

## ИСПОЛЬЗОВАНИЕ ЯЗЫКА R ДЛЯ РЕШЕНИЯ ЗАДАЧ КЛАССИФИКАЦИИ СОЦИАЛЬНО-ЭКОНОМИЧЕСКИХ ДАННЫХ С ПОМОЩЬЮ ИСКУССТВЕННОГО ИНТЕЛЛЕКТА

Н. Ю. Рудь • А. Н. Можельский

**Аннотация.** Недостаток структурированного подхода к применению машинного обучения в социально-экономической сфере серьезно замедляет инновационное развитие. Отсутствие четкой методологии внедрения искусственного интеллекта, основанного на машинном обучении, препятствует не только ускорению инновационных процессов, но и общему повышению научно-технического потенциала страны. Проблема заключается не только в отсутствии готовых решений, но и в недостатке понимания того, как эффективно адаптировать существующие алгоритмы машинного обучения к специфике социально-экономических задач. Классификация данных с использованием нейронных сетей представляет собой современный метод машинного обучения, который позволяет решать разнообразные задачи, включая распознавание изображений, анализ текстов, прогнозирование финансовых тенденций и диагностику заболеваний. В статье осуществляются систематизация и классификация аспектов машинного обучения, а также подчеркивается важность ускорения разработки и внедрения алгоритмов, которые являются основой искусственного интеллекта. Это необходимо для повышения эффективности управления социально-экономическими процессами.

**Ключевые слова:** искусственный интеллект; язык R; подход; метод; социально-экономический процесс.

### ВВЕДЕНИЕ

Множество вопросов, касающихся взаимодействия естественного и искусственного интеллекта, остаются на сегодняшний день открытыми. Ответы на эти вопросы меняются по мере прогресса науки в целом, что, в свою очередь, приводит к изменению подходов к пониманию и внедрению систем искусственного интеллекта, в том числе в социально-экономические сферы общества. Тем не менее основой для искусственного интеллекта является неоспоримое развитие современных информационных технологий и создание новейших вычислительных платформ.

Классификация данных с использованием нейронных сетей – это мощный и гибкий метод машинного обучения, который позволяет эффективно решать задачи, связанные с классификацией, такие как распознавание изображений, анализ текста и многое другое. Нейронные сети могут обучаться на больших объемах данных и выявлять сложные паттерны в данных. Также способны обеспечивать высокую точность классификации, что делает их популярным выбором для многих приложений в области машинного обучения.

### ОБЗОР ВОЗМОЖНОСТЕЙ ЯЗЫКА R

Язык R [1] разработан специально для статистической обработки и анализа данных, что делает его мощным инструментом для статистиков и аналитиков. В R доступно множество пакетов (более 21 000 на конец 2024 года), которые расширяют его функциональность и позволяют выполнять широкий спектр задач, от визуализации данных до машинного обучения. Пользователи могут легко создавать свои собственные функции и пакеты, что позволяет адаптировать язык под специфические задачи. Язык R предлагает мощные инструменты для визуализации данных, такие как `ggplot2`, которые позволяют создавать высококачественные графики и диаграммы. R может быть интегрирован с другими языками программирования,

такими как Python, C++ и Java, что позволяет использовать его в более широких контекстах. Язык R имеет активное сообщество пользователей и разработчиков, что обеспечивает обширную документацию, учебные материалы и форумы для получения помощи. R – это бесплатный язык программирования с лицензией GNU. Язык существует для всех популярных платформ и операционных систем, поэтому его можно установить на любой компьютер. R является эффективным инструментом для статистического анализа и визуализации данных, особенно в академической и исследовательской среде.

Базовая версия R включает в себя множество стандартных статистических методов, таких как линейная и нелинейная регрессия, анализ дисперсии (ANOVA), корреляционный анализ и т. д. В R есть базовые функции для создания графиков, такие как «plot()», «hist()», «boxplot()», которые позволяют визуализировать данные, но они могут быть ограничены по функциональности и эстетике. Базовая версия включает функции для работы с данными, такие как «data.frame», «list», и базовые операции с ними, но не предоставляет более сложных инструментов для манипуляции данными. В базовой версии доступны основные методы для обработки и анализа данных, такие как фильтрация, агрегация и сводные таблицы, но без расширенных возможностей.

