

(修士課程)  
Master's Program

令和 4 年 1 月 19 日

Date (yymmdd)

系・コース： Department of, Graduate major in	情報理工 知能情報	系 コース
学籍番号： Student ID Number	20M31190	
学生氏名： Student's Name	吉川 純平	
指導教員（主）： Academic Supervisor(main)	村田 剛志	
指導教員（副）： Academic Supervisor(sub)		

審査員主査： Chief Examiner	村田 剛志
審査員： Examiner	岡崎 直観
審査員： Examiner	徳永 健伸
審査員： Examiner	
審査員： Examiner	

## 論文要旨

THESIS SUMMARY

論文題目 Thesis Title	グラフ構造を用いて欠損値埋めを行うネットワークの半教師あり学習
----------------------	---------------------------------

要旨（和文 1000 字程度又は英文 400 語程度）

Thesis Summary (approx.1000 Japanese Characters or approx.400 English Words )

近年 Twitter や Facebook などの SNS の普及によりネットワーク上でのコミュニケーションが可能になったことや、化合物の物性推定の重要性が高まったなどにより、グラフ構造を含むデータの分析が注目を集めている。グラフとは「ノード」と、二つのノード間を結ぶ「エッジ」から構成されるデータ構造である。従来の研究において、グラフ構造とそれぞれのノードの特徴量を用いて予測する Graph Neural Network (GNN) が優れた成果を残している。その中でも、1 層で距離 1 の隣接ノードの特徴量の畳み込みを行う Graph Convolutional Network (GCN) が、半教師ありノード分類やリンク予測などの様々なタスクで高い精度を示している。

GCN は学習の際に、特徴量に欠損値を含まないことを前提とした手法である。しかし実世界において欠損値を含むデータは数多く存在する。例えば人的ミス、センサーのエラー、任意回答における未入力項目を含むアンケートなどにより欠損値を含むデータを得られることがある。従来の欠損値を含むグラフデータに欠損値補完をする手法の多くは、グラフ構造を無視した機械学習手法により欠損値を穴埋めし、得られたデータの特徴量として GCN などのモデルを学習させる手法が一般的である。この方法は代入法にグラフ構造を用いないため予測精度が低くなる可能性がある。

本研究ではこのような欠損値を含むグラフデータに対して、グラフ構造を用いて欠損値補完を行うことで、GCN を用いた予測精度を向上させる手法を提案する。提案手法はグラフの近接ノードの情報を再帰的に集約し更新する手法を用いて欠損値を補完し、GCN を用いて予測するモデルにより構成されている。提案手法はグラフ構造を用いて欠損値補完を用いることで、GCN を用いて予測するために適した特徴量を扱えるため、予測精度を向上させることができた。

提案手法の有効性を示すために半教師ありノード分類とリンク予測の 2 つのタスクを実験として行った。実験ではノードの特徴量は欠損値の割合を 10%から 90%まで 10%刻みで変化させて予測精度を調べた。実験の結果、主に欠損率が高い場合に提案手法は既存手法と比べて高い精度を得ることを確認した。

備考 上記の論文には、大学院学則第 3 4 条第 3 項に規定する特定の課題についての研究の成果を含む。  
note The above thesis includes the result of research on a specific theme noted in clause 3, article 34 of the Institute Regulations.  
**For Students of the Integrated Doctoral Education Program Only.**