



3. Ольшанский, А.М., Игнатенков, А.В. О построении квазихопфилдовской искусственной нейронной сети для решения задач упорядочения процессов [Текст]./А.М.Ольшанский, А.В. Игнатенков//В сб.: Перспективные информационные технологии (ПИТ 2015): труды Международной научно-технической конференции. – Самара, Самарский государственный аэрокосмический университет имени академика С.П. Королёва. 2015. С. 73-78.

С.А. Онисич, О.П. Солдатова

## РЕШЕНИЕ ЗАДАЧИ КЛАССИФИКАЦИИ С ИСПОЛЬЗОВАНИЕМ НЕЙРОННОЙ НЕЧЕТКОЙ СЕТИ МАМДАНИ-ЗАДЕ

(Самарский университет)

Использование методов машинного обучения для решения задач классификации, в том числе и на основе математической логики, давно стало эффективным методом решения данного класса задач. Однако за прошедшее время стали очевидны недостатки используемой классической логики, и поэтому очевидным шагом стал переход к применению нечеткой логики.

Нечеткой логикой называют логику многозначную, в отличие от логики бинарной, т.е. имеющей только два возможных значения (0, обозначающее ложь, и 1, обозначающее истину). В нечеткой логике значение находится на отрезке от 0 до 1, обозначает степень принадлежности чёткой переменной нечёткому множеству значений и рассчитывается при помощи функций фаззификации [1].

Нечеткая логика была разработана, чтобы позволить вычислительным системам определять различия между данными, которые не являются строго истинными или строго ложными. Преимущества нечеткой логики – в частности, более высокая помехоустойчивость – позволяют более эффективно решать задачу классификации в условиях пересекающихся классов.

Тем не менее, использование нечеткой логики ограничено неоднозначностью реализации ее основных операторов: T- и S-норм и импликации.

T-норма определяется как отображение  $T: [0,1] \times [0,1] \rightarrow [0,1]$ , удовлетворяющее следующим аксиомам:

– граничные условия

$$T(0,0) = T(0,1) = T(1,0) = 0,$$

$$T(x, 1) = x;$$

– коммутативность

$$T(x, y) = T(y, x);$$

– ассоциативность



$$T(T(x, y), z) = T(x, T(y, z));$$

– МОНОТОННОСТЬ

$$T(x, y) \leq T(x, z), \text{ если } y \leq z.$$

S-норма, или T-конорма, определяется как отображение  $S: [0,1] \times [0,1] \rightarrow [0,1]$ , удовлетворяющее следующим аксиомам:

– граничные условия

$$S(1,0) = S(0,1) = T(1,1) = 1,$$

$$S(x, 0) = x;$$

– КОММУТАТИВНОСТЬ

$$S(x, y) = S(y, x);$$

– АССОЦИАТИВНОСТЬ

$$S(S(x, y), z) = S(x, S(y, z));$$

– МОНОТОННОСТЬ

$$S(x, y) \leq S(x, z), \text{ если } y \leq z [2].$$

Исходя из приведенных определений,

$$\begin{aligned} \mu_A(x) \otimes \mu_B(y) = \min\{\mu_A(x), \mu_B(y)\} & \quad \mu_{A \rightarrow B}(x, y) = 1, \text{ если } \mu_A(x) \leq \mu_B(y), & \text{если} \\ \mu_A(x) > \mu_B(y) & \quad \mu_{A \rightarrow B}(x, y) = \mu_B(y), \text{ если } \mu_A(x) > \mu_B(y), & \text{если} \end{aligned} \quad (1)$$

$$\begin{aligned} \mu_A(x) \otimes \mu_B(y) = \mu_A(x) \cdot \mu_B(y) & \quad \mu_{A \rightarrow B}(x, y) = \frac{\mu_B(y)}{\mu_A(x)}, \text{ если } \mu_A(x) > \mu_B(y), & \text{если} \\ \mu_A(x) > \mu_B(y) & \quad \mu_{A \rightarrow B}(x, y) = \max\{0, \mu_A(x) + \mu_B(y) - \mu_A(x) \cdot \mu_B(y)\}, & \end{aligned} \quad (2)$$

