

УДК 519.862.6

МОДЕЛИРОВАНИЕ ДИНАМИКИ СТОИМОСТИ НЕДВИЖИМОСТИ НА ВТОРИЧНОМ РЫНКЕ В ГОРОДЕ ЕЛАБУГА РЕСПУБЛИКИ ТАТАРСТАН НА ОСНОВЕ НЕЙРОННЫХ СЕТЕЙ

Хлюпина М.А., Исавнин А.Г., Валеева З.Ф., Пивоваров В.А.

Набережночелнинский институт (филиал) ФГАОУ ВО «Казанский (Приволжский) федеральный университет», Набережные Челны, e-mail: marina150294@bk.ru

В данной работе применяется аппарат нейронных сетей с целью прогнозирования состояния рынка недвижимости в городе Елабуга. Объектом исследования является недвижимость в районном городе Елабуга Республики Татарстан, предметом исследования является метод нейронных сетей. В данной работе исследования проводились на основе модели сети с разными архитектурами (радиально-базисные функции и многослойные перцептроны) и были выбраны наилучшие сети по ряду характеристик. Целью проводимых экспериментов было построение нейросетевой прогностической системы с наименьшей ошибкой тестирования. Для достижения данной цели было проведено исследование влияния представления прогнозируемых данных на ошибку прогнозирования. Также были рассмотрены вопросы влияния структуры нейронной сети на скорость обучения и ошибку прогнозирования.

Ключевые слова: нейронные сети, факторы, прогнозирование, модель сети, радиально-базисные функции, многослойные перцептроны

MODELLING OF THE DYNAMICS OF THE COST OF REAL ESTATE IN THE SECONDARY MARKET IN THE CITY OF ELABUGA TATARSTAN BASED ON NEURAL NETWORKS

Khlyupina M.A., Isavnin A.G., Valeeva Z.F., Pivovarov V.A.

Institute (branch) of federal public autonomous educational institution of higher education the «Kazan (Volga region) federal university, Naberezhnye Chelny, e-mail: marina150294@bk.ru

In this paper we apply the neural apparatus to predict the real estate market in the city of Elabuga. The object of research is a property in the district of the city Elabuga Tatarstan, the subject of this study is the method of neural networks. In this paper, the research carried out on the basis of models with different network architectures (radial-basis functions and multilayer perceptrons) and best network for a number of characteristics were selected. The purpose of the experiments was to construct a neural network forecasting system with minimum-error testing. To achieve this goal, it was investigated the effect of the projected presentation of data on the prediction error. Also questions of the neural network structure influences were considered at the speed of learning and prediction error.

Keywords: neural networks, factors, forecasting, network model, radial basis functions, multi-layer perceptrons

Необходимым элементом рыночной экономики выступает институт независимой оценки собственности, без которого невозможно становление права собственности и демократизация экономической жизни.

В первую очередь это касается рынка недвижимости, развитие которого может определить в перспективе характер изменений всей экономики.

Оценка стоимости любого объекта собственности – упорядоченный, целенаправленный процесс определения в денежном выражении стоимости соответствующего вида с учетом потенциального и реального дохода, приносимого им в определенный момент времени в условиях конкретного сегмента рынка.

Особенностью процесса оценки стоимости объекта имущества является его рыночный характер. Это означает, что процесс оценки объекта не ограничивается учетом одних только затрат на создание или приобретение оцениваемого объекта собственности – необходим учет совокуп-

ности рыночных факторов, экономических особенностей, а также макроэкономического и микроэкономического окружения. Рыночная стоимость оцениваемого объекта непостоянна – изменяется во времени под воздействием многочисленных факторов. По этой причине она может быть определена только на данный конкретный момент времени. Это означает, что периодическая оценка объектов собственности является необходимым условием функционирования рыночной экономики.

