

Augnito Assignment Report

1. Data Understanding
 - 1.1. The data in dataset ECtHR B consists of text and label parts for both the training and test sets.
 - 1.2. The text part includes an array/list of strings that describe a case.
 - 1.3. Each case has a label part which consists of articles that were allegedly violated.
2. Data preprocessing : The text was converted to lowercase and non-alphanumeric characters were removed. The stop words were removed from the text.
3. Approach : I have attempted to use the BERT model.