**Описание функциональных характеристик**

**программа для ЭВМ**

**«Искусственный интеллект «Распознавание речи»**

# Термины

|  |  |
| --- | --- |
| Python | высокоуровневый язык программирования общего назначения с динамической строгой типизацией и автоматическим управлением памятью, ориентированный на повышение производительности разработчика, читаемости кода и его качества, а также на обеспечение переносимости написанных на нём программ. |
| mfcc | В обработке звука мел-частотный кепстр представляет собой представление краткосрочного спектра мощности звука, основанное на линейном косинусном преобразовании логарифмического спектра мощности по нелинейной мел-шкале частоты. |
| Stft | разновидность преобразования Фурье. |

# Аннотация

Настоящий документ содержит описание технических решений на основе существующих компетенции ЗАО «Сервис-Газификация» в области автоматизации распознавания речи, реализованных в программе для ЭВМ «Искусственный интеллект «Распознавание речи» (далее по тексту равнозначны полному наименованию: «ИИ «РР», «ИИ», «приложение «Распознавание речи»»).

# Описание программной платформы ИИ «РР»

Программная платформа представляет собой набор системных инструментов распознания речи. Основным программным компонентом используется Pytorch.

# Область применения платформы ИИ «РР»

На данный момент платформа используется для построения искусственного интеллекта на языке Python. За основу берется клиент-серверная архитектура. На основе данной архитектуры могут быть построены как приложения для закрытых сетей, так и публичные, доступные пользователям через интернет, так и гибридные варианты с распределением прав доступа к различным частям и данным приложения со стороны различных групп пользователей.

# Описание прикладных решений ИИ «РР»

**Цель разработки**

Распознавание речи.

Перевод речи в текстовый формат.

**Решаемые задачи**

Перевод речи в текстовый формат.

**Область применения**

Любая голосовая запись

# Описание модулей платформы ИИ «РР»

Основной модуль ИИ «РР» открывает wav файл с записанным голосом и распознает с помощью Мел-кепстрального коэффициента (mfcc) или Оконного преобразование Фурье (stft). На последнем этапе запускается decoder.

# Особенности разработки платформы ИИ «РР»

Разрабатываемое решение содержит различные уникальные алгоритмы обработки голоса.

Используются свои технологические решения для обработки голоса.

Была выработана концепция непрерывной аугментации при обучении нейронных сетей.