

УДК 004.89

Н.М. КОРАБЛЕВ, А.А. ФОМИЧЕВ
Харьковский национальный университет радиоэлектроники**ГИБРИДНЫЕ МОДЕЛИ КЛАССИФИКАЦИИ ДАННЫХ НА
ОСНОВЕ МОДЕЛИ ИСКУССТВЕННОЙ ИММУННОЙ СЕТИ**

Рассмотрено решение задач классификации и кластеризации данных на основе гибридных моделей искусственной иммунной сети, которая позволяет изменять как способ группировки исходных данных, так и структуру сети иммунных объектов. Использование адаптивной иммунной модели позволяет группировать данные или на основе обучающей выборки, или в процессе иммунной обучения. Гибридизация иммунных методов и моделей интеллектуальной обработки информации позволяет повысить их эффективность, что выражается в повышении быстродействия, либо точности их работы. Предложены гибридные модели классификации на основе метода k-ближайших соседей и кластеризации на основе метода k-средних с использованием модели искусственной иммунной сети. Проведены экспериментальные исследования, подтверждающие эффективность предложенных гибридных моделей.

Ключевые слова: классификация, кластеризация, искусственная иммунная сеть, обучающая выборка, иммунное обучения, аффинность, антитело, антиген.

М.М. КОРАБЛЬОВ, О.О. ФОМІЧОВ
Харківський національний університет радіоелектроніки**ГІБРИДНІ МОДЕЛІ КЛАСИФІКАЦІЇ ДАНИХ НА ОСНОВІ
МОДЕЛІ ШТУЧНОЇ ІМУННОЇ МЕРЕЖІ**

Розглянуто розв'язання задач класифікації і кластеризації даних на основі гібридних моделей штучної імунної мережі, яка дозволяє змінювати як спосіб групування вихідних даних, так і структуру мережі імунних об'єктів. Використання адаптивної імунної моделі дозволяє групувати дані або на основі навчальної вибірки, або в процесі імунного навчання. Гібридизація імунних методів і моделей інтелектуальної обробки інформації дозволяє підвищити їх ефективність, що виражається в підвищенні швидкодії, або точності їх роботи. Запропоновано гібридні моделі класифікації на основі методу k-найближчих сусідів і кластеризації на основі методу k-середніх з використанням моделі штучної імунної мережі. Проведено експериментальні дослідження, що підтверджують ефективність запропонованих гібридних моделей.

Ключові слова: класифікація, кластеризація, штучна імунна мережа, навчальна вибірка, імунне навчання, афінність, антитіло, антиген.

M.M. KORABLYOV, O.O. FOMICHOV
Kharkiv National University of Radio Electronics**HYBRID MODELS OF CLASSIFICATION OF DATA ON THE BASIS
OF THE MODEL OF THE ARTIFICIAL IMMUNE NETWORK**

The constant increase in the requirements for classification models that have the ability to adapt to changing conditions and combine the features of different learning methods requires the development of models that allow changing the grouping of input data in the process of work. The use of artificial immune systems to solve the classification problem allows for the grouping of data in various ways, as well as changing not only the structure of the network of immune objects, but also the methods for solving the task. The solution of the problem of classification based on the immune approach is reduced either to the separation of data between previously known groups, or to the formation of new groups from the original data set. Classification of data based on the immune approach can be performed with controlled and uncontrolled learning, as well as by automatic classification. The classification process is divided into three main stages: the preparatory stage, the stage of immune learning, and the stage of refinement of the boundaries of the original classes. Hybridization of immune methods and models of intelligent information processing makes it possible to increase their effectiveness. This helps to increase the speed and accuracy of their work. The possibilities of developing hybrid classification models that function on the basis of various immune models and other approaches have been explored. To this end, hybrid models of data classification based on the model of an artificial immune network are proposed in the work, allowing to change both the method of grouping the initial data and the structure of the network of immune objects. To improve the efficiency of data classification and the formation of immune hybrids in this work, the model of the immune network aiNET. Using a model of an artificial immune network, a hybrid data classification model based on the k-nearest-neighbor method

was proposed, as well as a hybrid data-clustering model based on the *k*-means method. The proposed models were investigated experimentally on different sets of data. Comparative experimental studies of the proposed hybrid models have been made to classify data with existing ones. They showed high efficiency of using for these purpose models of an artificial immune network.

