

МАШИННАЯ КЛАССИФИКАЦИЯ РЕЖИМА РАБОТЫ ЭЛЕКТРИЧЕСКОЙ РЕЛЬСОВОЙ ЦЕПИ МЕТОДОМ ОПОРНЫХ ВЕКТОРОВ

Системы мониторинга электрических рельсовых цепей, применяемые в настоящее время на железнодорожном транспорте, не обладают интеллектуальными функциями, что существенно снижает их потенциал. Эффективность этих систем может быть существенно повышена за счет реализации функций автоматического анализа собираемой информации. В рамках реализации этих функций ранее мы предложили автоматизировать классификацию режима работы электрической рельсовой цепи на основе логистической регрессии. Однако такой классификатор обладает рядом ограничений. В этой статье мы предлагаем более совершенный алгоритм классификации, построенный на основе метода опорных векторов. В статье представлена теоретическая база алгоритма, продемонстрирована его работа на синтезированных данных. Также показано, что применение гауссовского ядра позволяет уменьшить размерность пространства признаков при сохранении качества классификации состояния электрической рельсовой цепи.

Ключевые слова: железнодорожная автоматика, электрическая рельсовая цепь, машинное обучение, классификация, логистическая регрессия, метод опорных векторов.

Введение. В рамках стратегии холдинга ОАО «РЖД» до 2020 г. одним из ключевых проектов определена высокотехнологичная платформа «Цифровая железная дорога» [1].

Для инфраструктурного комплекса корпорации, к которому принадлежат устройства железнодорожной автоматики и телемеханики, приоритетной задачей является повышение потенциала систем технической диагностики и мониторинга (СТДМ) за счет расширения их функциональных возможностей по обработке диагностической информации, поступающей на верхний уровень СТДМ от низовых устройств [2–4]. Применяемые в настоящее время СТДМ не обладают интеллектуальными функциями, а лишь контролируют нахождение параметров объекта железнодорожной автоматики в допустимых пределах.

Таким образом, актуальна проблема разработки интеллектуальных систем, которые способны самостоятельно выявлять неочевидные человеку связи и зависимости внутри диагностических данных и, как следствие, оперативно информировать о реальном состоянии объекта автоматики и телемеханики.

Построение интеллектуальных систем обработки диагностических данных устройств железнодорожной автоматики и телемеханики может быть выполнено с применением методов машинного обучения.

В [5] представлено решение задачи классификации режима работы электрической рельсовой цепи (как объекта железнодорожной автоматики и теле-

механики) на основе логистической регрессии. Несмотря на успешное выполнение поставленной задачи, логистическая регрессия имеет существенное ограничение: для безошибочной классификации данные должны быть линейно разделимы [6]. Существуют альтернативные алгоритмы машинного обучения, не имеющие такого ограничения, которые способны решать подобные задачи достаточно эффективно. Одним из таких алгоритмов является рассматриваемый в данной статье метод опорных векторов (Support Vector Machine, далее SVM).

Основные понятия и используемая терминология. Рельсовая цепь работает в трех основных режимах — нормальном, шунтовом и контрольном [7], которые мы также будем называть классами. О режиме работы рельсовой цепи можно судить по параметрам электрических сигналов на ее входе и выходе. К таким параметрам относятся амплитуды и фазы напряжений и токов на входе и выходе рельсовой цепи. Значения всех или любой совокупности этих параметров, полученные в один и тот же момент времени, образуют вектор признаков.

SVM относится к группе алгоритмов обучения «с учителем», следовательно, каждый вектор признаков должен быть дополнен идентификатором соответствующего режима работы рельсовой цепи (ответом). Дополненный ответом вектор признаков называется обучающим примером, а совокупность нескольких обучающих примеров — выборкой.

Для построения классификаторов требуются обучающая и тестовая выборки [8]. Такие выборки синтезированы на основе [7] по цифровой модели

рельсовой цепи, описанной в [9]. Их структура, состав, а также особенности генерации по цифровой модели рельсовой цепи аналогичны тем, что использованы в [5] для построения классификатора на основе логистической регрессии.

