

Московский государственный университет имени М. В. Ломоносова  
Факультет Вычислительной математики и кибернетики  
Кафедра Математических методов прогнозирования

## ДИПЛОМНАЯ РАБОТА

### Анализ сигналов кровяного давления

**Выполнил:**

студент 517 группы

Харациди Олег Александрович

**Научный руководитель:**

д.ф-м.н., профессор

Дьяконов Александр Геннадьевич

Москва, 2015

# Содержание

<b>1</b>	<b>Введение</b>	<b>2</b>
<b>2</b>	<b>Постановка задачи</b>	<b>4</b>
<b>3</b>	<b>Описание данных</b>	<b>6</b>
<b>4</b>	<b>Меры близости между сигналами</b>	<b>7</b>
4.1	Косинусная мера . . . . .	7
4.2	DTW . . . . .	7
4.3	TWED . . . . .	9
<b>5</b>	<b>Исследование методов решения задачи и эксперименты</b>	<b>11</b>
5.1	Предобработка сигнала . . . . .	11
5.2	Выделение профиля пульсовой волны . . . . .	12
5.3	Классификация профилей пульсовой волны, полученных усреднением . . . . .	15
5.4	Классификация голосованием . . . . .	18
5.5	Усреднение с совмещением точек . . . . .	20
5.6	Тестирование на дополнительной выборке . . . . .	22
5.7	Выводы . . . . .	24
<b>6</b>	<b>Заключение</b>	<b>24</b>
	<b>Список литературы</b>	<b>25</b>

# 1 Введение

В последние несколько десятилетий роль информационных технологий в медицине стремительно растет. Появляются новые, часто более дешевые и менее болезненные, методы диагностики, исследований, а также повышается точность измерений.

Одна из многих сфер исследований в медицине, обязанных своим развитием информационным технологиям — анализ сигналов кровяного давления. Под сигналом кровяного давления понимается последовательность измерений давления или величины, характеризующей его, с некоторой достаточно высокой частотой. Наиболее распространенные методы измерения таких сигналов — *электрокардиография* и *фотоплетизмография*.

Принцип электрокардиографии основан на измерении напряженности электростатического поля, создаваемого сердцем, на поверхности тела. При сокращениях сердечной мышцы напряженность поля изменяется, формируя таким образом достаточно качественный сигнал — *электрокардиограмму*. Электрокардиография появилась в начале 20 века и широко применяется в медицине по сей день [1].

Метод фотоплетизмографии [2] заключается в просвечивании какого-либо органа (как правило, пальца руки) с помощью светодиода и измерения степени поглощения света датчиком с противоположной стороны или на некотором расстоянии. Полученный таким образом сигнал называется *фотоплетизмограммой* (*photoplethysmogram, PPG*). Значения фотоплетизмограммы отражают плотность кровяного потока в органе, а не значения давления в кровеносной системе. Тем не менее, эти величины достаточно сильно коррелируют, поэтому фотоплетизмо-

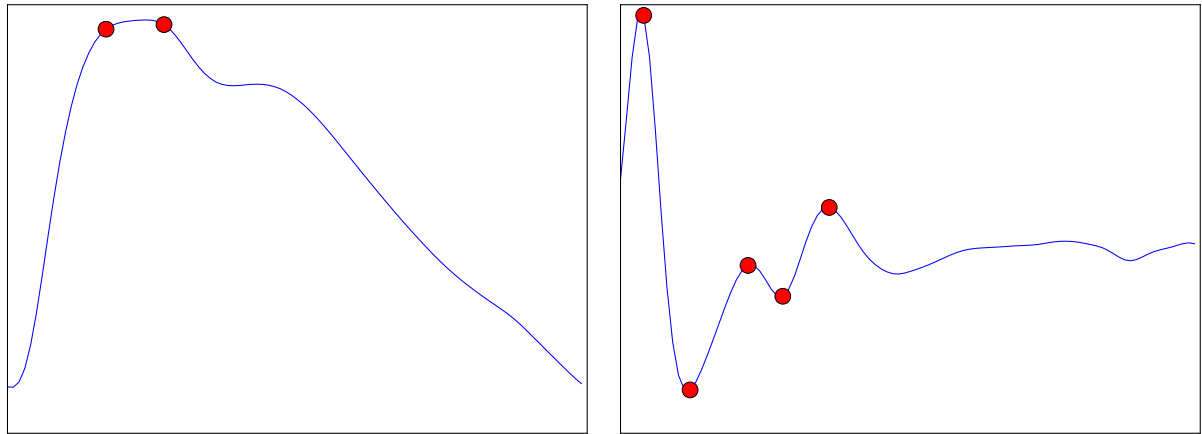
графия считается достаточно надежным способом получения сигналов кровяного давления.

При просвечивании пальца говорят о *периферийном давлении* (*peripheral pressure*), в то время как наиболее ценную информацию содержит *центральное*. В силу особенностей строения кровеносной системы сигнал периферийного давления существенно зашумлен и искажен. Однако, из-за своей простоты, дешевизны и безболезненности такой метод также нашел множество применений в медицине [3–7].

Структура сигнала кровяного давления не совсем проста. Считается, что сигнал состоит из нескольких компонент, в числе которых, например, волна Майера [8], волна дыхания, пульсовая волна. Последняя представляет немалый интерес для исследований: ее характеристики тесно связаны с состоянием сердечно-сосудистой системы человека. Помимо такого простого параметра пульсовой волны, как частота, рассматривают ее профиль, на исследовании которого основана данная работа.

К базовым численным характеристикам пульсовой волны обычно относят координаты некоторых «ключевых» точек (в частности, некоторых локальных экстремумов) ее профиля, а также второй производной профиля (Рис. 1).

На значениях координат этих точек основано множество рассчитываемых характеристик. Например, известно, что *индекс аугментации* (*augmentation index, AI*), имеет связь с риском сердечно-сосудистых заболеваний [9], жесткостью стенок аорты и наличием гиперхолестеринемии [10]. Также важным параметром является индекс жесткости (*stiffness index, SI*) [11, 12]. Известны и другие характеристики [13].



(a)

(б)

Рис. 1: Основные точки (а) профиля и (б) его второй производной, по которым рассчитываются некоторые характеристики пульсовой волны.

В последние годы наблюдается также активный интерес к сигналам кровяного давления в области обработки сигналов [14, 15], адаптируются и применяются многие классические методы, такие как метод главных компонент [16], преобразование Фурье [17], анализ независимых компонент [18] и другие. Разрабатываются алгоритмы для автоматической оценки параметров [19, 20]. Существуют даже методы оценки частоты пульса по фотоплетизмограммам, получаемым с помощью камеры и вспышки обычного смартфона [21]. Также известно множество исследований по применению методов статистики и машинного обучения [22–25].