Keywords: classification, clustering, artificial immune network, training sample, immune learning, affinity, antibody, antigen

Постановка проблеми

В настоящее время существует большое количество методов классификации, реализованных на основе различных принципов обучения и разделения группируемых объектов. Основными подходами, используемыми для решения задачи классификации, являются [1, 2]: иерархический, вероятностно-статистический и оптимизационный подходы, а также подходы на основе использования теории графов и методов моделирования систем искусственного интеллекта, функционирующие на основе биологических принципов организации вычислений [3-6]. Такими системами являются искусственные нейронные сети (ИНС), генетические и эволюционные алгоритмы, искусственные иммунные системы (ИИС), системы моделирования коллективного разума и др. Их основным преимуществом является возможность самообучения и гибкость управления процессом группировки объектов.

При решении практических задач возникает необходимость разработки модели классификации, позволяющей изменять способ группировки исходных данных в процессе работы. При этом модель должна быть адаптивной и сочетать особенности различных способов обучения. Аппарат ИИС позволяет проводить группировку данных различными способами, а также изменять не только структуру сети иммунных объектов, но и методы решения поставленной задачи. Использование адаптивной иммунной модели позволяет группировать данные либо на основе обучающей выборки (ОВ), либо в процессе иммунного обучения (ИМО). При классификации данных в процессе ИМО классифицируемые объекты представляются популяцией антител, а объекты ОВ – популяцией антигенов. В процессе ИМО антитела клонируются, подвергаются мутации и отбору, вследствие чего достигают состояния специфичности с антигенами ОВ [7].

Анализ последних исследований и публикаций

Классификация данных на основе иммунного подхода может быть выполнена с контролируемым и неконтролируемым обучением, а также путем автоматической классификации [8-10]. Особенностью классификации с контролируемым обучением является использование ОВ. Процесс классификации условно разбивается на три основных этапа: подготовительный этап PRP, этап иммунного обучения LRN и этап уточнения границ исходных классов CLS [8]. Подготовительный этап необходим для определения данных, которые могут быть классифицированы в процессе ИМО, и определения данных, которые классифицируются на завершающем этапе работы и не используются в процессе ИМО. Особенностью кластеризации на основе иммунного подхода является отсутствие ОВ, состоящей из исходных антигенов, что приводит к необходимости формирования популяции антигенов из набора антител. Процесс кластеризации на основе иммунного подхода также содержит три основных этапа [8]. При этом большое внимание уделяется подготовительному этапу, так как на данном этапе происходит формирование популяции антигенов для поиска центров формируемых кластеров. Задача автоматической классификации подразумевает не только распределение группируемых данных между множеством исходных классов, но и выделение кластеров для данных, которые не могут быть классифицированы. Автоматическая классификация разделяется на два этапа [10]: 1) классификация части исходных данных при использовании ОВ без проведения ИМО; 2) кластеризация оставшихся данных в процессе ИМО. Таким образом, решение задачи классификации на основе иммунного подхода сводится к решению задачи разделения данных между заранее известными группами, либо формированию новых групп из исходного множества данных.

