

ПРИМЕНЕНИЕ НЕЙРОСЕТЕВЫХ МОДЕЛЕЙ В ОЦЕНКЕ ГОРОДСКИХ ЗЕМЕЛЬ

Д. М. Жуков, Ю. Н. Галкина

*Воронежский государственный архитектурно-строительный университет
Управление Федерального агентства кадастра объектов недвижимости по Воронежской области*

Аннотация: адекватной современной рыночной ситуации является методика определения стоимости земельных участков, основанная на применении нейросетевых технологий. Нейросетевая модель современного состояния рынка земли может основываться на многослойных персептронах, представленных логистической функцией.

Нейросетевые методы позволяют проводить как массовую, так и индивидуальную оценку участков земель, основываясь на данных кадастрового учета и информации о ценах продаж.

Abstract: adequate modern market situation the technique of definition of cost of the ground areas, based on application of technologies neural networks is. The model of a neural network of a modern condition of the market of the ground can be based on the multilayered networks submitted by logistical function.

Methods neural networks allow to carry out both mass, and an individual estimation of sites of the grounds being based on the data of the cadastral account and the information on the prices of sales.

Ключевые слова: оценка земель, математические модели, нейронные сети.

Key words: an estimation of the grounds, mathematical models, neural networks.

Создание системы кадастровой оценки городских земель основывается на анализе стоимости земель в отдельных (эталонных) кадастровых кварталах, с последующей аппроксимацией комплекса полученных оценок на остальные кадастровые кварталы. При этом учитывается не только географическая близость кварталов, но и близость земель по физическим характеристикам, способам использования, взаимосвязям характеристик ближнего и дальнего порядка и т. п. Последовательность формирования системы оценки земель можно представить в виде следующих этапов:

- оптимальный выбор «эталонных» кварталов (если не проводится сплошная поквартальная оценка);
- получение наиболее достоверной оценки стоимости земли в эталонных кадастровых кварталах;
- перекрестная верификация исходных, промежуточных и результирующих данных по всем видам стоимости, методикам и слоям оценки;
- аппроксимация (экстраполяция) полученных оценок для эталонных кадастровых кварталов с учетом взаимосвязей их характеристик, на всю городскую территорию;

• территориально-экономическое зонирование на основе последовательной кластеризации с учетом основных элементов городской инфраструктуры, экономики и т. п. и связей различных фрагментов территорий ближнего и дальнего порядка, а также информации по реально сложившемуся земельному рынку.

Опишем необходимые переменные. Каждый земельный участок (кадастровый квартал) характеризуется определенным набором параметров, включающих физические свойства земельного участка, свойства расположенных на нем объектов недвижимости, свойства инженерных коммуникаций, характеристики транспортной и других видов городской инфраструктуры. Обозначим общее количество параметров для любого участка, которые мы в состоянии учесть, буквой n , а весь набор параметров для произвольного i -го участка вектором $X_i = (x_1^2, \dots, x_n^2)$.

Число кварталов обозначим через K , тогда земельные участки всех кварталов будут характеризоваться набором K векторов: $\{X_1, \dots, X_k\}$.

Любой i -й квартал можно оценить с помощью вектора оценок $Y_i = (y_1^2, \dots, y_m^2)$, где m — число методик. Наиболее достоверную оценку стоимости земли в этом квартале обозначим s .

С использованием этих обозначений можно сформулировать следующие задачи:

1. Для некоторого $k = K$ выбрать k векторов параметров $\{X_1, \dots, X_k\}$ — таких, чтобы эта выборка

была наиболее репрезентативной. Выбранные наборы будут соответствовать эталонным кварталам (при проведении сплошной оценки $k = K$).

2. По значениям m оценок стоимости земли кадастрового квартала (y_1, \dots, y_m) найти наиболее достоверную оценку c стоимости земель этого квартала.

3. Построить функцию $F(\cdot) \rightarrow C$ оценки стоимости земель произвольного кадастрового квартала (земельного участка) по его входным параметрам (используя информацию, полученную при решении предшествующих задач).

4. Объединить кварталы в кластеры по близости их параметров и найти территориальные границы кластеров.

Первую задачу будем называть задачей фрагментации, вторую — задачей получения наилучшей оценки, третью — задачей аппроксимации, а четвертую — задачей территориально-экономического зонирования.

Сложность поставленных задач определяется большой неточностью в задании исходных данных, которые к тому же часто неполны либо противоречивы. А поскольку общее число данных невелико, применение аппарата математической статистики не позволит получить приемлемые решения. В настоящее время для решения подобных задач разработаны ряд методов, — искусственные нейросети и нечеткая логика, которые относят к «мягким вычислениям». Ниже приведены разработанные нами алгоритмы, основанные на идеях нейроинформатики.

