УДК 004.032.26:004.89

ИСПОЛЬЗОВАНИЕ НЕЙРОННЫХ СЕТЕЙ ДЛЯ РАСЧЕТА КАЧЕСТВЕННЫХ ХАРАКТЕРИСТИК НЕФТЕПРОДУКТОВ

^{1,2}Кувыкин В.И., ²Кувыкина Е.В.

¹Институт проблем машиностроения РАН, Нижний Новгород;

²Нижегородский государственный университет имени Н.И. Лобачевского
(Национальный исследовательский университет), Нижний Новгород, e-mail: vkuvykin@yandex.ru

В статье исследованы способы уменьшения погрешности при использовании искусственных нейронных сетей для прогнозирования качественных характеристик нефтепродуктов. Рассмотрен пример вычисления предельной температуры фильтруемости дизельного топлива по известным физико-химическим показателям. При расчете применялись многослойный персептрон с различным числом скрытых слоев и сеть радиально-базисных функций. Математические модели включали одну зависимую переменную и десять предикторов. Исследовано влияние алгоритмов обучения, функций активации, количества нейронов на результаты прогнозирования. Многослойный персептрон, использующий на стадии обучения метод сопряженных градиентов, обеспечил лучшую точность, чем сеть радиально-базисных функций. Определено влияние входных параметров на низкотемпературные свойства дизельного топлива. Численные эксперименты показали, что уменьшение aberrации прогнозируемых значений базируется как на выборе оптимальных параметров модели, так и на статистическом анализе исходных данных. Точность прогноза существенно повышается при исключении ряда примеров, принадлежащих одной из границ обучающего множества. Предложенная модель может обеспечить необходимую для промышленного применения точность. Абсолютная погрешность вычисления предельной температуры фильтруемости соответствует требованиям стандартов измерения. Технологии автоматизированного расчета качественных характеристик нефтепродуктов с применением искусственных нейронных сетей могут успешно использоваться в лабораториях и на нефтеперерабатывающих предприятиях.

Ключевые слова: математическое моделирование, искусственные нейронные сети, многослойный персептрон, нефтепереработка, дизельное топливо, предельная температура фильтруемости

ARTIFICIAL NEURAL NETWORK APPLICATION TO PREDICT QUALITY CHARACTERISTICS OF PETROLEUM PRODUCTS

^{1,2}Kuvykin V.I., ¹Kuvykina E.V.

¹Mechanical Engineering Research Institute of RAS, Nizhny Novgorod;

²Lobachevsky State University (National Research University), Nizhny Novgorod, e-mail: vkuvykin@yandex.ru

This article deals with methods for reducing the error of predicting the qualitative characteristics of oil products using artificial neural networks. It was considered an example of calculating cold filter plugging point of diesel fuel using physicochemical parameters as an input. For calculating were used a multilayer perceptron with a different number of hidden layers and a radial basis function networks. Mathematical models included one dependent variable and ten predictors. The influence of training algorithms, activation functions, the number of neurons on the results of forecasting were investigated. The multilayer perceptron, which uses the conjugate gradient method at the training stage, ensured the better accuracy than a network with a radial basis function. The influence of input parameters on the cold filter plugging point of diesel fuel was determined. Numerical experiments showed that the reduction in the aberration of the predicted values was based both on the selection of the optimal model parameters and on a statistical analysis of the original data. The accuracy of the forecast significantly increases with the exception of a number of examples belonging to one of the boundaries from the training set. This model can provide the precision required for industrial applications. The absolute error for cold filter plugging point was less than the reproducibility of the presently fixed measurement standards. Automated calculation technologies with the neural network models can be successfully applied in laboratories and in oil refineries.

Keywords: mathematical modeling, artificial neural network, multilayer perceptron, refinery, diesel fuel, cold filter plugging point

Разработка производственных систем с искусственным интеллектом в нефтепереработке направлена на решение задач моделирования процессов оперативного управления, расчета показателей качества, управления технологическими процессами, разработки динамических баз данных и знаний [1]. Искусственные нейронные сети (ИНС) как составной компонент Industry 4.0 имеют огромный потенциал оптимизации производственных процессов

и снижения затрат нефтеперерабатывающих заводов (НПЗ). Цифровые технологии позволяют улучшить управление производственным процессом на основе достоверной информации, повышают производительность труда.

Следует отметить, что планирование, составление расписаний и оптимизация смешения продукции производят ощутимый эффект [2–4], который можно существенно увеличить за счет ИНС [5, 6]. При этом про-

гноз качества продукции в нефтепереработке выступает как один из основных элементов практического приложения ИНС [5, 7].