Система классификации режима работы рельсовой цепи. Поскольку рельсовая цепь имеет три режима работы, а SVM выполняет бинарное разделение пространства признаков [8], преобразуем бинарную задачу в мультиклассовую по аналогии с тем, как это сделано в [5]. Исследования на данную тему подробно описаны в [10] и рассматривают несколько вариантов такого преобразования («один против всех», «один против одного», «DAGSVM»).

Для классификации режима работы рельсовой цепи выберем метод «один против всех», возможности которого вполне удовлетворяют требованиям решаемой задачи. Выбор данного метода обусловлен относительно простой его программной реализации по сравнению с остальными, а также наиболее частым и успешным применением в работах других авторов, например, [8, 11].

Поясним, что метод «один против всех» обеспечивает принятие решения о принадлежности вектора признаков к соответствующему классу по наибольшему значению решающей функции для этого вектора.

Таким образом, необходимо построить систему машинной классификации, состоящую из трех самостоятельных структур (классификаторов), по одной для каждого режима работы рельсовой цепи. Каждый классификатор, в свою очередь, решает бинарную задачу отделения одного из классов от двух остальных в пространстве признаков.

Классификация режима работы рельсовой цепи методом опорных векторов. SVM был впервые представлен В. Н. Вапником [12] как алгоритм, способный выполнять бинарную классификацию данных в пространстве признаков и обеспечивающий при этом максимальный интервал между векторами разделяемых классов и границей принятия решения. Свойство разграничивать данные разных классов максимальным интервалом, тем самым обеспечивая их «уверенное» разделение, дает классификатору на основе SVM преимущество перед многими методами машинной классификации [6, 8, 11, 13].

Граница принятия решения, которая разделяет классы в пространстве признаков, представляется гиперплоскостью и описывается формулой (1). Решение о принадлежности вектора к классу принимается на основе ступенчатой решающей функции, которая описывается формулой (2).

$$z(\mathbf{x}) = \mathbf{w}^T \mathbf{x} + b, \quad (1)$$

$$h(\mathbf{x}) = \begin{cases} 1, & \text{при } z(\mathbf{x}) \geq 0 \\ -1, & \text{при } z(\mathbf{x}) < 0 \end{cases}, \quad (2)$$

где \mathbf{w} — вектор весовых коэффициентов, определяющих наклон границы принятия решения в пространстве признаков; b — параметр, определяющий смещение границы принятия решения относительно начала координат; \mathbf{x} — вектор признаков, подлежащий классификации.

Вектор \mathbf{w} и параметр b могут быть найдены в результате решения оптимизационной задачи (3), что и является обучением классификатора.

$$\min_{\mathbf{w}, b \in \{\xi^{(i)}\}} C \sum_{i=1}^m \xi^{(i)} + \frac{1}{2} \|\mathbf{w}\|^2,$$

$$\text{при } y^{(i)} (\mathbf{w}^T \mathbf{x}^{(i)} + b) \geq 1 - \xi^{(i)}, \quad i = 1 \dots m, \quad (3)$$

$$\xi^{(i)} \geq 0, \quad i = 1 \dots m,$$

где $\mathbf{x}^{(i)}$ — i -й обучающий пример из обучающей выборки; $y^{(i)}$ — идентификатор класса при i -м обучающем примере, равный 1 для примеров относящихся к выделяемому классу, и -1 для всех остальных примеров; m — размер обучающей выборки; C — параметр регуляризации, определяющий увеличение целевой функции для обучающих примеров вблизи границы принятия решения (при решении поставленной задачи принят равным 100); $\xi^{(i)}$ — переменная ослабления, являющаяся мерой ошибки, допускаемой при работе классификатора.

На сегодняшний день для решения задачи оптимизации при линейных ограничениях в SVM применяется алгоритм последовательной минимальной оптимизации (Sequential Minimal Optimization, далее SMO). Этот алгоритм предложен Д. Платтом в [14] вместо ранее используемого квадратичного программирования. Существенным достоинством SMO является кардинальное увеличение производительности за счет разбиения решаемой громоздкой оптимизационной задачи на набор небольших подзадач. Сформированные подзадачи разрешимы аналитически, что не требует серьезных затрат времени и вычислительных ресурсов, как было ранее при оптимизации по [12].

