

Министерство образования и науки Российской Федерации

УДК 620.179.1.05
ГРНТИ 59.45.00
Инв. № 2009-30К

ПРИНЯТО:	УТВЕРЖДЕНО:
Исполнитель Государственное образовательное учреждение высшего профессионального образования "Магнитогорский государственный технический университет им. Г.И. Носова"	Государственный заказчик Министерство образования и науки Российской Федерации
От имени Руководителя организации _____/ Колокольцев В.М./	От имени Государственного заказчика _____/ _____/

НАУЧНО-ТЕХНИЧЕСКИЙ ОТЧЕТ

о выполнении 2 этапа Государственного контракта
№ П2136 от 05 ноября 2009 г. и Дополнению от 02 апреля 2010 № 1

Исполнитель: Государственное образовательное учреждение высшего профессионального образования "Магнитогорский государственный технический университет им. Г.И. Носова"
Программа (мероприятие): Федеральная целевая программ «Научные и научно-педагогические кадры инновационной России» на 2009-2013 гг., в рамках реализации мероприятия № 1.3.2 Проведение научных исследований целевыми аспирантами.
Проект: Наименование исследований: Многофункциональное приборостроение для промышленных систем управления. Наименование проблемы исследования: Разработка прибора для диагностирования аномалий процессов в неоднородных средах
Руководитель проекта: _____/Мещеряков А.Ю. ПОДПИСЬ

Магнитогорск
2010 г.

СПИСОК ОСНОВНЫХ ИСПОЛНИТЕЛЕЙ

по Государственному контракту № П2136 от 05 ноября 2009 г. на выполнение поисковых научно-исследовательских работ для государственных нужд

Организация-Исполнитель: Государственное образовательное учреждение высшего профессионального образования "Магнитогорский государственный технический университет им. Г.И. Носова"

Руководитель темы:

без ученой степени, без
ученого звания _____ Мещеряков А. Ю.
подпись, дата

Исполнители темы:

доктор технических наук,
профессор _____ Лукьянов С. И.
подпись, дата

кандидат технических
наук, доцент _____ Швидченко Д. В.
подпись, дата

кандидат технических
наук, доцент _____ Суспицын Е. С.
подпись, дата

без ученой степени, без
ученого звания _____ Швидченко Н. В.
подпись, дата

без ученой степени, без
ученого звания _____ Аносов Р. А.
подпись, дата

без ученой степени, без
ученого звания _____ Евстегнеев А. А.
подпись, дата

без ученой степени, без
ученого звания _____ Федоров А. С.
подпись, дата

Реферат

Отчет 51 с., 4 ч., 8 рис., 7 табл., 16 источн., 0 прил.

ДИАГНОСТИРОВАНИЕ, ВИБРАЦИЯ, НЕОДНОРОДНАЯ СРЕДА, КИСЛОРОДНЫЙ КОНВЕРТЕР, ПРОДУВКА, ПРИБОРОСТРОЕНИЕ, УПРАВЛЕНИЕ, МАШИННОЕ ОБУЧЕНИЕ

В отчете представлены результаты исследований, выполненных по 2 этапу Государственного контракта № П2136 "Наименование исследований: Многофункциональное приборостроение для промышленных систем управления.

Наименование проблемы исследования: Разработка прибора для диагностирования аномалий процессов в неоднородных средах." (шифр "НК-394П") от 05 ноября 2009 по направлению "Многофункциональное приборостроение для промышленных систем управления" в рамках мероприятия 1.3.2 "Проведение научных исследований целевыми аспирантами.", мероприятия 1.3 "Проведение научных исследований молодыми учеными - кандидатами наук и целевыми аспирантами в научно-образовательных центрах", направления 1 "Стимулирование закрепления молодежи в сфере науки, образования и высоких технологий." федеральной целевой программы "Научные и научно-педагогические кадры инновационной России" на 2009-2013 годы.

Цель работы - разработка прибора для диагностирования аномалий процессов в неоднородных средах, обеспечивающего раннюю идентификацию тенденций к возникновению аномальных ситуаций.

Для составления аналитического обзора использовались методы теоретических исследований. Для анализа вибрационного сигнала применялся метод спектрального анализа. Для обработки массива диагностических данных применялись методы математической статистики. Синтез алгоритмов диагностирования аномалий процессов в неоднородных средах осуществлялся при помощи методов машинного обучения.

Для сбора и анализа данных о вибрации кислородного конвертера применялась система на основе платформы National Instruments PXI в составе контроллера PXI-

8104, платы АЦП РХІ-4072, модуля согласования SCXI-1530. В качестве датчика первичной информации применен акселерометр Н1С производства НПХ "Техноап".

Проведен статистический анализ собранного массива экспериментальных данных. Проведен анализ методов машинного обучения, исследованы области их применения. Выбраны и адаптированы методы машинного обучения для диагностирования аномалий процессов в неоднородных средах. Приведены методики экспериментальных исследований. Проведено экспериментальное сравнение результатов применения методов машинного обучения для диагностирования аномалий в технологическом процессе продувки кислородного конвертера. Разработаны обобщенные алгоритмы работы вычислительного блока прибора для диагностирования аномалий процессов в неоднородных средах. Проведена апробация разработанного прибора в условиях ККЦ ОАО "ММК". Проанализированы возможные области применения разработанного прибора в реальном секторе экономики.

Содержание

Реферат	3
Введение	6
1 Аннотированная справка по научным результатам НИР, полученным на I этапе....	7
2 Аналитический отчет о проведении теоретических и экспериментальных исследований.....	9
2.1 Методика сбора информации о значениях входных и выходных параметров....	9
2.2 Статистическая обработка экспериментальных данных	10
2.3 Постановка задачи автоматического синтеза диагностической модели	15
2.4 Выбор значимых признаков	16
2.5 Обзор методов машинного обучения.....	18
2.5.1 Многослойные нейронные сети.....	19
2.5.2 Радиальные базисные сети	23
2.5.3 Деревья решений	24
2.5.4 Машины опорных векторов	25
2.6 Методы оценки качества машинного обучения.....	26
2.6.1 Скользящий контроль	27
2.6.2 ROC-кривая.....	30
2.7 Экспериментальное сравнение методов машинного обучения.....	32
3 Отчет по обобщению и оценке результатов исследований	35
3.1 Разработка унифицированных алгоритмов диагностирования аномалий процессов в неоднородных средах	35
3.2 Апробация прибора для диагностирования аномалий в кислородном конвертере ОАО «ММК»	40
3.3 Рекомендации по возможности использования результатов НИР в реальном секторе экономики	41
4 Публикация результатов НИР	42
Заключение.....	49
Список использованных источников	50

Введение

Данная работа является завершающим этапом в разработке прибора для диагностирования аномалий процессов в неоднородных средах. На первом этапе была разработана структура прибора диагностирования аномалий процессов в неоднородных средах, выбрана аппаратная платформа для реализации прибора, выбран тип датчиков первичной информации, выбрана программная платформа для реализации алгоритмов функционирования прибора. Это позволило собрать массив диагностических данных по вибрации корпуса кислородного конвертера.

Применение методов машинного обучения для синтеза диагностической модели обуславливает необходимость обоснованного выбора алгоритмов машинного обучения и их адаптации к задаче классификации состояния процессов в неоднородных средах.

Внедрение прибора для диагностирования аномалий процессов в неоднородных средах на предприятиях позволит снизить простои, связанные с возникновением аномалий технологических процессов, увеличить выход годной продукции.

На данном этапе планируется выполнить:

- 1) провести статистическую обработку экспериментальных данных по диагностическим признакам аномалий в ходе продувки кислородного конвертера;
- 2) проанализировать и адаптировать методы машинного обучения для диагностирования аномалий процессов в неоднородных средах;
- 3) сравнить эффективность различных методов машинного обучения в задаче классификации состояния процессов в неоднородных средах;
- 5) разработать обобщенные алгоритмы работы вычислительного блока прибора для диагностирования аномалий процессов в неоднородных средах;
- 6) разработать программное обеспечение вычислительного блока прибора для диагностирования аномалий процессов в неоднородных средах;
- 7) апробировать и внедрить прибор диагностирования аномалий процессов в неоднородных средах в технологических условиях кислородно-конвертерного агрегата.

1 Аннотированная справка по научным результатам НИР, полученным на I этапе

В соответствии с детализированным предложением о качестве поисковых научно-исследовательских работ на 2009 год по I этапу выполнены научные исследования по разработке прибора диагностирования аномалий процессов в неоднородных средах.

Проведен аналитический обзор существующих приборов для диагностирования аномалий технологических процессов в неоднородных средах, исследованы области их применения. Установлено, что основными направлениями развития диагностических систем является анализ комплекса разнородных сигналов об исследуемом объекте, применение методов машинного обучения для автоматизированного построения классификаторов ситуаций.

Проведенные патентные исследования показали, что на данный момент нет никаких серьёзных препятствий для разработки многофункционального прибора для диагностирования аномалий процессов в неоднородных средах, унифицированная методика диагностирования аномалий в неоднородных средах, применяемая в различных отраслях промышленности до настоящего времени изучена недостаточно.

Рассмотрены физико-химические предпосылки для диагностирования аномалий в технологическом процессе продувки кислородного конвертера. Косвенное диагностирование указанного процесса возможно вследствие явлений барботажа и колебания зеркала расплава.

Выбран вибрационный метод косвенной диагностики аномалий процессов в неоднородных средах.

Определены требования к прибору диагностирования аномалий процессов в неоднородных средах в кислородно-конвертерном производстве. Разрабатываемый прибор для применения в кислородно-конвертерном производстве должен обладать следующими функциональными возможностями:

- Диагностирование процесса обезуглероживания металлического расплава по ходу продувки.
- Предупреждение о возможных выбросах шлака через горловину конвертера.

- Выработка предупреждающих и аварийных сигналов.

- Сбор, отображение и хранение данных технологического процесса с возможностью передачи данных в смежные системы.

Разработана структура прибора на основе агрегатных модулей. Выбрана платформа для построения системы сбора данных. Для сбора и анализа данных по вибрации кислородного конвертера применена система на основе платформы National Instruments PXI в составе контроллера PXI-8104, платы АЦП PXI-4072, модуля согласования SCXI-1530. В качестве датчика первичной информации применен акселерометр H1C производства НПХ "Техноап".

Для реализации программного обеспечения прибора в качестве среды программирования была выбрана SCADA-система LabView фирмы National Instruments, Эта система обеспечивает тесную интеграцию с оборудованием PXI и предоставляет широкий спектр возможностей по цифровой обработке сигналов.

Собран массив диагностических данных по вибрации корпуса кислородного конвертера в процессе продувки. Создана база данных по 150 продувкам конвертера №1 ККЦ ОАО «ММК», включающая технологические параметры, оцифрованные сигналы вибрации корпуса конвертера, информацию о наблюдаемых выбросах расплава.

Проведен предварительный анализ диагностического массива данных. Спектральный анализ вибросигналов показал сложность и многофакторность процессов в расплаве металла, что подтверждает необходимость применения методов машинного обучения для диагностирования возникающих аномалий.

Составлен план дальнейших экспериментальных и теоретических исследований, обеспечивающий систематизированный подход к решению поставленных задач.

2 Аналитический отчет о проведении теоретических и экспериментальных исследований

2.1 Методика сбора информации о значениях входных и выходных параметров

Входными данными для диагностирования аномалий процессов в неоднородных средах являются временные ряды, соответствующие дискретизированным величинам сигналов, измеряемых датчиками первичной информации. Параметры сигналов, использованных для диагностирования выбросов расплава из кислородного конвертера приведены в таблице 2.1:

Таблица 2.1 - Параметры входных сигналов.

№	Название	Частота дискретизации	Единицы измерения
1	Виброускорение вдоль вертикальной оси G1	512 Гц.	м/с ²
2	Виброускорение вдоль вертикальной оси G2	512 Гц.	м/с ²
3	Положение кислородной фурмы	1 Гц.	м
4	Мгновенный расход кислорода	1 Гц.	м ³ /мин
5	Концентрация СО в отходящих газах	1 Гц.	отн. ед.
6	Концентрация СО ₂ в отходящих газах	1 Гц.	отн. ед.
7	Температура чугуна	однократно	°С
8	Масса лома	однократно	т
9	Масса чугуна	однократно	т

Аналоговые входные сигналы оцифровываются при помощи плат АЦП, входящих в состав разработанного прибора. Технологические параметры поступают от серверов АСУ-ТП по интерфейсу Ethernet.

