
AUXILIARY DISCIPLINES

DOI: <https://doi.org/10.23649/jae.2021.4.20.13>

Galanina O.V.¹*, Golubev A.M.²

¹ Saint Petersburg State Agrarian University, Saint-Petersburg, Russia;

² T-Systems RUS LLC, Saint-Petersburg, Russia

* Corresponding author (olga_galanina[at]inbox.ru)

Received: 15.09.2021; Accepted: 27.09.2021; Published: 15.12.2021

FEEDFORWARD NEURAL NETWORK IN EXPOSING THE HIDDEN DATA OF THE STATE STATISTICS COMMITTEE

Research article

Abstract

Feedforward neural networks are effectively used with some information uncertainty to obtain forecasts and can be used to calculate intermediate values.

The aim of the study is to obtain the most probable quantitative estimates of hidden or unpublished statistical information.

For example, statistical collections repeatedly lack information on sugar beet yields in some regions. This information could have been deliberately hidden in order to ensure the confidentiality of primary statistical data in accordance with Federal Law No. 282-FZ dated November 29, 2007 "On Official Statistical Accounting and the System of State Statistics in the Russian Federation".

The feedforward neural network is able to solve the task set and evaluate the missing information.

The proposed research is suitable not only in assessing the missing numerical information of statistical indicators of the agro-industrial complex but also in any similar life situations.

Keywords: feedforward neural network, artificial intelligence, intelligent system, mathematical model, artificial neuron, perceptron, forecast, agro-industrial complex, agricultural sector of the economy, agricultural economy.

Галанина О.В.¹*, Голубев А.М.²

¹ Санкт-Петербургский государственный аграрный университет, Санкт-Петербург, Россия;

² ООО «Т-Системс РУС», Санкт-Петербург, Россия

* Корреспондирующий автора (olga_galanina[at]inbox.ru)

Получена: 15.09.2021; Доработана: 27.09.2021; Опубликована: 15.12.2021

НЕЙРОННАЯ СЕТЬ ПРЯМОГО РАСПРОСТРАНЕНИЯ (FEEDFORWARD NEURAL NETWORK) В ИЗОБЛИЧЕНИИ СОКРЫТЫХ ДАННЫХ ГОСКОМСТАТА

Научная статья

Аннотация

Нейронные сети прямого распространения эффективно используются при некоторой информационной неопределенности для получения прогнозов и могут быть использованы для расчета промежуточных значений.

Целью исследования является получение наивероятнейших количественных оценок скрытой или неопубликованной статистической информации.

Например, в статистических сборниках неоднократно отсутствует информация по урожайности сахарной свеклы в некоторых регионах. Эта информация сознательно могла быть скрыта с целью обеспечения конфиденциальности первичных статистических данных в соответствии с Федеральным законом от 29 ноября 2007 №282-ФЗ «Об официальном статистическом учете и системе государственной статистики в Российской Федерации».

Нейронная сеть прямого распространения способна решить поставленную задачу и оценить недостающие сведения.

Предлагаемое научное исследование пригодно не только в оценке недостающей числовой информации статистических показателей АПК, но и в любых аналогичных жизненных ситуациях.

Ключевые слова: нейронная сеть прямого распространения, искусственный интеллект, интеллектуальная система, математическая модель, искусственный нейрон, перцептрон, персептрон, прогноз, АПК, аграрный сектор экономики, аграрная экономика.

1. Введение

При анализе статистической информации по урожайности, посевным площадям, валовому сбору сахарной свеклы мы столкнулись с проблемой, что в статистических сборниках информация за два года – 2014 и 2015 не опубликована. Основанием к не публикации указано, что «данные не публикуются в целях обеспечения конфиденциальности первичных статистических данных, полученных от организаций, в соответствии с Федеральным законом от 29 ноября 2007 г. № 282-ФЗ «Об официальном статистическом учете и системе государственной статистики в Российской Федерации» (п.5 ст.4, ч.1 ст.9)». На рисунке 1 можно наблюдать описанную картину. Такие же проблемы отсутствия статистических данных прослеживаются и в некоторых других регионах (например, Оренбургская область за 2017-2019 гг.).

Конечно, эти данные могут быть аппроксимированы одним из традиционных методов, например, регрессионным уравнением зависимости. Но такие аппроксимации весьма неточны. Нам кажется, что получить наиболее точный результат возможно с использованием нейронной сети. Именно обоснованию этого положения и посвящена настоящая работа. Да и по сути метод прогнозирования и аппроксимации с использованием нейронных сетей прямого распространения является новейшим, весьма перспективным методом, процедура которого еще недостаточно описана.

