



# **Метод выделения звуков естественного языка в звучащей речи**

Студент: Левушкин Илья Кириллович, ИУ7-82Б

Научный руководитель: Градов Владимир Михайлович

Консультант: Строганов Юрий Владимирович

# Цель и задачи

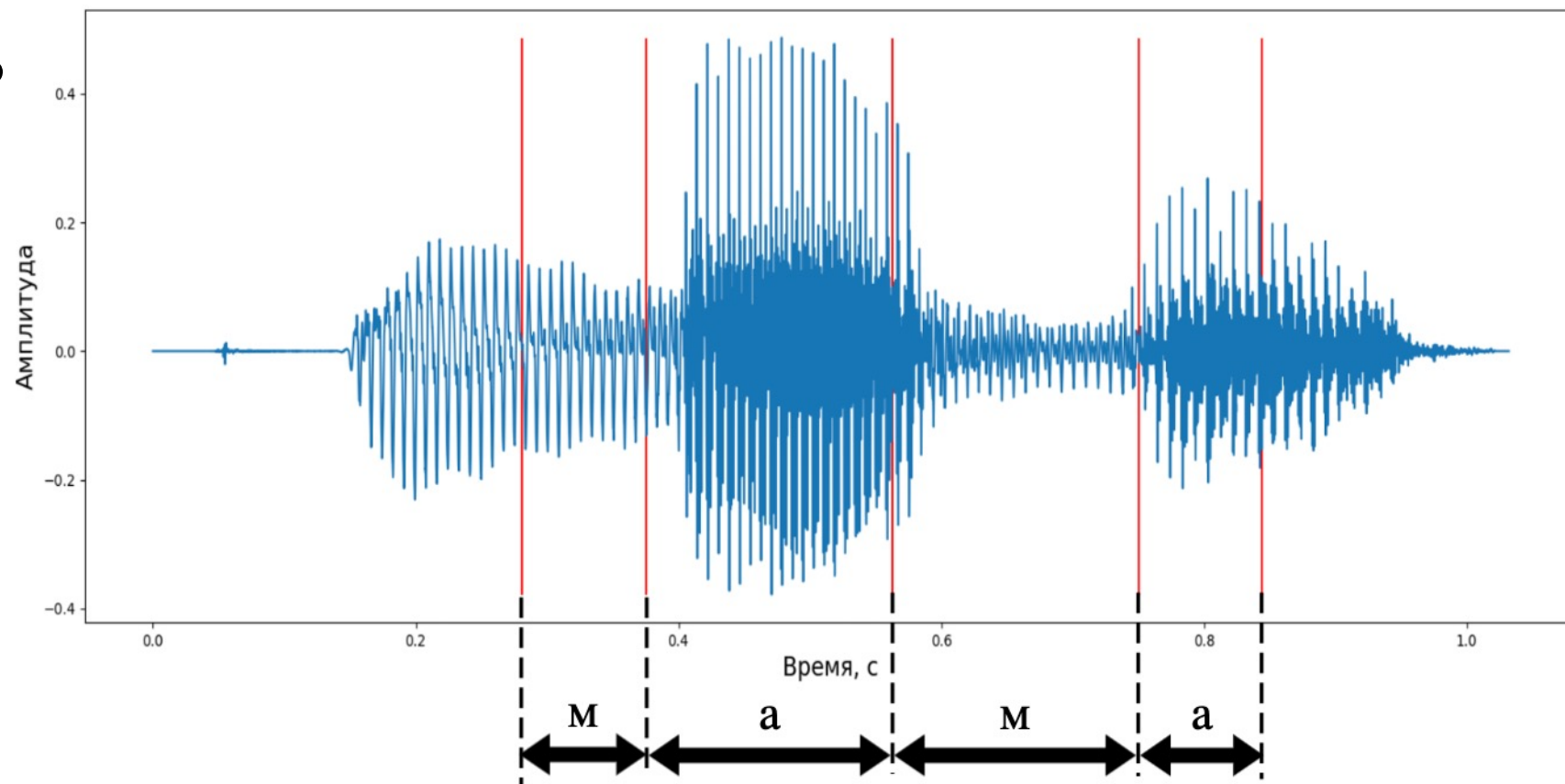
Цель – разработка метода выделения звуков естественного языка в звучащей речи.

