

УДК 519.254:004.032.26

МЕТОДЫ СТАТИСТИЧЕСКОГО АНАЛИЗА И НЕЙРОСЕТЕВЫЕ ТЕХНОЛОГИИ ДЛЯ ПРОГНОЗИРОВАНИЯ ПОКАЗАТЕЛЕЙ КАЧЕСТВА МЕТАЛЛУРГИЧЕСКОГО КОКСА

В. С. Дороганов, А. Г. Пимонов

METHODS OF STATISTICAL ANALYSIS AND NEURAL NETWORK TECHNOLOGIES FOR PREDICTION OF METALLURGICAL COKE QUALITY

V. S. Doroganov, A. G. Pimonov

В статье представлены результаты проведенного авторами статистического анализа зависимости показателей качества металлургического кокса от характеристик шихты. Предложена нейросетевая модель, позволяющая как осуществить прогноз характеристик кокса, так и подобрать состав шихты из имеющихся углей для получения кокса требуемого качества. Описаны особенности программной реализации предложенной нейросетевой модели.

The paper presents the results of the statistical analysis of the influence of coal concentrates mixture characteristics on metallurgical coke quality. The authors provide a neural network model, which allows forecasting coke characteristics. It helps to choose the mixture composition to obtain quality coke. The features of software implementation of the proposed neural network model are described.

Ключевые слова: статистический анализ, нейронные сети, генетические алгоритмы, металлургический кокс, прогнозирование, многопоточные вычисления.

Keywords: statistical analysis, neural networks, genetic algorithms, metallurgical coke, forecasting, multithreaded computing.

Строительство домов, машин, различной техники и механизмов требует огромных затрат твердых металлов – стали и чугуна. Металлы в природе, как правило, встречаются не в чистом, самородном состоянии, а в виде химических соединений. Чтобы добыть чистые металлы из руд, их подвергают соответствующей обработке. Кокс – главное топливо при выплавке чугуна в доменных печах. Он также используется в производстве теплоизоляционных материалов, в цветной металлургии и т. д. Одной из главных задач коксохимической промышленности является контроль и повышение качества получаемого кокса. Качественные характеристики кокса напрямую зависят от характеристик шихты (обогащенное рудное сырье с топливом – углем), из которой он изготавливается, а также от печи, в которой запекается шихта. Шихта имеет восемь основных параметров:

- 1) V_t – объемная доля витринита;
- 2) OK – содержание фюзенизированных компонентов;
- 3) Ad – зольность на сухое состояние;
- 4) V_{daf} – выход летучих веществ на сухое беззольное состояние;
- 5) PK – пористость;
- 6) I_B – объемная доля интернита;
- 7) R_0 – показатель отражения витринита;
- 8) W – влажность гигроскопическая.

Кокс же характеризуется двумя показателями: 1) CRI (Coke Reactivity Index) – индексом реакционной способности и 2) CSR (Coke Strength after Reaction) – индексом прочности кокса после взаимодействия. Несмотря на то, что производство кокса было изобретено еще в XVII веке, единой формулы расчёта его показателей на основе характеристик шихты до сих пор не существует.

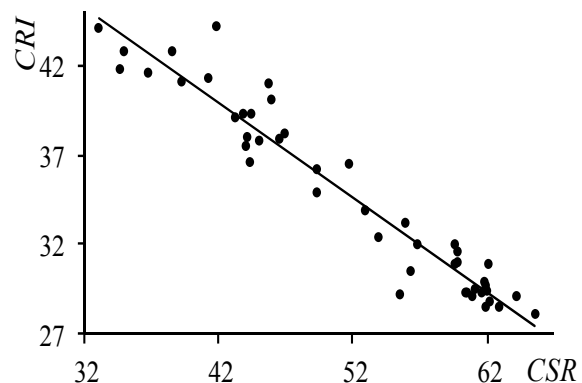


Рис. 1. Линейная корреляция CRI и CSR

Статистический анализ

Для построения такой формулы была предпринята попытка использования классических методов статистического анализа. На коксохимическом предприятии г. Кемерово для этой цели были собраны статистические данные, позволившие выполнить предварительный анализ. Для оценки степени влияния конкретной коксовой батареи на конечный результат к восьми вышеперечисленным числовым параметрам шихты была добавлена девятая качественная характеристика – название или номер батареи. Анализ проводился на выборке объемом в 112 наблюдений, 66 из которых были неполными.