Использование нейронных сетей может быть эффективным инструментом решения задачи оценки рыночной стоимости жилой недвижимости. Очевидно, что цена квартиры зависит от многих факторов, например общей и жилой площади, количества комнат, этажа, территориального расположения дома, его этажности, состояния, наличия коммуникаций и др. Опытные риэлторы справляются с задачей оценки без труда, применяя свои знания и интуицию, опираясь

на известные им аналоги и используя ассоциативное мышление. Все эти знания и умения относятся к числу плохо формализуемых, отчасти неосознаваемых, поэтому разработка однозначного алгоритма определения цены на основе значений влияющих факторов – крайне сложная и почти невыполнимая задача.

Вместе с тем существует значительное число примеров уже оцененных квартир. Используя массив сведений о них, можно попытаться извлечь интересующую зависимость.

Для этого создается нейронная сеть, в которой количество входных нейронов соответствует количеству входных факторов, которые влияют на цену. В выходном слое будет всего один нейрон, соответствующий выходному фактору – цене.

Для обучения необходим массив обучающих примеров. Количество примеров должно быть достаточно большим – по некоторым расчетам, в 10–15 раз больше числа нейронов в сети. Примеры предъявляются ИНС (искусственных нейронных сетей), при этом веса связей внутри нее постепенно изменяются, с тем, чтобы реальный выходной сигнал был как можно ближе к ожидаемому значению выходного фактора. Один цикл предъявления всех учебных образцов называется эпохой. Обычно требуется несколько тысяч эпох, чтобы обучить нейронную сеть, но на современных компьютерах такое обучение занимает несколько минут.

Часть примеров не участвует в обучении, а выделяется в так называемое тестовое множество. На каждой эпохе работа сети проверяется на тестовом множестве. Таким образом тестируется способность ИНС к обобщению: возможности распространить выявленную закономерность к данным, не участвующим в обучении.

Обучение ИНС заканчивается, когда достигнуто заданное значение средней (или минимальной) ошибки, когда сеть исчерпала

возможности обучения или же когда пройдено определенное число эпох. После этого веса связей фиксируются, и сеть может использоваться в рабочем режиме. Теперь, если в качестве входных сигналов сети указать параметры оцениваемой квартиры, значение на выходе будет представлять ее цену, рассчитанную на основе выявленной закономерности.

Первичный набор факторов, определявшийся экспертным путём с учётом наличия достаточного количества информации в основных риэлтерских базах, составил:

- выходная переменная: цена продажи объекта недвижимости;
- количественные факторы: общая площадь помещения (кв. м);
- географические факторы: расположение объекта.

Количественные факторы (с учётом преобразований) используются в модели в неизменном виде.

Преимущество нейронных сетей перед моделями множественной регрессии состоит в том, что нет необходимости преобразовывать упорядоченные категории в набор бинарных переменных, теряя порядок значений, обусловленный экономическими причинами. Так как зависимости в нейронных сетях нелинейны, достаточно указать произвольные числовые значения, монотонно связанные с уровнями фактора, например, последовательные целочисленные значения или усреднённые значения цены в разрезе соответствующих категорий.

Статистические данные цен продаж, индексы стоимости жилья города Елабуга Республики Татарстан, а также основная первичная информация была предоставлена агентством недвижимости «Престиж».

Данные о ценах продаж квартир на вторичном рынке города Елабуга приведены в табл. 1. Анализируя их, необходимо учитывать, что статистика цен продаж построена на основе ограниченного количества сделок.

Таблица 1

Статистические данные среднеарифметических цен продаж однокомнатных квартир в четвертом квартале 2015 года

Тип жилья	Дата	Однокомнатные	
		Цена в руб. за 1 кв. м	Изменение в % за рассматриваемый период
Ленинградка	01.10.2015	31 600	19,01
	01.12.2015	37 750	
Сталинка	01.10.2015	37 150	7,26
	01.12.2015	39 850	
Хрущевка	01.10.2015	27 100	12,2
	01.12.2015	30 400	
Брежневка	01.10.2015	29 050	13,4
	01.12.2015	32 950	

Реальная выборка, содержащая 13 входных параметров, которые необходимо использовать для построения прогностической системы оценки стоимости жилья в городе Елабуга, составила суммарно 100 наблюдений. Она была случайным образом разделена на обучающую (80%), валидационную (10%) и тестовую (10%).