При проведении оценок в странах с устоявшейся экономикой, как правило, используется нормативная база, позволяющая в каждом конкретном случае выбрать набор факторов, определяющих стоимость данного объекта. Например, Д. Эккерт [3] при проведении оценок, как правило, пользуется формулами вида:

$$C = b_0 + b_1 x_1 + \dots + b_n x_n, \quad (1)$$

где b_0, \dots, b_n — нормативные коэффициенты (цены), x_1, \dots, x_n — характеристики объектов. При оценке допускается калибровка модели, то есть небольшие поправки к коэффициентам, учитывающие индивидуальные особенности объекта, которые в результате могут дать не более 15 % изменения оценки.

В обозначениях, принятых в нейроинформатике, модели такого типа являются однослойной нейросетью с тождественными передаточными функциями:

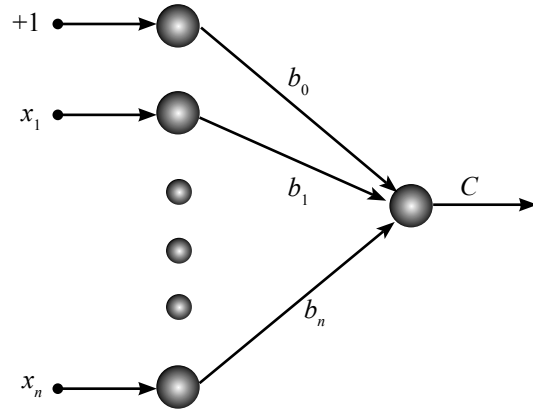


Рис. 1. Структура нейросети, осуществляющая оценку в рамках подходов Д. Эккерта

Входной сигнал (параметры объекта x_1, \dots, x_n) поступает на первый слой нейросети, который передает его к выходному слою. Выходной слой содержит лишь один нейрон, которые суммирует сигналы с весовыми множителями и b_0, \dots, b_n , формируя при этом величину оценки. Калибровка такой модели заключается в небольшой подстройке весовых коэффициентов.

Применение подобных моделей в переходной экономике не представляется возможным, поскольку, во-первых, не понятно, какие факторы являются ключевыми при формировании стоимости. Во-вторых, большинство доступных параметров не являются независимыми и степень их взаимного влияния неизвестна. Поэтому предлагается методика поэтапной оценки с последовательной верификацией параметров и усложнением моделей.

Для выделения наиболее важных параметров применяется метод выделения главных компонент

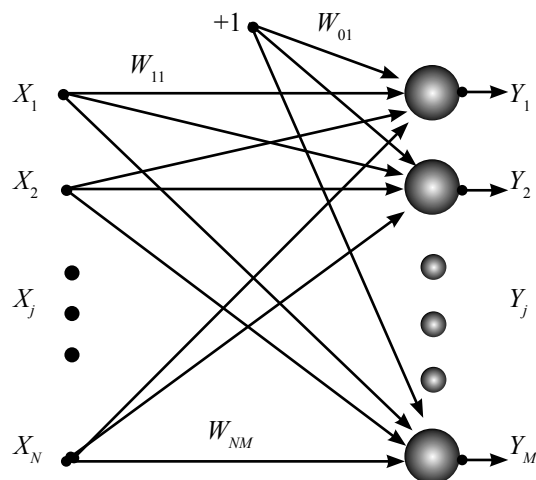


Рис. 2. Архитектура нейросети, осуществляющей выделение главных компонент

в модификации Зангера (Sanger)[2], который упорядочивает найденные главные компоненты по убыванию их значимости. Архитектура нейросети, позволяющей выделить главные компоненты, изображена на рисунке 2.

Она содержит один слой, состоящий из M нейронов; имеет N входов (плюс еще один вход, называемый смещением, на который всегда подается сигнал +1) и M выходов — число главных компоненты, которые следует выделить. При этом $M \leq N$. На вход нейросети подаются обрабатываемые векторы, а выход определяется следующим соотношением:

$$y_i = W_{0i} + \sum_{j=1}^N W_{ji} X_j, \quad i = 1, \dots, M \quad (2)$$

Метод настройки коэффициентов нейросети W_{ij} описывается рекуррентным соотношением:

$$W_{ji}(new) = W_{ji} + ay_i (x_j - \sum_{i=1}^j y_i W_{ji}),$$

$$j = 0, \dots, N, \quad i = 1, \dots, M. \quad (3)$$

После достижения заданной точности итерации прекращаются и с выхода нейросети можно получать главные компоненты обрабатываемых данных в порядке убывания их значимости. Для учета взаимных зависимостей в исходных параметрах следует применять многослойные искусственные нейросети, которые, в отличие от нейросети с постоянными весами, используемой Д. Эккертом, имеют дополнительные слои нейронов, называемые скрытыми слоями, а также допускают возможность автоматической настройки весов на основе обработки данных по уже оцененным участкам.