## 2 Постановка задачи

В данной работе исследуются методы анализа пульсовой волны в сигналах кровяного давления человека. Как было сказано ранее, основная

ценность пульсовой волны заключается в том, что она содержит информацию о состоянии сердечно-сосудистой системы человека. Поэтому преследуется цель оценить способность рассматриваемых методов учитывать данную специфику сигналов.

Для этого рассматривается задача идентификации состояний сердечно-сосудистой системы. Однако полностью описать состояния практически невозможно из-за большого количества параметров и сложных и дорогостоящих процедур диагностики. По этой причине делается естественное предположение о том, что у разных людей состояние сердечно-сосудистой системы различается, но у каждого оно претерпевает лишь незначительные изменения в пределах небольшого промежутка времени, в течение которого производится серия замеров давления. В рамках такого предположения задача идентификации состояний сердечно-сосудистой системы человека сводится к задаче идентификации человека, которая и будет использоваться для оценки исследуемых методов.

Поскольку наибольший интерес при исследовании пульсовой волны представляет ее профиль, будем производить идентификацию в два этапа:

1. Выделение профиля пульсовой волны из сигнала.
2. Классификация человека по его профилю.

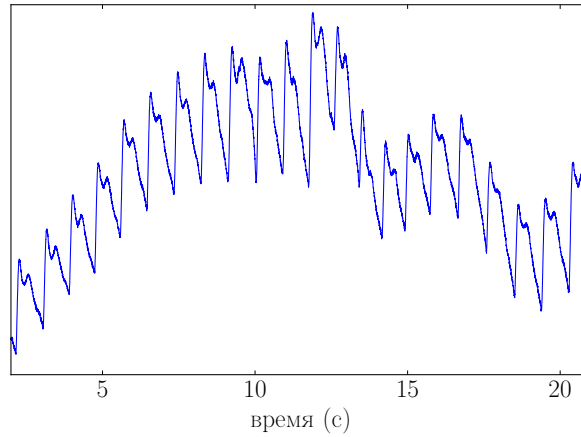


Рис. 2: Пример исходного сигнала.

### 3 Описание данных

Данные собраны при помощи измерительного прибора Ангиоскан-01М. Частота выходного сигнала — 1000Гц. Всего рассматривается два набора данных.

В первом наборе, предоставленном производителем прибора, содержатся фотоплетизмограммы 34 различных пациентов от 20 до 70 лет, всего 107 замеров, от 2 до 8 замеров на одного пациента. Каждый замер соответствует непрерывному отрезку времени длиной от 5 до 10 минут. Пример сигнала на отрезке в несколько секунд изображен на Рис. 2.

Дополнительная выборка собрана автором совместно с научным руководителем для повторения основных экспериментов с целью проверки полученных результатов. Она состоит из 36 замеров для 6 пациентов от 25 до 70 лет. Длина каждого замера от 50 до 80 секунд.

## 4 Меры близости между сигналами

Прежде чем перейти к описанию исследований, введем некоторые меры близости для сигналов. Для простоты будем определять их с точностью до домножения на  $-1$ . Предполагается, что все меры близости согласованы между собой путем явной замены значений на противоположные у некоторых из них.

### 4.1 Косинусная мера

Наиболее стандартная мера близости — косинусная мера:

$$\cos(X, Y) = \frac{\langle X, Y \rangle}{\|X\| \cdot \|Y\|}$$

Косинусная мера принимает значения из отрезка  $[-1, 1]$  и является одной из наиболее распространенных мер близости объектов в анализе данных.

### 4.2 DTW

Одна из мер, широко применяемых в анализе временных рядов, — Dynamic Time Warping Distance (DTW) [26, 27]. Существует множество ее модификаций [28]. Наиболее распространенная форма задается следующим образом.

Пусть  $d : \mathbb{N}_0 \times \mathbb{N}_0 \rightarrow \mathbb{R}$  — функция, определяемая по следующему правилу.

При  $i, j > 0$ ,  $|i - j| \leq \mu$  положим:



$$d(i, j) = (x_i - y_j)^2 + \min \begin{cases} d(i - 1, j) \\ d(i, j - 1) \\ d(i - 1, j - 1) \end{cases}$$

а также

$$\begin{aligned} d(0, 0) &= 0, \\ d(0, j) &= \infty, \quad 0 < j \leq l_y, \\ d(i, 0) &= \infty, \quad 0 < i \leq l_x, \\ d(i, j) &= \infty, \quad |i - j| > \mu. \end{aligned}$$

Здесь  $\mu \in \mathbb{N}_0$  — параметр, выполняющий роль регуляризатора, а  $l_x$  и  $l_y$  — соответственно длины сигналов  $X$  и  $Y$ .

Тогда значение DTW определим как

$$DTW(X, Y) = d(l_x, l_y).$$

Нетрудно видеть, что с помощью метода динамического программирования можно вычислить это значение и также восстановить «путь трансформации» — последовательность пар  $\{w_k = (u_k, v_k)\}_{k=0}^K$ , такую что

$$\begin{cases} w_0 = (0, 0), \\ w_K = (l_x, l_y), \\ d(w_k) = (x_{u_k} - y_{v_k})^2 + d(w_{k-1}), \quad 1 \leq k < K. \end{cases}$$

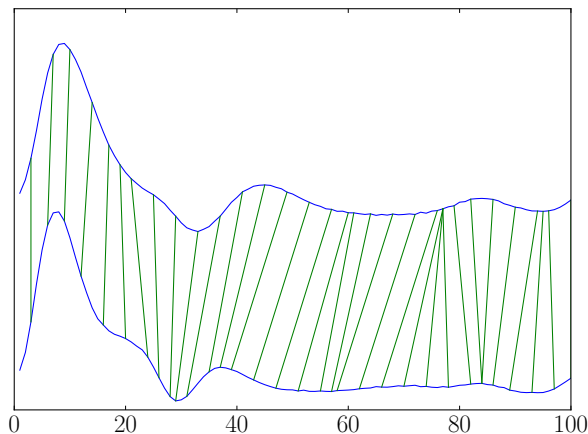


Рис. 3: Соответствие между точками двух сигналов.

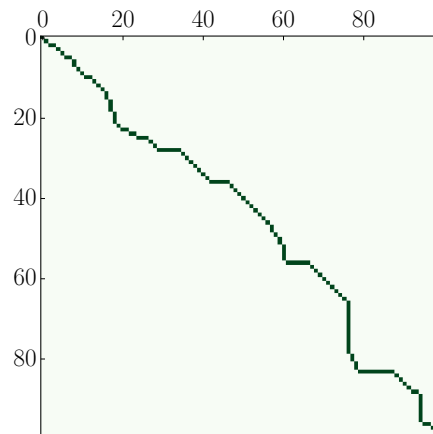


Рис. 4: Матрица пути трансформации.

