

*Петрова Е. А. канд. тех. наук, доц.,
Солодовник А. В., канд. тех. наук, доц.*

Харьковский национальный университет строительства и архитектуры

НЕЙРОСЕТЕВЫЕ ТЕХНОЛОГИИ В ЗЕМЕЛЬНОМ ДЕВЕЛОПМЕНТЕ

a.solodovnyk@gmail.com

Рассматривается процесс создания информационной модели, способной по входным известным рыночным параметрам спрогнозировать стоимость земельного участка в городе, пригороде или Харьковской области. Показано использование метода Data Mining – нейронная сеть, для определения стратегии поведения фирмы на рынке земли.

Ключевые слова: *земельный девелопмент, стоимость сотки земли, интеллектуальный анализ данных, нейронные сети.*

Постановка проблемы.

Уровень аналитической обработки данных в фирмах, занимающихся земельным девелопментом, и потребность в получении такими фирмами объективной информации о параметрах, влияющих на целевой показатель земельного девелопмента – стоимость сотки земли, определяет поведение фирмы на рынке. Поэтому актуальной задачей является разработка информационной модели, способной по входным известным рыночным параметрам спрогнозировать стоимость земельного участка в городе, пригороде или Харьковской области. Одним из способов решения такой задачи являются средства и методы обработки данных Data Mining [1,2]. Data Mining – это процесс обнаружения в данных ранее неизвестных, нетривиальных, практически полезных и доступных интерпретации знаний, необходимых для принятия решений в различных сферах человеческой деятельности [3].

Краткий анализ последних достижений и публикаций. В современных условиях экономического развития Украины за последние десятилетия наблюдался бурный рост компьютерных технологий, методов и моделей искусственного интеллекта, направленный на повышение эффективности и прибыльности бизнеса.

Информационные технологии позволяют решать сложные задачи экономического характера, связанные с анализом данных и принятием обоснованного управленческого решения. Известно, что для решения задач прогноза наряду с классическими методами статистического прогнозирования, моделирования, экспертного поиска решения, успешно используются нейросетевые технологии, представляющие перспективное направление развития искусственного интеллекта. Метод обнаружения данных – нейронные сети – используются в банковской сфере, промышленности, экономике, бизнесе, геолого-разведке, управлении, транспорте и других областях [4].

Авторами статьи нейросетевая технология используется для повышения эффективности и прибыльности бизнеса. Для любой фирмы, которая занимается развитием земельных участков, актуальность решения задач прогнозирования очень высока, так как успех в рыночных условиях невозможен без ясного представления о перспективах и последствиях принимаемых решений.

В предлагаемой статье авторами использовано концепцию и технологию современного компьютерного анализа данных для прогнозирования цены сотки земельного участка, расположенного в городе, пригороде или области.

Задача прогнозирования решалась методом многослойных нейронных сетей. После эмпирического выбора архитектуры нейросети и ее обучения были получены прогнозные значения результирующего признака – цены сотки земельного участка.

Постановка задачи. При использовании метода Data Mining авторы столкнулись с проблемой выбора значимых для модели входных признаков (в зарубежной литературе такая проблема известна как feature selection). Известно, что сокращение числа независимых переменных призвано уменьшить размерность модели не только с тем, чтобы удалить из нее все незначимые признаки, не несущие в себе какой-то полезной для анализа информации, и тем самым упростить модель, но и чтобы устранить избыточные признаки [1]. Авторы предположили, что с использованием нейронной сети формально нелинейный объект можно рассматривать как функцию зависимости цены сотки земли от таких показателей: расположение земельного участка, площадь земельного участка, целевое назначение земельного участка, права собственности на земельный участок, наличие дома, построек, наличие коммуникаций.

Методика нейросетевого прогнозирования стоимости сотки земельного участка представляет совокупность этапов, методов и правил,

которые обеспечивают задачу совершения прогноза.

Представленная база данных земельных участков хранит сведения и позволяет проводить бизнес-операции с недвижимостью. Сведения о цене сотки могут представлять интерес для выработки поведения при покупке земельного участка, исходя из целей покупки, будь-то участок для жилой постройки или для ведения дачного хозяйства, и района расположения участка.

Изложение основного материала. В качестве репрезентативной выборки для обучения

нейросети были выбраны 1000 записей про земельные участки в Харьковской области статистических баз данных ООО ИКК «Проконсул» и ООО «First Capital». С целью повышения эффективности обучения нейросети из характеристик земельных участков в качестве входных параметров отобраны наиболее значимые. В табл. 1 приведены отобранные характеристики земельных участков и их возможные значения или пояснения.

