

Торайғыров университетінің хабаршысы
ҒЫЛЫМИ ЖУРНАЛЫ

НАУЧНЫЙ ЖУРНАЛ
Вестник Торайғыров университета

Торайғыров университетінің ХАБАРШЫСЫ

Энергетикалық сериясы
1997 жылдан бастап шығады



ВЕСТНИК Торайғыров университета

Энергетическая серия
Издается с 1997 года

ISSN 2710-3420

№ 3 (2022)

ПАВЛОДАР

НАУЧНЫЙ ЖУРНАЛ
Вестник Торайгыров университета

Энергетическая серия
выходит 4 раза в год

СВИДЕТЕЛЬСТВО

о постановке на переучет периодического печатного издания,
информационного агентства и сетевого издания

№ 14310-Ж

выдано

Министерство информации и общественного развития
Республики Казахстан

Тематическая направленность

публикация материалов в области электроэнергетики,
электротехнологии, автоматизации, автоматизированных и
информационных систем, электромеханики и теплоэнергетики

Подписной индекс – 76136

<https://doi.org/10.48081/AFHU6838>

Бас редакторы – главный редактор

Кислов А. П.
к.т.н., доцент

Заместитель главного редактора

Талипов О. М., *доктор PhD*

Ответственный секретарь

Калтаев А.Г., *доктор PhD*

Редакция алкасы – Редакционная коллегия

Клецель М. Я., *д.т.н., профессор*
Новожилов А. Н., *д.т.н., профессор*
Никитин К. И., *д.т.н., профессор (Россия)*
Никифоров А. С., *д.т.н., профессор*
Новожилов Т. А., *д.т.н., профессор*
Алиферов А.И., *д.т.н., профессор (Россия)*
Кошеков К.Т., *д.т.н., профессор*
Приходько Е.В., *к.т.н., профессор*
Оспанова Н. Н., *к.п.н., доцент*
Нефтисов А. В., *доктор PhD*
Омарова А.Р., *технический редактор*

За достоверность материалов и рекламы ответственность несут авторы и рекламодатели

Редакция оставляет за собой право на отклонение материалов

При использовании материалов журнала ссылка на «Вестник Торайгыров университета» обязательна

© Торайгыров университет

<https://doi.org/10.48081/DALJ8073>

***С. Е. Шарипова¹, А. С. Аканова², Н. Н. Оспанова³,
Е. Б. Шарипов⁴**

^{1,2,4}Агротехнический университет имени С. Сейфуллина,
Республика Казахстан, г. Нур-Султан;

³Торайғыров университет, Республика Казахстан, г. Павлодар

НОРМАЛИЗАЦИЯ ВХОДНЫХ ДАННЫХ ДЛЯ ПРОГНОЗИРОВАНИЯ УРОЖАЙНОСТИ ПШЕНИЦЫ

В данной статье рассматривается нормализация входных данных для прогнозирования урожайности пшеницы. Объектом исследования в рамках статьи являются входные данные нейронной сети. В рамках данного исследования рассматривается влияние таких параметров как, внесение удобрений, содержащих фосфор и азот на урожайность пшеницы с 1997 года по 2021 год.

В ходе исследования были рассмотрены методы нормализации такие как минимаксная, z-средняя и робостная нормализация. В последующем была осуществлена нормализация имеющихся входных данных. С помощью полученных данных была обучена нейронная сеть прогнозирования урожайности пшеницы. В итоге был проведен сравнительный анализ способов нормализации по таким показателям, как ошибка обучения и точность обучения.

По результатам исследования робостный метод нормализации привел к наилучшему результату: ошибка 0,01. Полученные результаты объясняются тем, что, когда к одному и тому же набору данных применяются разные типы методов нормализации с использованием одного и того же метода машинного обучения, результат может отличаться. Робостная нормализация проста в использовании и дает улучшенный результат по сравнению с другими методами нормализации.