Эта последовательность задает метод трансформации одного сигнала в другой, сопоставляя пары точек, переходящих друг в друга (Рис. 3). Можно далее определить разреженную матрицу с единицами в ячейках, задаваемых парами индексов  $\{w_k\}_{k=1}^K$ , и нулями — в остальных. Пример такой матрицы проиллюстрирован на Рис. 4.

Мера сходства DTW имеет множество применений в различных областях, таких как распознавание речи [29], биоинформатика [30], распознавание рукописных текстов [31], обработка и анализ сигналов [32] и других.

### 4.3 TWED

Другая мера для сравнения сигналов — Time Warp Edit Distance (TWED) [33]. Она по своей природе схожа с DTW, но изначально была предложена именно для анализа профиля пульсовой волны. Приведем упрощенную ее формулировку — для случая равных временных интервалов между последовательными отсчетами.

В рамках аналогичных обозначений, при  $i, j > 0$  положим

$$d(i, j) = \min \begin{cases} d(i-1, j) + |x_i - x_{i-1}| + \lambda, \\ d(i, j-1) + |y_i - y_{i-1}| + \lambda, \\ d(i-1, j-1) + |x_{i-1} - y_{i-1}| + |x_i - y_i|; \end{cases}$$

а также

$$\begin{aligned} d(0, 0) &= 0, \\ d(0, j) &= \infty, \quad 0 < j \leq l_y, \\ d(i, 0) &= \infty, \quad 0 < i \leq l_x. \end{aligned}$$

Здесь  $\lambda > 0$  — параметр, выполняющий роль регуляризации. Для определенности также считается, что  $x_0 = y_0 = 0$ .

Значение TWED определяется как:

$$DTW(X, Y) = d(l_x, l_y).$$

TWED, так же как и DTW, вычисляется вместе со своим «путем трансформации» методом динамического программирования и имеет очень схожую интерпретацию.

Одно из отличий TWED от DTW состоит в том, что регуляризация накладывает более «мягкие» ограничения на путь трансформации. Другая особенность меры TWED заключается в том, что она является метрикой. Это позволяет, например, построить на ее основе ядро SVM [34].

## 5 Исследование методов решения задачи и эксперименты

Итак, в данной работе рассматривается задача классификации сигнала кровяного давления по принадлежности к конкретному человеку. Таким образом, объекты выборки — это замеры, а метки классов — номера людей.

В рамках предложенного метода исследования возникает несколько подзадач, каждая из которых рассмотрена в этом разделе. Предполагается, что все преобразования сигналов производятся над каждым отдельно взятым замером независимо и в терминах задачи классификации представляют собой, таким образом, отображение объектов (замеров) в признаковое пространство.

### 5.1 Предобработка сигнала

У исходного сигнала есть ряд существенных недостатков, поэтому над ним производится предобработка в несколько этапов.

Сперва от сигнала отсекаются первая и последняя секунды измерений, поскольку в это время возможны передвижения пациента, сильно искажающие показания измерительного прибора.

Также необходимо избавиться от низкочастотного шума, который возникает в следствие движений тела пациента в процессе измерений, разовых смещений датчика относительно пальца пациента, а также постепенных изменений давления со временем. Для этого с помощью свертки с гауссовым ядром шириной в 5000 отсчетов выделяется тренд,

который затем вычитается из исходного сигнала (Рис. 5,6). Обозначим полученный сигнал как  $s(t)$ .

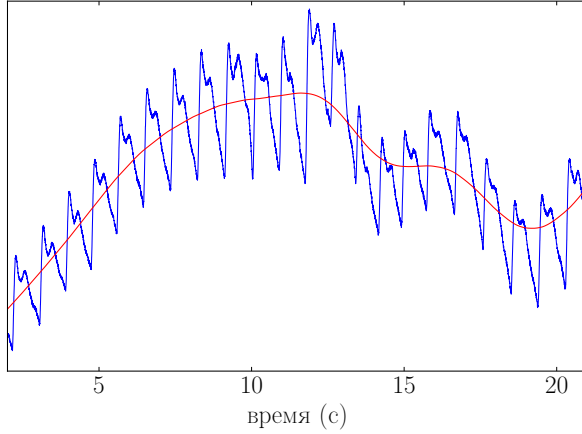


Рис. 5: Исходный сигнал и его тренд.

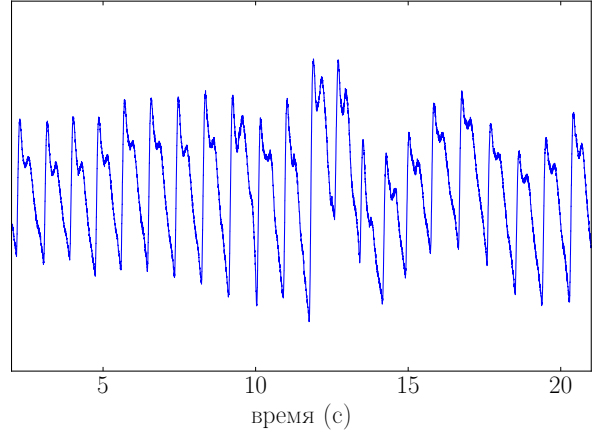


Рис. 6: Сигнал после вычитания тренда ( $s(t)$ ).

К сигналу  $s(t)$  независимо применяются свертки с гауссовым ядром шириной в 100 и  $w$  отсчетов, где  $w$  — параметр, процедура настройки которого будет описана в дальнейшем. В результате получается два новых сигнала:  $s_{100}(t) = s(t) * h_{100}(t)$ ,  $s_w(t) = s(t) * h_w(t)$  (здесь  $h_k(t)$  — гауссово ядро свертки шириной в  $k$  отсчетов). Далее в основном будет использоваться сигнал  $s_w(t)$ . Сигнал  $s_{100}(t)$  будет рассматриваться как вспомогательный и понадобится только для разметки пульсовых ударов.

## 5.2 Выделение профиля пульсовой волны

Выделение профиля пульсовой волны, а также его производных будем осуществлять в несколько этапов: сперва из замера длиной в несколько минут извлечем фрагменты, соответствующие отдельным ударам сердца, затем эти фрагменты независимо преобразуем, а затем усредним.

Ниже весь процесс изложен более подробно.

1. Сначала необходимо определить временные границы отдельных импульсных ударов. Импульсные удары, как правило, имеют непостоянную частоту в пределах 0.8–1.5 Гц.

