

УДК 681.518.5

АНАЛИЗ МАТЕМАТИЧЕСКИХ МОДЕЛЕЙ ПРИ ПОСТРОЕНИИ НЕЙРОННОЙ СЕТИ ДЛЯ ОЦЕНКИ РАБОТОСПОСОБНОСТИ АГРЕГАТА

© 2021 А.А. Санталов

Ульяновский государственный технический университет, Ульяновск, Россия

Статья поступила в редакцию 10.12.2020

Предупреждение аварийных ситуаций авиационной техники в значительной мере обеспечивается диагностикой функционирования ее агрегатов. Часто критерием качества функционирования является уровень вибраций, влияющий на решение о корректировке нагрузки или останове агрегата. В статье рассматриваются математические модели при применении нейросетевых методов для вибродиагностики. При использовании кросс-валидации исходная выборка с данными по вибрациям разбивается на несколько блоков, которые группируются в три выборки: обучающую, контрольную и тестовую. Для оценки эффективности диагностики использовались три разных критерия качества: средняя ошибка на тестовой выборке, AUC и F-мера. Для заданного набора исходных данных наилучшей подобранной конфигурацией оказалась нейронная сеть из трех слоев с 18 нейронами в каждом слое, реализованная в пакете MATLAB. В качестве функции обучения в ней используется алгоритм байесовской регуляризации. Процент средней ошибки распознавания состояния рассматриваемого агрегата с помощью нейронной сети оказался равным 4,85, значение AUC равно 0,885, а F-меры – 0,827. По сравнению с сетью, построенной в автоматическом режиме с помощью библиотек машинного обучения Statistics and Machine Learning Toolbox и Neural Network Toolbox, значение F-меры подобранной конфигурации сети выше на 6,7 %.

Ключевые слова: техническая диагностика, бинарная классификация, нейронная сеть, меры качества, вибродиагностика

DOI: 10.37313/1990-5378-2021-23-1-84-90

*Работа выполнена при поддержке гранта РФФИ
и Правительства Ульяновской области, проект 18-48-73001*

Диагностика функционирования технических объектов позволяет предупредить возникновение аварийных ситуаций. Для решения задач контроля технического состояния, поиска места и причин неисправности, а также прогнозирования состояния объекта широко используются статистические и интеллектуальные методы [1-5].

Технология вибромониторинга предназначена для анализа состояния технического объекта в условиях его эксплуатации. При этом контролируемые во время эксплуатации параметры часто не позволяют дать однозначное заключение о состоянии объекта мониторинга. Между тем принятие решения о работоспособности или неработоспособности объекта связано с риском ложной тревоги (когда исправный объект признается неисправным) или пропуском неисправности (наоборот, когда неисправный объект признается исправным).

Машинное обучение, в частности, нейронных сетей, активно используется при решении задач диагностики. Например, нейросетевые методы применены для решения задач контроля технического состояния авиационного газотурбинно-

го двигателя. Для решения поставленной задачи авторы разработали архитектуру нейросетевой экспертной системы «Эксперт Нейро» [6,7].

Для предупреждения возможных нарушений проводится вибромониторинг агрегатов, при этом получаемые в реальном времени данные по вибрациям рабочих узлов влияют на корректировку нагрузки и режима работы. Необходимость оперативного вмешательства и выработки профилактических мероприятий определяется по множеству показателей, характеризующих качество функционирования агрегата. Это показания датчиков относительной и абсолютной вибрации различных узлов агрегата, боя валов, датчиков измерения скорости вращения и других.

Решение задачи диагностики агрегата по результатам вибромониторинга возможно с применением различных подходов. Для анализа стабильности функционирования используются методы статистического управления процессами, для оценки исправности агрегата – методы машинного обучения, в частности, агрегированные классификаторы. В настоящей статье исследуется эффективность применения для вибродиагностики агрегата нейросетевых методов.

В процессе вибромониторинга агрегата контролируются десять показателей вибраций.

