Уважаемые кандидаты,

Поздравляем вас с успешным прохождением отборочного тестирования и приглашаем на обучение в Корпоративном университете Сбербанка на программу **«Apache Spark для задач анализа данных»**, которое будет проходить с 19.09 по 19.10.2023 на платформе СберМитап.

Предлагаем ознакомиться с программой обучения:

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Программа включает 10 модулей по 3 астрономических часа, которые будут проходить по вторникам и четвергам вечером в течение шести недель **в 19:00 по МСК**.   |  |  |  |  |  |  | | --- | --- | --- | --- | --- | --- | | **№ занятия** | **1** | **2** | **3** | **4** | **5** | | **Дата** | **19.09** | **21.09** | **26.09** | **28.09** | **03.10** | | **Время** | **19:00-22:00** | **19:00-22:00** | **19:00-22:00** | **19:00-22:00** | **19:00-22:00** | | **№ занятия** | **6** | **7** | **8** | **9** | **10** | | **Дата** | **05.10** | **10.10** | **12.10** | **17.10** | **19.10** | | **Время** | **19:00-22:00** | **19:00-22:00** | **19:00-22:00** | **19:00-22:00** | **19:00-22:00** |   **Программа занятий:**  *19.09.2023 Hadoop для Spark-юзера.*  *21.09.2023 Введение в Apache Spark.*  *Лаба 1. Дескриптивный анализ рейтингов фильмов на RDD*  *26.09.2023 Работа с DataFrames в Apache Spark.*  *Лаба 2. Content-based рекомендательная система на Dataframes*  *28.09.2023 Spark ML: основы и создание пайплайнов.*  *Лаба 3. Соревнование: рекомендация фильмов на основе данных об истории телесмотрения*  *Недельный перерыв для устранения задолженностей по лабораторным работам.*  *Лаба 4. Прогнозирование пола и возрастной категории пользователей на основе логов посещений в real-time*  *03.10.2023 Практический ML на Spark. Решение кейсов с использованием Spark ML.*  *Лаба 5. Прогнозирование оттока клиентов*  *05.10.2023 Spark Streaming.*  *10.10.2023 Spark Scala API.*  *12.10.2023 PySpark vs Python: Работа с графами.*  *17.10.2023 Разбор и обсуждение решений лабы 1, 2 и 3. Q&A.*  *19.10.2023 Разбор и обсуждение решений лабы 5 и 6. Q&A..* |  |

До встречи на обучении!

Коллеги, добрый день!

Вы номинированы на курс **"Apache Spark для анализа данных".** Программа содержит практические задания-лабы, максимально приближенные к реальным задачам дата инжиниринга, поэтому требует определенных входных знаний и времени на обучение в размере минимум 10 часов в неделю (на занятия и выполнение лаб). **Длительность программы – 7 недель. Пожалуйста, оцените все риски и свою возможность участия в обучении.**

Практическая часть реализуется на языке программирования Python.

Для попадания на курс необходимо пройти онлайн-тест. Тест откроется 31.08.2023 в 9:00 по МСК и будет доступен в течение 24 часов (01.09.2023 в 9:00 тест будет закрыт). Заполненные тесты позже этого времени проверяться не будут.

Входной тест будет доступен по ссылке: <https://se.sberbank-chool.ru/bQMYg6BJ>.  Время на прохождение – 17 минут. Всего в тесте 17 вопросов с вариантами ответов, которые смогут определить минимальный уровень знаний для успешного прохождения программы и выполнения лаб. Тест содержит вопросы по темам: Linux и работа в терминале, Python, Spark, HDFS, Hadoop, переобучение моделей, ML.

Большая просьба **не пытаться несколько раз пройти тест** за это время (это противоречит корпоративной этике). По итогам теста мы сформируем группу из 80 человек.

Старт обучения запланирован на **19.09.2023**, занятия будут проходить по вт и чт с 19:00 до 22:00 по Московскому времени в онлайн-формате в Zoom.

Тем, кто успешно пройдет входное тестирование, мы пришлем подробное письмо с информацией о программе, расписанием занятий и всей необходимой информацией – **не позже 10.09.2023**.

**Если обучение вам не актуально, отклоните встречу с тестированием.**

**Если кто-то из ваших коллег проявит интерес к программе, вы можете переслать встречу с тестированием.**