Задачи:

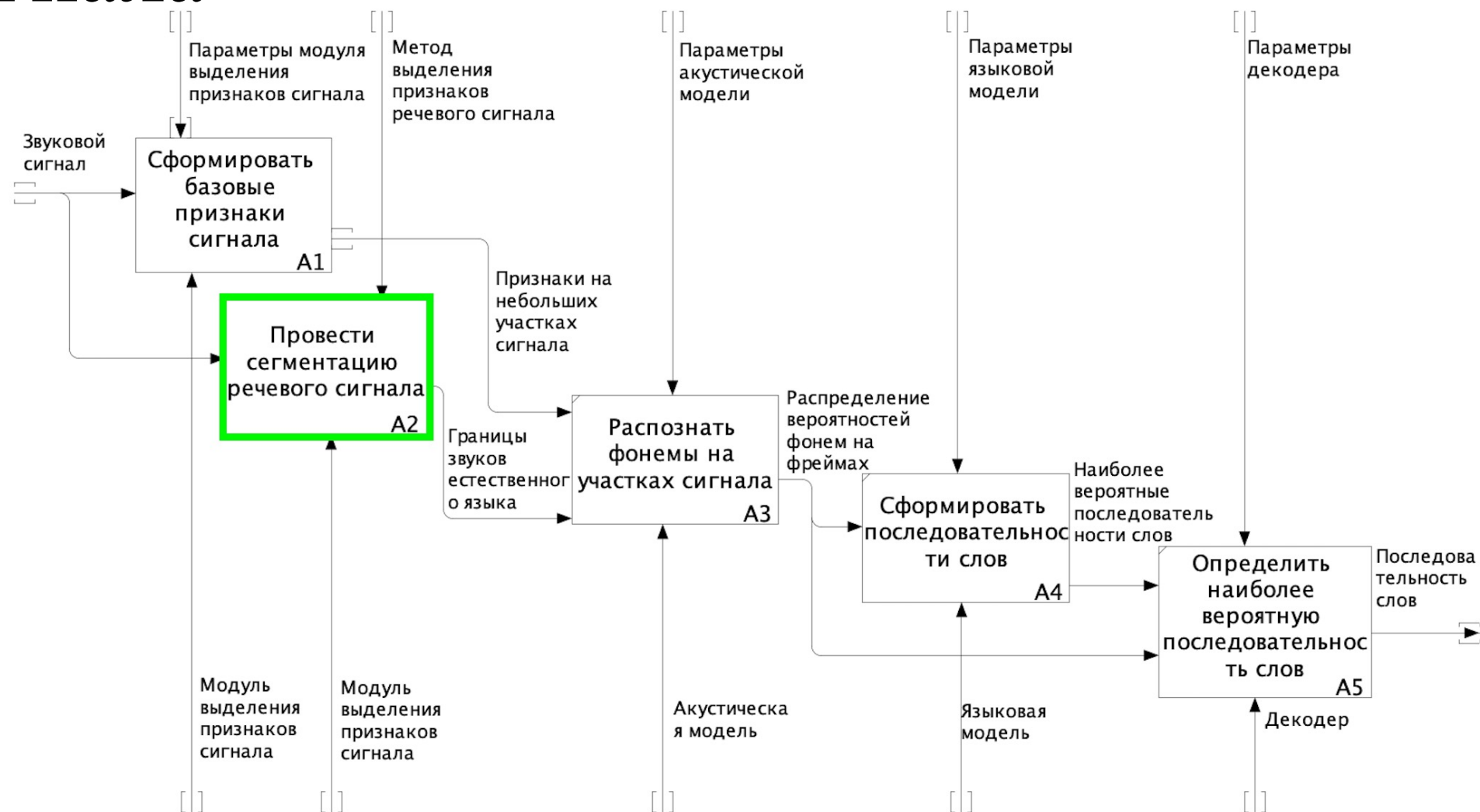
- анализ предметной области, существующих методов выделения признаков речевого сигнала и алгоритмов классификации;
- проектирование метода и разработка алгоритма, реализующего данный метод;
- проектирование системы для проверки работоспособности метода;
- определение характеристик предлагаемого метода.

