DOI <https://doi.org/10.18551/rjoas.2017-09.31>

**ОЦЕНКА ЭФФЕКТИВНОСТИ НОВЫХ ПРОДУКТОВ ПИТАНИЯ
С УЧЕТОМ ИСПОЛЬЗОВАНИЯ ИННОВАЦИЙ**
EVALUATION OF THE EFFECTIVENESS OF NEW FOOD PRODUCTS
TAKING INTO ACCOUNT THE USE OF INNOVATIONS

Лоза А.А., научный сотрудник
Loza A.A., Researcher

**Санкт-Петербургский национальный исследовательский университет
информационных технологий, механики и оптики, Санкт-Петербург, Россия**
St. Petersburg National Research University of Information Technologies, Mechanics and
Optics, St. Petersburg, Russia
E-mail: projecttechnic@mail.ru

АННОТАЦИЯ

Нейронные сети успешно применяются в экономике, бизнесе, технике, физике. С их помощью можно решать задачи прогнозирования, классификации или управления и, в более широком плане, задачи распознавания образов. Нейронные сети нелинейны по своей природе. В задачах, где линейная аппроксимация неудовлетворительна, линейные модели работают плохо. Нейронные сети учатся на примерах. Пользователь нейронной сети подбирает представительные данные, а затем запускает алгоритм обучения, который автоматически воспринимает структуру данных. Сравнительно простые нейронные сети, которые строит система, например, ST Neural Networks или NeuroShell 2, являются мощным инструментом в арсенале специалиста по прикладной статистике. В работе рассмотрен с названных позиций рынок продуктов питания, имеющие сбалансированные рецептуры.

ABSTRACT

The neural networks are successfully used in economics, business, engineering and physics. They can be used to solve the problems of prediction, classification and management, and generally, the problems of pattern recognition. The neural networks are non-linear in nature. The models do not work well in the problems where the linear approximation is not satisfactory. The neural networks learn by the examples. The user of neural network selects representative data, and then runs a learning algorithm which automatically accepts the data structure. Rather simple neural network, which is building the system, for example, ST Neural Networks or NeuroShell 2, is a powerful tool in the arsenal of an expert in the applied statistics. From the named position in this article we consider the food market with balanced formulas.

КЛЮЧЕВЫЕ СЛОВА

Нейронные сети, продукты питания, распознавание образов.

KEY WORDS

Neural networks, food, pattern recognition.

Современная индустрия продуктов питания, как и другие отрасли, реальной экономики стремится производить продукцию, удовлетворяющую как можно более широкому спектру потребителей, но с наименьшими производственными издержками. Специфика этой отрасли состоит в том, что ее продукция должна удовлетворять еще одному важному требованию – максимально соответствовать здоровому образу жизни.

Одним из наиболее апробированных путей решения такой задачи является выпуск продуктов питания сбалансированных по основным питательным веществам.

Пусть некоторая фирма, разрабатывающая и производящая новые продукты питания, имеющие необходимые для жизни человека пищевые вещества, такие как белки, жиры и углеводы вводит в их состав некоторые витаминные добавки.

Учитывая конкуренцию на рынке, целесообразно создать некоторую экономико-аналитическую систему, которая поможет продавать продукцию этой фирмы.

Стоимость продуктов питания в регионе, где работает фирма, относительно высока, и можно надеяться, что новые рецептуры продуктов питания существенно сэкономят деньги покупателей. Разрабатываемая система, должна предсказывать ежедневную стоимость продуктов питания, которые продает фирма, при условии удовлетворения запроса потребителя на отдельные группы пищевых веществ в зависимости от калорийности.

После того, как принято решение о цели прогноза, нужно решить, какие переменные следует принять во внимание для осуществления такого предсказания. Очевидно, что на стоимость продуктов питания потребленных за день, влияет несколько факторов, в том числе количество жиров, количество белков, количество углеводов и т.д. При первой попытке построить систему, например, на основе нейронной сети задача остается простой, поэтому можно использовать в качестве наиболее важной переменной среднее количество жиров потребляемых за день, для предсказания дневной стоимости продуктов питания одним потребителем.

