в почве перед посевом и на самом последнем месте обработка почвы.

Верификацию построенных моделей провели на данных 2018 года (табл. 5).

Таблица 5. Прогноз содержания нитратного азота в 0–40 см слое почвы в 2018 г.

Предшественник	Обработка почвы	Фактическое содержание N-NO ₃ кг/га	Содержание N-NO ₃ мг/кг (прогноз МЛР)	Содержание N-NO ₃ мг/кг (прогноз по НС)
Пар	Вспашка	107,7	Выше 50 кг/га	Выше 50 кг/га
Пар	Безотвальная	62,04	Выше 50 кг/га	Выше 50 кг/га /га
Пар	Нулевая	78,9	Выше 50 кг/га	Выше 50 кг/га /га
1-ая пшеница по пару	Вспашка	28,3	Низкая 25-50 кг/га	Низкая 25-50 кг/га
1-ая пшеница по пару	Безотвальная	17,5	Низкая 25-50 кг/га*	Очень низкая 0-25 кг/га
1-ая пшеница по пару	Нулевая	23,7	Низкая 25-50 кг/га*	Очень низкая 0-25 кг/га
2-ая пшеница по пару	Вспашка	18,5	Очень низкая 0-25 кг/га	Низкая 25-50 кг/га
2-ая пшеница по пару	Безотвальная	27,5	Очень низкая 0-25 кг/га*	Низкая 25-50 кг/га
2-ая пшеница по пару	Нулевая	27,5	Низкая 25-50 кг/га	Очень низкая 0-25 кг/га *
3-я пшеница по пару	Вспашка	28,5	Низкая 25-50 кг/га	Низкая 25-50 кг/га
3-я пшеница по пару	Безотвальная	28,7	Очень низкая 0-25 кг/га*	Очень низкая 0-25 кг/га*
3-я пшеница по пару	Нулевая	25,4	Низкая 25-50 кг/га	Низкая 25-50 кг/га

Примечание: *- неверно спрогнозированные градации

Результаты верификации показали, что наиболее точный прогноз был в модели НС, где из 12 наблюдений ошибочно было спрогнозировано 2. МЛР дала неправильный

прогноз в 4-х случаях. Таким образом, учитывая небольшой объем статистической выборки и малое количество предикторов, прогностические свойства проверенных моделей можно признать удовлетворительными.

Заключение

В ходе исследования с помощью методов машинного обучения были построены и обучены модели, позволяющие оперативно с допустимой достоверностью определять (прогнозировать) содержание нитратного азота перед посевом в 0–40 см слое почвы, используя только данные о погоде, обработке почвы, предшественнике и запасе продуктивной влаги перед посевом. На основе прогноза можно разрабатывать рекомендации по ориентировочным дозам внесения удобрений в зависимости от спрогнозированной градации, планировать возможную прибавку урожайности от удобрений.

В дальнейшем планируется улучшать качество работы моделей, добавляя другие предикторы, влияющие на результирующий признак, осуществлять поиск методов машинного обучения, допускающих анализ данных небольшого объема.

Список использованных источников:

1. Самойлов Л.Н., Чернова Л.С., Трушкин С.В. Комплекс факторов, влияющих на производство и качество зерна пшеницы // Плодородие. – 2018. – №6 (105). – С. 12–16.
2. Концевая С.Р. Анализ влияния абиотических факторов на урожайность зерновых культур // Бухучет в сельском хозяйстве. – 2018. – № 1. – С. 81–87.
3. Малыгина Н.С., Тихойкина И.М. Исследование факторов, влияющих на урожайность и качество зерна пшеницы (на примере Орловского района Орловской области) // Вестник Орел ГИЭТ. – 2015. – № 3. – С. 43–47.
4. Пасынкова Е.Н. Агрохимические приемы регулирования урожайности и качества зерна пшеницы: Автореф. дисс. докт. биол. наук. – М.: ВНИИА. – 2014. – 49 с.
5. Гамзиков Г.П. Почвенная диагностика азотного питания растений и применения азотных удобрений в севооборотах // Плодородие. – 2018. – №1 (100). – С. 8–14.
6. Кидин В.В. Основы питания растений и применения удобрений // Москва: РГАУ-МСХА. – 2008. – 415 с.
7. Воробьева Л.А. Теория и практика химического анализа почв // М.: ГЕОС. – 2006. – 400 с.
8. Grell M., Barandun G., Asfour T. Point-of-use sensors and machine learning enable

low-cost determination of soil nitrogen // *Nature Food* 2. – 2021. – P. 981-989.

