**Домашнее задание №1. «Требования в DS проектах aml»**

Возьмите за основу продукт компании “Нетология”: профессию Data Scientist.

1. Как бы Вы организовали аналитику по данному продукту?
2. Сформулируйте 2-3 ключевые метрики продукта
3. Какие данные Вам могут помочь в поиске инсайтов?
4. Опишите команду для работы с данными для данного продукта
5. Какие роли Вам будут нужны?
6. Чем они будут заниматься?

***Решение***

1. Продукт состоит из 7 блоков:
   1. Машинное обучение
   2. Менеджмент data-проектов
   3. Введение в нейронные сети
   4. Компьютерное зрение
   5. Рекомендательные системы
   6. Временные ряды
   7. Обработка естественного языка

По каждому блоку аналитические показатели делятся на **обще-курсовые** (сквозные, связанные с процессом онлайн обучения, в принципе) и **специальные** (обусловленные спецификой тематики блока, наверное, это то, что здесь называют «*инсайтом*»).

Количество показателей первого типа (сквозные) должно быть константой, одинаковой для всех блоков. Количество показателей 2-го типа (инсайты) привязывается именно к тематике блока, а именно, должно включать все ключевые навыки (компетенции), которые учащийся должен освоить во время прохождения данного блока. Именно эти компетенции, в целом, определяют стратегию и тактику дальнейшего развития курса.

1. Метрики эффективности обучения можно, например, взять из модели Киркпатрика (<https://www.ispring.ru/elearning-insights/kak-izmerit-effektivnost-obucheniya>):
   1. Реакция (понравилось ли обучение учащимся)
   2. Обучение (Какие знания и навыки были усвоены? Насколько хорошо (от 0 до 100 баллов)?)
   3. Поведение (Применяются ли полученные навыки на практике)
   4. Результаты (какое влияние обучение оказало на ученика: его дальнейший профессиональный рост, уровень жизни)
2. Ищем по формуле:

Данные -> Информация -> Инсайты

Для онлайн-курса важны, в частности, такие данные, как:

* 1. Результат онлайн-теста
  2. Результат опроса
  3. Письменные домашние работы (2 уровня сложности: Light и Pro, 3 уровня освоения: базовый (минимальный), средний, профессиональный)
  4. Письменные экзамены (уровень сложности – общий, 3 уровня освоения: базовый (минимальный), средний, профессиональный)
  5. Прогресс пользователей
  6. Количество попыток
  7. Качество ответов
  8. Количество и время просмотра обучающимся контента (посещаемость сайта учеником)
  9. Количество выполненных заданий
  10. Активность участия в обсуждениях (на вебинарах и в чате Slack)

1. Команда для работы с данными в курсе «Профессия DS» должна состоять из:
   1. лекторов,
   2. наставников,
   3. технической поддержки,
   4. менеджера поддержки курса (куратора).
2. Следующий список – роли в команде. Понятно, что эти роли могут назначаться как отдельному члену команды, так и совмещаться каким-либо специалистом, обладающим определенным набором проф. навыков (компетенций). Итак, необходимы специалисты в области:
   1. DS
   2. Аналитик
   3. BI-специалист
   4. Разработчик (разработка архитектуры, концепции)
   5. DE
   6. Специалист в области БД (реляционных и нереляционных)
   7. Владелец продукта
   8. TeamLead
   9. Приглашенные лекторы
   10. Наставники
3. Коротко опишем обязанности каждой роли.
   1. DS: перевод задач обучающей платформы на язык задач анализа (разработка моделей), разработка алгоритмов, оценка качества алгоритмов и моделей, создание A/B-тестов
   2. Аналитик: поиск закономерностей, инсайтов, выбор вида отчетов для показа экспертам
   3. BI: визуализация, поиск новых онлайн-инструментов, ускоряющих процессы обработки данных и представления их в удобном для анализа виде
   4. Разработчик – больше программист-прикладник, обладающий навыками разработки архитектур для выбранных моделей, подбора технологий для решения, решение проблем масштабирования (создание приложений для разных платформ).
   5. DE: В моем представлении, это сисадмин + специалист по сетевым технологиям (+облака, распределенные системы) + нереляционные БД, масштабирование продукта, интегрирование его на другие платформы (например, поддержка мобильного приложения)
   6. Специалист в области БД (создание и поддержка хранилищ данных различных архитектур, разработка запросов SQL и NoSql)
   7. Владелец: принятие решений, вопросы разработки концепций, стратегии развития, финансовые вопросы, заключение договоров
   8. TeamLead: в данном курсе – это координатор, менеджер, обладает широким спектром компетенций в области онлайн-обучения и, отчасти, в области тематики курса.
   9. Приглашенные лекторы: умение вести онлайн-конференции, запись видео, поддерживать контакт с аудиторией, экспертно отвечать на вопросы аудитории, умение работать с софтом, создающим и редактирующим видео, месенджерами типа Discord, Slack, Telegram
   10. Наставник: ***индивидуальная*** работа с обучающимся, ответы на вопросы, проверка ДЗ, контрольных и экзаменационных работ.