Заметим, что до того, как решиться использовать нейронную сеть, необходимо рассмотреть возможность использования регрессионного анализа. В рассматриваемом случае один момент делает построение линейной модели сложным: если содержание жиров лежит в диапазоне от 32 до 45%, то он весной и летом такие продукты питания плохо потребляются, поэтому дневная стоимость потребляемых продуктов питания нелинейна [1].

После того, как принято решение, о цели прогнозирования и составе переменных используемых для такого предсказания, можно приступить к использованию экономико-аналитической системы на базе нейронной сети.

Наиболее просто построить такую систему можно с помощью программы NeuroShell 2.

Обычный порядок действий для создания приложения в NeuroShell 2 состоит в последовательном переходе от значка к значку слева направо, сверху вниз.

В процессе введения начальных данных необходимо набрать данные за прошлый год. В рассматриваемом случае необходимо вводить ежедневные отчеты стоимости проданных продуктов питания за год, рассортированные в порядке от самой низкой до самой высокой стоимости.

Для успешного построения системы нужно проинформировать NeuroShell 2 о том, какие из столбцов являются входами, а какие - выходами.

Можно обозначить столбец "Среднее содержание жиров" как вход, I, а столбец "Стоимость/день" как столбец, который сеть пытается предсказать (выход, A).

Следующим шагом является ввод минимального и максимального значений для каждой переменной в строках Минимум и Максимум. Поскольку нейронные сети требуют перевода переменных путем масштабирования в диапазоны от 0 до 1 или от -1 до 1, сети необходимо знать истинный диапазон значений переменной. В этом модуле можно ввести минимальное и максимальное значения для каждой переменной, которая должна использоваться сетью, или можно вычислить диапазон автоматически из введенных данных, вызывая пункт «Расчет мин/макс» меню «Установки». Выбор этого пункта меню также приводит к вычислению среднего и стандартного отклонения для каждой переменной [2].

В общем случае, нужно использовать диапазон, границы которого вплотную примыкают к данным. (Можно указать значения минимума и максимума, которые будут чуть меньше и чуть больше соответствующих значений в файле данных, чтобы предусмотреть более широкий диапазон для будущих предсказаний, или можно предпочесть выбрать диапазон более узкий, чтобы исключить выбросы, которые могут повлиять на точность работы сети.) Если не установить значения минимума и

максимума вплотную к данным, сеть может потерять способность отслеживать мелкие различия в данных.

Деле следует перейти к тренировке сети, щелкая дважды по значку «Обучение», чтобы вызвать модуль тренировки. При появлении этот модуль уже должен знать о количестве входов и выходов из.MMX-файла (созданного в модуле «Выбор входов/выходов»). Тем не менее, есть еще несколько параметров, которые необходимо задать, прежде чем начать тренировку. Во-первых, нужно указать сложность задачи, щелкнув по одному из переключателей в левой верхней части экрана. Можно решить, что имеющиеся данные очень просты. Тогда следует щелкнуть мышью по кнопке «Очень простые». Следует обратить внимание, что при этом значения скорости обучения и момента автоматически устанавливаются равными 0,6 и 0,9.

Затем следует установить количество скрытых нейронов. Щелчок по кнопке вычисления по умолчанию вызывает вычисление количества по встроенной формуле NeuroShell 2, которая дает хорошее начальное приближение для будущих задач.

Можно использовать случайный порядок представления данных сети, так как данные отсортированы в порядке от низких стоимостей к высоким, а это позволяет сети хорошо предсказывать результат при всех стоимостях.

Для первого сеанса тренировки следует устанавливать интервал «Калибровки» равным 0, хотя это мощный инструмент, который следует использовать для большинства задач.

