

深層強化学習を用いたノード分類

柳田 雄輝

グラフは様々なデータを表現することが可能であり汎用性が高いため、これまでに多くのグラフ解析手法が提案されてきている。特にグラフ中のノード分類には様々な手法が考案されてきており、近年では、ニューラルネットワークに基づいた手法が主流である。しかしながら、既存のニューラルネットワークに基づくノード分類手法では、グラフ内の全ノードについて計算する必要があり計算に時間がかかってしまうことに加え、分類の予測に有用な情報だけを正確に利用できない可能性がある。

そこで、本論文では、分類に用いるノードを探索でき、かつ、どのようなノードに基づいてノードが分類されたのかを明らかにできる手法を提案する。この手法では分類対象となっているノードの特徴だけではなく、分類すべきノードを起点としたグラフ上のパスをノード分類の根拠として用いる。これにより、ノードの分類を行う上でグラフ内の核となる特徴を持つ部分だけを捉えることができ、かつ、多様な根拠に基づくノードラベルの分類を可能とする。パスを根拠とした分類には強化学習を用い、エージェントにグラフ上のパスを探索させることによって、分類根拠としてふさわしいパスを学習する。エージェントはノードのラベル数分作成し、エージェント l がラベル l を予測する多クラス分類問題を解く。各エージェントは分類したいノードを起点にグラフ上を探索し、そのラベルに分類するにあたって妥当な根拠となるパスを獲得する。そして、規定ステップ数以内に分類に適していると予測されるパスをたどった場合には観測成功という行動を取るようにする。規定回数探索を行い、各エージェントの中で最も多く観測成功という行動を採用したエージェントに基づいてノードラベルを予測する。エージェントがパスをたどる方策は Q 学習を用いて学習する。エージェントがたどったノード系列を入力とし、隣接する各ノードへの遷移確率と観測成功とする確率を出力とする関数を学習する。

実験にあたっては、次の 2 種類の人工データを作成し実験に用いた：(1) ノードが観測可能なノードラベルとエッジラベルを持つが特徴量は持たないデータ (2) ノードが特徴量を持つが観測可能なノードラベルやエッジラベルを持たないデータ。いずれのデータもおおよそ 2,000 個のノードを持つが、このうち 100 個のノードは観測できないノードラベルを持つ。本論文の実験においては、この観測できないノードラベルを予測できるかを評価した。実験の結果、既存手法とは異なる根拠に基づいて分類が行えることを確認することができた。

この論文における貢献は以下の 2 つである：(1) ノード分類タスクにおいて、分類根拠をパスに基づいて提示する手法を提案した。(2) 人工データを用いた実験の結果、提案手法が従来手法とは異なる分類および可視化ができることを示した。

(指導教員 加藤 誠)