

УДК: 004.8, 004.94, 51-74, 621.37

Моделирование промышленного процесса сушки с применением искусственных нейронных сетей

*Лихтенберг С.Р., студент
Россия, 105005, г. Москва, МГТУ им. Н.Э. Баумана,
кафедра «Робототехнические системы»*

*Мячин Н.В., студент
Россия, 105005, г. Москва, МГТУ им. Н.Э. Баумана,
кафедра «Робототехнические системы»*

*Научный руководитель: Александров Д.В., д.т.н, профессор
Россия, 105005, г. Москва, МГТУ им. Н.Э. Баумана
robot@bmstu.ru*

Постановка задачи

Тертый кокос - это продукт, полученный из кокосов с помощью длительного и непростого процесса трансформации и обработки исходного цельного продукта. В процессе обработки кокосы шелушатся, и их верхнее покрытие отслаивается и удаляется в отход, после чего их раскалывают и измельчают. Измельченный продукт пастеризуют и сушат с получением мелких кокосовых частиц размером около 3 мм. Сушка должна придать частицам окончательную влажность порядка 3 %. Тем не менее, в промышленных условиях это очень трудно обеспечить. В действительности, процент влажности конечного продукта выше этого значения (3 %), что ставит под угрозу сохранение продукта из-за возможного распространения микроорганизмов. Поэтому в реальных условиях осуществляют повторную сушку для решения этой проблемы. С другой стороны, когда значение влажности составляет менее 3 %, вес кокосовой стружки уменьшается, в результате чего возникают экономические потери. В рассматриваемом случае скорость отторжения продукта из-за невыполнения данного условия после первого цикла сушки превышает 90 %. Таким образом, необходимы второй, а иногда и третий цикл для той же партии с целью оптимизации качества выходного продукта. Цель выполненного исследования заключается в уменьшении этого недостатка путем моделирования процесса производства в целях снижения брака во время первого цикла сушки.

Обработка экспериментальных данных

Данные, используемые в этом исследовании, были предоставлены реальной компанией, специализирующейся на кокосовой обработке. Ее основной продукт - тертый кокосовый орех различных толщин (средних и мелких). Его получают с помощью процесса, который может быть разделен на унитарные операции. Наиболее важным из них является этап сушки, осуществляющийся в сушилке Proctor (Proctor Inc, Глазго), состоящей из трех отсеков (А, В и С). Эта сушилка имеет форму параллелепипеда и содержит псевдосжиженный слой с габаритами 20 х 4 х 3 м. Каждый из ее отсеков содержит один или два вентилятора, которые дуют горячим воздухом на конвейер, где сушится измельченный продукт. Отсеки подразделяются на семь отделений (А1, А2, А3, В1, В2, С1 и С2). Таким образом, процесс зависит от многих переменных.

Таблица 1

Экспериментальные данные

	IM (%)	TA1 (°C)	TA2 (°C)	TA3 (°C)	TB1 (°C)	TB2 (°C)	TC1 (°C)	TC2 (°C)	FT (°C)	FM (%)
Min	45.4	115.6	104.7	85.0	81.2	63.8	55.2	67.6	37.0	1.0
Max	53.9	130.1	130.6	118.4	120.7	121.2	100.6	96.3	45.8	4.8
Mean	50.5	124.8	122.8	107.7	103.6	89.2	83.0	82.9	42.0	2.1
SD	1.7	2.58	3.49	5.11	5.71	10.6	8.4	6.81	2.1	0.8
RSD (%)	3.3	2.1	2.8	4.8	5.5	11.9	10.2	8.2	5.1	38.3

