

С.А ЗАЙЦЕВ, асп., ЗНТУ, Запорожье,
С.А. СУББОТИН, канд. техн. наук, доц., ЗНТУ, Запорожье

МОДЕЛЬ ОТРИЦАТЕЛЬНОГО ОТБОРА С МАСКИРОВАНИЕМ ДЕТЕКТОРОВ ДЛЯ РЕШЕНИЯ ЗАДАЧ ДИАГНОСТИРОВАНИЯ С ВЕЩЕСТВЕННЫМ ПРЕДСТАВЛЕНИЕМ ПРИЗНАКОВ

Решена актуальная задача автоматизации диагностирования объектов и систем, описанных вещественными признаками. Получил дальнейшее развитие метод обучения модели отрицательного отбора с маскированием, который модифицирован для работы с вещественным представлением детекторов. Решена задача неразрушающего диагностирования лопаток газотурбинных авиадвигателей. Библиогр.: 10 назв.

Ключевые слова: диагностирование, вещественные признаки, отрицательный отбор, маскирование, детектор.

Постановка проблемы и анализ литературы. Техническое диагностирование помимо принятия решения о состоянии диагностируемого объекта или системы предполагает необходимость извлечения и интерпретации знаний, полученных диагностической моделью в процессе обучения. Благодаря простоте представления и организации логического вывода, наиболее удобной формой представления знаний для человека являются продукционные правила вида "если-то" [1].

Модели, основанные на принципе отрицательного отбора в искусственных иммунных системах, позволяют осуществлять диагностирование объектов, однако отличаются высокой сложностью извлечения знаний и низкой интерпретируемостью результатов обучения модели [2, 3].

Решением, лишенным описанных выше недостатков, может быть предложенная авторами модель отрицательного отбора с маскированием [4]. Модель способна обучаться на экземплярах одного класса, характеризуется прозрачностью процесса обучения, устойчива в задачах большой размерности и способна представлять знания в виде продукционных правил. Однако данная модель предназначена для диагностирования объектов, описанных только бинарными (или дискретными) признаками. На практике оказывается, что значительная часть диагностируемых объектов и систем представлена векторами вещественных чисел.

Цель данной работы заключается в модификации модели отрицательного отбора с маскированием и разработке метода ее обучения для работы с вещественными признаками.

Постановка задачи. В пространстве признаков U размерностью n задана обучающая выборка S' . Известно, что каждый экземпляр $x \in U$ можно отнести к классу "своих" (годных) или "чужих" (дефектных). Множества S и N , содержащие "свои" и "чужие" экземпляры соответственно, комплементарны: $S \cap N = \emptyset$, $S \cup N = U$. Необходимо построить на основе обучающей выборки $S' \subset S$ диагностическую модель, описанную набором детекторов D , чтобы на основании этого набора любой экземпляр $x \in U$ можно было однозначно отнести к классу "своих" или "чужих".

Отрицательный отбор. Модель отрицательного отбора, как правило, представляется в виде набора детекторов D . Детектором называют объект, способный определить, принадлежит ли неизвестный экземпляр к "чужому" классу или нет. Для этого к экземпляру x и детектору d применяется правило сопоставления $match(d, x)$, которое возвращает значение "1", если детектор активизировался при сопоставлении с экземпляром x , и "0" – в противном случае. В процессе обучения в модель добавляются толерантные к своим экземплярам детекторы: $\forall s \in S, match(d, s) = 0$.

Поскольку модель отрицательного отбора решает задачу двухклассовой классификации (каждый экземпляр может быть отнесен ко множеству S или N), и $S \cap N = \emptyset$, $S \cup N = U$, то это предполагает, что детекторы будут реагировать на "чужие" экземпляры $ns \in N$.

Первый метод обучения такой модели был описан в [5] и состоял из следующих этапов.