# Звуки естественного языка

Совместный НИР  
ИУ7, Л4 и МГЛУ



# Основные этапы распознавания речевого сигнала



# Сегментация речевого сигнала



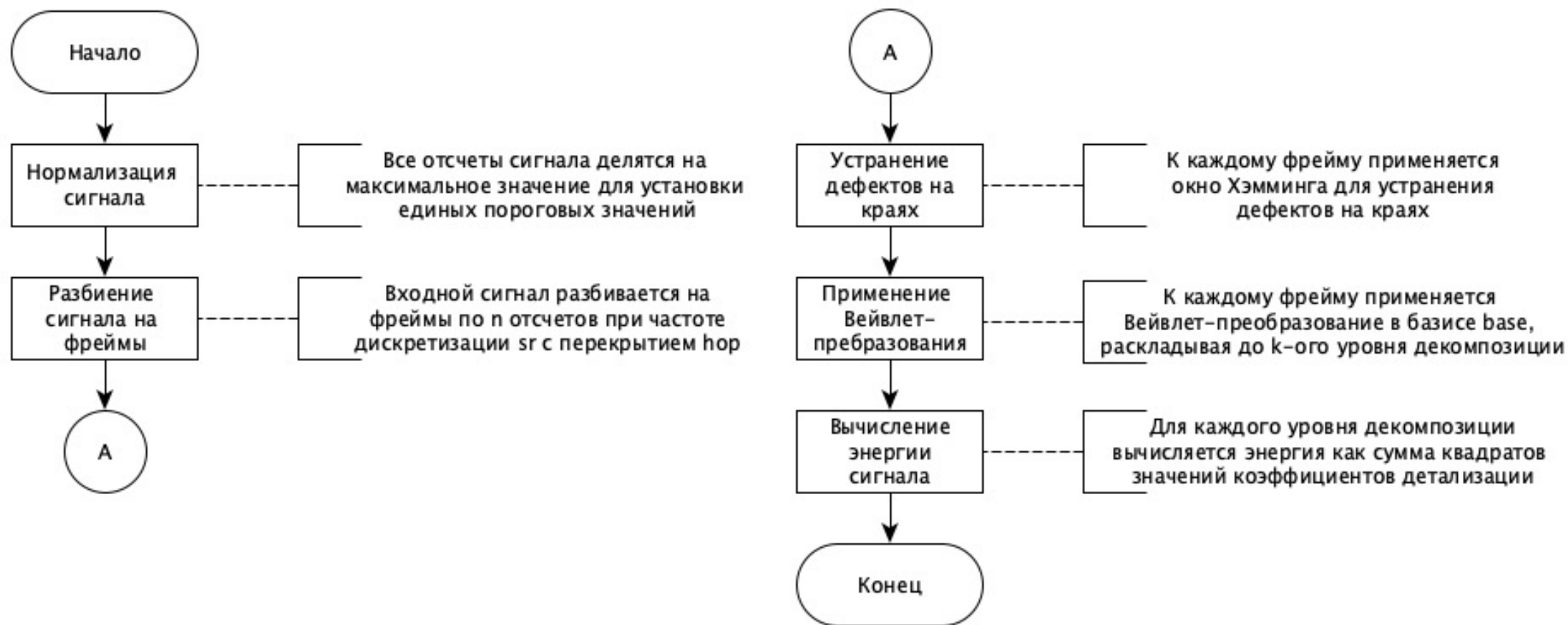
# Методы выделения признаков речевого сигнала

Метод	Базис	Представление	Нелинейность	Нестационарность	Вычислительная сложность
Преобразование Фурье	Априорный	Энергия-частота	Нет	Нет	Низкая
Вейвлет-преобразование	Априорный	Энергия-частота-время	Нет	Да	Средняя
Преобразование Гильберта-Хуанга	Адаптивный	Энергия-частота-время	Да	Да	Высокая

# Алгоритмы классификации при малом объеме обучающей выборки

Алгоритм	Качество обучения	Скорость работы	Поддержка инкрементного обучения
Наивный Байес (NB)	Низкое	Высокая	Есть
К-ближайших соседей (KNN)	Среднее	Низкая	Есть
Деревья решений (DT)	Среднее	Средняя	Нет
<b>Опорные вектора (SVM)</b>	Высокое	Средняя	Нет
Искусственные нейронные сети (NNs)	Среднее	Низкая	Есть

# Алгоритм формирования признаков речевого сигнала





# Математическая постановка задачи классификации

Пусть  $X$  – пространство признаков фреймов сигнала,  
 $Y^n$  – пространство меток классов.

