

著者識別を用いた文書の改ざん検知

研究駆動コース 松林由佑

背景



- ✓ 一部分が書き換えられた報告書
→ 原本がないと、どこが書き換えられたのかわからない！
- ✓ 書き換え部分は、本文と書いている人が違う
→ その微妙な違いが文章に現れるのでは？！

「書き手の特徴」を抽出

文章を品詞分解

「私/は/今日/の/朝/起きて/、/勉強/を/した」

内容語（文章の意味が分かる単語）を品詞に変換

[名詞・は・名詞・の・名詞・動詞・、・名詞・を・動詞・た]

手法①

リスト（ベクトル）に登場割合を入れる

[32%, 0%, 16%, 8%, 0%, 8%, 9%, 9%.....]

↑ ↑ ↑ ↑ ↑ ↑
名詞 動詞 、 。 の 名詞+は 名詞+を

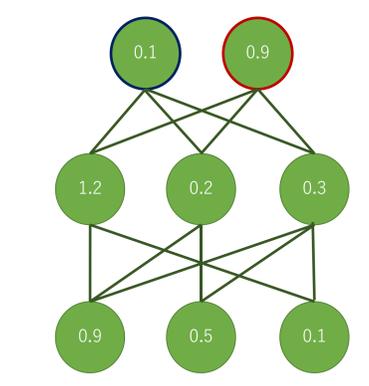
全結合ニューラルネットワーク / Random Forest

手法②

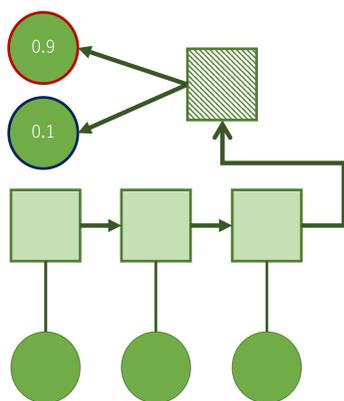
二次元配列（行列）に入れる

	名詞	は	...	た	_
名詞	1	0	...	0	0
助詞	0	1	...	1	0
動詞	0	0	...	0	0
...
。	0	0	...	0	0
の	0	1	...	0	0
...

RNN / CNN



全結合ニューラルネットワーク

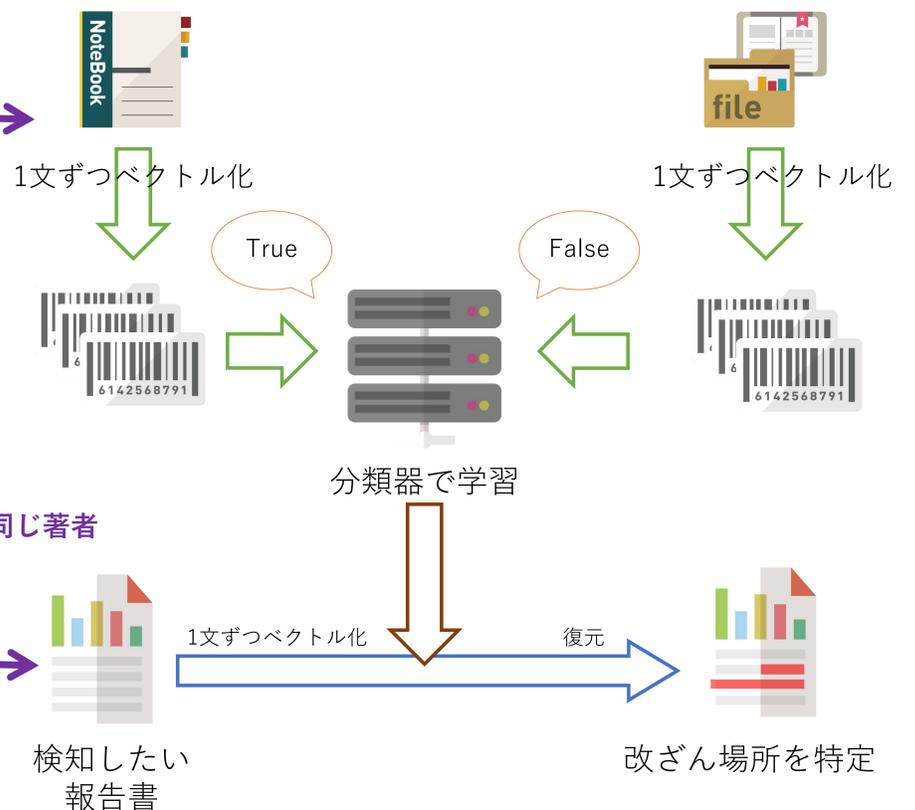


RNN

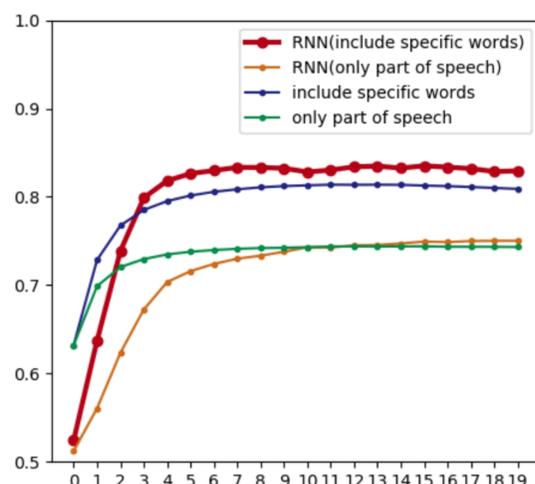
機械学習を用いて判定

報告書の著者が書いた文章

違う人が書いた文章



結果 & 今後の課題



- ・まだまだ、実用化レベルとは言えないが、目安程度にはなる精度
- ・ネットワークとベクトルの改善が今後の課題