Разметку границ импульсных ударов будем производить по значениям сигнала  $s_{100}$ : на каждом отрезке между последовательными моментами смены знака сигнала с положительного на отрицательный отмечается точка минимума. Обозначим через  $m < 0$  медиану значений сигнала в этих точках и уберем из рассмотрения те из них, которые превосходят  $\frac{m}{2}$ . Оставшиеся точки будем считать границами импульсных ударов (Рис. 7). Фрагменты сигнала между ними образуют совокупность единичных экземпляров профиля импульсной волны.

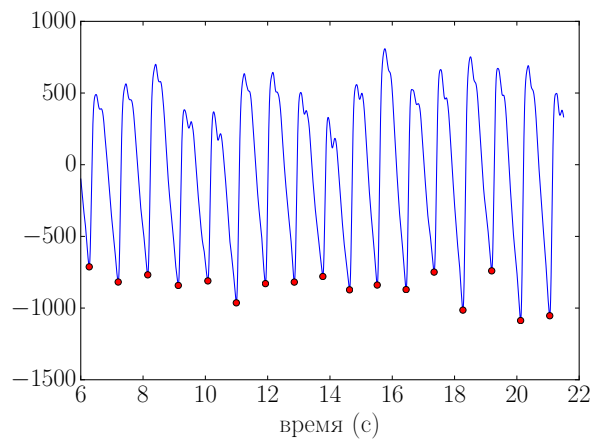


Рис. 7: Пример разметки границ фрагментов импульсной волны.

Однако иногда границы все еще расставляются неправильно: бывают «пропуски» и «лишние» точки. Поэтому из совокупности полученных фрагментов сигнала удалим 10% самых длинных, 10% самых коротких. Кроме того, удалим также 10% фрагментов с наибольшей по модулю разностью значений сигнала на концах —

таким образом отбрасываются самые «ненадежные» сегменты (предположительно, подвергшиеся сильным искажениям).

Разметка границ пульсовых ударов, полученная таким образом из сигнала  $s_{100}$ , применяется к  $s_w$ . Сигнал  $s_{100}$  в дальнейшем не используется.

2. Полученные фрагменты имеют, как правило, длину от 600 до 1200 отсчетов. Каждый из фрагментов выровним путем вычитания линейной функции (линейного тренда) так, чтобы значения на концах были равны нулю.
3. Затем применим разностное дифференцирование порядка  $n \geq 0$  (под дифференцированием порядка  $n = 0$  здесь и далее подразумевается тождественное преобразование).
4. После этого отмасштабируем фрагменты, приведя их к длине в  $T = 100$  отсчетов, а затем умножим на константу, чтобы привести к евклидовой норме, равной  $T$ . Пример набора таких фрагментов, соответствующих одному замеру, изображен на Рис. 8.

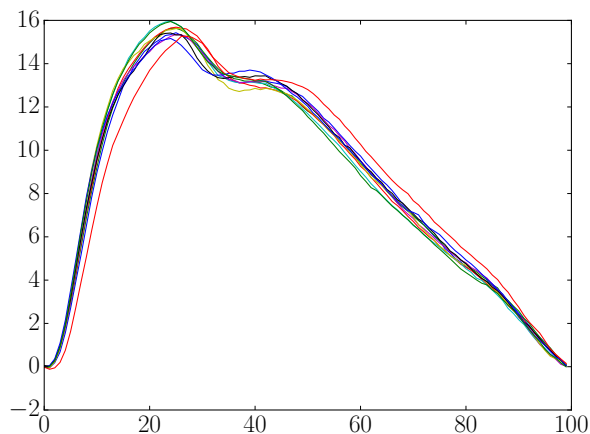


Рис. 8: Выделенные фрагменты пульсовой волны для  $n = 0$ .

5. Все полученные фрагменты, соответствующие одному исходному замеру, усредним между собой.

Таким образом, из одного замера сигнала кровяного давления получается профиль пульсовой волны или ее производной (Рис. 9).

Описанная эвристическая процедура дает фрагменты приемлемого качества. Пункты 1–3 достаточно естественны и в дальнейшем останутся без изменений, тогда как пункты 4–5 будут пересмотрены.

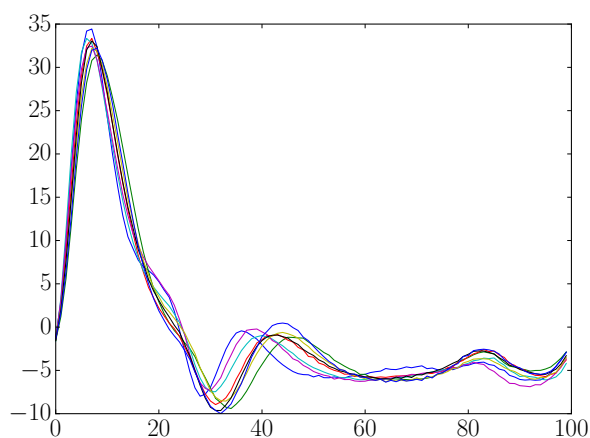


Рис. 9: Первые производные профилей пульсовой волны ( $n = 1$ ) для одного человека.

### 5.3 Классификация профилей пульсовой волны, полученных усреднением

В качестве признаков описаний при классификации будут рассматриваться профили пульсовой волны, процедура получения которых описана выше. Чтобы проанализировать и сравнить между собой меры близости сигналов, классификацию будем производить наиболее простым и естественным метрическим классификатором — методом



одного ближайшего соседа (Nearest Neighbour). Качество классификации будем оценивать на скользящем контроле по методу leave-one-out. В качестве численного результата будем рассматривать две величины:

- Ассигасу — доля совпадений ответов алгоритма с метками классов.
- Усредненное по всем классам значение функционала качества Area Under Curve (AUC). А именно — для каждой метки класса  $C_i$  исходная задача и полученный результат интерпретируются в рамках задачи бинарной классификации с классами « $C_i$ » и «не  $C_i$ », вычисляется AUC, а затем полученные значения AUC усредняются по всем  $i$ . В качестве оценки степени принадлежности объекта к классу здесь используется расстояние до ближайшего объекта из этого класса, умноженное на  $-1$ .

Второй функционал качества (усредненный AUC) будем считать основным, поскольку он, в отличие от ассигасу, не предвзят по отношению к классам с ббльшим числом объектов. Кроме того, множество значений AUC значительно шире, что в условиях небольшой выборки делает этот функционал более устойчивым к шуму и более чувствительным небольшим изменениям параметров классификатора, чем ассигасу.

Сначала определим оптимальный порядок дифференцирования, а также исследуем влияние ширины  $w$  ядра свертки при сглаживании сигнала. Зависимость качества классификации методом ближайшего соседа с косинусной мерой от параметра  $w$  для производных порядков  $n = 0, 1, 2$  изображена на Рис. 10.

