

De Onderzoeksgroep
Artificial Intelligence Lab

nodigt U graag uit op de openbare verdediging van het proefschrift van

Isel del Carmen Grau Garcia

ter behaling van de graad van Doctor in de Wetenschappen

Titel van het proefschrift:

Self-labeling Grey-box Model: An Interpretable Semi-supervised Classifier

Promotor:

Prof. dr. Ann Nowé

Co-promotor:

Prof. dr. Maria Matilde Garcia Lorenzo

De verdediging heeft plaats op

Vrijdag 9 oktober 2020 om 17u

De verdediging kan via een livestream gevolgd worden. Een online link voor live-streaming wordt hier binnenkort toegevoegd:

<https://bit.ly/36t0fxD>

Samenstelling van de jury

Prof. dr. Beat Signer (VUB, voorzitter)
Prof. dr. Wim Vranken (VUB, secretaris)
Prof. dr. Tom Lenaerts (VUB)
Prof. dr. Sonia Van Dooren (VUB)
Prof. dr. Koen Vanhoof (UHasselt)
Prof. dr. Chris Cornelis (UGent)

Curriculum vitae

Isel Grau behaalde in 2014 haar master diploma Computerwetenschappen aan de Central University of Las Villas (Cuba) en startte haar doctoraat in 2015 aan de VUB. Haar onderzoek is gericht op interpreteerbare machine learning en semi-gesuperviseerde classificatie technieken met toepassingen in de bioinformatica. Ze is co-auteur van 14 peer gereviewde artikels in internationale conferenties en tijdschriften, en presenteerde haar werk op meerdere internationale events.

Abstract van het doctoraatsonderzoek

In de context van Machine Learning is het verkrijgen van data punten meestal vrij eenvoudig, het labelen van de data daarentegen kan een duur of tijdrovend proces zijn. In dergelijke gevallen kunnen we werken met data sets waarvan slechts een aantal datapunten gelabeld zijn. Semi-supervised classificatie, laat namelijk toe om met deels gelabelde data en deels ongelabelde data te werken, en resulteert in betere modellen t.o.v. wanneer alleen de gelabelde data zou worden gebruikt. Een nadeel van deze benadering is de complexiteit van de modellen, en het feit dat de classificatie van nieuwe datapunten niet transparant is voor een gebruiker. Maw. de modellen gedragen zich als zwarte dozen. Er is echter een toenemend aantal applicatiedomeinen waar de gebruiker een transparante beslissing verwacht. Begrijpbare classificatie (white-boxes) bieden hier een antwoord. Het zijn transparante structuren die voorspellingen maken en daarbij een uitleg. Desalniettemin zijn deze voordelen meestal ten koste van performantie in termen van accuraatheid.

In dit proefschrift introduceren we een zelf-labeling in grijze modellen, d.i. een semi-gesuperviseerde classifier die een goede balans bieden tussen accuraatheid en interpreteerbaarheid. De zelf-labelende grijze dozen gebruiken een zwarte doos classifier voor het labelen van ongelabelde data en een white-box surrogaat classifier garandeert de interpreteerbaarheid. Omdat het zelf-labelings proces fouten kan propageren stellen we twee uitbreidingen voor. Een gebaseerd op gecalibreerde probabiliteiten en een tweede op basis van Rough Sets Theory. Experimenten tonen aan dat onze aanpak een belangrijke meerwaarde biedt wanneer er beperkte data is en de ratio van ongelabelde data versus de gelabelde data hoog is en dit voor verschillende benchmark datasets. Daarnaast vergelijken we ook het effect van verschillende zwarte en witte doos classificatoren en de twee voorgestelde aangepaste methoden in functie van zowel accuraatheid als interpreteerbaarheid. De resultaten tonen aan dat de classifier interpreteerbaar is en superieure predicties maakt vergeleken met state-of-the-art self-labeling classifiers. Daarnaast illustreren we ook praktische de toepasbaarheid van de zelf-labelende grijze doos classifier aan de hand van initiële resultaten in twee semi-gesuperviseerde classificatie taken uit de bio-informatica. De eerste betreft de classificatie van pathogeniciteit van genomische varianten bij een zeldzame ziekte en de tweede betreft de voorspelling van eiwitvouwing.