Формируя выборку определенного размера, можно всегда скорректировать количество входных и выходных данных (взять меньше, чем присутствует в таблице, таким образом, оставшиеся наборы просто не будут участвовать в обучении). То есть выборка не будет терпеть каких-либо изменений, что упростит работу при моделировании.

Проведем прогнозирование оценки рыночной стоимости недвижимости. Исследования проводились на основе модели сети с разными архитектурами (радиально-базисные функции и многослойные перцептроны), и были выбраны наилучшие сети по ряду характеристик. Целью проводимых экспериментов было построение нейросетевой прогностической системы с наименьшей ошибкой тестирования. Для достижения данной цели было проведено исследование влияния представления прогнозируемых данных на ошибку прогнозирования. Также были рассмотрены вопросы влияния структуры нейронной сети на скорость обучения и ошибку прогнозирования.

Каждый из экспериментов состоял из нескольких этапов:

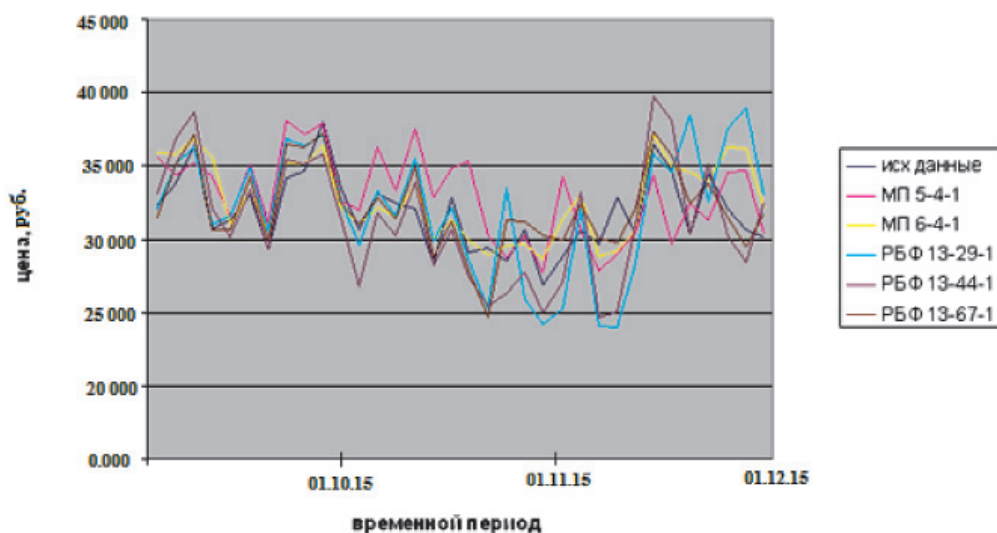
1. Формирование обучающей выборки. На этом этапе определялся вид представления прогнозируемых данных, осуществлялось формирование блока представительских (обучающих) выборок.

2. Обучение нейронной сети с использованием сформированного на первом этапе блока обучающих выборок. Качество обучения характеризовалось ошибкой обучения, определяемой как суммарное квадратичное отклонение значений на выходах нейронной сети в обучающей выборке от реальных значений, полученных на выходах нейронной сети. Критерий прекращения обучения – 200 итераций или уменьшение ошибки на выходах сети на два порядка, по сравнению с первичной ошибкой. В том случае, если при описании опыта не указано, что произошло снижение ошибки на два порядка, обучение останавливается по первому критерию.

3. Третий этап – тестирование нейронной сети. Определяется качество прогнозирования при подаче на вход 4,0–5,0% наборов из обучающей выборки. Эксперимент является успешным, если относительная достоверность не менее 80,0%.

4. На четвертом этапе осуществляется пробное прогнозирование. На входе нейронной сети – наборы, которые не были внесены в обучающую выборку, но результат по которым (прогноз) известен.

