

УДК 519.248:681.518.5

В.Н. Клячкин, Д.А. Жуков

АЛГОРИТМ ДИАГНОСТИКИ ФУНКЦИОНИРОВАНИЯ ТЕХНИЧЕСКОГО ОБЪЕКТА С ИСПОЛЬЗОВАНИЕМ АГРЕГИРОВАННЫХ КЛАССИФИКАТОРОВ¹

Клячкин Владимир Николаевич, доктор технических наук, профессор, окончил механический факультет Ульяновского политехнического института. Профессор кафедры «Прикладная математика и информатика» Ульяновского государственного технического университета. Имеет научные труды в области надежности и статистических методов. [e-mail: v_kl@mail.ru].

Жуков Дмитрий Анатольевич, окончил факультет информационных систем и технологий УлГТУ, специалист по базе данных Ульяновского филиала конструкторского бюро ПАО «Туполев», аспирант кафедры «Прикладная математика и информатика» УлГТУ. Имеет научные труды в области статистических методов и машинного обучения. [e-mail: zh.dimka17@mail.ru].

Аннотация

Рассматривается задача прогнозирования исправности технического объекта по известным показателям его функционирования. Исходными данными являются известные результаты оценки состояния объекта по информации о предшествующей эксплуатации: при заданных значениях контролируемых показателей техническая система исправна или неисправна. Такая задача может быть решена методами машинного обучения, она сводится к бинарной классификации состояний объекта. Проведенное исследование показало, что качество диагностики может быть повышено за счет выбора метода бинарной классификации, включая агрегированный подход, а также путем выбора объема контрольной выборки и способа отбора значимых показателей функционирования объекта. В рассмотренном примере (на котором отработывался предлагаемый алгоритм) значение F -меры, как наиболее объективной характеристики качества диагностики при несбалансированных классах, повысилось на 6% (с 0,86 до 0,91) по сравнению с методом опорных векторов и на 2% – по сравнению с лучшим для рассматриваемого примера из базовых методов – бэггингом деревьев решений, для которого значение F -меры составило 0,89. В некоторых ситуациях эти 2% могут быть весьма существенны с точки зрения обеспечения надежности при эксплуатации объекта.

Ключевые слова: техническая диагностика, исправность, показатели функционирования, машинное обучение, агрегированный подход, контрольная выборка, кросс-валидация.

doi: 10.35752/1991-2927-2019-2-56-37-43

ALGORITHM OF DIAGNOSTICS OF TECHNICAL OBJECT OPERATION USING AGGREGATED CLASSIFIERS

Vladimir Nikolaevich Kliachkin, Doctor of Science in Engineering; graduated from the Mechanical Faculty of the Ulyanovsk Polytechnic Institute; Professor at the Department of Applied Mathematics and Computer Science of Ulyanovsk State Technical University; an author of articles in the field of reliability issues and statistical methods. e-mail: v_kl@mail.ru.

Dmitrii Anatolevich Zhukov, graduated from the Faculty of Information System and Technologies of Ulyanovsk State Technical University; a database engineer at the Ulyanovsk Branch of PJSC 'Tupolev Design Bureau'; Postgraduate Student of the Department of Applied Mathematics and Computer Science; an author of articles in the field of the statistics methods and machine learning. e-mail: zh.dimka17@mail.ru.

Abstract

The paper addresses the issue of prediction of technical object's state of health using the known indicators of its operation. The basic data are the known results of the object state estimation by information about previous service: the technical system is healthy or faulty with predetermined values of specified indicators. Such problem may be solved using the machine learning methods, it reduces to binary classification of states of the object. The study showed that diagnostics quality may

¹ Исследование выполнено при финансовой поддержке РФФИ и Правительства Ульяновской области, проект № 18-48-730001.

be improved by means of the binary classification method including aggregated approach, as well as by means of selection of the volume of the validation set and the method of selection of relevant indicators of object operation. In the example (used for the algorithm testing), F -measure value, which is the most reliable diagnostics quality measure for unbalanced classes, has increased by 6% (from 0.86 to 0.91), compared with the Support Vector Machine, and by 2%, compared with the bagging of decision trees which is the best basic method for the example considered ($F = 0.89$). In some cases, this 2% may be of significance from the perspective of the object operation security.