Исходная информация о химическом составе чугуна, температуре чугуна собиралась из базы данных АСУ-ТП ОАО “ММК”.

При создании информационного массива были исключены плавки, технологическая информация по которым была внесена некорректно или отсутствовала.

Уровень шлака оценивался по трем критериям: выброс отмечался при наблюдении перелива шлакометаллической эмульсии через горловину конвертера, повышенный уровень шлака отмечался при истечении шлака через сталевыпускное отверстие, нормальный уровень шлака отмечался при отсутствии вышеперечисленных признаков.

Сигнал вибрации конвертера представляет собой реализацию нестационарного случайного процесса. Поэтому основой его анализа служит информация о динамике частотной структуры сигнала, получаемая с помощью спектрального анализа. Преобразование сигналов в частотную область позволяет оценить выраженность основных гармонических компонентов и выявить различия в спектральном составе сигнала. Получаемая при этом информация может служить основой для формирования набора признаков и построения регрессионной модели для предсказания возникновения выбросов. Для применения методов спектрального анализа исходный временной ряд разбивается на квазистационарные участки (окна) минимальной длины. Согласно исследованиям [1] период выделения пузырей монооксида углерода из вторичной зоны обезуглероживания составляет 0,2-0,3 с. Для обеспечения высокой разрешающей способности на соответствующих частотах размер окна выбран равным 8 с. Спектральная плотность мощности определяется при помощи преобразования Фурье для дискретных вещественных чисел с оконной функцией Ханнинга по формуле [2]:

$$G(k) = \sum_{n=0}^{N-1} x_n e^{-j\pi kn/N},$$

где $G(k)$ - результат преобразования; N - размер входного массива; k - номер спектральной составляющей; x - входной массив. Нулевой компонент спектра $G(0)$ представляет собой среднее значение измеренного виброположения, определяемого наклоном конвертера, и в дальнейшем не учитывается.

2.2 Статистическая обработка экспериментальных данных

В соответствии с разработанной методикой собран массив статистических данных. Фрагмент базы данных приведен в таблице 2.2.

Таблица 2.2 – Фрагмент базы данных параметров процесса продувки конвертера.

№	Время выбросов, с	Вес чугуна, т	Вес лома, т	Температура чугуна, °С	Средний расход O ₂ , мЗ/мин	S(g1)	S(g2)	S ₁₀₀ (g1)	S ₁₀₀ (g2)
1	508	296,6	110,7	1388	935,63	547,14	649,71	133,23	71,54
2	537	292,9	113,8	1340	1155,54	640,08	213,41	243,15	78,48
3	517	299,4	106,7	1401	1233,47	1091,21	339,38	220,60	82,89
4	290	311,4	93,9	1389	1357,08	1279,78	433,80	808,18	206,70
5	365	297,9	104,2	1392	1339,06	3106,18	341,29	158,61	67,85
6	445	297,7	111,2	1395	1164,70	716,67	221,78	161,27	78,16
7	162	300,1	111,5	1381	1174,17	409,33	149,58	228,08	99,73
8	562	291,4	109,6	1353	1099,40	394,59	269,66	124,96	62,35
9	412	298,1	125,5	1375	1159,60	603,43	335,61	172,27	86,47
10	114	324,2	80,3	1372	1082,84	723,84	258,39	156,23	77,51
11	643	303,7	107,4	1372	1149,84	651,23	304,86	165,95	142,39
12	355	292,9	117,1	1410	1261,39	1046,46	324,15	319,22	72,06
13	386	307,9	87,2	1368	1160,28	611,93	189,84	196,92	175,34
14	340	342,2	90,1	1382	1172,44	831,08	952,60	452,07	75,96
15	458	300,4	110,3	1379	1149,11	647,46	271,34	244,35	81,83
16	393	295,3	121,4	1367	1141,08	762,52	314,54	218,13	58,37
17	402	323,9	86,5	1415	1146,72	728,82	236,50	214,53	42,83
18	239	329,3	99,3	1394	1196,40	612,92	242,13	362,70	137,86
19	662	295,9	108,4	1376	1087,24	547,00	198,86	286,62	104,63
20	456	306,5	106	1415	1099,38	1022,06	348,08	183,66	83,17
21	458	298,8	116,5	1387	1036,90	353,98	123,12	247,49	46,65
22	362	327,5	105,1	1409	1243,94	917,48	370,09	225,96	47,70
23	326	310,5	93,7	1388	1300,42	990,57	389,21	330,12	93,06
24	345	300,5	109,4	1423	1172,64	632,25	177,86	217,21	108,64
25	263	302,1	111,4	1399	1154,02	976,46	365,42	252,16	86,32
26	183	288,5	110,8	1382	1144,62	415,47	116,77	291,80	74,57
27	241	309,9	103	1382	1144,28	1048,36	429,03	264,07	73,30
28	501	320,2	79,8	1385	1159,40	740,33	306,30	183,78	71,14
29	427	318,6	115,7	1387	1183,70	778,40	238,20	211,46	71,61
30	434	292,7	107,8	1360	1147,44	825,22	235,95	690,26	114,31
31	517	300,3	112,4	1372	1168,56	640,76	185,76	149,06	64,45
32	345	310,2	103,3	1403	1161,98	1135,79	480,59	256,98	73,62
33	415	336,4	84,2	1353	1144,24	983,16	422,42	248,86	56,73

Для оценки параметров вибрации корпуса кислородного конвертера в таблице 2.2 приведены оценки среднеквадратичного отклонения сигналов виброускорения в

течении всей продувки ($S(g1)$, $S(g2)$) и в течении 100 с от начала продувки ($S_{100}(g1)$, $S_{100}(g2)$).

Распределения частот выбросов расплава в зависимости от времени с начала продувки и суммарного расхода кислорода представлены на рисунках 2.1 и 2.2.

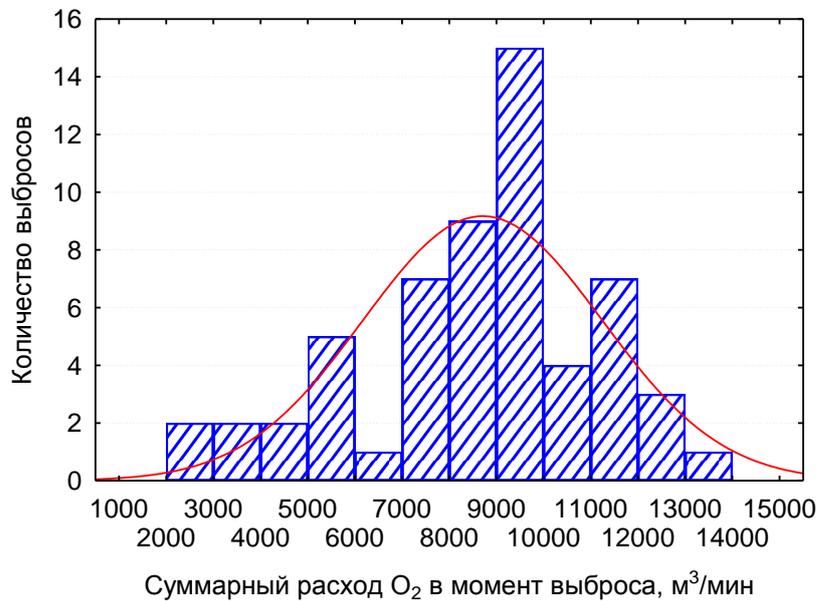


Рисунок 2.1 – Гистограмма распределения выбросов расплава по суммарному расходу кислорода.

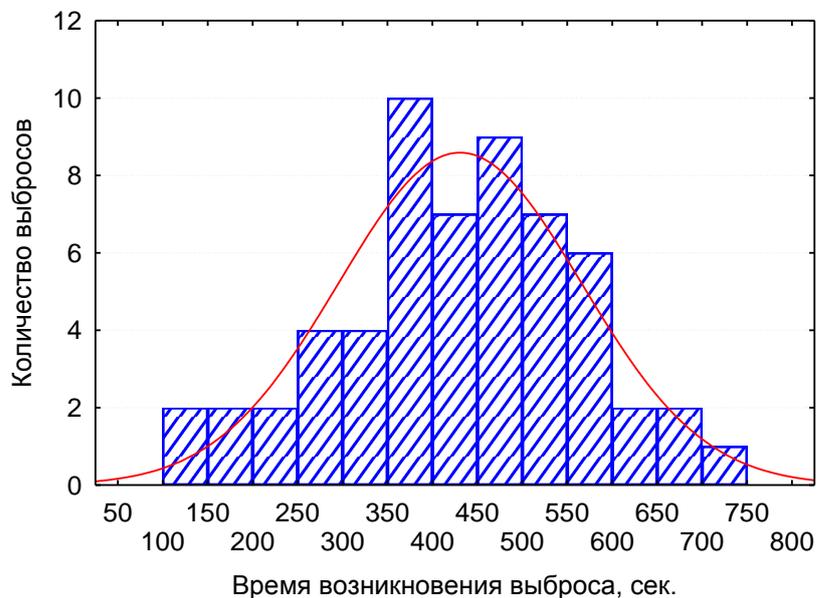


Рисунок 2.2 – Гистограмма распределения выбросов расплава по времени от начала продувки.

Оценка закона распределения величин технологических параметров P^k в момент возникновения выбросов проведена по критериям согласия. Выборочные

асимметрии \hat{A}_k и эксцессы \hat{E}_k для значений технологических параметров определены по (2.1-2.2) [3]:

$$\hat{A}_k = \frac{1}{N \cdot S_{p^k}^3} \sum_{j=1}^N (P_j^k - M(P^k))^3, \quad (2.1)$$

$$\hat{E}_k = \frac{1}{N \cdot S_{p^k}^4} \sum_{j=1}^N (P_j^k - M(P^k))^4 - 3, \quad (2.2)$$

где P_j^k – значения технологических параметров; N – объем выборки; $M(P^k)$ – математическое ожидание технологических параметров:

$$M(P_j^k) = \frac{1}{N} \cdot \sum_{j=1}^N P_j^k; \quad (2.3)$$

$S_{p^k} = \sqrt{S_{p^k}^2}$ – среднеквадратическое отклонение значений технологических параметров. Выборочная дисперсия $S_{p^k}^2$ – дисперсия рассеяния значений технологических параметров вычислена по выражению:

$$S_{p^k}^2 = \frac{1}{N-1} \cdot \sum_{j=1}^N (P_j^k - M(P^k))^2. \quad (2.4)$$

Дисперсии асимметрии и эксцесса для обучающей выборки вычислены по выражениям (2.5, 2.6):

$$DA = \frac{6(N-1)}{(N+1)(N+3)}, \quad (2.5)$$

$$DE = \frac{24N(N-2)(N-3)}{(N+1)^2(N+3)(N+5)}. \quad (2.6)$$

Результаты вычислений приведены в табл. 2.3.

Таблица 2.3 – Асимметрия и эксцесс технологических параметров в момент возникновения выбросов.

Параметр	$M(P^k)$	S_{P^k}	\hat{A}_k	\hat{E}_k	$3\sqrt{DA}$	$5\sqrt{DE}$
Суммарный расход кислорода дутья	8752,2	3216,7	-0,18	-1,22	1,26	3,52
Длительность продувки	443,5	215,4	-0,12	-0,92		
Амплитуда виброускорения вдоль вертикальной оси	0,083	0,027	0,14	-0,85		
Амплитуда виброускорения вдоль горизонтальной оси	0,024	0,008	0,11	-1,1		

Условие критериев согласия:

$$|\hat{A}_k| \leq 3\sqrt{DA} \text{ и } |\hat{E}_k| \leq 5\sqrt{DE} \quad (2.7)$$

выполняется для всех анализируемых параметров технологического процесса (таблица 2.2). Следовательно, распределения значений технологических параметров в момент возникновения выбросов подчиняются нормальному закону.