Дополнительные пакеты, такие как «caret», «randomForest», «glmnet», предлагают более сложные и специализированные методы статистического анализа и машинного обучения. Пакеты «ggplot2», «lattice» и «plotly» предоставляют более мощные и гибкие средства для визуализации данных, позволяя создавать сложные и интерактивные графики. Пакеты «dplyr» и «tidyr» предлагают удобные и эффективные функции для манипуляции данными, включая фильтрацию, сортировку, агрегацию и преобразование данных. Существуют пакеты, которые фокусируются на конкретных областях, таких как биоинформатика («Bioconductor»), текстовый анализ («tm», «text»), временные ряды («forecast») и многие другие. Некоторые пакеты позволяют интегрировать R с другими языками программирования (например, «reticulate» для Python) или инструментами (например, «shiny» для создания веб-приложений). Большинство пакетов имеют свою документацию и активные сообщества, что позволяет пользователям получать помощь и находить примеры использования.

В R имеется достаточно большое количество дополнительных пакетов, которые обеспечивают возможности для работы с нейронными сетями и глубоким обучением.

Представим краткий обзор пакетов, доступных на сегодняшний день, которые расширяют функциональность языка R в области создания нейросетей для выполнения задач классификации.

Таблица

Пакет	Назначение
1 deepNN	Создание и обучение глубоких нейронных сетей. Он предлагает пользователям инструменты для построения моделей, которые могут обрабатывать сложные данные и выполнять задачи, такие как классификация, регрессия и другие виды предсказаний
2 kerasformula	Упрощение процесса создания и обучения моделей глубокого обучения с использованием библиотеки Keras. Предоставляет удобный интерфейс, позволяющий пользователям использовать формулы для описания моделей, что делает его более доступным для тех, кто знаком с традиционными методами регрессии и анализа данных в R
3 keras	Интерфейс к библиотеке Keras, которая предназначена для разработки и обучения моделей глубокого обучения. Предоставляет пользователям возможность строить нейронные сети и выполнять обучение с использованием простого и интуитивно понятного синтаксиса R
4 kerasR	Является интерфейсом к библиотеке Keras, которая позволяет пользователям создавать и обучать модели глубокого обучения. Основная цель kerasR – предоставить удобный и интуитивно понятный способ работы с нейронными сетями в R, используя возможности Keras и TensorFlow
5 tensorflow	Является интерфейсом к библиотеке TensorFlow, предназначенной для разработки и выполнения моделей машинного обучения и глубокого обучения. Предоставляет пользователям возможность использовать мощные инструменты TensorFlow в среде R,