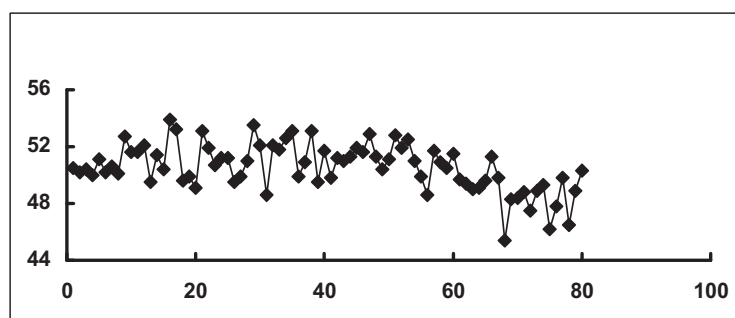


Рис. 1. Значения начальной влажности продукта

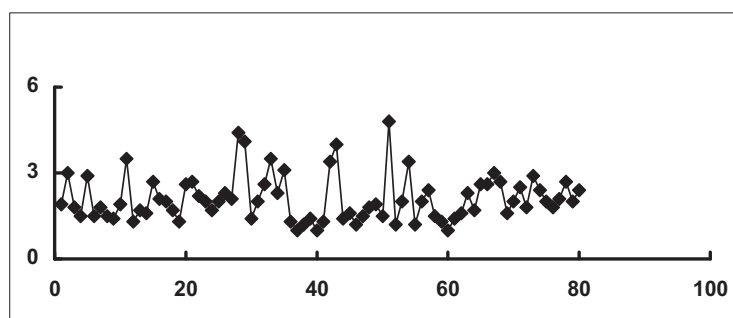


Рис. 2. Значения конечной влажности продукта

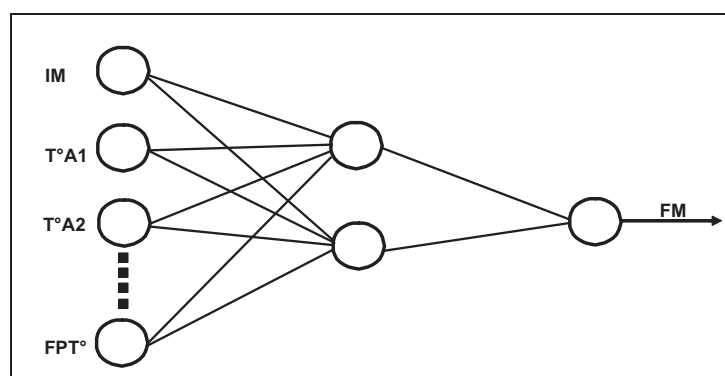


Рис. 3. Многослойный персептрон

Данные были собраны случайным образом. В таблице 1 представлены минимальное, максимальное и среднее значения параметров, стандартное отклонение (SD) и относительное стандартное отклонение (RSD), TA_i – значение температуры в i -м отделении сушильной камеры. Начальное и конечное значения влажности продукта представлены на гистограммах рис.1 и рис.2 соответственно.

Выбор топологии нейронной сети

В настоящее время в открытом доступе представлено множество вариаций нейронных сетей и алгоритмов. Одна из них – многослойный персептрон, который является наиболее широко используемым. Как правило, в персептроне нейроны группируются в три различных типа слоев, как показано на рис. 3: входной слой, где число узлов зависит от количества входных переменных, выходной слой, где число узлов равно количеству предполагаемых переменных скрытого слоя, который, в свою очередь, расположен между первыми двумя слоями.

Выбор нейросетевых моделей осуществляется на основе вычисляемых весов и предубеждений, которые сводят к минимуму расхождение между наблюдаемыми и моделируемыми данными.

Для обучения сети, реализованы следующие шаги:

1. данные разделены на три подмножества: учебные, проверочные и экспериментальные;

2. все входы нормализованы с помощью формулы

$$x_{ni} = \frac{2(x_i - x_{min})}{(x_{max} - x_{min})} - 1 \quad (1)$$

где X_{ni} – нормированные входные данные, лежащие в диапазоне от -1 до 1; X_i – исходные данные; X_{min} и X_{max} – минимальное и максимальное значения из набора данных;

3. веса определены как случайные значения;

4. каждый входной нейрон (i) получал входные сигналы и посылал выходной сигнал на все узлы в следующем слое (скрытом слое);

5. каждый скрытый блок (j) подсчитывал свои входные сигналы и добавлял смещение перед вычислением его выхода следующим образом:

$$y_j = \tanh\left(\sum w_{ij}x_i + b_j\right) \quad (2)$$

Полученный результат был направлен в узел выходного слоя;

6. блок выходного слоя вычислял остаточный член, веса, и поправочные члены смещения;

8. показатели ошибки распространялись в обратном направлении к сигналу, и тем самым осуществлялось корректирование весов нейронов.

Выбор удовлетворительной топологии сети – это довольно непростая задача. Нет жестких правил или теорем для нахождения оптимальной топологии для данного набора входных-выходных данных. Одним из подходов к выбору оптимальной топологии нейронной сети является определение байесовского информационного критерия Шварца (BIC), который получают следующим образом:

$$BIC = \log\left(\frac{V}{n}\right) + p \frac{\log(n)}{n} \quad (3)$$

где V является суммой квадратов ошибок аппроксимации, n – число учебных моделей и p – общее число весов нейронов.

Если N – это число наблюдений и k – число оцениваемых параметров, то степень свободы (DF) выглядит следующим образом:

$$df = N - k \quad (4)$$

В случае многослойного персептрона с одним выходом оцениваемые параметры связаны не только с весами соединения и смещения выходного слоя, но также и соединительными весами, и смещениями, которые соединяют скрытые слои (если их

несколько). Если предполагается, что число этих оценочных показателей определяется в данной задаче, то используется следующее уравнение для архитектуры скрытого слоя:

$$k = n_j(n_i + 2) + 1 \quad (5)$$

где n_j и n_i - число узлов в скрытом и входном слоях соответственно.