2.3 Постановка задачи автоматического синтеза диагностической модели

Диагностирование аномалий процессов в неоднородных средах связано с решением задачи классификации ситуаций по реализациям множества признаков. Классификации – формализованная задача, в которой имеется множество ситуаций, разделённых на классы [4]. Задано конечное множество объектов, для которых известны соответствующие классы – обучающая выборка $X^n = \{((x_1, y_1), \dots, (x_n, y_n))\}$, где (x_i, y_i) , $i = (1, 2, \dots, N)$ – прецеденты, вектор x_i – признаки, вектор y_i – номера классов. Требуется построить алгоритм, способный классифицировать произвольный объект из исходного множества $x \in X$, то есть аппроксимировать зависимость $y^* : X \rightarrow Y$, где X – множество описаний ситуаций, Y – множество номеров классов. Величину ошибки обучения характеризует функция потерь $L(y, f(x))$. Признаком называется отображение $f : X \rightarrow D_f$, где D_f – множество допустимых значений признака. Если заданы признаки f_1, \dots, f_n , то вектор $x = (f_1(x_1), \dots, f_n(x_n))$ называется признаковым описанием объекта $x \in X$. Множество $X = D_{f_1} \times \dots \times D_{f_n}$ называют признаковым пространством.

В зависимости от множества D_f выделяют следующие типы признаков:

- бинарный признак: $D_f = \{0, 1\}$;
- номинальный признак: D_f – конечное множество;
- порядковый признак: D_f – конечное упорядоченное множество;
- количественный признак: D_f – множество действительных чисел.

Типы классов:

- бинарные;
- множественные.

В задаче диагностирования аномалий процессов в неоднородных средах признаковое пространство формируется из реализаций временных рядов сигналов, характеризующих состояние анализируемого процесса. Поэтому признаковое пространство принадлежит множеству действительных чисел. Рассматриваются бинарные классы состояний процессов: «норма» и «аномалия».

Диагностирование аномалий процессов в неоднородных средах в разрабатываемом приборе осуществляется посредством классификации состояния процесса по параметрам входных сигналов в вычислительном блоке. Решение задачи автоматического синтеза классификатора для диагностирования аномалий процессов в неоднородных средах с применением методов машинного обучения разбивается на 5 основных этапов:

- 1) преобразование измеряемых сигналов
- 2) формирование обучающей и тестовой выборок
- 3) отбор значимых признаков
- 4) обучение классифицирующей модели
- 5) тестирование модели

При этом этапы 3-5 выполняются итеративно, с применением комбинации различных методов отбора признаков и методов обучения моделей.

2.4 Выбор значимых признаков

Решение задачи машинного обучения требует выбора ограниченного числа содержательных признаков. Отбор значимых признаков позволяет решить следующие задачи: улучшить предсказательную способность обучаемой модели, упростить структуру модели, обеспечить лучшее понимание моделируемых процессов.

Методы выбора значимых признаков делятся на следующие категории [5]:

- 1) методы фильтрации признаков;
- 2) методы полного поиска;
- 3) методы последовательного отбора (метод оберток);
- 4) методы встроенного отбора.

Методы встроенного отбора осуществляют выбор значимых признаков в процессе машинного обучения и специфичны для некоторых методов обучения.

При отборе признаков методом фильтрации осуществляется ранжирование признаков по степени их информативности до обучения классификатора. Методы ранжирования признаков:

- критерий Фишера;

- корреляционный критерий;
- тест Вальда;
- тест множителей Лагранжа;
- тест отношения правдоподобия;
- алгоритм Relief – эффективно обрабатывает зависимые признаки, недостатки:

применит только для бинарных классов, создает избыточный набор признаков при наличии значительной корреляции между ними [6].

При наличии выборки из m прецедентов $\{x_k, y_k\}$ ($k = 1, \dots, m$) состоящей из n входных переменных $x_{k,i}$ ($i = 1, \dots, n$) и одной выходной переменной y_k метод ранжирования использует функцию оценки $S(i)$, рассчитываемую по значениям $x_{k,i}$ and y_k , $k = 1, \dots, m$. Входные переменные сортируются в порядке уменьшения их значимости $S(i)$. Для обучения классификатора применяются выборки с возрастающим количеством входных переменных, начиная с наиболее значимых.

Методы итеративного отбора используют в качестве функции оценки результат тестирования алгоритма машинного обучения. Основные методы итеративного отбора признаков [7]:

- метод прямого отбора;
- метод обратного исключения;
- метод ступенчатого перебора;
- метод лучших подмножеств.

Выбор значимых признаков по методу прямого отбора заключается в поэтапном добавлении в обучающую выборку новых переменных:

Рассмотренные методы итеративного отбора значимых признаков имеют следующие недостатки:

- высокая вычислительная сложность – на каждой итерации производится обучение модели;
- подверженность достижению локальных минимумов ошибки обучения;
- не учитывается взаимодействие между признаками.

Для формирования оптимального признакового пространства при построении классифицирующей модели в вычислительном блоке прибора для диагностирования аномалий процессов в неоднородных средах разработана следующая методика:

- 1) Осуществляется ранжирование исходных признаков по степени их информативности, определяемой по корреляционному критерию и алгоритму ReliefF.
- 2) Производится обучение классификатора.
- 3) Оптимальное количество признаков определяется в ходе двухпараметрической оптимизации классификатора по критерию минимизации ошибки тестирования на обучающей и тестовой выборке.

2.5 Обзор методов машинного обучения

Из литературных источников известно множество методов машинного обучения. На каждой конкретной задаче различные методы машинного обучения показывают различные характеристики точности классификации, следовательно, необходимо выбирать оптимальный метод в каждом случае. Применение библиотеки алгоритмов машинного обучения позволяет обоснованно выбирать оптимальный метод классификации для каждой области применения разрабатываемого прибора для диагностирования аномалий процессов в неоднородных средах. Требования к библиотеке алгоритмов машинного обучения:

- совместимость по входным и выходным переменным;
- минимально необходимое количество методов;
- вычислительная эффективность каждого алгоритма;
- разнородность алгоритмов.

Для корректного сравнения различных алгоритмов машинного обучения их настройка и тестирование должно осуществляться на одних и тех же выборках данных. Практически применяемые выборки данных для диагностирования процессов в неоднородных средах, как правило, имеют большую размерность, поэтому одним из важнейших критериев при разработке вычислительного блока прибора является минимальная алгоритмическая сложность применяемых методов. Это условие обуславливает следующие требования: применение минимально необходимого количе-

ства методов машинного обучения, вычислительная эффективность каждого метода в отдельности. Применение разнородных алгоритмов машинного обучения позволяет выбрать из них наиболее эффективных для каждой специфической задачи.

Требования к алгоритмам машинного обучения для задачи классификации состояний диагностируемых процессов в неоднородных средах:

- возможность решения задачи классификации;
- оперирование с действительными входными переменными;
- аппроксимация нелинейных зависимостей.

Указанным требованиям соответствуют следующие основные методы машинного обучения:

- многослойные нейронные сети;
- радиальные базисные сети;
- деревья решений;
- машины опорных векторов.

Обоснованное применение указанных методов для решения задачи классификации состояния процессов в неоднородных средах потребовало их подробного анализа и адаптации.

2.5.1 Многослойные нейронные сети

Искусственная нейронная сеть — это математическая модель, а также ее программные или аппаратные реализации, построенная по образу сетей нервных клеток живого организма [8]. Нейронные сети выдают ответ вида:

$$y(x, w) = f\left(\sum_{j=1}^N w_j \phi_j(x)\right),$$

где f — нелинейная функция активации, w — вектор весов, ϕ — нелинейные базисные функции. Обучение нейронных сетей состоит в настройке весов и базисных функций. При обучении сети в задаче классификации предлагается множество прецедентов с указанием того, к какому классу они относятся. Прецедент представляет собой вектор значений признаков. При этом совокупность всех признаков должна

однозначно определять класс, к которому относится прецедент. В случае, если признаков недостаточно, сеть может ошибочно соотнести один и тот же образец с несколькими классами. По окончании обучения сети ей можно предъявлять неизвестные ранее образы и получать ответ о принадлежности к определённому классу. Топология такой сети характеризуется тем, что количество нейронов в выходном слое, как правило, равно количеству определяемых классов. При этом устанавливается соответствие между выходом нейронной сети и классом, который он представляет.

Этапы синтеза искусственной нейронной сети [9]:

- Сбор данных для обучения;
- Подготовка и нормализация данных;
- Выбор топологии сети;
- Экспериментальный подбор характеристик сети;
- Экспериментальный подбор параметров обучения;
- Собственно обучение;
- Проверка адекватности обучения;
- Корректировка параметров, окончательное обучение.

От правильного выбора параметров зависит скорость схождения нейронной сети. Выбор низкой скорости обучения увеличит время схождения, однако иногда позволяет избежать паралича сети. Увеличение момента обучения может привести как к увеличению, так и к уменьшению времени сходимости, в зависимости от формы поверхности ошибки. Следовательно, значения параметров нужно выбирать экспериментально, руководствуясь при этом критерием завершения обучения.

Многослойный персептрон (МП) – частный случай персептрона Розенблатта [10]. Особенностью является наличие более чем одного обучаемого слоя. Элемент многослойного персептрона полностью задается значениями своих весов и порогов, которые в совокупности определяют уравнение разделяющей прямой и скорость изменения функции при отходе от этой линии. Каждый элемент сети строит взвешенную сумму своих входов с поправкой в виде слагаемого и затем пропускает эту величину активации через передаточную функцию, и таким образом получается выходное значение этого элемента. Обучение проводится до стабилизации весовых ко-

эффиентов при обучении или прерывается ранее, чтобы избежать переобучения. Такая сеть может моделировать функцию практически любой степени сложности, причем число слоев и число элементов в каждом слое определяют сложность функции. Определение числа промежуточных слоев и числа элементов в них является важным вопросом при конструировании многослойного персептрона. Количество входных и выходных элементов определяется условиями задачи.

В качестве функций активации применяются следующие виды сигмоидальных функций [9]:

- экспоненциальная сигмоида: $f(s) = \frac{1}{1 + e^{-2cs}}$;
- рациональная сигмоида: $f(s) = \frac{s}{|s| + a}$;
- гиперболический тангенс: $f(s) = th \frac{s}{\alpha}$,

где s – выход сумматора нейрона, α – произвольная константа.

Наименьшими вычислительными затратами характеризуется расчет рациональной сигмоидальной функции. Поэтому в качестве функции активации выбираем рациональную сигмоидальную функцию.

Для обучения нейронной сети применен алгоритм обратного распространения. Основная идея этого метода состоит в распространении сигналов ошибки от выходов сети к её входам, в направлении, обратном прямому распространению сигналов в обычном режиме работы. Для возможности применения метода обратного распространения ошибки передаточная функция нейронов должна быть дифференцируема [10]. Блок-схема алгоритма обратного распространения приведена на рисунке 2.3.