Keywords: technical diagnostics, healthy state, operation indicators, machine learning, aggregated approach, validation set, cross-validation.

ВВЕДЕНИЕ

Для обеспечения надежности технического объекта проводится его диагностика в условиях эксплуатации. При этом часто диагностика сводится к распознаванию одного из двух состояний объекта, к разделению состояний на исправные или неисправные [1]. Это задача бинарной классификации, которая может быть решена с использованием методов машинного обучения [2–4].

Исходными данными являются известные результаты измерений косвенных показателей функционирования объекта по результатам предшествующей эксплуатации. Эти данные представляются в виде матрицы показателей функционирования системы X , элементы которой x_{ij} – результат i -го наблюдения по j -му показателю; $i = 1, \dots, l, j = 1, \dots, p$ (l – количество строк, или число наблюдений, p – количество столбцов, или число показателей функционирования), и вектор-столбец Y , состоящий из ответов 0 или 1, для опытов, где объект оказался неисправным и исправным соответственно. Совокупность (x_i, y_i) , где значение y_i вектора Y соответствует строке x_i матрицы X , представляет прецедент.

По множеству прецедентов необходимо построить модель, которая будет выдавать вероятность принадлежности нового состояния объекта к одному из двух классов. Существуют различные методы машинного обучения, обеспечивающие решение этой задачи с той или иной степенью точности. В работах [5, 6] для повышения качества классификации предложено использование агрегированного подхода, суть которого состоит в объединении результатов, полученных разными методами машинного обучения, в одно значение, которое будет учитывать особенности каждого. При этом важно учесть множество факторов, оказывающих влияние на точность решения.

Цель исследования – разработка алгоритма диагностики функционирования технического объекта, обеспечивающего наилучшее качество распознавания исправности этого объекта при заданном наборе прецедентов (показателей его работы по результатам эксплуатации) и используемых базовых методах машинного обучения.

1 ВЛИЯНИЕ ОСНОВНЫХ ФАКТОРОВ НА КАЧЕСТВО ДИАГНОСТИКИ

Для оценки качества диагностики исходную выборку случайным образом разбивают на два непере-

секающихся подмножества. Первое подмножество – собственно обучающая выборка – предназначена для решения задачи обучения, то есть для оценки параметров рассматриваемой модели классификации. Второе подмножество является контрольной (или тестовой) выборкой, не используемой для обучения, а предназначенной для оценки качества обучения по используемому критерию. Часто используется кросс-валидация: исходная выборка разбивается на k частей. При этом $(k - 1)$ часть используется для обучения, одна – для контроля. Последовательно перебираются все варианты. Для каждого разбиения вычисляется заданный критерий качества диагностики и усредняется.

Используется множество различных подходов к бинарной классификации. Это и классические статистические модели (например, логистическая регрессия), и методы, специально ориентированные на машинное обучение (метод опорных векторов, нейронные сети), композиционные методы (бэггинг, бустинг) и другие. Практическая реализация этих методов возможна, например, на базе библиотеки инструментов Statistics and Machine Learning Toolbox в пакете Matlab.

Часто эффективен агрегированный подход. Если в композиционных методах (бэггинг, бустинг) для построения ансамбля используется один и тот же метод классификации, построенный на разных подмножествах выборки или ориентированный на компенсацию ошибки предыдущей итерации [7, 8], то при агрегированном подходе используются совместно различные методы классификации, построенные на всей обучающей выборке.

Можно воспользоваться полным перебором наборов из H базовых методов. Тогда, например, при $H = 2$ получим три набора: два базовых и один агрегированный; при $H = 3$ наборов уже 7: три базовых, три агрегированных по два базовых и один агрегированный из всех трех базовых методов. Нетрудно видеть, что в общем случае число наборов равно $2^H - 1$.

