

ИСПОЛЬЗОВАНИЕ НЕЙРОННЫХ СЕТЕЙ И НЕЧЕТКОЙ ЛОГИКИ ДЛЯ ПРОГНОЗИРОВАНИЯ ФИЗИКО-ХИМИЧЕСКИХ СВОЙСТВ МАТЕРИАЛОВ

Н.И.Попок, М.В.Пята

Использование методов искусственного интеллекта позволяет получать точные результаты прогнозирования тех или иных физико-химических характеристик веществ.

Одной из важных задач при прогнозировании физико-химических параметров веществ является выбор адекватной модели, позволяющей строить нелинейные модели, превосходящие по точности прогноза традиционные регрессионные модели. В ходе проведенного исследования было отмечено, что значение прогнозируемых физико-химических параметров веществ зависит от структурных особенностей молекулы вещества и тех атомов, которые входят в данное вещество. Наиболее полно структура вещества описывается тремя топологическими индексами: индексом Виннера, индексом Балабана и индексом Рандича. Для описания тех атомов, которые содержатся в молекуле достаточно - молекулярной массы и кислородного баланса.

Выбор данных топологических индексов обусловлен простотой их расчета и хорошей описательной способностью. Данный метод позволяет описывать широкий круг молекулярных свойств, устанавливая корреляционные зависимости между свойствами и величинами топологических индексов.

Индексом Винера является число путей $W(G)$ для графа G , и он определяется как число связей между парами углеродных атомов в молекулярном графе молекулы насыщенного углеводорода.

$$W(G) = \frac{1}{2} \sum d_{ij},$$

где d_{ij} расстояние между i -ым и j -ым узлом графа.

Связность по сумме усредненных расстояний $J(G)$ называется индексом Балабана, и определяется как:

$$J(G) = \frac{2}{\mu+1} \sum_{i,j} (V_{D_i} V_{D_j})^{-1/2},$$

где μ - цикломатическое число графа G (число циклов в молекуле), V_{D_i}, V_{D_j} - суммы рас-

стояний; суммирование проводится по всем смежным вершинам.

Индекс Рандича вычисляется по формуле:

$$\chi = \sum \frac{1}{\sqrt{v_i v_j}},$$

где суммирование проводится по связям (ребрам молекулярного графа), а индексы i и j относятся к номерам вершинам молекулярного графа, соединенным данными ребрами.

Кислородный баланс это отношение количества кислорода, содержащегося во взрывчатом веществе (ВВ) к его количеству, необходимому для полного окисления всех остальных компонентов этого ВВ. Кислородный баланс зависит в основном от состава ВВ, а также от состояния заряда и условий взрывания. Он вычисляется по формуле:

$$KB = \frac{\left[d - \left(2a + \frac{b}{2} \right) \right] * 16}{12a + b + 14c + 16d} * 100\%,$$

где a - число атомов углерода, b - число атомов водорода, c - число атомов азота, d - число атомов кислорода.

Качественным прогнозом физико-химического параметра в исследовании будет считаться тот прогноз, у которого количество спрогнозированных значений, вышедших за пределы доверительного интервала, не превышает 10%, от общего числа элементов обучающей выборки.

ИСКУССТВЕННЫЙ НЕЙРОН

Искусственный нейрон имитирует в первом приближении свойства биологического нейрона. На вход искусственного нейрона поступает некоторое множество сигналов, каждый из которых является выходом другого нейрона. Каждый вход умножается на соответствующий вес, аналогичный синаптической силе, и все произведения суммируются, определяя уровень активации нейрона. На рисунке 1 представлена модель, реализующая эту идею. Хотя сетевые парадигмы весьма разнообразны, в основе почти всех их лежит эта конфигурация. Здесь множество