Вопросы взаимозависимости параметров кокса неоднократно обсуждались в современных исследованиях [2 – 4; 15; 17]. Для выявления степени взаимосвязи показателей качества металлургического кокса (CRI и CSR) был проведён корреляционный анализ. В качестве программного инструментария для анализа был использован пакет Statistica версии 10 [12]. Вычисленный коэффициент корреляции получился дос-

таточно высоким $R = 0,97$. Построенная линия регрессии $CSR = 113,43 - 1,77 CRI$ приведена на рис. 1.

Парная корреляция между характеристиками шихты и показателями качества металлургического кокса (таблица 1) была рассчитана без учета влияния коксовой батареи (ее номера), так как этот показатель представлен не в интервальной шкале, а в шкале наименований (номиналов). Для первых пяти характеристик подтверждается гипотеза о незначительной корреляции. Наличие высокой корреляции с тремя последними характеристиками (I_b , R_0 , W) хорошо иллюстрируется графически (рис. 2 – 4). Из графиков видно, что связь между характеристиками шихты и индексом реакционной способности кокса CRI носит ярко выраженный нелинейных характер. Аналогичные результаты получены и для CSR – индекса прочности кокса после взаимодействия.

Таблица 1

Коэффициенты парной корреляции R между характеристиками шихты и показателями качества кокса

Шихта \ Кокс	CRI	CSR
V_t	-0,05	0,07
OK	0,07	-0,10
Ad	-0,23	0,27
V_{daf}	0,21	-0,25
$ПК$	-0,35	0,41
I_b	-0,52	0,53
R_0	-0,52	0,53
W	0,51	-0,56

Степень влияния условий коксования (коксовой батареи) на показатели качества металлургического кокса была оценена с помощью дискриминантного анализа. С этой целью была построена матрица классификаций (таблица 2). Анализ результатов показал, что третья батарея определяется в 88,23 % случаев, пятая – в 37,50 %, а шестая – в 65,21 %.

Таблица 2

Матрица классификаций для определения номера батареи по значениям CRI и CSR

Эталонная батарея	Верных определений, %	Классифицированная батарея		
		№ 3	№ 5	№ 6
№ 3	88,24	15	1	1
№ 5	37,50	5	6	5
№ 6	69,23	1	3	9
Итого	65,22	21	10	15

