

## ПРОГНОЗ СПРОСА НА ТОВАР С ПОМОЩЬЮ НЕЙРОННЫХ СЕТЕЙ В УСЛОВИЯХ МЕНЯЮЩЕЙСЯ РАЗМЕРНОСТИ ВХОДНЫХ ДАННЫХ

*К.Д. Зюсько, Белорусский государственный университет информатики и радиоэлектроники, zyuusko.kirik@gmail.com*

**УДК 654.16**

**Аннотация.** Рассмотрена проблема интерпретации геоданных складской логистики при обучении нейронных сетей для предсказания спроса на товар и предложено решение, позволяющее снизить их размерность, избавиться от необходимости перестроения модели при добавлении новых точек, и позволяющее делать прогнозы, учитывающие изменение координат точки.

**Ключевые слова:** машинное обучение; комбинация алгоритмов; временные ряды; прогнозы спроса; нейронные сети; кластеризация; обработка геоданных.

### FORECAST OF GOODS DEMAND BY NEURAL NETWORKS IN CONDITIONS OF VARIABLE DIMENSION OF INPUT DATA

*Kiryl Ziusko, Belarusian state university of informatics and radioelectronics.*

**Annotation.** The problem of interpreting geodata during training neural networks for predicting demand for goods is considered, and a solution is proposed that reduces their dimensionality, eliminates the need to rebuild the model when adding new points, and allows to make forecasts that take into account changes in the coordinates of the point.

**Keywords:** machine learning; a combination of algorithms; time series; demand prediction; neural networks; clustering; geodata processing.

#### **Введение**

Проблема предсказания спроса с целью минимизации накладных расходов всегда была актуальной. В настоящее время существует огромное количество решений этой проблемы: добавление «буфера безопасности», расчет оптимального запаса на основе известного распределения вероятности, предсказание спроса с помощью временных рядов (*ARIMA*, *ARMA* и др.), использование алгоритмов машинного обучения, где, как правило, чаще всего используются нейронные сети.

Каждое из этих решений хорошо справляется со своей поставленной задачей, определенной условиями и ограничениями. Некоторые из них могут легко масштабироваться, но иметь довольно низкую точность прогнозирования, а некоторые, напротив имеют довольно высокую точность, но плохо масштабироваться (под масштабируемостью понимается возможность предсказывать спрос для нескольких складов).

В данной статье рассмотрена проблема использования и правильной интерпретации геоданных в процессе обучения нейронной сети.

#### **Применение нейронных сетей для прогнозирования**

На сегодняшний день использование нейронных сетей в различных областях приложения является наиболее популярным решением [1-3]. К основным преимуществам нейронных сетей относятся высокая гибкость и наиболее точные предсказания.

Одним из главных достоинств нейронных сетей является тот факт, что они могут обучаться на данных и искать взаимосвязь между входными данными и конечным результатом. Архитектура нейронных сетей позволяет находить

нелинейные зависимости, которые могут быть весьма неочевидными и довольно сложными для восприятия. Таким образом, после обучения, нейронная сеть представляет собой очень сложную функцию, на вход которой мы можем подать много параметров, а на выходе получить конечный результат. Причем во время обучения, нейронная сеть сама поймет насколько важен тот или иной параметр, и будет учитывать его влияние при подсчете итогового значения [4].

Для прогнозирования временных рядов используют различные архитектуры нейронных сетей: глубокие, рекуррентные и другие [5]. Рассмотрим здесь некоторые входные параметры, используемые при построении глубокой нейронной сети для предсказания спроса и их интерпретацию. В табл. 1 приведен анализ входных параметров.

Таблица 1.