$$\begin{aligned} \mu_A(x) \otimes \mu_B(y) = \max\{0, \mu_A(x) + \mu_B(y) - 1\} & \quad \mu_{A \rightarrow B}(x, y) = \min\{1, 1 - \mu_A(x) + \mu_B(y)\}, \\ \mu_A(x) + \mu_B(y) > 1 & \quad \mu_{A \rightarrow B}(x, y) = \min\{1, \mu_A(x) + \mu_B(y)\}, & \end{aligned} \quad (3):$$

$$\left\{ \begin{aligned} \mu_A(x) \otimes \mu_B(y) &= \min\{\mu_A(x), \mu_B(y)\} \\ \mu_{A \rightarrow B}(x, y) &= \begin{cases} 1, & \text{если } \mu_A(x) \leq \mu_B(y) \\ \mu_B(y), & \text{если } \mu_A(x) > \mu_B(y) \end{cases} \\ \mu_A(x) \oplus \mu_B(y) &= \max\{\mu_A(x), \mu_B(y)\} \end{aligned} \right. \quad (1)$$

$$\left\{ \begin{aligned} \mu_A(x) \otimes \mu_B(y) &= \mu_A(x) \cdot \mu_B(y) \\ \mu_{A \rightarrow B}(x, y) &= \begin{cases} 1, & \text{если } \mu_A(x) \leq \mu_B(y) \\ \frac{\mu_B(y)}{\mu_A(x)}, & \text{если } \mu_A(x) > \mu_B(y) \end{cases} \\ \mu_A(x) \oplus \mu_B(y) &= \mu_A(x) + \mu_B(y) - \mu_A(x) \cdot \mu_B(y) \end{aligned} \right. \quad (2)$$

$$\left\{ \begin{aligned} \mu_A(x) \otimes \mu_B(y) &= \max\{0, \mu_A(x) + \mu_B(y) - 1\} \\ \mu_{A \rightarrow B}(x, y) &= \min\{1, 1 - \mu_A(x) + \mu_B(y)\}, \\ \mu_A(x) \oplus \mu_B(y) &= \min\{1, \mu_A(x) + \mu_B(y)\} \end{aligned} \right. \quad (3)$$

где « $\otimes$ » обозначает T-норму. « $\oplus$ » - S-норму, « $\rightarrow$ » - импликацию, а  $\mu_A(x)$  и  $\mu_B(y)$ , степени принадлежности переменных  $x$  и  $y$  нечётким множествам  $A$  и  $B$  соответственно.

Для реализации была выбрана нейронная нечеткая сеть Мамдани-Заде, описанная в работе [3]. Структура используемой сети представлена на Рис. 1.

Для выбранной сети используются следующие ограничения:

- входные переменные являются четкими;
- функции фаззификации – функции Гаусса;



- декартово произведение – в форме минимума;
- дефаззификация не производится.

Используемая модель является четырехслойной. Первый слой реализует фаззификацию по функции Гаусса N-мерного (в модели на рисунке – четырехмерного) входного вектора. Следующий слой агрегирует условия правил вывода в соответствии с операцией T-нормы. Третий слой реализует операцию нечеткой импликации. Этот слой является параметрическим; в процессе обучения подбираются параметры  $v_k$ , соответствующий степени принадлежности заключения  $k$ -го правила вывода  $\mu_{B_k}(y^k)$ . Четвертый слой проводит агрегирование правил вывода для каждого класса выходных векторов в соответствии с используемой интерпретацией S-нормы. В отличие от описанной в [3] сети, используемая в данной работе сеть для вычисления выходного класса использует значения, полученные всеми функциями принадлежности, для чего используются веса связи, обозначенные  $z_{ij}$ . Таким образом, этот слой также является параметрическим.

