**Кейс: Распознавание эмоций в музыке для создания плейлистов настроения**

**Карта планирования ИИ-проекта**

**Задача системы**

**Кратко охарактеризуйте задачу, которую призвана решить система?**

Автоматическое определение эмоционального окраса музыки для создания плейлистов, соответствующих настроению пользователя.

**Ценность**

**Кто конечный пользователь системы?**

Пользователи аудио сервисов (Яндекс музыка, VK музыка, Spotify, Deezer и тд)

**Какие его проблемы мы решаем?**

* Сложность и время затратность поиска музыки
* Отсутствие автоматизированного инструмента для создания плейлистов под настроение
* Скучность плейлистов по интересам, а не по настроению

**Как он улучшит свой опыт при использовании данной системы?**

* Быстрый и удобный доступ к музыке, соответствующей настроению
* Экономия времени на поиск подходящих треков
* Возможность исследовать новую музыку, соответствующую настроению.

**Предсказание**

**На какой вопрос должна отвечать модель?**

Какая эмоция (радость, грусть, спокойствие, возбуждение и т.д.) преобладает в данной музыкальной композиции?

**Определите тип задачи обучения (классификация, регрессия, …)?**

Многоклассовая классификация.

**Входные данные (формат, атрибуты, шкалы):**

Аудиофайлы (MP3, WAV и т.д.); атрибуты - спектрограмма, мелодические характеристики, темп, тональность, ритм.

**Выходные данные (формат, возможные значения):**

Класс эмоции (например, "радость", "грусть", "спокойствие", "возбуждение", "злость", "нейтральное").

**Риски**

**Определите цену ошибки модели. Рассмотрите как ложноположительные, так и ложноотрицательные результаты.**

Ложноположительные: песня, не соответствующая настроению, может вызвать раздражение. Ложноотрицательные: песня, соответствующая настроению, не попала в плейлист, пользователь не получит ожидаемого эффекта.

**Определите уровень риска проекта в соответствии с AI Act. Опишите необходимые меры предосторожности:**

Низкий. Меры: прозрачность работы модели, возможность обратной связи, отсутствие критичных последствий от ошибок.

**Оцените этические риски при реализации проекта:**

Возможно искажение эмоционального восприятия музыки; потенциальное создание "пузыря" из музыки с определенной эмоцией.

**Оцените другие юридические риски, связанные с проектом:**

Нарушение авторских прав при использовании музыкальных произведений.

**Данные**

**Какие данные необходимо собрать для адекватного обучения модели?**

Большое количество музыкальных треков, размеченных по эмоциональному окрасу.

**Какие источники данных можно использовать?**

**Внутренние:**

Музыкальная библиотека сервиса (если есть).

**Внешние открытые:**

Датасеты по распознаванию эмоций в музыке, открытые API музыкальных сервисов (если есть).

**Внешние закрытые:**

Коммерческие музыкальные библиотеки.

**Генерация данных:**

Аугментация (изменение скорости, высоты тона, добавление шума), генерация синтетической музыки.

**Методы аугментации данных:**

Изменение скорости, высоты тона, добавление шума, нарезка на фрагменты.

**Необходимый объем данных:**

N \* 10^4 (N > 1) размеченных музыкальных треков (зависит от объема музыки предлагаемой конечным сервисом).

**Как часто обновляются данные?**

По мере появления новой музыки и по мере необходимости улучшения качества модели. В целом не часто, но необходимо дообучение на маленькой выборке новых треков, например, раз в месяц

**Какие требования к качеству данных являются критичными?**

Корректность разметки выборок, разнообразие музыкальных жанров.

**Насколько затратно хранение и обработка обучающей выборки?**

Затратно по дисковому пространству, тк в зависимости от частоты дискретизации и битности аудиофайлы могу весить достаточно много.

**Как сделать данные публично доступными другим системам?**

Создание API для доступа к предсказаниям модели (не к обучающей выборке).

**Обработка данных**

**Как извлекать необходимую информацию (признаки) из сырых данных?**

Использование библиотек для анализа аудио (librosa, pyAudioAnalysis); извлечение спектрограммы, мелодических характеристик, темпа, тональности.