Факторы для анализа	Значение каждого фактора	Количество нейронов во входном слое
Последние 14 продаж	Тренд	14
Неделя в году	Сезонность	53 (количество недель в году)
День недели	Изменчивость спроса в зависимости от дня недели	7 (количество дней в неделе)
Класс склада	Специфические паттерны	4 (A,B,C,D)
Цена за аренду	Изменчивость в зависимости от цены	1 (число)
Тип помещения	Специфические паттерны	5 (в соответствии с классификацией помещений для хранения)
Свободное место в помещении	Ограничивающий параметр (чтобы спрос не вышел больше, чем доступно фактического места)	1 (число)
Праздники	Влияние календарных праздников (8 Марта, 1 января и др.)	1 (да или нет)
		Итоговое число входных нейронов: 86

Итого выходит 86 параметров, которые можно менять и оценивать конечный результат. То есть, можно моделировать различные ситуации, например, как изменится спрос перед выходными днями, если мы повысим цену хранения и у нас остается мало места на складе, и, зная, как он изменится, принимать то или иное решение.

Такая нейронная сеть способна довольно точно предсказывать спрос для одного склада. Если создавать нейронную сеть, которая сможет делать предсказания для нескольких точек, то надо добавлять  $n$ -входных параметров [6]. Нейронные сети устроены так, что нельзя создать лишь один входной параметр, куда подавать на вход порядковый номер склада, для которого мы хотим сделать предсказание. Также нельзя создавать два входных нейрона (координаты (широта и долгота), которые могут однозначно идентифицировать точку на поверхности Земли) такие входные данные будут просто бессмысленными для нейронной сети, и она не сможет увидеть взаимосвязи между ними и конечными результатами. Поэтому необходимо создание  $n$ -входных параметров, которые будут получаться с помощью *one-hot* кодирования идентификатора склада. Для наглядности покажем, как меняется количество параметров при обучении модели с 847 складами. В табл. 2 приведено сравнение различного числа параметров.

Таблица 2.

Сценарий 1		Сценарий 2	
Входной параметр	Количество нейронов	Входной параметр	Количество нейронов
Предыдущая конфигурация сети	86	Предыдущая конфигурация сети	86
		Расположение склада	847
Количество входных нейронов	86	Количество входных нейронов	933

Очевидно, что одним из недостатков данной архитектуры является тот факт, что входное число параметров напрямую зависит от количества складов. То есть, при добавлении любого нового склада придется перестраивать нейронную сеть и заново ее обучать. Также, несмотря на то, что нейронная сеть одна, она будет расценивать каждый склад как индивидуальный «юнит». Таким образом, ответить на вопрос как изменится спрос, если склад куда-нибудь переместить, попросту будет невозможно, потому что мы не обладаем информацией – на какой вход нейронной сети пустить правильный сигнал.

Однако местоположение складов определяется не случайным образом [7]. На их расположение влияет очень большое количество факторов. Так, например, можно заметить, что много складов располагаются рядом с границами других крупных государств, от которых, очевидно, и идут поставки. В крупных городах, напротив, складов не очень много, поскольку, в этих городах, как правило, очень высокая плотность застройки и строительство склада где-нибудь в центре города является нецелесообразным ввиду слишком большой стоимости. Часто склады строят около больших городов, либо в мелких городах, которые находятся рядом с большим. Также не удастся встретить склад в глубине лесного массива, потому что с точки зрения логистики и цепочки поставок такое местоположение считается крайне невыгодным.

### **Кластеризация данных**

Таким образом, на местоположение склада влияет огромное количество факторов, которое очень трудно описать формальными правилами. Поэтому, построить определенную систему, которая смогла бы определять все возможные факторы, такие как местоположение склада относительно городов, оценка его клиентов, агрегировать все это и предоставлять результат в конечном виде, который впоследствии можно было бы разделить по каким-либо признакам на группы, довольно сложно.

Однако, если посмотреть на расположение складов на карте, то можно понять, что вся необходимая информация уже представлена. Дело в том, что склады, чаще всего, расположены группами. Это обусловлено тем, что те места, где они строились, были наиболее оптимальными с точки зрения многочисленных факторов: логистики, стоимости постройки и др. Поскольку первоначальной задачей было именно выделение тех самых групп, то теперь необходимость в построении сложной, агрегирующей системы отпадает, потому что наши исходные данные можно просто разбить на группы (кластеры). На рис. 1 показаны данные до и после кластеризации.