При формировании единого решения об исправности объекта на основе решений отдельных методов классификации возможно агрегирование результатов по среднему значению, по медиане, и с помощью процедуры голосования [5, 6]. В частности, результат агрегированного метода классификации по голосованию представляет собой среднее значение результатов базовых методов, которые определили исправность объекта с вероят-

ностью, например, не ниже, чем $0,1 \cdot (\hat{P}_K(X_r) \geq 0,1)$,

где \hat{P}_K – вероятность исправности объекта при исходных показателях функционирования объекта X_r . В противном случае вероятность того, что r -й объект исправен, равняется нулю. Таким образом, значения вероятностей классификации, оказавшиеся ниже 0,1, приравниваются к 0, а оставшиеся – к 1, и по этим значениям строятся агрегированные модели классификации.

Для оценки качества бинарной классификации наиболее распространенной метрикой является доля правильных ответов, или наоборот, процент ошибок на контрольной выборке.

При несбалансированных классах (как правило, в выборке количество данных об исправных состояниях технического объекта значительно больше, чем о неисправных) доля ошибок не может объективно оценивать качество классификации [9, 10]. Гораздо более информативна F -мера – это гармоническое среднее точности и полноты: чем ближе значение F к единице, тем выше качество классификации.

В качестве еще одного функционала качества может быть выбрана площадь под ROC-кривой (receiver operating characteristics): AUC (area under the curve) [9–13]. ROC-кривая образуется, если по оси абсцисс отложить значения $fp(c)$ – количество неправильно классифицированных исправных состояний в зависимости от заданного порога c , а по оси ординат $tp(c)$ – количество правильно классифицированных исправных состояний. Чем ближе значение AUC к единице, тем лучше качество диагностики.

Проведенное ранее исследование [14] показало, что качество диагностики может быть повышено путем рационального выбора объема контрольной выборки: различные методы машинного обучения по-разному реагируют на изменение этой величины. За счет ее правильного выбора возможно повышение точности диагностики (в рассмотренных примерах) на 5–7%. В связи с этим предлагается на стадии выбора алгоритма для проведения технической диагностики конкретного объекта провести соответствующие испытания.

Еще более существенно с точки зрения качества обучения влияние отбора значимых показателей функционирования технического объекта. Результаты исследований [12, 15] показывают, что удаление незначимых факторов оказывает неоднозначное влияние на эффективность машинного обучения, и для каждого метода подбор значимых факторов индивидуален. Путем варьирования этого фактора качество бинарной классификации объекта в рассмотренных примерах было повышено до 18%. Поэтому в системе диагностики состояния технического объекта необходимо предусмотреть подсистему отбора значимых показателей, обеспечивающую минимальную погрешность.

2 ПРЕДЛАГАЕМЫЙ АЛГОРИТМ ДИАГНОСТИКИ ФУНКЦИОНИРОВАНИЯ ТЕХНИЧЕСКОГО ОБЪЕКТА

После считывания файла исходных данных задается объем контрольной выборки. Для сокращения вычислений примем начальное значение объема контроля 25%, т. е. при кросс-валидации выборка разбивается на $k = 4$ части, три из которых используются для обучения.

Далее проводится обучение по всем 11 методам с использованием всех заданных показателей функционирования объекта (без отбора значимых показателей) на базе библиотеки инструментов Statistics and Machine Learning Toolbox в пакете Matlab с расчетом F -меры (эти методы перечислены в таблице 1). Выбирается заданное пользователем количество методов (исследования показывают, что достаточно два–три, хотя возможен и полный перебор), показавших наибольшее значение F -меры.

Таблица 1
Методы бинарной классификации в Matlab

Методы	Краткое обозначение
Логистическая регрессия	ЛР
Дискриминантный анализ	ДА
Наивный байесовский классификатор	БК
Нейронные сети	НС
Метод опорных векторов	МОВ
Бэггинг деревьев решений	БДР
Градиентный бустинг	GrB
AdaBoost	AB
LogitBoost	LB
GentleBoost	GB
RUSBoost	RB

Возможно и изменение порога – значения, которое определяет границы классов: например, если порог равен 0,6, то значения вероятностей от 0,6 до 1 будут относиться к классу исправных объектов, а меньше 0,6 – к классу неисправных.

