



В. В. Левшинский

## Многоклассовая классификация в задаче дифференциальной диагностики венозных заболеваний по данным микроволновой радиотермометрии

Аннотация. Работа посвящена применению математических моделей в задаче дифференциальной диагностики венозных заболеваний по данным микроволновой радиотермометрии. Описан модифицированный процесс трансформации признакового пространства в термометрических данных. После конструирования признаков выполняется решение задачи многоклассовой классификации несколькими способами: путем сведения к задачам бинарной классификации методами «один против всех» и «каждый против каждого», а также путем построения многомерной логистической регрессии. Лучшим алгоритмом классификации достигнут средний показатель сбалансированной точности 0.574. Ключевой особенностью подхода является то, что результат классификации обосновывается в понятных для врача-диагноста терминах. Представлены наиболее значимые закономерности и точность, с которой они могут выявить различные классы заболеваний.

*Ключевые слова и фразы:* микроволновая радиотермометрия, математическое моделирование, конструирование признаков, многоклассовая классификация.

### Введение

В настоящее время крайне актуальной является разработка интеллектуальных систем, основанных на применении различных методов искусственного интеллекта [1]. В медицинской диагностике подобные

---

Исследование выполнено при финансовой поддержке РФФИ в рамках научного проекта № 19-31-90153.

© В. В. Левшинский, 2021

© Волгоградский государственный университет, 2021

© Программные системы: теория и приложения (дизайн), 2021

 10.25209/2079-3316-2021-12-2-19-36



системы могут интерпретировать и анализировать данные обследования, выполнять функцию поддержки принятия решений. Наибольший интерес вызывают консультативные системы, не только использующие методы и алгоритмы машинного обучения, но и содержащие механизмы обоснования предлагаемых решений. Разработка таких систем требует применения методов математического моделирования, машинного обучения и анализа данных.

Микроволновая радиотермометрия является перспективным диагностическим методом, основанным на изучении собственного электромагнитного излучения тканей человека в микроволновом и инфракрасном диапазонах длин волн. Особенностью метода является абсолютная безвредность для пациента. Метод успешно применяется в различных областях медицины [2–4], в частности, в ранней диагностике и динамическом контроле заболеваний вен нижних конечностей [5], причисляемых к «болезням цивилизации», поскольку количество страдающих ими исчисляется миллиардами.

Методика обследования заключается в последовательном измерении внутренних и поверхностных (кожных) температур, регистрации температур в виде числовых данных и последующем анализе данных обследования. Специалист выполняет поиск аномалий в термометрических данных, что является крайне сложной интеллектуальной задачей, требующей длительной подготовки и многолетнего опыта. Разработка интеллектуальных систем в перспективе сможет не только улучшить качество диагностики в целом, но и разрешить проблему нехватки узкопрофильных специалистов, что сделает возможным массовое применение этого метода. Интерпретация и формализация знаний экспертов, а также извлечение знаний из данных, являются ключевыми этапами разработки моделей решения подобных задач.

В течение последнего десятилетия появились первые исследования, посвященные применению методов математического моделирования, машинного обучения и анализа данных в задаче диагностики заболеваний вен нижних конечностей по данным микроволновой радиотермометрии. Первые модели были основаны на байесовском классификаторе [6]. Признаковое пространство определялось значениями температур, а при постановке диагноза применялись критерии, непонятные врачу-диагносту. Все это создавало значительные трудности в обосновании диагностического решения.

Статистические модели стали предпосылками к созданию эффективных моделей и алгоритмов, допускающих интерпретацию и обоснование результата [7, 8]. Эти модели применялись для решения задачи бинарной классификации «Здоров»–«Болен».

В смежной сфере, в диагностике рака молочных желез по данным микроволновой радиотермометрии, в результате интеллектуального анализа термометрических данных было выявлено значительное количество закономерностей, описывающих аномалии поведения температурных полей [9]. На их основе выполнено построение модели, допускающей обоснование решения, которая также может быть применена и в диагностике венозных заболеваний [10, 11].

Целью исследования является применение модели динамического описания состояния пациента в задаче дифференциальной диагностики венозных заболеваний по данным микроволновой радиотермометрии.

Следует отметить, что задача дифференциальной диагностики рассматривается не впервые. Ранее для ее решения применялись нейронные сети [12]. Сравнительный обзор этого подхода представлен в заключении.

