

## Гетерогенные классификаторы с виртуальными потоками в интеллектуальных системах прогнозирования сердечно-сосудистых осложнений в реабилитационном периоде

### Аннотация

Рассмотрены виртуальные модели «слабых» классификаторов для интеллектуальных систем классификации риска повторного инфаркта миокарда. При построении моделей риска сердечно-сосудистых осложнений в качестве дополнительных факторов риска использовались результаты биоимпедансных исследований.

Исследовано пять классификаторов, из которых четыре являлись гетерогенными. Модели гетерогенных классификаторов получали путем последовательного увеличения числа решающих модулей, входящих в классификационную модель. При использовании всех решающих модулей в гетерогенном классификаторе получена диагностическая чувствительность 0,90 при диагностической специфичности 0,86. При конфигурации признакового пространства, включающей в себя только традиционные факторы риска, показатели качества классификации не уступают известным шкалам риска сердечно-сосудистых осложнений.

При решении задач синтеза математических моделей, обеспечивающих прогнозирование сердечно-сосудистых осложнений (ССО) в реабилитационном периоде, следует иметь в виду, что в распоряжении врачей поликлиник имеются стандартные, достаточно хорошо апробированные методы и методики, основывающиеся на данных опросов, осмотров, лабораторных и инструментальных исследований. Существуют также специфические методы исследований, «привязанные» к конкретным органам и заболеваниям. Однако стандартных перечней методов обследования, ориентированных на решение задач оценки течения периода реабилитации и прогнозирования осложнений по различным органам и заболеваниям, практически нет. Имеются лишь отдельные публикации, ориентированные в основном на исследование математических методов оценки информативных признаков по отношению к исследуемому классу задач с соответствующими практическими рекомендациями без достаточных обобщений, которые можно было бы распространить на различные заболевания сердечно-сосудистой системы (ССС) [1].

Данное исследование ориентировано на ведение пациентов с острым инфарктом миокарда (ИМ), при оценке степени риска ССО у которых большое значение имеет риск развития повторного ИМ (ПИМ). Известные методы прогнозирования ССО имеют ряд существенных недостатков: определение индивидуального прогноза у большей части больных остается неудовлетворительным; практически нет единых систем прогнозирования возникновения и исхода заболевания; составление прогноза является трудоемкой работой для врача, поэтому разработка методов прогнозирования ПИМ, основанных на автоматизации процесса анализа факторов риска (ФР) посредством гетерогенных классификаторов, и математических моделей, позволяющих учитывать фактор латентных переменных в риске ПИМ, является важной научно-технической задачей.

Гетерогенное признаковое пространство включает в себя множество подпространств информативных признаков (ИП), структурированных по принципу неоднородности источников получения данных. При этом в каждом из гетерогенных подпространств ИП могут быть получены свои гетерогенные подпространства. В каждом из гетерогенных подпространств синтезируется «слабый» классификатор. Для повышения диагностической эффективности «слабых» классификаторов используются виртуальные потоки [2]-[4]. Путем объединения «слабых» классификаторов по известным технологиям диагностическая эффективность итогового решающего правила приводится к требуемому значению. Таким образом, гетерогенные классификаторы отражают скрытые системные связи между реальными и виртуальными потоками.

В основу метода формирования гетерогенных классификаторов с виртуальными потоками положено структурирование вектора ИП на два подвектора, первый из которых соответствует реальным потокам, а второй – виртуальным потокам. Виртуальные потоки построены на основе предложенной мо-

дели универсального нейросетевого аппроксиматора и МГУА-моделирования [5]-[7].

Анализ литературных данных и серия специально проведенных исследований позволили установить, что достаточно высокими ФР являются отклонения энергетических характеристик биологически активных точек (БАТ), «связанных» с соответствующей патологией, от своих номинальных значений [8], что позволяет включить их в гетерогенный классификатор. Однако прогностические свойства энергетических показателей БАТ имеют ряд недостатков. Для их устранения было предложено использовать импульсное воздействие на БАТ:

$$U_i(t) = (U_0 + i \cdot \Delta U) \cdot \text{rect}(t), \quad (1)$$

где  $U_0$  – начальная амплитуда импульса;  $\Delta U$  – шаг квантования по амплитуде;  $t = i \cdot \Delta + (i - 1) \cdot \Delta$ ;  $\Delta$  – шаг квантования по времени;  $i = 1, 2, \dots, n$ .