Стоит заметить, что значения ассигасу и AUC достаточно сильно коррелируют, поэтому ассигасу в данной задаче можно также считать

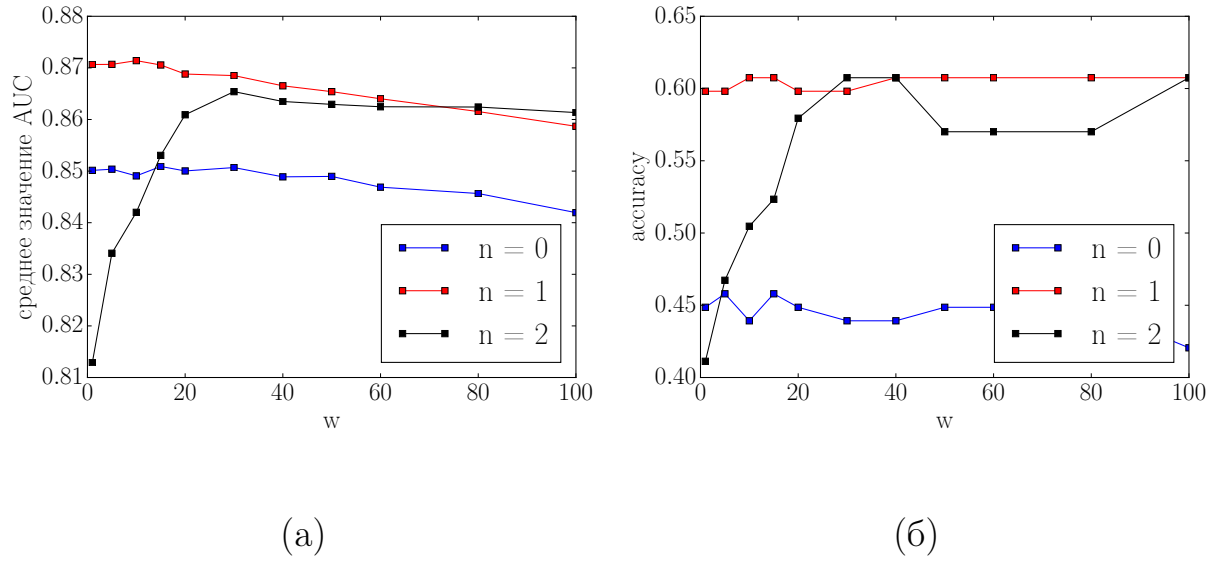


Рис. 10: Зависимость качества классификации, оцениваемого с помощью (а) среднего AUC и (б) ассигасу, от ширины  $w$  ядра свертки для производных порядков  $n = 0, 1, 2$ .

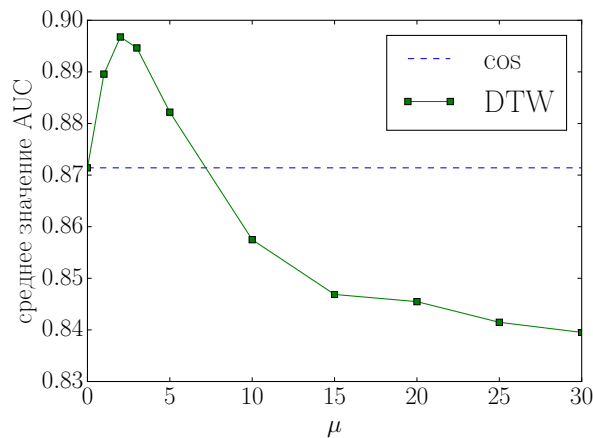
относительно достоверным, хотя и несколько зашумленным, функционалом качества. Это свойство понадобится в дальнейшем.

По значениям AUC видно, что первая производная показывает наилучшее качество, а оптимальное значение  $w$  для нее близко к 1. Поэтому зафиксируем значения  $n = 1$  и  $w = 1$ .

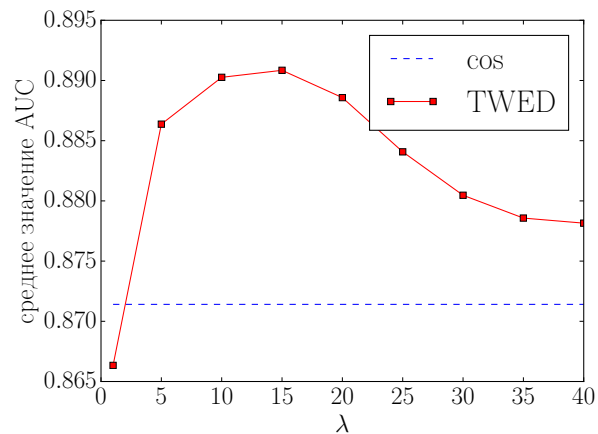
Далее, сравним меры близости DTW и TWED с косинусной мерой при разных значениях параметров  $\mu$  и  $\lambda$ , соответственно. Соответствующие значения качества классификации приведены на Рис. 11.

Если в данном эксперименте просуммировать матрицы путей трансформации TWED для всех пар профилей при разных значениях  $\lambda$ , получатся матрицы, изображенные на Рис. 12.

Итак, руководствуясь результатами эксперимента, установим значения параметров:  $\mu = 2$ ,  $\lambda = 15$ . Эти значения и будем использовать в дальнейшем.



(a)



(б)

Рис. 11: Зависимость среднего значения AUC от параметров для (а) DTW и (б) TWED в сравнении с косинусной мерой.

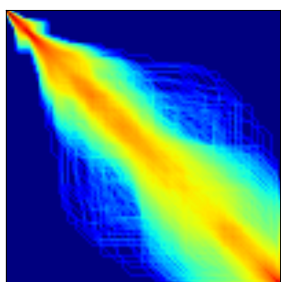
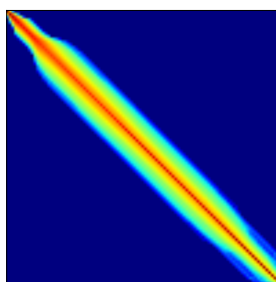
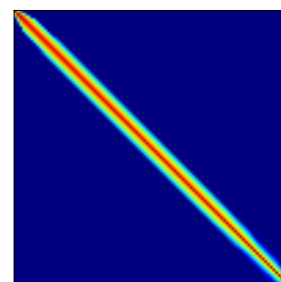
 $\lambda = 1$  $\lambda = 20$  $\lambda = 40$ 

Рис. 12: Теплокарта суммы матриц путей трансформации TWED при  $\lambda = 1, 20, 40$ .

Меры DTW и TWED, таким образом, дают схожее качество, лучшее, чем качество классификации с косинусной мерой.