Актуальным вопросом является моделирование низкотемпературных свойств дизельного топлива [5, 7, 8]. Во-первых, объем производства дизельного топлива в России максимальный среди других нефтепродуктов, и его выпуск поддерживается стабильным внешним и внутренним спросом. Во-вторых, климатические условия РФ вызывают большую потребность в высококачественных низкозастывающих дизельных топливах, которая удовлетворяется менее чем наполовину [9]. Эффективность нефтепереработки существенно вырастет, если математическое моделирование даст возможность вычислить низкотемпературные свойства дизельного топлива вместо проведения дорогостоящих физических измерений.

Основными параметрами, характеризующими низкотемпературные свойства, выступают температура помутнения и предельная температура фильтруемости (ПТФ). Если температура помутнения может быть вычислена с применением достаточно простой линейной регрессии, то расчет ПТФ требует построения сложных нелинейных моделей [7]. Вместе с тем ПТФ – это весьма важный показатель для потребителей, поскольку определяет минимальную температуру, при которой будет обеспечен бесперебойный поток в топливных системах [10]. В работах [7, 8] для некоторого заданного набора лабораторных данных отмечена приемлемая точность прогнозирования ПТФ с помощью ИНС. Тем не менее для ряда точек абсолютная погрешность остается высокой, и для практического применения ИНС требуются исследования, направленные на повышение точности прогнозирования.

Наиболее широкое распространение при анализе качества получили сети прямого распространения, такие как многослойный персептрон (МП) и сеть радиально-базисных функций (РБФ) [7, 8]. Они позволяют сравнить предсказанные моделью результаты с известными значениями зависимых переменных [11]. Нейросетевые модели не требуют формулировки гипотез о взаимосвязи между зависимыми и независимыми переменными, расчет происходит в процессе обучения. Хотя результаты моделирования ИНС представляют систему «черный ящик» и с трудом поддаются интерпретации, оценка важности значений предикторов позволяет анализировать зависимость выходной функции от входных переменных.

Целями работы являются разработка способов уменьшения погрешности прогнозирования качественных характеристик нефтепродуктов с помощью искусственных нейронных сетей на примере предельной температуры фильтруемости дизельного топлива; определение степени влияния входных параметров на итоговый прогноз.

Основные задачи могут быть сформулированы следующим образом: построение прогноза для определения ПТФ по другим известным физико-химическим параметрам; исследование влияния алгоритмов обучения, функций активации сетей и числа нейронов на точность прогнозирования; статистический анализ данных для выявления примеров, оказывающих отрицательное влияние на точность прогнозирования; оценка влияния предикторов на низкотемпературные свойства дизельного топлива.

Материалы и методы исследования

Для обучения ИНС использовались паспортные данные зимнего дизельного топлива, предусмотренные стандартами на готовую продукцию [12]. Входной слой включал 10 нейронов с показателями: цетановое число; цетановый индекс; плотность при 15 °C (кг/м³); массовая доля полициклических ароматических углеводородов (%); массовая доля серы (мг/кг); температура вспышки в закрытом тигле (°C); кинематическая вязкость при 40 °C (мм²/с); объем перегонки до 180 °C и до 360 °C (% об.); температура выкипания 95 % объема (°C). Поскольку изменение остальных характеристик было несущественным, то и в построении нейронной сети они не участвовали. На выходе имелся только один нейрон для расчета ПТФ (°C).

Вычисления и статистическая обработка информации осуществлялись с помощью компьютерной программы SPSS Statistics. В процессе построения модели множество данных делилось случайным образом на обучающее, тестирующее и контрольное. Контрольное множество ни в обучении, ни в оптимизации и тестировании сети не участвовало. Результаты сравнения для контрольного множества являлись основным фактором при сравнении архитектуры сетей, функций активации и выбора методов решения.

Расчет ПТФ проведен для МП с различным числом скрытых слоев и сетей РБФ. Созданы нейросетевые модели для одной зависимой переменной на основании значений переменных предикторов. Многослойный персептрон включает входной слой, ряд скрытых слоев и выходной уровень.

Для обучения сети варьировались алгоритмы, архитектура сети и число нейронов.

В отличие от многослойного персептрона, сеть РБФ включает только 3 уровня: входной, один скрытый и выходной. Функциями активации в скрытом слое являются радиальные базисные функции. Широкое применение нейронных сетей РБФ обусловлено простой структурой сети и высокой скоростью обучения.