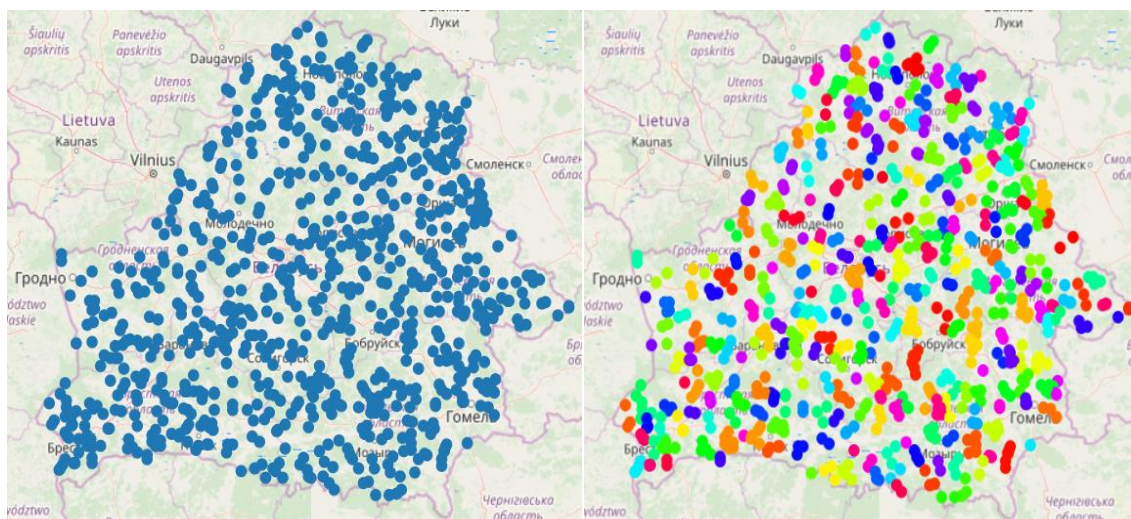


Рисунок 1

На сегодняшний день существует множество алгоритмов: иерархическая кластеризация, метод  $k$ -средних, нейронная сеть Кохонена и многие другие [8]. В данной статье будет рассмотрен метод  $k$ -средних, однако для решения этой задачи можно использовать и другие.

Для работы всех этих алгоритмов нужен единственный параметр – количество выделяемых групп. После его определения алгоритм сможет отнести наши входные данные к разным группам. Единственная проблема, которую можно заметить – мы не знаем заранее, сколько групп необходимо выделить, чтобы нейронная сеть смогла правильно обучиться, т. е. чтобы избежать проблемы недообучения.