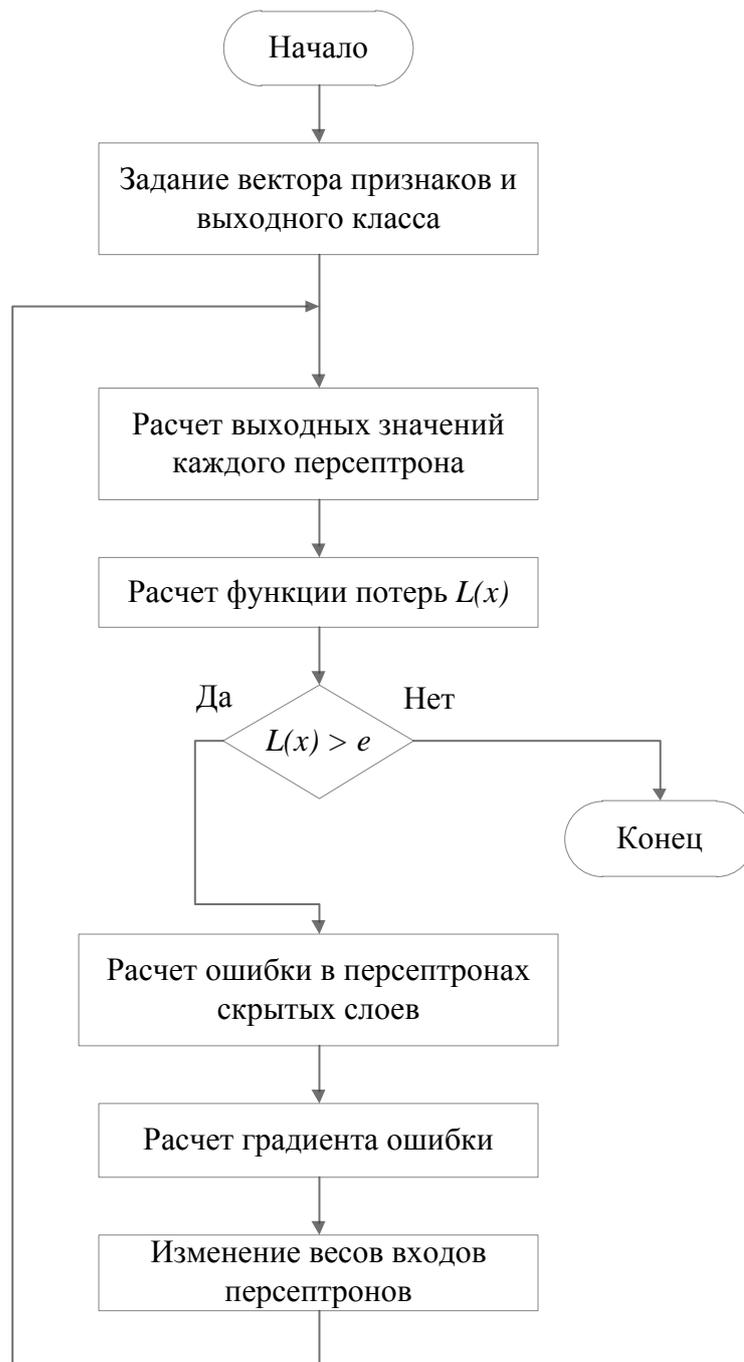


Рисунок 2.3 – Блок-схема алгоритма обратного распространения

Действие алгоритма заключается в следующем: блок обработки данных на некотором слое сети получает сигналы, поступающие с нижележащих слоев, эти сигналы обрабатываются по определенным правилам и передаются в блок обработки данных на более высокий уровень системы. Затем значения весов обновляются от самого верхнего к самому нижнему слою. Далее задаются новые значения весов. Конечная цель состоит в нахождении минимума функции ошибки. Процесс носит характер последовательного приближения к оптимальным значениям весов элемен-

тов при обратном движении от выходного слоя; направление распространения ошибок противоположно направлению поступления сигнала, веса изменяются через повторяющееся обучение на множестве фактических данных.

Недостатки метода обратного распространения ошибки:

- паралич сети,
- остановка обучения в локальных минимумах.

В процессе обучения сети значения весов могут достичь величины значительно большей 1. Это приводит к тому, что нейроны будут функционировать в области, где производная функции активации очень мала. Так как посылаемая обратно в процессе обучения ошибка пропорциональна этой производной, то процесс обучения может практически замереть, что и называется параличом сети. Избежать этого явления можно выбором низкой скорости обучения.

2.5.2 Радиальные базисные сети

Сети, использующие радиально-базисные функции (РБФ-сети), являются частным случаем двухслойной сети прямого распространения. В отличие от многослойного персептрона с его линейным подходом в сети на радиально-базисных функциях используется радиальная функция. Общий вид радиально-базисной функции:

$f(x) = \phi\left(\frac{x^2}{\sigma^2}\right)$, где x – вектор входных сигналов нейрона, σ – ширина окна функции, $\phi(y)$ – монотонно убывающая функция. Наиболее часто применяемой радиально-базисной функцией является: $f(x) = e^{-\frac{x^2}{\sigma^2}}$ [11].

Радиально-базисная сеть характеризуется тремя особенностями:

1. Единственный скрытый слой.
2. Только нейроны скрытого слоя имеют нелинейную активационную функцию.
3. Веса связей входного и скрытого слоев равны единице.

Сеть типа радиальной базисной функции имеет промежуточный слой из радиальных элементов, каждый из которых воспроизводит гауссову поверхность отклика. Поверхность отклика радиального элемента представляет собой гауссову функ-

цию (колоколообразной формы) с вершиной в центре и понижением к краям. Наклон гауссова радиального элемента можно менять подобно тому, как можно менять наклон сигмоидальной кривой в МП. Поскольку эти функции нелинейны, для моделирования произвольной функции нет необходимости брать более одного промежуточного слоя. Для моделирования любой функции необходимо лишь взять достаточное число радиальных элементов.

Сети RBF имеют ряд преимуществ перед сетями МП. Во-первых, они моделируют произвольную нелинейную функцию с помощью всего одного промежуточного слоя и тем самым обходят вопрос о числе слоев. Во-вторых, параметры линейной комбинации в выходном слое можно полностью оптимизировать с помощью хорошо известных методов линейного моделирования, которые работают быстро и не испытывают трудностей с локальными минимумами, так мешающими при обучении МП. Поэтому сеть РБФ обучается очень быстро, на порядок быстрее МП.

2.5.3 Деревья решений

Одним из эффективных методов машинного обучения является метод решающих деревьев (decision trees). Для задачи классификации подходит алгоритм построения решающего дерева «С4.5». Получаемая модель представляет собой направленный граф правил вида “если... то...”. В узлах графа производится сравнение входной переменной с константой, и в зависимости от результата производится переход к следующему узлу, либо выдается результат при достижении листьев графа. Построение модели производится итеративно путем выбора переменной и граничного уровня, максимизирующих относительное приращение информации $GR(S,A)$ [12]:

$$GR(S,A) = \frac{Gain(S,A)}{SI(S,A)}, \quad (2.8)$$

где $Gain(S,A)$ - приращение информации:

$$Gain(S,A) = E(S) - \sum_{v \in Values(A)} \frac{|S_v|}{|S|} E(S_v), \quad (2.9)$$

где $E(S)$ - энтропия, характеризующая непротиворечивость обучающей выборки S :

$$E(S) = - \sum_{i=1}^c p_i \log_2 p_i, \text{ где } c - \text{ множество возможных значений атрибута, } p_i - \text{ пропорция}$$

прецедентов в которых атрибут A имеет i -е значение, $\text{Values}(A)$ – множество возможных значений атрибута A , S_v - подмножество S , в котором $A = v$,

$$SI(S,A) = - \sum_{i=1}^c \frac{|S_i|}{S} \log_2 \frac{|S_i|}{S}.$$

На каждой итерации выбирается атрибут, максимизирующий выражение (2.8). Этот атрибут будет являться проверкой в текущем узле дерева, а затем по этому атрибуту производится дальнейшее построение дерева. Т.е. в узле будет проверяться значение по этому атрибуту и дальнейшее движение по дереву будет производиться в зависимости от полученного ответа. Если в процессе работы алгоритма получен узел, ассоциированный с пустым множеством прецедентов, то он помечается как лист, и в качестве решения листа выбирается наиболее часто встречающийся класс у непосредственного предка данного листа.

2.5.4 Машины опорных векторов

Машина опорных векторов – алгоритм машинного обучения, предложенный В. Н. Вапником. Этот метод заключается в переносе данных в пространство высокой размерности и поиске разделяющих гиперплоскостей, лежащей максимально удаленно от пограничных точек. Парадигмой машины опорных векторов можно считать выбор наиболее близких к границе классов объектов из обучающей выборки, по которым строится опорная гиперплоскость.

При наличии выборки из m прецедентов $\{x_k, y_k\}$ ($k = 1, \dots, m$), состоящей из n входных переменных $x_{k,i}$ ($i = 1, \dots, n$) и одной выходной переменной y_k , по методу SVM решается квадратичная оптимизационная задача [13]:

$$\min_{w,b,\xi} \frac{1}{2} w^T w + C \sum_{k=1}^m \xi_k, \text{ при условии что: } y_k (w^T \phi(x_k) + b) \geq 1 - \xi_k, \xi_k \geq 0, k = 1, \dots, m,$$

где обучающие данные преобразуются в пространство высокой размерности функцией $\phi(x_k)$, C – параметр штрафа на величину ошибки обучения. Для каждого тесто-

вого прецедента x_k классифицирующая функция будет иметь вид:

$f(x) = \text{sgn}(w^T \phi(x) + b)$. Для обучения SVM функция $\phi(x)$ может быть заменена на функцию ядра $k(x, x') = \phi(x)^T \phi(x')$. Наиболее перспективной функцией ядра является RBF-функция [10]:

$$k(x, x') = \exp(-\gamma \|x - x'\|^2).$$

Выбор параметров C и γ осуществляется экспериментально путем перекрестной проверки модели при варьировании параметров в заданных пределах.

2.6 Методы оценки качества машинного обучения

Качество машинного обучения оценивается по величине эмпирического риска. Эмпирический риск представляет собой среднюю ошибку алгоритма на обучающей выборке. Метод минимизации эмпирического риска наиболее часто применяется для построения алгоритмов обучения. Он состоит в том, чтобы в рамках заданной модели выбрать алгоритм, имеющий минимальное значение средней ошибки на заданной обучающей выборке. Средняя потеря алгоритма a на выборке $X^m = \{x_1, \dots, x_m\}$ [14]:

$$Q(a, X^m) = \frac{1}{m} \sum_{i=1}^m L(a, x_i), \quad (2.10)$$

где $L(a, x_i)$ - функция потерь.

В ходе обучения модели возможно возникновение такого негативного явления, как переобучение. Переобучение — это явление, при котором обучающий алгоритм выдает хорошие результаты на обучающей выборке, но имеет плохие обобщающие свойства. Величина переобученности алгоритма $a = \mu(X^m)$ относительно контрольной выборки X^k рассчитывается по формуле [12]:

$$\delta(a, X^m, X^k) = Q(a, X^k) - Q(a, X^m). \quad (2.11)$$

Как правило, усложнение модели приводит к уменьшению ошибки на обучающей выборке, на тестовой же выборке ошибка сначала уменьшается, затем возрастает, что является следствием переобучения модели. Основные причины переобучения – аппроксимация моделью шума в обучающей выборке, избыточность пространства

признаков, малое число прецедентов в обучающей выборке. Поэтому, для нахождения оптимальной классифицирующей модели необходимо проводить совместный анализ результатов множества моделей как на обучающей выборке, так и на тестовой.

2.6.1 Скользящий контроль

Скользящий контроль – процедура эмпирического оценивания обобщающей способности алгоритмов, обучаемых по прецедентам. Скользящий контроль является стандартной методикой тестирования и сравнения алгоритмов классификации, регрессии и прогнозирования. Если выборка независима, то средняя ошибка скользящего контроля даёт несмещённую оценку вероятности ошибки. Это выгодно отличает её от средней ошибки на обучающей выборке, которая может оказаться смещённой (оптимистически заниженной) оценкой вероятности ошибки, что связано с явлением переобучения [14].

Фиксируется некоторое множество разбиений исходной выборки X^m на две подвыборки: обучающую и контрольную. Для каждого разбиения выполняется настройка алгоритма по обучающей подвыборке, затем оценивается его средняя ошибка $Q(a, X^m)$ на объектах контрольной подвыборки. Оценкой скользящего контроля называется средняя по всем разбиениям величина ошибки на контрольных подвыборках.

Общий алгоритм скользящего контроля:

1) Исходная выборка X^L разбивается N различными способами на две непересекающиеся подвыборки: $X^L = X_n^m \cup X_n^k$, где X_n^m - обучающая подвыборка длины m , X_n^k - контрольная подвыборка длины $k = L - m$, $n = 1, \dots, N$.