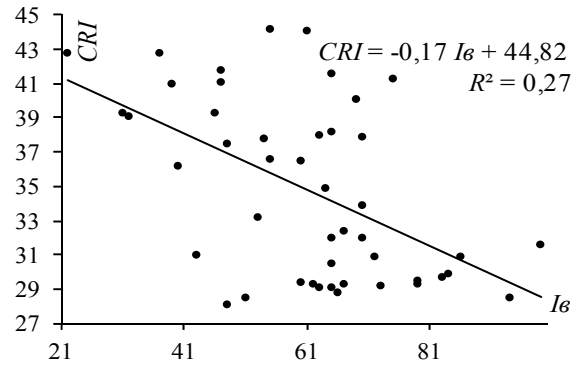


Рис. 2. Линейная корреляция CRI и I_b

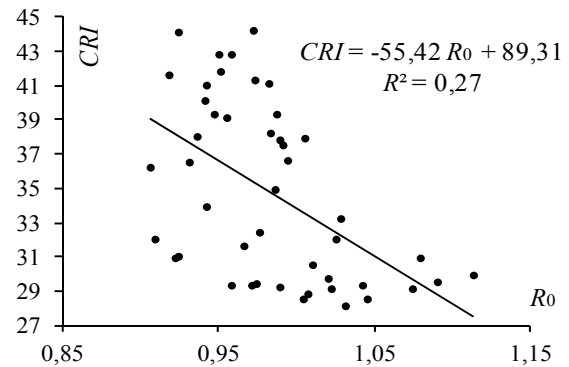


Рис. 3. Линейная корреляция CRI и R_0

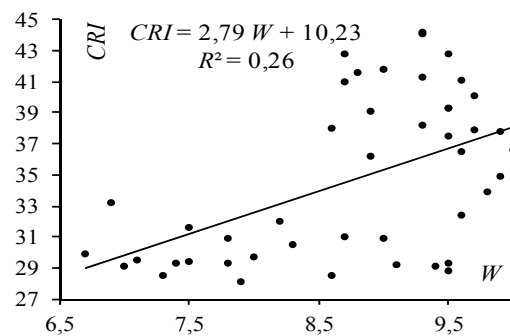


Рис. 4. Линейная корреляция CRI и W

Каждая из коксовых батарей определяет свой, характерный для нее, диапазон значений CRI и CSR (рис. 5 – 6). Диаграммы изменения характеристик CRI и CSR для каждой батареи на рис. 5 – 6 представлены в виде свечей, где точками обозначены медианные значения, прямоугольниками – нижние (25 %) и верхние (75 %) значения квартилей, отрезком – разброс (минимальные и максимальные значения).

При этом монотонность полученных зависимостей (рис. 5 – 6) указывает на возможность успешного использования ранговой корреляции. Вычисленный коэффициент ранговой корреляции Спирмена между номером коксовой батареи и показателем CRI получился равным 0,67. Просматривается наличие умеренной связи [1], но количественно установить степень этой связи не представляется возможным, так как один показатель измерен в шкале категорий (номер коксовой батареи), а второй – в интервальной шкале (индекс реакционной способности CRI).

