

STRESZCZENIE

Application of multimodal neural networks in solving problem of labeling bug reports

Zastosowanie modalnych sieci neuronowych w rozwiązaniu problemu
etykietowania raportów o błędach

Author: Łukasz Chmielowski, MSc, Eng.

Under the supervision of:

Prof. Robert Burduk, DSc, PhD, Eng.

Michał Kucharzak, PhD, Eng.

W dużych organizacjach zajmujących się rozwojem oprogramowania występują często złożone procesy związane z obsługą raportów o defektach powstałych podczas jego wytwarzania. Takie dokumenty o awariach systemu, zazwyczaj zawierają tytuł i opis rozbieżności w zachowaniu rozwiązania w stosunku do oczekiwanego w formie języka naturalnego. Dodatkowo są załączane informacje systemowe w formacie surowym lub przetworzonym przez narzędzia je analizujące. Głównym zagadnieniem jest przypisanie raportu do odpowiedniej komórki organizacyjnej oraz określenie czy dany raport jest powiązany z zagrożeniem dla bezpieczeństwa, nieprawidłowościami z pamięcią lub wydajnością. Wykorzystane w badaniach eksperymentalnych raporty pochodzą od operatorów telekomunikacyjnych bądź zostały zgłoszone w firmie zajmującej się wytwarzaniem oprogramowania i sprzętu telekomunikacyjnego. Dotyczą one stacji bazowych (BTS).

Badania potwierdziły tezę mówiącą, że *istnieje metoda automatycznego przypisywania raportu o błędzie oprogramowania do odpowiedniej grupy badawczo-rozwojowej, odpowiedzialnej za rozwiązanie błędu, która przewyższa dobrze znane metody przypisywania raportów o błędach.*



This work is licensed under a [Creative Commons Attribution-ShareAlike 4.0 International License](https://creativecommons.org/licenses/by-sa/4.0/).

Dokonano przeglądu literatury. Jako jeden z istotnych wniosków było zdefiniowanie luki badawczej wskazującej wady aktualnie używanych technik walidacyjnych oraz zaproponowanie alternatywnej pozbawionej wyżej wspomnianych negatywnych cech. Przeprowadzone eksperymenty przy użyciu protokołów wykorzystujących różne sposoby budowania zbiorów trenujących i testowych wykazały istotną różnicę w wynikach. Wykonano analizę możliwości zastosowania wyjaśnianej (wytłumaczalnej) sztucznej inteligencji.

Przeprowadzono szereg badań zarówno związanych z preprocessingiem, wektoryzacją danych źródłowych jak i na różnych poziomach struktury firmy. W ramach realizacji doktoratu zaproponowano nową metodę przypisywania raportów w kontekście kompozycji organizacji. Na postawie badań skuteczności wybranych algorytmów uczenia maszynowego, podjęto decyzję oraz wdrożono pilotażowe rozwiązanie w firmie. Ze względu na to, iż udało się zebrać w określonych warunkach zarówno predykcje modelu uczenia maszynowego jak i decyzje ludzi, zdecydowano się przeprowadzić analizę porównawczą wyników.

Badano możliwość zastosowania multimodalnych sieci neuronowych jak również innych sposobów rozwiązania problemu przypisywania raportów. Kluczowym aspektem było przygotowanie odpowiedniej reprezentacji danych wejściowych jak i zaprojektowanie wymaganej architektury multimodalnej sieci neuronowej. Przeprowadzone badania wykazały wyższość tej metody w porównaniu do zastosowanych metod referencyjnych. Zaświadczenie o pracach wdrożeniowych w firmie zostało umieszczone w załączniku. Zawiera ono informacje o trwających pracach związanych z wprowadzeniem multimodalnej sieci neuronowej. Pierwsze jej wersje są już wdrożone. Obecnie trwają prace związane z lepszym doбором cech i walką z brakującymi danymi. Ponadto zawarte są tam informacje na temat innych powiązanych wdrożeń związanych z raportami o błędach w oprogramowaniu.

Zubran Chmelowski