Основной статистический показатель тренировки - это внутренний средний "показатель ошибки". Предварительно вычислить или точно воспроизвести этот средний показатель на тренировочном наборе невозможно, да это и совершенно не нужно. Тем не менее, это среднее по всем примерам значение квадрата ошибки всех выходов, вычисленных в пределах внутреннего интервала NeuroShell 2. Значение этого числа само по себе бесполезно. Полезно в процессе тренировки видеть, улучшается ли качество сети, т.к. по мере улучшения сети показатель уменьшается.

Можно посмотреть, хорошие ли результаты дает построенная сеть. Нужно дважды щелкнуть по значку «Применение» к файлу и выбрать пункт «Начать» применение из меню «Работа». По умолчанию флажки "Вычислять R квадрат и т.д.", "Включать выходы в.OUT-файл" и "Включать в.OUT-файл разности между выходами и ответами сети" включены.

Модуль «Применение» по умолчанию обрабатывает.PAT-файл, который представляет собой первый введенный набор данных. Количество потребляемых жиров каждого дня (вход) обрабатывается натренированной сетью, которая вычисляет стоимость продуктов питания за этот день. На экране отображаются статистические показатели, измеряющие точность работы натренированной сети [3].

Натренированная сеть дает значение R квадрат на.PAT-файле около 0,95. Нужно записать значение R квадрат, чтобы сравнивать эту сеть с другими, которые можно создать позже.

R квадрат, коэффициент множественной детерминации, - это статистический индикатор, обычно используемый при анализе методом множественной регрессии. Он сравнивает точность модели с точностью тривиальной реперной модели, для которой предсказание представляет собой просто среднее по всем примерам. При безупречном совпадении предсказаний с желаемыми значениями R квадрат будет равен 1, при хорошем совпадении - близок к 1, а при очень плохом совпадении близок к 0. Если предсказания построенной нейросетевой модели хуже, чем можно было бы предсказать, просто используя среднее значение выхода по всем примерам, значение R квадрат будет равно 0.

Теперь нужно просмотреть.OUT-файл, в котором показаны реальные значения стоимости за день, которые вводились в файл первоначально, предсказания сети (ее выход) и разница между ними. Можно дополнительно осуществить "приписывание" выходного файла к исходному файлу данных, чтобы видеть исходный файл данных вместе с предсказаниями сети. Модуль «Приписывание» выходного файла сделает

это, создавая при этом новый -OUT-файл. Для вызова этого модуля можно щелкнуть по значку «Приписывание» выходного файла. Поскольку вероятно, нужно контролировать результаты работы сети рядом со значениями входной переменной, следует щелкнуть по значку «Приписать файл сбоку».

Теперь можно оценить.OUT-файл с помощью заполненной Таблицы, дважды щелкнув по значку «Просмотр данных».

Таблица не является электронной таблицей коммерческого класса, и при загрузке больших файлов работает довольно медленно. Если использован очень быстрый компьютер, то это не причинит неудобств, в противном случае необходимо использовать программу работы с электронными таблицами [4-5].

Можно оценить качество модели предсказания стоимости продуктов питания другим способом, не использующим R квадрат.

Это можно сделать, например, посмотрев на данные в графической форме.

Видно, что модель на основе нейронной сети близко следует реальным данным. Предположим, что ответы недостаточно хороши. Что следует делать в этом случае? Самое простое - это попробовать другие архитектуры сети из модуля «Нейронные сети для профессионала».

С другими архитектурами можно получить несколько лучшие результаты, но на практике, вероятно, необходимо сделать что-то из того, что перечислено ниже в порядке убывания вероятности положительного эффекта:

1. Использовать «Калибровку». NeuroShell 2 использует «Калибровку» для оптимизации сети путем применения текущей сети в процессе тренировки к независимому тестовому набору. Можно создать тестовый набор данных автоматически, используя модуль «Выделения» тестового набора. «Калибровка» позволяет найти оптимальную сеть для данных в тестовом наборе, (что означает, что сеть способна хорошо обобщать, и дает хорошие результаты на новых данных).