Будем рассматривать задачу на множестве  $Y^2: Y^2 = \{0, 1\}$ , где

- $1 = \{\text{фрейм содержит в себе границу звуков естественного языка}\},$
- $0 = \{\text{фрейм не содержит в себе границу звуков естественного языка}\}.$

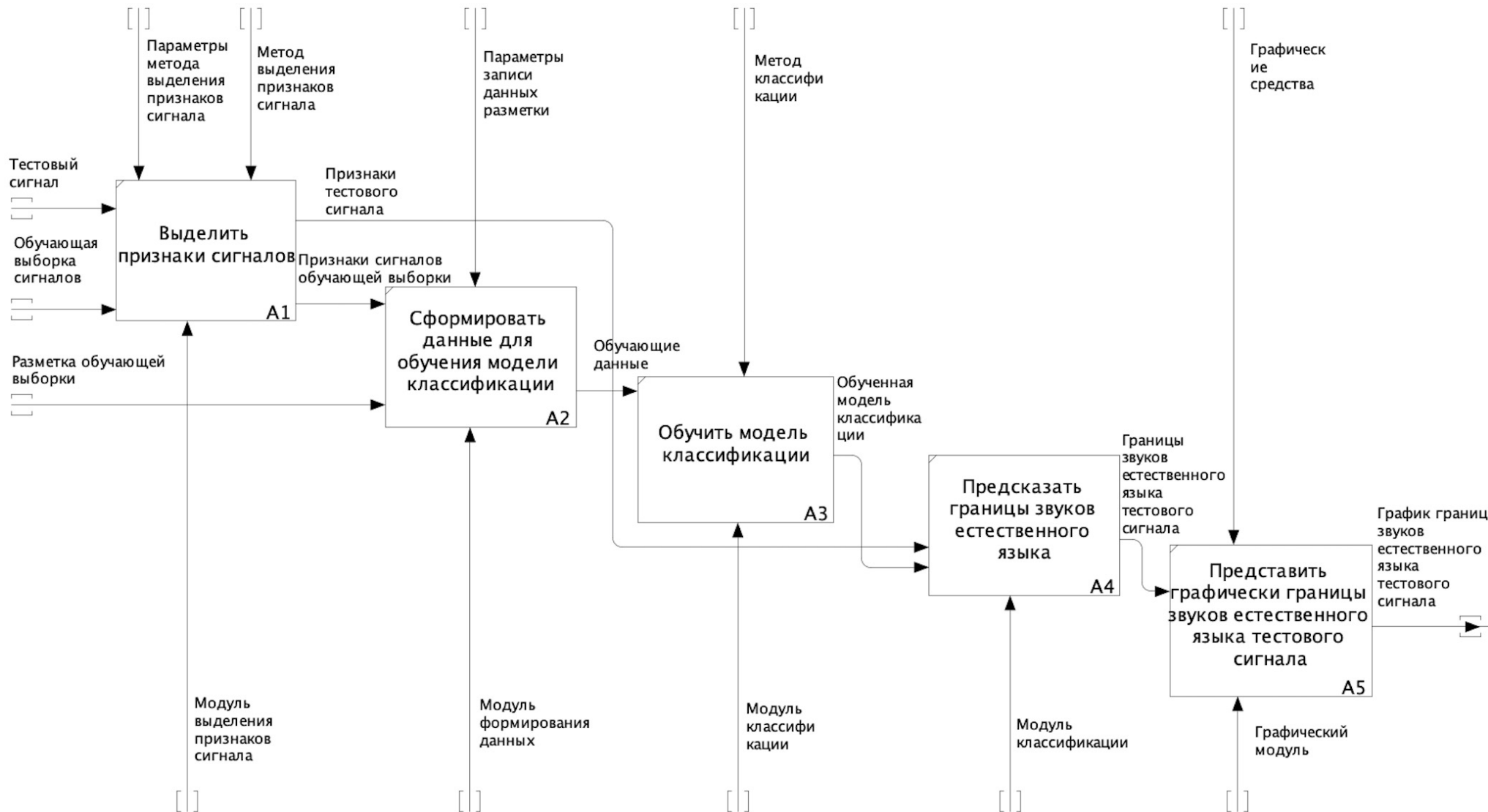
Пусть в распоряжении имеется обучающая выборка

$X_S: X_S = (x_1, y_1), \dots, (x_s, y_s),$  где  $x_i \in X, y_i \in Y^2; i, s \in N.$

