

**ИССЛЕДОВАНИЕ ЭФФЕКТИВНОСТИ РЕШЕНИЯ
ЗАДАЧИ КЛАССИФИКАЦИИ ГИБРИДНЫМИ СЕТЯМИ КОХОНЕНА**

© 2015 О.П. Солдатова, П.Д. Чайка

Самарский государственный аэрокосмический университет имени академика С.П. Королёва
(национальный исследовательский университет)

Поступила в редакцию 30.07.2015

В статье приводятся результаты решения задачи классификации стандартной гибридной сетью Кохонена и предложенными авторами модификациями данной модели: составной, распределённой и нечёткой гибридными сетями Кохонена. Исследования проводились на модельных задачах репозитория UCI. Приведены структуры предложенных нейросетевых моделей, описаны методики их обучения. Исследована эффективность решения задачи классификации гибридными сетями Кохонена на основе анализа двух критериев: погрешности классификации и суммарного числа нейронов в сети, при котором достигнута минимальная погрешность классификации.

Ключевые слова: гибридная сеть Кохонена, составная гибридная сеть Кохонена, распределённая гибридная сеть Кохонена, нечёткая сеть Кохонена, алгоритм WTA, алгоритм WTM, алгоритм обратного распространения ошибки, алгоритм C-means.

ВВЕДЕНИЕ

В настоящее время для решения задачи классификации широко используются нейронные сети различной архитектуры. Одним из самых известных классов нейронных сетей являются сети с самоорганизацией. Самоорганизация нейронных сетей основана на закономерности, в соответствии с которой глобальное упорядочение сети становится возможным в результате самоорганизующихся операций, независимо друг от друга реализуемых в различных локальных сегментах сети. В соответствии с входными сигналами осуществляется активация нейронов, в результате чего активным оказывается один нейрон в сети (или в группе). Выходной нейрон, который выиграл соревнование, называется нейроном-победителем. Сети с самоорганизацией обучаются без учителя. Типичным представителем сетей с самоорганизацией являются карты Кохонена, которая представляет собой однослойную сеть прямого распространения сигнала [1,2].

Также для решения задачи классификации часто используют многослойный персептрон (MLP). Обычно многослойная сеть состоит из множества входных узлов, которые образуют входной слой; одного или нескольких скрытых слоев вычислительных нейронов и одного выходного слоя. Входной сигнал распространяется по сети в прямом направлении от слоя к слою.

Многослойные персептроны обучаются с учителем, при этом обучение с учителем обычно выполняется с помощью градиентных алгоритмов и метода обратного распространения ошибки [2,3].

В качестве отдельного класса нейронных сетей выделяют сети с нечеткой логикой. Нечеткий многослойный персептрон является представителем данного класса. Он состоит из двух частей: нейронов «нечеткого слоя» и собственно многослойного персептрона. Функции активации нейронов «нечеткого слоя» такой сети являются радиальными базисными функциями (в виде функции Гаусса), моделирующими функции принадлежности. Эти нейроны предназначены для определения степеней принадлежности компонентов входных векторов. Использование нечеткого слоя позволяет уменьшить погрешность классификации пересекающихся классов.

Однако, как показывает практика, лучшие результаты достигаются при использовании гибридных сетей, которые строятся из сетей различных классов. Примером такой сети может служить гибридная сеть Кохонена, которая состоит из слоя Кохонена и многослойного персептрона [2].

Авторами были разработаны три собственные модификации сети на основе гибридной сети Кохонена – составная гибридная сеть Кохонена, распределенная гибридная сеть Кохонена, а также нечёткая гибридная сеть Кохонена [4-6].

1. МОДЕЛЬ ГИБРИДНОЙ СЕТИ КОХОНЕНА

Главная особенность сети с самоорганизацией – это очень высокая скорость обучения по сравнению с сетями, обучаемыми с учителем. Их

*Солдатова Ольга Петровна, кандидат технических наук, доцент кафедры информационных систем и технологий.
E-mail: op-soldatova@yandex.ru
Чайка Павел Дмитриевич, студент первого курса магистратуры факультета информатики.
E-mail: chaikapaoul@gmail.com*

недостатком считается сложность отображения пар обучающих данных (x, d) , поскольку сеть с самоорганизацией, выполняющая обработку только входного вектора x , не обладает свойствами хорошего аппроксиматора, присущими многослойному персептрону.