Для построения классификатора с виртуальными потоками формируется дополнительная псевдонезависимая переменная  $x_{n+1}$  как некоторая функция

$$x_{n+1} = f(x_1, \dots, x_n), \quad (2)$$

которая строится на парадигме, отличной от парадигмы, на которой построен классификатор.

Функция (2) подбирается таким образом, чтобы минимизировать ошибку классификатора, построенного с учетом виртуального потока.

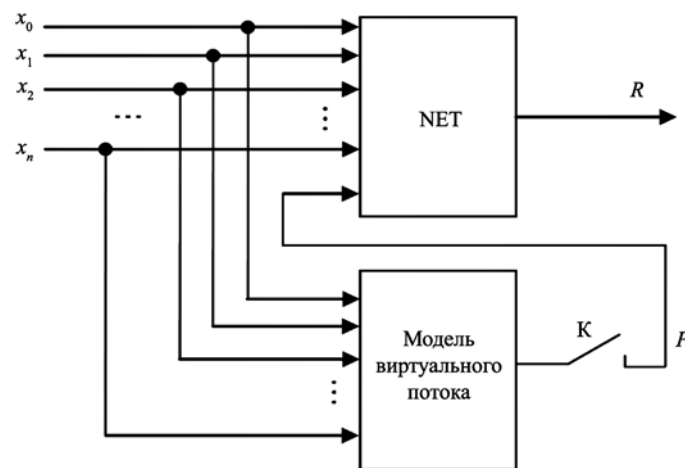


Рис. 1. Структура классификатора с виртуальным потоком

Модель «слабого» классификатора с виртуальными потоками представлена на рис. 1 и включает в себя три модуля. Модуль NET (нейронная сеть) настраивается сначала без виртуального потока (ключ К на схеме рис. 1 разомкнут), после чего выбирается модель виртуального потока и настраиваются ее параметры таким образом, чтобы минимизировать ошибку классификатора, при этом ключ К на рис. 1 замкнут.

Нейронная сеть NET (рис. 1) настраивается на обучающей выборке, включающей в себя пациентов с высоким риском ПИМ и с низким риском ПИМ при отключенном виртуальном потоке (ключ К разомкнут).

Если используется дихотомическая модель риска ПИМ, то классификация ведется по классу  $\omega_0$  (нет риска ПИМ) и классу  $\omega_1$  (есть риск ПИМ). Тогда каждый пациент характеризуется латентной переменной  $\theta_{\omega_i}$ , которая принимает значения 0 или 1 в зависимости от функционального состояния ССС. С другой стороны, имеется тест  $\beta_j$ , в результате которого получаем также два значения «реакции ответа»: 0 или 1. Для того чтобы тест  $\beta_j$  мог работать в качестве «слабого» классификатора, необходимо, чтобы его прогностическая способность превышала величину 0,5.

Если в импульсной посылке (1) имеется  $n$  импульсов и в качестве реакции на них имеем некоторое множество единиц и нулей, то уверенность в риске ПИМ по тесту для  $j$ -й БАТ увеличивается с ростом числа единиц в ответных реакциях на тестовую посылку.

В таком случае в качестве уверенности в риске ПИМ по  $j$ -й БАТ примем экспоненциальную модель

$$P_j = 1 - e^{-k\lambda_j}, \quad (3)$$

где  $\lambda_j = \sum_{i=1}^n \beta_{ij}$ ;  $k$  – параметр модели.

Однопараметровая модель (3) принята в качестве модели виртуального потока. Параметр  $k$  этой модели подбирается в процессе процедуры обучения.