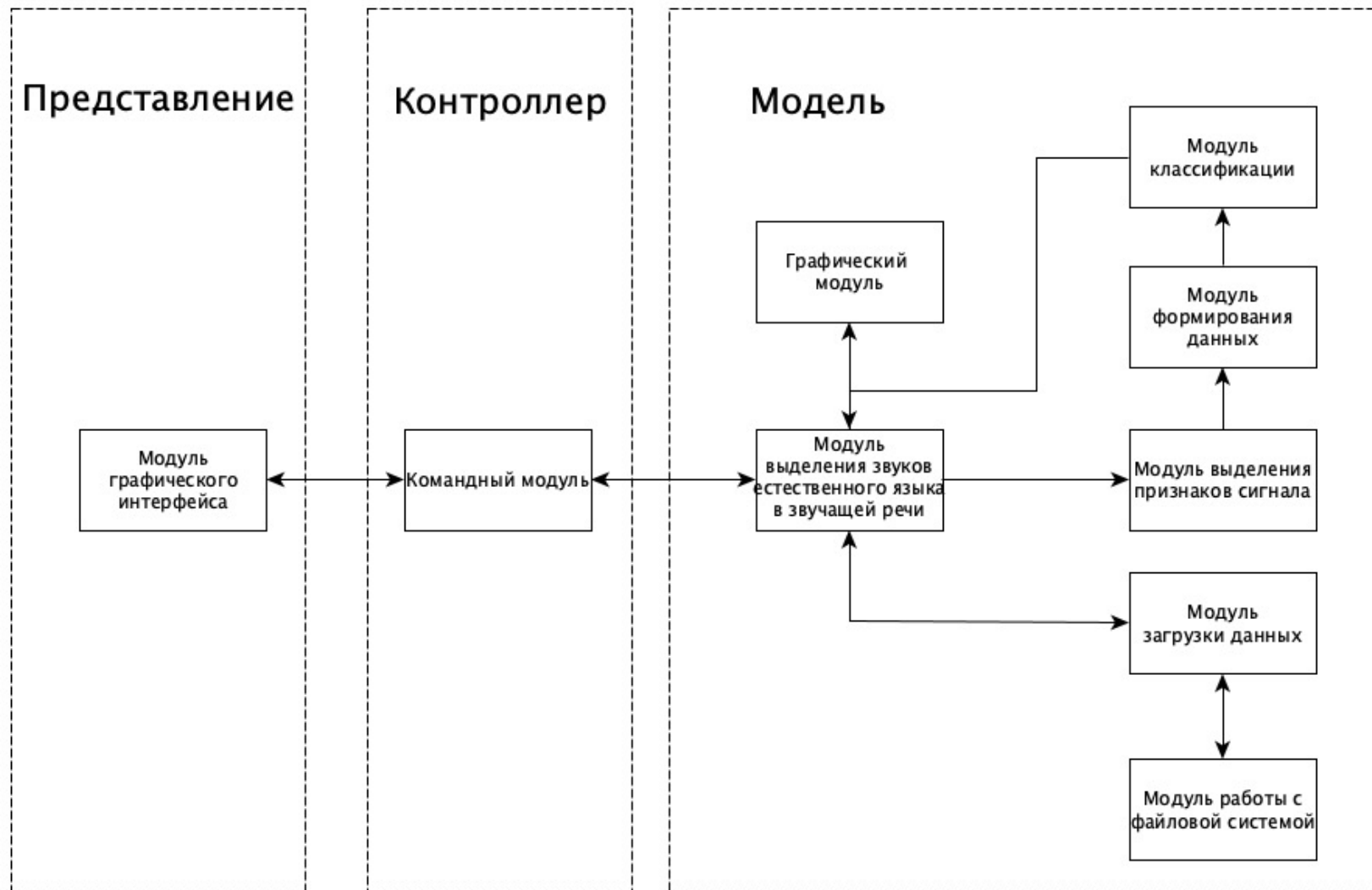
$$L = \sum_{i=1}^N |F(x_i) - y_i| \rightarrow 0$$

Требуется построить отображение  $F: X \rightarrow Y^2$ , способное классифицировать произвольный элемент  $x \in X$  с ошибкой  $L \rightarrow 0.$

# Функциональная модель реализации метода



# Архитектура разрабатываемой системы



# Формирование обучающей и тестовой выборки данных

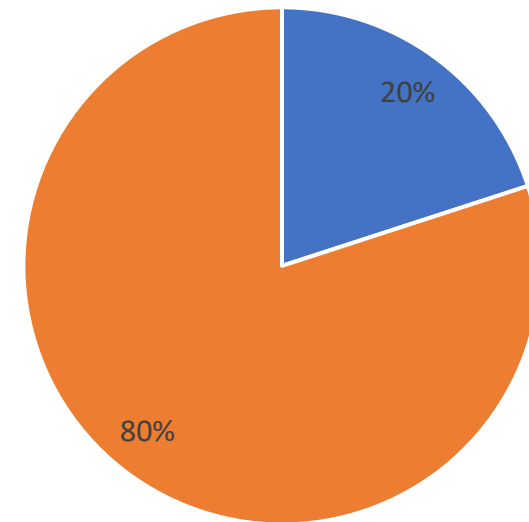
Начальные параметры:

- Количество отсчетов в фрейме  $n = 2500$ ;
- Частота дискретизации  $sr = 16000\text{Гц}$ ;
- Размер перекрытия фрейма  $hop = 40\%$ ;
- Ядро  $base = db4$  (Вейвлет Добеши 4);
- Уровень декомпозиции сигнала  $k = 6$ .