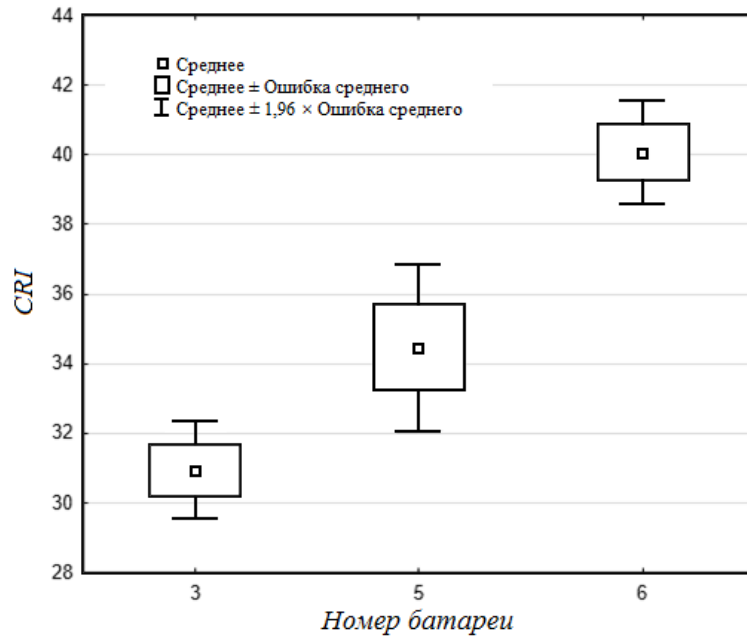


Рис. 5. Зависимость CRI от условий коксования

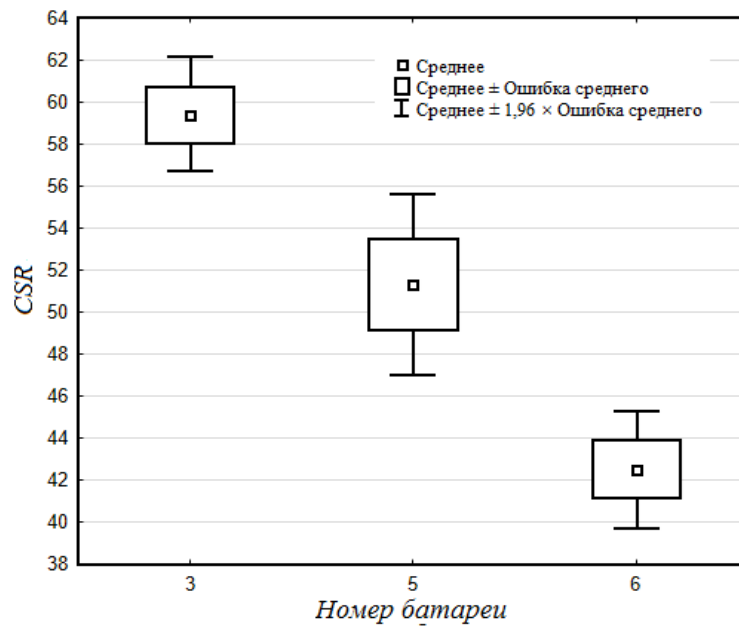


Рис. 6. Зависимость CSR от условий коксования

Для анализа связи между характеристиками шихты (независимые переменные) и индексом реакционной способности кокса (зависимая переменная) было построено уравнение множественной линейной регрессии (таблицы 3 – 4). Уравнение, построенное по результатам регрессионного анализа (таблица 4), объясняет 45,6 % вариации зависимой переменной. Значение коэффициента детерминации $R^2 = 0,456$ достаточно велико ($R = 0,675$, то есть 67,5 % всей изменчивости объясняется параметрами модели, в которой не учтено влияние коксовой батареи). В таблице 3 представлены рассчитанные значения коэффициентов детерминации. Это доля дисперсии зависимой переменной, объясняемая рассматриваемой моделью, то есть

независимыми переменными. Хотя в модели и не учтено влияние коксовых батарей, но из таблицы 3 видно, что большая часть дисперсий индекса реакционной способности и индекса прочности кокса после взаимодействия объяснима перечисленными в ней параметрами шихты. Таким образом, проведенный анализ показал, что прогноз качественных характеристик металлургического кокса на основе параметров запекаемой шихты с помощью классических методов математической статистики представляется весьма затруднительной процедурой.

Таблица 3

Множественная регрессия

	<i>Остаточная дисперсия</i>	<i>Коэффициент детерминации</i>
<i>CRI</i>	0,068	0,932
<i>CSR</i>	0,068	0,932
<i>OK</i>	0,981	0,019
<i>Ad</i>	0,913	0,087
<i>V_{daf}</i>	0,921	0,079
<i>I_в</i>	0,716	0,284
<i>R₀</i>	0,715	0,285
<i>ПК</i>	0,789	0,211
<i>W</i>	0,662	0,338

Нейронная сеть

Искусственные нейронные сети – набор математических и алгоритмических методов для решения широкого круга задач [8 – 11; 13 – 14]. Любая нейронная сеть представляет собой совокупность элементов (искусственных нейронов), связанных между собой синоптическими соединениями (весами). Принцип работы нейронной сети заключается в преобразовании входных сигналов, в результате чего изменяется внутреннее состояние сети и формируются выходные значения.

В основу разработанной программы для прогнозирования показателей качества металлургического кокса положена авторская топология нейронной сети, основанная на сети Ворда. Топология последней характеризуется тем, что внутренние (скрытые слои) нейронов разбиты на блоки с разными активационными функциями. Модификация заключалась в том, что каждый нейрон скрытого слоя был представлен группой, состоящей из пяти простых нейронов с разными передаточными функциями, разными параметрами функций и своими весовыми коэффициентами. Входной слой имеет 11 нейронов, выходной – 2. Число скрытых слоев и количество нейронов в каждом из них устанавливается пользователем.

Значение передаточной функции составного нейрона, определяющей нелинейное преобразование,

вычисляется как сумма значений передаточных функций простых нейронов. В качестве передаточных функций простых нейронов используются следующие: логистическая; тождественная; гиперболическая; экспоненциальная; синусоидальная.

Предложенная нейронная сеть реализована в виде программного продукта, созданного на платформе .NET 4.5, язык программирования – C#. Диаграмма классов программного инструментария представлена на рис. 7.

Поскольку алгоритм обратного распространения ошибки не предполагает изменения параметров передаточных функций, то для обучения сети был использован стохастический метод [16], в котором выполняются случайные изменения величин весов и характеристик передаточных функций составных нейронов [6 – 7]. При этом изменения, ведущие к улучшениям, сохраняются. Сохранение улучшающих изменений реализовано с помощью генетического алгоритма следующим образом. Изначально генерируется массив нейронных сетей. Каждая из полученных сетей дублируется в количестве 10 экземпляров. Все сгенерированные сети («потомки») обучаются стохастическим методом. Работа каждой сети на основе эмпирических данных оценивается по среднеквадратическому отклонению результатов прогнозирования показателей качества металлургического кокса. Из них отбираются лучшие, которые затем и становятся «родителями» для генерации новой выборки.