Так как предполагается нахождение пациентов под наблюдением, то в этом случае достаточно контролировать текущие значения биоимпеданса и сравнивать их с контрольными значениями, полученными в постинфарктный ремиссионный пе-

риод. Это позволяет построить две модели виртуального потока: одна привязана к индивидуальному риску, другая – к групповому риску.

Первая модель является моделью индивидуального риска и построена на основе аппроксимации вольтамперных характеристик (ВАХ) в биоактивных точках полиномами седьмого порядка с последующим нечетким логическим выводом [8], [9].

При групповой модели риска необходимо установить статистическое отличие между приращением электрического сопротивления в каждом  $i$ -м тесте  $j$ -й БАТ у пациентов с низким риском ПИМ и у пациентов с высоким риском ПИМ. Для формирования виртуального потока по модели на основе группового подхода построим регрессионную модель ВАХ  $j$ -й БАТ для пациентов с низким риском ПИМ и пациентов с высоким риском ПИМ. Регрессионные модели этих характеристик, аппроксимированных полиномом седьмого порядка, представлены на рис. 2. Имея две регрессионные модели, относим неизвестный образец (результат  $i$ -го теста) к одной из них в зависимости от индивидуального или группового риска, контролируемого ЛПР в интеллектуальной системе прогнозирования ПИМ.

Вторая модель включает в себя три блока. В первом блоке осуществляется нормирование данных относительно контрольных значений тестовых сигналов, во втором блоке строится модель по «сырым» данным на основе статистических исследований в виде аппроксимирующих полиномов (рис. 2), а в третьем блоке реализуется нейросетевая модель «слабого» классификатора, построенная на  $(n + 1)$  входных данных. Классификаторы строятся для каждой релевантной БАТ с последующей агрегацией на основе нечеткого логического вывода или нейросетевых моделей.

Модель принятия решений с гетерогенными классификато-

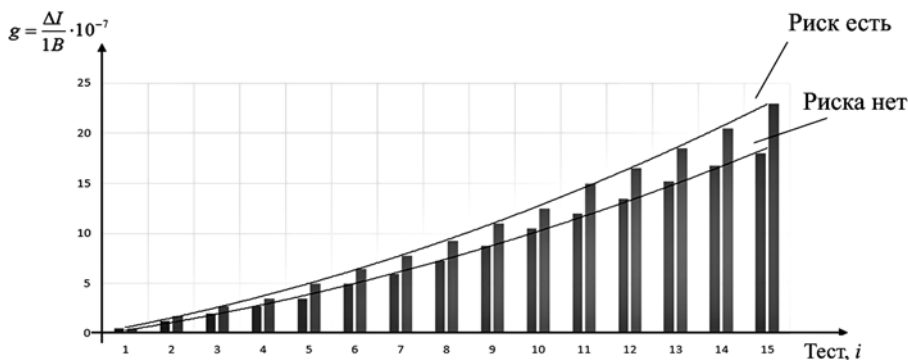


Рис. 2. Аппроксимирующие полиномы приращения токов в БАТ у пациентов с низким и высоким риском ПИМ для  $j$ -й БАТ

