

Automatically Learning Patterns for Self-Admitted Technical Debt Removal

Fiorella Zampetti

Alexander Serebrenik

Massimiliano Di Penta

2020 IEEE 27th International Conference on Software Analysis, Evolution and Reengineering (SANER)

背景・目的

- 多くの研究でSATDの除去について調査されている
- ZampettiらはSATDの除去時の6つの戦略を提案^[58]
- ただし、具体的な開発者の支援になっていない
- 各SATDに該当する除去戦略を推薦モデルの提案

[58] F. Zampetti, A. Serebrenik, and M. Di Penta, “Was self-admitted technical debt removal a real removal?: an in-depth perspective,” in Proceedings of the 15th International Conference on Mining Software Repositories, MSR 2018, Gothenburg, Sweden, May 28-29, 2018, 2018, pp. 526–536.

6つの戦略

TABLE I
SATD REMOVALS DATASET BY ZAMPETTI *et al.* [58].

Change	Camel	Gerrit	Hadoop	Log4J	Tomcat	Total
Method Calls	165	16	42	8	59	290
Conditionals	61	8	13	7	59	148
Try-Catch	9	3	2	0	5	19
Method Signature	36	3	6	0	17	62
Return Stmt	15	3	4	0	7	29
Other changes	145	23	67	10	94	339
OVERALL	431	56	134	25	241	887

