

This is today's agenda.

自然言語処理において、感情分析は注目を集めている研究である。

その中でもアスペクトベースセンチメント分析（ABSA）は文書をアスペクト情報に基づいて分析する研究である。

アスペクト情報とは文書のエンティティとアトリビュートからなる情報である。

ABSA を機械学習や深層学習を用いて解く際はマルチラベル分類モデルが構築される。

ABSA における MLC の精度を向上することは大きな問題である。

そこで、本研究ではその問題を改善するために、アンサンブル学習を導入したモデルを提案する。各アスペクトに特化した小分類器を作成して実験を行い、提案モデルの有効性を確認した。

モデルの作成にあたり、ABSA に取り組んでいる三浦らの提案手法を参考にした。

我々は Multi pre-trained models Transformer (Mpm+T) を提案する。

We describe the data flow in this model.

First, sentence tokens are generated from sentences.

Next, these tokens are input to multiple sub-classifiers, and CLS tokens are obtained from each of them.

Then each CLS token is fed to the Transformer Encoder with Positional embeddings.

The CLS tokens with MLC information are extracted and finally input to the Multilayer Perceptron to obtain the predictive labels.

Next, we introduce the model diagram of the sub-classifiers.

The general flow of data is the same as Mpm+T, but this one is a binary classifier.

We will explain the role of the sub-classifiers.

Entity classifiers are seven binary BERT models created for each category.

And attribute classifiers are 7 binary BERT models that determine whether a category is positive or negative.

我々は日本語のカスタマーレビューデータセットを用いて、提案手法の有効性を検証するために 5 分割交差検証をした。比較手法は従来手法である BERT+MLP と、先行研究である三浦らのモデルを選択する。

We use the " Rakuten Travel Reviews: Aspects and Sentiment-tagged corpus" published by Rakuten Group, Inc.

Here are the training parameters for each model.

表に Entity classifiers と attribute classifiers の各カテゴリにおける正解率と標準偏差を示す。いずれのカテゴリにおいても小分類器は高い精度で分類できていることがわかる。しかし、サービスと設備の Entity を 2 値分類するモデルの精度が低くなっていることがわかる。We believe this is due to the fact that the data is more imbalanced than for the other categories.

The results of the 5-fold cross-validation for each model are shown.

The effectiveness of the proposed model was confirmed by comparing each evaluation index.

新たな評価指標の定義をします.

We define a completely correct when the prediction and the correct label are completely matched. And we define a partially correct when the prediction and the correct label are partially matched.

表に 6000 のテストデータにおける completely correct と partially correct の数を示す.

比較手法と提案モデルの Completely correct の数を比較すると, 大きく向上していることがわかる.

表に, それぞれのモデルの場合のマルチラベルデータとシングルラベルデータにおける完全正解データの数と割合を示す. 特にマルチラベルデータの完全正解スコアが比較手法よりも向上していることがわかる. つまり, Mpm+T は従来手法に比べてマルチラベルデータの分類に特化した MLC モデルであるといえる.

マルチラベルデータの正解例と不正解例を紹介する.

表にラベル数が 3 の場合の正解例のテキストとラベルを示す.

表にラベル数が 7 の場合の正解例のテキストとラベルを示す.

ラベル数が 7 のデータの約 8 割が表に示すような, 全てのポジティブまたはネガティブラベルが付与されていた.

表にラベル数が 6 の場合の不正解例テキストと予測ラベルと正解ラベルを示す.

ラベル数が 6 の正解例は 0 件である.

ラベル数 6 のテストデータは人間が見ても分類が難しいテキストであることに加えて, 学習データ数が極端に少なかったためであると考えられる.