Для отобранных методов исследуется вначале влияние объема контрольной выборки, а затем – способа отбора значимых показателей. Объем контрольной выборки последовательно снижается с 25% ($k = 4$) до 5% (соответственно $k = 20$) с шагом 5%. Для вариантов, в которых F -мера оказалась максимальной, проводится отбор значимых показателей. При этом используются два подхода. Один – удаление показателей функционирования объекта, для которых их корреляция с его исправностью незначима; второй – по незначимости регрессоров (по критерию Стьюдента) в линейной регрессионной модели Y от X .

По результатам проведенных испытаний принимаются значение объема контрольной выборки и показатели, которые будут использованы для продолжения машинного обучения. С учетом этих двух факторов строятся агрегированные классификаторы трех типов – по среднему значению (АМС), по медиане (АММ) и с помощью процедуры голосования (АМГ). Выбирается модель (включающая вид агрегирования и компоненты агрегата: несколько базовых классификаторов), которая в дальнейшем будет использована для прогнозирования состояния технического объекта (рис. 1).



Рис. 1. Блок-схема алгоритма диагностики

3 ЧИСЛЕННОЕ ИССЛЕДОВАНИЕ

Для проведения испытаний использовались результаты наблюдений за функционированием системы водоочистки. Исправность системы Y оценивалась по показателям качества питьевой воды в зависимости от физико-химических показателей водоисточника: X_1 – температуры, X_2 – цветности, X_3 – мутности, X_4 – значения pH, X_5 – щёлочности, X_6 – окисляемости и доз добавляемых реагентов: X_7 – коагулянта и X_8 – флокулянта. Получено 527 наборов данных, из них в 127 случаях система была признана неисправной ($Y=0$) – показатели качества питьевой воды не соответствовали требованиям. Доля неисправностей (24%) свидетельствует о несбалансированности классов исправных и неисправных ситуаций. Для оценки качества диагностики использовалась F -мера.

На рисунке 2 показано окно программы с разделением выборки на обучающую и контрольную (25%) части, а на рисунке 3 – результаты расчета F -меры и площади AUC под ROC-кривой ошибок. Значение AUC целесообразно использовать для принятия решения в случае совпадения значений F -меры по разным методам.

Из рисунка 3 следует, что лучшие в данных испытаниях методы классификации по F -мере – это БДР, для которого F -мера составила 0,8917, LB (0,8863) и GrB (0,8863). Дальнейшие испытания проводились для этих трех методов.

В таблице 2 приведены результаты исследования влияния объема контрольной выборки: лучший результат получен при объеме 10% (расхождения в рассматриваемом примере оказались не слишком существенны, в других ситуациях они могут быть значительно большими [14]). При этом объеме контрольной выборки оценивалось влияние способа отбора значимых показателей, результаты которого приведены в таблице 3: лучший результат получен при использовании для обучения только значимых по регрессионной модели показателей.

Таким образом, для формирования агрегированных классификаторов использовалась кросс-валидация с разбиением выборки на $k=10$ частей (объем контрольной выборки 10%) с участием только тех показателей функционирования, которые оказались значимыми по регрессионной модели. Результаты расчета представлены на рисунке 4. Видно, что структура всех агреги-

Таблица 2

Значения F -меры в зависимости от объема контрольной выборки

Метод	5%	10%	15%	20%	25%
БДР	0,8956	0,8985	0,8974	0,8957	0,8917
GrB	0,8857	0,8851	0,8888	0,8883	0,8847
LB	0,8820	0,8786	0,8854	0,8920	0,8863

Выбор данных

Таблица с данными

	1	2	3	4	5	6	7
1	1	1.2000	36	13	7.5200	0.5500	7.1000
2	1	1.1000	36	12	7.6200	0.5700	8.6000
3	1	0.6000	41	9.2000	7.5900	0.5700	8.6000
4	1	0.4000	40	3.1000	7.4100	0.5800	8.6000
5	1	0.4000	37	5	7.6000	0.6200	8.2000
6	1	0.3000	39	4.2000	7.5400	0.5500	7.8000
7	1	0.7000	40	2.8000	7.4900	0.6300	8.6000
8	1	0.6000	36	1.5000	7.5500	0.5500	8.4000
9	1	0.3000	35	1.5000	7.4300	0.5500	7.9000
10	1	0.2000	38	1.7000	7.5000	0.6000	7.9000
11	1	0.2000	37	3.9000	7.3400	0.6000	8.6000
12	0	0.2000	37	1.6000	7.5100	0.5500	7.3000
13	1	0.2000	37	1.8000	7.3700	0.6000	7.9000

