

論文 / 著書情報  
Article / Book Information

|                   |   |
|-------------------|---|
| 題目(和文)            |   |
| Title(English)    | One-class Text Classification with Neural Networks  |
| 著者(和文)            | HuChenlong  |
| Author(English)   | Chenlong Hu   |
| 出典(和文)            | 学位:博士(工学),<br>学位授与機関:東京工業大学,<br>報告番号:甲第12070号,<br>授与年月日:2021年9月24日,<br>学位の種別:課程博士,<br>審査員:奥村 学,熊澤 逸夫,中山 実,篠崎 隆宏,船越 孝太郎  |
| Citation(English) | Degree:Doctor (Engineering),<br>Conferring organization: Tokyo Institute of Technology,<br>Report number:甲第12070号,<br>Conferred date:2021/9/24,<br>Degree Type:Course doctor,<br>Examiner:,,,,, |
| 学位種別(和文)          | 博士論文  |
| Category(English) | Doctoral Thesis   |
| 種別(和文)            | 審査の要旨   |
| Type(English)     | Exam Summary  |

(博士課程)

## 論文審査の要旨及び審査員

| 報告番号        | 甲第  | 号     | 学位申請者氏名 | Hu Chenlong |       |     |
|-------------|-----|-------|---------|-------------|-------|-----|
| 論文審査<br>審査員 |     | 氏名    | 職名      |             | 氏名    | 職名  |
|             | 主査  | 奥村 学  | 教授      |             | 船越孝太郎 | 准教授 |
|             | 審査員 | 熊澤 逸夫 | 教授      | 審査員         |       |     |
|             |     | 中山実   | 教授      |             |       |     |
|             |     | 篠崎隆宏  | 准教授     |             |       |     |

### 論文審査の要旨 (2000 字程度)

テキスト分類問題の1種である1クラステキスト分類は、特定のクラスの正例のみでモデルを学習する。正例は入手可能だが、負例を同時に入手するのが困難であるタスクは多く、そのような場合に1クラステキスト分類は有効である。本論文は「One-class Text Classification with Neural Networks」と題し、英文全5章より構成されている。

第1章「Introduction」ではまず、1クラステキスト分類問題の定義、研究の歴史、手法の種類について説明している。また、本論文では、1クラステキスト分類問題に対するニューラルネットワークを用いた手法に焦点を当てることを述べている。そして、ブートストラップ法に基づく意味知識獲得が、少量の目標クラスのデータを種として目標クラスのデータを漸次的に学習していくタスクであることから、1クラステキスト分類問題として解けることを説明し、ニューラルネットワークを用いてブートストラップ法の各プロセスをどのように改良するかが本研究で解くべき1つ目の課題であることを述べている。また、目標クラスのデータのより良い記述を構築するため、ニューラルネットワークを用いた1クラステキスト分類手法である、deep Support Vector Data Description (deep SVDD)を多峰型に拡張することが本研究で解くべき2つ目の課題であることを述べている。

第2章「Related Work」では、1クラステキスト分類、ブートストラップ法の先行研究について説明している。

第3章「AutoEncoder Guided Bootstrapping of Semantic Lexicon」ではまず、ブートストラップ法の1つであるBasiliskについて説明している。そして、Basiliskを元にしたブートストラップ法の枠組みが1クラステキスト分類問題の1種として定式化できることを述べ、その学習サイクルに対する教示を与えるためAutoEncoderを導入する提案手法について述べている。最後に、AutoEncoderを教示として用いて意味知識をブートストラップする提案手法が意味知識獲得の性能を全体的に向上できることを実験的に示している。

第4章「One-class Text Classification with Multi-modal Deep Support Vector Data Description」では、提案するmulti-modal deep Support Vector Data Description (mSVDD)の枠組みをまず述べ、複数の超平面を用いてどのように目標クラスのデータを記述するかを説明している。また、いくつかの命題を証明し、mSVDDと、ニューラルネットワークを元にした、いくつかの1クラステキスト分類モデルの関係を明らかにしている。さらに、1クラステキスト分類では入手が困難とされる負例を疑似的にmSVDDに教示する手法についても述べている。最後に、mSVDDが単峰型のdeep SVDDよりも性能的に優れており、提案する疑似的な負例の教示手法を用いることにより、さらに性能を向上させることができることを実験的に示している。

第5章「Conclusion」では、本研究の結論と今後の課題を説明している。今後の課題では、本論文で取り組んでいないものの、残された重要な課題を述べている。

以上を要するに、本論文は、ニューラルネットワークに基づく1クラステキスト分類モデルを高度化するとともに、意味知識獲得タスクを1クラステキスト分類問題として初めて定式化した上で、提案手法の有効性を検証しており、工学上貢献するところが大きい。よって博士(工学)の学位を授与するに十分な価値を持つものと認められる。

注意:「論文審査の要旨及び審査員」は、東工大リサーチポジトリ(T2R2)にてインターネット公表されますので、公表可能な範囲の内容で作成してください。