Ключевые слова: нейронная сеть, нормализация, прогнозирование, зерновые культуры, фосфор, удобрение.

Введение

В настоящее время использование нейронных сетей в интеллектуальных системах для агропромышленности является все более широко используемым

инструментом. Преимущество использования нейронных сетей заключается в том, что с помощью сети обнаруживаются скрытые от обычного человека закономерности влияния различных факторов на урожайность культур. К тому же есть возможность повысить точность прогноза за счет эффективного подбора параметров обучения нейронной сети. А также на точность может положительно повлиять нормализация входных данных. В рамках данного исследования рассматривается влияние таких параметров как, внесение удобрений, содержащих фосфор и азот на урожайность пшеницы с 1997 года по 2021 год. Когда значения входных данных очень различаются, а иногда даже разного типа, то это может отрицательно влиять на обучение нейронной сети. В связи с этим перед обучением нейронной сети ее входные данные должны быть нормализованы.

Целью данного исследования является нормализация входных данных для прогнозирования урожайности пшеницы и выполнить сравнительный анализ нескольких методов нормализации, в результате которого будет выбран наилучший способ.

Задачи исследования:

- рассмотреть методы нормализации и применить их к входным данным;
- провести обучение нейронной сети прогнозирования урожайности пшеницы;
- провести сравнительный анализ способов нормализации по выбранным показателям.

В мировой науке нормализация данных была исследована как сеть обратимой автоматической нормализации выбора (RASN), объединяющую уровень нормализации и перенормировки для оценки и выбора модуля нормализации модели прогнозирования для теплицы в интеллектуальной сельскохозяйственной системе [1].

При сравнительном анализе использования различных методов нормализации данных для прогнозирования индексов фондового рынка, результаты гибридной нормализации оказались более заметными по сравнению с другими методами. Приблизительная точность, полученная с использованием метода гибридной нормализации, составила 71 %, 60 % и 71 % в разных архитектурах нейронных сетей [2].

В работах [3], [4] отмечается, что во время исследовании данных предварительная обработка является важным шагом при использовании нейронных сетей.

Учеными было доказано, что разные степени нормализации данных могут привести к разным значениям точности обучения [5]. В исследованиях [6] предлагается теория под названием нормализация встраивания,

которая не только делает обучение модели стабильным, но и повышает производительность моделей прогнозирования показателя кликабельности.

В работе [7] представлен метод предварительной обработки нормализации высоты глобального разделения (GPEN), который предназначен для смягчения влияния ограниченного динамического диапазона данных. Результаты эксперимента показывают более быструю сходимость модели, уменьшение ошибок районирования и улучшенные показатели классификации по сравнению со стандартными методами предварительной обработки нормализации.

В исследовании [8] рассматривается влияние метода нормализации на прогнозирование химической потребности в кислороде, а также эксперименты для определения прогноза точности обратного распространения ошибки до/после методов нормализации.

Ученые с помощью добавления слоя нормализации достигли точности предсказания 81,5 %, что также подтверждает положительное влияние применения нормализации данных в нейронных сетях [9]. Такое же влияние нормализации на точность обучения отражается в работе [10]: результаты данного исследования показывают, что входные параметры узла нейронной сети обрабатываются методом нормализации Z-показателя, что улучшает обучаемость модели, повышает точность эффекта прогнозирования и уменьшает ошибку прогнозирования.

Таким образом, учеными в области нормализации данных было проведено достаточное количество исследований, которое доказывает, что нормализация входных данных значительно повышает точность обучения и уменьшает ошибку прогнозирования. Но вопрос нормализации входных данных для прогнозирования урожайности пшеницы не был достаточно изучен и рассмотрен.

Материалы и методы

В исследовании нормализации данных имеются различные способы и методы, которые имеют существенное влияние на результат нормализации.