Объем: 527

Обучающая выборка

	1	2	3	4	5	6	7
1	1	1.1000	36	12	7.6200	0.5700	8.6000
2	1	0.6000	41	9.2000	7.5900	0.5700	8.6000
3	1	0.4000	37	5	7.6000	0.6200	8.2000
4	1	0.3000	39	4.2000	7.5400	0.5500	7.8000
5	1	0.7000	40	2.8000	7.4900	0.6300	8.6000
6	1	0.6000	36	1.5000	7.5500	0.5500	8.4000
7	1	0.3000	35	1.5000	7.4300	0.5500	7.9000
8	1	0.2000	38	1.7000	7.5000	0.6000	7.9000
9	1	0.2000	37	3.9000	7.3400	0.6000	8.6000
10	0	0.2000	37	1.6000	7.5100	0.5500	7.3000
11	1	0.2000	37	1.8000	7.3700	0.6000	7.9000
12	1	0.3000	36	1.1000	7.4200	0.6300	8
13	1	0.3000	34	1	7.3800	0.5500	7.2000

Объем обучающей: 395

Процент контрольной выборки: Разделить выборку

Отбор значимых признаков:

с учетом коррелированности

по значимым регрессорам

Рис. 2. Разделение выборки на обучающую и контрольную

Методы машинного обучения

Порог:

Логистическая регрессия (ЛР)

Дискриминантный анализ (ДА)

Байесовский классификатор (БК)

Нейронная сеть (НС)

Метод опорных векторов (МОВ)

Бэггинг деревьев решений (БДР)

Методы бустинга:

RUSBoost (RB)

	Метод	F-мера	AUC
1	ЛР	0.8816	0.8778
2	ДА	0.8592	0.7150
3	БК	0.8616	0.7161
4	НС	0.8704	0.7556
5	МОВ	0.8556	0.8106
6	БДР	0.8917	0.9065
7	GrB	0.8847	0.8994
8	AB	0.8835	0.8009
9	LB	0.8863	0.8568
10	GB	0.8739	0.8467
11	RB	0.8666	0.8685

Рис. 3. Критерии качества диагностики по различным классификаторам

Таблица 3

Значения *F*-меры
в зависимости от показателей функционирования

Метод	Все показатели	Только показатели, коррелированные с исправностью	Только показатели, значимые по регрессионной модели
БДР	0,8985	0,8880	0,8990
GrB	0,8851	0,8864	0,8855
LB	0,8786	0,8876	0,8930

рованных классификаторов одинакова: включает БДР и метод LB (которые оказались и лучшими базовыми методами – см. рис. 3). При этом более высокое значение *F*-меры, равное 0,9095, соответствует одновременно двум методам агрегирования – и по среднему, и по медиане. Значение AUC у этих классификаторов также одинаково.

Для прогнозирования исправности объекта при новых значениях показателей функционирования $X_1 - X_8$ необходимо ввести эти значения из файла и выполнить прогноз: программа рассчитает, исправен объект ($Y = 1$) или неисправен ($Y = 0$) с помощью построенных агрегированных классификаторов (рис. 5).

	Метод	F-мера	AUC
1	AM-С: БДР+ LB	0.9095	0.9087
2	AM-М: БДР+ LB	0.9095	0.9087
3	AM-Г: БДР+ LB	0.9074	0.9031

Рис. 4. Выбор агрегированного классификатора

Агрегирование по среднему	1
Агрегирование по медиане	1
Агрегирование по голосованию	1

Рис. 5. Прогноз нового состояния: объект исправен

ЗАКЛЮЧЕНИЕ

Проведенное исследование показало, что качество диагностики может быть повышено за счет выбора метода бинарной классификации, включая агрегированный подход, а также путем выбора объема контрольной выборки и способа отбора значимых показателей функционирования объекта. В рассмотренном примере (на котором обрабатывался предлагаемый алгоритм) значение F -меры, как наиболее объективной характеристики качества диагностики при несбалансированных классах, повысилось на 6% (с 0,86 до 0,91) по сравнению с методом опорных векторов и на 2% – по сравнению с лучшим для рассматриваемого примера из базовых методов – БДР, для которого $F = 0,89$. В некоторых ситуациях эти 2% могут быть весьма существенны с точки зрения обеспечения надежности при эксплуатации объекта.