Таблица 1

Характеристики земельных участков

№	Название	Возможные значения
1	Харьков/область	область, пригород, Харьков
2	Площадь	общая площадь участка в сотках
3	Целевое назначение земельного участка	для садоводства, жилая и индивидуальная застройка
4	Право на земельный участок	приватизирован, не приватизирован
5	Наличие постройки	имеется постройка, постройки нет
6	Коммуникации	отсутствуют, рядом, частично и в плохом состоянии, частично, полностью в плохом состоянии, имеются полностью

Обработка исходных данных была проведена предварительно для устранения неточностей связанных с использованием характеристик имеющих большое количество вариантов текстовых значений. Так характеристики Право на земельный участок и Наличие постройки приобрели тип булевых переменных, а значения параметра Коммуникации были приведены к числовым в соответствии с табл. 2.

Таблица 2

Соответствия первоначальных значений обработанных характеристик числовым аналогам

Первоначальное значение	Числовое значение
Характеристика Коммуникации	
отсутствуют	0
рядом	1
частично и в плохом состоянии	2
частично	3
полностью в плохом состоянии	4
имеются полностью	5
Характеристика Право на земельный участок	
приватизирован	1
не приватизирован	0
Характеристика Наличие постройки	
имеется постройка	1
постройки нет	0

Для моделирования входные данные были представлены в виде файла формата .txt и импортированы в аналитическую платформу Deductor [1], в которой нейросетевые технологии представлены в виде нейронных сетей и их разновидности – карт Кохонена.

Был использован метод DataMining – нейронные сети, для определения прогнозного значения цены сотки земельного участка города, области и пригорода [2,3]. При импорте данных в аналитическую платформу Deductor в качестве входных полей выбирались: Харьков/область, площадь участка, целевое назначение, право на земельный участок, наличие построек, коммуникации, а выходным полем – цена за сотку.

На рис. 1 приведен фрагмент базы данных после импорта в аналитическую платформу.

Во время использования нейронных сетей предварительная обработка входных данных влияет на достоверность и точность прогноза не менее чем структура нейросети и способ обучения. Поэтому первым этапом сценария было уничтожение аномалий и сглаживание данных, приведенных в таблице (рис.1) с помощью парциальной обработки. На первом этапе нейросетевого прогнозирования цены сотки земельного участка определялись компоненты входного вектора X и выходной вектор Y. В данном случае на выходной показатель влияют такие независимые переменные: место расположения земельного участка, его площадь, наличие коммуникаций и построек на участке, целевое назначение участка, наличие права собственности на участок.

На втором этапе были определены назначения входных и выходного показателя и их нормализация. При построении нейросети настройка назначений полей выполнялась такая же, как и при импорте базы данных.

Целью третьего этапа было обучение нейронной сети с выполнением следующих шагов:

- разбивка входного набора данных на обучающее и тестовое подмножество;
- обучение нейросети методом обратного распределения ошибок.

Харьков/область	Площадь, сотка	Целевое назначение ЗУ	Права на з/у	Общая стоимость	Цена за сотку	НаличиеДома,Опост
3	15	1	<input type="checkbox"/>	5500	367	<input type="checkbox"/>
3	21,5	1	<input type="checkbox"/>	19999	930	<input type="checkbox"/>
3	25	1	<input type="checkbox"/>	10000	400	<input type="checkbox"/>
2	6	1	<input type="checkbox"/>	7000	1167	<input type="checkbox"/>
2	15	1	<input checked="" type="checkbox"/>	6000	400	<input type="checkbox"/>
2	15	1	<input type="checkbox"/>	5000	333	<input type="checkbox"/>
2	15,5	1	<input type="checkbox"/>	80000	5161	<input checked="" type="checkbox"/>
3	10	1	<input type="checkbox"/>	9000	900	<input type="checkbox"/>
3	70	1	<input type="checkbox"/>	100000	1429	<input type="checkbox"/>
2	15	1	<input type="checkbox"/>	7900	527	<input type="checkbox"/>
3	15	1	<input checked="" type="checkbox"/>	5200	347	<input type="checkbox"/>
3	15	1	<input type="checkbox"/>	4000	267	<input type="checkbox"/>
3	17	1	<input type="checkbox"/>	8000	471	<input type="checkbox"/>
3	20	1	<input type="checkbox"/>	10000	500	<input type="checkbox"/>
3	21	1	<input type="checkbox"/>	6999	333	<input type="checkbox"/>
2	10	1	<input checked="" type="checkbox"/>	11500	1150	<input type="checkbox"/>
2	20	1	<input checked="" type="checkbox"/>	15000	750	<input type="checkbox"/>
2	15	1	<input type="checkbox"/>	15000	1000	<input type="checkbox"/>
2	25	1	<input type="checkbox"/>	25000	1000	<input type="checkbox"/>
3	25	1	<input type="checkbox"/>	10000	400	<input type="checkbox"/>
3	25	1	<input checked="" type="checkbox"/>	8000	320	<input type="checkbox"/>
2	10,5	1	<input checked="" type="checkbox"/>	8500	810	<input type="checkbox"/>
2	15	1	<input type="checkbox"/>	6000	400	<input type="checkbox"/>
3	25	1	<input checked="" type="checkbox"/>	4200	168	<input type="checkbox"/>
2	10	2	<input type="checkbox"/>	5000	500	<input type="checkbox"/>
2	10	1	<input type="checkbox"/>	10000	1000	<input type="checkbox"/>
2	20	1	<input checked="" type="checkbox"/>	65000	3250	<input checked="" type="checkbox"/>
2	8	1	<input type="checkbox"/>	16500	2063	<input checked="" type="checkbox"/>
2	10	1	<input type="checkbox"/>	9000	900	<input type="checkbox"/>
2	14	1	<input type="checkbox"/>	7000	500	<input type="checkbox"/>
2	15	1	<input checked="" type="checkbox"/>	20000	1333	<input checked="" type="checkbox"/>
2	29	1	<input type="checkbox"/>	16000	552	<input type="checkbox"/>
3	41	1	<input checked="" type="checkbox"/>	45000	1098	<input type="checkbox"/>
3	10	1	<input checked="" type="checkbox"/>	10000	1000	<input type="checkbox"/>
3	10	1	<input type="checkbox"/>	8000	800	<input type="checkbox"/>

