

УДК 004.04

ПРОГНОЗИРОВАНИЕ ЧУВСТВИТЕЛЬНОСТИ К УДАРУ ВЗРЫВЧАТЫХ ВЕЩЕСТВ НЕЙРОННЫМИ СЕТЯМИ С ПРЕДВАРИТЕЛЬНОЙ ГРУППИРОВКОЙ ДАННЫХ

Стасюк В.В.

АО ФНПЦ «Алтай», Бийск, e-mail: ISVVI@yandex.ru

В работе представлены результаты прогноза чувствительности к удару h_{50} 226 взрывчатых веществ с общей формулой $C_aH_bN_cO_d$ с использованием простейших атомарно-молекулярных дескрипторов. Получены эмпирические выражения, используемые в качестве признаков при расчетах выходной характеристики. Показана эффективность применения группировки данных с учётом коэффициентов взаимной корреляции признаков и выходной характеристикой. Для разбиения данных на кластеры выбран алгоритм группировки Уорда, что привело к выделению группы веществ с высокой чувствительностью к удару. Прогнозирование проводилось на двух моделях нейронных сетей простой и сложной конфигурации, в которых варьировались размер обучающей выборки, количество нейронов и входные данные. Коэффициент корреляции при расчёте всей базы данных равен 0,785, а для второй группы – 0,892. Отношение средней ошибки результатов расчета и её среднеквадратического отклонения генеральной выборки к выделенной второй группе составило соответственно 6,00 и 9,33.

Ключевые слова: чувствительность к удару, дескрипторы, коэффициент корреляции, алгоритм группировки Уорда, нейронные сети

PREDICTION OF THE EXPLOSIVES IMPACT SENSITIVITY BY NEURAL NETWORKS WITH PRELIMINARY GROUPING DATA

Stasyuk V.V.

Federal Science-Production Centre «Altay», Biysk, e-mail: ISVVI@yandex.ru

This paper describes results of the calculating impact sensitivity h_{50} for $C_aH_bN_cO_d$ groups explosives using their atomic-molecular structure. New empiric expression including for prediction are introduced. Comparing of interdependence correlation coefficient values of various atomic-molecular descriptors several combinations are created. Ward's grouping algorithm was used for separation of data base. For calculating h_{50} for general and second bases used neural networks that have three neurons block with individual activation function in hidden layer and for the first one each layer connected only the previous layer. Best combinations of entrance variables for this case are found. Forecast's results for general samples have correlation coefficient $R = 0,785$, mean error value 35,01 cm and mean square of error deviation 33,80 cm. The predicted impact sensitivity for second group have better correlation coefficient $R = 0,892$, mean error value 5,80 cm and mean square of error deviation 3,62 cm. In case of first cluster mean error value 65,45 cm and mean square of error deviation 45,54 cm that are worse than calculated result for all base at once. But it is quite possible that repeated application of classify in second base reduce to improvement of prediction result.

Keywords: impact sensitivity, descriptors, correlation coefficient, Ward's grouping algorithm, neural networks

Из существующих характеристик, отражающих уровень опасности при обращении с взрывчатыми веществами (ВВ), можно выделить чувствительность к трению, электрическому разряду, тепловым импульсам и удару. Высокий уровень опасности представляет ударное механическое воздействие на материал. Существует несколько показателей, отражающих чувствительность к удару, среди которых наибольшее распространение получило значение высоты падения ВВ массой 2,5 кг, при котором наблюдается детонация с вероятностью 50%, характеризует чувствительность вещества и обозначается h_{50} . Работы по прогнозированию данной величины осуществлялись как с применением нейронных сетей [1], так и заданным математическим выражением [4]. В качестве входных данных для нейронных сетей служили молекулярные

дескрипторы, энергетические и физические характеристики определенных групп взрывчатых веществ.

Цель данной работы заключается в исследовании возможности прогноза h_{50} на основе атомарного состава ВВ и сравнение результатов прогноза для генеральной выборки и групп, выделенных по определенному алгоритму.

В качестве исходных данных взята информация о чувствительности h_{50} 226 ВВ с различной молекулярной структурой, рассмотренных в ряде статей [2, 3, 5, 6, 8]. Исходные данные представляют собой химические соединения с общей формулой $C_aH_bN_cO_d$, где индексы варьировались в пределах: $1 \leq a \leq 24$, $0 \leq b \leq 18$, $2 \leq c \leq 18$, $2 \leq d \leq 28$. Для исследования выбраны различные нитроароматические, нитроалифатические соединения и нитрамины.

Для исследования связи между характеристиками рассчитывались коэффициенты корреляции R между признаками и выходной величиной, а также непосредственно между дескрипторами – R_j .