2) Для каждого разбиения n строится алгоритм $a_n = \mu(X_n^m)$ и вычисляется значение функционала качества $Q_n = Q(a_n, X_n^k)$. Оценка скользящего контроля:

$$CV(\mu, X^L) = \frac{1}{N} \sum_{n=1}^N Q(\mu(X_n^m), X_n^k). \quad (2.12)$$

Доверительное оценивание среднего качества обучения может осуществляться параметрическими и не параметрическими методами. Параметрические оценки доверительного интервала основаны на априорном предположении о виде распределения случайной величины $Q(\mu(X_n^m), X_n^k)$. Если априорные предположения не выполняются, доверительный интервал может оказаться сильно смещённым. В общем случае в результате обучения по случайным подвыборкам X^m величина $Q(a(X^m), X^k)$ описывается неизвестной смесью биномиальных распределений. Аппроксимация смеси одним биномиальным распределением может приводить к ошибочному сужению доверительного интервала. Этому недостатка лишены непараметрические оценки доверительного интервала [15].

Непараметрическая оценка доверительного интервала осуществляется на основе вариационного ряда значений $Q_n = Q(a_n, X_n^k)$, $n = 1, \dots, N$:

$$Q^{(1)} \leq Q^{(2)} \leq \dots \leq Q^{(N)}. \quad (2.13)$$

Вероятность того, что значение случайной величины $Q(a(X^m), X^k)$ не выходит за границы вариационного ряда $Q^{(1)}, Q^{(N)}$ равна $\eta = \frac{2}{N+1}$ []. Таким образом, для получения двусторонней оценки с надёжностью 95% достаточно взять $N = 40$ разбиений.

Возможны различные варианты скользящего контроля, отличающиеся способами разбиения выборки. Наиболее эффективные:

- полный скользящий контроль;
- случайные разбиения;
- контроль по q блокам;
- контроль по r×q блокам.

В случае полного скользящего контроля оцениваются все $N = C_L^k$ разбиений. Даже при небольших значениях k число разбиений становится слишком большим, что затрудняет практическое применение данного метода. Способ случайных разбиений заключается в выборе разбиений случайно, независимо и равновероятно из множества всех N разбиений.

Методика контроля по q блокам (q-fold CV) [15]:

1) Выборка случайным образом разбивается на q непересекающихся блоков одинаковой длины k_1, \dots, k_q :

$$X^L = X_1^{k_1} \cup \dots \cup X_q^{k_q}, k_1 + \dots + k_q = L. \quad (2.14)$$

2) Каждый блок по очереди становится контрольной подвыборкой, при этом обучение производится по остальным $q-1$ блокам.

3) Оценка скользящего контроля определяется как средняя ошибка на контрольной подвыборке:

$$CV(\mu, X^L) = \frac{1}{q} \sum_{n=1}^q Q(\mu(X^L \setminus X_n^{k_n}), X_n^{k_n}). \quad (2.15)$$

Описанный метод является компромиссным между полным скользящим контролем и случайными разбиениями.

Методика контроль по $r \times q$ блокам:

1) Контроль по q блокам повторяется r раз.

2) Каждый раз выборка случайным образом разбивается на q непересекающихся блоков.

3) Рассчитывается средняя ошибка на контрольной выборке

Этот способ наследует все преимущества q -fold CV, при этом появляется дополнительная возможность увеличивать число разбиений. Данный вариант скользящего контроля, со стратификацией классов, является стандартной методикой тестирования и сравнения алгоритмов классификации.

Недостатки метода скользящего контроля:

1) Машинное обучение осуществляется N раз, что сопряжено со значительными вычислительными затратами.

3) Использование скользящего контроля при обучении в роли оптимизируемого критерия приводит к тому, что он утрачивает свойство несмещённости, и возникает риск переобучения.

4) Скользящий контроль дает несмещенную точечную, но не интервальную оценку риска.

2.6.2 ROC-кривая

ROC-кривая (Receiver Operator Characteristic) используется для представления результатов бинарной классификации в машинном обучении. Один из классов считается классом с положительными исходами, второй – с отрицательными исходами. ROC-кривая показывает зависимость количества верно классифицированных положительных примеров от количества неверно классифицированных отрицательных примеров. В терминологии ROC-анализа первые называются истинно положительным, вторые – ложно отрицательным множеством. Для построения указанной кривой классификатор должен иметь параметр, варьируя который, можно получать различные разбиения на два класса. Этот параметр называют порогом. В зависимости от него будут получаться различные величины ошибок I и II рода [16]:

- TP (True Positives) – верно классифицированные положительные примеры (так называемые истинно положительные случаи);
- TN (True Negatives) – верно классифицированные отрицательные примеры (истинно отрицательные случаи);
- FN (False Negatives) – положительные примеры, классифицированные как отрицательные (ошибка I рода).
- FP (False Positives) – отрицательные примеры, классифицированные как положительные (ошибка II рода).

Объективная ценность любого бинарного классификатора характеризуется критериями чувствительности и специфичности. Чувствительность определяется как доля истинно положительных случаев:

$$Se = \frac{TP}{TP + FN} 100\% . \quad (2.16)$$

Специфичность определяется как доля истинно отрицательных случаев, которые были правильно идентифицированы моделью:

$$Sp = \frac{TN}{TN + FP} 100\% . \quad (2.17)$$

В задаче диагностирования аномалий процессов в неоднородных средах указанные параметры имеют следующий смысл:

- Чувствительный диагностический тест проявляется в гипердиагностике – максимальном предотвращении пропуска аномальных ситуаций;

- Специфичный диагностический тест диагностирует только доподлинно аномальные ситуации.

ROC-кривая строится следующим образом:

1. Для каждого значения порога отсечения, которое меняется от 0 до 1 с шагом dx рассчитываются значения чувствительности Se и специфичности Sp .

2. Строится график зависимости: по оси Y откладывается чувствительность Se , по оси X : $100\% - Sp$.

Для идеального классификатора график ROC-кривой проходит через верхний левый угол, где доля истинно положительных случаев составляет 100% или 1.0 (идеальная чувствительность), а доля ложно положительных примеров равна нулю. Поэтому чем ближе кривая к верхнему левому углу, тем выше предсказательная способность модели. Наоборот, чем меньше изгиб кривой и чем ближе она расположена к диагональной прямой, тем менее эффективна модель. Диагональная линия соответствует полной неразличимости двух классов.

Для сравнения ROC-кривых рассчитывается оценка площади под кривыми (AUC). Эта оценка может быть получена непосредственно вычислением площади под фигуры, ограниченной справа и снизу осями координат и слева вверху – точками ROC-кривой. По методу трапеций [17]:

$$AUC = \int f(x)dx = \sum_i \left[\frac{x_{i+1} + x_i}{2} \right] (Y_{i+1} - Y_i). \quad (2.18)$$

Чем больше показатель AUC, тем лучшей прогностической способностью обладает модель. В литературе приводится экспертная шкала для значений AUC, по которой можно судить о качестве модели, приведенная в таблице 2.4 [16]:

Таблица 2.4 – Шкала для значений AUC.

Значение AUC	Качество модели
0.9 - 1.0	Отличное
0.8 - 0.9	Очень хорошее
0.7 - 0.8	Хорошее
0.6 - 0.7	Среднее
0.5 - 0.6	Неудовлетворительное

Компромисс между чувствительностью и специфичностью находится с помощью порога отсечения. Для определения оптимального порога выбран критерий баланса между чувствительностью и специфичностью $Se \approx Sp$. В этом случае порог есть точка пересечения двух кривых, когда по оси X откладывается порог отсечения, а по оси Y –чувствительность или специфичность модели.

2.7 Экспериментальное сравнение методов машинного обучения

Проведено экспериментальное сравнение описанных методов машинного обучения для диагностирования отклонений в процессе продувки кислородного конвертера, приводящих к выбросам расплава. Обучающая и тестовая выборки формировались из накопленного массива статистических данных измерения вибрации корпуса кислородного конвертера в ККЦ ОАО «ММК».

Схема эксперимента приведена в таблице 2.5.

Таблица 2.5 – Схема экспериментального сравнения методов машинного обучения.

Метод оценки качества обучения	Метод обучения			
	Нейронные сети	Радиально-базисные сети	Машины опорных векторов	Деревья решений
Скользкий контроль по 10 блокам	-	-	+	+
Отложенный скользящий контроль	+	+	-	-
ROC-кривая	-	-	+	+

Обучение нейронных сетей вида многослойный персептрон требует предварительного выбора следующих параметров: количество входных нейронов, количество выходных нейронов, количество скрытых слоев, количество нейронов в скрытых слоях, вид функции активации для каждого слоя. Количество входных нейронов k выбирается равным количеству входных переменных. Используется один выходной нейрон, что соответствует одному диагностируемому классу состояний процесса в неоднородной среде. Анализ литературных источников показал, что в большинстве случаев оптимальной оказывается нейронная сеть с количеством скрытых слоев от одного до двух, количеством нейронов в этих слоях в два раза меньшим, чем во входном слое. Выбран класс нейронных сетей с двумя скрытыми слоями и числом нейронов $m = \frac{k}{2}$.

Для применения алгоритма SVM для диагностирования аномалий процессов в неоднородных средах разработана следующая методика:

1) Выбирается пара значений (C, γ) на сетке в диапазоне $\log_2 C \in (-5, -3, \dots, 10)$, $\log_2 \gamma \in (-10, -7, \dots, 2)$. Применение логарифмической шкалы значений параметров C и γ позволяет протестировать больший диапазон значений при меньшем количестве итераций.

2) Производится обучение SVM и перекрестная проверка с числом разбиений 5 на обучающей выборке.

3) Сохраняется модель SMV с параметрами (C, γ) , обеспечивающими минимальную ошибку тестирования.

Расчет оценки скользящего контроля производился по формуле:

$$CV(\mu, X^L) = \frac{1}{N} \sum_{n=1}^N Q(\mu(X_n^m), X_n^k), \quad (2.19)$$

где $Q(\mu(X_n^m), X_n^k)$ - функционал качества:

$$Q(\mu(X_n^m), X_n^k) = 1 - |\mu(X_n^m) - X_n^k|.$$

Результаты обучения и тестирования алгоритмов диагностирования, построенных в соответствии с рассмотренными методами машинного обучения приведены в таблице 2.6.

Таблица 2.6 – Результаты тестирования алгоритмов диагностирования.

Оценка качества обучения	Метод обучения			
	Нейронные сети	Радиально-базисные сети	Машины опорных векторов	Деревья решений
$CV(\mu, X^L)$	0.67	0.51	0.53	0.82
AUC	-	-	0.67	0.89

Выводы:

1) Наиболее эффективным методом машинного обучения в задаче диагностирования выбросов расплава из кислородного конвертера является метод решающих деревьев (алгоритм C4.5).

2) Применение комплекса методов машинного обучения для синтеза классифицирующей модели позволяет адаптировать прибор для диагностирования аномалий процессов в неоднородных средах к широкому классу задач.

3 Отчет по обобщению и оценке результатов исследований

3.1 Разработка унифицированных алгоритмов диагностирования аномалий процессов в неоднородных средах

Структурная схема вычислительного блока прибора для диагностирования аномалий процессов в неоднородных средах представлена на рисунке 3.1.