*Санталов Антон Александрович, аспирант кафедры
«Прикладная математика и информатика».
E-mail: anton.santalov1995@gmail.com*

Необходимо распознать состояние агрегата – работоспособен он или неработоспособен. Это задача бинарной классификации [8-10]. Исходными данными являются известные результаты оценки состояния системы: при заданных значениях контролируемых показателей вибраций X_j ($j = 1, \dots, 10$) агрегат работоспособен ($Y = 1$) или неработоспособен ($Y = 0$). В настоящем исследовании исходная выборка составила 2000 наблюдений, при этом доля неработоспособных состояний составила 14,75 %.

Полученная выборка разбивается на три части: обучающую, контрольную и тестовую.

Обучающая часть предназначена для построения модели нейронной сети – зависимости между показателями функционирования агрегата и его состояниями.

Контрольная выборка используется для текущей оценки качества обучения и дает возможность предотвратить переобучение нейронной сети, что позволяет сохранить эффективность распознавания состояния агрегата на примерах, не участвовавших в обучении. Метод ранней остановки позволяет не допускать переобучения, останавливая процесс обучения нейронной сети, если за заданное количество эпох значение функции потерь, вычисленной на контрольной выборке, не уменьшается. По причине малого объема исходной выборки параметр «заданное количество эпох» равен 6.

По тестовой выборке оценивается качество классификации. При использовании кросс-валидации исходная выборка разбивается на несколько частей [11-13]. Например, при заданном объеме тестовой выборки 10 % исходная выборка разбивается на десять частей. Девять из них используются для обучения и контроля, а десятая для тестирования. Последовательно перебираются все десять вариантов, причем наблюдения в частях, используемых для обучающей и контрольной выборок, случайным образом перемешиваются и разделяются на две выборки в пропорции 8:1. Таким образом, итоговое соотношение выборок для обучения, контроля и тестирования, используемых в каждом варианте сочетаний частей исходной выборки, равно 8:1:1. Среднее по тестовым выборкам значение критерия качества диагностики характеризует точность алгоритма.

Стоит отметить, что не все функции обучения используют в своих алгоритмах контрольную выборку. Поэтому для алгоритма байесовской регуляризации, которая не применяет выборку для контроля, использовались только обучающая и тестовая выборки, соответственно, итоговое соотношение этих выборок составляло 9:1.

Для оценки эффективности диагностики использовались три разных критерия качества: средняя ошибка на тестовой выборке, AUC (area

under ROC curve) – площадь под ROC-кривой (кривой ошибок) и F-мера [14,15].

Кривая ошибок показывает зависимость количества верно классифицированных исправных объектов от количества неверно классифицированных неисправных объектов, поэтому площадь под кривой отражает долю объектов, верно распознанных нейросетью. Этот критерий эффективности является одним из наиболее популярных в задачах бинарной классификации.

F-мера вычисляется на основе двух метрик: точности и полноты. Точность (Precision) – процент верно определенных объектов одного класса среди всех объектов, отнесенных системой к этому классу; полнота (Recall) – процент верно определенных объектов одного класса среди всех объектов этого класса в тестовой выборке:

$$Precision = \frac{TP}{TP + FP},$$

$$Recall = \frac{TP}{TP + FN}$$

где TP (true-positive) – количество истинно-положительных решений (количество объектов 1-го класса, отнесенных к 1-му классу); FP (false-positive) – ложноположительное решение (количество объектов 2-го класса, отнесенных к 1-му классу); FN (false-negative) – ложноотрицательное решение (количество объектов 1-го класса, отнесенных ко 2-му классу). В общем случае F-мера определяется по формуле

$$F = \frac{(\beta^2 + 1) * Precision * Recall}{\beta^2 * Precision + Recall},$$

где β – коэффициент, задающий приоритет точности (при $0 < \beta < 1$) или полноты (при $\beta > 1$).

При $\beta = 1$ она сводится к гармоническому среднему между точностью и полнотой и называется F1-мерой, или сбалансированной F-мерой.

Именно такая мера является наиболее информативным показателем качества при несбалансированных классах (когда работоспособных состояний агрегата значительно больше, чем неработоспособных): чем ближе значение F к единице, тем качество классификации выше.

