## Задача 1. Классификация заявок в диспетчерскую

В диспетчерскую службу уездного городка Д поступает множество заявок: то трубу где прорвёт, то с крыши капает, то крышку люка кто-то позаимствовал. Dịch vụ điều độ thị trấn D nhận được rất nhiều yêu cầu: đường ống bị vỡ đâu đó, mái nhà rỉ nước, hoặc có người mượn nắp cống.

Администрация городка просит делить все полученные заявки по специальному классификатору, который состоит из пар «Группа работ»-«Категория работ в группе». Chính quyền thị trấn yêu cầu phân chia tất cả các đơn đăng ký nhận được theo một bộ phân loại đặc biệt, bao gồm các cặp “Nhóm công việc” - “Danh mục công việc trong nhóm”.

Городок продвинутый, поэтому диспетчерская служба может получать заявки как по телефону, так и через Интернет, и даже через мессенджеры. В любом случае это получается набор коротких текстов, которые надо каким-то образом оперативно классифицировать. Thị trấn rất tiên tiến nên dịch vụ điều phối có thể nhận yêu cầu bằng cả điện thoại và qua Internet, thậm chí qua tin nhắn tức thời. Trong mọi trường hợp, đây hóa ra là một tập hợp các văn bản ngắn cần được phân loại nhanh chóng bằng cách nào đó.

Данные для построения модели:

1. Сведения о ранее обработанных заявках (файл «*L.csv*») со следующими параметрами:
   * *Number* - Идентификатор заявки;
   * *Date* - дата и время регистрации;
   * *UID* - Уникальный идентификатор адреса дома;
   * *Desc* - Собственно текст заявки, который необходимо классифицировать;
   * *Group* - Группа работ;
   * *Cat* - Категория работ в группе.

Исходные данные:

1. Файл с новыми ещё не классифицированными заявками ***C.csv (предоставляется заранее)*** с параметрами, аналогичными файлу «*L.csv*» из данных для построения модели, но с пустыми *Group* и *Cat*. Một tệp chứa các ứng dụng mới, chưa được phân loại C.csv (được cung cấp trước) với các tham số tương tự như tệp “L.csv” từ dữ liệu để xây dựng mô hình, nhưng không có *Group* và *Cat*.

Выходные данные:

1. Файл ***A.csv***, который получился заполнением параметров *Group* и *Cat* (в исходном файле они пустые). Tệp A.csv, có được bằng cách điền vào các tham số Nhóm và Cat (chúng trống trong tệp nguồn).
2. (не обязательно) Файл ***A-2.csv*** аналогичного формата, где классификация проведена иначе. (tùy chọn) Tệp A-2.csv có định dạng tương tự, trong đó việc phân loại được thực hiện khác nhau.
3. (не обязательно) Файл ***A-3.csv*** аналогичного формата, где классификация проведена иначе.

Примечания:

* Формат всех файлов CSV: с заголовком, разделители - «;».
* Если будет реализовано несколько различных моделей классификаторов, засчитываться будет лучший результат.

На проверку отправляется архив, в котором присутствует:

* Файл ***README.pdf*** с кратким описанием применённого решения и инструкцией по сборке программ.
* Исходные тексты программы, формирующей модель, в папке ***/src-train***.
* Исходные тексты программы, преобразующей исходные данные в выходные на основании заранее сформированной (обученной) модели в папке ***/src-predict***.
* Собственно модель в папке ***/model***.
* Выходные файлы ***A.csv*** (***A-2.csv*** и ***A-3.csv*** при наличии) в папке ***/result***.

**Комиссия оставляет за собой право самостоятельно сформировать модель по предоставленным скрипам обучения и произвести классификацию на основе предоставленных исходных текстов и инструкции.**