14.15. УРОЖАЙНОСТЬ САХАРНОЙ СВЕКЛЫ
(в хозяйствах всех категорий; центнеров с одного гектара убранной площади)

	2005	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019
Российская Федерация	282	241	392	409	442	370	388	470	442	381	480
Центральный федеральный округ	293	204	407	431	445	355	391	470	445	401	486
Белгородская область	311	181	407	424	414	415	379	532	411	442	486
Брянская область	244	300	363	430	505	... ¹⁾	... ¹⁾	427	443	479	403
Воронежская область	281	179	382	438	439	395	437	484	470	395	509
Курская область	253	228	414	426	404	335	352	488	483	464	539
Липецкая область	374	218	406	490	459	299	380	445	420	348	455
Орловская область	253	272	419	406	446	326	333	403	405	404	463
Рязанская область	291	232	389	421	410	321	429	489	463	386	518
Тамбовская область	274	196	437	397	505	364	429	448	445	378	457

¹⁾ Данные не публикуются в целях обеспечения конфиденциальности первичных статистических данных, полученных от организаций, в соответствии с Федеральным законом от 29 ноября 2007 г. № 282-ФЗ «Об официальном статистическом учете и системе государственной статистики в Российской Федерации» (п.5 ст.4, ч.1 ст.9).

Рис.1 – Скрытые от широких масс данные по урожайности сахарной свеклы в Брянской области в 2014-2015 гг.

2. Методы

Для оценки степени актуальности прилагаемого исследования был произведен анализ научных публикаций в индексе научного цитирования на сайте <http://elibrary.ru/>. Оказалось, что запросу:

(искусственный интеллект | нейронная сеть | цифровые технологии) & (растениеводство | агрономия)

Соответствует общее число публикаций 218. Распределение публикаций по годам представлено на рисунке 2. Можно заметить, что резкий всплеск интереса к цифровизации растениеводства возник в 2018 году. С 2017 года наблюдается экспоненциальный рост интереса к данной проблеме. Ожидаемое прогнозное значение числа публикаций к концу 2021 года составит не менее 100 публикаций.

С другой стороны, количество публикаций по теме запроса 218 – ничтожно малое количество. Что говорит о небывалой перспективности исследований в этой сфере.

Нейросетевые модели, как неотъемлемая часть цифровизации, стали интенсивно описываться и внедряться в сферу АПК. Математики всего мира акцентируют свое внимание на этом разделе математической науки. Например, исследователь из Волгограда И.С. Белоусов описал использование нейронных сетей в сельском хозяйстве, а именно сбор клубники, выращивание огурцов в автоматизированных теплицах и построение почвенных карт [1]. Теоретическими формулировками особенных агроэкономических задач занимается профессор Парфенова В.Е. из Санкт-Петербурга [2], [3], [4], [5].

В работах Галаниной О.В. [7], [8] установлено, что нейронная сеть с 4-мя входными параметрами выдает результаты прогнозирования урожайностей с коэффициентом вариации 1-4%. Причем для прогнозирования урожайности свеклы этот коэффициент составил 2%. Это очень высокий показатель качества прогноза, говорящий о том, что нейронная сеть достаточно точно предсказывает урожайность свеклы.

Ученый из Екатеринбурга Скворцов Е.А. проанализировал и перечислил основные направления применения технологий искусственного интеллекта и нейронных сетей в сельском хозяйстве [9], а именно, для прогноза урожайности сельскохозяйственных культур в зависимости от различных факторов; для мониторинга почв; в диагностике болезней растений, в прогнозе климатических рисков.

Не только в Московской области, но и во многих регионах РФ для мониторинга посевов используют малую авиацию и дистанционное зондирование посевов [10] в комплексе с интеллектуальным модулем.