входных сигналов, обозначенных x_1, x_2, \dots, x_n , поступает на искусственный нейрон. Эти входные сигналы, в совокупности обозначаемые вектором \mathbf{X} , соответствуют сигналам, приходящим в синапсы биологического нейрона. Каждый сигнал умножается на соответствующий вес w_1, w_2, \dots, w_n , и поступает на суммирующий блок, обозначенный Σ . Каждый вес соответствует «силе» одной биологической синаптической связи. (Множество весов в совокупности обозначается вектором \mathbf{W} .) Суммирующий блок, соответствующий телу биологического элемента, складывает взвешенные входы алгебраически, создавая выход, который мы будем называть NET. В векторных обозначениях это может быть компактно записано следующим образом:

$$NET = \mathbf{XW},$$

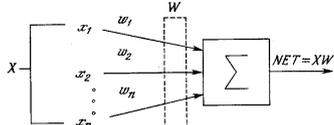


Рис. 1. Искусственный нейрон

АКТИВАЦИОННЫЕ ФУНКЦИИ

Сигнал NET далее, как правило, преобразуется активационной функцией F и дает выходной нейронный сигнал OUT. Активационная функция может быть обычной линейной функцией

$$OUT = K(NET),$$

где K – постоянная, пороговой функции

$$OUT = 1 \text{ если } NET > T$$

$$OUT = 0 \text{ в остальных случаях,}$$

где T – некоторая постоянная пороговая величина, или же функцией, более точно моделирующей нелинейную передаточную характеристику биологического нейрона и представляющей нейронной сети большие возможности.

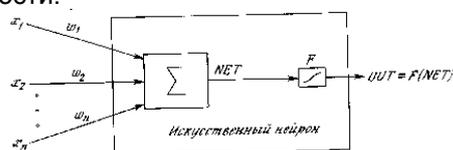


Рис. 2. Искусственный нейрон с активационной функцией

На рисунке 2 блок, обозначенный F , принимает сигнал NET и выдает сигнал OUT. Если блок F сужает диапазон изменения величины NET так, что при любых значениях NET значения OUT принадлежат некоторому конечному интервалу, то F называется «сжи-

мающей» функцией. В качестве «сжимающей» функции часто используется логистическая или «сигмоидальная» (S-образная) функция, показанная на рисунке 3 Эта функция математически выражается как:

$$F(x) = \frac{1}{(1 + e^{-x})}.$$

Таким образом,

$$OUT = \frac{1}{1 + e^{-NET}}.$$

По аналогии с электронными системами активационную функцию можно считать нелинейной усилительной характеристикой искусственного нейрона. Коэффициент усиления вычисляется как отношение приращения величины OUT к вызвавшему его небольшому приращению величины NET. Он выражается наклоном кривой при определенном уровне возбуждения и изменяется от малых значений при больших отрицательных возбуждениях (кривая почти горизонтальна) до максимального значения при нулевом возбуждении и снова уменьшается, когда возбуждение становится большим положительным. Гроссберг (1973) обнаружил, что подобная нелинейная характеристика решает поставленную им дилемму шумового насыщения. Каким образом одна и та же сеть может обрабатывать как слабые, так и сильные сигналы? Слабые сигналы нуждаются в большом сетевом усилении, чтобы дать пригодный к использованию выходной сигнал. Однако усилительные каскады с большими коэффициентами усиления могут привести к насыщению выхода шумами усилителей (случайными флуктуациями), которые присутствуют в любой физически реализованной сети. Сильные входные сигналы в свою очередь также будут приводить к насыщению усилительных каскадов, исключая возможность полезного использования выхода. Центральная область логистической функции, имеющая большой коэффициент усиления, решает проблему обработки слабых сигналов, в то время как области с падающим усилением на положительном и отрицательном концах подходят для больших возбуждений. Таким образом, нейрон функционирует с большим усилением в широком диапазоне уровня входного сигнала.

