Description of Dataset:

For detecting events from Bengali and Banglish Facebook posts, a dataset named **"A dataset for event detection from Bengali and Banglish Facebook posts",** has been created. In this dataset, 200 Banglish (File 1) and 364 Bengali (File 2) Facebook posts have been included, and these posts were collected from various public pages, groups, and individual accounts. Besides, 8 features i.e., common event words, celebrating event words, protesting event words, religious events words, celebrating phrases, protesting phrases, religious phrases, and sentiment score, were extracted, and their associated binary values i.e., either 0 or 1, are also included in the dataset. Eventually, each post's designated output class i.e., celebrating, protesting, religious or neutral, is mentioned in the last column of the dataset. These classes are labeled using numbers from 0 to 3 where 0, 1, 2, and 3 define protesting, celebrating, neutral, and religious events respectively. For avoiding conflicts with persons and places, we replaced some letters in those words in posts with an asterisk sign (*).

As Bengali and Banglish both employ specific phrases to describe events, we have developed separate event specific phrase lists for Bengali and Banglish Facebook posts. These phrase lists with their corresponding English translations (File 3) represent event specific phrases for celebrating, protesting, and religious events.