Учитывая значения коэффициентов корреляции между чувствительностью к удару h_{50} и различными простыми выражениями типа C/O, (C + H)/O и т.д., получены эмпирические признаки AM_1 , AM_2 и AM_3 :

$$AM_1 = \log \left[\left(100 \frac{MW_c \cdot a + b/c}{MW \cdot d} \right)^2 + 1 \right]; \quad (1)$$

$$AM_2 = 100 \frac{a + b}{MW \cdot d}; \quad (2)$$

$$AM_3 = 100 \frac{a + c}{MW \cdot d}; \quad (3)$$

где AM_1 , AM_2 , AM_3 – атомарно-молекулярные дескрипторы; MW – молекулярная масса ВВ; MW_c – молекулярная масса углерода; a , b , c , d – число атомов в молекуле $C_a H_b N_c O_d$.

Следует отметить невысокую корреляцию между данными параметрами и кислородным балансом O_b , что позволяет использовать их совместно при расчетах. Выборочные числовые значения коэффициентов корреляции приведены в табл. 1.

Применение нейронных сетей для прогноза чувствительности к удару по всей базе веществ не приводит к удовлетворительным результатам. Для уменьшения ошибки прогноза принято решение о разбиении всей выборки на несколько частей с учётом различных наборов признаков.

Разбиение данных проводилось с помощью процедур кластерного анализа. Применение метода ближнего соседа, дальнего соседа и других аналогичных приводит к последовательному объединению объектов и, соответственно, невозможности выделения нескольких множеств. Среди различных алгоритмов выделения групп выбран метод Уорда [7], т.к. результатом его работы является разбиение на четко выраженные кластеры. Сформированные группы N_1 и N_2 представлены на рис. 1. В качестве характеристики близости объектов выбрано евклидово расстояние, а в качестве входных признаков рассматривались AM_1 (выражение (1)), $AM_1 \cdot AM_2$, (C + H)/O, (C + H)/(N + O), O/MW и O_b .

В результате группировки получены данные по связи признака AM_1 и h_{50} . Так при значениях данного параметра, больших 2, для всех веществ характерна низкая чувствительность к удару ≈ 300 см.

Распределение объектов в группах по исследуемому параметру представлено на рис. 2.

Как видно из графика, во вторую группу вошли взрывчатые вещества с высокой чувствительностью h_{50} . Для первой группы характерно наличие объектов во всем диапазоне исследуемого параметра и снижение коэффициента корреляции, что указывает на необходимость дальнейшего разбиения на подмножества в данном кластере. Для второй группы коэффициент корреляции между параметрами увеличился до 0,892, что с учётом существенного сужения диапазона h_{50} приводит к снижению абсолютной ошибки прогноза.

Таблица 1

Значения выборочных коэффициентов корреляции для генеральной совокупности

	AM_1	AM_2	$AM_1 \cdot AM_2$	(C + N)/O	(C + H)/O	...	C*H/O	O_b	h_{50}	$\lg(h_{50})$
AM_1	1,000	0,788	0,793	0,712	0,710	...	0,172	-0,749	0,574	0,652
AM_2	0,788	1,000	0,897	0,644	0,790	...	0,213	-0,481	0,519	0,503
$AM_1 \cdot AM_2$	0,793	0,897	1,000	0,790	0,721	...	0,169	-0,462	0,524	0,472
C/O	0,710	0,432	0,570	0,856	0,761	...	0,583	-0,890	0,564	0,631
(C + N)/O	0,712	0,644	0,790	1,000	0,815	...	0,445	-0,698	0,563	0,572
(C + H)/O	0,710	0,790	0,721	0,815	1,000	...	0,710	-0,792	0,600	0,649
...
C*H/O	0,172	0,213	0,169	0,445	0,710	...	1,000	-0,647	0,411	0,490
O_b	-0,749	-0,481	-0,462	-0,698	-0,792	...	-0,647	1,000	-0,603	-0,758
h_{50}	0,574	0,519	0,524	0,563	0,600	...	0,411	-0,603	1,000	0,882
$\lg(h_{50})$	0,652	0,503	0,472	0,572	0,649	...	0,490	-0,758	0,882	1,000