	Пакет	Назначение
6	nnet	что позволяет интегрировать машинное обучение в анализ данных и статистические вычисления. Предоставляет функции для создания и обучения нейронных сетей, а также для выполнения различных задач машинного обучения. Он является частью более широкого набора инструментов для статистического анализа и моделирования в R и сосредоточен на простоте использования и эффективности
7	neuralnet	Создание и обучение нейронных сетей, а также для выполнения задач регрессии и классификации. Предоставляет пользователям возможность строить многослойные нейронные сети с различными параметрами и функциями активации
8	NeuralNetTools	Упрощение работы с нейронными сетями, созданными с помощью других пакетов, таких как nnet и neuralnet. Предоставляет функции для визуализации, оценки и интерпретации нейронных сетей
9	fastai	Обертка над библиотекой fastai, изначально разработанной для Python, и для упрощения процесса обучения и использования моделей глубокого обучения. Предоставляет высокоуровневый интерфейс для работы с различными задачами машинного обучения и глубокого обучения, включая классификацию, регрессию и обработку изображений
10	torch	Интерфейс к библиотеке Torch, которая предназначена для выполнения вычислений с использованием графов вычислений и предоставляет мощные инструменты для разработки и обучения моделей глубокого обучения. Он обеспечивает высокоуровневый доступ к возможностям, которые предлагает Torch, и позволяет пользователям R эффективно использовать GPU для ускорения вычислений
11	nntlib2Rcpp	Упрощение интеграции и использования библиотек глубокого обучения, написанных на C++, в рамках R. Предоставляет интерфейс для работы с нейронными сетями и алгоритмами машинного обучения, реализованными на C++, что позволяет пользователям R эффективно использовать высокопроизводительные вычисления
12	nntrf	Работа с нейронными сетями, в частности, для выполнения задач регрессии и классификации с использованием методов глубокого обучения. Предоставляет инструменты для построения, обучения и оценки нейронных сетей, что позволяет пользователям эффективно применять эти модели для анализа данных
13	nnetsauce	Создание и обучение нейронных сетей с использованием подхода, основанного на ансамблях. Предоставляет инструменты для построения сложных моделей, которые комбинируют несколько нейронных сетей, что позволяет улучшить точность предсказаний и повысить устойчивость моделей
14	RSNNS	Интерфейс для работы с нейронными сетями, используя библиотеку Stuttgart Neural Network Simulator (SNNS). Этот пакет позволяет пользователям создавать, обучать и тестировать нейронные сети различной архитектуры и назначения
15	FCNN4R	Создание и обучение полносвязных нейронных сетей (Fully Connected Neural Networks, FCNN). Предоставляет пользователям инструменты для работы с нейронными сетями с различными архитектурами и функциями активации
16	softmaxreg	Выполнение регрессии с использованием функции активации softmax, что делает его особенно полезным для задач многоклассовой классификации
17	deeplearning	Разработка и обучение глубоких нейронных сетей. Предоставляет пользователям инструменты для создания, настройки и обучения моделей глубокого обучения, что делает его полезным для задач, связанных с анализом данных, классификацией и прогнозированием
18	deepnet	Работа с глубокими нейронными сетями и предоставляет инструменты для создания, обучения и оценки моделей глубокого обучения. Он ориентирован на исследователей и разработчиков, желающих применять методы глубокого обучения к различным задачам
19	rbm	Работа с ограниченными бозмановскими машинами (Restricted Boltzmann Machines, RBM). RBM – это тип вероятностной графической модели, которая может использоваться для обучения представлений данных и для выполнения задач, связанных с машинным обучением, таких как уменьшение размерности, классификация и генерация данных
20	RcppDL	Упрощение работы с глубоким обучением (Deep Learning) в среде R, используя C++ для повышения производительности. Предоставляет интерфейс для создания и обучения нейронных сетей, а также для выполнения операций, связанных с глубоким обучением, что позволяет пользователям использовать преимущества скорости и эффективности C++ в своих R-приложениях

	Пакет	Назначение
21	h2oDeepLearning	Является частью платформы H2O.ai и для выполнения задач глубокого обучения. Предоставляет инструменты для создания, обучения и оценки нейронных сетей, а также интеграцию с другими функциональными возможностями платформы H2O
22	mxnet	Интерфейс к популярной библиотеке глубокого обучения MXNet, разработанной для создания и обучения нейронных сетей. Предоставляет пользователям мощные инструменты для работы с глубоким обучением, включая поддержку различных архитектур нейронных сетей и эффективное выполнение вычислений
23	autoencoder	Инструменты для создания и работы с автокодировщиками (autoencoders) – типом нейронных сетей, используемых для обучения эффективных представлений данных, обычно для задач уменьшения размерности или обучения без учителя
24	SAENET	Работа со структурами автокодировщиков, которые используют механизм саморегуляции для обучения представлений данных. Он фокусируется на создании и обучении нейронных сетей, которые могут эффективно извлекать признаки из входных данных, а также восстанавливать их
25	nnetpredint	Создание и оценка предсказаний с использованием нейронных сетей, разработанных с помощью пакета nnet. Основная цель этого пакета – предоставление инструментов для вычисления предсказаний и интервалов предсказаний, что может быть полезно в задачах регрессии и классификации
26	deepr	Работа с глубоким обучением и предоставляет пользователям инструменты для создания, обучения и оценки моделей глубоких нейронных сетей. Он упрощает процесс разработки и внедрения моделей глубокого обучения, делая его доступным даже для тех, кто не имеет глубоких знаний в этой области
27	pnn	Работа с вероятностными нейронными сетями (Probabilistic Neural Networks, PNN). PNN – это тип нейронной сети, который используется для классификации данных, основываясь на вероятностных оценках. Этот пакет предоставляет функции для построения и обучения PNN, а также для предсказания классов на основе обученной модели
28	validann	Валидация и оценка производительности искусственных нейронных сетей. Предоставляет инструменты для проверки качества моделей, созданных с использованием нейронных сетей, а также для анализа их точности и надежности
29	neural	Создание и обучение искусственных нейронных сетей. Предоставляет пользователям инструменты для построения моделей нейронных сетей, которые можно применять в задачах регрессии и классификации
30	learNN	Обучение и применение нейронных сетей. Предоставляет инструменты для создания, настройки и оценки моделей нейронных сетей, что делает его полезным для задач машинного обучения
31	GMDH	Реализация метода генетической модульной зависимости (GMDH, Group Method of Data Handling), который используется для построения моделей регрессии и классификации. Этот метод позволяет автоматически создавать модели на основе предоставленных данных, выявляя зависимости между входными и выходными переменными.
32	elmNNRcpp	Реализация метода Extreme Learning Machine (ELM) с использованием C++ для повышения производительности. ELM – это метод обучения нейронных сетей, который отличается высокой скоростью обучения и простотой в реализации. Этот пакет позволяет пользователям эффективно применять ELM для задач регрессии и классификации
33	brnn	Реализация метода "Bayesian Regularized Neural Networks" (BRNN), который является подходом к обучению нейронных сетей с регуляризацией на основе байесовских методов. Этот пакет позволяет строить нейронные сети, которые могут эффективно обрабатывать задачи регрессии и классификации
34	AMORE	Реализация нейронных сетей и методов машинного обучения, особенно для задач регрессии и классификации. Он предлагает гибкие инструменты для создания и обучения нейронных сетей, а также для обработки и анализа данных
35	simpleNeural	Создание и обучение простых нейронных сетей. Предоставляет пользователям удобный интерфейс для работы с нейронными сетями, что делает его подходящим для тех, кто хочет быстро начать использовать методы машинного обучения без необходимости в глубоком понимании сложных архитектур и алгоритмов