## 5.4 Классификация голосованием

Вернемся к методу выделения пульсовой волны, описанному выше в главе 5.2, и поставим вопрос о целесообразности усреднения отдельных

фрагментов пульсовой волны (п.5). Будем классифицировать фрагменты по отдельности, используя тот же самый метод, а затем классифицировать исходные замеры голосованием по принципу большинства. Однако, в силу чрезмерной вычислительной сложности такого эксперимента, вместо всех фрагментов, количество которых в одном замере колеблется от 150 до 400, будем использовать только по  $f$  из каждого замера, взятых примерно через равные интервалы.

Кроме того, заметим также, что DTW и TWED, в отличие от косинусной меры, не требуют равенства длин сравниваемых сигналов. Поэтому для этих мер сходства, в п.4 той же главы положим значение  $T$  равным не 100, а  $\frac{1}{10}$  от исходной длины соответствующего фрагмента. В этом случае после масштабирования фрагменты будут иметь разные длины, но в среднем по-прежнему около 100 отсчетов.

Итак, применяя классификацию голосованием, при различных значениях  $f$  получаем результат, изображенный на графике на Рис. 13.

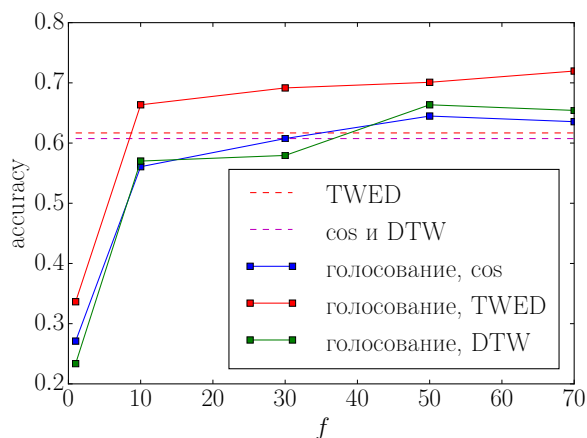


Рис. 13: Сравнение значений accuracy для усреднения и голосования при различных значениях  $f$ .

Видно, что при  $s \geq 50$  такая схема несколько улучшает результат. Для меры TWED это улучшение достаточно существенное.

## 5.5 Усреднение с совмещением точек

Теперь рассмотрим альтернативный способ усреднения фрагментов для формирования профиля пульсовой волны.

Как было сказано ранее, меры близости DTW и TWED при сравнении двух сигналов помимо численного значения расстояния предоставляют множество пар «похожих» точек сигналов. При сравнении двух фрагментов из одного измерения логично предположить, что это соответствие между точками будет достаточно достоверным, поскольку такие фрагменты, как правило, имеют схожую структуру (например, как на Рис. 3).

Опираясь на это предположение, предложим несколько иной метод усреднения фрагментов: будем «совмещать» соответствующие точки сигналов, предварительно применив нелинейное преобразование к их абсциссам.

Более формально, рассмотрим множество фрагментов  $\{X^{(i)}(t)\}$ ,  $i = 1, \dots, f$  из одного замера. Рассмотрим отдельно взятый фрагмент  $X^{(k)}(t)$  ( $t = 1, \dots, l$ ,  $k$  фиксировано) и все его матрицы  $M^{(1)}, \dots, M^{(f)}$  путей трансформации при сравнении с фрагментами  $X^{(1)}, \dots, X^{(f)}$ , соответственно. Будем считать, что  $\forall i$

$$M_{u,v}^{(i)} = \begin{cases} 1, & \text{если путь трансформации между } X^{(k)} \text{ и } X^{(i)} \\ & \text{содержит пару } (u, v), \\ 0, & \text{иначе.} \end{cases}$$

Затем для фиксированного  $t$  положим

$$t' = \frac{1}{f} \sum_i \frac{\sum_j M_{t,j}^{(i)} \cdot j}{\sum_j M_{t,j}^{(i)}},$$

то есть,  $t'$  — среднее значение абсциссы точки, в которую при трансформации переходит точка  $(t, X^{(k)}(t))$  сигнала  $X^{(k)}$ . Введем новый сигнал  $\tilde{X}^{(i)}$ , такой что  $\tilde{X}^{(k)}(t') = X^{(i)}(t)$ . Определим таким образом значения сигнала  $\tilde{X}^{(i)}(t')$  для всех  $t$ , а затем с помощью линейной интерполяции вычислим его значения при  $t' = 1, \dots, [l']$ .

Таким образом, из фрагмента  $X^{(i)}$  получен новый фрагмент  $\tilde{X}^{(i)}$ . Повторив эту процедуру для всех  $i = 1, \dots, f$ , продолжим преобразования в соответствии с исходным методом усреднения: приведем все фрагменты к длине в  $T = 100$  отсчетов, пронормируем и усредним. Получим профиль пульсовой волны.

В качестве меры близости в таком методе усреднения можно использовать и DTW, и TWED. Ограничимся рассмотрением меры TWED.

Итак, рассмотрим классификацию методом ближайшего соседа с косинусной мерой сходства, описанным в главе 5.3, с тем лишь отличием, что вместо обычного усреднения фрагментов одного замера будем усреднять фрагменты новым способом с совмещением соответствующих точек сигналов. Будем варьировать параметр  $f$  и выбирать из каждого замера  $f$  по фрагментов примерно через равные интервалы, либо все фрагменты замера, если их меньше, чем  $f$ . Сравним предложенный метод с исходным. Результаты экспериментов приведены на Рис. 14.

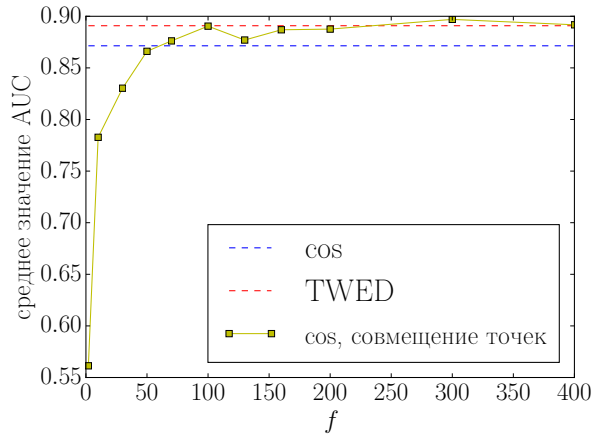


Рис. 14: Сравнение методов усреднения.

Таким образом, предложенный метод усреднения дает лучшее качество классификации, чем исходный, и примерно равное качеству исходного метода ближайшего соседа с мерой близости TWED.