## **1. Данные и методы**

### **1.1. Микроволновая радиотермометрия**

Метод микроволновой радиотермометрии — это биофизический метод неинвазивного обследования, заключенный в последовательном измерении внутренних и поверхностных температур в определенных точках и регистрации температур в виде числовых данных. Специалист выполняет анализ полученных данных, которые могут быть отображены в виде термограмм или карт температурных полей с целью выявить температурные аномалии, и выносит заключение о состоянии здоровья, либо о необходимости дальнейшего обследования. Метод основывается на факте, что температурные аномалии предшествуют структурным изменениям.

В качестве примера на рисунке 1 представлены карты полей внутренних и поверхностных температур голени пациента, левая нога которого поражена венозным заболеванием.

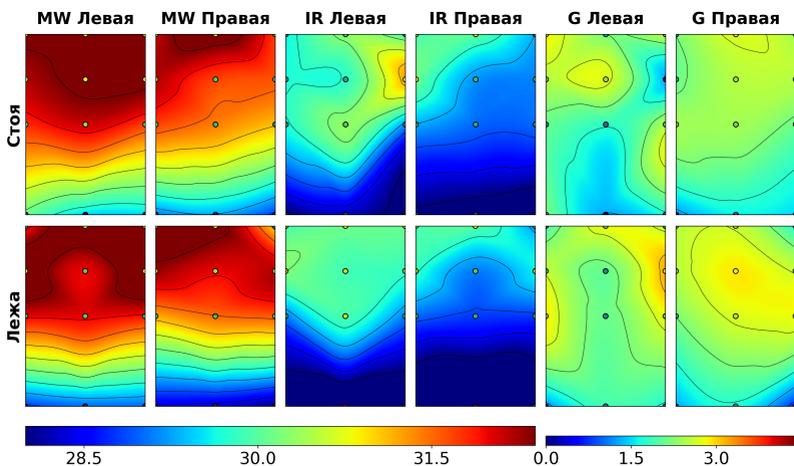


Рисунок 1. Поля температур пациента, левая нога которого поражена венозным заболеванием. Обозначения: MW — внутренние, IR — поверхностные, G — внутренние градиенты температур. Выполнена интерполяция кубическими сплайнами.

В процессе термометрического обследования голеней специалист последовательно измеряет внутренние и поверхностные температуры в 12 симметричных точках, расположенных по задней поверхности обеих голеней пациента в соответствии со схемой на рисунке 2. Выполняется несколько измерений при различных положениях пациента: лежа на животе и стоя.

## 1.2. Набор данных

Исследуется набор термометрических данных нижних конечностей 146 пациентов (292 голени). Каждая голень имеет метку в зависимости от наличия того или иного заболевания:

- 0 (Здоровые) — данные измерений голеней пациентов без заболеваний, 36 голеней (12.3%);
- 1 (Норма 2) — здоровые голени пациентов, имеющих венозное заболевание на другой голени, 67 голеней (22.9%);
- 2 (ХВН) — голени пациентов с хронической венозной недостаточностью, 100 (34.2%);

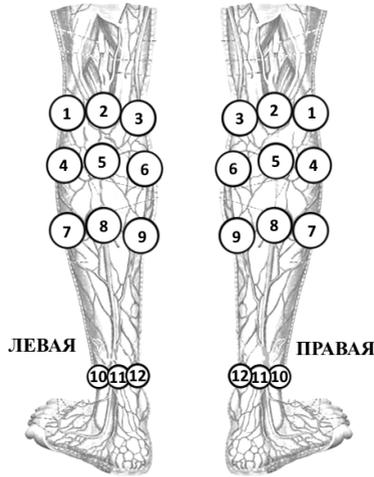


Рисунок 2. Схема измерений, точки 1–12.

- 3 (ПТБ) — постромботическая болезнь, 69 голеней (23.6%);
- 4 (ОВТ) — острый венозный тромбоз, 20 (6.8%).

В формальном виде набор данных может быть представлен в виде матрицы

$$(1) \quad X = \begin{bmatrix} t_1^1 & t_2^1 & \dots & t_n^1 \\ t_1^2 & t_2^2 & \dots & t_n^2 \\ \dots & \dots & \dots & \dots \\ t_1^m & t_2^m & \dots & t_n^m \end{bmatrix}, y = \begin{bmatrix} y_1 \\ y_2 \\ \dots \\ y_m \end{bmatrix}, Y = \{1, 2, \dots, C\},$$

где  $m$  — количество объектов в наборе данных,  $n$  — количество признаков,  $x^i = (t_1^i, \dots, t_n^i)$  — признаковое описание объекта  $i$ ,  $Y$  — множество меток классов, а  $y_i \in Y$  — метка класса.

### 1.3. Конструирование признаков

В несколько шагов выполняется построение модели состояния каждой голени и преобразование признакового пространства.