Очень хорошие результаты удалось получить при объединении самоорганизующегося слоя и персептронной сети, что позволяет совместить способности сети Кохонена к локализации и возможности аппроксимации, свойственные многослойному персептрону [2]. Подобная структура, называемая гибридной сетью Кохонена, изображена на рис. 1. Она представляет собой каскадное подключение слоя Кохонена к персептронной сети. Самоорганизующийся слой улавливает значимые признаки процесса (локализует их на основе входных данных x), после чего им приписывается входной вектор в персептронном слое. Обучение гибридной сети состоит из двух отдельных этапов, следующих друг за другом. Сначала на множестве входных векторов x обучается слой Кохонена. В результате нейроны этого слоя организуются таким образом, что векторы их весов наилучшим образом отображают распределение данных обучающих векторов x . Обучение слоя Кохонена происходит по алгоритмам WTA («победитель получает всё») или WTM («победитель получает больше») [2].

Обучающими сигналами для нее является множество пар (z_i, d_i) , где z_i – это вектор, составленный из выходных сигналов нейронов слоя Кохонена, а d_i – вектор, ожидаемых значений оригинального отображения (x_i, d_i) . Сеть обучается по методу обратного распространения ошибки. Если принять во внимание то, что каждому вектору x соответствует конкретный нейрон-победитель, то можно осуществить предварительную инициализацию весов персептронного слоя таким образом, что значения весов, соединяющих нейрон-победитель с выходными нейронами гибридной сети, будут усредняться по тем векторам d , на которых были одержаны победы.

При такой предварительной инициализации весов, обучение персептронного слоя сводится к незначительной корректировке их значений, отражающей влияние проигравших нейронов слоя Кохонена на окончательный результат. Эта корректировка, как правило, требует небольшого количества итераций и ведет к достижению глобального минимума функции погрешности [3].

Обучение гибридной сети состоит из двух последовательных этапов. Сначала обучается слой Кохонена. В результате обучения векторы весов нейронов данного слоя с минимальной погрешностью отображают распределение данных обучающих векторов x . После завершения обучения, значения весов нейронов слоя Кохонена замораживаются, и производится анализ выходных сигналов. Выход нейрона-победителя переводится в состояние 1, а выходы остальных нейронов переводятся в состояния из интервала (0-1). Переход от фактических выходных сигналов к нормализованным сигналам может производиться по различным формулам. Хорошие результаты можно получить, используя формулу:

$$y_r = \exp\left(-\frac{|u_r - u_{\max}|^2}{\sigma^2}\right),$$

где u_r – реальный выход r -го нейрона слоя Кохонена, y_r – нормализованный выход r -го нейрона слоя Кохонена, u_{\max} – выход нейрона-победителя, σ – параметр, подбираемый для конкретной задачи.

Затем обучается персептронная сеть по одному из алгоритмов обучения с учителем. Обучающими сигналами является множество пар (y_i, d_i) , где y_i – нормализованный выходной вектор слоя Кохонена, а d_i – это вектор эталонных значений выходных нейронов персептрона, соответствующих входному вектору x_i , $t = 1, 2, \dots, p$.

Чаще всего используются градиентные алгоритмы и метод обратного распространения ошибки. Процесс обучения будет протекать гораздо быстрее, чем для обычного многослойного персептрона, благодаря хорошей локализации

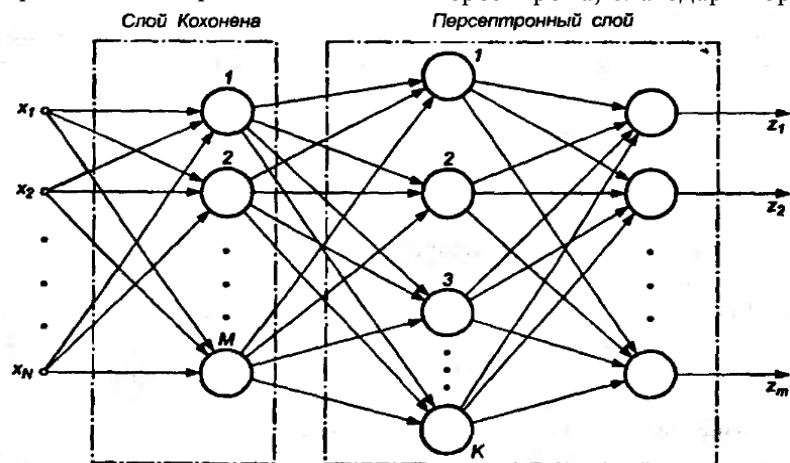


Рис. 1. Структура гибридной сети Кохонена

данных в слое Кохонена. Кроме того, при этом, как правило, достигается глобальный минимум функции погрешности.