Наиболее распространёнными моделями ИИС, используемыми для классификации данных, являются модели клонального отбора и искусственной иммунной сети [7]. Модель клонального отбора является одной из наиболее простых в реализации моделей, функционирующих на основе теории ИИС. Наиболее распространёнными алгоритмами, реализующими модель клонального отбора, являются алгоритмы CLONALG и VCA [7]. Основное различие между данными алгоритмами заключается в способе организации обработки антител в процессе ИМО, использовании оператора первичного отбора и особенностях клонирования антител. В [11] рассмотрено решение задач классификации и кластеризации объектов с использованием иммунной модели клонального отбора. Модель искусственной иммунной сети является одной из наиболее распространённых моделей, используемых при решении различных практических задач. Это обуславливается большими возможностями, выраженными в возможности взаимодействия классифицируемых антител не только с обучающими антигенами, но и между собой. Наиболее распространёнными алгоритмами, реализующими данную иммунную модель, являются алгоритмы aiNET и opt-aiNET [7]. Алгоритм opt-aiNET является модификацией алгоритма aiNET, предназначенной для решения специфических задач оптимизации. В [12] для решения задач классификации, кластеризации и автоматической классификации на основе модели иммунной сети базовым был выбран

алгоритм aiNET, который был модифицирован для повышения эффективности классификации.

Гибридизация иммунных методов и алгоритмов интеллектуальной обработки информации проводится с целью повышения их эффективности, выражающейся в повышении скорости, либо точности работы. Применительно к задаче классификации на основе иммунного подхода модификация иммунных моделей заключается в использовании различных не иммунных подходов к решению задачи группировки объектов. В [13-16] были предложены гибридные алгоритмы классификации, кластеризации и автоматической классификации данных, функционирующие на основе алгоритма CLONALG в сочетании с методами k-ближайших соседей (kNN), k-средних (k-means) и нечеткой логикой. Однако возможности алгоритма CLONALG ограничены по сравнению с другими иммунными моделями, предполагающими не только межпопуляционные взаимодействия, но и взаимодействия внутри популяции антител или клонов. Поэтому для повышения эффективности классификации с целью формирования иммунных гибридов, в данной работе за основу был выбран модифицированный алгоритм aiNETm [12].

Цель исследования

Целью работы является разработка гибридных моделей классификации данных при использовании метода kNN и кластеризации данных при использовании метода k-means на основе модели иммунной сети.

Изложение основного материала исследования

Классификация данных на основе гибридного иммунного метода kNN

Метод KNN является одним из простейших и наиболее распространенных методов классификации. Поэтому для повышения точности классификации в [16] был предложен модифицированный алгоритм kNN с иммунным обучением на основе алгоритма CLONALG. Поскольку модель искусственной иммунной сети предоставляет больше возможностей для организации взаимодействия между иммунными объектами, принцип работы гибридного алгоритма, приведенного в [16], был использован для алгоритма aiNETm. Таким образом, полученный иммунный гибрид aiNETmkn является модификацией алгоритма aiNETm с иммунным обучением на основе модели искусственной иммунной сети. Работу гибридного алгоритма aiNETmkn можно условно разделить на следующие основные этапы:

- 1) определение стимулирующих антител, формирующих множество «ближайших соседей» и их областей стимуляции;
- 2) определение целевых антигенов для стимулирующих антител;
- 3) классификация стимулирующих антител в процессе ИМО;
- 4) классификация антител из областей стимуляции;
- 5) восстановление популяции исходных объектов и их группировка по результатам классификации последней популяции антител.

Для определения степени взаимодействия между иммунными объектами (антителами, клонами и антигенами) используется специальный критерий аффинности [7]:

$$aff_{ij} = \frac{1}{1 + d_{ij}}, \tag{1}$$

где d_{ij} – расстояние между i -м антителом и j -м антигеном.