Рис. 1. Импортированная база данных

В результате проведенного моделирования были получены следующие результаты: после обучения корректный прогноз был осуществлен на 98,1 % обучающей выборки и на 98 % тестовой выборки.

На рис.2 представлена нейросеть, в которой в качестве входных показателей использовались

5 показателей. Характеристика Целевое назначение было системой воспринято как неиспользуемое. Это можно объяснить тем, что цена сотки земельного участка, предназначенного под жилую застройку, и постройку дачи отличаются несущественно.

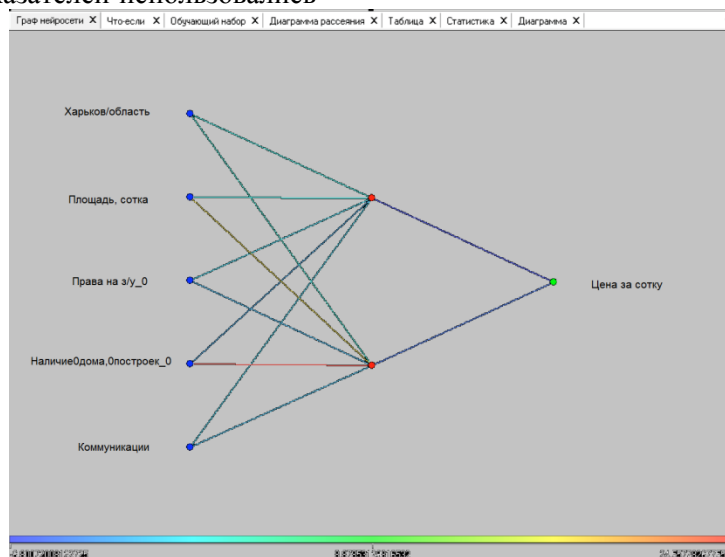


Рис. 2. Граф нейросети

При вычислении прогнозного значения, в визуализаторе «Что-если», определена зависимость, путем сочетания различных

исходных данных, влияние которых производит вычисление показателя цены Y (рис3).

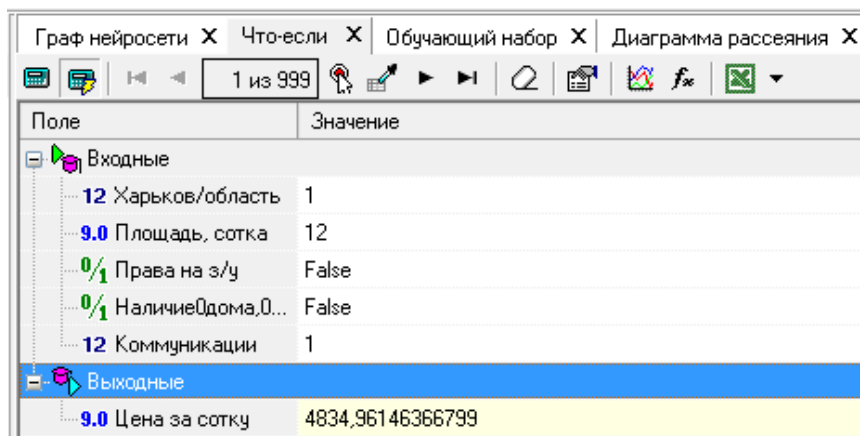


Рис. 3. Визуализатор «Что-если»

Выбор модели — это всегда бизнес-решение. Следующий шаг состоит в том, чтобы использовать модель в реальной жизни. На четвертом этапе, на основании построенной нейросети, строился прогноз изменения цены сотки земельного участка при условии, что

земельный участок выбран в пригороде Харькова площадью 15 соток с имеющимися постройками, коммуникациями и отсутствием права собственности на этот участок. Прогнозируемое значение цены приведено на рис. 4.