$$OUT = \frac{1}{1 + e^{-NET}} = F(NET).$$

ИСПОЛЬЗОВАНИЕ НЕЙРОННЫХ СЕТЕЙ И НЕЧЕТКОЙ ЛОГИКИ ДЛЯ ПРОГНОЗИРОВАНИЯ ФИЗИКО-ХИМИЧЕСКИХ СВОЙСТВ МАТЕРИАЛОВ

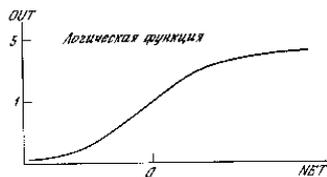


Рис. 3. Сигмоидальная логистическая функция

МНОГОСЛОЙНЫЕ ИСКУССТВЕННЫЕ НЕЙРОННЫЕ СЕТИ

Более крупные и сложные нейронные сети обладают, как правило, и большими вычислительными возможностями. Хотя созданы сети всех конфигураций, какие только можно себе представить, послойная организация нейронов копирует слоистые структуры определенных отделов мозга. Оказалось, что такие многослойные сети обладают большими возможностями, чем однослойные, и в последние годы были разработаны алгоритмы для их обучения.

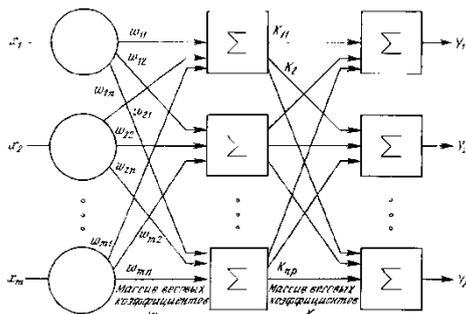


Рис. 4. Двухслойная нейронная сеть

Многослойные сети могут образовываться каскадами слоев. Выход одного слоя является входом для последующего слоя. Подобная сеть показана на рисунке 4 и снова изображена со всеми соединениями.

НЕЛИНЕЙНАЯ АКТИВАЦИОННАЯ ФУНКЦИЯ

Многослойные сети не могут привести к увеличению вычислительной мощности по сравнению с однослойной сетью лишь в том случае, если активационная функция между слоями будет нелинейной. Вычисление выхода слоя заключается в умножении входного вектора на первую весовую матрицу с последующим умножением (если отсутствует нелинейная активационная функция) результирующего вектора на вторую весовую матрицу.

$$(XW_1)W_2$$

Так как умножение матриц ассоциативно, то:

$$X(W_1W_2).$$

Это показывает, что двухслойная линейная сеть эквивалентна одному слою с весовой матрицей, равной произведению двух весовых матриц. Следовательно, любая многослойная линейная сеть может быть заменена эквивалентной однослойной сетью. Показано, что однослойные сети весьма ограничены по своим вычислительным возможностям. Таким образом, для расширения возможностей сетей по сравнению с однослойной сетью необходима нелинейная активационная функция.

СЕТИ С ОБРАТНЫМИ СВЯЗЯМИ

У сетей, рассмотренных до сих пор, не было обратных связей, т. е. соединений, идущих от выходов некоторого слоя к входам этого же слоя или предшествующих слоев. Этот специальный класс сетей, называемых сетями без обратных связей или сетями прямого распространения, представляет интерес и широко используется. Сети более общего вида, имеющие соединения от выходов к входам, называются сетями с обратными связями. У сетей без обратных связей нет памяти, их выход полностью определяется текущими входами и значениями весов. В некоторых конфигурациях сетей с обратными связями предыдущие значения выходов возвращаются на входы; выход, следовательно, определяется как текущим входом, так и предыдущими выходами. По этой причине сети с обратными связями могут обладать свойствами, сходными с кратковременной человеческой памятью, сетевые выходы частично зависят от предыдущих входов.

АЛГОРИТМЫ ОБУЧЕНИЯ

Большинство современных алгоритмов обучения выросло из концепций Хэбба. Им предложена модель обучения без учителя, в которой синаптическая сила (вес) возрастает, если активированы оба нейрона, источник и приемник. Таким образом, часто используемые пути в сети усиливаются и феномен привычки и обучения через повторение получает объяснение.