Обозначим через Y множество значений ПТФ, G_+ , G_- – граничные точки данного множества. При оценке результатов работы ИНС вычислялась средняя абсолютная погрешность MAE (Mean Absolute Error) по формуле:

$$MAE_k = \frac{1}{N} \sum_{n=1}^N |y_{nk} - \hat{y}_{nk}|, \quad (1)$$

где $\hat{y}_{nk} \in Y$ – измеренное, y_{nk} – прогнозируемое значение ПТФ для k -ой группы n -го значения, N – количество примеров. Следующим важным показателем, характеризующим работу сети, является средняя абсолютная процентная погрешность MAPE (Mean absolute percentage error)

$$MAPE_k = \frac{1}{N} \sum_{n=1}^N \left| \frac{y_{nk} - \hat{y}_{nk}}{y_{nk}} \right| \times 100. \quad (2)$$

Были проведены расчеты с разным числом нейронов. Определено оптимальное количество нейронов, которое обеспечивает наименьшую относительную погрешность для контрольных данных и наименьшие значения показателей (1) и (2).

В основе предлагаемой методологии лежит статистический анализ влияния исходного обучающего множества на точность прогноза. Известно, что в исходной информации могут быть данные, которые выпадают из общей закономерности. Их исключение приводит к уменьшению погрешности обучения сети и улучшению прогноза [11]. Проведено последовательное исключение примеров из обучающего множества и вычислена погрешность ИНС, обученной на таких усеченных множествах.

Результаты исследования и их обсуждение

На первом этапе исследования рассмотрено 188 примеров в диапазоне изменения ПТФ от -34°C до -28°C (Ряд 1 на рис. 1). Расчеты показали, что уменьшение абберации прогнозируемых значений ПТФ происходит при исключении ряда примеров из обучающего множества. Для дальнейшего исследования были удалены примеры со значениями, принадлежащими нижней

границе интервала $1,5 \cdot \text{IQR}$ (IQR – межквартильный размах). Они вносят наибольшие погрешности, и включать их для обучения ИНС нецелесообразно. С целью уменьшения погрешности прогнозирования ПТФ выбран измененный массив данных $\Psi = Y \setminus G_-$, содержащий 180 точек (Ряд 2 на рис. 1). Отметим, что МП позволяет идентифицировать примеры, которые выпадают из закономерности, характерной для изучаемого массива данных.

Рис. 1. Диаграмма ящик с усами исходного массива данных предельной температуры фильтруемости (ПТФ), обозначенных Ряд 1, и с исключенными примерами – Ряд 2

Обучающее множество для МП включало 102 (56,7%) примера, тестовое – 23 (12,7%) и контрольное – 55 (30,6%). Расчет ПТФ с максимальной точностью обеспечивает нейронная сеть с 2 скрытыми слоями, 7 нейронов было задействовано на первом скрытом слое и 3 – на втором. Функция активации – гиперболический тангенс. При обучении применялся метод масштабируемых сопряженных градиентов. Для сети РБФ обучающее множество содержало 104 (57,8%) примера, тестовое – 35 (19,4%) и контрольное – 41 (22,8%). Лучшие результаты для контрольного множества обеспечила нормализованная радиальная базисная функция с 9 нейронами скрытого слоя.

Как МП, так и сети РБФ дают погрешности, которые находятся в пределах, предусмотренных национальным стандартом [10]. Расчет показателей (1) и (2) представлен в табл. 1, в которой K указывает номер модели, N – количество

примеров, показатели MAE и $MAPE$ вычислены по формулам (1), (2). Как следует из приведенной таблицы, МП с показателями $MAE_3 = 0,4^\circ\text{C}$ и $MAPE_3 = 1,3\%$ дает существенно лучшие результаты, чем сеть РБФ (табл. 1). Этот вывод совпадает с результатами, полученными в работе [8] для других входных данных и сети с другим числом нейронов. Многослойный персептрон для примеров, исключающих нижнюю границу интервала, отделяющего выбросы, имеет показатели в 2,5 раза меньшие (табл. 1). Таким образом, исключение 4% точек из обучающего множества существенно улучшает прогноз.

Таблица 1

Сравнение статистических показателей для МП и сети РБФ

K	Нейросеть	N	$MAE, ^\circ\text{C}$	$MAPE, \%$
1	МП	188	1,0	3,2
2	РБФ	180	1,0	3,6
3	МП	180	0,4	1,3

Для МП (табл. 1, $K = 3$) получены наилучшие результаты и обеспечено вычисление ПТФ с абсолютной погрешностью, соответствующей стандартам измерения [10]. На рис. 2 приведены исходные значения ПТФ и прогнозируемые значения, рассчитанные ИНС. Для приведенного в работе диапазона температур расхождение менее $2,0^\circ\text{C}$, что соответствует стандарту по определению ПТФ [10].