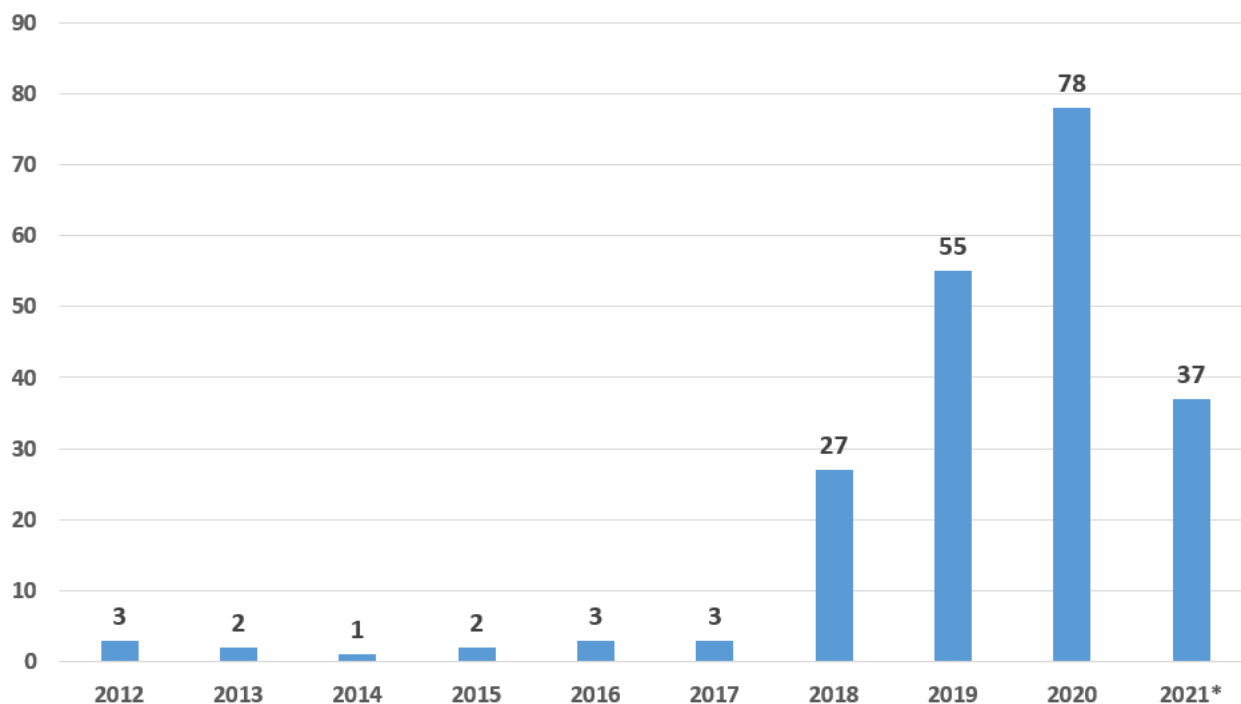


Рис. 2 – Динамика роста числа научных публикаций по исследуемому запросу

Примечание: * – год 2021 не полный

Поскольку уже было установлено, что нейронные сети с высокой степенью точности предсказывают результат [7],[8], можно получить недостающие сведения с их помощью.

Будем использовать нейронную сеть прямого распространения для заполнения пустых ячеек по урожайности сахарной свеклы в 2014-2015 гг. по Брянской области. Для этого в качестве исходной информации будем использовать следующие статистические значения:

- x1 – внесенное количество минеральных удобрений (в пересчете на 100% пит.в-в) на 1 га посева с.-х. культур, кг;
- x2 – внесенное количество органических удобрений на 1 га посева с.-х культур, т;
- x3 – среднемесячная температура воздуха за 3 месяца (май, июнь, июль), град.С;
- x4 – количество осадков за 6 месяцев (январь, февраль, март, апрель, май, июнь), мм;
- y – урожайность сахарной свеклы, ц/га.

Данные были сведены в таблицу (табл. 1).

Проведем корреляционно-регрессионный анализ статистической информации. Будем считать переменную y – зависимой величиной, переменные x1, x2, x3 – независимыми переменными.

Таблица 1 – Исходные данные для регрессионного и нейросетевого прогноза (кроме 2014 и 2015 гг)

Год	X1	X2	X3	X4	Y
2000	48,2	1,5	290	14,4	147
2001	35,2	1,4	272	15,4	149
...
2013	56,0	1,0	238	15,98	505
2014	72,9	1,1	160	15,5	?
2015	90,3	1,0	292	14,5	?
2016	119,5	0,9	290	15,58	427
...
2018	126,8	1,7	225	15,75	479
2019	124,6	2,6	187	15,68	403

Коэффициенты корреляции Пирсона между внесенным количеством минеральных удобрений и урожайностью сахарной свеклы $r_{x1,y}=0,55$; коэффициент корреляции Пирсона между внесенным количеством органических удобрений и урожайностью сахарной свеклы $r_{x2,y}=0,10$; коэффициент корреляции Пирсона между количеством осадков за 4 месяца и урожайностью свеклы $r_{x3,y}=-0,29$ и коэффициент корреляции Пирсона между средней температурой за 4 месяца и урожайностью свеклы $r_{x4,y}=0,27$. Таким образом, какая-то вероятностная связь между этими показателями прослеживается.