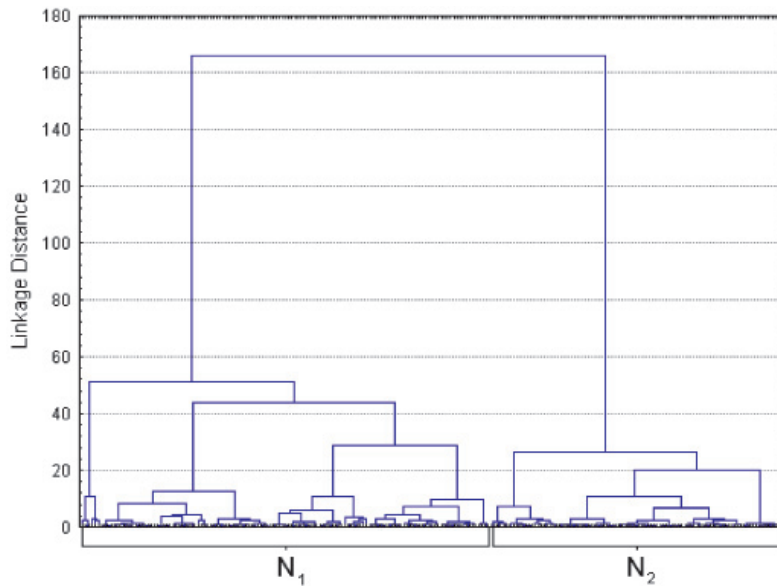


Рис. 1. Дендрограмма для генеральной выборки

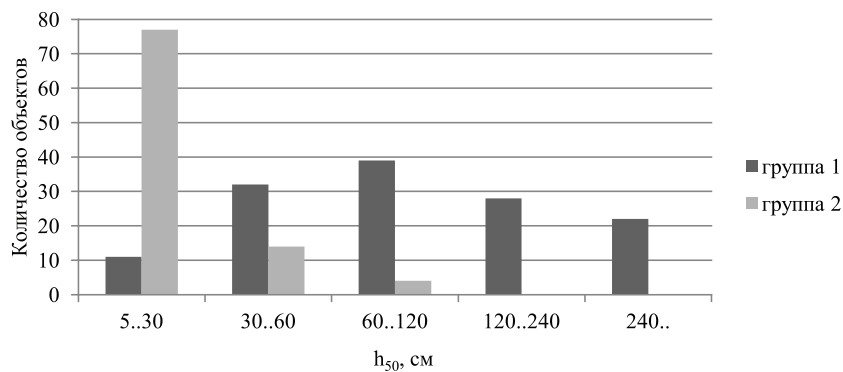


Рис. 2. Распределение веществ в группах по диапазонам чувствительности h_{50}

Таблица 2

Выбор входных параметров для первой группы

1,1	1,2	1,3
$R_i < 0,9$	$R_i < 0,8$	$R_i < 0,8$
$R > 0,3$	$R > 0,35$	$R > 0,35$
AM_1 AM_2 AM_3 $AM_1 \cdot AM_2$ $(C + N)/O$ $(C + H)/O$ $(C + H + N)/O$ O/MW O_b	$AM_1 \cdot AM_2$ $(C + N)/O$ $(C + H)/O$ O_b	AM_1 $(C + H)/O$ O_b
h_{50}		$lg(h_{50})$

Для определения влияния вариаций входных параметров рассматривались коэффициенты взаимной корреляции между дескрипторами и корреляция с h_{50} . В табл. 2 и 3 приведены рассмотренные наборы признаков.

При исследовании рассматривались результаты прогноза нейронными сетями с архитектурой, представленной на рис. 3.

Структура «а» является линейной нейронной сетью с входным слоем данных X , выходным – Y и одним промежуточным слоем M , в котором число нейронов с функцией активации логистического вида варьировалось от 1 до 9. Схема «б» характеризует нейронные сети Уорда с промежуточным слоем из трех блоков. В каждом из блоков используются различные передаточные функции: Гауссова, Гауссова комплементарная и гиперболическая тангенсальная. Количество нейронов между блоками одинаковое, т.к. отклонение от данного соотношения приводит к ухудшению результатов прогноза.

Таблица 3

Выбор входных параметров для второй группы

2,1	2,2	2,3	2,4	2,5	2,6	2,7
	$R_i < 0,9$	$R_i < 0,8$		$R_i < 0,9$	$R_i < 0,8$	
$R > 0,5$	$R > 0,55$		$R > 0,5$			$R > 0,55$
AM ₁ AM ₂ AM ₁ ·AM ₂ C/O (C+N)/O (C+H+N)/O O/MW O _b	AM ₁ AM ₁ ·AM ₂ C/O (C+N)/O (C+H+N)/O O _b	AM ₁ ·AM ₂ (C+N)/O (C+H+N)/O O _b	AM ₁ AM ₁ ·AM ₂ C/O (C+N)/O (C+H)/O (C+H+N)/O C/MW O/MW O _b	AM ₁ AM ₁ ·AM ₂ C/O (C+N)/O (C+H)/O (C+H+N)/O C/MW O _b	AM ₁ (C+N)/O (C+H)/O O _b	(C+N)/O O _b
h_{50}			$\lg(h_{50})$			