Степень свободы должна иметь положительное значение; этот факт накладывает ограничение на размер сети:

$$n_j(\max) = \frac{N - 1}{n_i + 2} \quad (6)$$

Это условие (с использованием степеней свободы) редко используется из-за того, что искусственные нейронные сети изначально разрабатывались для непараметрического моделирования.

В данной работе был выбран алгоритм обратного распространения ошибки, который представляет собой градиент метода спуска, с помощью которого производят регулирование весов нейронов и значений смещения для минимизации суммы квадратов разности между заданным выходом (Y_o) и вычисленным (Y_c), который может быть получен из:

$$e = \frac{1}{2} \sum (y_c - y_o)^2 \quad (7)$$

Обучение подмножества используется для вычисления и обновления весов нейронов. Ошибка при проверке точности нейросетевой модели также контролируется в процессе обучения. Ошибка, как правило, уменьшается на начальном этапе обучения. Но когда ошибка проверки увеличивается в течение определенного количества итераций, обучение останавливается и веса и предубеждения переписываются, присваивая себе минимальные значения ошибки. Этот метод (ранней остановки) позволяет повысить эффективность обучающего процесса.

Результаты моделирования

Эксперименты по моделированию проводились около 1000 раз, чтобы минимизировать ошибку вычислений. Затем фиксировался лучший результат с наименьшей ошибкой. Линия регрессии характеризовалась ее наклоном (m), коэффициентом (b) и коэффициентом корреляции (r). Результаты сравнения представлены в таблице 2.

Анализируя эту таблицу, видно, что, когда количество узлов в скрытых слоях увеличивается, то наклон и значения коэффициентов корреляции значительно

изменяются. Но являются ли эти изменения достаточными, чтобы установить расхождения между этими значениями?

Для того чтобы ответить на этот вопрос, был проведен дополнительный эксперимент, в процессе которого получен коэффициент Фишера (0,72) - ниже чем свое предельное значение (2.46) на уровне 0,05 и, следовательно, видно, что количество нейронов в скрытом слое существенно влияет на поведение нейронной сети.

Таблица 2

Результаты сравнения смоделированных данных ИНС (с архитектурой 9-9-1) с реальными

	Моделирование			Реальные данные		
Узлы в скрытом слое	m	b	r	m	b	r
1	0.489	-0.217	0.774	0.738	-0.026	0.699
2	0.656	-0.138	0.814	0.866	-0.017	0.750
3	0.736	-0.101	0.857	0.764	-0.035	0.695
4	0.745	-0.082	0.865	0.783	0.032	0.783
5	0.815	-0.065	0.909	0.575	-0.184	0.635
6	0.836	-0.070	0.919	0.690	-0.131	0.649
7	0.938	0.002	0.980	0.847	-0.102	0.598
8	0.883	-0.035	0.954	0.643	-0.195	0.530
9	0.993	0.004	1.000	1.154	0.119	0.725
10	0.982	-0.008	0.996	0.291	-0.456	0.228
11	0.971	-0.015	0.990	0.737	-0.086	0.490
12	0.993	-0.005	1.000	0.319	-0.150	0.294
13	1.000	0.000	1.000	0.649	-0.280	0.405
14	1.000	0.000	1.000	0.840	0.031	0.461
15	1.000	0.000	1.000	0.585	-0.094	0.300

Коэффициент корреляции r колебался от 0,774 до 1,000 при подготовке подмножества и от 0,229 до 0,783 при испытании выбранной нейросетевой модели. Чем выше его абсолютная величина (т.е. около 1.000), тем выше точность между

прогнозируемыми значениями и наблюдаемыми. Таким образом, можно заметить, что, когда количество узлов в скрытом слое возрастает, коэффициент корреляции увеличивается (до 1,000). Кроме того, при рассмотрении тестового подмножества, оказалось, что коэффициент корреляции не увеличивался больше в том же направлении, что и количество нейронов в скрытом слое. Его значение для архитектуры с четырьмя узлами - $R = 0,78$ или девятью узлами - $R = 0,723$ в скрытом слое.

Лучшая топология нейронной сети в данном случае представляет собой компромисс между результатами, полученными во время проведения опытных экспериментов, и реальными данными, полученными в процессе сушки тертого кокосового ореха (с наименьшей погрешностью). Это варианты архитектуры: 9-4-1 и 9-9-1. Максимальное приемлемое количество нейронов было рассчитано с помощью уравнения (6). Можно наблюдать (табл. 3), что сеть с четырьмя скрытыми слоями нейронов имеют наименьшую погрешность и байесовский информационный критерий Шварца (BIC).