Практическая реализация поставленной задачи проводилась в среде Neural Network Toolbox пакета MATLAB, а также в этом же пакете на базе библиотеки инструментов Statistics and Machine Learning Toolbox. При этом в связи с недостаточным объемом выборки не использовался активно развиваемый в последние годы подход глубокого обучения.

Цель исследования – повышение качества диагностики агрегата за счет разработки конфигурации нейронной сети и анализа методов ее обучения.

Вначале с помощью библиотеки машинного обучения Statistics and Machine Learning Toolbox была автоматически построена нейронная сеть. Она состоит из одного скрытого слоя с 10 нейронами, использует в качестве функции активации *tansig* (гиперболический тангенс) и обучается с помощью функции обучения *trainlm*. Функция обучения *trainlm* реализует простой и эффективный алгоритм, использующий идею доверительной области и являющийся частным случаем метода градиентного спуска. Алгоритм Левенберга-Марквардта с функцией обратного распространения ошибки предназначен для оптимизации параметров нелинейных регрессионных моделей путем итеративного аппроксимирования поверхностью второго порядка и решения системы линейных уравнений для обновления оценки.

Как и квазиньютоновские методы оптимизации, алгоритм Левенберга-Марквардта был разработан для достижения высокой скорости обучения без вычисления матрицы Гессе – симметрической квадратичной формы *n*-го порядка, элементами которой являются частные производные целевой функции второго порядка. Для этого представим функцию ошибки следующим образом

$$e = d - o = \begin{bmatrix} e_{11} \\ \vdots \\ e_{m1} \\ e_{12} \\ \vdots \\ e_{mp} \end{bmatrix},$$

где *o* – полученный выход нейросети, *d* – ожидаемый (учебный) выход сети, *m* – количество выходов сети, *p* – количество наблюдений в обучающей выборке.

Чтобы избежать вычисления гессиана, алгоритм подразумевает его аппроксимирование с помощью якобиана, который может быть вычислен с помощью обычного метода обратного распространения ошибки, который гораздо более прост в вычислении, чем сам гессиан.

$$H \approx J^T \cdot J + \mu \cdot I,$$

где *J* – матрица Якоби, которая содержит отношения первых производных ошибок сети к весам и смещениям, μ – параметр, отвечающий за точность и скорость схождения, а $I = (J^T \cdot J) \circ E$ – диагональная матрица из элементов главной диагонали $(J^T \cdot J)$, поэлементно умноженная на *E*.

Градиент может быть представлен следующим образом

$$g = J^T \cdot e,$$

где *e* – вектор ошибок нейронной сети.

Для получения изменения весов сети используется следующий квазиньютоновский способ аппроксимирования матрицы Гесса:

$$\Delta W = (J^T \cdot J + \mu \cdot I)^{-1} \cdot J^T \cdot e.$$

Особенность алгоритма Левенберга-Марквардта заключается в том, что алгоритм превращается в метод Ньютона, аппроксимирующий гессиан, когда $\mu = 0$, но чем больше становится μ , тем больше алгоритм становится похожим на метод градиентного спуска с маленьким шагом. Идея заключается в том, чтобы как можно быстрее переключиться на метод Ньютона, так как метод Ньютона быстрее и более точен вблизи минимума ошибки. По этой причине μ уменьшается после каждого шага, который привел к уменьшению функции качества обучения, и увеличивается только тогда, когда пробный шаг увеличивает функцию качества обучения. Таким образом, функция качества обучения всегда будет уменьшаться с каждой итерацией *trainlm*.

Критерии качества построенной с помощью Statistics and Machine Learning Toolbox сети: процент ошибок – 5,95, AUC равна 0,848, а F-мера – 0,775.

При разработке сети в среде Neural Network Toolbox учитывалось, что количество скрытых слоев в нейронной сети и нейронов в каждом слое определяет количество связей между нейронами соседних слоев, что, в свою очередь, определяет обобщающую способность нейросети. Если обобщающая способность будет мала, средняя ошибка прогноза будет велика. Вместе с этим, если обучающая способность будет больше требуемой, избыточные способности могут привести к переобучению, что обусловит высокую среднюю ошибку прогноза.