Уравнение регрессионной зависимости значимо по критерию Фишера и имеет вид: $Y=361.37+2.46x_1-131.06x_2-0.42x_3+7.95x_4$ при $R^2=0.46$. Это уравнение будем использовать для альтернативного (регрессионного) прогноза.

Для нейросетевого прогноза была выбрана архитектура нейронной сети прямого распространения в четырем входными нейронами (x_1, x_2, x_3, x_4), одним выходным нейроном (y) и со скрытым слоем нейронов, состоящим из 3-х слоев по 18, 10 и 10 нейронов в каждом (рис. 3).

В качестве обучающей выборки для простоты выбрали весь массив данных (18 векторов). Качеством нейро- и регрессионного прогнозов будем считать коэффициент вариации $V\sigma$.

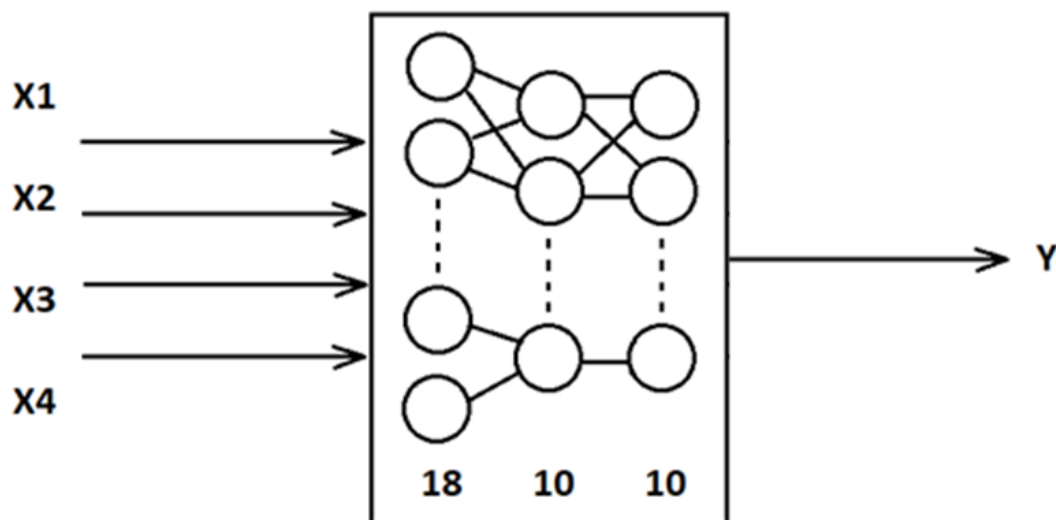


Рис.3 – Нейронная сеть прямого распространения, содержащая 4 нейрона во входном слое, 1 нейрон в выходном слое, скрытый слой, состоящий из трех слоев нейронов по 18, 10 и 10 нейронов в каждом

3. Результаты

Результаты вычислений были сведены в таблицу 2.

Таблица 2 – Результаты прогнозирования урожайности свеклы

	Фактическое значение урожайности сахарной свеклы, ц/га	Прогноз по регрессионной модели	Прогноз по нейросетевой модели
	Y	Y	Y
2000	147	276,1	150,0
2001	149	266,0	147,9
...
2013	505	395,3	498,1
2014	?	452,8	484,0
2015	?	445,3	441,7
2016	427	539,8	426,9
...
2018	479	481,7	479,0
2019	403	373,7	402,2
σ_i	0	87,30	5,63
μ_i		329,5	328,2
$V\sigma_i$		26%	2%

Коэффициент вариации $V\sigma$ прогноза по регрессионной модели составил 26%, это говорит о низком качестве прогноза. Коэффициент вариации $V\sigma$ по нейронной сети составил всего 2%. Таким образом, с высокой долей вероятности можно утверждать, что в 2014-м году урожайность сахарной свеклы в Брянской области составила 484,0 ц/га, а в 2015-м году – 441,7 ц/га.