Нормализация Min-max. Данный метод широко используется при нормализации данных. По правилам данного метода сначала находятся минимальные и максимальные элементы списка, затем они приравниваются к 0 и 1 соответственно, а остальные элементы будут находиться между 0 и 1. Формула данного метода указана в таблице 1 [11].

Минимаксная нормализация не учитывает аномальные значения, это может отрицательно повлиять на анализ данных и на последующее обучение нейронной сети в целом. В Z-нормализации данный случай уже учитывается путем использования среднего значения. Формула данного метода также

указана в таблице 1 [11], как и формула третьего способа нормализации – робостная (межквартильная нормализация).

Робостная нормализация рассматривает межквартильные интервалы значений, в котором находятся «центральные» 50% данных набора. Эта величина не зависит от «нормальности» распределения наличия/отсутствия асимметрии и уже устойчива к выбросам.

Таблица 1 – методы нормализации

Метод	Формула
Min-max	
Z-среднее	
Робостная	

A_i – i элемент

μ – минимальное значение функции

σ – стандартное отклонение функции

В данном исследовании входными данными являются объем удобрения фосфором, азотом, а выходными данными – урожайность за последние 25 лет [12, 13].

Применив формулу минимаксной нормализации нормализуем входные данные, для наглядности можно увидеть в Таблице 2 результаты вычислений 2002-2004 годов.

Таблица 2 – Входные нормализованные данные

Год	Фосфор	Азот	Урожай	ММФ	ММА	ZФ	ZФ	РФ	РА
2002	198,5	803,4	10,8	0,33435	0,208828	0,008964	0,136985	1,604119	0,685614
2003	158,4	759,8	11,6	0,198509	0,177485	0,000477	0,127758	0,686499	0,466298
2004	169,1	702,4	8,4	0,234756	0,136223	0,002742	0,115609	0,93135	0,177565

Обозначения в таблице 2:

- ММФ – Минимаксная нормализация для фосфора
- ММА – Минимаксная нормализация для азота
- ZФ – Z-среднее для фосфора
- ZA – Z-среднее для азота
- РФ – робостная для фосфора
- РА – робостная для азота

Проведем нормализованные входные данные через нейронную сеть. Нейронная сеть имеет структуру как на рисунке 1.

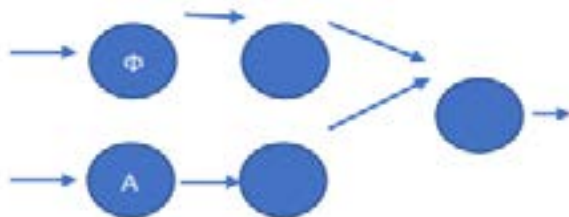


Рисунок 1 – Структура нейронной сети

Данная нейронная сеть имеет два нейрона на входе, где проходят такие данные, как объем внесенного фосфора и азота в почву. В скрытом слое входные данные нормализуются и передаются на выход, где результатом является урожайность пшеницы.

Результаты и обсуждение

Мы сравнили несколько методов нормализации входных данных по показателям ошибки и точности обучения нейронной сети. Результаты можно увидеть в таблице 3.

Таблица 3 – Результаты исследования

Метод	Ошибка	Точность
Без нормализации	0,025	78 %
Min-max	0,02	85 %
Z-среднее	0,01	87 %
Робостная	0,001	91 %

Полученные результаты объясняются тем, что, когда к одному и тому же набору данных применяются разные типы методов нормализации с использованием одного и того же метода машинного обучения, результат может отличаться. Робостная нормализация проста в использовании и дает улучшенный результат по сравнению с другими методами нормализации. В дальнейшем предлагается изучить данные методы нормализации для других случаев и большего количества данных.

Выводы

В ходе исследования были рассмотрены методы нормализации и была осуществлена нормализация имеющихся входных данных. С помощью полученных данных была обучена нейронная сеть прогнозирования урожайности пшеницы. В итоге был проведен сравнительный анализ способов нормализации по таким показателям, как ошибка обучения и точность обучения.

