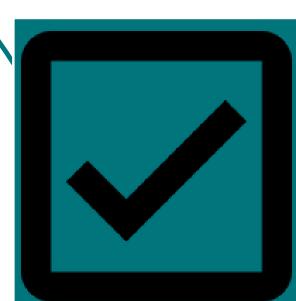
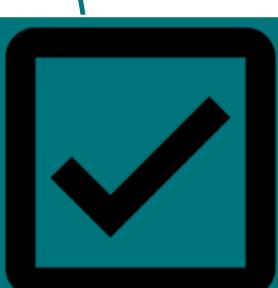
Classification of articles into the Sustainable Development Goals

Aims



Train a classification model to classify articles into 17 SDGs.



Improve the process of classification of the articles into the 169 targets.



Enhance the display of the classification model results for users of the SA SDG Hub website.

Approach

Visualisation

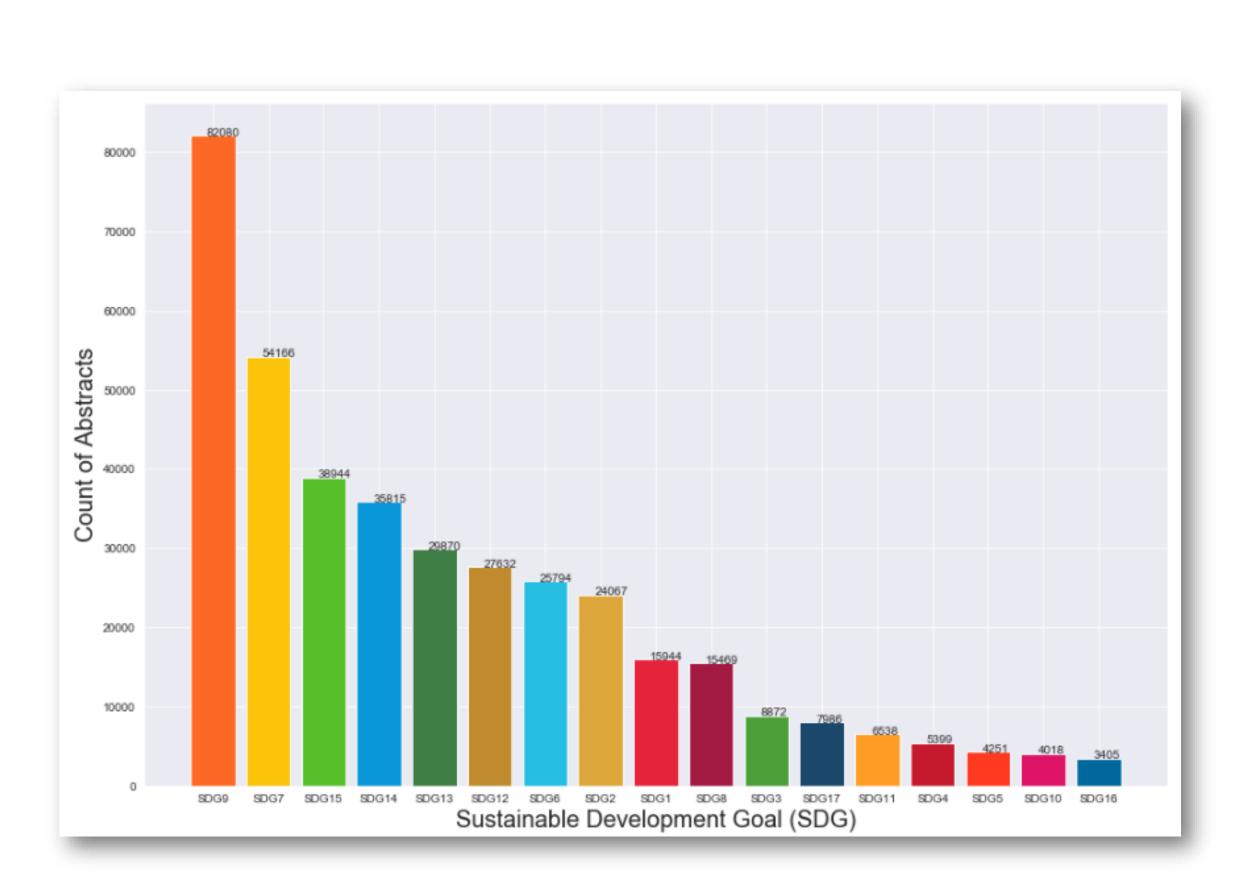
- > Explore data
- > NLP processes for data preparation
- > <u>Limitation:</u> Final classifier only trained on articles with less than two SDGs associated with them

Fitting the models

- > Test & train set split
- > Fit classifiers
- > Validate results
- > Perform predictability, computability & stability checks

Deployment

- > Deploy a Streamlit App for visualisation of classifier
- > Demonstrate how the classifiers work & value to be added from study



Performance metrics from the Classification Models				
Model Name	Accuracy	Precision	Recall	F1 Statistic
Linear Support Vector Classifier	68.56%	58.83%	56.38%	57.22%
Logistic Regression	70.09%	67.08%	55.50%	59.27%
Multinomial Naïve Bayes	66.04%	53.44%	42.52%	43.89%
Random Forest Classifier	35.39%	5.70%	5.89%	3.10%



Contact: Team Rivonia Rodeo

Cindy Hayward u13060334@tuks.co.za

Frederik Pretorius ffcpretorius@gmail.com



Faculty of Engineering, Built Environment and Information Technology

Department of Computer Science

Fakulteit Ingenieurswese, Bou-omgewing en Inligtingtegnologie / Lefapha la Boetšenere, Tikologo ya Kago le Theknolotši ya Tshedimošo Capstone Project - MIT 808

Course Coordinators: Dr. Vukosi Marivate (vukosi.marivate@cs.up.ac.za) Abiodun Modupe (abiodun.modupe@cs.up.ac.za)