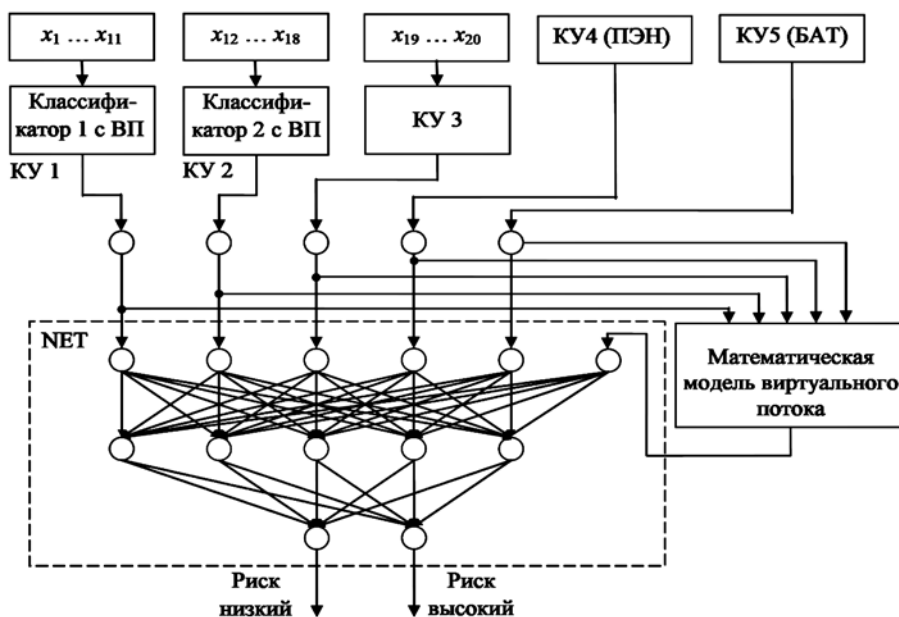


Рис. 3. Модель принятия решений с гетерогенными классификаторами с виртуальными потоками

рами с виртуальными потоками, построенная посредством такой системы для прогнозирования ПИМ, представлена на рис. 3. В модели использовались гетерогенные классификаторы, построенные на ФР, традиционно принятых в медицинской практике ( $x_1 \dots x_{11}$ ), сопутствующих заболеваниях ( $x_{12} \dots x_{18}$ ), дополнительных ФР, влияющих на положительный прогноз риска ПИМ, например, показателей перекисного окисления липидов и антиокислительной активности ( $x_{19} \dots x_{20}$ ); психоэмоциональном статусе (ПЭН), энергетических характеристиках релевантных БАТ. Разработанные интерфейсные окна позволяют управлять показателями качества «слабых» классификаторов как путем обучения, так и посредством экспертного оценивания.

Выполнен ROC-анализ качества принятия решений гетерогенными классификаторами по оценке риска ПИМ на репрезентативных контрольных выборках (рис. 4). Исследовано пять гетерогенных классификаторов с последовательным увеличением числа решающих модулей, входящих в классификационную модель.

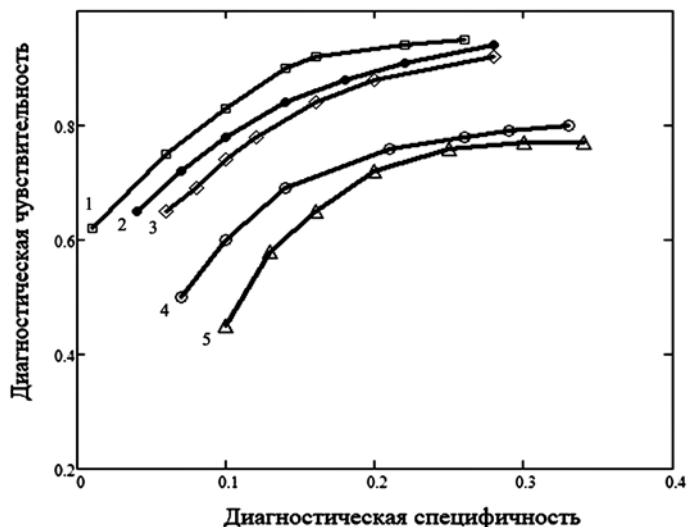


Рис. 4. ROC-кривые пяти гетерогенных классификаторов: 1 – по всему пространству ИП; 2 – без учета виртуальных потоков; 3 – без виртуальных потоков и биоимпедансных исследований; 4 – без виртуальных потоков, биоимпедансных исследований и анализа ПЭН; 5 – без виртуальных потоков, биоимпедансных исследований, анализа ПЭН и показателей ПОЛ и АОА

При использовании всех решающих модулей в гетерогенном классификаторе получена диагностическая чувствительность 0,90 при диагностической специфичности 0,86. При конфигурации признакового пространства, включающей в себя только традиционные факторы риска, показатели качества классификации не уступают известным шкалам риска сердечно-сосудистых осложнений.