В [8] показано, что оптимизационная задача (3) сводится к задаче (4), которая решается с помощью алгоритма SMO [14].

$$\max_{\alpha} \sum_{i=1}^m \alpha_i - \frac{1}{2} \sum_{i=1}^m \sum_{j=1}^m y^{(i)} \alpha_i \alpha_j K(\mathbf{x}^{(i)}, \mathbf{x}^{(j)}),$$

$$\text{при } 0 \leq \alpha_i \leq C, \quad i = 1 \dots m, \quad (4)$$

$$\sum_{i=1}^m \alpha_i y^{(i)} = 0,$$

где α_i — множитель Лагранжа при i -м обучающем примере; $K(\mathbf{x}^{(i)}, \mathbf{x}^{(j)})$ — функция, результат которой равен скалярному произведению векторов i -го и j -го обучающих примеров.

В формулах задачи (4) скалярное произведение векторов приведено в виде отдельной функции. Это сделано для удобства перехода при классификации к пространствам более высокой размерности, что, в свою очередь, позволяет реализовать нелинейную границу принятия решения и повысить точность классификатора. Функция K называется ядром, и в простейшем случае, при отсутствии необходимости перехода к пространству более высокой размерности, определяется формулой (5). Такое ядро называется линейным.

$$K(\mathbf{x}^{(i)}, \mathbf{x}^{(j)}) = (\mathbf{x}^{(j)})^T \cdot \mathbf{x}^{(i)}. \quad (5)$$

Решением задачи (4) является совокупность значений множителей Лагранжа, причем в [8] показано, что лишь для нескольких обучающих примеров соответствующий множитель Лагранжа отличается от нуля.

Такие обучающие примеры называются опорными векторами. После решения задачи (4) весовые коэффициенты \mathbf{w} определяются по формуле (6).

$$\mathbf{w} = \sum_{i \in S} \alpha_i y^{(i)} \mathbf{x}^{(i)}, \quad (6)$$

где S — множество номеров обучающих примеров, для которых множители Лагранжа, полученные при решении задачи (4), больше нуля.

При подстановке (6) в (1) с учетом обозначения (5) получаем выражение (7), описывающее границу принятия решения.

$$z(\mathbf{x}) = \left(\sum_{i \in S} \alpha_i y^{(i)} \mathbf{x}^{(i)} \right)^T \mathbf{x} + b = \\ = \sum_{i \in S} \alpha_i y^{(i)} (\mathbf{x}^{(i)})^T \mathbf{x} + b = \sum_{i \in S} \alpha_i y^{(i)} K(\mathbf{x}^{(i)}, \mathbf{x}) + b. \quad (7)$$

В [8] также показано, что параметр b , фигурирующий в формулах (1) и (7), определяется на основе выявленных опорных векторов по формуле (8):

$$b = \frac{1}{N_M} \sum_{i \in M} \left(y^{(i)} - \sum_{j \in S} \alpha_j y^{(j)} K(x^{(i)}, x^{(j)}) \right), \quad (8)$$

где M — множество номеров обучающих примеров, для которых, полученные в результате решения задачи (4) множители Лагранжа удовлетворяют неравенству $0 < \alpha_i < C$; N_M — число элементов в множестве M .

Таким образом, обучение классификатора заключается в определении множителей Лагранжа α_i для обучающей выборки путем решения задачи (4) с помощью алгоритма SMO. После этого по формуле (8) может быть определен параметр b . Классификация произвольного вектора \mathbf{x} в пространстве признаков заключается в расчете формулы (7) и подстановке полученного значения в формулу (2). Максимальное значение ступенчатой функции $h(\mathbf{x})$, полученное по формуле (2), для одного из классов свидетельствует о принадлежности вектора \mathbf{x} к этому классу.