По результатам исследования можно сделать вывод о том, что применение метода межквартильной нормализации приводит к уменьшению ошибки обучения нейронной сети прогнозирования урожайности пшеницы.

СПИСОК ИСПОЛЬЗОВАННЫХ ИСТОЧНИКОВ

1 **Xuebo Jin et al.** A Reversible Automatic Selection Normalization (RASN) Deep Network for Predicting in the Smart Agriculture System [Text] // *Agronomy*, 2022. – № 12 (3). – P. 591.

2 **Kumari, B., Swarnkar, T.** Stock movement prediction using hybrid normalization technique and artificial neural network [Text] // *International Journal of Advanced Technology and Engineering Exploration*, 2021, – № 8 (83). – P.1336-1350. DOI:10.19101/IJATEE.2021.874387.

3 **Jain, S., Shukla, S., Wadhvani, R.** Dynamic selection of normalization techniques using data complexity measures [Text] // *Expert Systems with Applications*. – 2018. – Vol. 106. – P. 252–262.

4 **Kumari, B., Swarnkar, T.** A Comparative Analysis of Data Standardization Methods on Stock Movement. [Text] // *Intelligent and Cloud Computing. Smart Innovation, Systems and Technologies 2022*, – Vol. 286.

5 **Ramon-Julvez, U., Hernandez, M., Mayordomo E.** and Adni, Analysis of the Influence of Diffeomorphic Normalization in the Prediction of Stable [Text] // *VS Progressive MCI Conversion with Convolutional Neural Networks IEEE 17th International Symposium on Biomedical Imaging*. – 2020. – P. 1120–1124, doi: 10.1109/ISBI45749.2020.9098445.

6 **Yi, J., Kim, B. and Chang, B.** Embedding Normalization: Significance Preserving Feature Normalization for Click-Through Rate [Text] // *Prediction. International Conference on Data Mining Workshops*. – 2021. – P. 75-84, doi: 10.1109/ICDMW53433.2021.00016.

7 **Fafard, A., Van Aardt, J., Coletti, M. and Page, D. L.** Global Partitioning Elevation Normalization Applied to Building Footprint Prediction [Text] // *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*. – 2020. – № 13. – P. 3493-3502. doi: 10.1109/JSTARS.2020.3002502.

8. **Vijayabhanu, R., Radha V.** Statistical Normalization techniques for the prediction of COD level for an anaerobic wastewater treatment plant [Text] // *CCSEIT 12: Proceedings of the Second International Conference on Computational Science, Engineering and Information Technology*. 26 October. – 2012. – P. 232–236.

9 **Zeghoud, S., Ali, S. G., Ertugrul, E.** et al. Real-time spatial normalization for dynamic gesture classification [Text] // *The Visual Computer*. – 2022. – № 38, – P. 1345–1357.

10 Dingran Zhang et al. The Research on Improving the Precision of The Polymer IFHI By BP Neural Network of The Method Data Normalization [Text] // IOP Conference Series: Earth and Environmental Science – 2021. Vol. 692. – P. 032067.

11 **Shalabi, L.A.** Data Mining: A Preprocessing Engine [Text] // Journal of Computer Science, 2006. Vol 2. – P. 735–739.

12 stats.gov.kz [Electronic resource]

13 **Какезанова З.** Влияние биогумуса «Павлодарский» на свойства почвы [Text] // XI Международная научно-практическая конференция «Аграрная наука – сельскому хозяйству» 2016. – № 2. – С. 109–110.

REFERENCE

1 **Xuebo Jin et al.** A Reversible Automatic Selection Normalization (RASN) Deep Network for Predicting in the Smart Agriculture System [Text] // Agronomy. 2022. – № 12 (3). – P. 591.