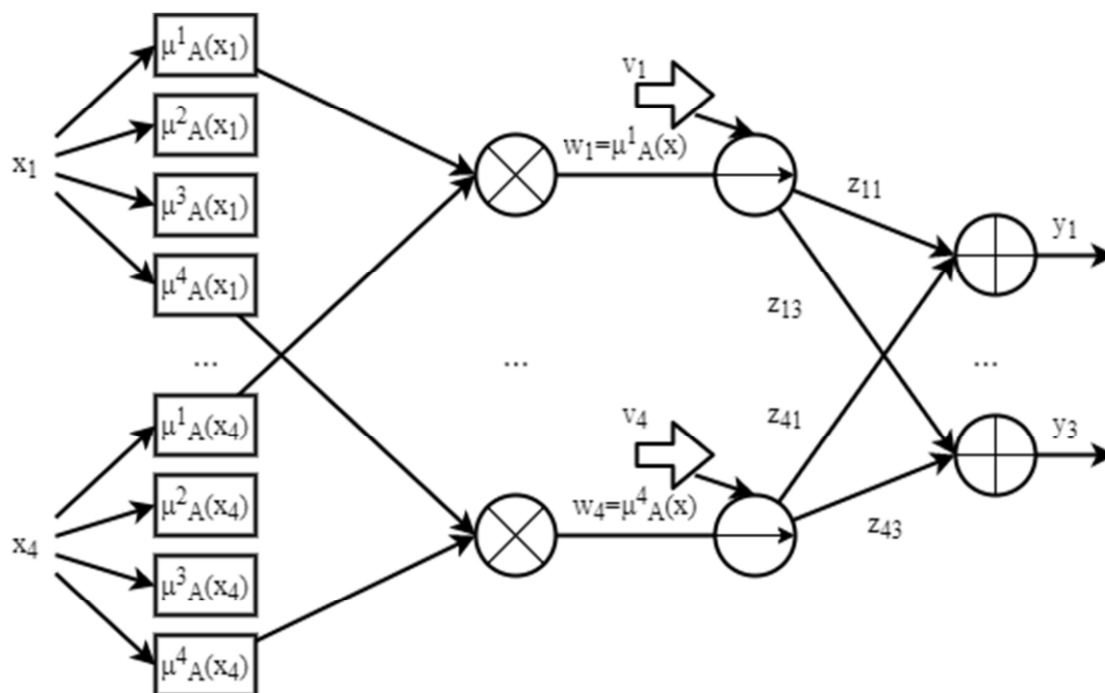


Рис. 1. Структура нейронной нечеткой сети Мамдани-Заде

Для экспериментальных исследований был выбран набор данных «Ирисы Фишера». Ирисы Фишера состоят из данных о 150 экземплярах ирисов, по 50 экземпляров трёх видов – ирис щетинистый (англ. *Iris setosa*), ирис виргинский (англ. *Iris virginica*) и ирис разноцветный (англ. *Iris versicolor*). Для каждого экземпляра измерялись четыре характеристики. Один из классов (*Iris setosa*) линейно отделим от двух остальных. Для обучения использовались 90 образцов ( $\approx 60\%$ ), а оставшиеся 60 образцов – использовались для тестирования качества решения задачи.



В качестве алгоритма обучения используется алгоритм роя частиц. Этот алгоритм является эвристическим и не использует значение производной, что позволяет использовать в нечетких операторах не дифференцируемые функции.

В качестве функции погрешности используется значение бинарной кросс-энтропии. Кросс-энтропия показывает расхождение между двумя вероятностными распределениями. В случае бинарной кросс-энтропии (т.е. если каждое из распределений может иметь лишь одно из двух значений), ее можно найти по формуле:  $H_p(q) = -\frac{1}{N} \sum_{i=1}^N y_i \cdot \ln(p(y_i)) + (1 - y_i) \cdot \ln(1 - p(y_i))$  [4], где  $H_p(q)$  – значение кросс-энтропии между распределениями  $p$  и  $q$  (т.е. между целевым и полученным векторами);  $N$  – размер выходного вектора;  $y_i$  – значение, находящееся в целевом векторе на  $i$ -том месте;  $p(y_i)$  – значение, находящееся в результирующем векторе на  $i$ -том месте.

Полученные результаты представлены на рисунках 2 – 4.