## ПРИМЕР ИСПОЛЬЗОВАНИЯ R

В настоящей статье в качестве примера использования рассматриваемой технологии для решения задач классификации различных данных, проведена классификация цитрусовых по следующим параметрам: цвет плода RGB, диаметр и вес<sup>1</sup>. Построены модели нейронных сетей с использованием дополнительных пакетов `nnet` [2] и `neuralnet` [3]. Проведено сравнение по времени создания модели, точности, чувствительности и специфичности.

Скачиваем файл `citrus.csv` по вышеприведенной ссылке и помещаем его в рабочую директорию.

С помощью команды

```
df <- read.table("citrus.csv", header = T, sep = ",")
```

создаем dataframe `df`.

Подгружаем библиотеку `caret` [4] (инструмент для упрощения процесса создания, оценки и настройки моделей машинного обучения) следующей командой:

```
library(caret)
```

Для обеспечения воспроизводимости результатов устанавливаем начальное значение генератора случайных чисел, введя команду

```
set.seed(111)
```

Так как dataframe содержит 10 000 записей, и построение моделей может занять значительное время, уменьшим количество записей до 1000, выбрав случайным образом с помощью следующей команды:

```
df <- df[sample(nrow(df), nrow(df) * 0.1),]
```

Создаем тренировочную и тестовую выборки:

```
indexes <- createDataPartition(df$name, p = 0.8, list = F)
train <- df[indexes,]
test <- df[-indexes,]
```

Подгружаем библиотеку `neuralnet` для создания и обучения искусственной нейронной сети:

```
library(neuralnet)
```

Проводим обучение сети и создаем модель следующей командой:

```
system.time(model <- neuralnet
             (name ~., train, hidden = c(3), linear.output = F, stepmax=1e7))
```

Для оценки времени выполнения данная команда была выполнена внутри функции `system.time()`, которая показывает время выполнения команды. В данном случае время составило 93.014 сек.

Изобразим полученную модель командой `plot(model)` (рис. 1). Мы видим, что у нас 5 нейронов на входном слое, 3 нейрона в скрытом слое и 2 нейрона в выходном слое. Для выполнения прогнозирования на основе уже обученной нейронной сети выполним следующие команды:

```
xtest <- test[, -1]
ytest <- test[, 1]
ypred <- compute(model, xtest)
yhat <- ypred$net.result
yhat <- data.frame("yhat"=ifelse(max.col(yhat[, 1:2]) == 1, "grapefruit", "orange"))
```

---

<sup>1</sup> Dataset с сайта <https://www.kaggle.com/datasets/joshmcadams/oranges-vs-grapefruit>.