В признаковом описании содержатся 48 значений внутренних и поверхностных температур, измеренных в определенных точках

голеней в положениях лежа и стоя. Точки измерений изображены на рисунке 2. На первом шаге в температурных данных выделяются следующие группы:

1. Внутренние температуры, положение стоя

$$T^{i,mw,st} = (T_1^{i,mw,st}, \dots, T_{12}^{i,mw,st})$$

2. Внутренние температуры, положение лежа

$$T^{i,mw,ly} = (T_1^{i,mw,ly}, \dots, T_{12}^{i,mw,ly})$$

3. Поверхностные температуры, положение стоя

$$T^{i,ir,st} = (T_1^{i,ir,st}, \dots, T_{12}^{i,ir,st})$$

4. Поверхностные температуры, положение лежа

$$T^{i,ir,ly} = (T_1^{i,ir,ly}, \dots, T_{12}^{i,ir,ly})$$

Здесь верхний индекс  $mw$  или  $ir$  определяет диапазон температур (внутренние или поверхностные), а индекс  $st$  или  $ly$  (лежа или стоя) определяет положение, в котором находился пациент в процессе измерений. Нижний индекс — номер точки.

Отдельно добавляется группа внутренних градиентов температур — разностей внутренних и поверхностных температур в соответствующих точках. Например, внутренние градиенты температур, измеренных в положении лежа

$$T^{i,g,ly} = (T_1^{i,g,ly}, \dots, T_{12}^{i,g,ly}).$$

Далее, для каждой группы точек и отдельно для пар групп вычисляются характеристики, представленные в виде гипотез о поведении температурных полей и соответствующих им обобщенных математических описаний [9, 10]:

1. Гипотеза о незначительном перепаде температур, согласно которой для здоровых голеней характерны невысокие значения следующих функционалов:

- 1.1. Осцилляция температур

$$(2) \quad F_1(T) = \max_{t \in T} t - \min_{t \in T} t$$

где  $T$  — температуры.

1.2. Разброс температур

$$(3) \quad F_2(T) = ST_{dev}(T) = \sqrt{\frac{\sum_{t \in T} (t - \bar{T})^2}{|T| - 1}},$$

где  $\bar{T}$  — среднее значение температур в  $T$ ,  $|T|$  — количество значений температур в  $T$ .

1.3. Отклонение значений температур относительно средних

$$(4) \quad F_3(T) = \max_{t \in T} |\bar{T} - t|$$

1.4. Разброс внутренних градиентов. В качестве меры разброса внутренних градиентов используются максимальные и минимальные значения, (2), (3), (4), а также  $L^p$  нормы:

$$(5) \quad F_4(T) = \|T\|_1, \quad F_5(T) = \|T\|_2, \quad F_6(T) = \|T\|_\infty,$$

где

$$\|T\|_p = \left( \sum_{t \in T} |t|^p \right)^{\frac{1}{p}}, \quad \|T\|_\infty = \max_{t \in T} |t|.$$

Более конкретно:

1.4.1. Максимум разности внутренних температур голени и средней температуры, положение стоя

$$f_1(x^i) = F_3(T^{i,mw,st}) = \max_{t \in T^{i,mw,st}} |T^{i,mw,st} - t|$$

1.4.2. Разброс внутренних температур голени, измеренных в положении лежа

$$f_2(x^i) = F_2(T^{i,mw,ly}) = \sqrt{\frac{\sum_{t \in T^{i,mw,ly}} (t - \overline{T^{i,mw,ly}})^2}{|T^{i,mw,ly}| - 1}}$$

1.4.3. Осцилляция поверхностных температур голени, измеренных в положении стоя

$$f_3(x^i) = F_1(T^{i,ir,st}) = \max_{t \in T^{i,ir,st}} t - \min_{t \in T^{i,ir,st}} t$$

2. Гипотеза о симметрии температурных полей, согласно которой для здоровых голеней характерно незначительное отклонение температур в соответствующих точках (подобластях), а также незначительная разность сопутствующих характеристик.

В качестве мер симметрии используются характеристики

вида

$$(6) \quad F(T_c, T_p) = \|T_c - T_p\|, \quad F(T_c, T_p) = \|T_c\| - \|T_p\|,$$

где  $\|z\|$  — функционал,  $T_c - T_p$  — поэлементная разность,  $T_c$  — «текущая», а  $T_p$  — «парная» группа температур. Эти характеристики требуют дополнительного этапа предобработки данных, а также наличия пары для каждого объекта в выборке. Например, в процессе предобработки данных голеней, если просматривается левая голень, то «текущая» группа температур — это внутренние или поверхностные температуры левой голени, а «парной» ей группой будет соответственно группа внутренних или поверхностных температур правой голени.