## Выводы

Разработаны виртуальные модели классификации риска повторного инфаркта миокарда по результатам биоимпедансных исследований, первая из которых является моделью индивидуального риска и построена на основе полиномиальной аппроксимации вольтамперных характеристик в биоактивных точках с последующим нечетким логическим выводом, а вторая модель является моделью группового риска и использует экспоненциальную аппроксимацию тестовых сигналов для формирования виртуального потока, что позволяет агрегировать факторы риска по релевантным биологически активным точкам.

ROC-анализ качества принятия решений предложенными гетерогенными классификаторами по оценке риска повторного инфаркта миокарда на репрезентативных контрольных выборках показал целесообразность использования в качестве факторов риска сердечно-сосудистых осложнений результатов биоимпедансного анализа в аномальных зонах электропроводимости. Исследовано пять гетерогенных классификаторов с

последовательным увеличением числа решающих модулей, входящих в классификационную модель. При использовании всех решающих модулей в гетерогенном классификаторе получена диагностическая чувствительность 0,90 при диагностической специфичности 0,86. При конфигурации признакового пространства, включающей в себя только традиционные факторы риска, показатели качества классификации не уступают известным шкалам риска сердечно-сосудистых осложнений.

## Список литературы:

1. Гурьева М.Э., Журавлева М.В., Алеева Г.Н. Критерии качества жизни в медицине и кардиологии // Русский медицинский журнал. 2006. Т. 14. № 10. С. 761-763.
2. Киселев А.В., Петрова Т.В., Дегтярев С.В. и др. Нейросетевые модули с виртуальными потоками для классификации и прогнозирования функционального состояния сложных систем // Известия Юго-Западного государственного университета. 2018. Т. 22. № 4. С. 123-134.
3. Киселев А.В., Филист С.А., Шаталова О.В. и др. Виртуальные потоки в гибридных решающих модулях классификации сложноструктурируемых данных // Прикаспийский журнал: управление и высокие технологии. 2018. № 2 (42). С. 137-149.
4. Киселев А.В., Шаталова О.В., Петрунина Е.В. и др. Слабые классификаторы с виртуальными потоками в интеллектуальных системах прогнозирования сердечно-сосудистых осложнений // Известия Юго-Западного государственного университета. Серия «Управление, вычислительная техника, информатика. Медицинское приборостроение». 2019. Т. 9. № 1 (30). С. 6-19.
5. Филист С.А., Емельянов С.Г., Рыбочкин А.Ф. Нейросетевой решающий модуль для исследования живых систем // Известия Курского государственного технического университета. 2008. № 2 (23). С. 77-82.
6. Филист С.А., Шаталова О.В., Ефремов М.А. Гибридная нейронная сеть с макрослоями для медицинских приложений // Нейрокомпьютеры. Разработка и применение. 2014. № 6. С. 35-39.
7. Томакова Р.А., Филист С.А., Яа Зар До Универсальные сетевые модели для задач классификации биомедицинских данных // Известия Юго-Западного государственного университета. 2012. № 4 (43). С. 44-50.
8. Суржикова С.Е., Шаталова О.В., Богданов А.С. Применение автоматизированной системы для исследования вольтамперных характеристик биоматериалов // Биомедицинская радиоэлектроника. 2014. Вып. 9. С. 43-46.
9. Мохаммед Авад А.А., Филист С.А., Шаталова О.В. Моделирование импеданса биоматериалов в среде MATLAB // Известия Юго-Западного государственного университета. Серия «Управление, вычислительная техника, информатика. Медицинское приборостроение». 2013. № 4. С. 61-66.

Елена Валерьевна Петрунина,  
канд. техн. наук, доцент, декан,  
факультет прикладной математики и информатики,  
ФГБОУ ИВО «Московский государственный  
гуманитарно-экономический университет»,  
Ольга Владимировна Шаталова,  
канд. техн. наук, доцент,  
Денис Станиславович Забанов,  
аспирант,  
кафедра биомедицинской инженерии,  
Вадим Владимирович Серебровский,  
д-р техн. наук, профессор,  
кафедра программной инженерии,  
ФГБОУ ВО «Юго-Западный  
государственный университет»,  
г. Курск,  
e-mail: shatolg@mail.ru