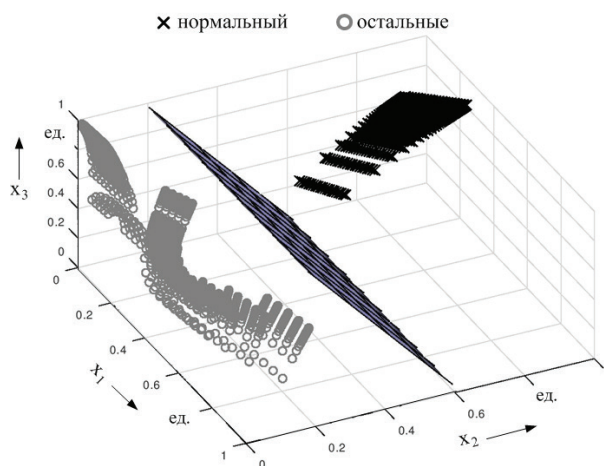
Демонстрация результатов работы SVM на примерах. Продемонстрируем классификацию режима работы рельсовой цепи с помощью SVM для нескольких случаев в вычислительной среде Octave. Для этого зададимся трехмерным пространством признаков, образованным следующими параметрами рельсовой цепи: x_1 — модуль тока на входе рельсовой цепи, x_2 — модуль напряжения на выходе рельсовой цепи, x_3 — фаза напряжения на выходе рельсовой цепи, следовательно, $\mathbf{x} = \{x_1, x_2, x_3\}$.

Поскольку выбранные для обучения параметры имеют различные диапазоны изменения, необходимым этапом является их нормализация, то есть приведение к единому интервалу [6]. Выполним нормализацию в интервале [0,1], так же как это сделано в [5].

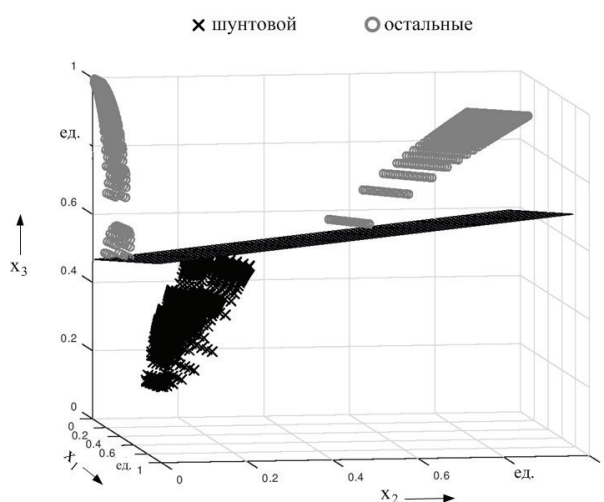
После обучения классификаторов, согласно (1) — (8), для каждого режима работы рельсовой цепи получены следующие гиперплоскости (рис. 1), являющиеся границами принятия решений по принципу «один против всех».

На рис. 1 видно, что полученные границы принятия решений являются плоскостями и расположены, во-первых, на максимальном, а во-вторых, на одинаковом расстоянии от ближайших друг к другу векторов выделяемого и остальных классов. Данное положение границ отличает SVM от логистической регрессии из [5], при которой они устанавливались произвольно в пространстве между классами.

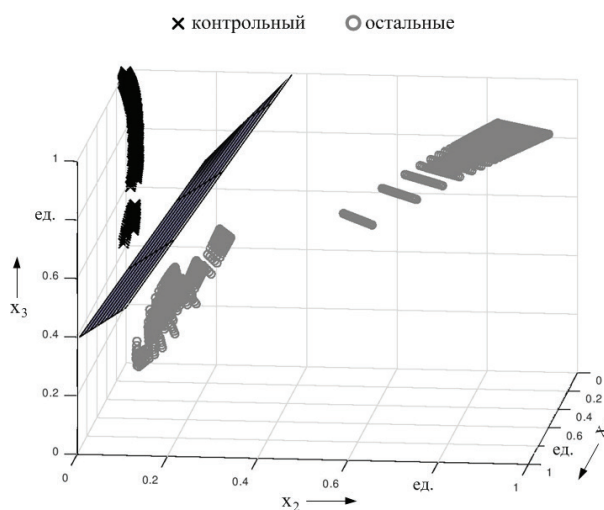
Точность классификации как на обучающей, так и на тестовой выборках составила 100 %, что объясняется линейной разделительностью классов в пространстве трех признаков.



