```
In [1]:    1  from datasets import load_dataset
           2  dataset = load_dataset("qanastek/MASSIVE", trust_remote_code=True)
```

C:\Users\MANOJ\.conda\envs\ProjectAutoKeras\lib\site-packages\tqdm\auto.p
y:21: TqdmWarning: IProgress not found. Please update jupyter and ipywidg
ets. See https://ipywidgets.readthedocs.io/en/stable/user_install.html (h
ttps://ipywidgets.readthedocs.io/en/stable/user_install.html)
  from .autonotebook import tqdm as notebook_tqdm

```
In [2]:    1  dataset.keys()
```

Out[2]: dict_keys(['train', 'validation', 'test'])

```
In [3]:    1  dataset.values()
```

Out[3]: dict_values([Dataset({
             features: ['id', 'locale', 'partition', 'scenario', 'intent', 'utt',
         'annot_utt', 'tokens', 'ner_tags', 'worker_id', 'slot_method', 'judgment
         s'],
             num_rows: 587214
         }), Dataset({
             features: ['id', 'locale', 'partition', 'scenario', 'intent', 'utt',
         'annot_utt', 'tokens', 'ner_tags', 'worker_id', 'slot_method', 'judgment
         s'],
             num_rows: 103683
         }), Dataset({
             features: ['id', 'locale', 'partition', 'scenario', 'intent', 'utt',
         'annot_utt', 'tokens', 'ner_tags', 'worker_id', 'slot_method', 'judgment
         s'],
             num_rows: 151674
         })])

```
In [5]:    1  locales = ['af-ZA', 'da-DK', 'de-DE', 'en-US', 'es-ES', 'fr-FR', 'fi-F
           2             'it-IT', 'jv-ID', 'lv-LV', 'ms-MY', 'nb-NO', 'nl-NL', 'pl-P
           3             'ru-RU', 'sl-SL', 'sv-SE', 'sq-AL', 'sw-KE', 'tl-PH', 'tr-T
```

```
In [6]:    1  import os
```

```
In [8]:    1  os.mkdir("Dataset")
```

```
In [9]:    1  for locale in locales:
           2      with open(f"Dataset/{locale}.txt",'w', encoding='utf-8') as file:
           3          for j in ['train','test','validation']:
           4              for i in dataset[j]:
           5                  if(i['locale'] == locale):
           6                      file.write(i['utt']+'\n')
           7
```

```
In [ ]:    1
```