Формат данных:

$\langle \text{признак}_1 \rangle \cdots \langle \text{признак}_7 \rangle \langle \text{метка класса} \rangle$

Количество объектов=5806



■ Тестирование ■ Обучение

# Определение характеристик предлагаемого метода

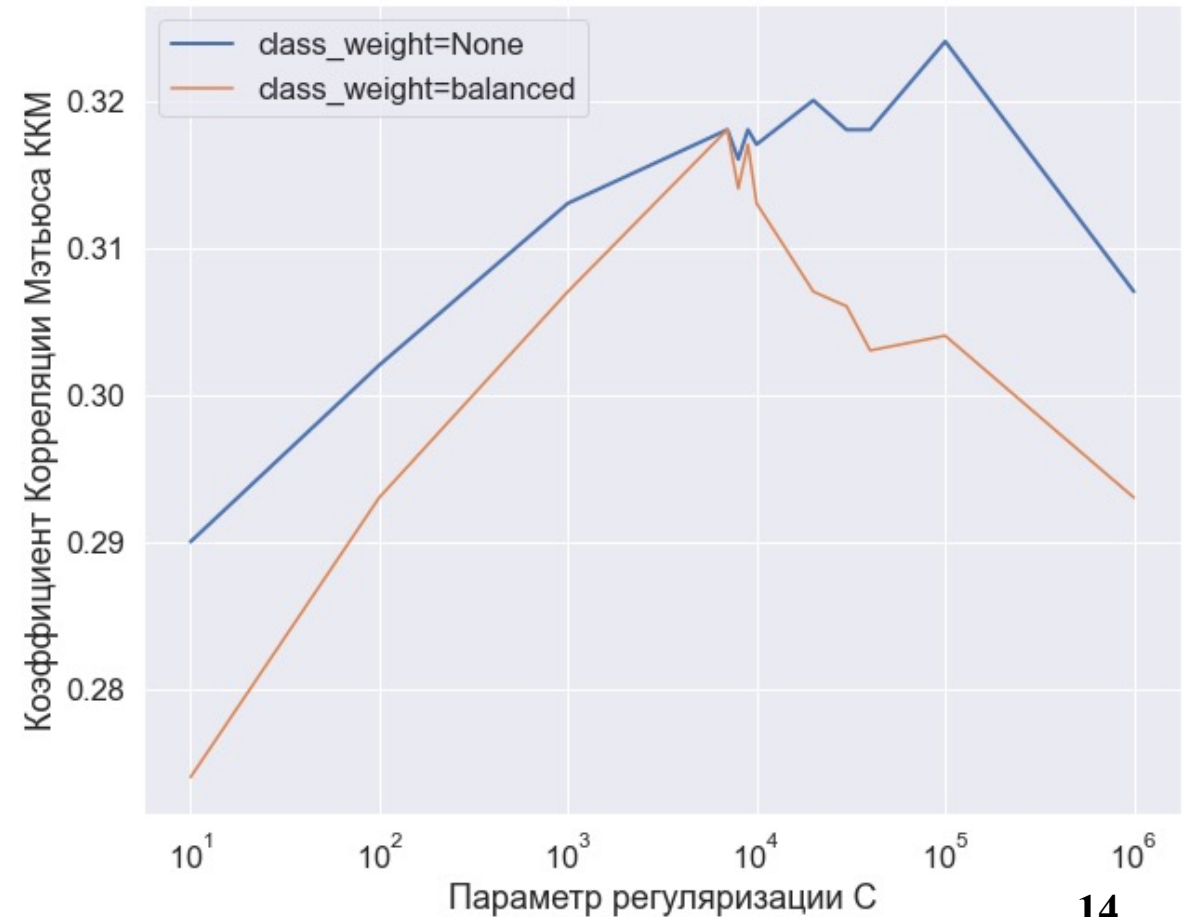