При использовании «Калибровки» это делается путем вычисления среднего квадратичного отклонения между реальными и предсказанными значениями для всех выходов по всем примерам. При «Калибровке» вычисляется квадратичное отклонение для каждого выхода в данном примере, они суммируются, и затем вычисляется среднее значение этой величины по всем примерам в тестовом наборе [6-8].

Для сетей с обратным распространением ошибки сеть сохраняется всякий раз при достижении нового минимума средней ошибки (или среднего квадратичного отклонения). Чтобы использовать «Калибровку», необходимо установить интервал проверки «Калибровки», т.е. как часто производится оценка на тестовом наборе. Можно устанавливать его в диапазоне от 50 до 200. В этом случае также целесообразно выбрать "Автоматическую запись сети при наилучшем результате на тестовом наборе".

2. Включить в рассмотрение переменные, которые лучше предсказывают то, что сеть пытается предсказать. Другим выходом являются лучшие способы представления переменных, которые уже есть. Можно получить лучшие результаты, включив в рассмотрение переменную для количества минеральных добавок в продукте питания. Увеличение этого показателя непосредственно влияет на калорийность продукта, возможность использования его в качестве лечебно-профилактического и т.д.

Если бы у сети было большое количество входов, возможно, стоило бы преобразовать некоторые входы в отношения. Это дает больше информации при меньшем количестве переменных. Нейронные сети похожи на людей: чем проще сделать входы, которые они должны выучить, тем легче сети выучить задание. Отношения служат этой цели.

3. Подойти заново к вопросу о том, что именно нужно предсказать. Для некоторых вещей это сделать проще, чем для других. Иногда можно получить более высокую точность, предсказывая процентное изменение стоимости продуктов питания, а не саму стоимость.

4. Можно собрать более качественный набор исторических данных или более представительный тестовый набор. Убедитесь в том, что выбранные переменные

нормированы, если это необходимо. Иногда это означает необходимость убедиться в том, что уровни, характерные для ситуации несколько лет назад, переведены в тот же диапазон, что и сегодняшние. В научной области нормировка может означать многие другие вещи. Если нет возможности нормировать данные, придется предъявить сети гораздо большее количество примеров, с соответствующим увеличением времени обучения.

5. Следует подобрать скорость обучения, момент и количество скрытых нейронов и посмотрите, не получатся ли сети более высокого качества. Можно использовать Turboprop, который не требует установки скорости обучения и момента. Этот метод включен в модуль «Проектирования» в системе для профессионала и работает для сетей с обратным распространением ошибки.

Описанный подход построения системы экономико-аналитического анализа стратегии поведения фирмы-производителя на рынке продуктов питания не исчерпывает возможностей применения приемов нечеткой логики в экономических целях.

Не менее эффективным в некоторых случаях являются подходы построения многомерных моделей с помощью модуля Fuzzi Logic пакета программ Matlab или MathCad. В этом случае удастся проанализировать комплексное влияние на выбранную целевую функцию сразу нескольких изменяемых параметров.

Органолептические оценки можно рассматривать как лингвистические переменные теории нечетких множеств, что позволяет одновременно в одной серии экспертных исследований оптимизировать биотехнологические и экономический параметр продукта. В результате органолептической оценки дегустаторами продукта были получены две матрицы оценок.

Таблица 1 – Результаты органолептической оценки продукта

Обозначение матриц	Показатели	Средние арифметические величины показателей						
μ_1 -матрица ингредиента 1	Содержание мг/100 г, g	20	21.4	23	25	27.2	28.6	30
	Средние оценки экспертов	0.12	0.32	0.68	0.88	0.82	0.38	0.06
μ_2 - матрица цены	Цена руб/100 г, h	3.5	4	5	6	10	20	30
	Средние оценки экспертов	1	1	1	1	0.64	0.35	0

Такого рода экономико-аналитические системы часто оказываются эффективны в практике работы пищевых производств и используются для обучения и повышения профессиональной квалификации менеджеров крупных фирм-производителей [9-12].