Полученные результаты приведены ниже (рисунок, табл. 2).



Результат прогнозирования:

МП 5-4-1 – это многослойный перцептрон с пятью входами, четырьмя нейронами в скрытом слое и одним нейроном выходного слоя; МП 6-4-1 – это многослойный перцептрон с шестью входами, четырьмя нейронами в скрытом слое и одним нейроном выходного слоя; нейронная сеть РБФ содержит три слоя: входной слой; слой скрытых нейронов с радиально-симметричной активационной функцией, каждый j -й из которых предназначен для хранения отдельного эталонного вектора в виде вектора весов $w_j(h)$; выходной слой

Было выявлено, что радиальные базисные сети обладают рядом преимуществ перед сетями типа многослойный персептрон. Во-первых, они моделируют произвольную нелинейную функцию с помощью одного промежуточного слоя. Тем самым отпадает вопрос о числе слоев. Во-вторых, параметры линейной комбинации в выходном слое можно полностью оптимизировать с помощью известных методов моделирования, которые не испытывают трудностей с локальными минимумами, мешающими при обучении МП. Поэтому сеть РБФ обучается очень быстро (на порядок быстрее МП).

С другой стороны, до того как применить линейную оптимизацию в выходном слое сети РБФ, необходимо определить число радиальных элементов, положение их центров и величины отклонений. Для устранения этой проблемы предлагается использовать автоматизированный конструктор сети, который выполняет за пользователя основные эксперименты с сетью.

Другие отличия работы РБФ от МП связаны с различным представлени-

ем пространства модели: «групповым» в РБФ и «плоскостным» в МП. Опыт показывает, что для правильного моделирования типичной функции сеть РБФ требует несколько большего числа элементов. Следовательно, модель, основанная на РБФ, будет работать медленнее и потребует больше памяти, чем соответствующий МП (однако она гораздо быстрее обучается, а в некоторых случаях это важнее).

С «групповым» подходом связано и неумение сетей РБФ экстраполировать свои выводы за область известных данных. При удалении от обучающего множества значение функции отклика быстро падает до нуля. Напротив, сеть МП выдаст более определенные решения при обработке сильно отклоняющихся данных, однако в целом склонность МП к некритическому экстраполированию результата считается его слабостью. Сети РБФ более чувствительны к «проклятию размерности» и испытывают значительные трудности, когда число входов велико.

Таблица 2

Результаты поиска оптимальных нейросетевых структур при проведении исследования (Методы искусственных нейронных сетей и влияние экзогенных переменных)

№ п/п	Архитектура	Производительность обучения	Ошибка обучения	Контрольная ошибка	Корреляция
1	МП 5-4-1	0,401670	0,084714	0,085163	0,90129
2	МП 6-4-1	0,409401	0,085963	0,082306	0,89730
3	РБФ 13-29-1	0,399905	0,042725	0,046866	0,87924
4	РБФ 12-44-1	0,372236	0,039769	0,044508	0,89125
5	РБФ 12-67-1	0,370119	0,039542	0,039268	0,89041

Таблица 3

Результаты прогноза пяти наилучших сетей

Номер наблюдения	Выход	МП 5-4-1	МП 6-4-1	РБФ 13-29-1	РБФ 12-44-1	РБФ 12-67-1
1	19,30000	16,46174	17,52021	18,15556	18,69394	20,23986
2	22,00000	19,18554	21,64104	20,24270	24,02081	22,60867
3	20,30000	20,37075	22,07099	20,99243	23,81311	22,53081
4	20,50000	20,07585	20,98084	19,75282	21,61238	20,09558
5	17,30000	20,59252	20,83783	17,01615	18,12504	16,49583
6	18,80000	19,35636	20,82702	20,12393	21,78268	20,12268
7	21,40000	20,18651	21,81011	21,23228	23,69920	22,15571
8	15,70000	19,24575	20,63956	15,99494	16,97535	15,58635
9	16,20000	16,47351	15,98440	16,54179	15,09492	15,63252
10	18,00000	20,13308	18,21978	19,95714	18,36202	19,22542
11	14,30000	16,09037	15,45824	15,49104	14,27741	14,79159
12	19,20000	23,05850	20,23653	21,77788	20,33284	21,45920
.....
100	23,10000	15,28950	16,22822	18,22000	21,10664	22,01762