Для парных групп температур вычисляются характеристики, в основном определенные в рамках предыдущей гипотезы, например:

- 2.1. Максимальное абсолютное значение разности температур одноименных точек

$$F_7(T_c, T_p) = F_6(T_c - T_p)$$

- 2.2. Разность минимальной и максимальной температур голеней

$$F_8(T_c, T_p) = \max_{t \in T_c} t - \min_{t \in T_p} t$$

- 2.3. Разность стандартных отклонений температур голеней

$$(7) \quad F_9(T_c, T_p) = F_2(T_c) - F_2(T_p)$$

- 2.4. Разность средних значений и т.д.

$$F_{10}(T_c, T_p) = \overline{T_c} - \overline{T_p}$$

3. Гипотеза о стабильности температурных полей, согласно которой для здоровых голеней характерны незначительные различия температур, измеренных при различных положениях пациента.

Признаки этой группы характеризуют степень близости температурных полей в различных положениях и практически аналогичны признакам, определенным в рамках гипотезы о симметрии. Например:

- 3.1. Разность средних значений поверхностных температур голеней, измеренных в положениях стоя и лежа

$$f_4(x^i) = F_{10}(T^{i,ir,st}, T^{i,ir,ly}) = \overline{T^{i,ir,st}} - \overline{T^{i,ir,ly}}$$

- 3.2. Максимальное абсолютное значение разности внутренних градиентов температур голеней, измеренных в положениях

стоя и лежа

$$f_5(x^i) = F_7(T^{i,g,st}, T^{i,g,ly}) = \left\| T^{i,g,st} - T^{i,g,ly} \right\|_{\infty}$$

4. Гипотезы, связанные с физиологическим строением голеней [7, 8]. Рассматриваются значения латерально-медиального и осевого градиентов для различных групп температур, а также их разности для соответствующих групп правой и левой голеней:

- 4.1. латерально-медиальный градиент

$$F_{11}(T) = LMG(T) = \overline{T_{ext}} - \overline{T_{int}},$$

где  $T_{ext}$  — это подгруппа температур точек наружной части голени (точки 1, 4, 7, 10), а  $T_{int}$  — подгруппа температур точек внутренней части голени (точки 3, 6, 9, 12).

- 4.2. осевой градиент

$$(8) \quad F_{12}(T) = AG(T) = \overline{T_{top}} - \overline{T_{bot}},$$

где  $T_{top}$  — это подгруппа температур точек верхней части голени (точки 1, 2, 3), а  $T_{bot}$  — подгруппа температур точек нижней части голени (точки 10, 11, 12).

Возвращаясь к примеру на рисунке 1, наблюдается следующее:

1. Схожесть полей внутренних и поверхностных температур правой голени при измерениях стоя или лежа. Для выявления и описания можно применить признаки вида (6);
2. Схожесть полей внутренних температур правой голени при измерениях стоя и лежа. Аналогично для поверхностных температур. Тоже признаки вида (6);
3. Асимметрия температурных полей правой и левой голеней. Применяются признаки вида (6), включающие все прочие характеристики, например, (7).
4. Различия полей внутренних и поверхностных температур левой голени при измерениях стоя или лежа. Аналогично п. 2;
5. У левой голени наблюдаются колоколообразные контуры. Выявить такие данные можно, например, используя меры разброса, признаки вида (5), а также осевой градиент (8) и др.

#### 1.4. Термометрические признаки

Для каждого объекта в наборе данных вычисляются значения функций  $f$  и конструируется 128 новых признаков. Далее путем

бинаризации [13] полученных значений выполняется построение множества термометрических признаков

$$S = (\phi_1, \phi_2, \dots, \phi_s),$$

где  $s$  — количество.

Термометрический признак это тройка  $\phi = (f, I, W)$ , где  $I$  — интервал и  $W$  — вес (информативность  $f$  на  $I$ ), или количественный показатель, определяющий насколько хорошо признак отделяет объекты одного класса от других классов. Термометрический признак считается выполненным (наблюдается у объекта  $x^i$ ), если  $f(x^i) \in I$ .

Использовалась статистическая информативность [13], которая для случая нескольких классов определяется как

$$I(\phi, X) = -\ln \frac{C_{P_1}^{p_1} \dots C_{P_K}^{p_K}}{C_m^p},$$

где  $C_n^k$  — биномиальный коэффициент,  $P_i$  — кол-во объектов класса  $i$  в выборке  $X$ ,  $p_i$  — кол-во объектов класса  $i$ , у которых наблюдается признак  $\phi$ ,  $p = p_1 + \dots + p_K$ . Эта мера является достаточно справедливой и отлично подходит для несбалансированных выборок небольшого размера.