БИБЛИОГРАФИЯ

1. Алексеев Г.В, Аксенова О.И. Использование нечеткой логики в пищевых производствах [Текст] / Сборник материалов докладов IV международной научно-практической конференции «Фундаментальная наука и технологии – перспективные разработки». Северный Чарльстон. США, 2014. – С. 4-8.
2. Алексеев Г.В. Основы теории решения изобретательских задач // учеб. пособие / Г.В. Алексеев, Н.Б. Жарикова; Федер. агентство по образованию, Гос. образоват. учреждение высш. проф. образования С.-Петерб. гос. ун-т низкотемператур. и пищевых технологий. СПб., 2004.
3. Алексеев Г.В., Бриденко И.И. Виртуальный лабораторный практикум по курсу «механика жидкости и газа». Учебное пособие / Саратов, 2013.
4. Алексеев Г.В., Верболоз Е.И., Кондратов А.В. Возможности управления процессом измельчения путем изменения структурно-механических свойств пищевой смеси. Известия Санкт-Петербургского государственного университета низкотемпературных и пищевых технологий. 2008. № 4. С. 54-58.
5. Арет В.А. Реологические основы расчета оборудования производства жиродержащих пищевых продуктов. Учеб. пособие для студентов вузов, обучающихся по направлению подгот. бакалавров и магистров 551800 - "Технол. машины и оборуд." для дисциплины "Технол. оборуд. при пр-ве жировой продукции"

- / Федер. агентство по образованию, Гос. образоват. учреждение высш. проф. образования С.-Петерб. гос ун-т низкотемператур. и пищевых технологий. СПб., 2004
6. Алексеев Г.В., Аксенова О.И. Использование математического моделирования для ресурсосберегающих пищевых производств / Процессы и аппараты пищевых производств – № 3 – 2014.
 7. Аксенова О.И., Шубенкова В.А. Проектирование многокомпонентных продуктов с использованием теории нечетких множеств / Сборник трудов II заочной Международной научно-практической конференции «Техника и технологии: роль в развитии современного общества». Краснодар. 2013. – С. 34-3.
 8. Аксенова О.И., Шубенкова В.А. Составление рецептурной смеси проектируемых продуктов при неопределенности структурных факторов показателей качества / Сборник трудов Международной научно-технической конференции «Инновационные технологии в пищевой промышленности: наука, образование и производство». Воронеж. 2013. – С.816-820.
 9. Арет В.А., Кулаев Д.Х., Малявко Д.П., Морозов Е.А. О подготовке учебных материалов для обучения инженеров в интернете Научный журнал НИУ ИТМО. Серия: Процессы и аппараты пищевых производств. 2006. № 1. С. 20-25.
 10. Арет В.А., Байченко Л.А., Байченко А.А. Применение нечеткой логики в управлении предприятием пищевой промышленности // Научный журнал НИУ ИТМО. Серия: Экономика и экологический менеджмент / Санкт-Петербургский национальный исследовательский университет информационных технологий, механики и оптики, 2014. - Вып. 15. - № 2. - URL: <http://economics.ihbt.ifmo.ru/file/article/10476.pdf>
 11. Арет В.А., Байченко Л.А., Арет Г.М., Байченко А.А. Рост численности населения Земли и эволюция продовольственных технологий. Научный журнал НИУ ИТМО. Серия: Экономика и экологический менеджмент / Санкт-Петербургский национальный исследовательский университет информационных технологий, механики и оптики, - 2010. - Вып. март 2010. - URL: <http://economics.open-mechanics.com/articles/169.pdf>.
 12. Арет В.А., Пирогов А.Н., Полякова Г.В., Шокальская Л.П. К интенсификации технологических процессов формования на основе распознавания реологических образов // Межвузовский сборник научных работ "Интенсификация существующих и разработка новых технологических процессов в пищевой промышленности. - Москва, 1978.

© 2017 by the author. Licensee RJOAS, Orel, Russia. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license: <http://creativecommons.org/licenses/by/4.0/>