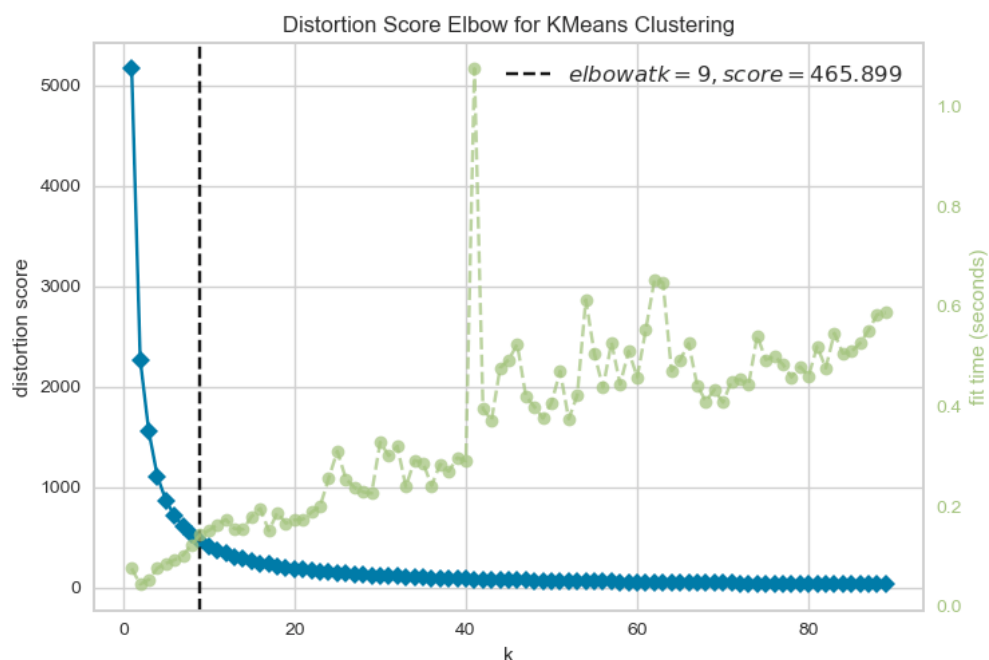


Рисунок 2

Одним из известных методов для определения количества кластеров является локтевой метод, где мы определяем диапазон кластеров и дальше в цикле,

используя каждое значение из этого диапазона как количество кластеров, обучаем модель и считаем ошибку [9]. Далее строится график, где наглядно можно определить, какое количество кластеров необходимо. Обычно ищется то значение, когда ошибка довольно мала, а последующие бóльшие размеры кластеров незначительно уменьшают ее. Данный подход используется повсеместно, однако для того, чтобы использовать его для наших целей нам придется его немного изменить. А именно: считать ошибку не от алгоритма обучения без учителя (метод *k*-средних, в нашем случае), а от обучения нейронной сети на данном количестве кластеров. Таким образом, можно понять, какое количество кластеров будет достаточно для корректного обучения. На рис. 2 показаны результаты применения локтевого метода. В табл. 3. приведена зависимость ошибки от количества кластеров.

Таблица 3.

Количество кластеров	Ошибка ( <i>MSE</i> )
9	1238519,6392
20	1018743,8351
30	814891,1352
50	402840,7529
100	245921,2301
200	162547,9165
300	137994,0376
500	135204,4719
847	132994,0153

Согласно графику и табл. 3, оптимальное количество кластеров в каждом случае разное и зависит от того, от чего считается ошибка. Т. е., судя по графику и по *distortion score k*-средних, девяти кластеров достаточно, чтобы правильно разбить данные. Однако, этого количества кластеров явно недостаточно для обучения нейронной сети, потому что ошибка обучения при девяти кластерах явно большая. По мере увеличения количества кластеров видно, что ошибка уменьшается, и если разбить данные на 300 кластеров, то ошибка уже отличается незначительно от той ошибки, когда нейронная сеть обучалась на некластеризованных данных (847 кластеров соответствуют общему количеству складов). Таким образом, можно сделать вывод, что 300 кластеров является оптимальным количеством. Возможно, это количество является даже немного бóльшим, чем нужно, однако, это не является проблемой, поскольку нейронная сеть сама сможет найти взаимосвязь между кластерами и выходными данными. Гораздо бóльшей проблемой будет ситуация, когда выбрано меньшее число кластеров и нейронная сеть не сможет правильно обучиться на этих данных.

После выбора оптимального количества кластеров можно запускать обучение нейронной сети, оптимизировать гипер-параметры и проверять результаты обучения. На рис. 3 можно увидеть, как обученная нейронная сеть предсказывает значения (зеленый цвет – реальные значения, красный цвет – предсказанные).

Построить систему, идеально предсказывающую спрос, практически невозможно, потому что иногда некоторые факторы проявляются случайным образом, и установить взаимосвязь между этими факторами, спросом, и коэффициентом влияния этого фактора довольно сложно. И, как было сказано выше, добавление кластеризации для уменьшения размерности входных данных, скорее всего, добавит чуть больше неточности к предсказаниям, однако вместе с

ЭТИМ ПОЗВОЛИТ РЕШИТЬ МНОГИЕ ВОПРОСЫ, ТАКИЕ КАК: ДИНАМИЧЕСКАЯ МНОГОМЕРНОСТЬ ДАННЫХ, БОЛЬШИЕ НАКЛАДНЫЕ РАСХОДЫ ПАМЯТИ ПРИ ОБУЧЕНИИ, ДОЛГОЕ ОБУЧЕНИЕ, УЧЕТ КОНТЕКСТА ГЕОЛОКАЦИИ ДЛЯ ПРОГНОЗИРОВАНИЯ ВЫХОДНЫХ ЗНАЧЕНИЙ И МНОГИЕ ДРУГИЕ.