Таблица 4 ясно показывает, что добавление второго скрытого слоя не повысит производительность сети.

Модель сети 9-4-1 была выбрана на основе анализа ошибок. Рис. 4 показывает, что ошибки равномерно распределяются вокруг среднего значения (0) в диапазоне от -0,35 до 0,34 %. Поскольку ошибки имеют однородную дисперсию и нет наблюдаемой систематической погрешности, модель архитектуры 9-4-1 является достаточной для выполнения предъявленных требований.

Таблица 3

Погрешность и байесовский информационный критерий Шварца (BIC)

Узлы в скрытом слое	Значение ошибки	BIC
1	0.164	-2.473
2	0.139	-2.331
3	0.136	-2.128
4	0.085	-2.752
5	0.101	-1.830
6	0.097	-1.631
7	0.162	-1.194
8	0.157	-1.918
9	0.153	-0.790

10	0.274	-1.023
11	0.221	-0.203
12	0.211	-0.010
13	0.295	0.349
14	0.334	0.619
15	0.454	0.965

Таблица 4

Сравнение параметров ИНС-модели (с архитектурой 9-4-1) с реальными данными

	Моделирование			Реальные данные		
Узлы во втором скрытом слое	m	b	r	m	b	r
1	0.663	-0.124	0.866	0.878	0.077	0.743
2	0.757	-0.092	0.877	0.651	-0.138	0.742
3	0.750	-0.103	0.881	0.365	-0.308	0.428
4	0.863	-0.052	0.934	0.631	-0.072	0.495
5	0.816	-0.075	0.904	0.173	-0.541	0.057
6	0.948	-0.028	0.982	0.192	-0.561	0.091
7	0.913	-0.063	0.955	0.762	-0.205	0.676
8	0.903	-0.047	0.965	0.588	-0.136	0.623
9	0.950	-0.024	0.982	0.506	-0.145	0.339
10	0.996	-0.002	0.999	0.435	-0.194	0.391

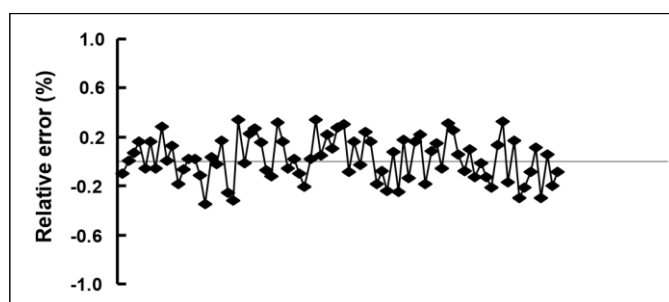


Рис. 4. Точечная диаграмма остатков между наблюдаемыми и расчетными выходами

Заключение

В настоящем исследовании были разработаны нейронные сети, прогнозирующие и корректирующие значение влажности тертого кокоса в процессе промышленной сушки. Метод Левенберга-Марквардта для обучения сети оказался довольно эффективным в данном случае и позволил предсказать процент влажности конечного высушенного продукта. С помощью результатов дисперсионного анализа проведенного теста было установлено, что число скрытых узлов слоя существенно влияет (на уровне 0,05) на производительность нейронной сети. Конфигурация сети с использованием метода обратного распространения ошибки дала самый точный прогноз. Она имеет один скрытый слой, состоящий из четырех нейронов. Нейронная сеть предсказала результаты очень близкие к экспериментальным значениям. СКО достигла значения 0,01 и максимальный процент относительной ошибки лежал в диапазоне от 0,35 до 0,34 %. Таким образом, прогностическая способность нейронных сетей может быть использована в качестве перспективного метода для моделирования, оценки и прогнозирования процесса сушки тертого кокосового ореха.

Список литературы

1. Desheng (Dash), Wu., Zijiang, Y. and Liang L., Using DEA-neural network approach to evaluate branch efficiency of a large canadian bank, Expert Systems with Applications, 31 (1), 108- 115 (2006).
2. Jend, S. and Das, H., Modelling for vacuum drying characteristics of coconut presscake, J. Food Eng., 79 (1), 92-99 (2007).
3. Khamis, A., Ismail, Z., Haron, K. and Tarmizi, A., Neural network model for oil palm yield modelling, J. Applied Sciences, 6 (2), 391-399 (2006).
4. Pramanik, K., Use of artificial neural networks for prediction of cell mass and ethanol concentration in batch fermentation using saccharomyces cerevisiae yeast, IE (I) Journal.CH, 85, 31-35 (2004).