На уровне ИО предлагаемый гибридный алгоритм aiNETmkn представляется следующим образом:

$$aiNETmkn(AB, AG, T, k) = \left[\begin{array}{l} NatCalculation(AG, NAT) \rightarrow \\ \rightarrow Presentation(AB, AG) \rightarrow \\ \rightarrow Selection(AB, AG^T) \rightarrow \\ \rightarrow Selection(AB, AB^S, AB^C, AB') \rightarrow \\ \rightarrow SpacesSelection(AB', AB^S, k) \end{array} \right]^{PRP} \rightarrow$$

$$\rightarrow \left[\begin{array}{l} AB^S \\ \left\{ \begin{array}{l} Cloning(ab_i^S, CL_i, k) \rightarrow \\ \rightarrow Mutation(CL_i) \rightarrow \\ \rightarrow Presentation(CL_i, AG_i^T) \rightarrow \\ \rightarrow Suppression(ab_i^S, CL_i) \end{array} \right\} \\ \rightarrow ClassDetection(AB^S, AG^T) \end{array} \right]^{LRN} \rightarrow \tag{2}$$

$$\rightarrow \left[\begin{array}{l} \text{kNNclassDetection}(AB', AB^S, k) \rightarrow \\ \rightarrow \text{Re construction}(AB, AB^C) \end{array} \right]^{CLS}$$

где $\text{NatCalculation}(AB, NAT)$ – ИО определения значений пороговых аффинностей NAT для популяции группируемых объектов; $\text{Presentation}(AG, ab_i)$ – ИО представления антигенов i -му антителу; $\text{Selection}(AB, AG^T)$ – ИО отбора целевых антигенов AG_i^T для каждого классифицируемого антитела; $\text{Selection}(AB, AB^S, AB^C, AB')$ – ИО определения классифицированных антител AB^C , стимулирующих антител AB^S , и антител, которые будут классифицированы после проведения иммунного обучения AB' ; $\text{SpacesSelection}(AB', AB^S, k)$ – ИО определения k ближайших стимулирующих антител для каждого антитела из множества AB' ; $\text{Suppression}(ab_i^S, CL_i)$ – ИО супрессии клонов и стимулирующего антитела; $\text{ClassDetection}(AB^S, AG^T)$ – ИО завершения классификации стимулирующих антител; $\text{kNNclassDetection}(AB', AB^S, k)$ – ИО классификации антител из областей стимуляции на основе метода ближайших соседей; $\text{Re construction}(AB, AB^C)$ – ИО восстановления и классификации исходных объектов по результатам классификации популяции антител.

Таким образом, на первом этапе работы алгоритма aiNETmkn благодаря значениям пороговых аффинностей антигенов происходит определение стимулирующих антител, которые являются ближайшими соседями для объектов, характеризующихся низкими аффинностями к исходным классам. Следует отметить, что большинство стимулирующих антител классифицируются посредством ИМО, однако, для некоторых антител предусматривается возможность классификации без обучения в случае, если они характеризуются максимальными значениями аффинности к обучающим антигенам. В соответствии с этим в процессе ИМО участвуют только стимулирующие антитела, т.е. антитела, входящие в области стимуляции, не подвергаются действию операторов клонирования, мутации и редактирования популяции иммунных объектов. Классификация антител AB' , входящих в области стимуляции, происходит после завершения ИМО и классификации всех стимулирующих антител при использовании метода k ближайших соседей. В соответствии с этим, для каждого антитела, находящегося в области стимуляции на начальном этапе работы алгоритма aiNETmkn определяется k ближайших стимулирующих антител. После завершения ИМО антитела множества AB' классифицируются на основании значений начальных аффинностей со своими ближайшими стимулирующими антителами. При этом значения аффинностей используются в качестве весовых коэффициентов. Таким образом, решение о принадлежности антитела одному из исходных классов определяется путем определения максимальной суммы аффинностей принадлежащих ему стимулирующих антител, входящих во множество ближайших соседей.

Для оценки работы предложенного алгоритма aiNETmkn использовались наборы данных, представленные в табл. 1. Каждый набор характеризуется количеством объектов и размерностью матриц признаков. При этом, если размерность матрицы признаков характеризуется значением 3×3 , то каждый классифицируемый объект и объект ОВ описывается девятью признаками.