Для исследования зависимости качества прогноза и времени обучения от топологии сети было проведено пробное обучение в 10000 эпох. При этом анализировались результаты работы двух версий программного продукта (таблица 5), в одной из которых были реализованы последовательные вычисления, в другой – параллельные. На основе анализа этих результатов получено, что значительный прирост производительности из-за распараллеливания вычислений наблюдается при значительном увеличении количества нейронов в отдельном слое и не зависит от количества слоёв.

Таблица 4

Параметры уравнения множественной линейной регрессии

	<i>Значения оценок неизвестных коэффициентов регрессии</i>	<i>Стандартные ошибки оценки коэффициентов</i>	<i>Значение статистики Стьюдента для проверки гипотезы о нулевом значении коэффициента</i>	<i>Уровень значимости принятия этой гипотезы</i>
<i>Свобод. член</i>	54,60	117,24	0,47	0,64
<i>V_t</i>	0,30	1,23	0,24	0,81
<i>OK</i>	0,45	1,26	0,35	0,73
<i>Ad</i>	-1,32	1,64	-0,80	0,43
<i>V_{daf}</i>	-0,26	0,90	-0,29	0,78
<i>I_в</i>	-0,08	0,06	-1,39	0,17
<i>R₀</i>	-34,14	21,11	-1,62	0,11
<i>ПК</i>	0,27	0,43	-0,63	0,53
<i>W</i>	1,11	1,02	1,09	0,28

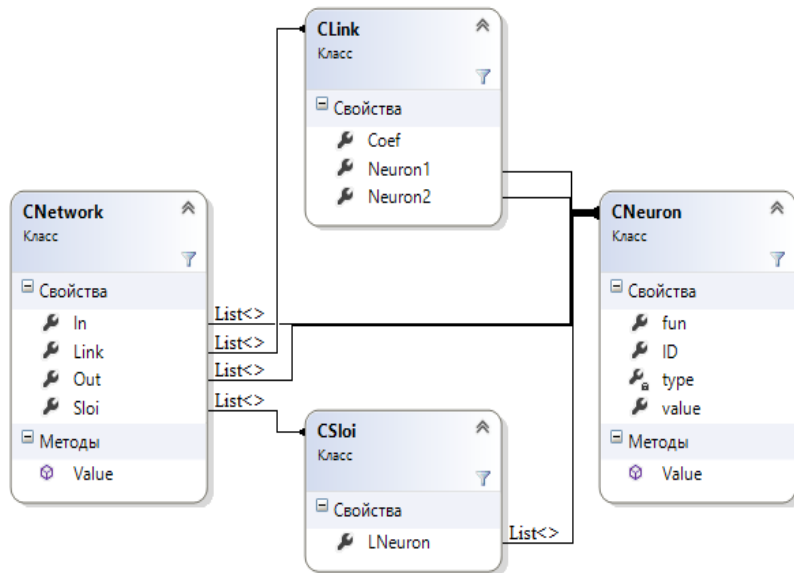


Рис. 7. Диаграмма классов

Таблица 5
 Результаты работы сетей разного размера

Количество, шт.		Время обучения, с		Ошибка прогноза, %	
Слоёв	Нейронов	Последовательно	Параллельно	CRI	CSR
1	1	7	10	22	15
2	22	1055	400	11	10
7	7	701	256	12	15
22	2	429	388	20	14

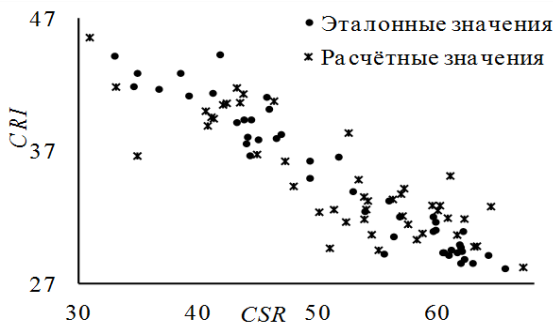


Рис. 8. Результаты прогноза показателей качества металлургического кокса

В результате проведения вычислительных экспериментов было получено, что лучшее соотношение точности прогноза показателей качества кокса и времени обучения нейронной сети достигается при 3 скрытых слоях и 15 нейронах в каждом из них (рис. 8). В результате работы нейронной сети, обученной в течение 24 часов, средняя относительная ошибка прогноза показателей составила 9 % и 7 % для CRI и CSR соответственно.