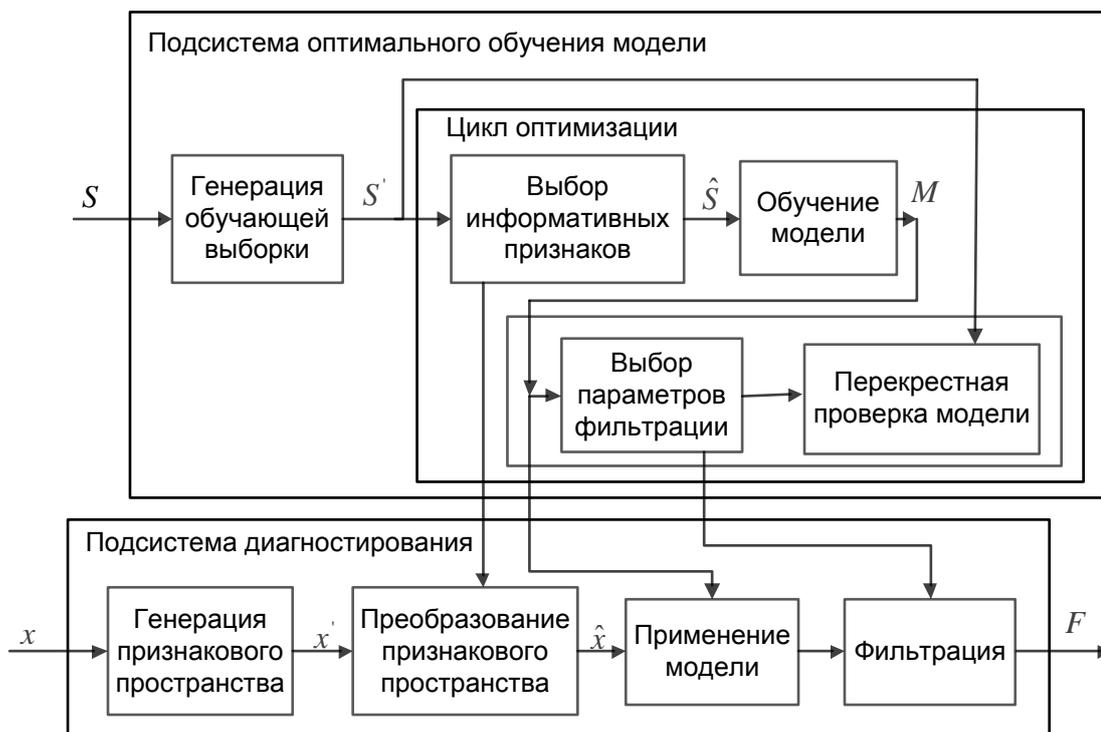


Рисунок 3.1 – Структурная схема вычислительного блока прибора для диагностирования аномалий процессов в неоднородных средах.

В режиме синтеза параметров системы диагностирования осуществляется генерация эффективного признакового пространства на основе обучающей выборки S . В качестве обучающей выборки S выступает множество диагностических сигналов (СПМ вибрации, высота фурмы, расход кислорода), сопоставленное с диагностируемой величиной. Для уменьшения размерности задачи обучения осуществляется выбор репрезентативных признаков. Последовательно выбираются различные алгоритмы машинного обучения. На созданной обучающей выборке осуществляется обучение классификатора. Параметры генерации признаков и модель классификации сохраняются и передаются в блок диагностирования, который на основе непрерывно поступающих данных рассчитывает диагностируемую величину.

В соответствии с указанной структурной схемой разработаны алгоритмы работы вычислительного блока прибора для диагностирования аномалий процессов в неоднородных средах (рисунок 3.2).

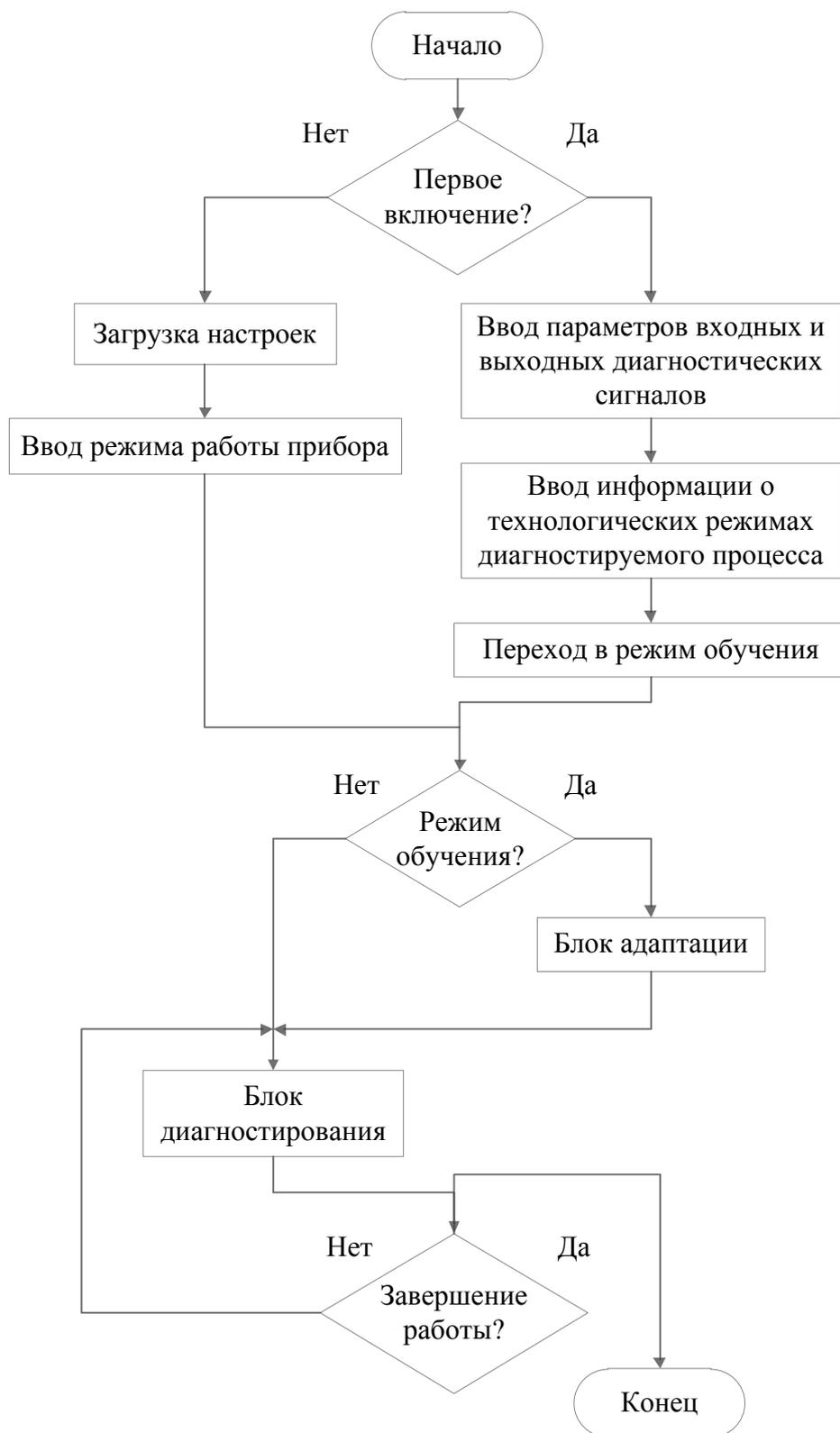


Рисунок 3.2 – Блок-схема общего алгоритма работы прибора для диагностирования аномалий процессов в неоднородных средах.

При первом включении прибора оператором задаются настройки. В блоке ввода параметров входных и выходных сигналов задаются параметры, приведенные в таблице 3.1. После завершения настройки прибор переходит в режим обучения.

Таблица 3.1 – Настраиваемые параметры входных переменных.

№	Параметр	Возможные значения
1	Тип переменной	скалярный; векторный
2	Частота опроса	1Гц...100кГц
3	Источник сигнала	АЦП; База данных
4	Номер канала АЦП	1...16
5	IP-адрес сервера баз данных	0.0.0.0 ... 255.255.255.255
6	Имя базы данных	текстовое значение
7	Имя таблицы	текстовое значение
8	Имя столбца данных	текстовое значение

Блок-схема блока адаптации приведена на рисунке 3.3. После поступления сигнала начала диагностируемого технологического процесса производится накопление массива статистических данных для формирования обучающей выборки. Класс текущего состояния процесса может быть введен оператором либо в режиме реального времени, либо в режиме редактирования архивных данных в произвольное время. При достижении заданного объема обучающей выборки производится обучения модели классификатора состояния процесса. По результатам тестирования полученной модели определяется качество обучения. Если ошибка обучения меньше заданной величины, диагностическая модель сохраняется, иначе – продолжается накопление статистических данных.

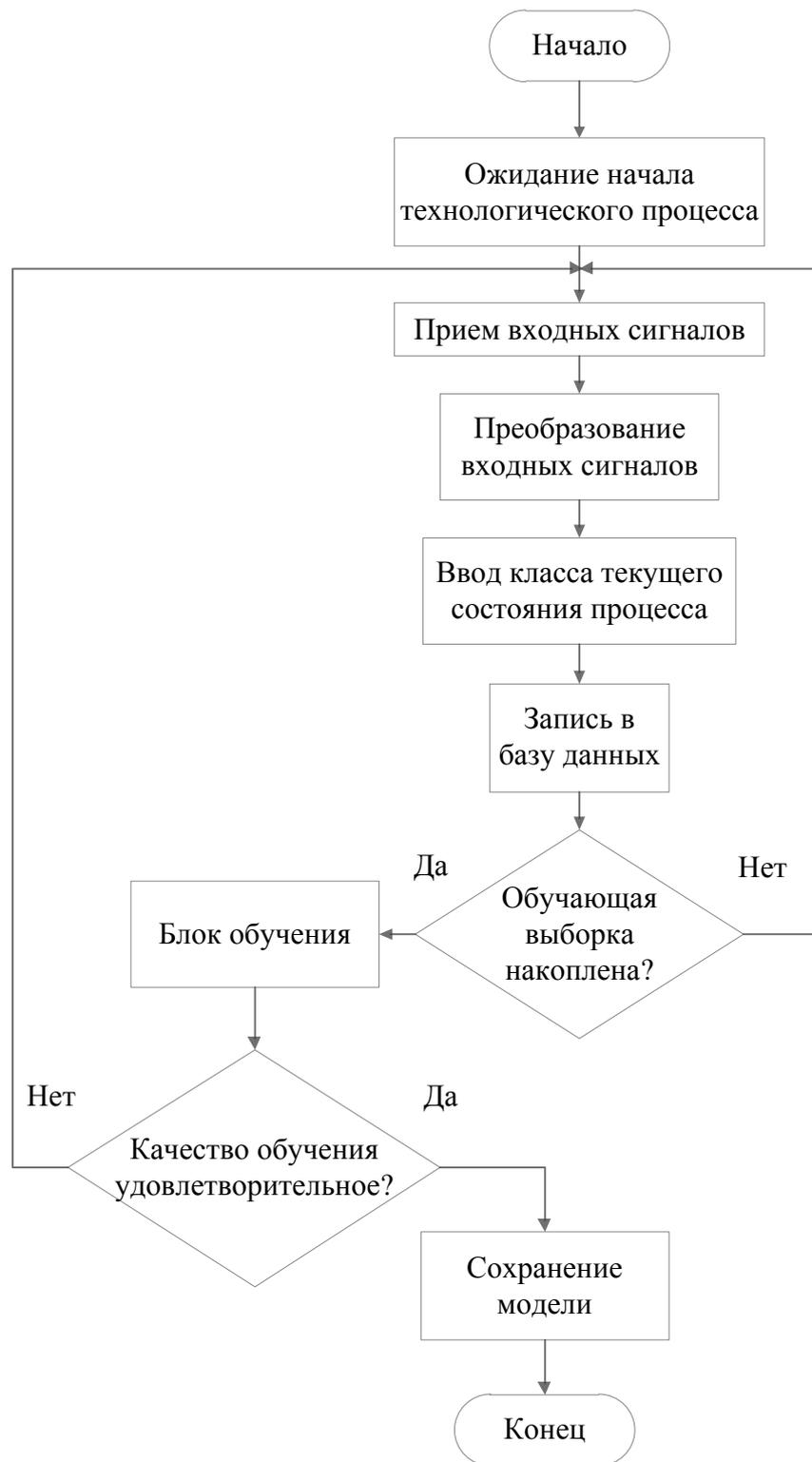


Рисунок 3.3 – Блок-схема алгоритма работы блока адаптации.

Блок-схема блока диагностирования приведена на рисунке 3.4. После начала технологического процесса производится прием входных сигналов и их преобразование в вектор диагностических признаков.

