

先端科学技術研究科 修士論文要旨

所属研究室 (主指導教員)	知能コミュニケーション (中村 哲 (教授))		
学籍番号	2011273	提出日	令和 4年 1月 20日
学生氏名	矢野 祐貴		
論文題目	構造情報を利用するData-to-text生成		
要旨			
<p>近年の事前学習済みTransformerモデルであるT5を用いたData-to-text生成は, その入力方法に諸般の問題を抱えている. 本稿では, まず線形化においてレコード順序を無作為化することで, それが生成文に与える影響について検証する.</p> <p>また, 線形化における構造情報の欠落の問題を解決するため, Token列としてデータを入力する事に加え, レコードを単位とする入力を加えることで, Data-to-textを拡張する方法について検討する.</p> <p>実験結果から, レコード順序の無作為化を行わない場合, T5はデータセットの頒布された元の並び順を手掛かりとしており, これを無作為化することで, 未知のドメインデータに対する生成精度を改善することがわかった.</p> <p>また, レコードを単位とする入力を加える手法は従来のモデルと同等以上の性能を達成し, 更なる改善の方向性を示した. また, レコードを単位とする入力のみで構成されたモデルであっても, 既知ドメインでは元のモデルから大きく精度が劣らない生成文を出力できることがわかった.</p>			