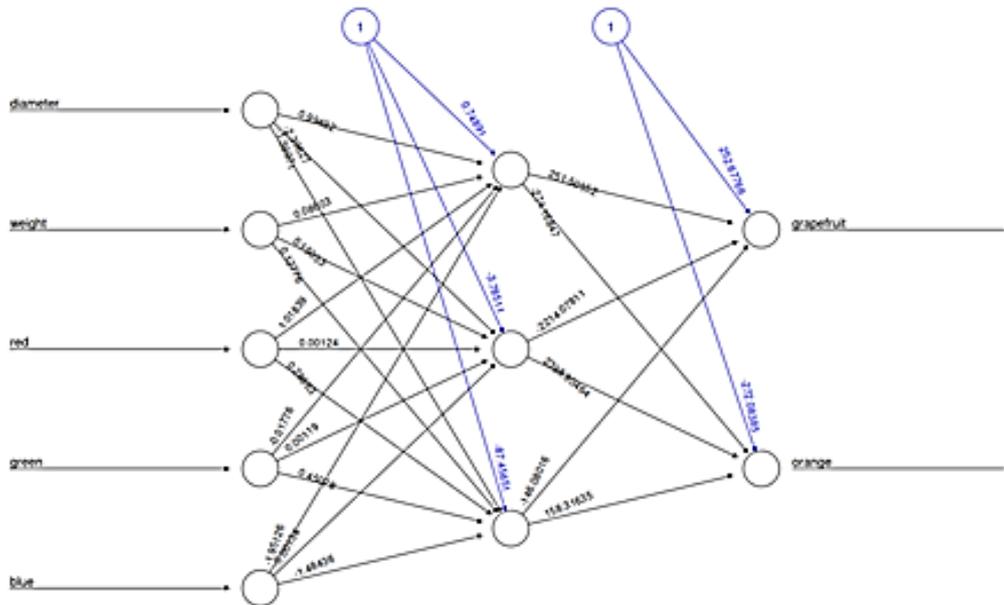


Рис. 1 Архитектура нейронной сети

Для предоставления сводной информации о правильности предсказаний модели и сравнения предсказанных классов с фактическими классами выполним следующую команду:

```
cm <- confusionMatrix(as.factor(ytest), as.factor(yhat$yhat))
print(cm)
```

Получаем следующий результат:

Confusion Matrix and Statistics

	Reference	
Prediction	grapefruit	orange
grapefruit	90	3
orange	3	103

Accuracy: 0.9698

95 % CI: (0.9355, 0.9889)

No Information Rate: 0.5327

P-Value [Acc > NIR]: <2e-16

Kappa: 0.9394

Mcnemar's Test P-Value: 1

Sensitivity: 0.9677

Specificity: 0.9717

Pos Pred Value: 0.9677

Neg Pred Value: 0.9717

Prevalence: 0.4673

Detection Rate: 0.4523

Detection Prevalence: 0.4673

Balanced Accuracy: 0.9697

'Positive' Class: grapefruit

Точность модели составила – 0.9698 с доверительным интервалом 95 % CI: (0.9355, 0.9889), чувствительность – 0.9677 и специфичность – 0.9717.

Теперь на тех же исходных данных построим еще одну модель с помощью пакета `nnet` и сравним полученные результаты. Подгружаем библиотеку с помощью команды:

```
library(nnet)
```

Преобразуем переменную `name` в факторную и проводим обучение модели:

```
df$name <- as.factor(df$name)
train <- df[indexes,]
test <- df[-indexes,]
system.time(newNet <- nnet(name ~ ., data = train, size = 3, rang = 1.0e-06,
decay = 1.0e-04, maxit = 4000))
```

Данная библиотека справилась с обучением модели гораздо быстрее – за 0.146 сек. Выводим на экран архитектуру полученной модели командой (рис. 2).

```
plotnet(newNet)
```