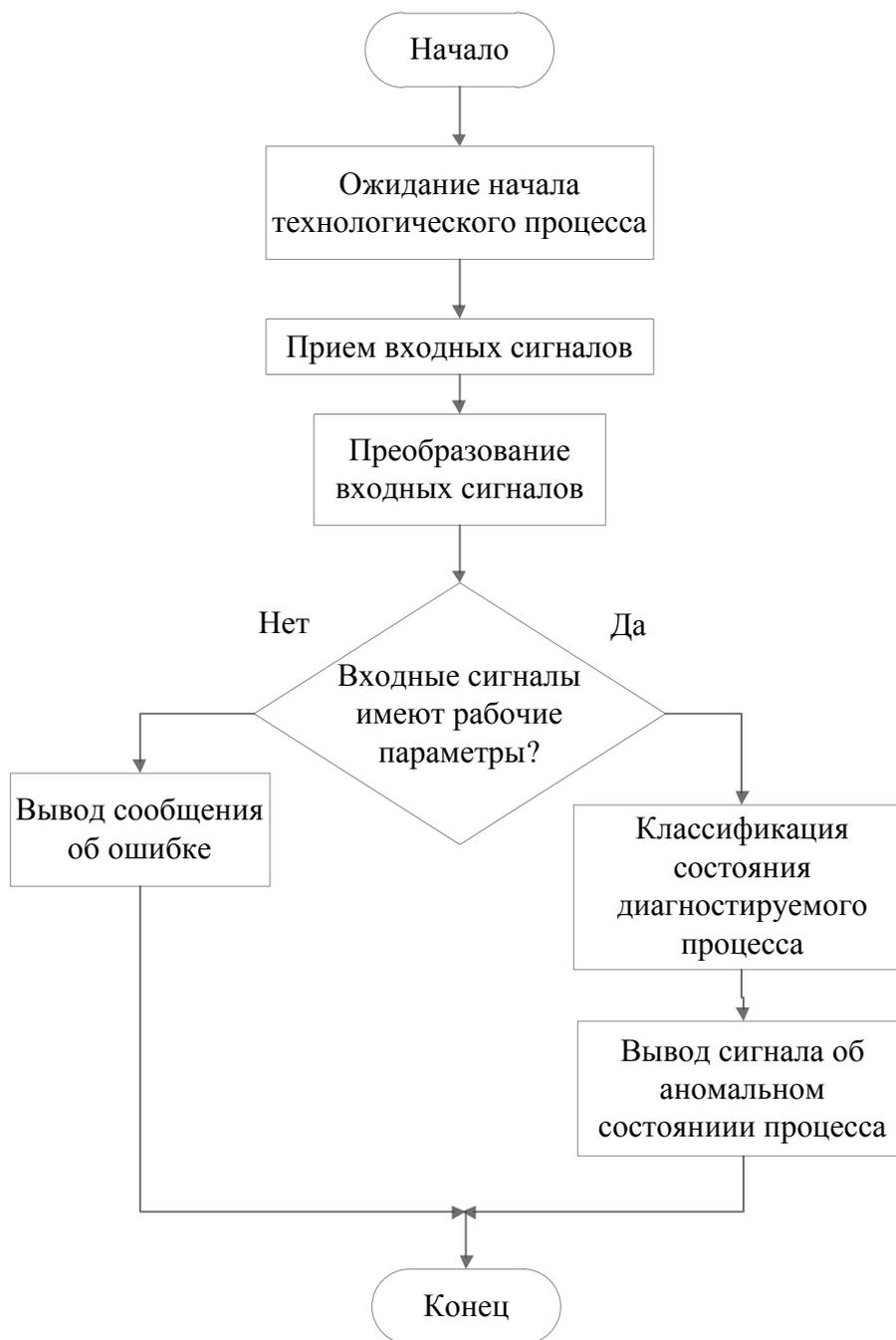


Рисунок 3.4 – Блок-схема алгоритма работы блока диагностирования.

Производится оценивание работоспособности датчиков первичной информации и каналов передачи данных по параметрам полученных входных сигналов. В случае выявленных нарушений оператору прибора выводится сообщение об ошибке. Если параметры сигналов признаны удовлетворительными, производится применение диагностической модели к вектору признаков. Полученный класс состояния диагностируемого процесса выводится на дисплей и выходной порт прибора.

3.2 Апробация прибора для диагностирования аномалий в кислородном конвертере ОАО «ММК»

Разработанный прибор для диагностирования аномалий процессов в неоднородных средах был применен для раннего предупреждения выбросов расплава через горловину конвертера 1 в ККЦ ОАО «ММК» (рисунок 3.5).

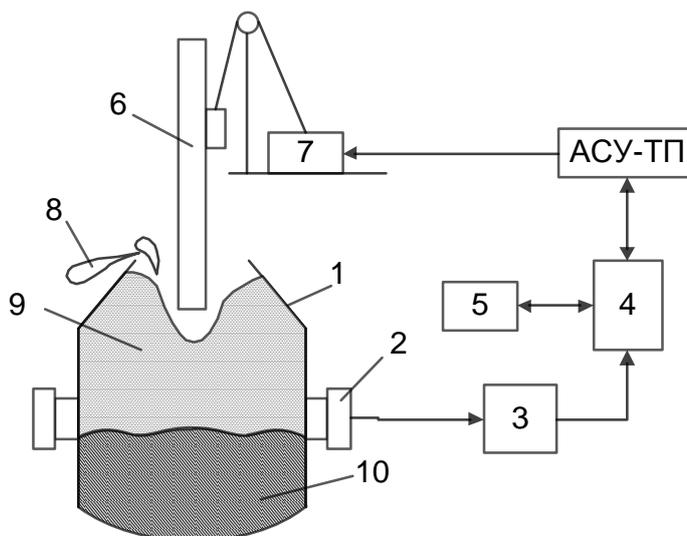


Рисунок 3.5 – Схема установки прибора для диагностирования аномалий процессов в неоднородных средах в ККЦ ОАО «ММК»:

1 – корпус конвертера; 2 – цапфа; 3 – вибродатчики; 4 – прибор для диагностирования аномалий процессов в неоднородных средах; 5 – пульт оператора; 6 – кислородная фурма; 7 – электропривод фурмы; 8 – выброс расплава; 9 – шлак; 10 – жидкий металл.

В качестве датчиков первичной информации применены вибродатчики 2, установленные на цапфе конвертера 3. При помощи вибродатчиков измеряется виброускорение корпуса конвертера вдоль трех ортогональных осей. Сигналы передаются в прибор для диагностирования аномалий процессов в неоднородных средах, где оцифровываются с частотой дискретизации 512 Гц. Кроме того, в прибор в качестве входных переменных поступают данные от АСУ-ТП «Плавка» о технологических параметрах (расходе кислорода, положения фурмы, температуры и массы чугуна и

пр.) При диагностировании возможности возникновения выбросов расплава оператору продувки 5 выдается визуальный сигнал.

В режиме обучения были обработаны данные 260 продувок конвертера №1. Применение прибора показало его высокую эффективность (таблица 3.2).

Таблица 3.2 – Результаты тестирования прибора для диагностирования аномалий процессов в неоднородных средах.

Диагностированы	Выбросы наблюдались	Выбросы не наблюдались
выбросы	32	4
отсутствие выбросов	7	217

Возникновение выбросов диагностировалось за время от 24 до 9 с до визуального наблюдения. Таким образом, разработанный прибор позволяет эффективно диагностировать тенденции к возникновению выбросов расплава из кислородного конвертера.

3.3 Рекомендации по возможности использования результатов НИР в реальном секторе экономики

Внедрение прибора для диагностирования аномалий процессов в неоднородных средах на производственных предприятиях позволит предотвратить аварийные ситуации, связанные с возникновением аномалий технологических процессов, увеличить выход годной продукции. Применительно к кислородно-конверторному производству внедрение указанного прибора позволит снизить простои, связанные с выбросами расплава из горловины, увеличить выход годной стали, повысит качество выпускаемой продукции за счет поддержания нормального шлакового режима плавки.

4 Публикация результатов НИР

Основные результаты выполненной научно-исследовательской НИР опубликованы в следующих статьях:

№	Название статьи	Сборник	Авторы
1	Разработка системы диагностирования выбросов расплава из кислородного конвертера	Сборник трудов VI международной (XVII всероссийской) конференции по автоматизированному электроприводу. – Тула: ГОУ ВПО «Тульский государственный университет», 2010 г.	Лукьянов С.И., Суспицын Е.С., Мещеряков А.Ю.

Экспертное заключение о возможности опубликования в печати

на рукопись С.И. Лукьянова, Е.С. Суспицына, А.Ю. Мещерякова

«РАЗРАБОТКА СИСТЕМЫ ДИАГНОСТИРОВАНИЯ ВЫБРОСОВ РАСПЛАВА ИЗ КИСЛОРОДНОГО КОНВЕРТЕРА»

Одной из важнейших задач в кислородно-конвертерном производстве является контроль за ходом процесса продувки. Основными динамическими показателями протекания процесса продувки являются скорость обезуглероживания расплава и уровень шлака. Повышение уровня шлака может повлечь его выброс из горловины конвертера, что приводит к простоею производства и ухудшению качества выпускаемой продукции.

В настоящее время широкое распространение получили системы диагностирования хода продувки при помощи вспомогательной фурмы, измерения уровня шума, измерения состава отходящих газов и пр. Системы, применяемые на ККЦ ОАО «ММК» не удовлетворяют требованиям современного металлургического производства.

Авторами были подробно изучены технология конвертерного производства стали, конструктивные особенности конвертера и сопутствующих систем. Проведены работы по сбору и анализу экспериментальных данных.

Авторами были разработаны алгоритмы диагностирования выбросов расплава по характеристикам вибрации корпуса конвертера с применением методов машинного обучения. Для апробации разработанных алгоритмов в условиях реального производства (ККЦ ОАО «ММК») был разработан комплекс программно-аппаратных средств.

Направление проведенного исследования является перспективным в области автоматизации конвертерного производства и электропривода металлургических агрегатов, поэтому рукопись может быть рекомендована к опубликованию.

Рецензент _____ доцент каф. ЭиМЭ, к.т.н. Швидченко Д.В.

РАЗРАБОТКА СИСТЕМЫ ДИАГНОСТИРОВАНИЯ ВЫБРОСОВ РАСПЛАВА ИЗ КИСЛОРОДНОГО КОНВЕРТЕРА *

Лукьянов С.И., Суспицын Е.С., Мещеряков А.Ю.

(г. Магнитогорск, ГОУ ВПО «Магнитогорский Государственный Технический Университет им. Г.И. Носова»)

Аннотация. Разработана система диагностирования выбросов расплава из кислородного конвертера по параметрам вибрации корпуса конвертера в ККЦ ОАО «ММК». Синтез диагностической модели осуществлялся при помощи методов машинного обучения. На основе указанной системы диагностирования разрабатывается система автоматического управления положением кислородной фурмы, обеспечивающая предотвращение выбросов расплава.

Abstract. A diagnostic system of slopping in a BOF process according to the vibration parameters of the body was developed in the oxygen converter steel shop of OJSC "MMK". Synthesis of a slopping prediction model was implemented with machine learning methods. The developed diagnostic system can be used for slopping prevention through the correction of lance position by automatic electric drive control system.

Кислородно-конвертерный способ передела жидкого чугуна – основной сталеплавильный процесс в мире. Доля стали, полученной кислородно-конвертерным способом, достигает 60% от общего количества. Основным направлением дальнейшего развития кислородно-конвертерного процесса является увеличение производительности и повышение качества выплавляемого металла. Одним из способов решения этих задач является автоматизация процессов управления выплавкой стали. Отсутствие объективных оценок текущей информации, характеризующей динамику окисления углерода, затрудняет управление плавкой, что приводит к выбросам расплава. Нарушение технологического режима во время выбросов приводит к ухудшению качества выпускаемой стали, потерям металла и, в крайних случаях, к потере всей плавки [1].

В настоящее время в ККЦ ОАО «ММК» действуют три кислородных конвертера 1 объемом 370 тонн каждый с верхним способом продувки (рис. 1). Основой автоматизации процесса продувки является АСУТП «Плавка». В ее функции входит сбор и визуализация данных о состоянии процесса, управление электроприводом 2-6 положения кислородной фурмы (КФ) 7, расходом кислорода и присадкой сыпучих материалов. Управление технологическим процессом осуществляется оператором. Основным способом ведения продувки заключается в ступенчатом регулировании высоты фурмы и расхода кислорода по утвержденным шаблонам для каждого типа стали. Динамическая коррекция управляющих воздействий осуществляется оператором на основании данных о химическом составе отходящих газов по показаниям газоанализаторов, визуального наблюдения горловины конвертера и данным системы вибрационного контроля производства НПО «Техноап» в составе трех вибродатчиков 8 и цифровой системы сбора данных 9-11. Вибродатчики установлены на цапфе конвертера 12. При помощи вибродатчиков измеряется виброперемещение корпуса конвертера вдоль трех ортогональных осей. Сигналы оцифровываются с частотой дискретизации 512 Гц и передаются в ПК. Оператору на посту 13 выдается визуальный сигнал, пропорциональный мощности вибрации.