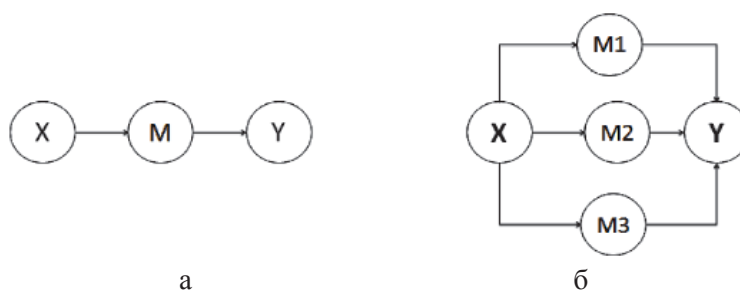


Рис. 3. Рассмотренные архитектуры нейронной сети

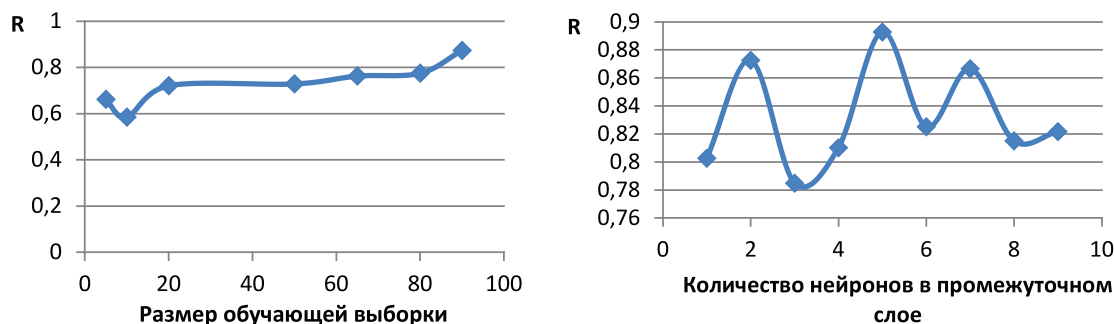


Рис. 4. Графики изменения коэффициента корреляции прогнозируемой величины для второй группы

Изменение коэффициента корреляции в зависимости от количества нейронов в промежуточном слое и размера обучающей выборки для второй группы представлено на рис. 4.

Изменение коэффициента корреляции при увеличении размера промежуточного слоя позволяет выбрать значение, равное 5.

Наилучшие результаты расчета при различных комбинациях дескрипторов получены для следующих условий:

– 1 группа – входные параметры 1,2, выходные параметры – h_{50} ; 90% обучающее множество; архитектура: нейронная

сеть Уорда с 5 нейронами в промежуточных блоках.

– 2 группа – входные параметры 2,3, выходные параметры – h_{50} ; 80% обучающее множество; архитектура: последовательная нейронная сеть с 2 нейронами в промежуточном слое.

Прогноз генеральной выборки в целом проводился при входных параметрах – (AM₁, C/O, (C+N)/O, (C+H)/(N+O), O/MW, O_b), выходным параметром взят $\lg(h_{50})$; 80% обучающее множество; архитектура: нейронной сети Уорда с 5 нейронами в промежуточных блоках.

Сравнительные результаты прогноза генеральной выборки и полученных кластеров приведены в табл. 4.

Таблица 4
Показатели погрешности результатов прогноза

Данные	Ср. ошибка, см	СКО, см
Генеральная выборка	35,01	33,80
1 группа	65,45	45,54
2 группа	5,80	3,62

Представленные выше данные показывают, что учёт коэффициентов корреляции и выделение групп с помощью процедур кластерного анализа приводит к улучшению результатов прогноза, полученных с помощью нейронных сетей.

В табл. 5 представлены результаты прогноза для некоторых взрывчатых веществ.