2 **Kumari, B., Swarnkar, T.** Stock movement prediction using hybrid normalization technique and artificial neural network [Text] // International Journal of Advanced Technology and Engineering Exploration. – 2021; – № 8(83). – P.1336-1350. DOI:10.19101/IJATEE.2021.874387.

3 **Jain, S., Shukla, S., Wadhvani, R.** Dynamic selection of normalization techniques using data complexity measures [Text] // Expert Systems with Applications. – 2018. – Vol. 106. – P. 252-262.

4 **Kumari, B., Swarnkar, T.A.** Comparative Analysis of Data Standardization Methods on Stock Movement. [Text] // Intelligent and Cloud Computing. Smart Innovation, Systems and Technologies. – 2022, –Vol. 286.

5 **Ramon-Julvez, U., Hernandez, M., Mayordomo E.** and Adni. Analysis of the Influence of Diffeomorphic Normalization in the Prediction of Stable [Text] // VS Progressive MCI Conversion with Convolutional Neural Networks IEEE 17th International Symposium on Biomedical Imaging. – 2020. – P. 1120-1124, doi: 10.1109/ISBI45749.2020.9098445.

6 **Yi, J., Kim, B. and Chang, B.** Embedding Normalization: Significance Preserving Feature Normalization for Click-Through Rate [Text] // Prediction. International Conference on Data Mining Workshops. – 2021. – P. 75-84, doi: 10.1109/ICDMW53433.2021.00016.

7 **Fafard, A., Van Aardt, J., Coletti, M. and Page, D. L.** Global Partitioning Elevation Normalization Applied to Building Footprint Prediction [Text] // IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing. – 2020. – № 13. – P. 3493-3502. doi: 10.1109/JSTARS.2020.3002502.

8 **Vijayabhanu, R., Radha V.** Statistical Normalization techniques for the prediction of COD level for an anaerobic wastewater treatment plant [Text] //

CCSEIT 12: Proceedings of the Second International Conference on Computational Science, Engineering and Information Technology. 26 October. – 2012. – P. 232–236.

9 **Zeghoud, S., Ali, S. G., Ertugrul, E.** et al. Real-time spatial normalization for dynamic gesture classification [Text] // The Visual Computer. – 2022. – № 38, – P. 1345–1357.

10 Dingran Zhang et al. The Research on Improving the Precision of The Polymer IFHI By BP Neural Network of The Method Data Normalization [Text] // IOP Conference Series: Earth and Environmental Science– 2021. Vol. 692.– P. 032067.

11 **Shalabi L.A. Data Mining: A** Preprocessing Engine [Text] // Journal of Computer Science – 2006. Vol 2. – P. 735–739.

12 stats.gov.kz [Electronic resource]

13 **Kakezhanova Z.** Vliyanie biogumusa «Pavlodarskiy» na svoystva pochvy [The influence of «Pavlodarskiy» biohumus on soil properties] // XI Mezhdunarodnaya nauchno-prakticheskaya konferenciya «Agrarnaya nauka – sel'skomu hozyajstvu» (XI International Scientific and Practical Conference «Agrarian Science for Agriculture» – 2016. – № 2. – P. 109–110.

Материал поступил в редакцию 15.09.22.

*С. Е. Шарипова¹, А. С. Аканова², Н. Н. Оспанова³, Е. Б. Шарипов⁴

^{1,2,4}С. Сейфуллин атындағы агротехникалық университет,

Қазақстан Республикасы, Нур-Сұлтан қ.,

³Торайғыров университет, Қазақстан Республикасы, Павлодар қ.

Материал баспаға 15.09.22 түсті.

БИДАЙ ӨНІМДІЛІГІН БОЛЖАУ ҮШІН КІРІС ДЕРЕКТЕРДІ ҚАЛЫПҚА КЕЛТІРУ

Бұл мақалада бидай өнімділігін болжау үшін деректерді қалыпқа келтіру қарастырылады. Мақала шеңберіндегі зерттеу нысаны нейрондық желінің кірісі болып табылады. Осы зерттеу аясында 1997 жылдан 2021 жылға дейін бидай өнімділігіне құрамында фосфор мен азот бар тыңайтқыштарды енгізу сияқты параметрлердің әсері қарастырылады.