В искусственной нейронной сети, использующей обучение по Хэббу, наращивание весов определяется произведением уровней возбуждения передающего и принимающего нейронов. Это можно записать как

$$w_{ij}(n+1) = w(n) + \alpha OUT_i * OUT_j,$$

где $w_{ij}(n)$ – значение веса от нейрона i к нейрону j до подстройки, $w_{ij}(n+1)$ – значение веса от нейрона i к нейрону j после подстройки, α – коэффициент скорости обучения, OUT_i – выход нейрона i и вход нейрона j , OUT_j – выход нейрона j .

Сети, использующие обучение по Хэббу, конструктивно развивались, однако за последние 20 лет были развиты более эффективные алгоритмы обучения.

В настоящее время используется огромное разнообразие обучающих алгоритмов. В приложении Б[3] представлен общий обзор, в определенной мере более обширный, хотя и не очень глубокий. В нем дан исторический контекст алгоритмов обучения, их общая таксономия, ряд преимуществ и ограничений. В силу необходимости это приведет к повторению части материала, оправданием ему служит расширение взгляда на предмет.

НЕЧЕТКАЯ ЛОГИКА

Характеристикой нечеткого множества выступает функция принадлежности (Membership Function). Обозначим через $MF_c(x)$ – степень принадлежности к нечеткому множеству C , представляющей собой обобщение понятия характеристической функции обычного множества. Тогда нечетким множеством C называется множество упорядоченных пар функций $MF_c(x) \in [0,1]$. Значение $MF_c(x)=0$ означает отсутствие принадлежности к множеству, 1 – полную принадлежность[4].

Для описания нечетких множеств вводятся понятия нечеткой и лингвистической переменных.

Нечеткая переменная описывается набором (N, X, A) , где N – это название переменной, X – универсальное множество (область рассуждений), A – нечеткое множество на X . Значениями лингвистической переменной могут быть нечеткие переменные, т.е. лингвистическая переменная находится на более высоком уровне, чем нечеткая переменная. Каждая лингвистическая переменная состоит из:

- названия;
- множества своих значений, которое также называется базовым терм-множеством. Элементы базового терм-множества представляют собой названия нечетких переменных;
- универсального множества X ;

- синтаксического правила G , по которому генерируются новые термы с применением слов естественного или формального языка;

- семантического правила P , которое каждому значению лингвистической переменной ставит в соответствие нечеткое подмножество множества X .

В данной нечетко-логической модели будет использоваться функция принадлежности гауссова типа описываемая формулой:

$$MF(x) = e^{-\frac{(x-c)^2}{\sigma}},$$

и оперирующая двумя параметрами. Параметр c обозначает центр нечеткого множества, а параметр σ отвечает за крутизну функции.

НЕЧЕТКИЙ ЛОГИЧЕСКИЙ ВЫВОД

Основой для проведения операции нечеткого логического вывода является база правил, содержащая нечеткие высказывания в форме 'Если-то' и функции принадлежности для соответствующих лингвистических термов. При этом должны соблюдаться следующие условия:

1. Существует хотя бы одно правило для каждого лингвистического термина выходной переменной.
2. Для любого термина входной переменной имеется хотя бы одно правило, в котором этот терм используется в качестве предпосылки (левая часть правила).

В противном случае имеет место неполная база нечетких правил.

Результатом нечеткого вывода является четкое значение переменной y^* на основе заданных четких значений x_k , $k=1..n$.

В общем случае механизм логического вывода включает четыре этапа: введение нечеткости (фазификация), нечеткий вывод, композиция и приведение к четкости, или дефазификация.

Алгоритмы нечеткого вывода различаются главным образом видом используемых правил, логических операций и разновидностью метода дефазификации. Разработаны модели нечеткого вывода Мамдани, Сугено, Ларсена, Цукамото.