СПИСОК ЛИТЕРАТУРЫ

1. Биргер И.А. Техническая диагностика. – М. : Машиностроение, 1978. – 240 с. (2-е изд.: М. : URSS, 2019).
2. Witten I.H., Frank E. Data mining: practical machine learning tools and techniques. 2nd ed. – San Francisco : Morgan Kaufmann Publishers, 2005. – 525 p.
3. Мерков А.Б. Распознавание образов. Введение в методы статистического обучения. – М. : Едиториал УРСС, 2011. – 256 с.
4. Теория и практика машинного обучения : учеб. пособие / В.В. Воронина, А.В. Михеев, Н.Г. Ярушкина, К.В. Святков. – Ульяновск : УлГТУ, 2017. – 290 с.
5. Клячкин В.Н., Кувайскова Ю.Е., Жуков Д.А. Диагностика технического состояния аппаратуры с использованием агрегированных классификаторов // Радиотехника. – 2018. – № 6. – С. 46–49.
6. Klyachkin V.N., Kuvayskova Yu.E., Zhukov D.A. The use of aggregate classifiers in technical diagnostics, based on machine learning // CEUR Workshop Proceedings. Data Science. Information Technology and Nanotechnology. – 2017. – Vol. 1903. – pp. 32–35.
7. Neykov M., Jun S. Liu, Tianxi Cai. On the Characterization of a Class of Fisher-Consistent Loss Functions and its Application to Boosting // Journal of Machine Learning Research. – 2016. – № 17 (70). – pp. 1–32.
8. Explaining the Success of AdaBoost and Random Forests as Interpolating Classifiers / A.J. Wyner, M. Olson, J. Bleich, D. Mease // Journal of Machine Learning Research. – 2017. – № 18 (48). – pp. 1–33.
9. Соколов Е.А. ФКН ВШЭ. Лекция 4. Линейная классификация. – URL: <https://github.com/esokolov/ml-course-hse/blob/master/2018-fall/lecture-notes/lecture04-linclass.pdf> (дата обращения: 01.03.2019).
10. Davis J., Goadrich M. The relationship between Precision-Recall and ROC curves // Proceedings of the 23rd international conference on Machine learning. – Pittsburgh, 2006. – pp. 233–240.

11. Kuvayskova Yu.E. The prediction algorithm of the technical state of an object by means of fuzzy logic inference models // Procedia Engineering. 3rd International Conference «Information Technology and Nanotechnology», ITNT 2017. – 2017. – pp. 767–772.

12. Воронцов К.В. Машинное обучение. Композиция классификаторов. – URL: <https://yadi.sk/i/Fitlu6V0beBmF> (дата обращения: 01.03.2019).

13. Дьяконов А.М. AUC ROC (площадь под кривой ошибок). – URL: <https://dyakonov.org/2017/07/28/auc-gos-ploshchad-pod-krivoy-oshibok/#more-5362> (дата обращения: 01.03.2019).

14. Жуков Д.А., Клячкин В.Н. Влияние объема контрольной выборки на качество диагностики состояния технического объекта // Автоматизация процессов управления. – 2018. – № 2 (52). – С. 90–95.

15. Клячкин В.Н., Кувайскова Ю.Е., Жуков Д.А. Влияние способа отбора значимых показателей на качество диагностики состояния технического объекта // Автоматизация. Современные технологии. – 2019. – Т. 73, № 1. – С. 32–36.