Зерттеу барысында минимаксті, z-орташа және робастикалық сияқты қалыпқа келтіру әдістері қарастырылды. Одан кейін қолда бар деректерді қалыпқа келтіру жүзеге асырылды. Алынған мәліметтер арқылы бидай өнімділігін болжаудың нейрондық желісі оқытылды. Нәтижесінде оқу қателігі және оқытудың дәлдігі сияқты

көрсеткіштер бойынша қалыпқа келтіру әдістеріне салыстырмалы талдау жүргізілді.

Зерттеу нәтижелері бойынша қалыпқа келтірудің робастикалық әдісі жақсы нәтижеге әкелді: қате 0,01. Нәтижелер бірдей мәліметтер жиынтығына бірдей Машиналық оқыту әдісін қолдана отырып, қалыпқа келтіру әдістерінің әр түрлі түрлері қолданылған кезде, нәтиже әр түрлі болуы мүмкін екендігімен түсіндіріледі. Тұрақты қалыпқа келтіруді қолдану оңай және басқа қалыпқа келтіру әдістерімен салыстырғанда жақсартылған нәтиже береді.

Кілтті сөздер: нейрондық желі, қалыпқа келтіру, болжау, дақылдар, фосфор, тыңайтқыш.

*S. Sharipova¹, A. Akanova², N. Ospanova³, Ye. Sharipov⁴

^{1,2,4}S. Seifullin Kazakh Agrotechnical University,

Republic of Kazakhstan, Nur-Sultan;

³Toraighyrov University, Republic of Kazakhstan, Nur-Sultan.

Material received on 15.09.22.

NORMALIZATION OF INPUT DATA FOR WHEAT YIELD PREDICTION

This article discusses the normalization of input data for predicting wheat yield. The object of research within the framework of the article is the input data of the neural network. This study examines the impact of parameters such as the application of fertilizers containing phosphorus and nitrogen on wheat yields from 1997 to 2021.

In the course of the study, normalization methods such as minimax, z-score and robust normalization were considered and the normalization of the available input data was carried out. Using the obtained data, a neural network was trained to predict wheat yield. As a result, a comparative analysis of normalization methods was carried out in terms of such indicators as training error and training accuracy.

According to the results of the study, the robust normalization method led to the best result: an error of 0.01. The results obtained are explained by the fact that when different types of normalization methods are applied to the same dataset using the same machine learning method, the result may differ. Robust normalization is easy to use and gives better results than other normalization methods.

Keywords: neural network, normalization, forecasting, crops, phosphorus, fertilizer.

Теруге 15.09.2022 ж. жіберілді. Басуға 30.09.2022 ж. қол қойылды.

Электронды баспа

3 Mb RAM

Шартты баспа табағы 19,8. Таралымы 300 дана. Бағасы келісім бойынша.

Компьютерде беттеген: А. К. Мыржикова

Корректор: А. Р. Омарова

Тапсырыс № 3989

Сдано в набор 15.09.2022 г. Подписано в печать 30.09.2022 г.

Электронное издание

3 Mb RAM

Усл. печ. л. 19,8. Тираж 300 экз. Цена договорная.

Компьютерная верстка: А. К. Мыржикова

Корректор: А. Р. Омарова

Заказ № 3989

«Toraighyrov University» баспасынан басылып шығарылған

Торайғыров университеті

140008, Павлодар қ., Ломов к., 64, 137 каб.

«Toraighyrov University» баспасы

Торайғыров университеті

140008, Павлодар қ., Ломов к., 64, 137 каб.

67-36-69

E-mail: kereku@tou.edu.kz

www.vestnik-energy.tou.edu.kz