а)



б)



в)

Рис. 1. Границы принятия решений для режимов работы рельсовой цепи в пространстве трех признаков: а) нормальный; б) шунтовой; в) контрольный

Классификация режима работы рельсовой цепи в пространстве трех признаков системой трех классификаторов может быть визуализирована при помощи развертки этого пространства (рис. 2).

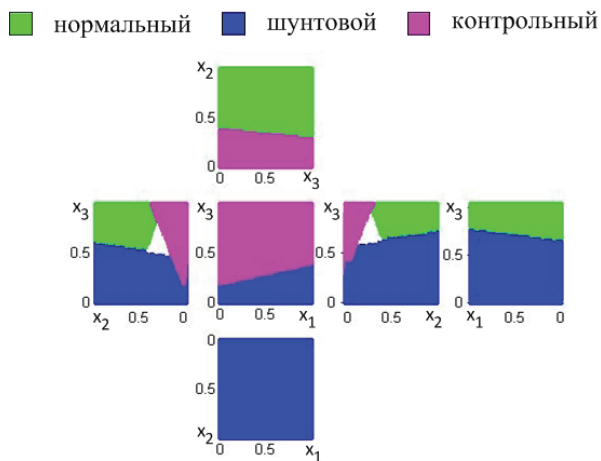


Рис. 2. Развертка трехмерного пространства признаков

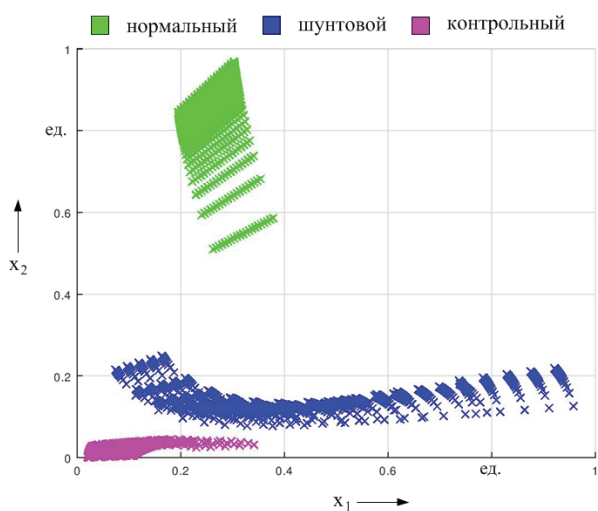


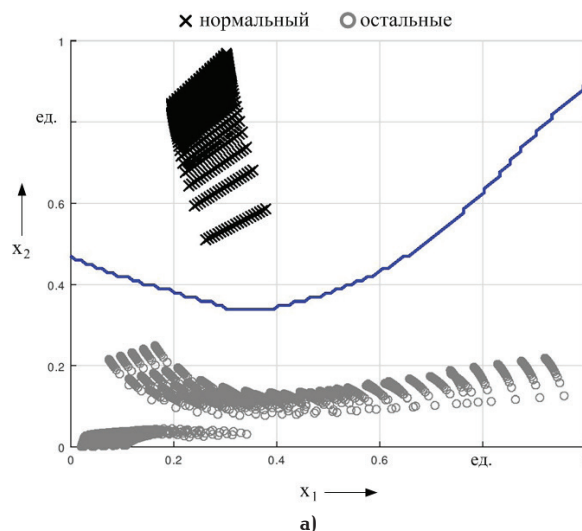
Рис. 3. Визуализация обучающей выборки в пространстве двух признаков

Построение пространства на рис. 2 реализовано за счет генерации векторов признаков, равномерно распределенных внутри него в диапазоне от 0 до 1 с постоянным интервалом. Визуализация границ принятия решений обеспечена различной подсветкой векторов, а выбор оттенка подсветки — решением классификатора.