Рассмотрим подробнее нечеткий вывод на примере механизма Мамдани (Mamdani). Это наиболее распространенный способ логического вывода в нечетких системах. В нем используется минимаксная композиция нечетких множеств. Данный механизм включает

ИСПОЛЬЗОВАНИЕ НЕЙРОННЫХ СЕТЕЙ И НЕЧЕТКОЙ ЛОГИКИ ДЛЯ ПРОГНОЗИРОВАНИЯ ФИЗИКО-ХИМИЧЕСКИХ СВОЙСТВ МАТЕРИАЛОВ

в себя следующую последовательность действий.

1. Процедура фазификации: определяются степени истинности, т.е. значения функций принадлежности для левых частей каждого правила (предпосылок). Для базы правил с m правилами обозначим степени истинности как $A_{ik}(x_k)$, $i=1..m$, $k=1..n$.

2. Нечеткий вывод. Сначала определяются уровни 'отсечения' для левой части каждого из правил:

$$\alpha_i = \min(A_{ik}(x_k)).$$

Далее находятся 'усеченные' функции принадлежности:

$$B_i^* = \min(\alpha_i, B_i(y)).$$

3. Композиция, или объединение полученных усеченных функций, для чего используется максимальная композиция нечетких множеств:

$$MF(y) = \max(B_i^*(y)).$$

где $MF(y)$ – функция принадлежности итогового нечеткого множества.

4. Дефазификация, или приведение к четкости. Существует несколько методов дефазификации. Например, метод среднего центра, или центроидный метод:

$$MF(y) = \max(B_i^*(y)),$$

Геометрический смысл такого значения – центр тяжести для кривой $MF(y)$.

ЭКСПЕРИМЕНТАЛЬНАЯ ЧАСТЬ

Для построения нейронных сетей использовался пакет Deductor 4(Demo). Данный пакет позволяет создавать и обучать нейронные сети, он прост в использовании и позволяет работать со многими популярными форматами.

Для прогноза таких физико-химических характеристик взрывчатых веществ как: энтальпия образования и скорость детонации, использовались нейронные сети, так как в описание данных параметров отсутствует сложность и неопределенность.

В результате численного моделирования были получены три нейронных сети. Оценка результатов прогнозирования осуществлялась следующим образом: для набора истинных значений прогнозируемого параметра рассчитывался доверительный интервал, и определялось какое количество спрогнозированных значений выходило за границы данного доверительного интервала [1]. Полученные результаты удовлетворяют критерию качества прогноза.

Для прогнозирования более сложных физико-химических параметров взрывчатых веществ использовалась нечеткая логика как показано в статье [2], хороший результат прогнозирования чувствительности к удару достигается при указании к какой вершине основания присоединяется та или иная функциональная группа. Данный метод описания вещества весьма сложен при описании разветвленных структур имеющих несколько циклов основания.

Для того чтобы в каждой нечетко-логической модели присутствовали только заведомо близкие между собой вещества, было решено провести предварительную кластеризацию на основании метода нечетких с-средних. Кластеризация проводилась по чувствительности к удару, а для того чтобы в дальнейшем можно было отнести вещество к тому или иному классу было построено дерево решений. Дерево решений построено при помощи трех дескрипторов которыми описывается молекула: молярная масса, индекс Рандича и кислородный баланс.

В результате проведенной кластеризации были получены три класса веществ. Первый кластер содержит такие вещества как: 1,3,5-триамино-2,4,6-тринитробензол, 1-динитрометил-3-нитробензол, 1,3,5-триамино-2,4,6-тринитробензол, 1,3-диамино-2,4,6-тринитробензол и 1,3-диамино-2,4,6-тринитробензол. Второй кластер имеет следующих характерных представителей: гексанитростильбен, гексанитродифенил, бензотрифуроксан, 5-нитро-1-пикрил-4-пикриламинопипразол, и 2,3,4,6-тетранитроанилин. В третьем кластере имеются следующие представители: гексанитробензол, додеканитроокватерфенил, 1,2,4,5-тетранитробензол, 2',2',2'-тринитроэтил-3,5-динитросолицилат и 2',2',2'-тринитроэтил-2,4,6-тринитробензоат.

Для построения деревьев решений так же использовался пакет Deductor 4(Demo).