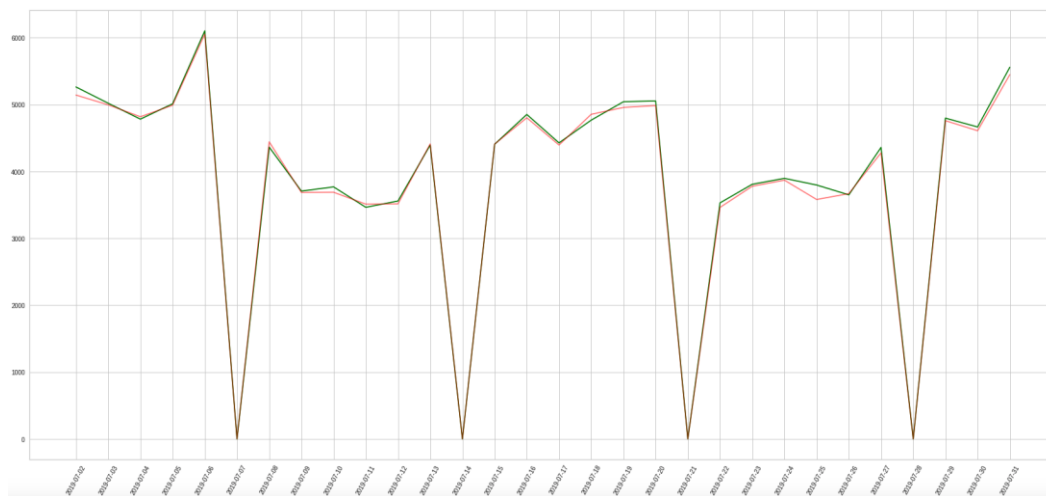


Рисунок 3

### Заключение

В заключении хочется отметить, что в данной статье была рассмотрена проблема интерпретации геоданных в процессе обучения нейронных сетей, рассмотрены существующие подходы, их проблемы и ограничения. Была показана особенность этих данных, и на основе этого предложено решение (применение кластеризации к данным), которое позволяет обойти ограничения. В статье также уделено внимание эффективности этого решения и рассказано о некоторых нюансах его применения, в частности, раскрыто условие, что выбор оптимального количества кластеров должен рассчитываться не от ошибки кластеризации, а от ошибки обучения нейронной сети.

### Литература

1. <https://www.hindawi.com/journals/jam/2014/614342.pdf> (дата обращения – февраль 2020 г.).
2. Максименко В.Н., Волошина Т.С. Анализ системы распознавания лиц по алгоритму нейронной сети // Экономика и качество систем связи, 2018. – № 4. – С. 31-37.
3. Данков Н.И. Исследование возможностей нейросетевых технологий в области идентификации голоса // Экономика и качество систем связи, 2018. – № 3. – С. 47-52.
4. Саймон Х. Нейронные сети: полный курс, 2-е издание. Пер. с англ. – М.: Издательский дом «Вильмс», 2006. – 1104 с.
5. <https://arxiv.org/abs/1809.04356.pdf> (дата обращения – февраль 2020 г.).
6. <http://www.diva-portal.se/smash/get/diva2:1259073/FULLTEXT01.pdf> (дата обращения – апрель 2020 г.).
7. [https://rep.bntu.by/bitstream/handle/data/24352/Logistika\\_skladirovaniya.pdf](https://rep.bntu.by/bitstream/handle/data/24352/Logistika_skladirovaniya.pdf) (дата обращения - апрель 2020 г.)
8. <https://pdfs.semanticscholar.org/7e5e/88e931f511a24b156e49e62263e5aae49f50.pdf> (дата обращения – февраль 2020 г.).
9. <https://www.scikit-yb.org/en/latest/api/cluster/elbow.html> (дата обращения – февраль 2020 г.).