**Отчет о разработке модели.**

1. **Обоснование степени значимости модели**
2. **Финансово-экономическое обоснование (ФЭО)**
3. **Описание модели**

Тип задачи: классификация бинарная

Бизнес цели: сократить нагрузку на рекрутеров в плане прослушивания и классификации видео/аудиоинтервью по качеству речи соискателя, тибо полностью автоматизировать данный процесс, исключив из него человеческий фактор

Целевая переменная: качество речи (априори оцененное рекрутером)

Краткое описание: модель представляет собой бинарный классификатор, на выходе которого предоставляются вероятности отнесения аудиоряда, предоставленного соискателем, к классам качественной речи и некачественной речи, для выбора последующего действия относительно данного соискателя. В качестве основных критериев оценки аудиоряда в тренировочном наборе данных с различной степенью строгости и лояльности использовались такие признаки как наличие усугубленных дефектов речи, наличие акцента и степень его выраженности, свазанность и логичность сказанного. Настоящая модель способна реагировать на 2 первых признака.

1. **Подготовка данных**
2. Алгоритм выгрузки данных

Выгрузка данных осуществлялась с помощью парсинга видеоинтервью с платформы skillaz с последующим выделением аудиоряда из видеоинтервью

1. Статистика по выборкам: описание обучающего и валидационного дата сетов (периоды времени, на которых формировались эти дата сеты, количество наблюдений).

Обучающий датасет содержал 20000 самплов аудиодорожек и был получен из видеоинтервью, загруженных за последние полгода относительно начала работ над моделью, валидационный и тестовый(out-of-sample) датасеты содержали по 7000 сэмплов аудиодорожек. К каждому самплу использвоанных датасетов было применено несколько методик препроцессинга.

1. **Моделирование**
2. Математическая обработка входных переменных: основные шаги, которые использовались при feature engineering, особенно если это были какие-то математические преобразования на основе других моделей, однофакторный анализ фич, биннинг, если он проводился и т. п.

В качестве входных данных использовался дискретизированный аудиоряд, частота дискретизации данного сигнала была подобрана в соответсвии со стандартами использованной библиотеки Python по работе с аудиорядами, а также в удовлетворяла требованиям теоремы Котельникова по однозначности и точности восстанновленного из дискретных отсчетов сигнала, т.е. использованная частота дискретизации более, чем в 2 раза превышала верхнюю(значимую) предельную частоту используемый аудиорядов и принималась равной 22кГц. Далее полученный аудиоряд проходил 3 типа предобработки для дальнейшего использования в различных алгоритмах:

1. Дискретизированный сигнал разбивался на не-/перекрывающиеся субвектора(эмбединги), последовательность из которых в последствии передавалась в рекуррентную нейронную сеть (или в сеть с долгосрочно-краткосрочной памятью) для выполнения классификации, данный метод, не включал в себя математическую предобработку непосредственно самих субвекторов и исходил из предположения возникновения некачественной-дефективной речи не на базе частотной основы сигнала, а на базе последовательностей амплитуд сигнала.
2. Для дискретизированного сигнала на протяжении его действия рассчитывались физические параметры:
   1. Спектральный центроид(spectral\_centroids) - указывает, на какой частоте сосредоточена энергия спектра, или, другими словами, указывает, где находится "центр масс" звука
   2. Спектральный Спад Частотной Характеристики(spectral\_rolloff) – это мера формы сигнала. Он представляет частоту, на которой высокие частоты снижаются до 0. Чтобы получить его, мы должны вычислить долю частотных спектров, где 85% мощности сигнала приходится на более низкие частоты.
   3. Спектральная Полоса Пропускания(spectral\_bandwidth) - спектральная полоса пропускания определяется как ширина полосы просвета при прохождении половины пика-максимума
   4. Скорость Пересечения Нуля(zero\_crossings) - очень простой способ измерения гладкости сигнала - вычисление числа пересечений нуля в пределах сегмента этого сигнала. Речевой сигнал колеблется медленно — например, сигнал 100 Гц будет пересекать ноль 100 в секунду — тогда как беззвучный фрикатив может иметь 3000 нулевых пересечений в секунду.
   5. Коэффициенты mel frequency cepstral (MFCCs) сигнала представляют собой небольшой набор признаков (обычно около 10-20), которые кратко описывают общую форму спектральной огибающей. Он моделирует характеристики человеческого голоса.
3. Выбор алгоритма: работа каких алгоритмов, помимо выбранного, проверялась, почему был выбран тот или иной финальный алгоритм.
4. Результат обучения (в том числе подобранные гиперпараметры): параметры сетки значений для отбора гиперпараметров, метрика отбора гиперпараметров, финальные значения.
5. Определить ключевую метрику качества и ограничения (если они есть), указать значения на обучающей и тестовых выборках

*Пример. Ключевая метрика* ***precision*** *при ограничении* ***recall*** *> 0.7*

1. Переменные, вошедшие в модель: