



- ※1：入力された単語(seq\_len)のうちマスクされた不明単語がvocab(全単語)のどれかをSoftMax関数で計算する。(マスクされた単語の前後何単語とは指定しない点がCBOWと違う)
- ※2：「連続的に存在する意味が深い文章」or「全く関係がなく文脈のつながりがない2つの文章」のどちらかを判定する。
- ※3：最後のBertLayerの出力では文章の中にある単語とその周辺単語の関係をベースにその単語の意味にあった特徴ベクトルが生成される。
- 例えば、銀行のbankと土手のbankは同じ単語でも前後の文章によって意味が異なってくる。