По этой причине количество скрытых слоев и нейронов в них было выбрано с некоторым запасом. Стояла задача определения наилучшего количества скрытых слоев и нейронов в каждом из слоев путем перебора различных сочетаний. Для перебора были заданы ограничения на количество слоев в нейросети (от 1 до 4) и на количество нейронов (от 5 до 20). Таким образом, количество возможных сочетаний параметров с учетом четырех функций обучения (*trainlm* – алгоритм Левенберга-Марквардта, *trainbr* – алгоритм байесовской регуляризации, *trainrp* – алгоритм эластичного обратного распространения ошибки и *trainscg* – алгоритм масштабируемых сопряженных градиентов) составляло 256.

Программа работала по следующему принципу: вначале исходная выборка разбивалась на 10 кросс-валидационных подвыборок, после чего для каждого сочетания этих подвыборок обучалась нейронная сеть, конфигурация которой была сформирована в соответствии с тремя циклами перебора параметров сети – количества нейронов в слое, количества скрытых слоев и функции обучения. Нейронная сеть обучалась на 9 подвыборках. После этого производился тест на оставшейся подвыборке, а критерии

качества усреднялись. На выходе программы получалась таблица данных с конфигурациями нейронных сетей и соответствующими значениями критериев качества. В таблице приведены полученные результаты для десяти вариантов сетей с сортировкой по возрастанию значений средней ошибки.

Наилучшей подобранной конфигурацией является нейронная сеть из 3 слоев с 18 нейронами в каждом слое (рис. 1).

В качестве функции обучения в ней используется `trainbr` – алгоритм байесовской регуляризации. Эта функция использует алгоритм оптимизации Левенберга-Марквардта для обновления сдвигов и весов нейронов, но за счет того, что функция качества обучения сети минимизирует не только линейную комбинацию ошибок, но и линейную комбинацию весов и смещений, обученная нейросеть хорошо справляется с задачей обобщения. Такой алгоритм называется алгоритмом байесовской регуляризации и подразумевает оптимизацию модифицированной функции качества обучения для обеспечения свойства генерализации.

Обычно функцией качества обучения для нейронных сетей прямого распространения выбирают среднюю сумму квадратов ошибок:

$$F = mse = \frac{1}{N} \sum_{i=1}^N (e_i)^2 = \frac{1}{N} \sum_{i=1}^N (t_i - a_i)^2.$$

Для улучшения выраженности свойства генерализации можно модифицировать функцию

качества обучения, добавив часть, содержащую среднюю сумму квадратов весов и смещений сети:

$$mse_{reg} = \gamma \cdot mse + (1 - \gamma) \cdot msw,$$

где mse_{reg} – модифицированная функция качества нейронной сети с учетом регуляризации, γ – коэффициент качества, mse – средняя сумма квадратов ошибок, msw – средняя сумма квадратов весов сети:

$$msw = \frac{1}{n} \sum_{j=1}^n w_j^2.$$

Использование такой функции качества обучения нейронной сети позволяет уменьшить значения параметров (весов и смещений) сети, что снижает вероятность ее переобучения. В использовании такой модифицированной функции качества и заключается метод байесовской регуляризации.

Процент средней ошибки распознавания состояния агрегата с помощью подобранной нейронной сети равен 4,85, AUC равна 0,885, а F-мера – 0,827.

На рис. 2 показана отображаемая для каждой нейросети информация на примере сети, показавшей наилучшее значение по F-мере: Data Division: Random – случайный способ деления данных на выборки для обучения, контроля и теста (поскольку алгоритм байесовской регуляризации не использует метод ранней остановки, а из-за применения метода кросс-валидации проверка на тестовой выборке производится

Таблица. Отсортированный по увеличению средней ошибки список нейросетей

Функция обучения	Количество слоев в нейросети	Количество нейронов в слое	Средняя ошибка, %	AUC	F-мера
trainbr	3	18	4,85	0,885	0,827
trainlm	2	8	4,90	0,888	0,828
trainlm	3	20	4,90	0,893	0,828
trainbr	4	10	4,95	0,874	0,818
trainlm	2	12	4,95	0,890	0,823
trainbr	3	6	5,00	0,883	0,819
trainbr	3	12	5,05	0,876	0,816
trainbr	2	18	5,05	0,882	0,822
trainbr	3	17	5,05	0,891	0,826
trainlm	3	14	5,10	0,876	0,816