Таблица 1 – Характеристики наборов данных

Наборы	Классифицируемые данные	Обучающая выборка	Количество классов	Матрица признаков
Набор 1	500	100	3	3×3
Набор 2	500	100	3	5×5
Набор 3	1000	200	5	5×5
Набор 4	2000	400	5	5×5
Набор 5	5000	400	5	5×5
Набор 6	5000	400	5	10×10
Набор 7	10 000	500	10	5×5
Набор 8	15 000	500	10	5×5
Набор 9	20 000	500	10	5×5
Набор 10	20 000	500	10	10×10

При определении эффективности предложенного алгоритма aiNETmkn было проведено сравнение

результатов его работы с результатами работы алгоритмов kNN, CLONALGm и aiNETm, полученными в [16]. Результаты сравнения перечисленных алгоритмов приведены в табл. 2, в которой используются следующие условные обозначения: «Т» характеризует быстрдействие алгоритма классификации (Time – в % к самому медленному алгоритму); «А» используется для отображения точности группировки (Accuracy – в % к классифицируемым исходным данным); «S» характеризует устойчивость алгоритма (Stability – в % к исходным данным).

Таблица 2 – Результаты классификации данных

Методы		Группируемые наборы данных									
		Н 1	Н 2	Н 3	Н 4	Н 5	Н 6	Н 7	Н 8	Н 9	Н 10
kNN (k = 5)	T	11.7	12.1	12.5	12.6	12.5	14.2	14.8	14.7	14.7	16.0
	A	96.2	96.3	93.8	91.6	90.5	90.5	87.2	86.1	84.3	84.4
	S	99.8	99.7	98.6	98.7	97.5	97.4	95.3	95.4	95.4	95.3
CLONALGm	T	39.5	39.6	39.8	39.9	39.8	41.3	41.5	41.4	41.5	43.4
	A	98.2	98.3	98.2	98.2	98.1	98.1	97.9	97.8	97.8	97.7
	S	99.8	99.7	99.2	99.2	99.1	99.1	98.8	98.8	98.7	98.7
aiNETm	T	38.9	38.9	39.0	39.0	39.1	39.4	39.4	39.4	39.5	39.6
	A	98.7	98.7	98.6	98.5	98.5	98.4	98.3	98.3	98.2	98.0
	S	99.8	99.8	99.7	99.7	99.6	99.6	99.5	99.5	99.4	99.4
aiNETmkn (k = 5)	T	37.9	37.9	38.1	38.1	38.1	38.2	38.2	38.3	38.3	38.4
	A	98.7	98.7	98.6	98.5	98.5	98.4	98.3	98.3	98.2	98.1
	S	99.8	99.8	99.7	99.7	99.6	99.6	99.5	99.5	99.4	99.4

Из результатов экспериментов следует, что алгоритм aiNETmkn превосходит алгоритм aiNETm только по быстрдействию, при этом, не уступая ему по точности и устойчивости классификации. Следует отметить, что при сравнении данного алгоритма с алгоритмом kNN, алгоритм aiNETmkn превосходит kNN по точности и устойчивости классификации, уступая ему по быстрдействию на 20-25%. Таким образом, использование гибридного алгоритма aiNETmkn для решения задачи классификации с контролируемым обучением, более предпочтительно, чем использование других иммунных алгоритмов, таких как CLONALGm, aiNETm или RLAIsm.

Кластеризация данных на основе гибридного иммунного метода k-means

Алгоритм k-means является одним из простейших и наиболее распространённых методов кластеризации. Данный алгоритм часто используется как основа для формирования гибридных методов интеллектуальной обработки информации. Одним из наиболее эффективных методов кластеризации данных является гибридный метод FCM, функционирующий на основе алгоритма k-means, и использующий принципы теории нечетких множеств при выделении кластеров [8]. Поэтому для повышения точности кластеризации в [17] был предложен гибридный алгоритм k-means, функционирующий на основе иммунной модели клонального отбора, и использующий ИМО для выделения кластеров. В связи с тем, что модель искусственной иммунной сети предоставляет больше возможностей для организации обучения, принцип работы гибридного алгоритма, приведенного в [12], был использован для алгоритма aiNETm.