Для МП осуществлен анализ чувствительности, при котором производилось вычисление важности каждого входного параметра ИНС на основе обучающих и проверочных примеров. В табл. 2 приведены данные, полученные с помощью МП, показывающие важность и нормализованную важность для каждого предиктора при расчете ПТФ.

Наиболее существенное влияние, как следует из табл. 2, оказали показатели объема перегонки до 360°C и температуры выкипания 95% объема. Информация представляет интерес при анализе производственных процессов и может служить основой для построения регрессионных моделей.

Заключение

Модель с применением ИНС способна обеспечить необходимую для промышленного применения точность, поскольку абсолютная погрешность вычисления ПТФ соответствует требованиям стандартов измерения.

Численные эксперименты показали, что уменьшение аберрации прогнозируемых значений базируется как на выборе оптимальных параметров модели, так и на статистическом анализе исходных данных. В задачах определения низкотемпературных свойств дизельных топлив МП дает лучшие результаты, чем сеть РБФ, и рекомендуется для применения. Исключение ряда примеров, принадлежащих одной из границ обучающего множества, существенно повышает точность прогноза.

Рис. 2. Предсказанные и исходные значения предельной температуры фильтруемости зимнего дизельного топлива

Таблица 2

Важность независимых переменных

Показатели качества дизельного топлива	Важность	Нормализованная важность
Цетановое число	0,102	78,5%
Цетановый индекс	0,114	87,4%
Плотность при 15°C, кг/м ³	0,072	54,8%
Массовая доля полициклических ароматических углеводородов (%)	0,098	74,9%
Массовая доля серы, мг/кг	0,113	86,2%
Темп вспышки	0,065	49,8%
Кинематическая вязкость при 40°C, мм ² /с	0,055	42,4%
Объем перегонки до 180°C	0,124	95,0%
Объем перегонки до 360°C	0,131	100,0%
Температура выкипания 95% объема	0,127	97,1%

Разработанные модели могут применяться в лабораториях и на заправочных станциях для быстрого вычисления и контроля предельной температуры фильтруемости дизельного топлива и на нефтеперерабатывающих заводах для непрерывного контроля качества, в системах календарного планирования и управления смешением топлив.

Список литературы

1. Технология, экономика и автоматизация процессов переработки нефти и газа / Под ред. С.А. Ахметова. М.: Химия, 2005. 736 с.
2. Кувыкин В.И. Использование моделей бизнес-процессов НПЗ в системах планирования и учёта // Мир нефтепродуктов. Вестник нефтяных компаний. 2013. № 7. С. 47–48.
3. Кувыкин В.И., Кувыкина Е.В., Петухов М.Ю. Анализ оптимальных решений в задачах нелинейного программирования // Вестник Нижегородского университета им. Н.И. Лобачевского. 2011. № 4–5. С. 2285–2286.
4. Кувыкин В.И., Мелешкевич М.А., Наумова С.В. Системный подход к оптимизации управления смешением // Международный научно-исследовательский журнал. 2016. № 10–2 (52). С. 133–136. DOI: 10.18454/IRJ.2016.52.075.

5. Weimin Li, Wenkai Li, Hui C.-W. Integrating neural network models for refinery planning. Computer Aided Chemical Engineering. 2003. V. 15. No. 12. P. 1304–1309. DOI: 10.1016/S1570-7946(03)80490-X.

6. Gao X., Shang C., Jiang Y., Huang D., Chen T. Refinery scheduling with varying crude: A deep belief network classification and multimodal approach. Aiche Journal. 2014. V. 60 (7). P. 2525–2532. DOI: 10.1002/aic.14455.

7. Wu C., Zhang J., Li W., Wang Y., Cao H. Artificial neural network model to predict cold filter plugging point of blended diesel fuels. Fuel Process. Technol. 2006. No. 87. P. 585–590.

8. Маринович С., Боланча Т., Юкич С., Рукавина В., Юкич А. Вычисление низкотемпературных характеристик дизельного топлива с помощью искусственных нейронных сетей // Химия и технология топлив и масел. 2012. № 1. С. 47–52.

9. Камешков А.В., Гайле А.А. Получение дизельных топлив с улучшенными низкотемпературными свойствами // Изв. СПбГТИ(ТУ). 2015. № 29. С. 49–60.

10. ГОСТ 33755-2016. Топливо дизельное и мазут топочный. Определение предельной температуры фильтруемости на холодном фильтре. М.: Стандартинформ, 2016. 16 с.

11. Ясницкий Л.Н. Интеллектуальные системы. М.: Лаборатория знаний, 2016. 221 с.

12. ГОСТ 35511-2013. Топливо дизельное ЕВРО. Технические условия. М.: Стандартинформ, 2014. 16 с.