|  |  |  |  |
| --- | --- | --- | --- |
| **Filtration** | **Number of Tweets** | **Number of Trained Tokens** | **Number of unique tokens** |
| **Scrapped Tweets** | **377,524,076** | **7,983,061,048** | **2,027,042** |
| **Arabic Tweets** | **267,201,072** |
| **Remove Duplicates** | **249403256** |
| **Detect Arabic Language** | **189141269** |