Нечетко-логическая модели Мамдани были построены в пакете MatLab 7.0. Данный пакет позволяет наглядно строить нечетко-логические модели и настраивать их.

В результате проведенных численных экспериментов были получены три нечетко-логические модели позволяющие прогнозировать чувствительность к удару. Полученные результаты удовлетворяют условиям качества прогноза, что можно увидеть на рисунках 7-9.

Выводы

Полученные результаты приближены к линейной зависимости действительного и спрогнозированного значений:

$$X_{\text{прогноз}} = X_{\text{действительное}} + \delta.$$

Все результаты укладываются в соответствующий доверительный интервал.

В совокупности это позволяет говорить о высокой точности методов нейронных сетей и нечеткой логики для прогноза свойств взрывчатых материалов на основе описания топологических и молекулярных дискрипторов.

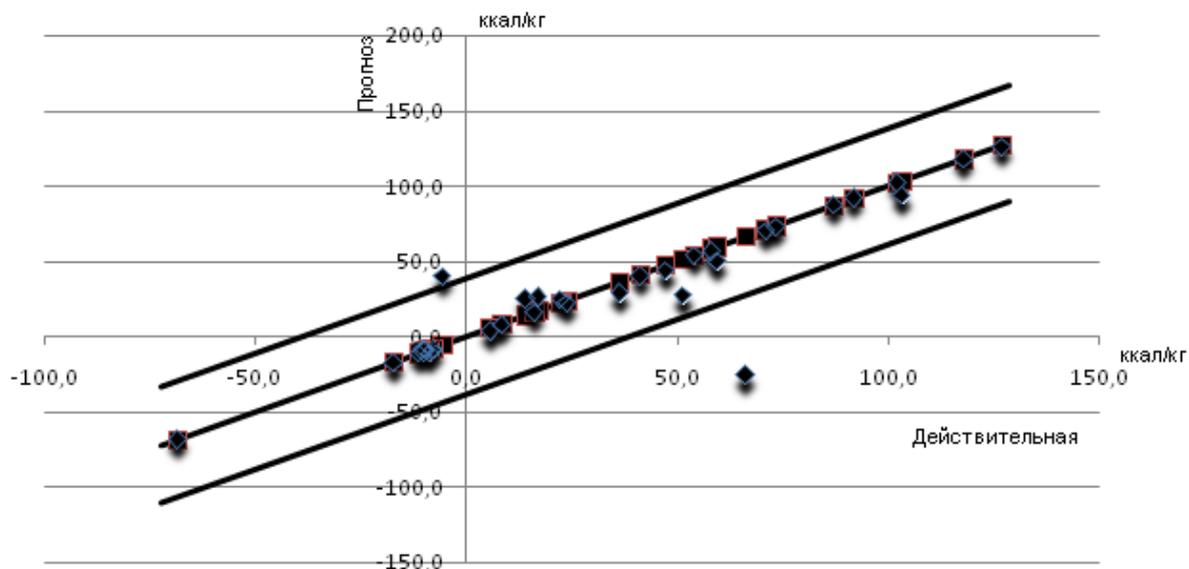


Рис. 6. Результаты прогноза энтальпии образования

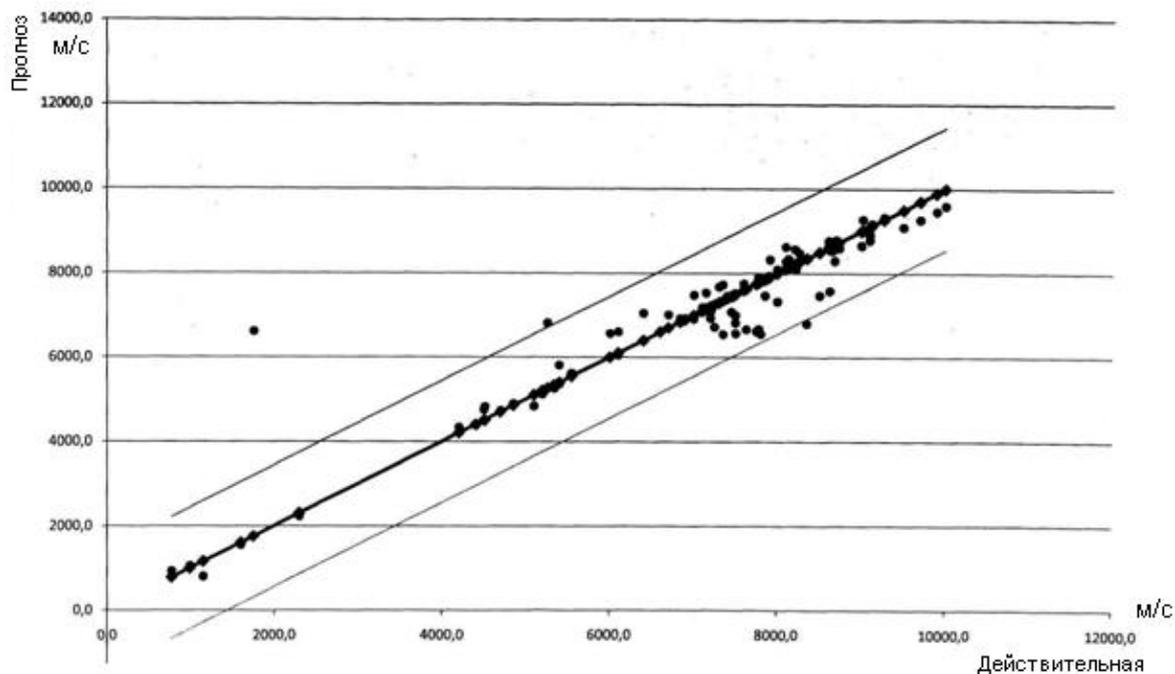


Рис. 5. Результаты прогноза скорости детонации

ИСПОЛЬЗОВАНИЕ НЕЙРОННЫХ СЕТЕЙ И НЕЧЕТКОЙ ЛОГИКИ ДЛЯ ПРОГНОЗИРОВАНИЯ ФИЗИКО-ХИМИЧЕСКИХ СВОЙСТВ МАТЕРИАЛОВ

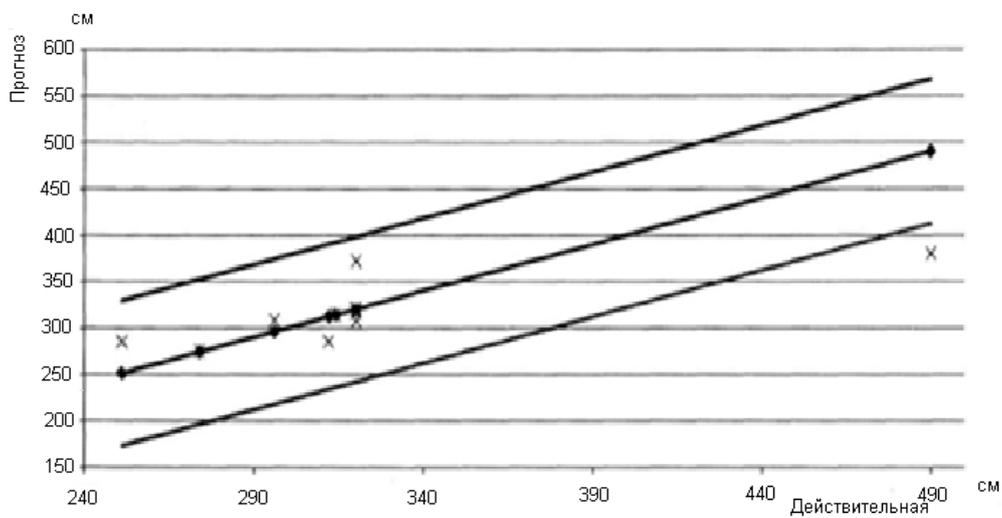


Рис. 7. Результаты прогноза чувствительности к удару для первого кластера

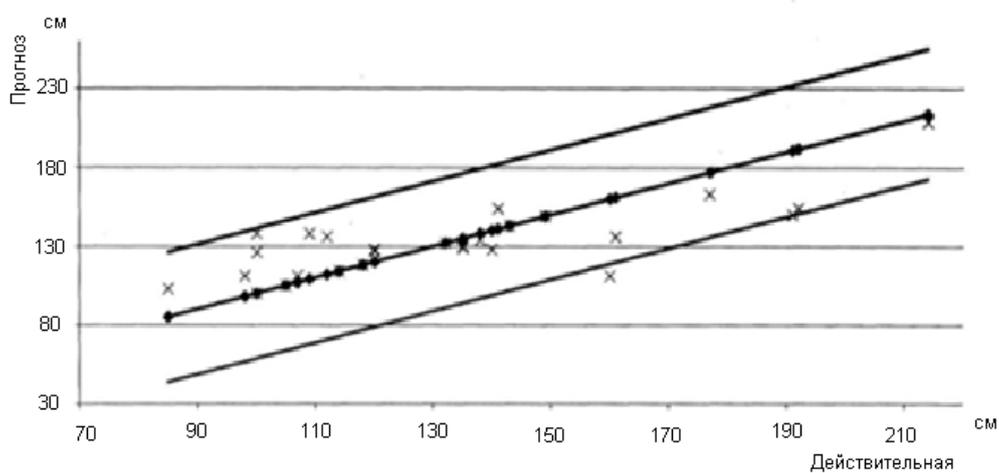


Рис. 8. Результаты прогноза чувствительности к удару для второго кластера

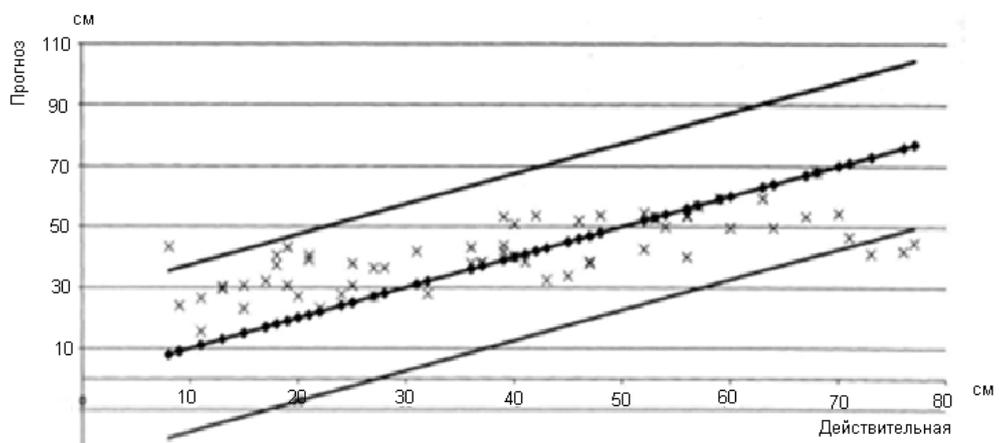


Рис. 9. Результаты прогноза чувствительности к удару для третьего кластера

ЛИТЕРАТУРА

1. Крылова Г.Д. Основы стандартизации, сертификации, метрологии: Учебник. – М.: Изд-во Юнити-Дана, 2000. – 712 с.

2. Пята М.В. Использование методов искусственного интеллекта для прогнозирования взрывчатых свойств производных бензола // Сб. трудов НПК «Исследование, разработка и применение высоких технологий в промышленности» под ред. А.П. Кудинова, Г.Г. Матвиенко. – СПб.: Изд-во Политехнического университета, 2007, Т.11, с.225-227.

3. Уоссермен Ф. Нейрокомпьютерная техника: Теория и практика. – Электронное издание.

4. Штовба С.Д. Введение в теорию нечетких множеств и нечеткую логику – Электронное издание.