Описание тестового задания «Recist»

Задача:

1. Написать скрипт для обработки выгрузки «Приложение 1»

2. Сформировать отчет по форме «Приложения 2»

3. Визуализировать данные на Dashboard Yandex DataLens

I. Условия обработки (скрипт):

1. Отобрать все записи по процедуре "Компьютерная томография органов грудной клетки, брюшной полости и малого таза с контрастированием"

2. В столбце «CONCLUSION» выбрать ключевые поля

1) RECIST

2) Тип ответа

\*могут быть как оба ключевых слова, либо только одно ключевое слово

3. Создать ДОП столбец «keyword» и записать выбранное ключевое слово (приоритетным является RECIST)

4. Создать ДОП поле «estimation\_Recist». В это поле записать «Тип ответа» из поля «CONCLUSION» проработав заключение по ключевым словам (перечислены ниже).

 В строках только с заполненным полем «keyword» провести запись по найденным типам ответа:

1) стабилизация (Stable Disease)

2) прогрессирование (Progressive Disease)

3) частичный ответ (Partial Response )

4) полный ответ (Complete Response).

5. Если в «CONCLUSION» имеются такие записи как:

- Оценка по RECIST невозможна

- baseline не указан

- по критериям RECIST не выявлены

- не попадают под критерии RECIST

- данные nadir и baseline не указаны

и прочие, которые указывают на невозможность оценить Тип ответа по Recist, оставляем запись в поле «estimation\_Recist» - Пустой

6. Если в «CONCLUSION» имеется ключевое слово «Recist», но не указан тип ответа, то в поле «estimation Recist» указываем «Ошибка описания»

Примеры ошибок:

1) «Выбраны таргетные очаги, RECIST-сумма 125 мм»

2) Выбраны таргетные очаги, RECIST-сумма таргетных очагов 62 мм. При сравнении с КТ от 25.08.2022. В остальном, без динамики

3) умеренная положительная динамика (RECIST -16%).

4) уменьшение размеров очагов в печени, наиболее крупного на 39%. (регрессия по RECIST 1.1)

и прочие аналогичные записи

7. Если в CONCLUSION есть «Тип ответа», но нет «Recist» - это не является ошибкой. Выбираем тип ответа.

8. Все ответы в поле «estimation\_Recist» должны быть записаны строго в формате:

1) стабилизация (Stable Disease)

2) прогрессирование (Progressive Disease)

3) частичный ответ (Partial Response )

4) полный ответ (Complete Response)

5) ошибка описания

Если тип ответа не содержит полный ответ из справочника, то необходимо идентифицировать по совпадающим ключевым словам и указать в требуемом формате

9. Создать ДОП поле «Date\_conclusion» в котором вывести дату заключения в формате дд.мм.гггг

10. Создать ДОП поле «Time\_till\_conclusion» и рассчитать значение как:

LAST\_VALIDATION\_1 – (минус) STUDY\_DATETIME.

Значение вывести в формате: чч.мм.сс

Все ДОП поля создаем после основных в выгрузке.

II. Выгрузка и форма отчета в приложении 1 и 2

**III. Условия визуализации данных на Dashboard в Yandex DataLens**

1. Фильтр:

1) Период (дата)

2) Наименование МО

3) Наличие записи в поле С

2. Вывести общую информацию:

1) Всего исследований

2) Количество исследований по Recist (наличие записей в поле «keyword») и Доля исследований по Recist

3) Количество и Доля (от recist) стабилизация (Stable Disease)

4) Количество и Доля прогрессирование (Progressive Disease)

5) Количество и Доля частичный ответ (Partial Response )

6) Количество и Доля полный ответ (Complete Response)

7) Количество и Доля ошибка описания

3. График динамики количества всех исследований, количества исследований по Recist, доли Recist

4. График динамики Recist по типам ответа

5. Таблица с возможностью сортировки Количество описаний по Recist с градацией по типам ответа

6. График среднего времени до заключения (по расчетному полю «Time\_till\_conclusion») по каждой организации. Дополнительные фильтра к графику

1) «Оценка по Recist» и

2) «Наименование Медицинской Организации»

Итоговые результаты выслать обратным письмом.

Мы от Вас ждем:

1. Скрипт для обработки выгрузки на Python

2. Отчет по форме «Приложения 2» (без пароля)

**3. Dashboard на Yandex DataLens с рабочей ссылкой.**