Обученная нейронная сеть позволяет проводить прогноз показателей качества металлургического кок-

са на основе вводимых характеристик шихты (рис. 9). Изменяя характеристики шихты и условия ее запекания (номер коксовой батареи), можно оценить качество изготовляемого кокса.

Полученная обученная нейронная сеть позволяет осуществить подбор угольной смеси (рис. 10). Угольная смесь может быть подобрана как для достижения требуемых характеристик шихты, так и показателей качества металлургического кокса. Смесь подбирается из состава автоматизированной базы данных имеющихся в наличии углей. В результате работы программа предоставляет процентное соотношение углей в составе возможной смеси. Подбор смеси также реализован с использованием генетического алгоритма. Программа [5] генерирует набор возможных вариантов угольных смесей и оценивает каждую из них, оставляя лучшие. Затем каждый вариант подвергается изменению с целью улучшения качества смеси. Из числа полученных отбираются лучшие, и процесс повторяется до достижения приемлемых расхождений между заданными и получаемыми характеристиками.

Рис. 9. Окно ввода характеристик шихты и вывода результатов прогноза показателей качества кокса

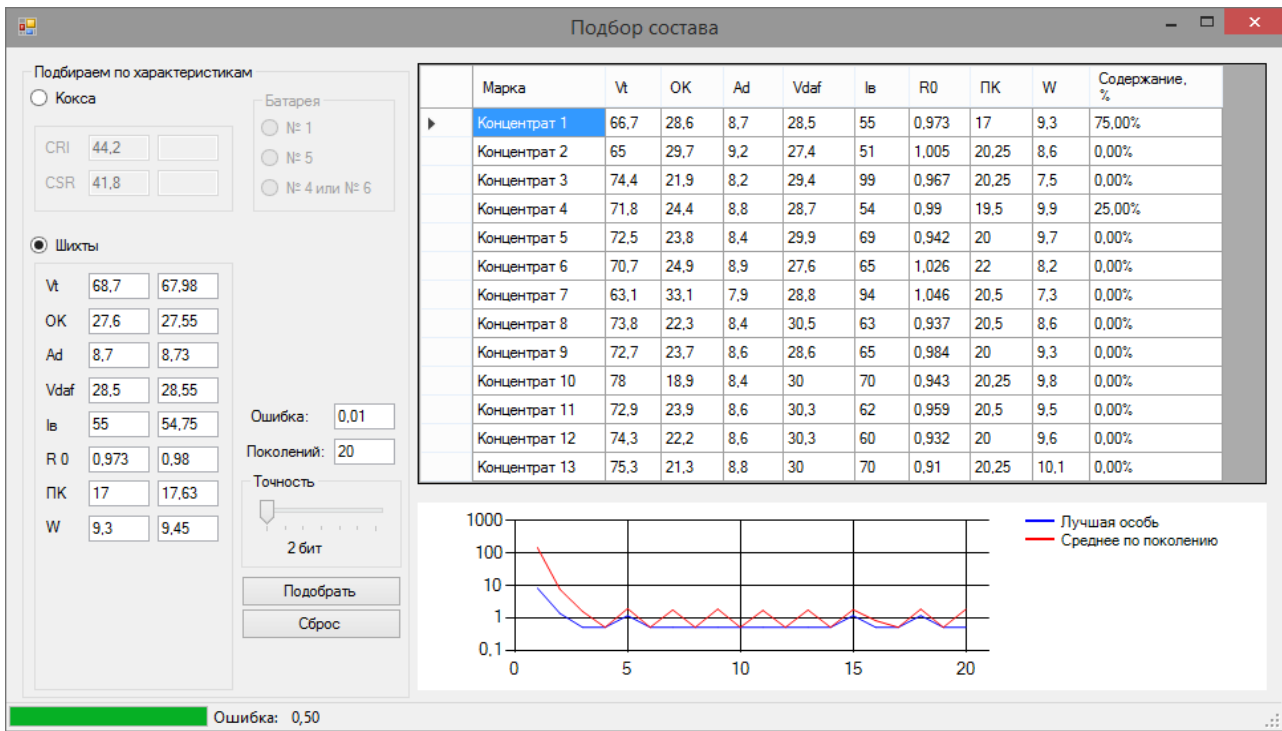


Рис. 10. Форма подбора состава угольной смеси

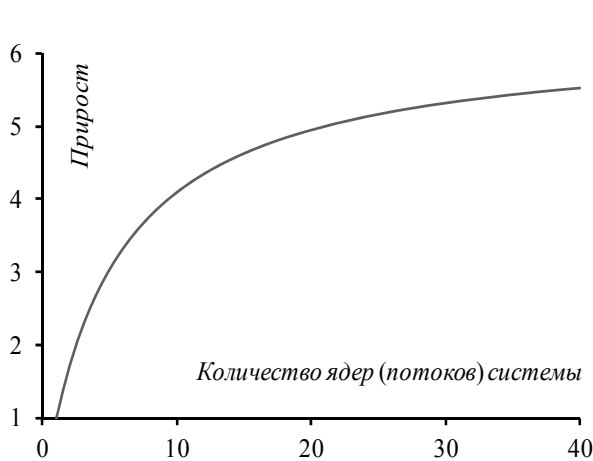


Рис. 11. Зависимость прироста производительности программы от количества ядер

Особенности программной реализации

Нейронные сети являются идеальным объектом для реализации параллельных вычислений из-за обособленности каждого нейрона. Для ускорения обучения сетей в программном продукте были использованы возможности платформы .Net 4.5. Все слои обрабатываются последовательно, начиная с первого. Внутри каждого слоя нейроны обчитываются параллельно. Время одной эпохи уменьшилось в 2,7 раза (таблица 5) – пропорционально количеству