9. Анищенко А.Н., Шутьков А.А. Agriculture 4.0 как перспективная модель научно-технологического развития аграрного сектора современной России // *Продовольственная политика и безопасность*. – 2019. – №6 (3). – С. 129–140.

10. Ганиева И.А. Цифровая трансформация сельского хозяйства России: консолидация государства и агробизнеса // *Достижения науки и техники АПК*. – 2019. – №4 (33) – С. 5–7.

11. Федоренко Н.П., Мишуров Д.С., Буклагин Д.С., Гольцяпин В.Я., Голубев И.Г. Цифровое сельское хозяйство: состояние и перспективы развития. – Москва: ФГБНУ "Росинформагротех", 2019. – 316 с.

12. Степных Н.В., Нестерова Е.В., Заргарян А.М. Цифровизация управления агротехнологиями. – Куртамыш: ООО «Куртамышская типография», 2018. – 43 с.

13. Yuefen Li, Shuo Liang, Yiying Zhao, Wenbo Li, Yuejiao Wang. Machine learning for the prediction of *L. chinensis* carbon, nitrogen and phosphorus contents and understanding of mechanisms underlying grassland degradation // *Journal of Environmental Management*. – 2017. – Volume 192. – P. 116-123.

14. Каличкин В.К., Лужных Т.А., Риксен В.С. Прогнозирование содержания нитратного азота в почве с использованием машинного обучения // *Сибирский вестник сельскохозяйственной науки*. – 2021. – Т. 51. – № 5. – С. 91–100.

15. Xinbing Wang, Yuxin Miao, Rui Dong, Hainie Zha, Tingting Xia, Zhichao Chen, Krzysztof Kusnierek, Guohua Mi, Hong Sun, Minzan Li. Machine learning-based in-season nitrogen status diagnosis and side-dress nitrogen recommendation for corn // *European Journal of Agronomy*. – 2021. – Volume 123. – P. 126-193.

16. Osco L.P., Junior J.M., Ramos A.P., Furuya D.E.G., Santana, D.C., Teodoro, L.P.R., Gonçalves, W.N., Baio F.H.R., Pistori H. Leaf Nitrogen Concentration and Plant Height Prediction for Maize Using UAV-Based Multispectral Imagery and Machine Learning Techniques // *Remote Sens*. – 2020. – Vol .12. – P. 32-37.

17. Chlingaryan A., Sukkariéh S., Whelan B. Machine learning approaches for crop yield prediction and nitrogen status estimation in precision agriculture // *A review, Computers and Electronics in Agriculture*. – 2018. – Vol. 151. – P. 61-69.

18. Скворцов Е.А., Набоков В.И., Некрасов К.В., Скворцова Е.Г., Кротов М.И. Применение технологий искусственного интеллекта в сельском хозяйстве // *АВУ*. – 2019. – №8 (187). – С. 91–98.

19. Бююль А., Цеффель П. SPSS: искусство обработки информации. – СПб.: ООО «Диа-СофтЮП», 2005. – 608 с.

20. Николенко С., Кадурич А., Архангельская Е. Глубокое обучение. — СПб.: Питер, 2018. — 480 с.

21. Cohen A., Rom M.A. Method for Hypothesis Tests in Polychotomous Logistic Regression // *Computational Statistics and Data Analysis*, 1994. – Vol. 17. – P. 277-288.

Кизимова Т.А., Риксен В.С., Шпак В.А., Максимович К.Ю., Галимов Р.Р.

Использование методов машинного обучения для прогнозирования нитратного азота в почве

Электронный научно-производственный журнал

«АгроЭкоИнфо»

22. Long J.S. Regression models for categorical and limited dependent variables, Thousand Oaks: Sage Publ, 1997. – 328 p.

Цитирование:

Кизимова Т.А., Риксен В.С., Шпак В.А., Максимович К.Ю., Галимов Р.Р. Использование методов машинного обучения для прогнозирования нитратного азота в почве [Электрон. ресурс] // АгроЭкоИнфо: Электронный научно-производственный журнал. – 2022. – № 5. – Режим доступа: http://agroecoinfo.ru/STATYI/2022/5/st_521.pdf. DOI: <https://doi.org/10.51419/202125521>.