Ключевой особенностью термометрических признаков является интерпретируемость, что позволяет формировать заключение о состоянии объекта на основе значений термометрических признаков. Вектор  $(\phi_1, \phi_2, \dots, \phi_s)$  определяет диагностическое состояние объекта в выборке. Элемент вектора с индексом  $j$  будет равен 1, если признак  $j$  наблюдается у объекта  $x^i$ , и 0 в противном случае.

После всех преобразований матрица (1) принимает вид бинарной матрицы

$$(9) \quad X' = \begin{bmatrix} \phi_1(x^1) & \phi_2(x^1) & \dots & \phi_s(x^1) \\ \phi_1(x^2) & \phi_2(x^2) & \dots & \phi_s(x^2) \\ \dots & \dots & \dots & \dots \\ \phi_1(x^m) & \phi_2(x^m) & \dots & \phi_s(x^m) \end{bmatrix}$$

и далее выполняется построение алгоритмов классификации. При этом каждый признак из (9) может быть описан на языке, понятном врачу-диагносту.

В результате бинаризации получается большое количество термометрических признаков, при этом многие признаки не дают новой

информации в совокупности друг с другом, поэтому здесь возникает задача отбора признаков. Для решения этой задачи далее в работе применяется логистическая регрессия с  $L_1$ -регуляризацией [14]. Процесс преобразования признакового пространства проиллюстрирован на рисунке 3.



Рисунок 3. Иллюстрация процесса преобразования признакового пространства.

### 1.5. Оценка модели и построение алгоритмов

Для оценки эффективности термометрических признаков при решении задачи дифференциальной диагностики было выполнено построение нескольких алгоритмов классификации. В качестве базы для всех моделей использовался традиционный метод логистической регрессии.

Алгоритм классификации определяется как

$$a(x^i) = \begin{cases} 1, & \text{если } h_W(x^i) \geq 0.5, \\ 0, & \text{иначе} \end{cases}$$

где  $h_W(x^i) = g(W_0 + \sum_{j=1}^s W_j \phi_j(x^i))$  — сумма весов термометрических признаков,  $W_j$  — вес признака  $\phi_j$ , а  $g(z) = \frac{1}{1 + e^{-z}}$  — сигмоида.

В совокупности с термометрическими признаками логистическая регрессия представляет собой алгоритм взвешенного голосования признаков. Построение обоснования решения такого алгоритма состоит в объединении описаний признаков, наблюдаемых у объекта.

Рассмотрено несколько подходов к решению задачи многоклассовой классификации:

1. Логистическая регрессия (LR), один против всех (OvR). Для каждого класса из набора данных строится модель, определяющая, относится ли объект к выбранному классу. Итоговый класс определяется наиболее уверенной моделью. Всего обучается  $C$  классификаторов,  $C$  — кол-во классов.
2. Логистическая регрессия (LR), каждый против каждого (OvO). Для каждой пары классов из набора данных строится отдельный алгоритм классификации. Итоговый класс определяется большинством голосов. Всего обучается  $\frac{C(C-1)}{2}$  классификаторов.
3. Полиномиальная логистическая регрессия (MLR), являющаяся обобщением логистической регрессии для случая нескольких классов.

Помимо общей задачи многоклассовой классификации также рассмотрены случаи без объектов класса «Здоровые», которых можно достаточно эффективно отделить от всех остальных, используя бинарный классификатор [10]. Одним из возможных способов построения алгоритма дифференциальной диагностики видится применение иерархии классификаторов. Самая первая модель предназначена для

классификации «Здоров»–«Болен», а все последующие — для уточнения класса заболевания, если оно имеется.

Для оценки эффективности классификации и сравнения моделей между собой применялся метод вложенной кросс-валидации с сохранением баланса классов [15]. Данные разбивались на 9 блоков на внешнем уровне и 8 блоков на внутреннем. В качестве метрики эффективности использовалась сбалансированная точность [16], определяемая как среднее значение критериев полноты для каждого из классов:

$$\text{Acc}_b = \sum_{i=1}^C \frac{\text{Recall}_i}{C}, \quad \text{Recall}_i = \frac{\text{correct}_i}{\text{total}_i},$$

где  $\text{correct}_i$  — количество объектов класса  $i$ , отнесенные алгоритмом к классу  $i$ ,  $\text{total}_i$  — общее количество объектов класса  $i$ ,  $C$  — количество классов.