потоков процессора (использовался двухядерный процессор Intel Core i5-2410M с технологией Hyper-threading, видеокарта NVidia GeForce GT520M). Согласно закону Амдала ускорение, которое может быть получено на вычислительной системе из p процессоров, по сравнению с однопроцессорным решением не будет превышать величины

$$S_p = \frac{1}{\alpha + \frac{1-\alpha}{p}}$$

где S_p – максимальное теоретическое ускорение, α – объём вычислений, которые могут быть выполнены только последовательно, p – количество процессоров (потоков). В результате подстановки в эту формулу значений $S_p = 2,7$ и $p = 4$ получается $\alpha = 0,16$. То есть теоретический объём вычислений, которые можно осуществить параллельно в рассматриваемой задаче, равняется 84 %.

Для увеличения скорости вычислений планируется реализация программного продукта на более мощном компьютере с применением GPGPU – техники использования графического процессора видеокарты. Значение величины производительности флопс (FLOPS от англ. *F*loating-*O*perations *P*er *S*econd – количество операций с плавающей точкой в секунду) GPU относительно CPU на планируемом компьютере более чем в 15 раз больше (77,76 гигафлопс против 5 гигафлопс), а количество ядер CUDA составляет 48 штук. Это теоретически должно привести к увеличению скорости обучения нейронной сети в 5,5 (в сравнении с однопроцессорной системой) и двукратному увеличению по отношению к реализованному решению (рис. 11).

Увеличение скорости работы программного продукта откроет новые возможности для повышения качества прогнозирования показателей качества металлургического кокса за счет увеличения размера нейронной сети и подбора более сложной передаточной функции составного нейрона. Для достижения более высокого прироста производительности при многопоточных вычислениях ведется работа над оптимизацией исходного кода программы.

