

2 Инструмент анализа и агрегации различных по структуре данных

Ваш заказчик - развивающийся Банк. Руководство Банка хочет усовершенствовать систему скоринга кредитных заявок физ. лиц. Сейчас принятие решения о выдаче кредита принимается только на основании данных из анкет.

Банк хочет обогатить анкеты данными из других источников.

Вам будет предоставлена документация текущей скоринговой и других банковских систем, а также виртуальная машина (VM) с используемыми СУБД и тестовым набором данных. Для использования VM понадобится установить Oracle VM VirtualBox.

2 Инструмент анализа и агрегации различных по структуре данных

У банка есть скоринговая система, которая состоит из компонент:

База данных “Заявки” в которой хранятся только анкеты к заявкам заемщиков (ФЛ).

Модуль принятия решения МПР, который на основе простейшего логического анализа анкет принимает решение о выдаче кредитов. В ряде случаев, решение принимает специалист. МПР лучше рассматривать как черный ящик.

Также, есть ряд систем, которые не задействованных в процессе скоринга:

БД Кредитов (реляционная в 3NF) в которой ведутся сами кредитные продукты ФЛ

БД Зарплатных проектов (графовая аналитическая) предназначенная для анализа портфеля зарплатных карт и ЗП проектов.

БД Страховок (реляционная в 6NF) содержит информацию об ипотечных, авто и других видах договоров страхования, оформленных в банке.

БД Интервью (документарная БД) содержит транскрибированные записи телефонных разговоров сотрудников Банка и потенциальных заемщиков.

Excel-файлы со справочной информацией.

2 Инструмент анализа и агрегации различных по структуре данных

Процесс AS IS: решение о выдаче кредита принимается только на основе анкеты.

Заказчик хочет оживить процесс скоринга. Обогатить скудные анкеты интересными данными из остальных систем.

Заказчик готов доработать модуль принятия решения МПР в соответствии с рекомендациями Команды участников, но для проведения доработок требуется подготовить модель данных, архитектуру решения и ETL-процесс который будет подаваться на вход МПР.

Заказчик просит уделить внимание потенциальным возможностям использования информации из разных прикладных областей, а также рациональному использованию вычислительных ресурсов.

2 Инструмент анализа и агрегации различных по структуре данных

Задачи Команд:

1. Продумать ТО ВЕ, процесс скоринга.
2. Продумать архитектуру решения.
3. Составить новую модель данных, позволяющую решить определенную задачу.
4. Провести интеграцию между источниками и придумать, как написать ETL.

При этом сам модуль не нужно разрабатывать или дорабатывать. Достаточно написать рекомендации к использованию новой модели данных модулем.

2 Инструмент анализа и агрегации различных по структуре данных

Критерии оценки:

1. Процесс TO BE и функциональность Модель данных – вес 30%
2. Архитектура решения и ETL-процесс – вес 30%
3. Рoad маpa развития продукта 20%
4. Презентация решения – вес 20%

До дедлайна отправки заданий **26.10.2021 12:00 МСК** необходимо отправить на почту **hackathon@glowbyte.com**

- Презентацию разработанного продукта в формате pdf - **обязательно**
- Видео с презентацией продукта продолжительностью не больше 5 минут – **по желанию**

Желаем удачи!