На основании приведенных ниже данных можно прийти к следующим выводам:

- при использовании алгебры Гёделя после 200 эпох обучения значение кросс-энтропии в среднем на 0,1 – 0,15 выше, чем при использовании алгебр Гогена и Лукашевича;
- использование алгебры Гогена позволяет добиться более быстрого уменьшения значения кросс-энтропии;
- использование алгебры Лукашевича позволяет добиться меньшего значения кросс-энтропии при продолжительном обучении.

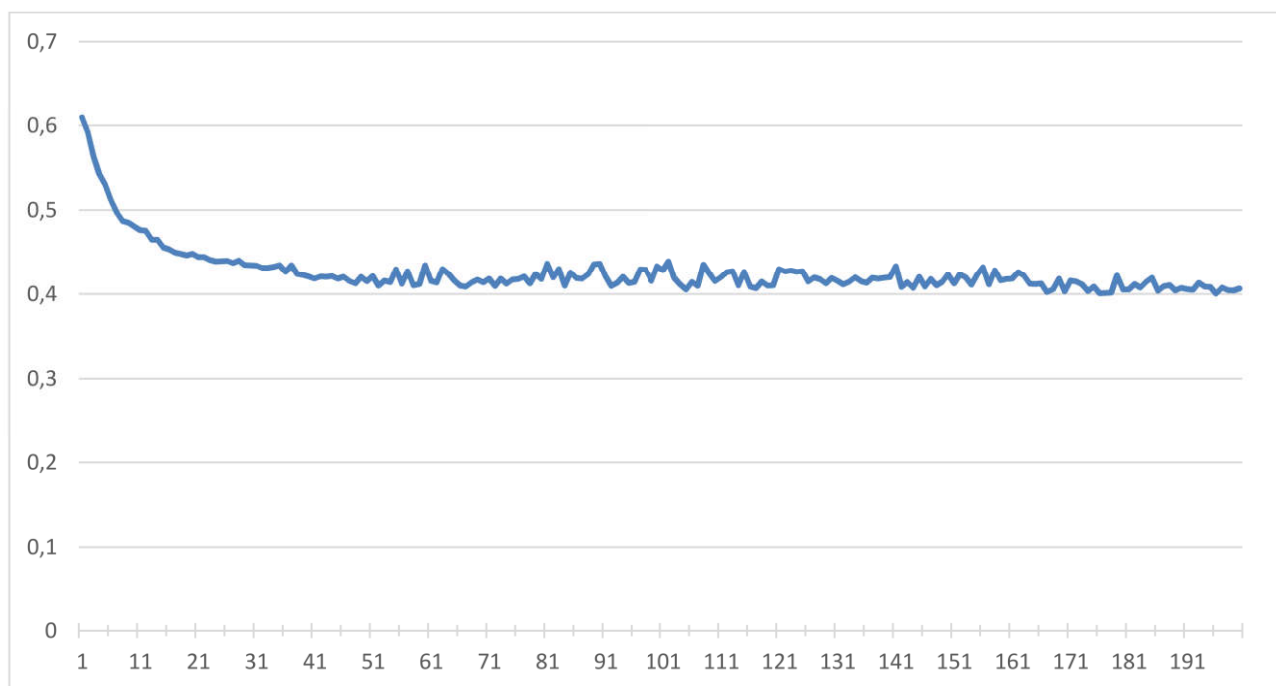


Рис. 2. Уменьшение функции погрешности в зависимости от числа эпох обучения сети при использовании алгебры Гёделя