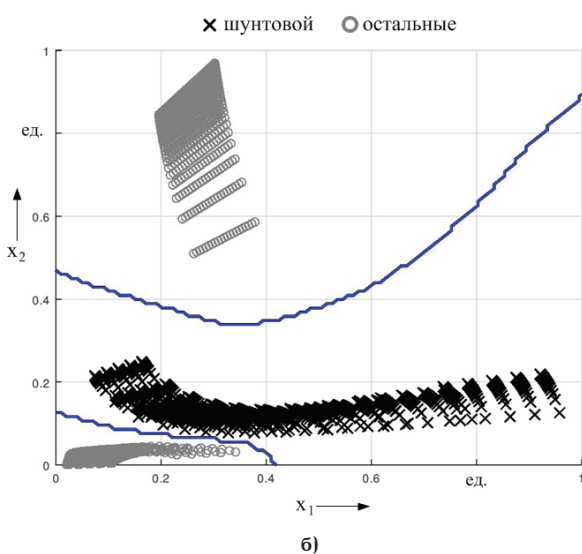
С точки зрения физической реализации системы машинной классификации, сбор информации о параметрах электрических сигналов рельсовой цепи осуществляется с помощью датчиков различного назначения. Число требуемых датчиков определяется числом используемых для классификации признаков.

Очевидно, что желательным является снижение числа датчиков, так как это позволит уменьшить затраты на реализацию и функционирование системы машинной классификации.

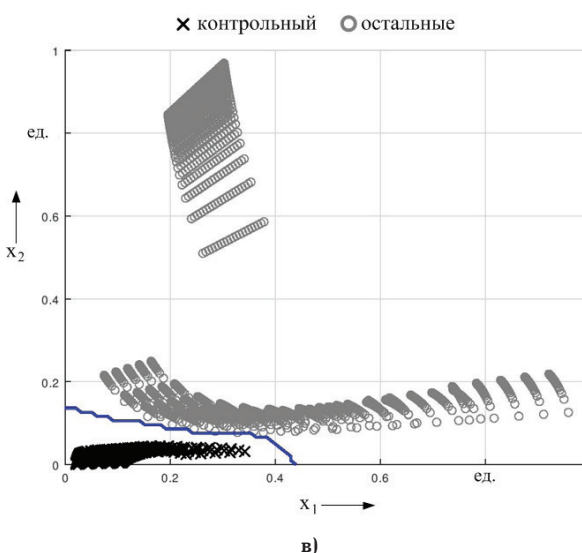
Уменьшение числа необходимых датчиков может быть достигнуто за счет уменьшения числа признаков, по которым выполняется классификация. Следствием этого может стать ухудшение разделимости классов. Так, например, если исключить из вектора \mathbf{x} признак x_3 (фаза напряжения на выходе рельсовой цепи), то обучающая выборка может быть визуализирована в виде рис. 3.



а)



б)



в)

Рис. 4. Границы принятия решений для режимов работы рельсовой цепи в пространстве двух признаков: а) нормальный; б) шунтовой; в) контрольный

Как видно из рис. 3, при снижении размерности пространства признаков разделимость классов сохранилась, однако стала нелинейной. А именно класс, соответствующий шунтовому режиму работы рельсовой цепи, не может быть отделен от осталь-

ных классов одной прямой линией. Следовательно, для сохранения прежнего качества классификации границы принятия решений должны стать некоторыми кривыми.

Для достижения указанного эффекта можно использовать прием преобразования исходного двумерного пространства признаков, в котором линейная разделимость классов сохраняется. Тогда классификация при линейной разделимости в многомерном пространстве приведет к получению нелинейной границы принятия решения в исходном (двумерном) пространстве признаков. Такой прием можно выполнить путем замены линейного ядра (5) на гауссовское, имеющее вид (9):

$$K(x^{(i)}, x^{(j)}) = \exp\left(-\frac{\|x^{(i)} - x^{(j)}\|^2}{2\sigma^2}\right), \quad (9)$$

где σ — параметр сглаживания (в данном исследовании выбран равным 0,03 для наилучшей обобщающей способности классификатора).

Границы принятия решения в пространстве двух признаков, формируемые классификаторами, использующими гауссовское ядро (9), приведены на рис. 4.

На рис. 4 видно, что границами принятия решений являются кривые. Точность работы системы машинной классификации по-прежнему составила 100 % как на обучающей, так и на тестовой выборках.

Расчет гауссовского ядра (9) для одного опорного вектора по вычислительной сложности сопоставим с расчетом сигмоидальной функции и линейного полинома при логистической регрессии из [5]. Однако для выполнения классификации используется несколько опорных векторов, что приводит к пропорциональному увеличению времени вычислений по сравнению с логистической регрессией. Это обстоятельство не является критичным с учетом уменьшающейся стоимости вычислительных ресурсов и невысокой скорости изменения режимов работы рельсовых цепей.

