

Streszczenie

W ostatnich latach odnotowano gwałtowny wzrost w obszarze modeli głębokiego uczenia, które zostały zaprojektowane w celu wspierania lekarzy. Jednak ze względu na to, że opieka zdrowotna jest szczególnie wrażliwą dziedziną, pojawiła się potrzeba wyjaśniania tych modeli. Wspomniane zapewnienie jakości modeli i opracowanie bardziej odpowiedzialnych rozwiązań ma zasadnicze znaczenie w ich potencjalnym zastosowaniu do danych klinicznych.

Celem niniejszej pracy doktorskiej jest zbadanie pozytywnego wpływu na metryki jakości modeli wieloaspektowych i wyjaśnień zorientowanych na człowieka. Główna teza jest poparta wprowadzonymi innowacjami, utworzonymi modelami, przeprowadzonymi eksperymentami i opracowanymi nowymi metodami, opartymi na sześciu artykułach.

W tej rozprawie kluczowy wkład w uczenie głębokie obejmuje stworzenie modelu wielozadaniowego zdolnego do wykrywania nawet małych zmian chorobowych oraz wprowadzenie pojęcia polireprezentacji, które obejmuje łączenie reprezentacji uzyskanych na różne sposoby z danych tej samej modalności. Ponadto w rozprawie przedstawiono kompleksowy przegląd artykułów naukowych wykorzystujących modele uczenia głębokiego do analizy obrazów płuc w celu identyfikacji COVID-19. Prace te doprowadziły do opracowania listy kontrolnej wspierającej tworzenie odpowiedzialnych rozwiązań uczenia głębokiego na obrazach medycznych.

Znaczący wkład w wyjaśnialną sztuczną inteligencję obejmuje stworzenie interaktywnej metody LIMCraft, która poprawia jakość wyjaśnień, umożliwiając użytkownikom wybieranie semantycznie spójnych obszarów obrazu i sprawdzanie przewidywań modelu po modyfikacji obrazu. Praca wprowadza również dwie różne metody łączenia wyjaśnień: NormEnsembleXAI i metodę opartą o sieci konwolucyjne.

Podsumowując, niniejsza rozprawa odpowiedziała na wszystkie postawione pytania badawcze oraz osiągnęła wszystkie cele, zapewniając znaczący postęp w wyjaśnialności i niezawodności modeli uczenia głębokiego pracujących na danych medycznych, dokonując zmiany w kierunku bezpieczniejszych i bardziej godnych zaufania systemów sztucznej inteligencji.

Abstract

The area of deep learning models that were designed to support medical practitioners has faced unprecedented growth in recent years. Furthermore, due to the fact that healthcare is a particularly sensitive domain, the need for interpretability of deep learning models is prominent. Ensuring the quality of these models and developing more responsible approaches is essential for their potential applications on clinical data.

The purpose of this PhD thesis is to investigate the positive impact of multifaceted models and human-oriented explanations on quality metrics. The main thesis is supported through introduced innovations, created models, conducted experiments, and developed new methods, grounded in six research papers.

The key deep learning contributions to this thesis include a multi-task model capable of detecting even small lesions and introduces a concept of polyrepresentation, which involves combining representations obtained in different ways from data of the same modality. Additionally, the thesis presents a comprehensive review of research papers using deep learning models for lung image analysis to identify COVID-19. This led to the development of a checklist for creating responsible deep learning solutions in healthcare.

Significant contributions to explainable artificial intelligence include a creation of interactive LIMEcraft that enhances explanation quality by allowing users to select semantically consistent areas and examine model predictions by manually perturbing image. The thesis also introduces two different explanation ensembling methods: NormEnsembleXAI and CNN-based explanation ensembling.

In conclusion, this thesis answers all research questions and achieves all objectives, providing significant advancements in the explainability and reliability of deep learning models in healthcare, making a shift towards safer and more trustworthy artificial intelligence systems.

Keywords: Deep Learning, Explainable Artificial Intelligence, medical imaging

Słowa kluczowe: uczenie głębokie, wyjaśnialna sztuczna inteligencja, obrazowanie medyczne

Wojciech