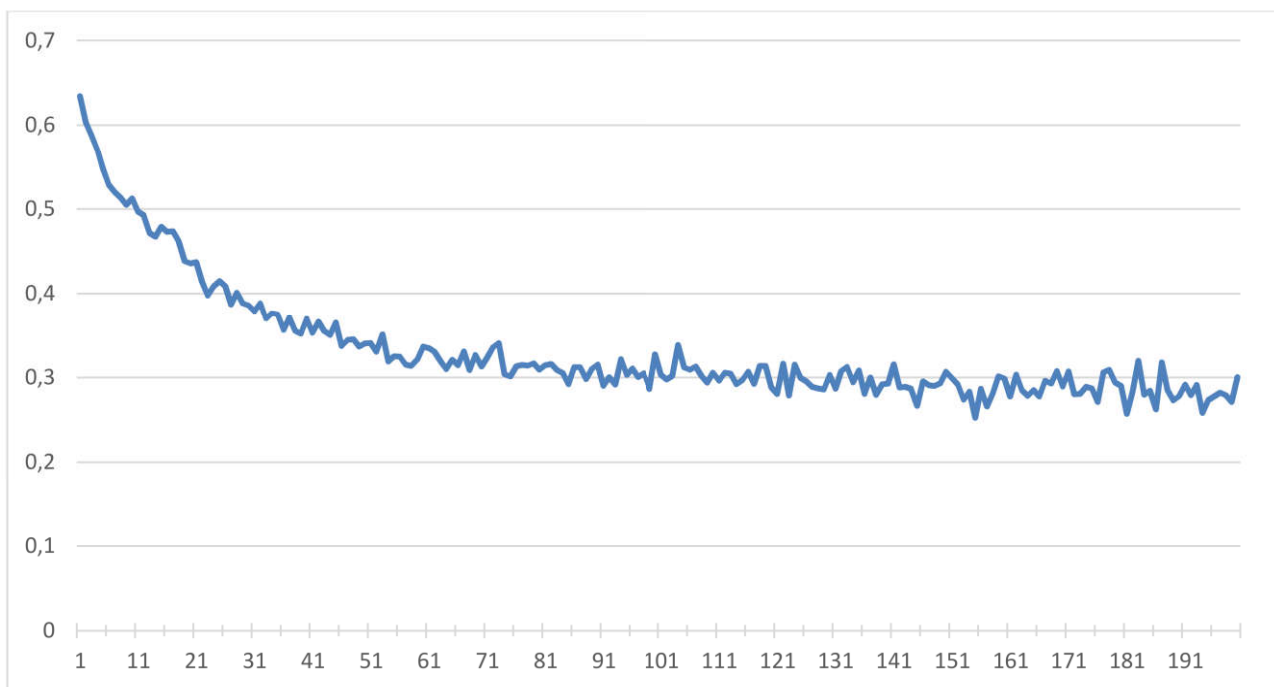


Рис. 3. Уменьшение функции погрешности в зависимости от числа эпох обучения сети при использовании алгебры Гогена