1. Этап генерации детекторов. На этом этапе случайным образом формируется набор кандидатов в детекторы C , как правило, превышающий по объему число экземпляров в обучающей выборке:

$$|C| > |S'|.$$

2. Цензурирование. На данном этапе каждый кандидат в детекторы $c \in C$ сопоставляется с каждым "своим" экземпляром из обучающей выборки $s \in S'$. Если кандидат активизировался (значение правила сопоставления стало равным "1"), то он уничтожается, поскольку не может обеспечить толерантность модели к "своим" экземплярам. Все

оставшиеся кандидаты формируют набор детекторов D , который может быть использован для обнаружения "чужих" экземпляров.

Несмотря на очевидную ресурсоемкость данного метода, он получил дальнейшее развитие, которое заключалось в разработке более оптимальных методов генерации набора детекторов (жадного метода, метода генерации детекторов за линейное время [6]).

Маскирование бинарных детекторов. Основная идея модели отрицательного отбора с маскированием заключается в том, что часть признаков содержат замаскированные значения, т. е. во время применения правила сопоставления значения таких признаков у экземпляра игнорируются. Незамаскированными, как правило, остаются наиболее информативные признаки.

Для обучения модели отрицательного отбора с использованием маскирования детекторов предлагается применять описанный ниже метод.

1. Установить номер итерации: $I = 0$. Положить число детекторов, подвергнутых цензурированию: $N_c = 0$. Задать: $D = \emptyset$.

2. Сформировать замаскированный детектор $d = \{d_i\}$, где d_i – значение маски для n -го разряда. Для этого всем битам строки, описывающей детектор d , необходимо установить значение: $\forall n: d_i = Z$, где Z – символ маски. Увеличить номер итерации: $I = I + 1$.

3. Если $\exists s \in S: match(d, s) = 1$, тогда перейти к этапу 3, в противном случае – к этапу 6.

4. Выбрать произвольным образом бит d_i , $i = 1, \dots, l$, $d_i = Z$. Если такого бита не существует, тогда установить $N_c = N_c + 1$ и перейти к этапу 2, в противном случае – перейти к этапу 5.

5. Установить значение i -го бита детектора: $d_i = -s_i$. Перейти к этапу 3.

6. Добавить детектор d в набор детекторов: $D = D \cup \{d\}$. Если выполняется критерий останова (который может быть основан на значениях N_c и I), тогда перейти к этапу 7, в противном случае – перейти к этапу 2.

7. Останов.

Важно отметить, что обучение детектора является "жадным" по своей природе. Поэтому для нахождения оптимального набора детекторов рекомендуется во время обучения отбирать "свои" экземпляры в случайном порядке.

Модификация метода для работы с вещественным представлением детекторов. Основная идея предлагаемого метода заключается в том, что детектору соответствует не точка в пространстве признаков, а гиперкуб максимально возможного объема.

Пусть дана обучающая выборка S' . Перед обучением модели необходимо установить значения коэффициента α , который определяет значение допуска при классификации детектора, и минимально допустимое значение порога g_{min} , который определяет способность модели к обобщению. Оба коэффициента должны принимать значения в диапазоне $(0,1]$.

Для данной модели рекомендуется использовать следующую метрику:

$$match(d, s) = \begin{cases} 1, & \sum_{i=1}^n \{1 \mid d_{min}^i < s_i \wedge d_{max}^i > s_i\} = n; \\ 0, & \text{в противном случае.} \end{cases}$$

Метод обучения модели отрицательного отбора с маскированием в случае использования вещественных признаков включает в себя следующие этапы.

1. Установить: $D = \emptyset$.

2. Сформировать кандидата в детекторы $d = \langle d_{min}, d_{max} \rangle$, $d_{min} = \{s_{min}^1, s_{min}^2, \dots, s_{min}^n\}$, $d_{max} = \{s_{max}^1, s_{max}^2, \dots, s_{max}^n\}$, где s_{min}^i и s_{max}^i – минимальное и максимальное значения i -го признака среди экземпляров выборки соответственно, $i = 1, \dots, n$.

3. Для каждого экземпляра $s \in S'$ выполнять подэтапы 3.1 – 3.4.

3.1. Если $match(d, s) = 1$, тогда перейти к подэтапу 3.2, в противном случае – взять следующий экземпляр и перейти к подэтапу 3.4.