**Как привлечь экспертность в предметной области для повышения качества признаков?**

Привлечение музыкальных критиков или психологов для уточнения разметки и выявления неочевидных признаков.

**Модель и Предсказания**

**Опишите процесс превращения предсказания в действие или решение пользователя?**

Модель предсказывает эмоцию всех треков сервиса → система формирует плейлист на основе предсказанных эмоций и выборе пользователя → пользователь получает плейлист.

**Есть ли скрытые издержки в процессе принятия решения?**

* Зависимость от точности модели
* Возможность смещения в сторону популярных эмоций.

**Как пользователь взаимодействует с результатом работы модели?**

* Пользователь указывает желаемые эмоции -> Приложение отображает плейлисты по эмоциям.

**Необходимо потоковое или пакетное моделирование?**

Потоковое при обучении (тк тяжелые данные) и пакетное при дообучении (можно тоже потоковое). Можно везде пакетное, если небольшими батчами

**Насколько вычислительно сложно прямое действие модели?**

Зависит от вида и архитектуры модели.

**Инжиниринг модели**

**Как часто потребуется повторное обучение модели?**

По мере появления новых треков или новых источников данных для обучения (например, раз в месяц).

**С какими издержками связано повторное обучение модели?**

* Появление новых источников данных
* Появление новых треков
* Изменение архитектуры модели

**Насколько этот процесс длителен?**

До суток, в зависимости от мощности инфраструктуры.

**Насколько этот процесс может быть автоматизирован?**

Можно сделать стойку для обучения, и автоматизировать до той степени, что мы просто складываем новый батч данных, модель самостоятельно дообучается/переобучается и после этого улетает новая версия в registry

**Планируются ли изменения технологического стека?**

Зависит от бизнес-требований

**Сколько моделей в проде нам понадобится?**

В зависимости от реализации. Если хотим отдельные модели под каждое настроение (регрессия/бинарная классификация) то ровно столько, сколько эмоций поддерживаем Если хотим многоклассовую, то зависит от требований к масштабируемости

**Какова стратегия обновления моделей?**

Тестирование на тех же тестовых данных -> Выкатка в тестовый стенд -> Тестирование альфа группой -> Переключение на неё основного сервиса

**Оценка и мониторинг моделей**

**Какие метрики эффективности отражают качество работы модели?**

Точность, полнота, F1-мера, AUC-ROC (тк классификация)

**Какие другие показатели критичны для работы модели?**

Загрузка ресурсов, удовлетворенность пользователей.

**Какой триггер будет использован для инвалидизации модели?**

Снижение метрик эффективности ниже порогового значения, жалобы пользователей.

**Как мы оцениваем удовлетворенность пользователей системой?**

Сбор обратной связи, сбор метрик использования (частота использования, время прослушивания).

**Предусмотрена ли обратная связь в процессе работы системы?**

Возможность оценить предложенный плейлист, сообщить об ошибке классификации, добавить эмоцию.

**Интеграция и развертывание модели**

**Как модель будет взаимодействовать с прикладной системой?**

REST API или gRPC.

**Как соотносятся их технологические стеки?**

Модель - Python/PyTorch/TensorFlow, прикладная система - зависит от имеющейся архитектуры.

**Какова схема развертывания прикладной системы?**

Микросервисы, контейнеризация (Docker), оркестрация (Kubernetes).

**Каково продуктовое окружение прикладной системы?**

Мобильное приложение, веб-приложение, API для сторонних сервисов.

**Какая схема интеграции наиболее целесообразна?**

REST API для интеграции с фронтендом, брокер сообщений для асинхронных задач. Registry моделей - MLFlow и inference (mlserver)

**Возможно ли для развертывания модели применять облачные системы?**

Зависит от требований и имеющейся инфраструктуры/стоимости обновления инфраструктуры. В теории, возможно, тк чувствительность данных низкая и не содержит конфиденциальных данных пользователей.

**Какие шаги развертывания можно полностью автоматизировать?**

Сборка контейнеров, деплой в кластер, тестирование, дообучение, снятие метрик в реальном времени.