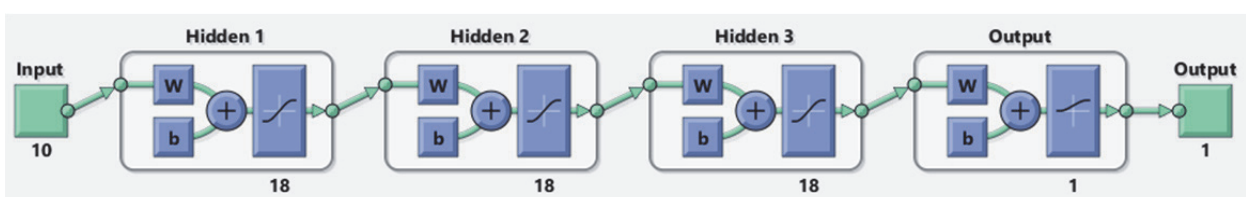


Рис. 1. Архитектура подобранной нейронной сети

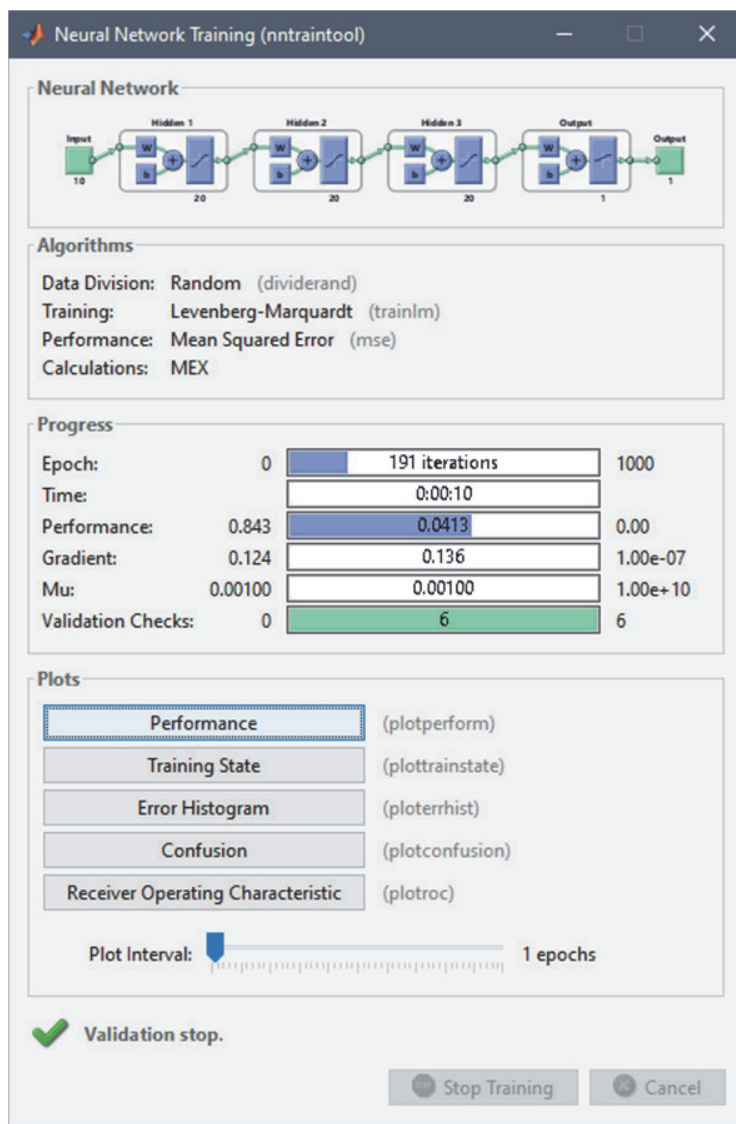


Рис. 2. Обучение нейронной сети

отдельно, все 9 частей исходной выборки используются в обучающей выборке); Training: Bayesian Regularization – в качестве функции обучения выбран алгоритм байесовской регуляризации; Performance: Mean Squared Error – из-за особенностей регуляризации в качестве функции потерь используется среднеквадратическая ошибка; Calculations: MEX – выбранная вычислительная модель, позволяющая эффективно подключать написанные на C++ программы и использовать их с экономией памяти.