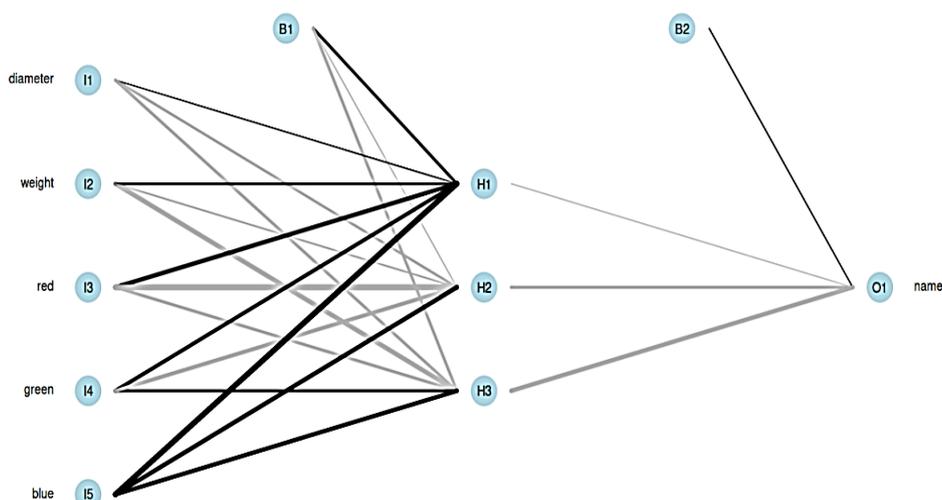


Рис. 2 Архитектура нейронной сети

Архитектура нейронной сети также содержит 5 нейронов на входном слое, 3 нейрона в скрытом слое и 1 нейрон в выходном слое. Для выполнения прогнозирования на основе уже обученной нейронной сети выполним следующие команды:

```
myPrediction <- predict(newNet, newdata = test, type = "class")
cm <- confusionMatrix(as.factor(myPrediction), as.factor(test$name))
print(cm)
```

Получаем следующий результат:

Confusion Matrix and Statistics

Prediction	Reference	
	grapefruit	orange
grapefruit	97	4
orange	1	97

Accuracy: 0.9749

95 % CI: (0.9423, 0.9918)

No Information Rate: 0.5075

P-Value [Acc > NIR]: <2e-16

Карра: 0.9498

Mcnemar's Test P-Value: 0.3711

Sensitivity: 0.9898  
 Specificity: 0.9604  
 Pos Pred Value: 0.9604  
 Neg Pred Value: 0.9898  
 Prevalence: 0.4925  
 Detection Rate: 0.4874  
 Detection Prevalence: 0.5075  
 Balanced Accuracy: 0.9751

'Positive' Class: grapefruit

Точность модели составила – 0.9749 с доверительным интервалом 95 % CI: (0.9423, 0.9918), чувствительность – 0.9898 и специфичность – 0.9604.

### ЗАКЛЮЧЕНИЕ

Как мы видим, модели, созданные с помощью пакетов `nnet` и `neuralnet`, показали примерно одинаковый результат. Хотя по скорости создания модели, точности и чувствительности пакет `nnet` оказался немного впереди.

Таким образом, создание и обучение нейросетевых моделей на языке R с помощью дополнительных пакетов не является сложной задачей, овладеть данной технологией может каждый, и не обязательно нужно быть программистом.

В контексте данного исследования авторы считают полезным упомянуть работы [5–12] из смежных областей применения нейронных сетей.