REFERENCES

1. Birger I.A. *Tekhnicheskaja diagnostika* [Technical Diagnostics]. Moscow, Mashinostroenie Publ., 1978. 240 p. (2-e izd.: M. : URSS, 2019).
2. Witten I.H., Frank E. *Data Mining: Practical Machine Learning Tools and Techniques*. 2nd ed. San Francisco, Morgan Kaufmann Publ., 2005. 525 p.
3. Merkov A.B. *Raspoznavanie obrazov. Vvedenie v metody statisticheskogo obucheniia* [Image Recognition. Methods of Statistical Learning]. Moscow, Editorial URSS Publ., 2011. 256 p.
4. Voronina V.V., Mikheev A.V., Yarushkina N.G., Sviatov K.V. *Teoriia i praktika mashinnogo obucheniia. Ucheb. posobie* [Theory and Practice of Machine Training]. Ulyanovsk, UISTU Publ., 2017. 290 p.
5. Klyachkin V.N., Kuvayskova Yu.E., Zhukov D.A. Diagnostika tekhnicheskogo sostoianiia apparatury s ispolzovaniem agregirovannykh klassifikatorov [Diagnostics of Technical State of the Equipment Using Aggregated Classifiers]. *Radiotekhnika* [Radioengineering], 2018, no. 6, pp. 46–49.
6. Klyachkin V.N., Kuvayskova Yu.E., Zhukov D.A. The Use of Aggregate Classifiers in Technical Diagnostics Based on Machine Learning. *CEUR Workshop Proceedings Data Science. Information Technology and Nanotechnology*. 2017, vol. 1903, pp. 32–35.
7. Neykov M., Jun S. Liu, Tianxi Cai. On the Characterization of a Class of Fisher-Consistent Loss Functions and its Application to Boosting. *Journal of Machine Learning Research*, 2016, no. 17 (70), pp. 1–32.
8. Wyner, A.J., M. Olson, J. Bleich. D. Mease. Explaining the Success of AdaBoost and Random Forests as Interpolating Classifiers. *Journal of Machine Learning Research*, 2017, no. 18 (48), pp. 1–33.
9. Sokolov E.A. *FKN VShE. Lektsiia 4. Lineinaia klassifikatsiia* [FKN VShE. Lesson 4, Linear Classification].

Available at: <https://github.com/esokolov/ml-course-hse/blob/master/2018-fall/lecture-notes/lecture04-linclass.pdf> (accessed: 01.03.2019).

10. Davis J., Goadrich M. The Relationship Between Precision-Recall and ROC Curves. *Proceedings of the 23rd Intern. Conf. on Machine Learning*. Pittsburgh, 2006, pp. 233–240.

11. Kuvaiskova Iu.E. The Prediction Algorithm of the Technical State of an Object by Means of Fuzzy Logic Inference Models. *Procedia Engineering. 3rd Int. Conf. 'Information Technology and Nanotechnology', ITNT 2017*. 2017, pp. 767–772.

12. Vorontsov K.V. *Mashinnoe obuchenie. Kompozitsiia klassifikatorov* [Machine Learning. Classifiers Structure]. Available at: <https://yadi.sk/i/Ftlu6V0beBmF> (accessed: 01.03.2019).

13. Diakonov A.M. *AUC ROC* (ploshchad pod krivoi oshibok) [AUC ROC. Area under Error Curve]. Available at:

<https://dyakonov.org/2017/07/28/auc-roc-ploshchad-pod-krivoy-oshibok/#more-5362> (accessed: 01.03.2019).

14. Zhukov D.A., Kliachkin V.N. Vliianie obiema kontrolnoi vyborki na kachestvo diagnostiki sostoianiiia tekhnicheskogo objekta [The Effect of the Control Sample Volume on the Quality of Diagnostics of the Technical Objects State]. *Avtomatizatsiia protsessov upravleniia* [Automation of Control Processes], 2018, no. 2 (52), pp. 90–95.

15. Kliachkin V.N., Kuvaiskova Iu.E., Zhukov D.A. Vliianie sposoba otbora znachimykh pokazatelei na kachestvo diagnostiki sostoianiiia tekhnicheskogo objekta [The Influence of the Significant Indicators Selection Method on the Quality of the Diagnostics]. *Avtomatizatsiia. Sovremennye tekhnologii* [Automation. Modern Technologies], 2019, vol. 73, no. 1, pp. 32–36.