Method Calls

メソッドの呼び出しに関する変更

Conditional Statement

条件文に関する変更

Try-Catch Blocks

例外処理に関する変更

Method Signature

メソッド名、引数、データ型などの変更

Return Statements

Return文に関する変更

Other

5つに該当しなかった変更

SARDELEの概要

SATD Removal Deep Learning

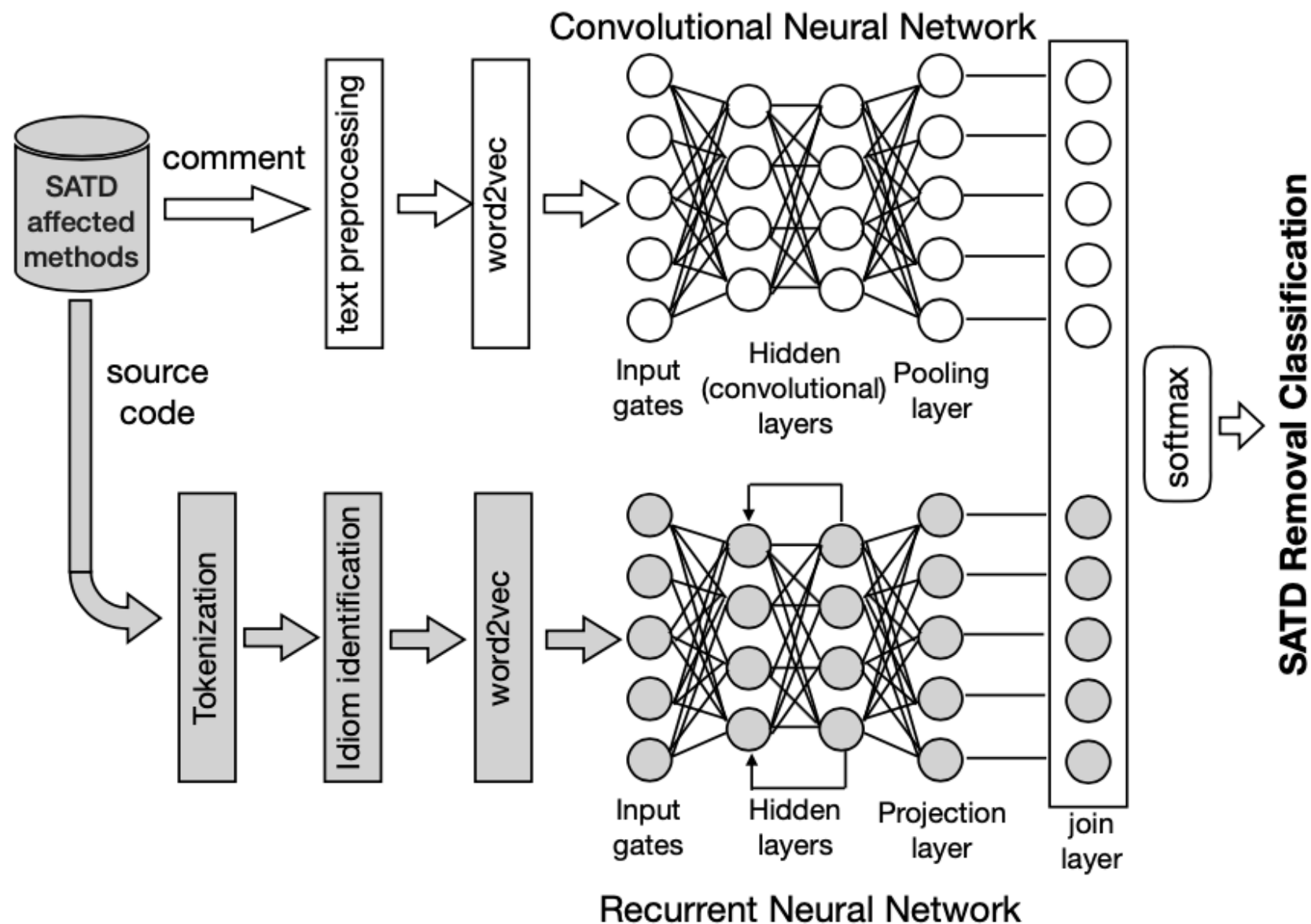


Fig. 1. SATD classification approach (the number of hidden layers and of nodes in the neural network varies based on the calibration).

コメントのモデル構築

1. 前処理

- 特殊な文字や数字の除去
- 小文字に統一
- Stop wordの除去やステミング処理

2. 単語のベクトル化

- skip-gramを利用

3. 分類器の構築

- CNN (Convolutional Neural Network) を利用

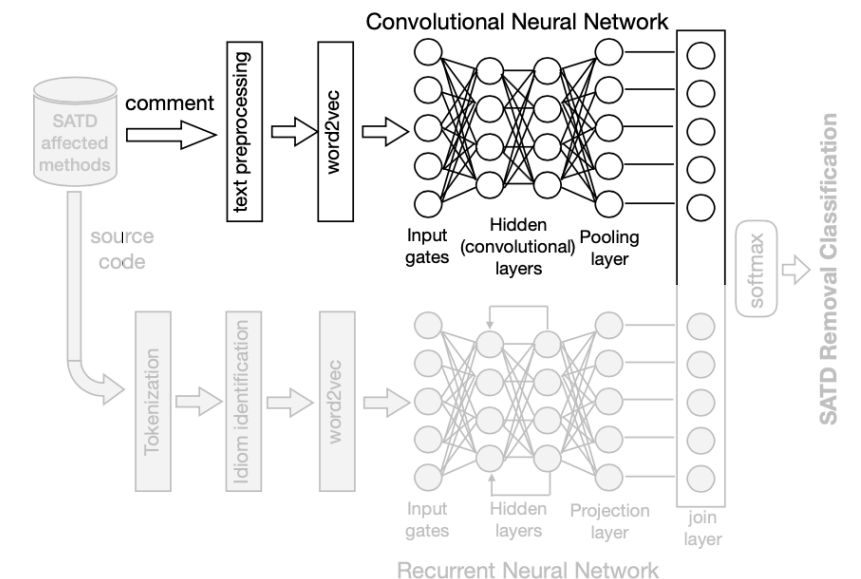


Fig. 1. SATD classification approach (the number of hidden layers and of nodes in the neural network varies based on the calibration).

ソースコードのモデル構築

1. 前処理

- GumTreeを用いてTokenization
- Stop wordの除去やステミング処理

2. 単語のベクトル化

- skip-gramを利用

3. 分類器の構築

- RNN（Recurrent Neural Network）であるLSTMを利用

- CNNとRNNの出力層をsoftmax関数に入力し分類

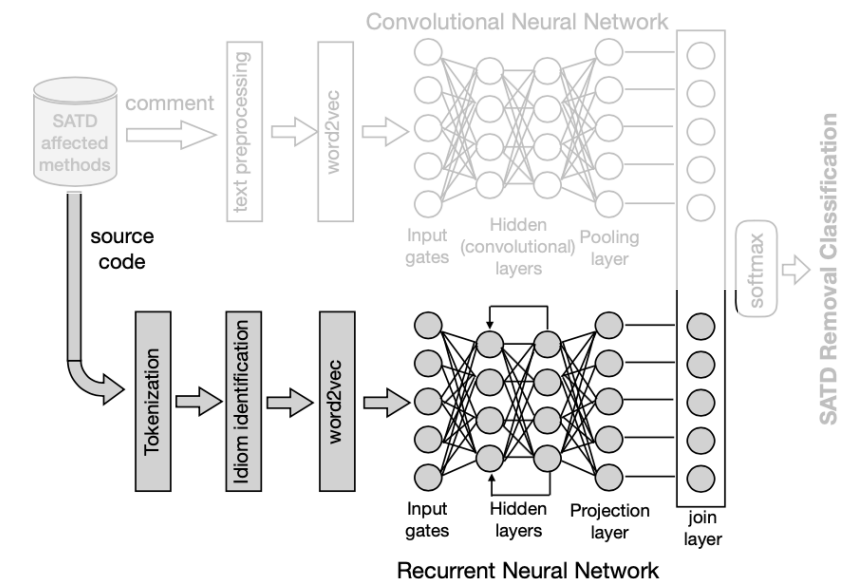


Fig. 1. SATD classification approach (the number of hidden layers and of nodes in the neural network varies based on the calibration).