3.2. Найти такой i -ый признак, чтобы $d_{min}^i < s_i$ и $d_{max}^i > s_i$.

3.3. Если $rand > 0,5$, где $rand$ – функция, возвращающая случайное число из диапазона $(0,1]$, тогда принять:

$d_{min}^i = s_i + \alpha(s_{max}^i - s_{min}^i)$, в противном случае – принять:
 $d_{max}^i = s_i - \alpha(s_{max}^i - s_{min}^i)$.

3.4. Если $match(d, s) = 0$, тогда перейти к подэтапу 3.5, в противном случае – перейти к подэтапу 3.2.

- 3.5. Выбрать следующий экземпляр $s \in S'$ и перейти к подэтапу 3.1.
 4. Оценить способность детектора к обобщению:

$$g(d) = \frac{\sum_{i=1}^n \frac{d_{max}^i - d_{min}^i}{s_{max}^i - s_{min}^i}}{n}.$$

5. Если $g(d) > g_{min}$, тогда добавить детектор в набор: $D = D \cup \{d\}$.

В результате обучения будет сформирован набор детекторов, описывающих пространство, комплементарное к пространству "своих" экземпляров, при этом все детекторы обладают требуемой способностью к обобщению, что позволяет избежать проблемы переобученности модели.

Также, изменяя значение коэффициента α , можно установить порог толерантности модели к своим экземплярам – чем больше значение α , тем больше расстояние между детекторами модели и "своими" экземплярами.

Эксперименты и результаты. С целью проведения экспериментов предложенная модель отрицательного отбора и метод ее обучения были реализованы на языке Python.

С помощью разработанного математического обеспечения решалась задача диагностирования лопаток газотурбинных авиадвигателей [7]. Выборка экспериментально полученных наблюдений содержала экземпляры, характеризовавшиеся спектрами мощностей свободных затухающих колебаний лопаток после ударного возбуждения. Исходя из этих данных, требовалось осуществлять классификацию лопаток на группы кондиционных и дефектных (некондиционных). Каждый экземпляр выборки характеризовался 10240 признаками. Выборка содержала 32 экземпляра. Для обучения модели использовались свертки [8], что позволило сократить размерность задачи. В результате точность классификации составила 91%.

Для обеспечения объективности сравнения предложенного метода с известными методами [9] решалась задача классификации ирисов Фишера [10]: выборка данных содержала 150 экземпляров, характеризовавшихся значениями четырёх признаков. В результате применения разработанного математического обеспечения были сформированы отдельные модели для каждого из трёх классов, позволяющие отличать соответствующий класс от двух остальных. Совокупная модель содержала всего 7 детекторов и обеспечивала

точность классификации 97,3%. В случае применения модели V-Detector для решения данной задачи точность классификации составила до 90%, при этом модель содержала 16 детекторов.

Результаты проведенных экспериментов и реализующего его программного обеспечения, свидетельствуют о работоспособности разработанного метода, а также позволяют рекомендовать его для использования при решении практических задач диагностирования.

Выводы. С целью решения актуальной задачи автоматизации процесса диагностирования объектов и систем разработано математическое обеспечение, позволяющее строить диагностические модели по прецедентам на основе иммунокомпьютинга.

Научная новизна результатов работы заключается в том, что получил дальнейшее развитие метод обучения модели отрицательного отбора с маскированием, который модифицирован путем добавления значений верхней и нижней границы допустимых значений каждого признака детектора, что позволяет решать задачи с вещественным представлением признаков. Также, благодаря введению коэффициентов α и g_{min} , повышается контролеспособность метода.

Практическая ценность работы заключается в том, что разработана программная реализация модели отрицательного отбора с использованием маскированных детекторов, которая может использоваться для решения практических задач технического диагностирования.

Работа выполнена в рамках госбюджетной научно-исследовательской темы "Методы, модели и устройства принятия решений в системах распознавания образов" (номер гос. регистрации 0111U000059).