## 2. Результаты и обсуждение

Результаты представлены в таблице 1. Здесь «w/o H» отмечены алгоритмы, которые обучались на выборке без класса «Здоровые», Avg — средняя оценка, Std Dev — стандартное отклонение. Наивысшие показатели сбалансированной точности достигнуты с применением LR, OvO в наборе данных всех классов, и с применением LR, OvR в наборе данных без класса 0 («Здоровых»).

Таблица 1. Эффективность классификации

Метрика		LR, OvR w/o H		LR, OvO w/o H		MLR w/o H	
Acc <sub>b</sub>	Avg	0.557	0.548	0.574	0.537	0.56	0.541
	Std Dev	0.102	0.074	0.065	0.05	0.073	0.078
Recall <sub>0</sub>	Avg	0.844	-	0.781	-	0.838	-
	Std Dev	0.136	-	0.162	-	0.096	-
Recall <sub>1</sub>	Avg	0.609	0.641	0.583	0.684	0.597	0.608
	Std Dev	0.138	0.176	0.13	0.092	0.12	0.151
Recall <sub>2</sub>	Avg	0.614	0.579	0.617	0.549	0.595	0.641
	Std Dev	0.197	0.142	0.09	0.11	0.163	0.167
Recall <sub>3</sub>	Avg	0.45	0.576	0.535	0.561	0.519	0.519
	Std Dev	0.18	0.136	0.104	0.189	0.143	0.177
Recall <sub>4</sub>	Avg	0.271	0.396	0.354	0.354	0.25	0.396
	Std Dev	0.333	0.249	0.227	0.227	0.204	0.249

Класс «Здоровые» лучше всего выделяется моделью LR, OvR, средняя точность составляет 0.844.

Разброс оценок LR, OvO меньше, чем у других моделей. Эта модель выделяет класс 2, характеризующий венозные заболевания, в среднем лучше, чем остальные модели. Однако средняя точность диагностики прочих классов выше у других моделей.

Алгоритмы, построенные для набора данных без класса 0 имеют более высокие показатели точности классификации различных заболеваний, чем те же самые алгоритмы для полного набора данных. В этом случае модель LR, OvR лидирует и более эффективно чем остальные выделяет класс 3. В сравнении с MLR эта модель выделяет класс 1 лучше, класс 2 хуже, а класс 3 — с одинаковой точностью. В сравнении с LR, OvO эта модель выделяет класс 1 хуже, а все остальные лучше. Для класса 4 характерен значительный разброс средних оценок точности.

Примеры наиболее информативных термометрических признаков, на основе которых строятся алгоритмы, представлены в таблице 2, в которой  $W$  — информативность,  $R_i$  — доля объектов класса  $i$ , у которых наблюдается признак. Все приведенные признаки описывают симметрию температурных полей голеней.

Таблица 2. Термометрические признаки

Признак	W	$R_0$	$R_1$	$R_2$	$R_3$	$R_4$
$\overline{T_c^{i,ir,st}} - \overline{T_p^{i,ir,st}} \in (-\infty, -0.288)$	48.32	0.0	0.7	0.18	0.16	0.2
$\ T_c^{i,ir,st} - T_p^{i,ir,st}\ _2 \in (1.967, 7.72)$	44.91	0.0	0.75	0.64	0.62	0.85
$\ T_c^{i,mw,ly} - T_p^{i,mw,ly}\ _1 \in (4.25, \infty)$	44.65	0.0	0.69	0.57	0.55	0.95
$\overline{T_c^{i,ir,st}} - \overline{T_p^{i,ir,st}} \in (-0.288, 0.279)$	44.44	1.0	0.25	0.48	0.39	0.15
$\ T_c^{i,mw,ly} - T_p^{i,mw,ly}\ _1 \in (1.45, 3.85)$	42.81	0.94	0.22	0.34	0.32	0.05
$\ T_c^{i,ir,ly} - T_p^{i,ir,ly}\ _1 \in (4.45, \infty)$	42.61	0.17	0.85	0.64	0.9	0.85
$\ T_c^{i,ir,st} - T_p^{i,ir,st}\ _\infty \in (1.25, \infty)$	41.62	0.0	0.73	0.64	0.54	0.7
$\overline{T_c^{i,ir,ly}} - \overline{T_p^{i,ir,ly}} \in (0.221, \infty)$	40.39	0.11	0.04	0.38	0.54	0.7
$\overline{T_c^{i,mw,ly}} - \overline{T_p^{i,mw,ly}} \in (-\infty, -0.096)$	39.98	0.14	0.75	0.31	0.16	0.25
$\ T_c^{i,mw,st} - T_p^{i,mw,st}\ _1 \in [0, 5.55)$	39.93	1.0	0.49	0.67	0.62	0.05

Три признака с наивысшей информативностью не наблюдаются у класса «Здоровые», т.е. позволяют эффективно выявлять голени с заболеваниями. Такими признаками являются разность средних и мера разброса поверхностных температур голеней, измеренных в положении стоя, а также мера разброса внутренних температур, измеренных в положении лежа.