Таблица 5
Результаты прогноза чувствительности некоторых ВВ

Название	h_{50} , см	
	Эксп.	Прогноз
2, 4, 6, 8, 10, 12-Гексани- троhexoazaisowurzitane (ε-polymorph)	12	21,17
Hexanitrobenzene	11	15,08
Trinitroethylnitroguanidine	15	15,41
2, 3, 4, 5, 6-Pentonitroaniline	15	20,12
Benzotrifuroxan	53	50,21
1, 1, 1, 7, 7, 7-Hexanitroheptanone-4	34	31,48
Tetranitrate pentaerythritol	13	16,76
2', 2', 2'-trinitroethyl-2,4,6- trinitrobenzoate	24	28,80
Ammonium 3,5-dinitro-1,2,4- triazolate	110	110
Bis-(trinitroethyl)-oxalate	15	13,77

В результате проведенного исследования показана эффективность использования процедур кластерного анализа при прогнозе чувствительности h_{50} . Определено наилучшее сочетание входных признаков, тип и строение нейронной сети. Коэффициент корреляции при расчёте всей базы данных равен 0,785, а для второй группы – 0,892. Отношение средней ошибки результатов расчета и её среднеквадратического отклонения генеральной выборки к выделенной

второй группе составило соответственно 6,00 и 9,33. Для получения результатов прогноза генеральной выборки с меньшей погрешностью целесообразно исследовать применение алгоритма разбиения на подгруппы и расширить диапазон изменения характеристик нейронной сети.

Список литературы

1. Chen D.S. Neural Network Correlations of Detonation Properties of High Energy Explosives / D.S. Chen, D.S. H. Wong, C.-Y. Chen // *Propellants, Explos. – Pyrotech.* 23, 1998. – P. 296–301.
2. Keshavarz M.H. Simple Relationship for Prediction Impact Sensitivity of Nitroaromatics, Nitramines, and Nitroaliphatics // *Propellants, Explos. – Pyrotech.* 35, 2010. – P. 175–181.
3. Keshavarz M.H., Investigation of Various Structure Parameters for Predicting Impact Sensitivity of Energetic Molecules via Artificial Neural Network / M.H. Keshavarz and M. Jaafari // *Propellants, Explos. – Pyrotech.* 31. – 2006. – № 3. – P. 216–225.
4. Mullay J.A Relationship between Impact Sensitivity and Molecular Electronegativity / J. Mullay // *Propellants, Explos. – Pyrotech.* 12, 1987. – P. 60–63.
5. Nefati H. Prediction of the Impact Sensitivity by Neural Networks / H. Nefati, J.-M. Cense, and J.-J. LegJ. *Chem. Inf. Comput. Sci.* 36. – 1996. – P. 804–810.
6. Song X.-Sh. Relationship between the Bond Dissociation Energies and Impact Sensitivities of Some Nitro-Explosives / X.-S. Song, X.-L. Cheng, X.-D. Yang, B. He // *Propellants, Explos., Pyrotech.* 31. – 2006. – № 4. – P. 306–310.
7. Ward J. H. Hierarchical Grouping to Optimize an Objective Function / J.H. Ward, Jr. // *Journal of American Statistical Association*, 58. – 1963. – P. 236–244.
8. Wingborg N. Triaminoguanidine Dinitramide, TAGDN: Synthesis and Characterization / N. Wingborg and N.V. Latypov // *Propellants, Explos., Pyrotech.* 28. – 2003. – № 6. – P. 314–318.

References

1. Chen D.S., Wong D.S.H., Chen Ch.-Y., Neural Network Correlations of Detonation Properties of High Energy Explosives, *Propellants, Explos., Pyrotech.* 23, 1998, pp. 296–301.
2. Keshavarz M.H., Jaafari M., Investigation of Various Structure Parameters for Predicting Impact Sensitivity of Energetic Molecules via Artificial Neural Network, *Propellants, Explos., Pyrotech.* 31, no. 3, 2006, pp. 216–225.
3. Keshavarz M.H., Simple Relationship for Prediction Impact Sensitivity of Nitroaromatics, Nitramines, and Nitroaliphatics, *Propellants, Explos., Pyrotech.* 35, 2010, pp. 175–181.
4. Mullay J., A Relationship between Impact Sensitivity and Molecular Electronegativity, *Propellants, Explos., Pyrotech.* 12, 1987, pp. 60–63.
5. Nefati H., Cense J.-M., Legendre J.-J., Prediction of the Impact Sensitivity by Neural Networks, *J. Chem. Inf. Comput. Sci.* 36, 1996, pp. 804–810.
6. Song X.-Sh., Cheng X.-L., Yang X.-D., He B., *Propellants, Explos., Pyrotech.* 31, no. 4, 2006, pp. 306–310.
7. Ward J.H. Hierarchical Grouping to Optimize an Objective Function / J.H. Ward, Jr. *Journal of American Statistical Association*, 58, 1963. pp. 236–244.
8. Wingborg N., Latypov N.V., Triaminoguanidine Dinitramide, TAGDN: Synthesis and Characterization, *Propellants, Explos., Pyrotech.* 28, no. 6, 2003, pp. 314–318.