Следует отметить, что данная сеть считается обобщением сети встречного распространения. Учёт активности многих нейронов позволяет лучше локализовать входной вектор x в многомерном пространстве и получить лучшее отображение данных нейронной сетью в целом.

2. МОДЕЛЬ СОСТАВНОГО ГИБРИДНОГО ПЕРСЕПТРОНА

Главное отличие составного гибридного перцептрона от гибридной сети Кохонена является измененное начальное условие инициализации весов в слое Кохонена. Также в составном гибридном перцептроне предполагается возможность использования двумерного слоя Кохонена в отличие от стандартной гибридной сети [4]. Структура составного гибридного перцептрона представлена на рис. 2.

В обычной гибридной сети Кохонена веса в слое Кохонена инициализируются случайными значениями от 0 до 1. В составном гибридном перцептроне применяется метод начального предобучения. При этом создается заданное количество сетей Кохонена меньшей размерности, веса которых инициализируются случайным об-

разом. Далее эти сети обучаются на обучающей выборке, но с меньшим количеством итераций, чем задано у составного гибридного перцептрона. Выбираются сети с минимальными погрешностями, а затем они склеиваются в одну сеть, т.е. их матрицы весов объединяются в одну.

После этого составной гибридный перцептрон обучается как обычная гибридная сеть Кохонена. Подобная начальная инициализация позволяет снизить погрешность обучения, а также уменьшить число итераций обучения.

3. МОДЕЛЬ РАСПРЕДЕЛЕННОГО ГИБРИДНОГО ПЕРСЕПТРОНА

Проблема оптимизации архитектуры сети остаётся острой и для составного гибридного перцептрона. Несмотря на хорошие результаты, гибридная сеть содержит большее число нейронов и весов, чем аналогичная сеть Кохонена и многослойный перцептрон, что сказывается на времени обучения и объеме занимаемой памяти.

Поэтому авторами было выдвинуто предложение: уменьшить число связей в гибридной сети Кохонена, путем объединения нейронов в слое Кохонена в группы [4]. Такая сеть получила название распределенная гибридная сеть Кохонена. Ее структура представлена на рис. 3.