* Работа выполняется при поддержке Федерального агентства по образованию в рамках федеральной целевой программы «Научные и научно-педагогические кадры инновационной России» на 2009-2013 годы, контракт № П2136

$$p_{\text{дан}} = (4200\text{ м}^3 < Q_{O_2} < 14700\text{ м}^3) \wedge (H_{\text{ф}} < 3,5\text{ м}) \wedge (t_{\text{пр}} > 1337 - 11,25M_{\text{л}}), \quad (1)$$

где Q_{O_2} - суммарный расход кислорода, м^3 ; $H_{\text{ф}}$ - высота фурмы, м; $t_{\text{пр}}$ - длительность продувки, с; $M_{\text{л}}$ - масса лома в шихте. Анализ сигнала вибрации проводился при помощи дискретного преобразования Фурье. Выбран размер окна для анализа сигнала вибрации 8 с для обеспечения высокой разрешающей способности на частотах, соответствующих периоду выделения пузырей монооксида углерода из вторичной зоны обезуглероживания (0,2-0,3 с согласно [6]). Проведено сравнение нескольких методов машинного обучения для идентификации выбросов расплава. Наименьшую ошибку классификации показал метод решающих деревьев [7].

На основе анализа данных, приведенных в [8], определен способ управления положением кислородной фурмы для предотвращения выбросов расплава из конвертера. При чрезмерном вспенивании шлака необходимо опускание фурмы в диапазоне от 0,3 до 1 м относительно текущего уровня.

На основе проведенных исследований разработана методика расчета диагностических правил нарушения режима шлакообразования, приводящего к выбросам расплава из кислородного конвертера:

1. Формируется массив статистических данных, включающий сигналы вибродатчиков, технологические параметры (расход кислорода, положение кислородной фурмы, концентрация CO и CO₂ в отходящих газах), сигнал визуального наблюдения выбросов. Частота дискретизации сигналов вибродатчиков составляет 512 Гц.

2. Рассчитывается спектральная плотность мощности (СПМ) сигнала вибрации на квазистационарных участках длительностью 8 с при помощи дискретного преобразования Фурье. Для уменьшения размерности вектора СПМ применяется нелинейное сжатие спектра. Полученный вектор СПМ объединяется с вектором технологических параметров формируя признаковое пространство с размерностью N .

3. Осуществляется ранжирование диагностических признаков по степени их информативности по критерию Пирсона χ^2 .

4. Производится обучение решающих деревьев на выборках, содержащих от 2 до N наиболее информативных признаков по методу [7].

5. Полученные модели деревьев преобразуются в логически эквивалентные пропозициональные правила. Каждое правило тестируется на обучающей и тестовой выборках. Фильтруются правила, выдающие ложные сигналы о возможности выбросов расплава. Оставшиеся правила сохраняются в ПЗУ вычислительного устройства.

Разработана методика диагностирования выбросов расплава из кислородного конвертера:

1. Формируется вектор диагностических признаков, включающий сигналы вибродатчиков, технологические параметры. Частота дискретизации сигналов вибродатчиков составляет 512 Гц.

2. Рассчитывается СПМ сигнала вибрации на окнах длительностью 8 с при помощи дискретного преобразования Фурье, производится нелинейное сжатие спектра. Полученный вектор СПМ объединяется с вектором технологических параметров.

3. Осуществляется сортировка диагностических признаков в порядке, определенном на этапе обучения диагностической модели.

4. В процессе продувки определяется возможность диагностирования выбросов по выражению (1). Диагностирование выбросов осуществляется в период времени, когда $p_{\text{дан}}=1$.

5. При помощи ранее сохраненной модели рассчитывается максимум функции подобия $p_B(t)$ для текущей ситуации. Приближение величины этой функции к 1 свиде-

тельствует о необходимости осаживания шлака путем погружения кислородной фурмы.

По предложенным методикам разработана система диагностирования выбросов расплава из кислородного конвертера (рис. 2.).



Рис. 2. Структурная схема системы диагностирования выбросов расплава

В подсистему обучения модели подается массив архивных данных, состоящий из сигналов вибродатчиков $G1, G2$, параметров технологического процесса D_{ASU} , сигнала наблюдения выбросов V . После преобразования исходных данных формирует выборка \hat{S} , используемая для обучения модели. В режиме диагностирования на основе поступающих сигналов $d_{ASU}, g1, g2$ производится идентификация выбросов моделью M и формируется предупреждающий сигнал F' .

Разработанная система диагностирования была обучена на архивных данных 260 продувок конвертера №1. Применение системы показало ее высокую эффективность (табл. 1).

Таблица 1

Результаты тестирования системы диагностирования выбросов.

Диагностированы выбросы	Выбросы наблюдались	Выбросы не наблюдались
выбросы	32	4
отсутствие выбросов	7	217

Возникновение выбросов предсказывалось за время от 24 до 9 с до визуального наблюдения. Таким образом, по изменению вибрационных характеристик корпуса конвертера возможно заблаговременно идентифицировать тенденции к возникновению выбросов расплава. Сформулировано требование к электроприводу положения КФ: с целью предотвращения выбросов электропривод фурмы должен обеспечивать ее перемещение на 0,3 м из неподвижного состояния за 3 с. Проведенные исследования позволяют разработать и внедрить систему автоматического управления электроприводом кислородной фурмы конвертера для предотвращения выбросов расплава из конвертера.

Литература

1. Технология производства стали в современных конвертерных цехах / С.В. Колпаков, Р.В. Старов, В.В. Смоктий и др. – М.: Машиностроение, 1991. – 464 С.
2. Козловский А.В., Марковский С. Н., Федичкин И. Л. Газоаналитический информационно-контрольный комплекс ЭМГ-21 в цикле АСУ ТП «Плавка» // Контрольно-измерительные приборы и системы. – 2000 – №4. – С.17 – 18.
3. Яценко А. К., Каменев Ю. С. Адаптивное управление в АСУТП выплавки стали в конвертерах. // «Автоматизация конвертерного и прокатного производств», Киев, 1987. – С. 27 – 30.
4. Metallurgy of steel / Под ред. Явойского В.И. и Кряковского Ю.В. – М.: Металлургия, 1983. – 329 С.
5. Контроль и оценивание конвертерной плавки по косвенным параметрам / М.И. Волович, В.И. Авдеев, Я.Г. Парпаров // Кемеровское книжное издательство, 1989. – 124 С.
6. В. И. Баптизманский, А. Г. Величко. Идентификация колебательных процессов при обезуглероживании металла в ходе конвертерной плавки // Металлы. – 1993 – №6. – С. 18 – 22.
7. Применение методов машинного обучения при анализе вибрации кислородного конвертера / Лукьянов С.И., Суспицын Е.С., Мещеряков А.Ю. и др. // Материалы 67-й научно-технической конференции: сб. докл., Магнитогорск: ГОУ ВПО «МГТУ», 2009. – С. 108 – 111.
8. Комбинированное управление конвертерной плавкой / М.И. Волович, В.П. Авдеев и др. – Кемеровское книжное издательство, 1990. – 141 С.



Лукьянов Сергей Иванович – д-р техн. наук, проф. кафедры “Электроника и микроэлектроника” Магнитогорского государственного технического университета им. Г.И. Носова.



Суспицын Евгений Сергеевич – к-т техн. наук, доцент кафедры “Электроника и микроэлектроника” Магнитогорского государственного технического университета им. Г.И. Носова.



Мещеряков Александр Юрьевич – аспирант кафедры “Электроника и микроэлектроника” Магнитогорского государственного технического университета им. Г.И. Носова, E-mail: magalm7@mail.ru

Заключение

В ходе выполнения научно-исследовательской работы был разработан прибор для диагностирования аномалий процессов в неоднородных средах. Применение методов машинного обучения в качестве основы для классификации состояния технологического процесса позволило автоматизировать настройку указанного прибора для функционирования в условиях различных производств. Проведенный анализ нескольких методов машинного обучения при решении задачи диагностирования выбросов расплава из кислородного конвертера показал наибольшую эффективность алгоритма решающих деревьев С4.5. Применение комплекса разнородных алгоритмов синтеза диагностической модели в разработанном приборе позволяет эффективно адаптировать его для решения широкого круга задач.

В ходе выполнения заключительного этапа работы были достигнуты следующие основные результаты:

- 1) проведен статистический анализ экспериментальных данных по диагностическим признакам аномалий в ходе продувки кислородного конвертера;
- 2) проведен анализ и адаптация методов машинного обучения для диагностирования аномалий процессов в неоднородных средах;
- 3) проведено сравнение эффективности различных методов машинного обучения при решении задачи классификации состояния процессов в неоднородных средах;
- 4) разработаны обобщенные алгоритмы работы вычислительного блока прибора для диагностирования аномалий процессов в неоднородных средах;
- 5) проведена апробация прибора для диагностирования аномалий процессов в неоднородных средах в технологических условиях кислородно-конвертерного цеха ОАО «ММК».

Внедрение прибора для диагностирования аномалий процессов в неоднородных средах на производственных предприятиях позволит снизить простои, связанные с возникновением аномалий технологических процессов, увеличить выход годной продукции.

Список использованных источников

1. М.И. Волович, В.И. Авдеев, Я.Г. Парпаров. Контроль и оценивание конвертер-ной плавки по косвенным параметрам. – Кемеровское книжное издательство, - 1989. – 124 с.
2. Айфичер Э.С., Джервис Б.У. Цифровая обработка сигналов: практический под-ход. – М.: Издательский дом «Вильямс», 2004. – 992 с.
3. Вентцель Е.С., Овчаров Л.А. Теория вероятностей и ее инженерные приложе-ния. Учеб. пособие для вузов. - 2-е изд. стер. – М.: Высш. шк., 2000. – 480 с.
4. Айвазян С. А., Бухштабер В. М., Енюков И. С., Мешалкин Л. Д. Прикладная статистика: классификация и снижение размерности. — М.: Финансы и стати-стика, 1989. – 272 с.
5. Mitchell T. Machine Learning. — McGraw-Hill Science/Engineering/Math, 1997. – 524 с.
6. Kononenko I. Estimating attributes: analysis and extension of relief / In Proceedings of 7th European Conference on Machine Learning. – Springer-Verlag, Catania, Italy, 1994. – С. 171–182.
7. Ruan D., Guoqing C.. Intelligent Data Mining. Techniques and Applications. – Springer, 2005. – 523 с.
8. Осовский С. Нейронные сети для обработки информации. – М.: Финансы и ста-тистика, 2002. – 344 с.
9. Круглов В.В., Борисов В.В. Искусственные нейронные сети. Теория и практика. – М.: Горячая линия–Телеком, 2002. – 382 с.
10. Комашинский В.И., Смирнов Д.А. Нейронные сети и их применение в системах управления и автоматике. – М.: Горячая линия–Телеком, 2003. – 94 с.
11. Галушкин А.И. Теория нейронных сетей. Кн. 1: Учеб. пособие для вузов. – М.: ИПРЖР, 2000. – 416 с.
12. Witten I. H. Data mining : practical machine learning tools and techniques. – Morgan Kaufmann, 2005. – 525 с.

13. Вапник В.Н., Червоненкис А.Я. Теория распознавания образов. Статистические проблемы обучения. – М.: Наука, 1974. – 215 с.
14. Paliouras G., Karkaletsis V. Machine learning and its applications: advanced lectures. – Springer, 2001. – 332 с.
15. Larose D. T. Data mining methods and models . – John Wiley & Sons Inc., 2006. – 289 с.
16. Manning C.D., Raghavan P., Schütze H. Introduction to Information Retrieval. – Cambridge University Press, 2008. – 496 с.