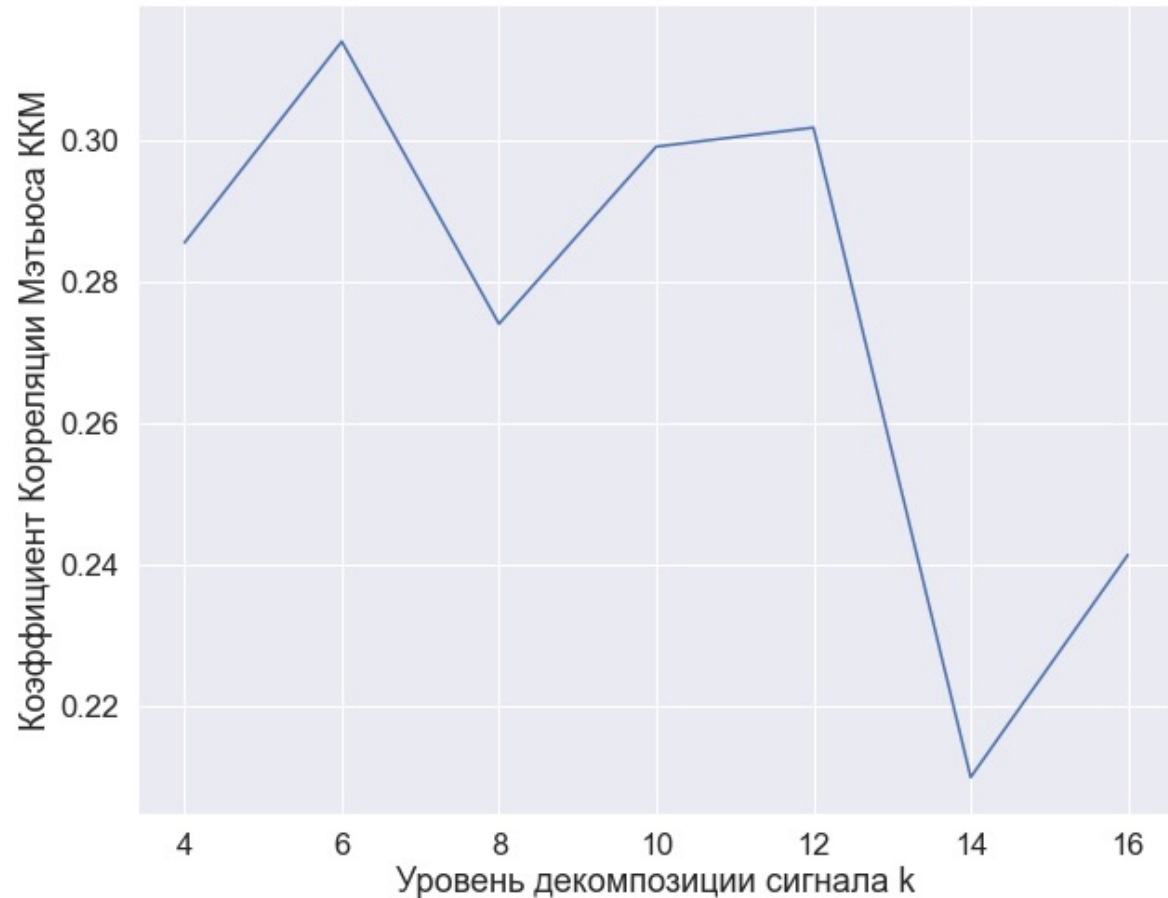
- $C$  – параметр регуляризации классификатора SVM;
- `class_weight` – параметр SVM, регулирующий размер штрафа при неудачном выборе класса;
- $k$  – уровень декомпозиции сигнала.