### СПИСОК ЛИТЕРАТУРЫ / REFERENCES

- [1] R Core Team. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL: <https://www.R-project.org/> (visited on 19.10.2024).
- [2] Venables W. N., Ripley B. D. Modern Applied Statistics with S, 4th ed by W. N. Venables and B. D. Ripley. Springer. 2002. DOI [10.1007/978-0-387-21706-2](https://doi.org/10.1007/978-0-387-21706-2).
- [3] Fritsch S., Guenther F., Wright M. (2019). `_neuralnet` Training of Neural Networks\_. R package version 1.44.2. URL <https://CRAN.R-project.org/package=neuralnet> (visited on 19.10.2024).
- [4] Kuhn M. Building predictive models in R using the caret package // Journal of Statistical Software. 2008. 28(5). P. 1–26. DOI [10.18637/jss.v028.i05](https://doi.org/10.18637/jss.v028.i05) (visited on 19.10.2024). DOI [10.18637/jss.v028.i05](https://doi.org/10.18637/jss.v028.i05).
- [5] Маркина Н. В., Беленкова Э. И., Диденко Г. А., Касюк С. Т., Степанова О. А., Шамаева Т. Н. Искусственный интеллект в диагностике пневмонии на рентгенологических снимках. Челябинск: ЮУМГУ, 2023. 142 с. DOI [10.22138/2500-0918-2023-20-4-162-166](https://doi.org/10.22138/2500-0918-2023-20-4-162-166). EDN [VKMMOT](https://doi.org/10.22138/2500-0918-2023-20-4-162-166). [[Markina N. V., Belenkova E. I., Didenko G. A., Kasyuk S. T., Stepanova O. A., Shamaeva T. N. Artificial Intelligence in the Diagnosis of Pneumonia on X-ray Images. Chelyabinsk: South Ural State University, 2023. DOI [10.22138/2500-0918-2023-20-4-162-166](https://doi.org/10.22138/2500-0918-2023-20-4-162-166). EDN [VKMMOT](https://doi.org/10.22138/2500-0918-2023-20-4-162-166). (In Russian).]]
- [6] Касюк С. Т., Диденко Г. А., Степанова О. А. Классификация данных медицинских исследований с использованием языка R // Современная наука: актуальные проблемы теории и практики. Серия: Естественные и технические науки. 2022. № 09. С. 96–102. DOI [10.37882/2223-2966.2022.09.14](https://doi.org/10.37882/2223-2966.2022.09.14). EDN [BSJRLE](https://doi.org/10.37882/2223-2966.2022.09.14). [[Kasyuk S. T., Didenko G. A., Stepanova O. A. “Classification of medical research data using the R language” // Modern Science: Current Problems of Theory and Practice. Series: Natural and Technical Sciences. 2022. No. 09, pp. 96–102. DOI [10.37882/2223-2966.2022.09.14](https://doi.org/10.37882/2223-2966.2022.09.14). EDN [BSJRLE](https://doi.org/10.37882/2223-2966.2022.09.14). (In Russian).]]
- [7] Коровин Е. А., Чиглинцева С. А., Сазонова Е. Ю., Сметанина О. Н. Медицинская рекомендательная система на основе автоматического извлечения знаний из текстов // СИИТ. 2024. Т. 6. № 4(19). С. 111–121. EDN [OTVTXR](https://doi.org/10.22138/2500-0918-2023-20-4-111-121). [[Korovina E. A., Chiglintseva S. A., Sazonova E. Yu., Smetanina O. N. “Medical recommendation system based on automatic knowledge extraction from texts” // СИИТ. 2024. Vol. 6, no. 4(19), pp. 111–121. EDN [OTVTXR](https://doi.org/10.22138/2500-0918-2023-20-4-111-121). (In Russian).]]
- [8] Муравьева Е. А., Николаева А. И. Разработка нейронной сети для управления процессом пиролиза дихлорэтана с целью повышения эффективности процесса // СИИТ. 2024. Т. 6. № 4(19). С. 69–76. EDN [BJQURV](https://doi.org/10.22138/2500-0918-2023-20-4-69-76). [[Muravyova E. A., Nikolaeva A. I. “Development of a neural network for controlling the dichloroethane pyrolysis process to improve the efficiency of the process” // СИИТ. 2024. Vol. 6, no. 4(19), pp. 69–76. EDN [BJQURV](https://doi.org/10.22138/2500-0918-2023-20-4-69-76). (In Russian).]]
- [9] Зиновьев М. С., Нургаянова О. С. Прогнозирование вероятности развития диабетической ретинопатии у пациентов с сахарным диабетом: анализ методов машинного обучения // СИИТ. 2024. Т. 6. № 3(18). С. 95–101. EDN [VLFFLP](https://doi.org/10.22138/2500-0918-2023-20-3-95-101). [[Zinoviev M. S., Nurgayanova O. S. “Predicting the probability of developing diabetic retinopathy in patients with diabetes mellitus: analysis of machine learning methods” // СИИТ. 2024. Vol. 6, no. 3(18), pp. 95–101. EDN [VLFFLP](https://doi.org/10.22138/2500-0918-2023-20-3-95-101). (In Russian).]]