## 5.6 Тестирование на дополнительной выборке

Чтобы проверить полученные результаты, проведем также некоторые эксперименты на дополнительной выборке. Однако, она существенно меньше, поэтому сравнения результатов по функционалу качества ассигасу не имеют смысла. Проверим только основные результаты, основанные на функционале качества AUC.

Результат эксперимента по настройке параметра  $w$  для  $n = 1$  представлен на Рис. 15.

Оптимальное значение  $w = 25$ . В отличие от основной выборки, где оптимальная ширина  $w$  свертки при сглаживании была равна 10, в данном случае количество фрагментов в одном замере колеблется

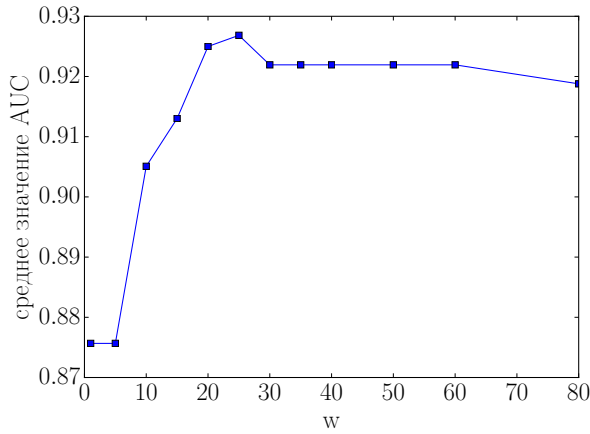


Рис. 15: Подбор оптимального значения  $w$ .

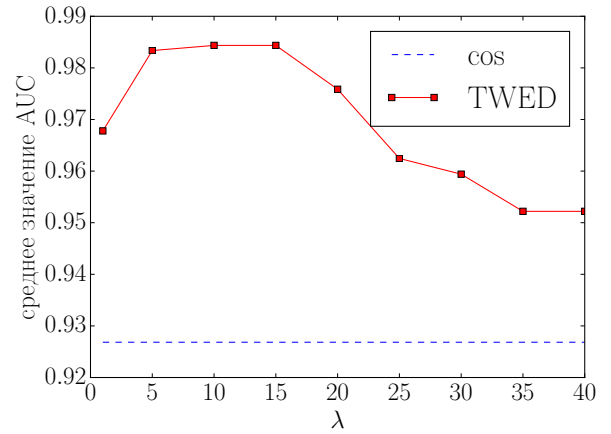


Рис. 16: Подбор оптимального значения  $\lambda$ .

в пределах от 40 до 80, что существенно меньше, поэтому при их усреднении эффект подавления шума проявляется в меньшей степени.

Далее, фиксируя  $n = 1$ ,  $w = 25$ , настроим параметр  $\lambda$  для метода ближайшего соседа с мерой сходства TWED (Рис. 16) и получим  $\lambda \in [10, 15]$ , что примерно соответствует значению  $\lambda = 15$ , полученному на основной выборке. Положим, таким образом,  $\lambda = 15$ .

Наконец, сравним методы усреднения. Результаты экспериментов (Рис. 17) подтверждают, что предложенный в разделе 5.5 метод превосходит обычное усреднение. В то же время, в данном случае он оказался несколько хуже, чем стандартный метод ближайшего соседа с мерой сходства TWED.



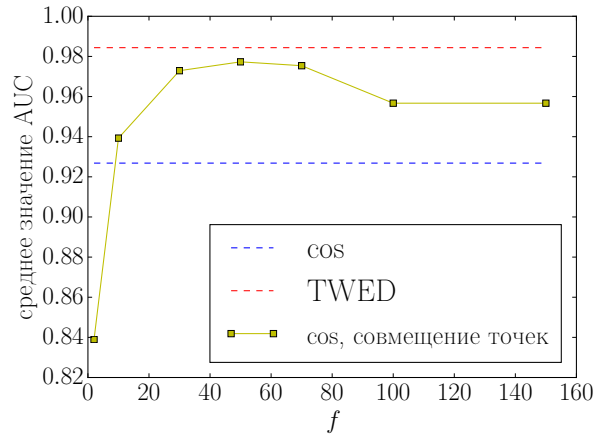


Рис. 17: Сравнение методов усреднения.

## 5.7 Выводы

По результатам проведенных исследований можно сделать следующие выводы:

- меры сходства DTW и TWED имеют некоторое превосходство над косинусной мерой при сравнении профилей пульсовой волны;
- простое усреднение — неоптимальный метод выделения профиля пульсовой волны из множества фрагментов;
- усреднение с совмещением точек более эффективно, чем простое усреднение.

## 6 Заключение

В ходе проделанной работы было предложено использование задачи идентификации человека для сравнения методов анализа сигналов кровяного давления. Экспериментально проведено сравнение трех мер

близости сигналов, одна из которых легла в основу предложенного автором метода усреднения сигналов с совмещением соответствующих точек. В ходе экспериментов также показана эффективность этого метода в задаче выделения профиля пульсовой волны.

Для контроля полученных результатов была собрана дополнительная выборка, эксперименты на ней подтвердили основные выводы.

Все описанные методы реализованы на связке языков MATLAB и C++.

## Список литературы

1. Marriott H. J. L., Wagner G. S. Practical electrocardiography. — Williams & Wilkins Baltimore, 1988.
2. Wearable photoplethysmographic sensors — past and present / Toshiyo Tamura, Yuka Maeda, Masaki Sekine, Masaki Yoshida // Electronics. — 2014. — Vol. 3, no. 2. — P. 282–302.
3. Allen J. Photoplethysmography and its application in clinical physiological measurement // Physiological measurement. — 2007. — Т. 28, № 3. — С. R1.
4. Эндотелиальная функция у пациентов с артериальной гипертензией высокого риска / О.М. Драпкина, О.Н. Дикур, Я.И. Ашихмин et al. // Артериальная гипертензия. — 2010. — Vol. 16, no. 2. — P. 156–163.
5. Assessment of vasoactive agents and vascular aging by the second derivative of photoplethysmogram waveform / Kenji Takazawa,