Практически для всех здоровых голеней характерна невысокая разница средних значений кожных температур, измеренных в положении стоя, невысокий разброс внутренних температур, измеренных как стоя, так и лежа.

У класса 1 практически не наблюдается высокая разность средних значений внутренних или поверхностных температур, измеренных как стоя, так и лежа. При этом разница средних значений поверхностных температур стоя у них обычно выше, чем у класса 0.

У класса 4 не наблюдается невысокий разброс внутренних градиентов температур, измеренных в положении лежа. Практически не наблюдаются невысокий разброс внутренних температур, измеренных как стоя так и лежа.

Эти и другие свойства используются в алгоритмах взвешенного голосования признаков. А приведенные признаки сигнализируют о том, что для различных классов заболеваний, характерны высокие меры разброса внутренних и поверхностных температур голеней.

## **Заключение**

Наиболее эффективным универсальным алгоритмом для решения поставленной задачи оказался LR, OvO, имеющий среднюю сбалансированную точность 0.574. Однако, при применении иерархии классификаторов и сведении задачи к бинарной классификации «Здоров»–«Болен» с последующим уточнением заболевания, наилучший результат диагностики будет достигнут с LR, OvR со средней оценкой уточнения класса заболевания равной 0.548.

В работе [12] для решения задачи дифференциальной диагностики последовательно применялись три двухслойные нейронные сети. Каждая сеть отделяла один из классов от всех остальных, а оставшиеся голени помечались как «Здоровые». Точность диагностики составила 0.59. Детальное сравнение результатов не представляется возможным,

поскольку различаются как методы проверки алгоритмов, так и выборки. Несмотря на это, можно выделить следующие преимущества подхода, представленного в работе:

1. Возможность обосновать результат классификации. Каждый термометрический признак может быть интерпретирован;
2. Не используются данные анамнеза (наличие отека, боли, кожных изменений). Не исключено, что их применение может значительно повысить эффективность классификации. Впрочем, это представляет интерес для дальнейших исследований.

Результаты показывают применимость модели динамического описания состояния пациента в задаче дифференциальной диагностики венозных заболеваний. Ключевой особенностью построенных алгоритмов является возможность обоснования диагностического решения.

## Список литературы

- [1] *Дорожная карта развития «Сквозной» цифровой технологии «Нейро-технологии и искусственный интеллект»*, Министерство цифрового развития, связи и массовых коммуникаций Российской Федерации, 2019.  <sup>↑</sup><sub>19</sub>
- [2] M. Sedankin, A. Gudkov, V. Leushin, S. Vesnin, I. Sidorov, D. Chupina, S. Agasieva, V. Skuratov, S. Chizhikov. “Microwave radiometry of the pelvic organs”, *Biomedical Engineering*, **53**:4 (2019), pp. 288–292.  <sup>↑</sup><sub>20</sub>
- [3] A.V. Tarakanov, A.A. Tarakanov, S. Vesnin, V. Efremov, N. Roberts, I. Goryanin. “Influence of ambient temperature on recording of skin and deep tissue temperature in region of lumbar spine”, *European Journal of Molecular & Clinical Medicine*, **7**:1 (2020), pp. 21–26.  <sup>↑</sup><sub>20</sub>
- [4] I. Goryanin, S. Karbainov, O. Shevelev, A. Tarakanov, K. Redpath, S. Vesnin, Y. Ivanov. “Passive microwave radiometry in biomedical studies”, *Drug Discovery Today*, **25**:4 (2020), pp. 757–763.  <sup>↑</sup><sub>20</sub>
- [5] Т.В. Замечник, С.И. Ларин, А.Г. Лосев. *Комбинированная радиотермометрия как метод исследования венозного кровообращения нижних конечностей*, монография, Изд-во ВолгГМУ, Волгоград, 2015, ISBN 978-5-9652-0400-7, 252 с. <sup>↑</sup><sub>20</sub>
- [6] Т.В. Замечник, Н.С. Овчаренко, С.И. Ларин, А.Г. Лосев. «Изучение надежности комбинированной термографии как метода диагностики состояния вен нижних конечностей», *Флебология*, **4**:3 (2010), с. 23–26.   
 <sup>↑</sup><sub>20</sub>