Заключение. Таким образом, удалось показать, что SVM может применяться для классификации режима работы рельсовой цепи как альтернатива логистической регрессии. В сочетании с гауссовским ядром этот метод позволяет снизить число признаков, по которым выполняется классификация, что обеспечивает снижение стоимости системы классификации. Следует, однако, отметить, что дальнейшему исследованию подлежит вопрос об оценке вычислительной эффективности SVM и логистической регрессии.

Библиографический список

1. Чаркин Е. И. Новая технологическая реальность // Автоматика, связь, информатика. 2018. № 1. С. 2–5.
2. Насонов Г. Ф., Сусленникова Е. О., Дзюба Ю. В. Развитие информационных технологий в инфраструктурном

комплексе // Автоматика, связь, информатика. 2018. № 1. С. 14–16.

3. Розенберг Е. Н., Дзюба Ю. В., Батраев В. В. О направлениях развития цифровой железной дороги // Автоматика, связь, информатика. 2018. № 1. С. 9–13.
4. Урусов А. В. Цифровая железная дорога // Автоматика, связь, информатика. 2018. № 1. С. 6–8.
5. Борисенко Д. В., Присухина И. В., Лунёв С. А. Машинная классификация режима работы электрической рельсовой цепи на основе логистической регрессии // Омский научный вестник. 2018. № 4 (160). С. 67–72. DOI: 10.25206/1813-8225-2018-160-67-72.
6. Harrington P. Machine learning in action. NY: Manning Publications, 2012. 354 p. ISBN 1617290181; 9781617290183.
7. Брылеев А. М., Кравцов Ю. А., Шишляков А. В. Теория, устройство и работа рельсовых цепей. 2-е изд., перераб. и доп. М.: Транспорт, 1978. 344 с.
8. Bishop C. M. Pattern recognition and machine learning. NY: Springer, 2010. 738 p. ISBN 978-0-387-31073-2.
9. Борисенко Д. В., Присухина И. В., Лунёв С. А. Математическая модель рельсовой цепи для генерации обучающей выборки при решении задач машинной классификации // Известия Транссиба. 2017. № 4 (32). С. 111–121.
10. Hsu C.-W., Lin C.-J. A comparison of methods for multiclass support vector machines // IEEE Transactions on Neural Networks. 2002. Vol. 13, Issue 2. P. 415–425. DOI: 10.1109/72.991427.
11. Murphy K. P. Machine learning: a probabilistic perspective. Cambridge, Massachusetts: MIT Press, 2012. 1067 p. ISBN 0262018020; 978-0262018029.
12. Вапник В. Н. Восстановление зависимостей по эмпирическим данным. М.: Наука, 1979. 448 с.
13. Hastie T. The elements of statistical learning. 2nd ed. NY: Springer, 2009. 745 p. ISBN 0387848576, 978-0387848570.
14. John C. Platt. Fast training of support vector machines using sequential minimal optimization // Advances in Kernel Methods: Support Vector Learning. MIT Press. 1998. P. 185–208. ISBN 0262194163; 978-0262194167.

ПРИСУХИНА Илона Вадимовна, аспирантка кафедры «Автоматика и телемеханика».

SPIN-код: 8429-0259

AuthorID (РИНЦ): 907740

Адрес для переписки: prisukhinaiv@omgups.ru

БОРИСЕНКО Дмитрий Владимирович, кандидат технических наук, доцент (Россия), доцент кафедры «Автоматика и телемеханика».

SPIN-код: 9697-9830

AuthorID (РИНЦ): 533908

Адрес для переписки: borisenkodv@yandex.ru

Для цитирования

Присухина И. В., Борисенко Д. В. Машинная классификация режима работы электрической рельсовой цепи методом опорных векторов // Омский научный вестник. 2018. № 6 (162). С. 126–130. DOI: 10.25206/1813-8225-2018-162-126-130.

Статья поступила в редакцию 19.09.2018 г.

© И. В. Присухина, Д. В. Борисенко