Литература

1. Афанасьев В. В. Теория вероятностей в вопросах и задачах. Режим доступа: <http://cito-web.yspu.-org/link1/metod/theory/node42.html>, свободный.
2. Базегский А. Е., Рябиченко А. Д., Станкевич А. С. Взаимосвязь показателей CSR и CRI кокса с генетическими характеристиками шихты и технологическими параметрами коксования // *Материалы научно-практической конференции «Химия и химическая технология на рубеже тысячелетий»*. Томск: ТПУ, 2000. С. 10 – 12.
3. Буланов Е. А., Зайнутдинов В. Н., Кузнецов В. Я., Зиновьева Л. А., Цекот Ю. Б. Прогноз горячей прочности (CSR) и реакционной способности (CRI) кокса // *Кокс и химия*. 2005. № 5. С. 23 – 26.
4. Гуляев В. М., Барский В. Д. Реакционная способность, «горячая» и «холодная» прочность доменного кокса // *Теория и практика металлургии*. 2010. № 1 – 2'. С. 7 – 10.
5. Дороганов В. С. Интеллектуальная информационная система подготовки шихты для производства кокса заданного качества // *Сборник материалов II Всероссийской научно-практической конференции «Современные проблемы производства кокса и переработки продуктов коксования»*. Кемерово: КузГТУ. 2014. С. 5 – 7.
6. Дороганов В. С. Использование элементов генетического алгоритма в обучении нейронной сети стохастическим методом // *Перспективы развития информационных технологий: труды Всероссийской молодежной научно-практической конференции*, г. Кемерово, 29 – 30 мая 2014 г. Кемерово, 2014. С. 214 – 215.
7. Дороганов В. С., Суханова Е. Ю. Прогнозирование характеристик кокса на основе показателей шихты // *Труды Всероссийской молодежной школы-семинара «Анализ, геометрия и топология»*. Барнаул: АлтГУ, 2013. С. 46 – 50.
8. Евсюткин Т. В., Тунгусова А. В. Классификация облачности по спутниковым снимкам вероятностной нейронной сетью // *Материалы 51-й Международной научной студенческой конференции «Студент и научно-технический прогресс»*. Новосибирск: НГУ, 2013. С. 131.
9. Корниенко И. Л. Обучение нейросетевой информационной системы для определения состава угольного концентрата с использованием базы данных уникальных углей // *Информационные системы и технологии в образовании, науке и бизнесе (ИСИТ-2014)*. Материалы Всероссийской молодежной научно-практической школы, г. Кемерово, 19 – 21 июня 2014 г. С. 127 – 128.
10. Кулманакоева Е. В. Отслеживание движения людей в видеопотоке // *Сборник трудов XVIII Международной научно-практической конференции студентов, аспирантов и молодых ученых «Современные техника и технологии»*. Томск: ТПУ, 2012. С. 335 – 336.
11. Миндияров Н. И., Дороганов В. С. Программа обучения искусственного интеллекта при помощи нейронной сети на примере игры «гонки» // *Материалы Всероссийской молодежной конференции «Информационно-телекоммуникационные системы и технологии» (ИТСИТ-2012)*. Кемерово: КузГТУ, 2012. С. 132 – 133.
12. Официальный сайт «StatSoft». Режим доступа: <http://www.statsoft.ru/products/trial/>, свободный.
13. Пахарев Р. С. Разработка программного обеспечения для задач неразрушающего контроля материалов с неоднородностями // *Материалы 51-й Международной научной студенческой конференции «Студент и научно-технический прогресс»*. Новосибирск: НГУ, 2013. С. 53.
14. Погорелов Н. Е., Рейзенбук К. Э., Пимонов А. Г. Интеллектуальная информационная система для анализа и прогнозирования биржевых котировок акций // *Вестник Кузбасского государственного технического университета*. 2012. № 6. С. 118 – 122.
15. Станкевич А. С., Гилязетдинов Р. Р., Попова Н. К., Кошкаров Д. А. Модель прогноза показателей CSR и CRI кокса на основе химико-петрографических параметров угольных шихт и условий их коксования // *Кокс и химия*. 2008. № 9. С. 37 – 44.
16. Уоссермен Ф. *Нейрокомпьютерная техника: Теория и практика*. М.: Мир, 1992. 118 с.
17. Díez M. A., Alvarez R., Barriocanal C. Coal for metallurgical coke production: predictions of coke quality and future requirements for cokemaking // *International Journal of Coal Geology*. 2002. № 50. P. 389 – 412.

Информация об авторах:

Дороганов Виталий Сергеевич – аспирант кафедры прикладных информационных технологий Кузбасского государственного технического университета им. Т. Ф. Горбачева, DoroganovV@mail.ru.

Vitaly S. Doroganov – post-graduate student at the Department of Applied Information Technologies, Kuzbass State Technical University named after T. F. Gorbachev, Kemerovo.

(**Научный руководитель – А. Г. Пимонов**).

Пимонов Александр Григорьевич – доктор технических наук, профессор кафедры прикладных информационных технологий КузГТУ им. Т. Ф. Горбачева, pag_vt@kuzstu.ru.

Alexander G. Pimonov – Doctor of Technical Science, Professor, Professor at the Department of Applied Information Technologies, Kuzbass State Technical University named after T. F. Gorbachev, Kemerovo.

Статья поступила в редколлегию 24.10.2014 г.