Research Questions

- RQ1

それぞれの分類器は精度はどうか？

- RQ2

単なる機械学習モデルと比べてSARDELEはどうか？

- RQ3

人手に比べてSARDELEはどうか？

RQ1：精度はどんな感じ？

コメント

TABLE II

COMMENT CLASSIFICATION WITH CNN: PERFORMANCES ACROSS THE SIX SATD REMOVAL CATEGORIES.

Category	Pr	Rc	F ₁	AUC	MCC
Method Calls	50.32	34.05	40.62	0.56	0.13
Conditionals	38.02	38.66	38.33	0.60	0.19
Try-Catch	21.05	26.67	23.53	0.61	0.20
Method Signature	34.09	30.00	31.91	0.61	0.23
Return	34.62	39.13	36.73	0.67	0.33
Other	58.26	73.63	65.05	0.63	0.26
OVERALL	39.39	41.03	39.04	0.61	0.22

ソースコード

TABLE III

SOURCE CODE CLASSIFICATION WITH RNN: PERFORMANCES ACROSS THE SIX SATD REMOVAL CATEGORIES.

Category	Pr	Rc	F ₁	AUC	MCC
Method Calls	58.68	30.60	40.23	0.59	0.21
Conditionals	47.48	55.46	51.16	0.69	0.35
Try-Catch	33.33	33.33	33.33	0.65	0.31
Method Signature	52.00	26.00	34.67	0.61	0.31
Return	33.33	30.43	31.82	0.63	0.28
Other	59.59	85.35	70.18	0.68	0.38
OVERALL	47.40	43.53	43.56	0.64	0.31

SARDELE

TABLE IV

COMBINED CLASSIFICATION (SARDELE) PERFORMANCES ACROSS THE SIX SATD REMOVAL CATEGORIES.

Category	Pr	Rc	F ₁	AUC	MCC
Method Calls	73.13	63.36	67.90	0.74	0.50
Conditionals	58.47	57.98	58.23	0.73	0.47
Try-Catch	38.89	46.67	42.42	0.72	0.41
Method Signature	50.00	48.00	48.98	0.71	0.44
Return	42.31	47.83	44.90	0.72	0.42
Other	69.10	76.19	72.47	0.75	0.49
OVERALL	55.32	56.67	55.82	0.73	0.46

SARDELE > ソースコード > コメント

RQ2 : 機械学習と比べてどう？

SARDELE

TABLE IV

COMBINED CLASSIFICATION (SARDELE) PERFORMANCES ACROSS THE SIX SATD REMOVAL CATEGORIES.

Category	Pr	Rc	F ₁	AUC	MCC
Method Calls	73.13	63.36	67.90	0.74	0.50
Conditionals	58.47	57.98	58.23	0.73	0.47
Try-Catch	38.89	46.67	42.42	0.72	0.41
Method Signature	50.00	48.00	48.98	0.71	0.44
Return	42.31	47.83	44.90	0.72	0.42
Other	69.10	76.19	72.47	0.75	0.49
OVERALL	55.32	56.67	55.82	0.73	0.46

Random Forest

TABLE VI

RANDOM FOREST PERFORMANCES ACROSS THE SIX SATD REMOVAL CATEGORIES.

Category	Pr	Rc	F ₁	AUC	MCC
Method Calls	58.70	27.90	37.90	0.62	0.21
Conditionals	72.20	8.80	15.70	0.60	0.21
Try-Catch	11.80	10.50	11.10	0.54	0.09
Method Signature	69.20	14.50	24.00	0.64	0.30
Return	6.90	6.90	6.90	0.52	0.03
Other	51.40	37.50	43.30	0.59	0.11
OVERALL	45.03	17.68	23.15	0.59	0.16

SARDELE > Random Forest

RQ3 : 人と比べてどう？

TABLE VII
CONFUSION MATRIX COMPARING THE MANUAL CLASSIFICATION AND THE
CNN ON THE SATD COMMENTS.

		CNN is correct		Total
		No	Yes	
Manual classification is correct	No	288	283	571
	Yes	70	71	141
Total		358	354	712

人手の分類：SATDのコメントを読んで推測

人手の分類の正解数：141（712件中）

分類器の正解数：354（712件中）

推薦モデルの方が優れている

まとめ

- これまでのSATD除去の調査をもとにSATDの除去戦略推薦モデルを構築
- コメント，ソースコードの2つを組み合わせたSARDELEが一番精度が良かった
- SATDコメントから除去戦略を推測する場合，人手に比べて推薦モデルの方が精度が良かった

所感

- どうやってSATDが関係するソースコードと結びつけているのか？
- 人手と推薦モデルで、正確に分類できる戦略はことにならないのか？
- SATDコメントから推測できればかなり有用だと思うが、コメントは情報量が少ないので無理があるのでは？