На рисунке 3 изображена нейросеть прямого распространения с одним скрытым слоем нейронов Rosenblatt [1]. На вход сети подается обрабатываемый

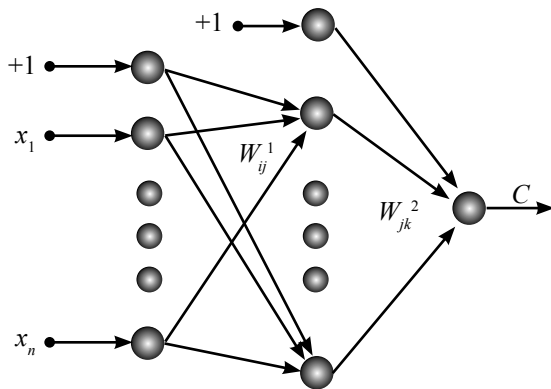


Рис. 3. Архитектура нейросети прямого распространения с одним скрытым слоем, n входами и одним выходом

вектор параметров, с выхода снимается оценка стоимости. Нейроном является простейший «решающий» элемент, устройство которого изображено на рисунке 4:

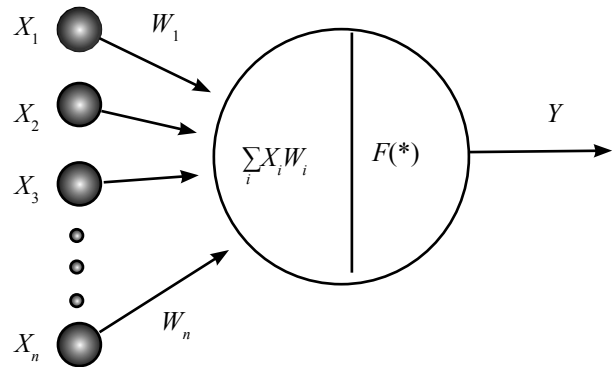


Рис. 4. Схематическое изображение «искусственного нейрона», содержащего n входов, на которые подается возбуждающий сигнал $X_1, \dots, X_n, W_1, \dots, W_n$ — веса входов — настраиваемые параметры. $F(*)$ — передаточная функция нейрона

Сигналы, поступающие на входы X_1, \dots, X_n (входной вектор), умножаются на соответствующие им коэффициенты (веса) связей W_1, \dots, W_n и поступают на суммирующий блок, который определяет уровень возбуждения нейрона:

$$S = \sum_{i=1}^n X_i W_i \quad (4)$$

Выходной сигнал Y определяется путем пропускания возбуждения нейрона S через нелинейную функцию.

$$Y = F(s - d), \quad (5)$$

где d — постоянное смещение. В качестве нелинейной функции чаще всего используют сигмоидную функцию:

$$Y = \frac{1}{1 + \exp(d - S)} \quad (6)$$

которая является непрерывной и сколько угодно раз дифференцируемой. Область ее значений лежит в интервале $(0,1)$.

Настройка весов нейронной сети прямого распространения происходит в результате так называемого обучения с учителем. Это означает, что сети предъявляется как входной образец, так и выходные данные, которые нейросеть должна выдавать на этом образце.

Процесс настройки нейросети, как правило, состоит из следующих этапов:

1. Разделение данных на два множества: обучающее и тестирующее (которые будут далее называться, соответственно, обучающей и тестирующей выборками).

2. Обучение нейросети на обучающей выборке.
3. Тестирование нейросети с помощью тестирующей выборки.
4. Если результаты тестирования неудовлетворительны, вернуться к шагу 4, а возможно, и к шагу 1.

По сути, ошибка тестирования является главной характеристикой обученной сети. Ошибки такого порядка надо ожидать в ответах нейросети при обработке новых данных — вопросов, ответы на которые нам неизвестны. Естественно, если результаты тестирования не устраивают (например, по смыслу решаемой задачи точность прогнозов не должна быть менее 1 % , а ошибка тестирования — порядка 5 %), следует вернуться к предшествующим шагам с целью добиться нужного поведения сети. Причин плохой обученности существует много, наиболее часто это происходит из-за малого числа итераций (недообученность), из-за попадания в локальный минимум, из-за недостаточного количества нейронов в скрытых слоях либо даже из-за использования для настройки данных, которым не свойственна зависимость.

СПИСОК ЛИТЕРАТУРЫ

1. *Rosenblatt F. Principles of Neurodynamics : Perceptrons and Theory of Brain Mechanisms / F. Rosenblatt. — Spartan Press, Washington DC, 1961.*
2. *Sanger T. D. Optimal Unsupervised Learning in a Single-Layer Linear Feedforward Neural Network / T. D. Sanger. — Neural Networks 2. — 1989. — P. 459—473.*
3. Организация оценки и налогообложения недвижимости / под ред. Д. Эккерта. — М., 1997.

THE LITERATURE LIST

1. *Rosenblatt F. Principles of Neurodynamics : Perceptrons and Theory of Brain Mechanisms / F. Rosenblatt. — Spartan Press, Washington DC, 1961.*
2. *Sanger T. D. Optimal Unsupervised Learning in a Single-Layer Linear Feedforward Neural Network / T. D. Sander. — Neural Networks 2. — 1989. — P. 459—473.*
3. The organization of an estimation and the taxation of the real estate / edited by D. Ekkert. — M., 1997.