- [7] Е.В. Анисимова, Т.В. Замечник, А.Г. Лосев. «Об одной математической модели диагностики венозных заболеваний по данным комбинированной термометрии», *Вестник новых медицинских технологий*, **19:2** (2012), с. 239–241.   [↑<sub>21,27</sub>](#)
- [8] Т.А. Ставров, Е.В. Букина, А.Г. Лосев, Т.В. Замечник. «Математическая модель верификации ранних рецидивов варикозной болезни по данным радиотермометрии», *Вестник новых медицинских технологий*, **20:2** (2013), с. 14–18.   [↑<sub>21,27</sub>](#)
- [9] А.Г. Лосев, В.В. Левшинский. «Интеллектуальный анализ термометрических данных в диагностике молочных желез», *Управление большими системами*, **70** (2017), с. 113–135.   [↑<sub>21,24</sub>](#)
- [10] V. Levshinskii. “Intelligent system for diagnostics of venous diseases based on the microwave radiothermometry data”, “Smart Technologies” for Society, State and Economy, ISC 2020, Lecture Notes in Networks and Systems, vol. **155**, eds. E. G. Popkova, B. S. Sergi, Springer, Cham, 2021, pp. 212–219.  [↑<sub>21,24,30</sub>](#)
- [11] Е.А. Мазепа, О.В. Гришина, В.В. Левшинский, Х.М. Сулейманова. «Об унификации метода анализа данных микроволновой радиотермометрии», *Математическая физика и компьютерное моделирование*, **20:6** (2017), с. 38–50.  [↑<sub>21</sub>](#)
- [12] Д.А. Веденяпин, А.Г. Лосев. «Об одной нейросетевой модели диагностики венозных заболеваний», *Управление большими системами*, **39** (2012), с. 219–229.  [↑<sub>21,33</sub>](#)
- [13] К.В. Воронцов. *Лекции по логическим алгоритмам классификации*, 2007, 53 с.  [↑<sub>28</sub>](#)
- [14] P. Flach. *Machine Learning. The Art and Science of Algorithms that Make Sense of Data*, 1st Edition, Cambridge University Press, 2012, ISBN 978-1107422223, 409 pp. [↑<sub>29</sub>](#)
- [15] S. Raschka, V. Mirjalili. *Python Machine Learning: Machine Learning and Deep Learning with Python, scikit-learn, and TensorFlow 2*, 3rd Ed., Packt Publishing, Birmingham, UK, 2019, ISBN 978-1789955750, 770 pp. [↑<sub>31</sub>](#)
- [16] M. Grandini, E. Bagli, G. Visani. *Metrics for multi-class classification: an overview*, 2020, 17 pp. arXiv:  2008.05756 [stat.ML] [↑<sub>31</sub>](#)

Поступила в редакцию 16.03.2021

Переработана 24.03.2021

Опубликована 14.04.2021

Рекомендовал к публикации

д.т.н. В. М. Хачумов

*Пример ссылки на эту публикацию:*

В. В. Левшинский. «Многоклассовая классификация в задаче дифференциальной диагностики венозных заболеваний по данным микроволновой радиотермометрии». *Программные системы: теория и приложения*, 2021, **12**:2(49), с. 19–36.

 10.25209/2079-3316-2021-12-2-19-36

 [http://psta.psiras.ru/read/psta2021\\_2\\_19-36.pdf](http://psta.psiras.ru/read/psta2021_2_19-36.pdf)

*Об авторе:*



### **Владислав Викторович Левшинский**

Аспирант Волгоградского государственного университета.  
Интересы: программирование, искусственный интеллект, анализ данных, машинное обучение, математическое моделирование.



0000-0002-3587-0124

e-mail: [v.levshinskii@volsu.ru](mailto:v.levshinskii@volsu.ru)

*Sample citation of this publication:*

Vladislav V. Levshinskiy. “Multiclass Classification in the Problem of Differential Diagnosis of Venous Diseases Based on Microwave Radiometry Data”. *Program Systems: Theory and Applications*, 2021, **12**:2(49), pp. 19–36. (*In Russian*).

 10.25209/2079-3316-2021-12-2-19-36

 [http://psta.psiras.ru/read/psta2021\\_2\\_19-36.pdf](http://psta.psiras.ru/read/psta2021_2_19-36.pdf)

*The same article in English:*



10.25209/2079-3316-2021-12-2-37-52