Подбор параметров осуществлялся при помощи Коэффициента Корреляции Мэтьюса KKM:

$$KKM = \frac{TP * TN - FP * FN}{\sqrt{(TP + FP) * (TP + FN) * (TN + FP) * (TN + FN)}}, \text{ где}$$

- TP - верно классифицированные положительные объекты;
- TN - верно классифицированные отрицательные объекты;
- FN - неверно классифицированные отрицательные объекты;
- FP - неверно классифицированные положительные объекты.

# Определение характеристик предлагаемого метода

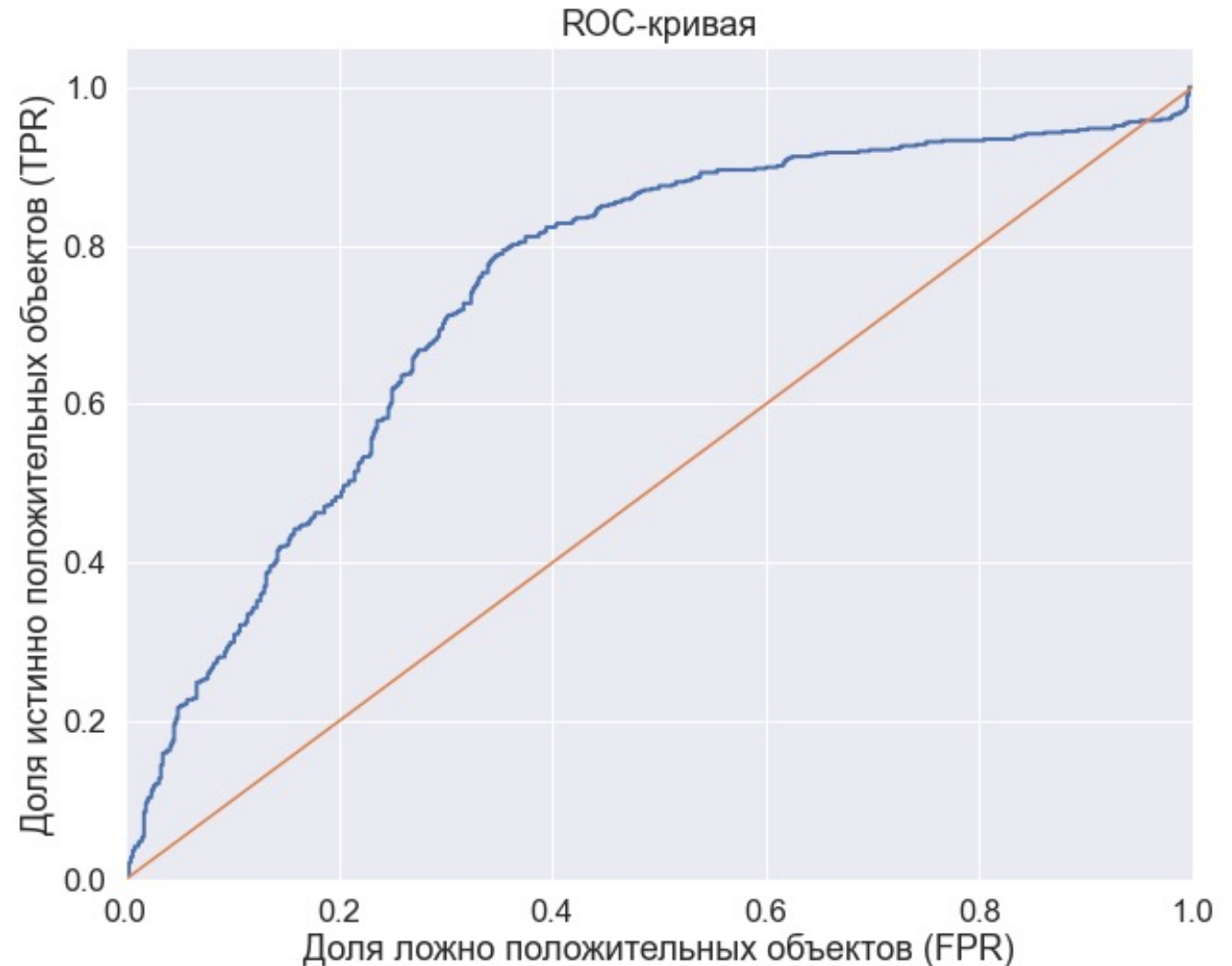


# Оценка качества предлагаемого метода

- Коэффициент Корреляции Мэтьюса  $KKM = 0.33$
- AUC–площадь ROC–кривой = 0.7325

Интервал AUC	Качество модели
0.9-1.0	отличное
0.8-0.9	очень хорошее
0.7-0.8	хорошее
0.6-0.7	среднее
0.5-0.6	удовлетворительное

Экспертная шкала значений AUC



# Заключение

- Выполнен анализ предметной области, существующих методов выделения признаков речевого сигнала и алгоритмов классификации;
- Спроектирован метод и разработан алгоритм, реализующий данный метод;
- Спроектирована система для проверки работоспособности метода;
- Определены характеристики разрабатываемого метода.



# Направления дальнейшего развития

- Подготовка большего количества обучающей выборки, имеющей меньшую погрешность измерений для повышения точности метода;
- Применение альтернативных алгоритмов классификации, позволяющих повысить качество распознавания с применением большего количества данных.