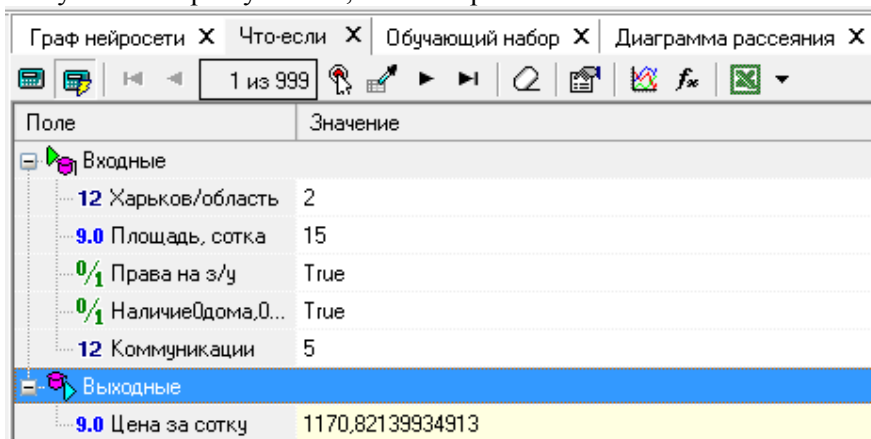


Рис. 4. Прогноз нейросети

Необходимо отметить, что спрогнозированное нейросетью значение стоимости сотки земли в пригороде с указанными показателями несущественно отличается от значения стоимости сотки земельного участка в пригороде, выставленного на продажу (данные реальной базе данных 1333 грн.). Данные визуализатора «Что-если» (рис. 3) также подтверждают правильность работы нейросети: земельный участок в центре

Харькова площадью 12 соток даже без наличия государственного права собственности, отсутствием построек и коммуникаций стоит в реальной жизни 5000 грн, а спрогнозированное значение равно 4835 грн.

Стоимость сотки выставленных на продажу реальных земельных участков с помощью фильтра отобрано из существующей базы данных и представлено на рис.5,6.

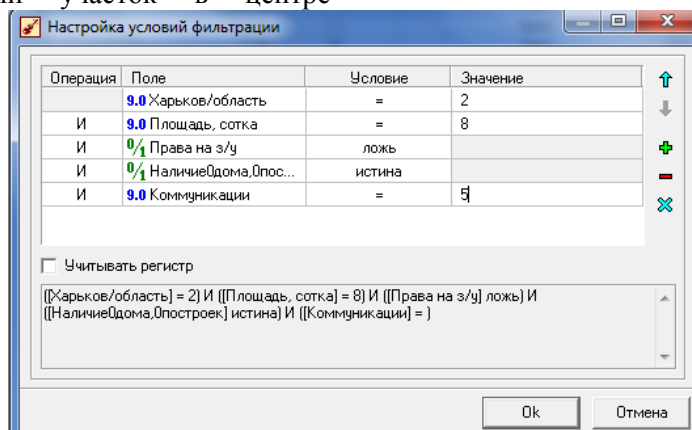


Рис. 5. Настройка условий фильтра

Целевое назначение ЗЧ	Права на з/у	Общая стоимость	Цена за сотку	НаличиеОдома,Опостроек	Коммуникации	Цена за сотку_01
1	<input type="checkbox"/>	16500	2063	<input checked="" type="checkbox"/>	5	1692.993563

Рис. 6. Результат фильтрации

Выводы. В данной статье была освещена тема создания прогнозирующей модели на примере определения цены земельного участка и продемонстрировано решение проблем при помощи инструментов Data Mining платформы Deductor. Для построения модели была проведена предварительная обработка данных, а в дальнейшем применен математический метод анализа – нейронная сеть. Построенную модель можно использовать для принятия решений, пояснения причин, оценки значимости факторов, моделирования различных вариантов сценариев в сфере развития земельных участков различных форм собственности в Харькове и Харьковской области.

БИБЛИОГРАФИЧЕСКИЙ СПИСОК

1. <http://www.basegroup.ru>
2. Дюк В.А., Самойленко А.П. DataMining: учебный курс. СПб.: Питер, 2001. 368 с.
3. Чубукова И. А. Data Mining: Учеб. пособие. М.: Интернет-университет информационных технологий: БИНОМ: Лаборатория знаний, 2006. 382 с.
4. Паклин, Н. Б., Орешков В. И. Бизнес-аналитика: от данных к знаниям (+CD) : учеб. пособие - 2-е изд., перераб. и доп. СПб.: Питер, 2010. 704 с.