Полученный иммунный гибрид aiNETmkm является модификацией алгоритма k-means, использующего этап ИМО алгоритма aiNETm для повышения скорости формирования кластеров. Работа гибридного алгоритма aiNETmkm начинается с выделения стимулирующих антител как центров формируемых кластеров и определения областей стимуляции при использовании значения предельной аффинности NAT. Кластеризация с помощью данного алгоритма не подразумевает использование популяции антигенов, т.к. модель искусственно иммунной сети позволяет использовать взаимодействие между антителами в процессе ИМО. Таким образом, при использовании стимулирующих антител формируются пересекающиеся области стимуляции, содержащие группируемые объекты.

В процессе ИМО антитела, находящиеся в областях стимуляции, взаимодействуют со стимулирующими объектами как с антигенами. Стимулирующие антитела не подвергаются клонированию и мутации, и изменяются только после переопределения центроида для каждого выделяемого кластера. За счет использования пересекающихся областей стимуляции в процессе супрессии сети клонированным антителам и их клонам представляется несколько стимулирующих антител-центроидов. Обучение модели aiNETmkm завершается при достижении состояния специфичности между всеми антителами, находящимися в областях стимуляции и выделенными центрами кластеров, либо при достижении максимального количества популяций иммунных объектов, формируемых в процессе ИМО.

В соответствии с этим метод aiNETmst на уровне ИО представляется следующим образом:

$$\begin{aligned}
 \text{aiNETmkm}(AB, T, k) = & \left[\begin{array}{l} \text{NatCalculation}(AB, NAT) \rightarrow \\ \rightarrow \text{Selection}(AB, k, AB^S, AB') \rightarrow \\ \rightarrow \text{SpacesSelection}(AB', AB^S) \end{array} \right]^{\text{PRP}} \rightarrow \\
 \rightarrow, & \left[\begin{array}{l} AB' \left\{ \begin{array}{l} \text{Cloning}(ab'_i, CL_i, k) \rightarrow \\ \rightarrow \text{Mutation}(CL_i) \rightarrow \\ \rightarrow \text{Presentation}(CL_i, AB^S_i) \rightarrow \\ \rightarrow \text{Suppression}(ab'_i, CL_i) \end{array} \right\} \\ \rightarrow \text{ClassDetection}(AB', AB^S) \end{array} \right]^T \rightarrow \\
 & \rightarrow \text{Reconstruction}(AB, AB^C),
 \end{aligned} \tag{3}$$

где $\text{NatCalculation}(AB, NAT)$ – ИО определения значений пороговых аффинностей NAT для популяции группируемых объектов; $\text{Selection}(AB, k, AB^S, AB')$ – ИО определения k стимулирующих антител AB^S , являющихся начальными центрами кластеров; $\text{SpacesSelection}(AB', AB^S)$ – ИО определения антител в областях стимуляции; $\text{Cloning}(ab'_i, CL_i, k)$ – ИО клонирования антител в областях стимуляции; $\text{Presentation}(CL_i, AB^S_i)$ – ИО представления клонов стимулирующим антителам; $\text{Suppression}(ab'_i, CL_i)$ – ИО супрессии антител и клонов; $\text{ClassDetection}(AB', AB^S)$ – ИО уточнения центроидов AB^S в областях стимуляции.

Для оценки работы алгоритма aiNETmkm использовались наборы данных, представленные в табл. 1, преобразованные для проведения кластеризации. При определении эффективности алгоритма aiNETmkm было проведено сравнение результатов его работы с результатами работы алгоритмов k -means, CLONALGmc и aiNETmc , полученными в [12]. Результаты сравнения перечисленных алгоритмов приведены в табл. 3.