Таблица 4

Ошибки регрессии исходного ряда и ряда, построенного выбранной сетью

	МП 5-4-1	МП 6-4-1	РБФ 13-29-1	РБФ 12-44-1	РБФ 12-67-1
Среднее данных	22,59536	22,59536	22,59536	22,59536	22,59536
Ст. откл. данных	9,25768	9,25768	9,25768	9,25768	9,25768
Среднее ошибки	-0,28934	-0,04780	-0,11828	-0,01311	-0,04417
Ст. откл. ошибки	4,01236	4,08657	4,41488	4,19897	4,21786
Среднее абсолютной ошибки	2,86810	2,75566	2,91148	2,81306	2,56776
Отношение ст. откл.	0,43341	0,44143	0,47689	0,45357	0,45561
Корреляция	0,90129	0,89730	0,87924	0,89125	0,89041

В результате получили 5 обученных сетей с определенной архитектурой, которые могут прогнозировать оценку рыночной стоимости недвижимости (рисунок) при 13-ти входах исключительно по данным изменения стоимости. Как видим, коэффициент корреляции примерно одинаков для всех пяти сетей, что говорит о малой точности прогноза. По результатам опыта можно сказать, что все сети справились с поставленной задачей одинаково. Однако на некоторых значениях выхода радиально-базисная сеть имеет значительные отклонения от ожидаемого значения.

Построенная модель позволяет повысить эффективность управления комплексами недвижимости в масштабах города и сделать этот механизм более прозрачным.

Список литературы

1. Галушкина А.И., Цыпкина Я.З. Нейронные сети: история развития теории. – М.: ИПРЖР, 2001. – 576 с. – С. 347–352.
2. Исавнин А.Г., Галиев Д.Р. Модели портфельного инвестирования с применением асимметричных мер риска

и генетических алгоритмов // Финансовая аналитика: проблемы и решения. – 2011. – № 48. – С. 32–38.

3. Исавнин А.Г., Фархутдинов И.И. Метод оценки экономической эффективности применения производственного аутсорсинга на автомобилестроительном предприятии России // Региональная экономика: теория и практика. – 2012. – № 13. – С. 16–21.

4. Круглов В.В., Дли М.И., Голунов Р.Ю. Нечёткая логика и искусственные нейронные сети. – М.: Физматлит, 2001. – 70 с. – С. 40–44.

5. Тархов Д.А. Нейронные сети. Модели и алгоритмы: справочник. – М.: Радиотехника, 2005. – 250 с. – С. 19–23.

References

1. Galushkina A.I., Cypkina Ja.Z. Nejrionnye seti: istorija razvitija teorii. M.: IPRZhR, 2001. 576 p. pp. 347–352.
2. Isavnin A.G., Galiev D.R. Modeli portfel'nogo investirovaniya s primeneniem asimmetrichnyh mer riska i geneticheskikh algoritmov // Finansovaja analitika: problemy i reshenija. 2011. no. 48. pp. 32–38.
3. Isavnin A.G., Farhutdinov I.I. Metod ocenki jekonomicheskoj jeffektivnosti primeneniya proizvodstvennogo aoutsorsinga na avtomobilestroitel'nom predpriyatii Rossii // Regional'naja jekonomika: teorija i praktika. 2012. no. 13. pp. 16–21.
4. Kruglov V.V., Dli M.I., Golunov R.Ju. Nechjotkaja logika i iskusstvennye nejronnye seti. M.: Fizmatlit, 2001. 70 p. pp. 40–44.
5. Tarhov D.A. Nejrionnye seti. Modeli i algoritmy: spravochnik. M.: Radiotehnika, 2005. 250 p. pp. 19–23.