На рис. 2 также отражена информация о процессе обучения нейросети: Epoch – количество эпох; Time – время, затраченное на обучение; Performance – значение функции потерь (среднеквадратической ошибки); Gradient – значение градиента; Mu – значение параметра, используемого в формуле обновления весов нейронов и препятствующего застреванию нейросети в локальных минимумах; Effective #Param – количество эффективно настроенных параметров нейро-

сети; Sum Squared Param – остаточная сумма квадратов.

Из окна с информацией можно открыть график функции потерь (рис. 3), из которого видно уменьшение ошибки во время процесса обучения на каждой эпохе. На графике кружком отмечена точка минимума, соответствующая эпохе, в которой ошибка была меньше всего.

Разработанная методика диагностики работоспособности агрегата с использованием нейронных сетей обеспечивает повышение на 6,7% значения F-меры (являющейся наиболее информативным показателем качества бинарной классификации при несбалансированных классах) в рассматриваемой выборке.

СПИСОК ЛИТЕРАТУРЫ

1. Биргер, И.А. Техническая диагностика / И.А. Биргер. М.: Машиностроение, 1978. 240 с. (2-е изд.: М.: URSS, 2019).

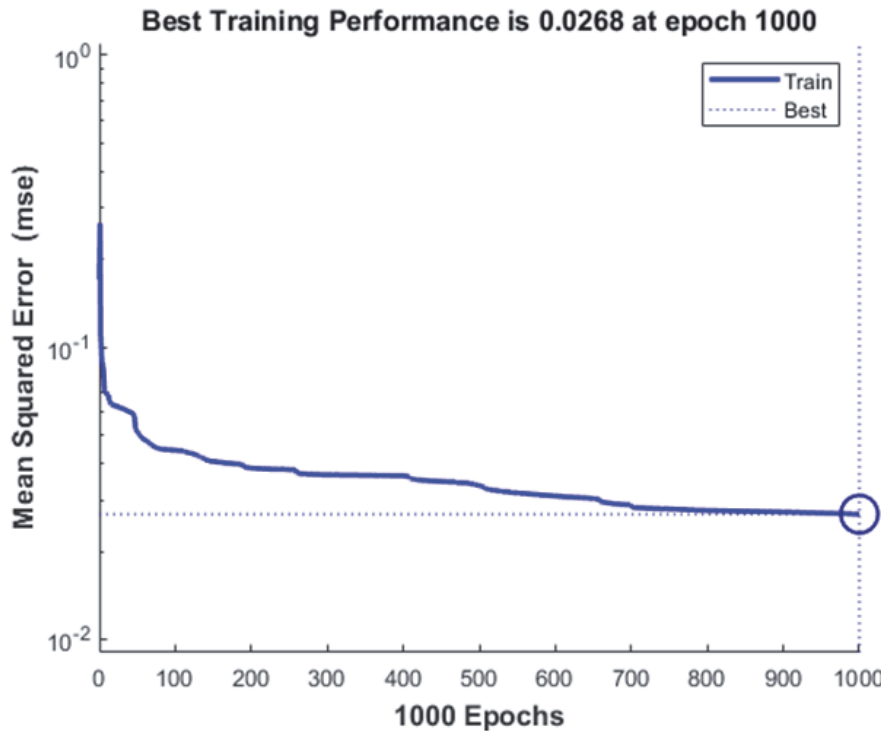


Рис. 3. Процесс обучения сети

2. Клячкин, В.Н. Прогнозирование и диагностика стабильности функционирования технических объектов: монография / В.Н. Клячкин, В.Р. Крашенинников, Ю.Е. Кувайскова. М.: РУСАЙНС, 2020. 200 с.
3. Wyner A. J. et al. Explaining the success of adaboost and random forests as interpolating classifiers // The Journal of Machine Learning Research. 2017. Т. 18. №. 1. P. 1558-1590.
4. Jose Orozco, Carlos A. Reyes Garcia. Detecting Pathologies from Infant Cry Applying Scaled Conjugate Gradient Neural Networks // European Symposium on Artificial Neural Networks, Bruges (Belgium), 23-25 April 2003, d-side publ. P. 349-354.
5. Санталов А.А., Жуков Д.А. Диагностика технического состояния системы с применением нейросетевых методов // Перспективные информационные технологии: труды Международной научно-технической конференции / под ред. С.А. Прохорова. Самара: Издательство Самарского научного центра РАН, 2018. С. 202-205.
6. Васильев В.И., Жернаков С.В. Классификация режимов работы ГТД с использованием технологии нейронных сетей // Вестник Уфимского государственного авиационного технического университета. 2009. Т. 12(1). С.53–56.
7. Жернаков С.В., Гильманшин А.Т. Применение интеллектуальных алгоритмов на основе нечеткой логики и нейронных сетей для решения задач диагностики отказов авиационного ГТД // В сборнике: Intelligent Technologies for Information Processing and Management (ITIPM'2014) Proceedings of the 2nd International Conference. 2014. С. 112–115.
8. Клячкин В.Н., Кувайскова Ю.Е., Жуков А.А. Выбор метода бинарной классификации при технической диагностике с применением машинного обучения // Известия Самарского научного центра РАН, 2018. Т. 20. №4-3(84). С. 494-497.