Таблица 3 – Результаты кластеризации данных

Методы		Группируемые наборы данных									
		Н 1	Н 2	Н 3	Н 4	Н 5	Н 6	Н 7	Н 8	Н 9	Н 10
k-means	T	24.2	24.1	24.2	27.7	27.8	27.8	30.2	30.4	30.5	30.5
	A	99.8	99.7	99.7	99.6	99.5	99.4	99.5	99.4	99.4	99.3
	S	99.9	99.8	99.8	99.7	99.7	99.7	99.7	99.7	99.7	99.8
CLONALGmc	T	25.5	25.5	25.7	25.8	25.9	26.2	26.8	26.9	27.0	27.1
	A	94.8	94.8	94.6	94.7	94.6	94.1	94.0	94.0	94.1	94.0
	S	98.1	98.0	97.9	97.9	97.8	97.6	97.5	97.5	97.5	97.4
aiNETmc	T	25.3	25.3	25.5	25.7	25.7	26.0	26.5	26.5	26.9	26.9
	A	94.8	94.8	94.7	94.7	94.7	94.2	94.0	94.0	94.1	94.0
	S	98.0	98.0	98.0	97.7	97.9	97.7	97.6	97.6	97.6	97.5
aiNETmkm	T	20.0	20.0	20.1	22.4	22.4	22.7	22.9	23.0	23.2	23.2
	A	94.8	94.8	94.7	94.7	94.7	94.2	94.0	94.0	94.1	94.0
	S	98.0	98.0	98.0	97.7	97.9	97.7	97.6	97.6	97.6	97.5

Из результатов кластеризации следует, что модифицированный алгоритм aiNETmkm сопоставим с алгоритмом aiNETmc и превосходит его по быстродействию, совпадая при этом с ним по остальным характеристикам. При этом метод aiNETmkm превосходит алгоритм k -means по быстродействию, но уступает ему в точности и устойчивости кластеризации. Кроме того, при использовании метода aiNETmkm отсутствует необходимость в использовании априорного количества формируемых кластеров, так как их количество определяется иммунными методами. Полученный в результате модификации метод aiNETmkm превосходит другие иммунные методы, использованные в исследованиях.

Выводы

Постоянное повышение требований к моделям классификации, обладающих способностью адаптироваться к изменяющимся условиям и сочетать особенности различных способов обучения, требует разработки моделей, позволяющих изменять группировку исходных данных в процессе работы. Использование ИИС для решения задачи классификации позволяет проводить группировку данных различными способами, а также изменять не только структуру сети иммунных объектов, но и методы решения поставленной задачи. Решение задачи классификации на основе иммунного подхода сводится либо к решению задачи разделения данных между заранее известными группами, либо формированию новых групп из исходного множества данных. При этом процесс классификации условно разбивается на три основных этапа: подготовительный этап PRP, этап иммунного обучения LRN и этап уточнения границ исходных классов CLS

Гибридизация иммунных методов и моделей интеллектуальной обработки информации позволяет повысить их эффективность, что выражается в повышении быстродействия, либо точности их работы. Исследованы возможности разработки гибридных методов классификации, функционирующих на основе различных иммунных моделей и других подходов. С этой целью в работе предложены гибридные модели классификации данных на основе модели искусственной иммунной сети, позволяющей изменять как способ группировки исходных данных, так и структуру сети иммунных объектов. С использованием модели искусственной иммунной сети предложена гибридная модель классификации данных на основе метода k-ближайших соседей, а также гибридная модель кластеризации данных на основе метода k-средних.

Характеристики предложенных моделей были исследованы экспериментальным путем на различных наборах данных. Проведены сравнительные экспериментальные исследования предложенных гибридных моделей классификации данных с существующими, которые показали высокую эффективность использования для этих целей модели искусственной иммунной сети.

