Набор данных представляет из себя зарегистрированные при помощи датчиков реакции на вопросы. Они размечены специалистами по трем состояниям 0, 1 и 2. Каждая проверка представляет из себя несколько групп вопросов, каждый из которых задается не более 4 раз. Для каждого человека была следующая структура тестирования:

Группа1 Повторение1: Вопрос1, Вопрос2, ..., Вопрос12

Группа1 Повторение2: ВопросI1, ВопросI2, ..., ВопросI12 (где I1, I2, ..., I12 - вопросы из Повторения1 в случайном порядке)

Группа1 Повторение3: ВопросI1, ВопросI2, ..., ВопросI12 (где I1, I2, ..., I12 - вопросы из Повторения1 в случайном порядке)

Группа2 Повторение1: Вопрос1, Вопрос2, ..., Вопрос12

Группа2 Повторение2: ВопросI1, ВопросI2, ..., ВопросI12 (где I1, I2, ..., I12 - вопросы из Повторения2 в случайном порядке)

Группа2 Повторение3: ВопросI1, ВопросI2, ..., ВопросI12 (где I1, I2, ..., I12 - вопросы из Повторения2 в случайном порядке)

и так далее ...

Количество групп от 1 до 7, в группах может быть от 1 до 14 вопросов.

Соответственно, в наборе данных присутствуют:

Obfuscated name - уникальный id испытуемого.

Test\_index - номер группы, нумерация с 1

Presentation - номер повторения, нумерация с 1

Question - номер вопроса, нумерация с 1

Data - равномерно зарегистрированные полиграфом 12 секунд сигнала с частотой 20 Гц (240 значений) (метка датчика регистрации сигнала - фотоплетизмограмма)

Class\_label\_FPG - метка класса 0,1 или 2

Важное примечание: согласно методике полиграфолог при выставлении баллов оценивает реакции на вопросы при помощи попарного сравнения в рамках каждого повторения отдельно - не более 14 вопросов. Такая процедура выполняется независимо по каждой группе вопросов по каждому повторению.

Метрика будет подсчитываться подобным образом. Будут начисляться баллы за ошибки. Начисление баллов описывается таблицей.

|  |  |  |
| --- | --- | --- |
| Ваши предсказания | Настоящие результаты | Количество баллов, начисленных за ошибку |
| 2 | 1 | +1 |
| 2 | 0 | +2 |
| 1 | 2 | +1 |
| 1 | 0 | +1 |
| 0 | 1 | +1 |
| 0 | 2 | +2 |

Далее ваши баллы за ошибки делятся на максимально возможную ошибку и вычитаются из 1.