Импорт библиотек ¶

In [1]:

```
from subprocess import check_output
from tqdm import tqdm_notebook
import os
import pandas as pd
import numpy as np
import json
```

Настройка для работы с Kaggle API

In [ ]:

```
!mkdir ~/.kaggle
!touch ~/.kaggle/kaggle.json

api_token = {"username":"student","key":"abcdef0123456789"}

with open('/root/.kaggle/kaggle.json', 'w') as file:
    json.dump(api_token, file)

!chmod 600 ~/.kaggle/kaggle.json
```

Подготовка данных

In [ ]:

```
subs = []
files = os.listdir('/content/drive/My Drive/understanding_cloud_organization/new')
```

In [ ]:

```
for i in tqdm_notebook(range(len(files)), total=len(files)):
    subs.append(pd.read_csv('/content/drive/My Drive/understanding_cloud_organization/new/' + files[i], keep_default_na=True))
    subs[i]['ImageId'] = subs[i]['Image_Label'].apply(lambda x: x.split('_')[0])
    subs[i]['EncodedPixels'].replace({np.nan: ''}, inplace=True)
```

Вспомогательные функции для работы с масками

In [ ]:

```
def rle2mask(rle, input_shape):
    width, height = input_shape[:2]
    mask = np.zeros( width*height ).astype(np.uint8)
    array = np.asarray([int(x) for x in rle.split()])
    starts = array[0::2]
    lengths = array[1::2]
    current_position = 0
    for index, start in enumerate(starts):
        mask[int(start):int(start+lengths[index])] = 1
        current_position += lengths[index]
    return mask.reshape(height, width).T
```

In [ ]:

```
def mask2rle(img):
    pixels= img.T.flatten()
    pixels = np.concatenate([[0], pixels, [0]])
    runs = np.where(pixels[1:] != pixels[:-1])[0] + 1
    runs[1::2] -= runs[::2]
    return ' '.join(str(x) for x in runs)
```

Объединение предсказаний и отправка на Kaggle для подсчета метрики

In [ ]:

```
for i in range(len(subs)-1):
    for j in range(i+1, len(subs)-1):
        sub_1 = subs[i].copy(deep=True)
        sub_2 = subs[j].copy(deep=True)
        enc_pixels_list = []
        for index, row in sub_1.iterrows():
            mask_rle_1 = row['EncodedPixels']
            mask_rle_2 = sub_2.loc[index, 'EncodedPixels']

            mask_1 = rle2mask(mask_rle_1, (350, 525, 3))
            mask_2 = rle2mask(mask_rle_2, (350, 525, 3))
            mask = np.logical_and(mask_1, mask_2)
            enc_pixels_list.append(mask2rle(mask))

        sub_2['EncodedPixels'] = enc_pixels_list
        sub_2.drop('ImageId', axis=1, inplace=True)
        sub_2.to_csv(f'blended_mask_{i}_{j}.csv', index=None)
        fname = f'blended_mask_{i}_{j}.csv'
        msg = f"blend {files[i][28:]} + {files[j][28:]}"
        os.system(f"kaggle competitions submit -c understanding_cloud_organization -f
{fname} -m \"{msg}\"")
```