А.А. Тарасов, И.В. Лёзина

ИССЛЕДОВАНИЕ АППРОКСИМАТИВНЫХ ВОЗМОЖНОСТЕЙ НЕЧЁТКОГО ПЕРСЕПТРОНА ПРИ ИСПОЛЬЗОВАНИИ РАЗЛИЧНЫХ ЯДЕРНЫХ ФУНКЦИЙ

(Самарский университет)

Задача аппроксимации является актуальной математической задачей. Для моделирования различных процессов, в частности экономических или физических, необходимо иметь функциональную зависимость, описывающую исследуемый процесс. На практике, для получения такой зависимости необходимо решить задачу аппроксимации, позволяющую на основании имеющихся данных построить функцию, описывающую исследуемый процесс [1].

В данной работе для решения задачи аппроксимации плотности распределения вероятности используется нейронная сеть архитектуры «нечёткий» персептрон. Эта сеть является многослойной и однонаправленной, состоящей из следующих слоёв: слой с нечёткой самоорганизацией (препроцессор) и многослойный персептрон.

Использование препроцессора обосновывается теоремой Ковера, утверждающей, что преобразование задачи в пространство более высокой размерности повышает вероятность линейной разделимости образов. Это позволяет подавать на вход многослойного персептрона уже откластеризованные данные [2].

Структура сети, используемая в данной работе, представлена на рисунке 1.

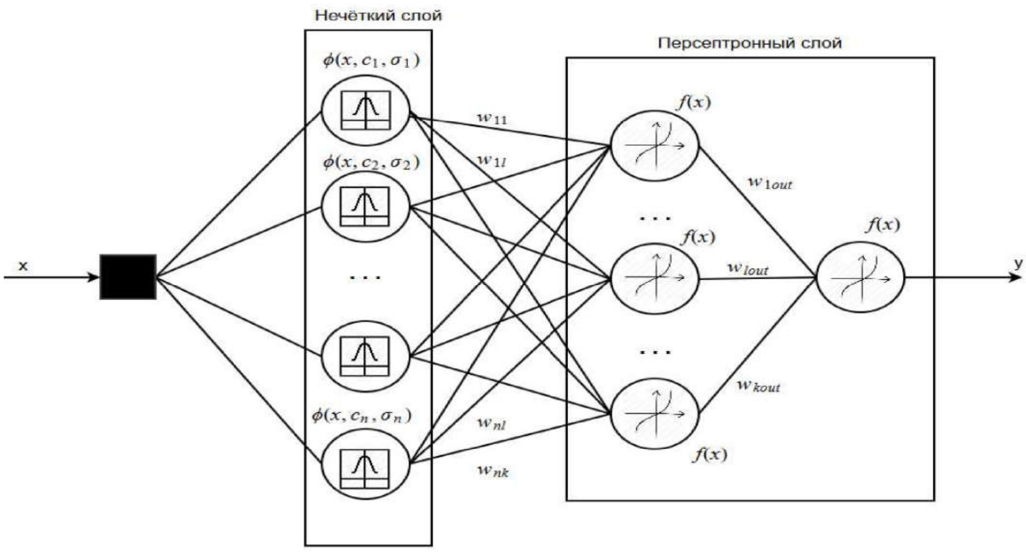

Рис. 1. Структура «нечёткого» персептрона

Аппроксимирующее выражение можно представить в виде:

$$h(x) = f\left(\sum_{i=1}^{k} w_{iout} f\left(\sum_{j=1}^{n} w_{ji} \varphi(x, c_j, \sigma_j)\right)\right) \quad (1)$$

где f – функция активации персептронного слоя, k – число нейронов в персептронном слое, n – число нейронов в нечётком слое, φ – ядерная функция нечёткого слоя, c_j – центр j-го нейрона, σ_j – радиус j-го нейрона.

С учётом ярко выраженной двухкомпонентной структуры для обучения нейронной сети используется алгоритм, состоящий из двух этапов.