4. Заключение

Цифровые технологии все интенсивнее проникают в сферу сельскохозяйственного производства. Формулируется и ставится большое количество задач нейросетевого интеллектуального характера в АПК, теоретических исследованиях. Сферы применения нейросетевого программирования и искусственного интеллекта постоянно расширяются.

На основе нейронных сетей можно находить недостающие или скрытые от посторонних глаз значения. Аналогично, нейронные сети прямого распространения можно применять не только для вычисления урожайностей каких-либо культур, но и для выявления ошибок, мошенничеств, спекуляций, недобросовестностей во всех сферах деятельности.

Conflict of Interest

None declared.

Конфликт интересов

Не указан.

References

1. Белоусов, И. С. Использование нейронных сетей в сельском хозяйстве / И. С. Белоусов // Развитие АПК на основе принципов рационального природопользования и применения конвергентных технологий: Материалы Международной научно-практической конференции, проведенной в рамках Международного научно-практического форума, посвященного 75-летию образования Волгоградского государственного аграрного университета, Волгоград, 30 января – 01 2019 года. – Волгоград: Волгоградский государственный аграрный университет, 2019. – С. 331-335.
2. Парфенова, В. Е. Нечеткое регрессионное моделирование в задачах управления аграрным производством / В. Е. Парфенова // Инновации. – 2019. – № 7(249). – С. 88-92. – DOI 10.26310/2071-3010.2019.249.7.013.
3. Парфенова, В. Е. Нечеткие модели принятия оптимальных решений в управлении аграрным производством / В. Е. Парфенова // Инновации. – 2018. – № 10(240). – С. 88-92.
4. Парфенова, В. Е. Нечеткая модель оптимизации структуры посевных площадей / В. Е. Парфенова // Известия Санкт-Петербургского государственного аграрного университета. – 2017. – С. 176-183.
5. Fuzzy modelling for tasks of management of the agricultural-industrial complex / V. E. Parfenova, G. G. Bulgakova, Y. G. Amagaeva, K. V. Evdokimov // IOP Conference Series: Materials Science and Engineering : International Scientific-Practical Conference on Quality Management and Reliability of Technical Systems 2019, St. Petersburg, 20–21 июня 2019 года. – St. Petersburg: Institute of Physics Publishing, 2019. – P. 012067. – DOI 10.1088/1757-899X/666/1/012067.
6. Forecasting models of agricultural process based on fuzzy time series / V. E. Parfenova, G. G. Bulgakova, Y. G. Amagaeva [et al.] // IOP Conference Series: Materials Science and Engineering : 3rd International Scientific-Practical Conference on Quality Management and Reliability of Technical Systems, St. Petersburg, 27–29 августа 2020 года. – BRISTOL, 2021. – P. 012013. – DOI 10.1088/1757-899X/986/1/012013.
7. Галанина, О. В. Преимущества использования интеллектуальных систем прогнозирования в экономике сельского хозяйства в условиях недостатка информации / О. В. Галанина // Известия Международной академии аграрного образования. – 2020. – № 51. – С. 61-64.
8. Галанина, О. В. Нейронная сеть прямого распространения в исследовании экономики сельского хозяйства / О. В. Галанина, Ю. П. Золотарёва // Известия Международной академии аграрного образования. – 2021. – № 56. – С. 61-64.
9. Скворцов, Е. А. Перспективы применения технологий искусственного интеллекта в сельском хозяйстве региона / Е. А. Скворцов // Экономика региона. – 2020. – Т. 16. – № 2. – С. 563-576. – DOI 10.17059/2020-2-17.
10. Применение аэрофотосъёмки с применением мультиспектральной камеры в комплексе с нейросетью в сельском хозяйстве / А. С. Сыч, В. Балык, А. И. Поляков, А. В. Карталов // Chronos. – 2019. – № 9(36). – С. 26-28.