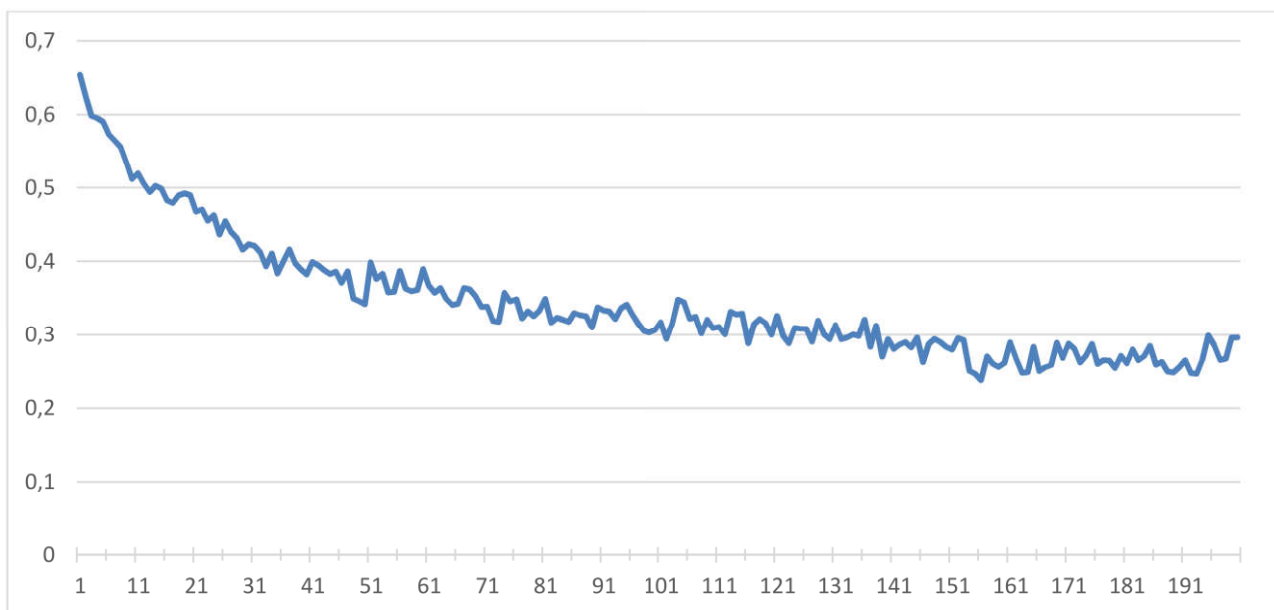


Рис. 4. Уменьшение функции погрешности в зависимости от числа эпох обучения сети при использовании алгебры Лукашевича

### Литература

1. What is Fuzzy Logic? - Definition from Techopedia [Электронный ресурс] // Techopedia: [сайт]. URL: <https://www.techopedia.com/definition/1809/fuzzy-logic> (дата обращения: 02.04.2020).

2. Коновалов А.С., Шумилов П.Е. Параметрические максиминные операторы конъюнкции и дизъюнкции в нечеткой логике // Информационно-управляющие системы, № 1, 2003. С. 4-10.



3. Солдатова О.П., Лёзин И.А. Исследование качества решения задачи классификации нейронными нечёткими продукционными сетями на основе модели вывода Мамдани–Заде // Вестник Самарского государственного технического университета. Серия «Технические науки», № 2(35), 2014. С. 136-148.

4. Understanding binary cross-entropy / log loss: a visual explanation [Электронный ресурс] // Towards Data Science: [сайт]. - URL: <https://towardsdatascience.com/understanding-binary-cross-entropy-log-loss-a-visual-explanation-a3ac6025181a> (дата обращения: 02.04.2020).

Е.В. Орлова

## СРАВНИТЕЛЬНЫЙ АНАЛИЗ МЕТОДОВ МАШИННОГО ОБУЧЕНИЯ ПРИ РЕШЕНИИ ЗАДАЧИ КЛАСТЕРИЗАЦИИ В КРЕДИТНОМ СКОРИНГЕ

(Уфимский государственный авиационный технический университет)

**Аннотация.** Осуществляется сравнительный анализ методов машинного обучения при формировании однородных кластеров кредиторов в задаче кредитного скоринга. Показана большая устойчивость метода  $k$ -средних по сравнению с методов кластеризации, основанным на самоорганизующихся картах Кохонена.

**Ключевые слова:** машинное обучение, метод  $k$ -means, метод ANN.

Для проведения мониторинга платежеспособности клиентов кредитные организации используют скоринговые модели, с помощью которых на основе анализа предыдущих кредитных историй составляется рейтинг заемщиков и определяется вероятность возврата кредита в срок потенциальным заемщиком. Одной из основных проблем при этом и как как составная часть модели кредитного скоринга кредитная организация должна разделить своих клиентов на однородные группы, имеющих одинаковые кредитные портреты, с последующим определением рисковой составляющей кредиторов в каждой однородной группе, а также дальнейшего выявления и моделирования влияния существенный факторов на возможные риски неплатежей у клиентов каждой группы [1-4]. Исследование и сравнительный анализ методов кластеризации, пригодных для решения поставленной задачи показал следующие результаты.

Модельные эксперименты проводятся с использованием статистической информации о кредитных историях клиентов кредитной организации с использование программы статистической обработки данных Statistica 7.0. Особенности изучаемых данных состоят в их разнородности и многомерности. Исследуемая модельная выборка включает данные по 38 клиентам и состоит из следующих показателей, характеризующих заемщиков: срок кредита (*month*), сумма кредита (*value*), пол (0 – женщина, 1 – мужчина), возраст (*age*), количество детей (*children*), средний доход (*income*), а также каждому заемщику присвоим переменную *problems*, характеризующую наличие или