**Introduction:**

Vaccination is emerging as a critical tactic in the continuous fight to protect public health to reduce the dangers and spread of many diseases. Recent events highlight the crucial role that immunisation plays in limiting the COVID-19 pandemic's widespread effects. However, the need of mass vaccination goes beyond the current emergency and includes a wider range of disorders like childhood illnesses, recurring flu outbreaks, and more. Nevertheless, the scepticism that clouds the vaccination adoption landscape is a complicated scenario made of political machinations, worries about potential adverse effects, and other nuanced issues. Understanding the various issues that surround vaccines is essential to effectively addressing this complex topic. In this age of digital communication, the breadth of social media appears as a formidable library of viewpoints, enabling the quick gathering of insights on vaccine-related conversations.

**Problem Statement:**

The primary goal at hand is to create a strong multi-label categorization system capable of assigning meaningful labels to tweets. These labels are intended to encompass the authors' wide range of vaccine-related concerns. It's worth noting that a single tweet may have many categories, indicating the author's plethora of concerns. The classification task's taxonomy of concerns covers the following:

Unnecessary, Mandatory, Pharma, Conspiracy, Political, Country, Rushed, Ingredients, Side-effect, Ineffective, Religious, None.

**Method Description:**

We started with a methodical approach, isolating the tweets and their related labels, an essential step that provided the groundwork for following operations. Using the Multi-level Binarizer approach, we took advantage of one-hot encoding to convert the labels into a structured format suitable for computer analysis. The preprocessing step now comes into the picture, when a number of modifications were performed on the tweets. First, the tweets were converted to lowercase, which aligned them uniformly and improved their manageability. In order to achieve language coherence, emoticons, and emojis were replaced by written equivalents conveying analogous meanings. Furthermore, unnecessary features like hyperlinks and unneeded expressions were removed, resulting in a streamlined dataset ready for deeper analysis.

To overcome the aforementioned issues, we turned to the BERT-base model, a transformer architectural exemplar known for its contextual comprehension capabilities. The BERT model was used as the foundation of our solution, and it was configured to handle a contextual window of 150 tokens. Following that, the raw input data was encoded using the BERT Tokenizer, a tool capable of encapsulating text in a manner understandable by the BERT model. This encoding effort resulted in the extraction of critical components, including the Input Word ID, Attention Mask, and Input Type ID, which collectively permitted a sophisticated understanding of the structure and relevance of the input data.

In order to evaluate the model's efficacy, the dataset was partitioned, resulting in discrete training and validation sets. The primary model was a multi-layered composition resembling an elaborate mesh of computer capabilities. Its creation required the addition of three unique input layers: Input ID, Attention Mask, and Segment IDs. Following that, a Keras layer evolved as the hub of neural computing, with a Dense layer, synonymous with a fully connected layer, augmenting the depth of processing.

Recognizing the need of reducing overfitting and improving model adaptability, a careful deployment of Dropout layers was carried out. These layers reduce the focus on single data points and encourage balanced model performance. A Dense layer iteration was followed by the reintegration of a Dropout layer.

The last architectural layer took the shape of the output layer, which was meant to communicate the model's predictions. This layer formed as a channel via which the model passed on its insights and apprehensions, aided by still another Dense layer.

After methodically structuring the model, the next phase required systematic training of the model using the assigned training dataset. This preliminary stage is intended to implant in the model a thorough comprehension of the subtle patterns and attitudes included in the training data. This updated knowledge was subsequently applied when the model delved into providing predictions on the specified test dataset, encapsulating the overall goal.

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

Layer (type) Output Shape Param # Connected to

===================================================================================

input\_ids (InputLayer) [(None, 150)] 0 []

attention\_mask (InputLayer) [(None, 150)] 0 []

segment\_ids (InputLayer) [(None, 150)] 0 []

keras\_layer (KerasLayer) [(None, 768), 109482241 ['input\_ids[0][0]',

(None, 150, 768)] 'attention\_mask[0][0]',

'segment\_ids[0][0]']

dense (Dense) (None, 128) 98432 ['keras\_layer[0][0]']

dropout (Dropout) (None, 128) 0 ['dense[0][0]']

dense\_1 (Dense) (None, 64) 8256 ['dropout[0][0]']

dropout\_1 (Dropout) (None, 64) 0 ['dense\_1[0][0]']

output (Dense) (None, 12) 780 ['dropout\_1[0][0]']

===================================================================================

Total params: 109,589,709

Trainable params: 109,589,708

Non-trainable params: 1

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_