References in English

1. Belousov, I. S. Ispol'zovanie nejjronnykh setejj v sel'skom khozjajjstve [The use of neural networks in agriculture] / I. S. Belousov // Razvitie APK na osnove principov racional'nogo prirodopol'zovanija i primenenija konvergentnykh tekhnologijj: Materialy Mezhdunarodnoj nauchno-prakticheskoy konferencii, provedennoj v ramkakh Mezhdunarodnogo nauchno-prakticheskogo foruma, posvjashhennogo 75-letiju obrazovanija Volgogradskogo gosudarstvennogo agrarnogo universiteta, Volgograd, 30 janvarja – 01 2019 goda [The development of agriculture based on the principles of rational nature management and the use of convergent technologies: Materials of the International Scientific and Practical Conference held within the framework of the International Scientific and Practical Forum dedicated to the 75th anniversary of the Volgograd State Agrarian University, Volgograd, January 30 - 01, 2019]. - Volgograd: Volgograd State Agrarian University, 2019. - pp. 331-335 [in Russian]
2. Parfenova, V. E. Nechetkoe regressionnoe modelirovanie v zadachakh upravlenija agrarnym proizvodstvom [Fuzzy regression modeling in agricultural production management problems] / V. E. Parfenova // Innovacii [Innovation]. – 2019. – № 7(249). – pp. 88-92– DOI 10.26310/2071-3010.2019.249.7.013 [in Russian]
3. Parfenova, V. E. Nechetkie modeli prinjatija optimal'nykh reshenijj v upravlenii agrarnym proizvodstvom [Fuzzy models of optimal decision-making in agricultural production management] / V. E. Parfenova // Innovacii [Innovation]. – 2018. – № 10(240). – pp. 88-92 [in Russian]
4. Parfenova, V. E. Nechetkaja model' optimizacii struktury posevnykh ploshhadejj [Fuzzy model of optimization of the structure of acreage] / V. E. Parfenova // Izvestija Sankt-Peterburgskogo gosudarstvennogo agrarnogo universiteta [Proceedings of the St. Petersburg State Agrarian University]. - 2017. - pp. 176-183 [in Russian]
5. Fuzzy modelling for tasks of management of the agricultural-industrial complex / V. E. Parfenova, G. G. Bulgakova, Y. G. Amagaeva, K. V. Evdokimov // IOP Conference Series: Materials Science and Engineering : International Scientific-Practical Conference on Quality Management and Reliability of Technical Systems 2019, St. Petersburg, 20–21 июня 2019 года. – St. Petersburg: Institute of Physics Publishing, 2019. – P. 012067. – DOI 10.1088/1757-899X/666/1/012067.
6. Forecasting models of agricultural process based on fuzzy time series / V. E. Parfenova, G. G. Bulgakova, Y. G. Amagaeva [et al.] // IOP Conference Series: Materials Science and Engineering : 3rd International Scientific-Practical Conference on Quality Management and Reliability of Technical Systems, St. Petersburg, 27–29 августа 2020 года. – BRISTOL, 2021. – P. 012013. – DOI 10.1088/1757-899X/986/1/012013.
7. Galanina, O. V. Preimushhestva ispol'zovanija intellektual'nykh sistem prognozirovanija v ehkonomike sel'skogo khozjajjstva v uslovijakh nedostatka informacii [Advantages of using intelligent forecasting systems in the agricultural economy in conditions of lack of information] / O. V. Galanina // Izvestija Mezhdunarodnoj akademii agrarnogo obrazovanija [Proceedings of the International Academy of Agrarian Education]. - 2020. - No. 51. - pp. 61-64 [in Russian]

8. Galanina, O. V. Nejjronnaja set' prjamogo rasprostraneniya v issledovanii ehkonomiki sel'skogo khozjajstva [Neural network of direct distribution in the study of agricultural economics] / O. V. Galanina, Yu. P. Zolotareva // Izvestija Mezhdunarodnoj akademii agrarnogo obrazovanija [Proceedings of the International Academy of Agrarian Education]. - 2021. - No. 56. - pp. 61-64 [in Russian]

9. Skvortsov, E. A. Perspektivy primeneniya tekhnologij iskusstvennogo intellekta v sel'skom khozjajstve regiona [Prospects for the use of artificial intelligence technologies in agriculture in the region] / E. A. Skvortsov // Ehkonomika regiona [The economy of the region]. - 2020. - Vol. 16. - No. 2. - pp. 563-576– DOI 10.17059/2020-2-17 [in Russian]

10. Primenenie aehrofotos"jomki s primeneniem mul'tispektral'noj kamery v komplekse s nejjroset'ju v sel'skom khozjajstve [The use of aerial photography using a multispectral camera in combination with a neural network in agriculture] / A. S. Sych, V. Balyk, A. I. Polyakov, et al. // Chronos. – 2019. – № 9(36). – pp. 26-28 [in Russian]