Список литературы: 1. *Субботин С.О.* Подання й обробка знань у системах штучного інтелекту та підтримки прийняття рішень: навч. посібник / *С.О. Субботин*. – Запоріжжя: ЗНТУ, 2008. – 341 с. 2. The effect of binary matching rules in negative selection / *F. Gonzalez, D. Dasgupta, J. Gomez* // Genetic and Evolutionary Computation: Conference GECCO-2003: Chicago, July 12-16 2003: proceedings. – Berlin-Heidelberg: Springer-Verlag, 2003. – P. 195-206. 3. *Ji Z.* Revisiting Negative Selection Algorithms / *Z. Ji, D. Dasgupta* // Evolutionary Computation. – 2007. – Vol. 15. – P. 223-251. 4. Модель отрицательного отбора с использованием маскированных детекторов и метод ее обучения для решения задач диагностирования / *С.А. Зайцев, С.А. Субботин* // Бионика интеллекта. – 2011. – № 2. 5. Self-Nonsell Discrimination in a Computer / *S. Forrest, A.S. Perelson, R. Cherkuri, L. Allen* // Research in Security and Privacy: IEEE Symposium, Oakland, 16-18 May 1994: proceedings. – Los Alamitos: IEEE, 1994. – P. 202-212. 6. An immunological approach to change detection: algorithms, analysis, and implications / *P. D'haeseleer, S. Forrest, P. Helman* // Computer Security and Privacy: IEEE Symposium, Oakland, 6-8 May 1996 : proceedings. – Los Alamitos: IEEE, 1996. – P. 110-119. 7. Интеллектуальные средства диагностики и прогнозирования надежности авиадвигателей: монография / *В.И. Дубровин, С.А. Субботин, А.В. Богуслаев*,

В.К. Яценко]. – Запорожье: ОАО "Мотор-Сич", 2003. – 279 с. **8.** Диагностика лопаток авиадвигателей на основе многослойной логически прозрачной нейронной сети / *А.В. Богуслаев, В.И. Дубровин, С.А. Субботин, В.К. Яценко* // Вісник двигунобудування. – 2001. – № 1. – С. 85-90. **9.** Ji Z. V-Detector: An Efficient Negative Selection Algorithm with "Probably Adequate" Detector Coverage / *Z. Ji, D. Dasgupta* // Information Sciences. – 2009. – Vol. 179. – P. 1390-1406. **10.** Fisher R.A. Iris data set [Electronic resource] / *R.A. Fisher*. – Massachusetts: University of Massachusetts Amherst, 2007. – Access mode: <http://archive.ics.uci.edu/ml/datasets/Iris>.

Статью представил д.т.н., проф. НТУ "ХПИ" Дмитриенко В.Д.

УДК 004.93

Модель негативного відбору з маскуванням детекторів для вирішення задач діагностування з дійсним представленням ознак / Зайцев С.О., Субботін С.О. // Вісник НТУ "ХПІ". Серія: Інформатика та моделювання. – Харків: НТУ "ХПІ". – 2012. – № 38. – С. 70 – 76.

Вирішено актуальне завдання автоматизації діагностування об'єктів та систем, описаних дійсними ознаками. Дістав подальшого розвитку метод навчання моделі негативного відбору з маскуванням, модифікований для роботи з дійсними ознаками детекторів. Вирішено задачу неруйнівного діагностування лопаток газотурбінних авіадвигунів. Бібліогр.: 10 найм.

Ключові слова: діагностування, дійсні ознаки, негативний відбір, маскування, детектор.

UDC 004.93

A negative selection model with masked detectors for solving diagnostic tasks with real-valued feature representation / Zaitsev S.A., Subbotin S.A. // Herald of the National Technical University "KhPI". Subject issue: Information Science and Modeling. – Kharkov: NTU "KhPI". – 2012. – № 38. – P. 70 – 76.

The topical problem of automated diagnostics for objects and systems described with real-valued features had been solved. The training method for the negative selection model with masked detector has been developed. It was modified for handling real-valued features. The problem of turboshaft airplane engine non-destructive diagnostics has been done using the suggested model. Refs: 10 titles.

Keywords: diagnostics, real-valued features, negative selection, masking, detector.

Поступила в редакцію 09.06.2012