Список использованной литературы

1. Duda R.O. Pattern classification / R.O. Duda, P.E. Hart, D.G. Stork. – Wiley & Sons, 2010. – 738 p.
2. Ершов К.С. Анализ и классификация алгоритмов кластеризации / К.С. Ершов, Т.Н. Романова // Новые информационные технологии в информационных системах, 2016. – № 1. – С.274-279.
3. Потапов А.С. Искусственный интеллект и универсальное мышление / А.С. Потапов. – СПб.: Политехника, 2012. – 711 с.
4. Samarasinghe S. Neural networks for applied sciences and engineering: from fundamentals to complex pattern recognition / S. Samarasinghe – Auerbach publications, 2016. – 317 p.
5. Самаркин М.Е. Классификация данных телекоммуникационной компании с помощью технологии data-mining / М.Е. Самаркин, В.Н. Тарасов // Инфокоммуникационные технологии. – ПГУТИ, Самара, 2017. – Том 14. – № 3. – С. 258-263.
6. Жетимекова Г.Ж. Нечеткая классификация с использованием нечеткого анализа кластеризации / Г.Ж. Жетимекова // Вестник КГУСТА. – 2014. – № 4. – С. 117-120.
7. Dasgupta D. Immunological computation, Theory and applications / D. Dasgupta, L.F. Nino – Taylor & Francis Group, 2009. – 278 p.
8. Кораблев Н.М. Классификация объектов на основе искусственных иммунных систем / Н.М. Кораблев, А.А. Фомичёв // Системи обробки інформації. – 2010. – Вип. 6 (87). – С. 13–17.
9. Кораблев Н.М. Кластеризация данных на основе искусственных иммунных систем / Н.М. Кораблев, А.А. Фомичев // Системи обробки інформації. – 2009. – Вип. 4 (78). – С. 77–82.
10. Кораблев Н.М. Автоматическая классификация данных на основе иммунного подхода / Н.М. Кораблев, А.А. Фомичев // Бионика интеллекта: науч.-техн. журнал. – 2014. – № 2 (83). – С. 83-90.
11. Кораблев Н.М. Использование иммунной модели клонального отбора для кластеризации объектов / Н.М. Кораблев, А.А. Фомичев, Д.Н. Соловьев // Інтелектуальні системи прийняття рішень і проблеми обчислювального інтелекту. Матеріали міжнар. наук. конференції. – Херсон: Вид-во ФОП Вишемирський В.С., 2018. – С. 234-236.
12. Кораблев Н.М. Классификация данных с использованием модели искусственной иммунной сети / Н.М. Кораблев, А.А. Фомичев // Інформаційні технології: сучасний стан та перспективи: монографія / Альошин Г.В., Безсонов О.О., Білецький А.Я. [та ін.]; за заг. ред. В.С. Пономаренка. – Х.: ТОВ «ДІСА ПЛЮС», 2018. – С. 86-101.
13. Кораблев Н.М. Нечеткая кластеризация данных на основе искусственных иммунных систем / Н.М. Кораблев, А.А. Фомичёв // Вестник ХНТУ. – 2010. – № 2 (38). – С. 323–328.
14. Кораблев, Н.М. Кластеризация данных методом k-means с использованием иммунных операторов / Н.М. Кораблев, А.А. Фомичев // Бионика интеллекта. – 2011. – № 3 (77). – С. 102–106.

15. Кораблев Н.М. Нечеткая классификация объектов на основе искусственных иммунных систем / Н.М. Кораблев, А.А. Фомичев, М.В. Кушнарев // Комп'ютерні системи та компоненти. – 2010. – Том 1. – Вип.2. – С. 88–94.
16. Кораблев Н.М. Классификация объектов с помощью иммунного метода ближайших соседей / Н.М. Кораблев, А.А. Фомичев // Системи управління, навігації та зв'язку. – 2011. – Вип. 2 (18). – С. 114–118.
17. Korablyov M. The immune method for classifying objects on the basis of the target clonal selection (Immunologiczne metody klasyfikacji obiektów bazujące na selekcji klonalnej) / M. Korablyov, O. Fomichov, M. Kushnaryov, W. Wójcik // Elektronika (LIV). – 2013. – № 8. – P. 36-39.