- [10] Макарова Е. А., Габдуллина Э. Р., Юсупов М. М., Вагапова Г. Р. Алгоритм интеллектуального анализа региональных данных об инвестиционном риске // СИИТ. 2024. Т. 6. № 1(16). С. 77–86. EDN [EBASQU](#). [[Makarova E. A., Gabdullina E. R., Yusupov M. M., Vagarova G. R. "Algorithm for intellectual analysis of regional data on investment risk" // SIIT. 2024. Vol. 6, no. 1(16), pp. 77–86. EDN [EBASQU](#). (In Russian).]]
- [11] Елизарова А. В., Сaitова Г. А., Момзи́ков Н. В. Выбор архитектуры нейронной сети для прогнозирования состояния заряда аккумулятора // СИИТ. 2023. Т. 5. № 4(13). С. 123–131. EDN [RXSDIA](#). [[Elizarova A. V., Saitova G. A., Momzikov N. V. "Selection of neural network architecture for predicting the state of charge of a battery" // SIIT. 2023. Vol. 5, no. 4(13), pp. 123–131. EDN [RXSDIA](#). (In Russian).]]
- [12] Закиева Е. Ш. Методология поддержки принятия решений при управлении социетальной системой на основе динамического моделирования и интеллектуальных технологий // СИИТ. 2023. Т. 5. № 3(12). С. 69–92. EDN [UWIPDO](#). [[Zakieva E. Sh. "Methodology of decision support in managing a societal system based on dynamic modeling and intelligent technologies" // SIIT. 2023. Vol. 5, no. 3(12), pp. 69–92. EDN [UWIPDO](#). (In Russian).]]

*Поступила в редакцию 3 февраля 2025 г.*

#### МЕТАДАННЫЕ / METADATA

**Title:** Using R to solve socio-economic data classification problems with artificial intelligence.

**Abstract.** The lack of a structured approach to the application of machine learning in the socio-economic sphere seriously slows down innovative development. The absence of a clear methodology for the implementation of artificial intelligence based on machine learning hinders not only the acceleration of innovation processes, but also the overall increase in the scientific and technical potential of the country. The problem lies not only in the lack of ready-made solutions, but also in the lack of understanding of how to effectively adapt existing machine learning algorithms to the specifics of socio-economic problems. Data classification using neural networks is a modern machine learning method that allows you to solve a variety of problems, including image recognition, text analysis, forecasting financial trends and disease diagnosis. The article systematizes and classifies aspects of machine learning and emphasizes the importance of accelerating the development and implementation of algorithms that are the basis of artificial intelligence. This is necessary to improve the efficiency of managing socio-economic processes.

**Key words:** artificial intelligence, R language, approach, method, socio-economic process.

**Язык статьи / Language:** Русский / Russian.

**Поддержка/Support:** Государственное задание для высших учебных заведений № FEUE-2023-0007.

#### Об авторах / About the authors:

##### **РУДЬ Наталья Юрьевна**

Северо-Кавказский федеральный университет, Россия.  
Зав. каф. мировой экономики и таможенного дела  
(Пятигорск. филиал), профессор. Дипл. экономист по  
мировой экономике (Всеросс. акад. внеш. торговли, 1995).  
E-mail: rud-td@yandex.ru

##### **МОЖЕЛЬСКИЙ Андрей Николаевич**

Пятигорский научно-исследовательский институт  
курортологии (филиал ФФГБУ СКФНКЦ ФМБА), Россия.  
Мл. науч. сотр. Дипл. специалист (Ростовск. высш. воен.  
ком. инж. училище ракетн. войск).  
E-mail: man-cmw@yandex.ru

##### **RUD Natalia Yuryevna**

North Caucasus Federal University, Russia.  
Head of the Department of World Economy and Customs  
Affairs (Pyatigorsk branch), professor. Diploma in Economist in  
World Economy (All-Russian Academy of Foreign Trade, 1995).  
E-mail: rasulovsakit@gmail.com

##### **MOZHELSKY Andrey Nikolaevich**

Pyatigorsk Research Institute of Balneology,  
Russia.  
Junior research fellow. Certified specialist (Rostov Higher  
Military Command Engineering School of Missile Troops).  
E-mail: man-cmw@yandex.ru