На первом этапе проводится обучение, самоорганизующегося слоя, состоящего в подборе параметров центра и радиуса каждого нейрона нечёткого слоя. Как правило, центр и радиус функций активации нейронов нечёткого слоя должны иметь характеристики, подобные обучающим данным. Для определения центров нейронов обычно используются алгоритмы кластеризации. В данной работе используется алгоритм нечёткой кластеризации C-Means, а за ширину ядерной функции будет принято среднее расстояние между данными кластера [3].

Для обучения слоя многослойного персептрона будет использоваться алгоритм обратного распространения ошибки.

Функция фуззификации нечёткого слоя должна удовлетворять следующим требованиям [4]:

$$\varphi(u) = -\varphi(u) \quad (2)$$

$$0 \le \varphi(u) \le \infty \quad (3)$$

$$\int \varphi(u) du = 1 \quad (4)$$

$$\int u^2 \varphi(u) du = 1 \quad (5)$$

В данной работе будут исследованы аппроксимативные возможности «нечёткого» персептрона при использовании следующих функций фуззификации [5]:

– Ядро Епанечникова;
– Гауссово ядро;
– Триквадратное ядро;
– Косинусоидальное ядро;
– Логистическое ядро;
– Треугольное ядро;
– Сигмоидальное ядро;
– Ядро Сильвермана.

Исследования, проводились на случайных последовательностях. Объём выборки составлял 10000 отсчётов, число дифференциальных коридоров – 15. Количество нейронов в нечётком слое 15, в персептронном слое – 20. Функция активации персептронного слоя – сигмоидальная.

В таблице 1 представлены типовые законы распределения плотности вероятности и СКО аппроксимации в зависимости от вида ядерной функции.

Таблица 1 – СКО при тестировании сети для различных законов распределения с помощью различных ядерных функций

Распределение вероятности/ Ядерная функция	Нормальное	Экспоненциальное	Лапласа
Гауссова	0.00159	0.0045	0.0014
Ядро Епанечникова	0.0268	0.0098	0.0163
Биквадратное ядро	0.0235	0.0054	0.0271
Триквадратное ядро	0.01235	0.1108	0.01131
Косинусоидальное ядро	0.0174	0.0307	0.101
Логистическое ядро	0.000582	0.0041	0.006
Треугольное ядро	0.004	0.0179	0.073
Сигмоидальное ядро	0.000563	0.00265	0.003
Ядро Сильвермана	0.000592	0.00297	0.00354

Наилучшие результаты были получены при использовании логистического ядра, формула которого приведена ниже [5].

$$\varphi(x) = \frac{1}{\exp\left(\frac{(x-c)^2}{2\sigma^2}\right)+2+\exp\left(-\frac{(x-c)^2}{2\sigma^2}\right)} \qquad (6)$$

Литература

1. Прохоров, С.А. Аппроксимативный анализ случайных процессов. [Текст]/А.С.Прохоров. –2-е изд., перераб. и доп./СНЦ РАН, 2001. – 125с.
2. Хайкин, С. Нейронные сети: полный курс, 2-е издание [Текст]/ Пер. с англ. – М.: Издательский дом «Вильямс», 2006. – 1104 с.: ил.
3. Осовский, С. Нейронные сети для обработки информации [Текст]/ Пер. с польского И.Д. Рудинского. – М.: Финансы и статистика, 2002. – 344с.: ил.
4. Дьёрфи, Л., Деврой Л. Непараметрическое оценивание плотности. L1-подход [Текст] / Пер. с английского А.Б. Цыбакова. – М.: Наука, 1988. – 407 с.: ил.
5. Kernel Functions for Machine Learning Applications [Электронный ресурс]/ – Электрон. текстовые дан. –,. – Режим доступа: – http://crsouza.com/2010/03/17/kernel-functions-for-machine-learning-applications, свободный.