- Nobuhiro Tanaka, Masami Fujita et al. // Hypertension. — 1998. — Vol. 32, no. 2. — P. 365–370.
6. Utility of second derivative of the finger photoplethysmogram for the estimation of the risk of coronary heart disease in the general population / Toshiaki Otsuka, Tomoyuki Kawada, Masao Katsumata, Chikao Ibuki // Circulation Journal. — 2006. — Vol. 70, no. 3. — P. 304–310.
  7. Assessment of vascular aging and atherosclerosis in hypertensive subjects: second derivative of photoplethysmogram versus pulse wave velocity / Luiz A Bortolotto, Jacques Blacher, Takeshi Kondo et al. // American journal of hypertension. — 2000. — Vol. 13, no. 2. — P. 165–171.
  8. Julien C. The enigma of mayer waves: facts and models // Cardiovascular research. — 2006. — Vol. 70, no. 1. — P. 12–21.
  9. Augmentation index is associated with cardiovascular risk / Jens Nürnbergger, Ayten Keflioglu-Scheiber, Anabelle M Opazo Saez et al. // Journal of hypertension. — 2002. — Vol. 20, no. 12. — P. 2407–2414.
  10. Increased central pulse pressure and augmentation index in subjects with hypercholesterolemia / Ian B Wilkinson, Krishna Prasad, Ian R Hall et al. // Journal of the American College of Cardiology. — 2002. — Vol. 39, no. 6. — P. 1005–1011.
  11. Ambulatory arterial stiffness index as a predictor of cardiovascular mortality in the dublin outcome study / Eamon Dolan, Lutgarde Thijs, Yan Li et al. // Hypertension. — 2006. — Vol. 47, no. 3. — P. 365–370.

12. Increased ambulatory arterial stiffness index is associated with target organ damage in primary hypertension / Giovanna Leoncini, Elena Ratto, Francesca Viazzi et al. // *Hypertension*. — 2006. — Vol. 48, no. 3. — P. 397–403.
13. Arterial stiffness contributes to coronary artery disease risk prediction beyond the traditional risk score (rama-egat score) / Teerapat Yingchoncharoen, Thosaphol Limpijankit, Sutipong Jongjirasiri et al. // *Heart Asia*. — 2012. — Vol. 4, no. 1. — P. 77–82.
14. A review on recent patents in digital processing for cardiac electric signals (i): From basic systems to arrhythmia analysis / Rebeca Goya-Esteban, Óscar Barquero-Pérez, Felipe Alonso-Atienza et al. // *Recent Patents on Biomedical Engineering*. — 2009. — Vol. 2, no. 1. — P. 22–31.
15. A review on recent patents in digital processing for cardiac electric signals (ii): Advanced systems and applications / Óscar Barquero-Pérez, Rebeca Goya-Esteban, Felipe Alonso-Atienza et al. // *Recent Patents on Biomedical Engineering*. — 2009. — Vol. 2, no. 1. — P. 32–47.
16. Reddy K. A., Kumar V. J. Motion artifact reduction in photoplethysmographic signals using singular value decomposition // *Instrumentation and Measurement Technology Conference Proceedings, 2007. IMTC 2007. IEEE / IEEE*. — 2007. — P. 1–4.
17. Multiparameter respiratory rate estimation from the photoplethysmogram / Walter Karlen, Srinivas Raman, J Mark Ansermino, Guy A Dumont // *Biomedical Engineering, IEEE Transactions on*. — 2013. — Vol. 60, no. 7. — P. 1946–1953.

18. Gbaoui L., Kaniusas E. Decomposition of photoplethysmographical arterial pulse waves by independent component analysis: possibilities and limitations // *Advances in Biomedical Sensing, Measurements, Instrumentation and Systems*. — Springer, 2010. — P. 166–185.
19. Estimating respiratory and heart rates from the correntropy spectral density of the photoplethysmogram / Ainara Garde, Walter Karlen, J Mark Ansermino, Guy A Dumont // *PloS one*. — 2014. — Vol. 9, no. 1. — P. e86427.
20. Deriving respiration from the pulse photoplethysmographic signal / Jesús Lázaro, Eduardo Gil, Raquel Bailón, Pablo Laguna // *Computing in Cardiology, 2011 / IEEE*. — 2011. — P. 713–716.
21. Non-invasive heart rate measuring smartphone applications using on-board cameras: A short survey / Tarikul Islam Papon, Ishtiyaque Ahmad, Nazmus Saquib, Ashikur Rahman // *Networking Systems and Security (NSysS), 2015 International Conference on / IEEE*. — 2015. — P. 1–6.
22. *Advances in cardiac signal processing* / Rajendra Acharya, Shankar M Krishnan, Jos AE Spaan, Jasjit S Suri. — Springer, 2007.
23. Jerrard-Dunne P., Mahmud A., Feely J. Ambulatory arterial stiffness index, pulse wave velocity and augmentation index—interchangeable or mutually exclusive measures? // *Journal of hypertension*. — 2008. — Vol. 26, no. 3. — P. 529–534.

24. González Rubio T., Salgado Castillo A. Evaluation of ann and svm for the classification and prediction of patients with diabetic neuropathy // XVIII Congreso Argentino de Ciencias de la Computación. — 2012.
25. Salem A., Revett K., El-Dahshan E.-S. Machine learning in electrocardiogram diagnosis // Computer Science and Information Technology, 2009. IMCSIT'09. International Multiconference on / IEEE. — 2009. — P. 429–433.
26. Berndt D. J., Clifford J. Using dynamic time warping to find patterns in time series. // KDD workshop / Seattle, WA. — Vol. 10. — 1994. — P. 359–370.
27. Ratanamahatana C. A., Keogh E. Everything you know about dynamic time warping is wrong // Third Workshop on Mining Temporal and Sequential Data. — 2004. — P. 22–25.
28. Keogh E. J., Pazzani M. J. Derivative dynamic time warping. // SDM / SIAM. — Vol. 1. — 2001. — P. 5–7.
29. Muda L., Begam M., Elamvazuthi I. Voice recognition algorithms using mel frequency cepstral coefficient (mfcc) and dynamic time warping (dtw) techniques // arXiv preprint arXiv:1003.4083. — 2010.
30. Aach J., Church G. M. Aligning gene expression time series with time warping algorithms // Bioinformatics. — 2001. — Vol. 17, no. 6. — P. 495–508.
31. Bahlmann C., Haasdonk B., Burkhardt H. Online handwriting recognition with support vector machines—a kernel approach // Frontiers

in handwriting recognition, 2002. proceedings. eighth international workshop on / IEEE. — 2002. — P. 49–54.

32. Classification of motor activities through derivative dynamic time warping applied on accelerometer data / Rossana Muscillo, Silvia Conforto, Maurizio Schmid et al. // Engineering in Medicine and Biology Society, 2007. EMBS 2007. 29th Annual International Conference of the IEEE / IEEE. — 2007. — P. 4930–4933.
33. Marteau P.-F. Time warp edit distance with stiffness adjustment for time series matching // Pattern Analysis and Machine Intelligence, IEEE Transactions on. — 2009. — Vol. 31, no. 2. — P. 306–318.
34. Jia D., Zhang D., Li N. Pulse waveform classification using support vector machine with gaussian time warp edit distance kernel // Computational and mathematical methods in medicine. — 2014. — Vol. 2014.