1 Global COVID-19 X (Twitter) Dataset

The **Global COVID-19 X (Twitter) Dataset** is a comprehensive collection of tweets related to the COVID-19 pandemic, spanning from March 2020 to February 2022. It includes tweets from six countries: Australia, Brazil, India, Indonesia, Japan, and the United Kingdom, offering a wide array of public opinions and sentiments about the pandemic. This dataset is invaluable for conducting a detailed analysis of sentiments across different cultural and regional contexts, enabling a global perspective on public sentiment during the pandemic. The dataset includes the following tweet counts by country: Australia contributed 3,212,464 tweets, Brazil 943,913 tweets, India 5,411,294 tweets, Indonesia 229,935 tweets, Japan 644,510 tweets, and the United Kingdom 16,958,866 tweets. The COVID-19 pandemic had a profound impact on mental health, economies, and political environments in these countries. Significant events, such as lockdowns, vaccine distribution efforts, and elections, led to a surge in social media activity, all of which are captured in this dataset. This resource provides valuable insights into public sentiment during a time of global crisis.

In our research, we have utilized the **Global COVID-19 X (Twitter) Dataset** to analyze Hindu-phobic content using our specialized HP-BERT model. This dataset is publicly available on Kaggle and can be accessed via the following link **Global COVID-19 X (Twitter) Dataset**[1].

---

[1] https://www.kaggle.com/datasets/rohitashchandra/global-covid19-twitter-dataset