Towards Standardization of Geospatial AI and Geospatial Foundation Model

Piotr Gramacki, Piotr Szymański
Wrocław University of Science and Technology
Department of Artificial Intelligence
Kraina.AI Research Group

Introduction

OpenStreetMap is an open-source map service with detailed metadata. It can be used to download vector data with geospatial information for any place on the planet. We have created an **open source library (srai)** which makes **accessing data** easier and integrates **representation learning models** for geospatial domain.

Motivation

We believe that geospatial domain needs standardization in the form of:

- benchmark datasets and tasks
- unified pipeline for GeoAI tasks
- reproducible and extendable experimental code
- *HuggingFace-like* hub for models and data

The SRAI library

With the *srai* library we make it possible to compose GeoAI pipeline from unified components:

1. Loaders - accessing various data from OSM and beyond
2. Regionalizers - various approaches to space discretization
3. Embedders - unified format for embedding models

Transfer learning

With our research, we show that thanks to wide coverage of OSM data it is possible to build models transferable between distant parts of the world.

Foundation model

When trained at scale, on large samples from OSM we want to build a foundation model for GeoAI based on vector data. Such model could be fine-tuned for different tasks and used on different parts of the world. We're looking for collaborators interested in training such a model.

Wrocław University of Science and Technology

Vector data from **OpenStreetMap** is great for **geospatial tasks** and could be used to train a **Geospatial Foundation Model**

Take a picture to download the full paper

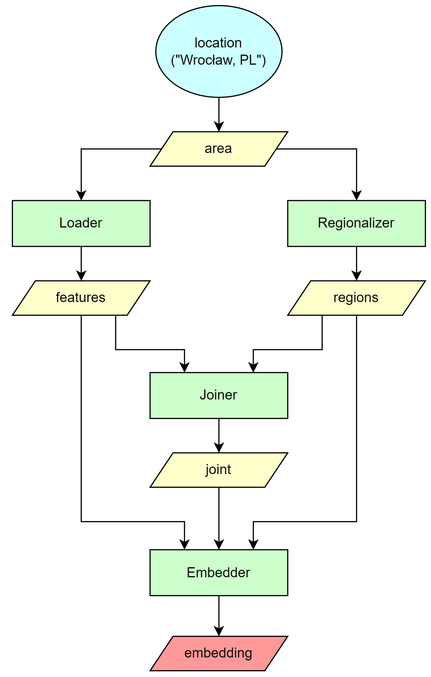

Table 1: Geospatial toolboxes and their capabilities.

Library	Spatial files	OSM	Trajectories	GTFS	Raster	Visualization	Regionalization	Geocoding	ML	Datasets
geowrangler[1]	✓	✓			✓		✓		✓	
teapy[2]		✓				✓	✓	✓		
geomancer[3]		✓								
Mosaic[4]	✓				✓		✓		✓	
PySal[30]					✓	✓	✓		✓	
Verde[36]	✓						✓		✓	✓
WhiteboxTools[6]	✓				✓				✓	
Pandana[5]	✓	✓								
MovingPandas[12]			✓							
Scikit-mobility[23]			✓							
segment-geospatial[41]	✓				✓				✓	
TorchGeo[33]	✓				✓				✓	✓
srai	✓	✓		✓	✓*	✓	✓	✓	✓	

[1] https://github.com/thinkingmachines/geowrangler, [2] https://github.com/scravath-oda/teapy, [3] https://github.com/thinkingmachines/geomancer, [4] https://github.com/databricks/mosaic, [5] https://github.com/UDST/pandana, [6] https://github.com/jblindsay/whitebox-tools, * only for data downloading and preparation