В данной сети производится разделение об-

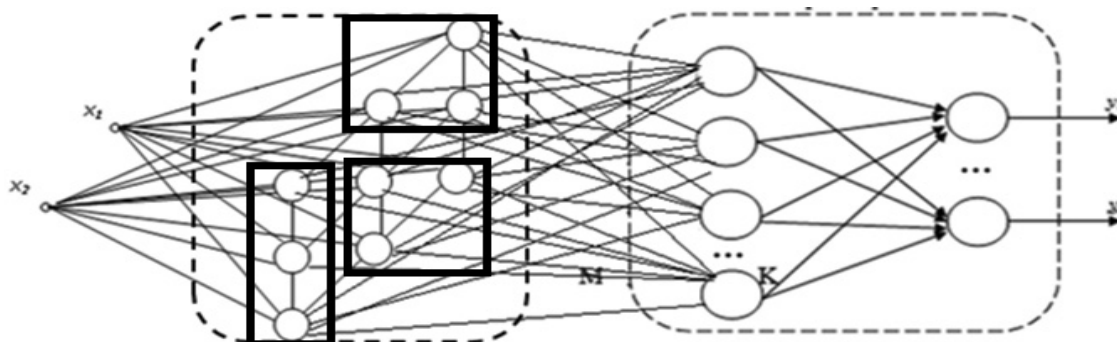


Рис. 2. Структура составного гибридного перцептрона

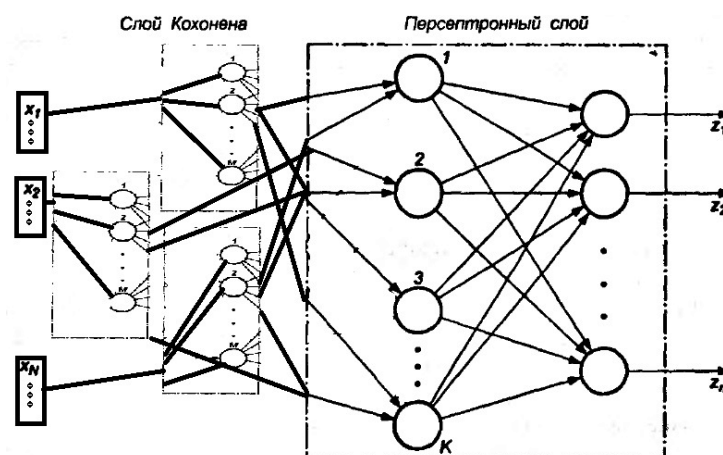


Рис. 3. Структура распределенной гибридной сети Кохонена

учающих данных на классы и в слое Кохонена создаётся своя группа нейронов для каждого отдельного класса. Затем каждая группа обучается только на своём классе при помощи алгоритмов WTA или WTM. После обучения слоя Кохонена обучается персептронная часть сети при помощи алгоритма обратного распространения ошибки. В качестве входного вектора для персептронного слоя используются не значения сигналов от каждого нейрона слоя Кохонена, а усредненные значения выходных сигналов каждой группы нейронов.

Так как каждая группа нейронов настроена только на свой класс, то в случае непересекающихся классов входных данных, в результате самоорганизации определяется не нейрон-победитель, а группа нейронов-победителей. Таким образом, количество групп нейронов равно числу классов, вследствие чего существенно уменьшается количество связей в сети, а значит, уменьшается количество операций на каждом шаге итерации. В данном случае вместо самоорганизации на уровне входных данных реализуется самоорганизация на уровне классов входных данных.

Данное действие не является эффективным в случае обычной карты Кохонена, так как неизвестно предварительное разбиение входных данных на классы. В случае гибридной карты это предварительное разбиение данных существует, поэтому его можно использовать для оптимизации структуры сети.

Однако пользователь может заранее не знать точного количества обучающих выборок для каждого класса. Поэтому для генерации структуры сети авторами вводится понятие так называемого коэффициента уменьшения. При использовании обычной карты Кохонена можно заметить, что минимально число нейронов, необходимое для эффективной работы сети равно количеству классов входных данных (по одному нейрону на класс). Меньшее количество нейронов приведет к отсутствию навыков распознавания некоторых классов входных данных. Максимальное количество нейронов будет равно количеству обучающих примеров (по нейрону на каждый пример). Больше количество нейронов приведет к неизбежному появлению мертвых нейронов. Поэтому оптимальное количество нейронов, необходимое для эффективного решения задачи классификации, лежит внутри этих границ. Для этих целей и вводится коэффициент уменьшения a ($0 \leq a \leq 1$), при котором количество нейронов в каждой группе $N_i = a * D_i$, где D_i количество обучающих примеров принадлежащих данному классу.

Перед началом обучения множество обучающих векторов разделяют на классы. Затем в качестве слоя Кохонена создается n подсетей Кохонена, где n – это количество классов в обучающей выборке.

Далее происходит обучение слоя Кохонена.

Основной особенностью является то, что каждая из составляющих этот слой подсетей обучается на своем классе входных векторов, т.е. каждый конкретный входной вектор подается только на одну конкретную подсеть.

После обучения слоя Кохонена происходит обучение персептронной сети методом обратного распространения ошибки. В качестве входных векторов для персептронной сети будут использоваться усредненные выходы подсетей, составляющих слой Кохонена. Входной вектор подается на каждую подсеть, затем считается выходное значение каждого нейрона подсети, затем суммируется и делится на количество нейронов в подсети.

Благодаря такому распределению входных векторов существенно снижается количество связей между слоями Кохонена и персептронной сетью, что уменьшает количество вычислений, производящихся на каждом шаге итерации. Также снижается погрешность обучения и количество итераций обучения.

4. МОДЕЛЬ НЕЧЕТКОЙ ГИБРИДНОЙ СЕТИ КОХОНЕНА

В качестве отдельного класса нейронных сетей выделяют сети с нечеткой логикой. Нечеткий многослойный персептрон является представителем данного класса. Он состоит из двух частей: нейронов «нечеткого слоя» и собственно многослойного персептрона. Функции активации нейронов «нечеткого слоя» такой сети являются радиальными базисными функциями (в виде функции Гаусса), моделирующими функции принадлежности. Эти нейроны предназначены для определения степеней принадлежности компонентов входных векторов.

Нечеткая гибридная сеть Кохонена состоит из трех связанных друг с другом слоев. Первый слой – это сама сеть Кохонена. С ней соединяется второй слой нечеткой самоорганизации. Третий слой – это многослойный персептрон [5]. Структура сети представлена на рис. 4.