9. Воронина, В.В. Теория и практика машинного обучения: учеб. пособие / В.В. Воронина, А.В. Михеев, Н.Г. Ярушкина, К.В. Святлов. Ульяновск: УлГТУ, 2017. 290 с.
10. Witten I.H., Frank E. Data mining: practical machine learning tools and techniques. SF: Morgan Kaufmann Publ., 2005. 525 p.
11. 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.
12. Санталов А.А., Клячкин В.Н. Разработка нейронной сети для оценки исправности гидроагрегата по результатам вибромониторинга // Программные продукты и системы. 2020. Т. 33. №4. С.561-566.
13. Применение агрегированных классификаторов при машинном обучении для оценки стабильности функционирования технических объектов / А.В. Алексеева, Д.А. Жуков, В.Н. Клячкин, А.А. Санталов // Материалы XIII Международной конференции по прикладной математике и механике в аэрокосмической отрасли, Алушта. М.: МАИ, 2020. С. 622-624.
14. Davis J., Goadrich M. The relationship between Precision-Recall and ROC curves / Proceedings of the 23rd international conference on Machine learning. Pittsburgh. 2006. P. 233-240.
15. Hand D.J., Till R.J. A simple generalisation of the area under the ROC curve for multiple class classification problems // Machine learning. 2001. Vol. 45. №. 2. P. 171–186.

**ANALYSIS OF MATHEMATICAL MODELS IN THE DESIGN
OF NEURAL NETWORK FOR ESTIMATING THE UNIT PERFORMANCE**

© 2021 A. A. Santalov

Ulyanovsk State Technical University, Ulyanovsk, Russia

Prevention of emergencies in aviation technology is largely ensured by diagnostics of the functioning of its units. Often the performance criterion is the level of vibration that influences the decision to adjust the load or shut down the unit. The article discusses mathematical models when applying neural network methods for vibration diagnostics. When using cross-validation, the initial data set with vibration data is divided into several blocks, which are grouped into three sets: training, validation, and test. To assess the effectiveness of diagnostics, three different quality criteria were used: mean error in the test set, AUC, and F-measure. For a given set of initial data, the best fitted configuration turned out to be a neural network of three layers with 18 neurons in each layer, implemented in the MATLAB package. It uses a Bayesian regularization algorithm as a learning function. The percentage of the average error in recognizing the state of the considered aggregate using the neural network turned out to be 4.85, the AUC value was 0.885, and the F-measure was 0.827. Compared to a network built in automatic mode using the Statistics and Machine Learning Toolbox and Neural Network Toolbox machine learning libraries, the F-measure of the fitted network configuration is 6.7% higher.

Keywords: technical diagnostics, binary classification, neural network, quality measures, vibration diagnostics.

DOI: 10.37313/1990-5378-2021-23-1-84-90