Входной вектор $x = [x_1, x_2, \dots, x_n]^T$ подается на вход слоя Кохонена, в котором формируется вектор расстояний от этого вектора до каждого из нейронов слоя Кохонена $z = [z_1, z_2, \dots, z_p]^T$. Далее этот вектор попадает на вход слоя нечеткой самоорганизации, на выходе которого формируется вектор, состоящий из степеней принадлежности z к конкретным центрам (радиально базисным функциям); $\mu(z) = [\mu_1(z), \mu_2(z), \dots, \mu_k(z)]^T$. Конкретные компоненты $\mu_i(z)$ рассчитываются таким образом, чтобы удовлетворять усло-

вию нормализации $\sum_{i=1}^m \mu_i(z^{(t)}) = 1$ для каждого

вектора $z^t, t = 1, \dots, p$, где p – число векторов в обучающей выборке [2]. Выходы нечеткого

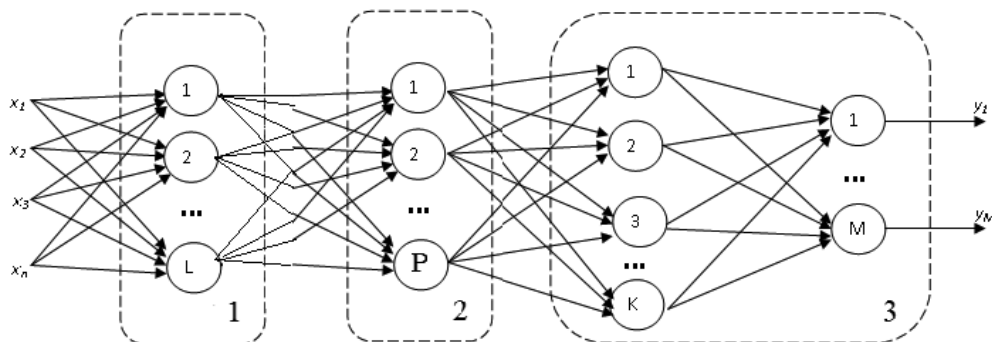


Рис. 4. Структура гибридной сети Кохонена:

1 – слой Кохонена, 2 – слой нечеткой самоорганизации, 3 – многослойный перцептрон

многослойного перцептрона трактуются как степени принадлежности предъявленного объекта соответствующему классу. После этого вектор степеней принадлежности поступает на вход многослойного перцептрона, на выходе которого мы получаем код класса, к которому принадлежит $x = [x_1, x_2, \dots, x_n]^T$.

Обучение сети происходит в три этапа. Сначала обучается слой Кохонена. В результате обучения векторы весов нейронов данного слоя с минимальной погрешностью отображают распределение данных обучающих векторов x [7]. После этого происходит обучение слоя нечеткой самоорганизации с помощью алгоритма C-means [2]. И наконец, происходит обучение многослойного перцептрона алгоритмом наискорейшего спуска и методом обратного распространения ошибки.

5. ЭКСПЕРИМЕНТАЛЬНЫЕ ИССЛЕДОВАНИЯ

Целью проведённых исследований является определение эффективности построенной сети как универсального классификатора и выявление влияния параметров сети, алгоритмов обучения и характеристик входных данных на качество классификации. Исследования проведены при помощи разработанной авторами автоматизированной системы [6].

В качестве тестовых данных были взяты три на-

бора данных из репозитория UCI [8]: ирисы Фишера, сорта вин и заболевания печени. В табл. 1 производится сравнение погрешности классификации многослойного перцептрона, нечеткого многослойного перцептрона, гибридной сети Кохонена, составной гибридной сети Кохонена, распределенной гибридной сети Кохонена и нечеткой гибридной сети Кохонена на заданных наборах данных. В качестве погрешности классификации использовалась доля неверно распознанных векторов в тестовой выборке.

В табл. 2 производится сравнение суммарного количества нейронов в гибридной сети Кохонена, составной гибридной сети Кохонена, распределенной гибридной сети Кохонена, и нечеткой гибридной сети Кохонена при решении задач классификации на заданных наборах данных с результатами, описанными в табл. 1.

ЗАКЛЮЧЕНИЕ

На основании проведённых исследований можно сделать вывод, что распределенная гибридная сеть Кохонена имеет более адекватную структуру для решения данных задач, чем обычная гибридная сеть Кохонена и составная гибридная сеть Кохонена, при аналогичных погрешностях классификации. Это позволяет уменьшить вычислительную сложность используемых алгоритмов, а именно: количество опе-

Таблица 1. Сравнение погрешности классификации на различных наборах данных

Тип сети	Ирисы Фишера	Сорта вин	Заболевания печени
Многослойный перцептрон	0,0370	0,2450	0,4850
Нечеткий многослойный перцептрон	0,0000	0,0723	0,4140
Гибридная сеть Кохонена	0,0000	0,0723	0,3860
Составная гибридная сеть Кохонена	0,0000	0,0552	0,2160
Распределенная гибридная сеть Кохонена	0,0000	0,0693	0,2300
Нечеткая гибридная сеть Кохонена	0,0000	0,0552	0,1530

Таблица 2. Сравнение суммарного числа нейронов на различных наборах данных

Тип сети	Ирисы Фишера	Сорта вин	Заболевания печени
Гибридная сеть Кохонена	100	144	256
Составная гибридная сеть Кохонена	58	63	201
Распределенная гибридная сеть Кохонена	27	41	181
Нечеткая гибридная сеть Кохонена	43	55	185

раций, выполняемых за одну итерацию, время обучения и объем занимаемой памяти.

Однако самой эффективной сетью для решения задачи классификации является нечеткая гибридная сеть Кохонена, которая с минимальной погрешностью решает задачи классификации для классов с нелинейно разделимостью (ирисы Фишера), для классов с вероятностной разделимостью (сорта вин), в том числе и для сильно пересекающихся классов (заболевания печени).

СПИСОК ЛИТЕРАТУРЫ

1. Kohonen T. Self-organizing maps // Springer Science & Business Media, 2001. V. 30.
2. Осовский С. Нейронные сети для обработки информации [пер. с польского И.Д. Рудинского]. М.: Финансы и статистика, 2002. 344с.
3. Haykin S.S. et al. Neural networks and learning machines // Upper Saddle River : Pearson Education, 2009. V. 3.
4. Солдатова О.П., Чайка П.Д. Исследование эффективности решения задачи классификации распределенной

гибридной сетью Кохонена // Труды международной научно технической конференции «Перспективные информационные технологии (ПИТ 2014)» (Самара, СГАУ, 30 июня–2 июля 2014 г.). Самара: Издательство Самарского научного центра РАН, 2014. с.170-173.

5. Солдатова О.П., Чайка П.Д. Использование нечёткой гибридной сети Кохонена для решения задачи классификации // Труды международной научно технической конференции «Перспективные информационные технологии (ПИТ 2015)» (Самара, СГАУ, 28–30 апреля 2015 г.). Самара: Издательство Самарского научного центра РАН, 2015. Т.1. с.218-220.
6. Солдатова О.П., Чайка П.Д. Программный комплекс генерации гибридных нейронных сетей для решения задачи классификации эмпирических данных // Свидетельство о государственной регистрации программы для ЭВМ. Рег.№ 2015611353 от 28.01.2015.
7. Нейронные сети [Электронный ресурс]. URL: <http://www.aiportal.ru/articles/neural-networks/neural-networks.html> (дата обращения 12.05.2015).
8. Bache K., Lichman M. UCI Machine Learning Repository [available at: <http://archive.ics.uci.edu/ml>]. Irvine, CA: University of California, School of Information and Computer Science, 2013.

EFFICIENCY ANALYSIS OF SOLUTION OF CLASSIFICATION USING HYBRID KOHONEN NEURAL NETWORKS

© 2015 O.P. Soldatova, P.D. Chayka

Samara State Aerospace University named after Academician S.P. Korolyov
(National Research University)

The results of solution of classification using standard hybrid kohonen neural network and modifications of this model created by authors - composite, distributed and fuzzy hybrid kohonen neural network were obtained in this paper. Researches were performed on dataset models from UCI repository. Neural networks' structures and teaching algorithms were described. The efficiency of neural networks based on two criteria: classification error and neurons number in network when minimum classification error was archived was researched.

Keywords: hybrid kohonen neural network, composite hybrid kohonen neural network, distributed hybrid kohonen neural network, fuzzy kohonen neural network, WTA algorithm, WTM algorithm, backpropagation, C-means algorithm.

Olga Soldatova, Candidate of Technics, Associate Professor at the Information Systems and Technologies Department. E-mail: op-soldatova@yandex.ru
Pavel Chayka, First-Year Undergraduate Student of Information Systems and Technologies Department. E-mail: chaikapaul@gmail.com