Land Cover Classification of Fused Optical Radar Data Using Three Machine Learning Algorithms

David Junor

Student ID: 501074017

January 28, 2022

Remote sensing is the method of acquiring information about the physical characteristics of the earth’s surface from a distance, such as from satellites or aircraft, by measuring its reflected and emitted energy (NASA, 2021). Remote sensors can be either passive or active (NOAA, 2021). Passive sensors respond to external stimuli and record natural energy that is reflected or emitted from the Earth’s surface (NOAA, 2021). The most common source of reflected energy detected by passive sensors is sunlight (NOAA, 2021). Active sensors emit a stimulus to collect data about the Earth. For example, LiDAR emits a laser-beam onto the Earth’s surface and measures the time it takes for the beam to reflect back to the sensor (NOAA, 2021). Remote sensors provide an abundance of data about the earth’s systems which facilitates data-informed decisions based on the current and future state of the globe (NASA, 2021). Remotely sensed information is used in military, intelligence, commercial, economic, planning and humanitarian applications, among others (“Remote sensing”, 2022).

Remote sensing and data science communities have begun to merge due to several factors (Abdi, 2020). First, competitions held by companies such as Kaggle have demonstrated that high classification accuracies are possible using advanced machine learning algorithms (Abdi, 2020). Second, there is now an abundance of user-friendly open-source programming tools (Abdi, 2020). Third, a reduction in the cost of high-end computing power (Abdi, 2020). Fourth, and last, the availability of free remotely sensed data such as satellite data (Abdi, 2020).

One of the core applications of remotely sensed data is in land cover and land use classification. Land cover classification is the classification of the observed biophysical cover of the earth’s surface (NASA, 2021). Examples of land cover classes found on the earth’s surface include vegetation, urban infrastructure, water, rocks, and bare soil (Government of Canada, 2015). Land use classification is the classification of the purpose the land serves such as recreation, agriculture, or wildlife habitat (Government of Canada, 2015). Mapping land use/land cover (LULC) is important for global environmental monitoring studies, natural resources management, and landscape planning activities such as watershed management to name a few applications (Government of Canada, 2015).

The accuracy of LULC maps is dependent on several factors such as the training data used, the implementation of the classification, the classification algorithm, and the environmental heterogeneity of the area being classified (Hermosilla et al., 2022). Of these factors there is current interest in the application of advanced machine learning algorithms to LULC classification.

The research objectives of this project are to assess the performance of the K-Nearest Neighbours, Random Forest, and Support Vector Machine machine learning algorithms in the multi-class land cover classification of a fused optical-radar dataset (Kosravi, 2020). Using Python Pandas, Numpy, and Scikit-learn libraries each machine learning algorithm experiment will be applied to the same data set. Each model will be evaluated, and the machine learning algorithms compared to determine which one is best for our use case of classifying cropland from a fused optical-radar dataset for a relatively flat and environmentally homogenous area around Winnipeg, Manitoba.

References

Abdulhakim Mohamed Abdi. (2020). Land cover and land use classification performance of machine learning algorithms in a boreal landscape using Sentinel-2 data*.* *GIScience & Remote Sensing, 57(1), 1-20.* <https://doi.org/10.1080/15481603.2019.1650447>

Government of Canada. (2015, 11 20). *Land Cover and Land Use.* Land Cover/Biomass Mapping. <https://www.nrcan.gc.ca/maps-tools-and-publications/satellite-imagery-and-air-photos/tutorial-fundamentals-remote-sensing/educational-resources-applications/land-cover-biomass-mapping/land-cover-land-use/9373>

Hermosilla, T., Wulder, M. A., White, J. C., Coops, N. C. (2022). Land cover classification in an era of big and open data: Optimizing localized implementation and training data selection to improve mapping outcomes. *Remote Sensing of Environment. 268:112780*. <https://doi.org/10.1016/j.rse.2021.112780>

Kosravi, I. (2020). *Crop mapping using fused optical-radar* [Data set]. UCI Machine Learning Repository. [https://archive.ics.uci.edu/ml/datasets/Crop+mapping+using+fused+optical-radar+data+set#](https://archive.ics.uci.edu/ml/datasets/Crop+mapping+using+fused+optical-radar+data+set)

NASA. (2021, 12 6). *What is Remote Sensing?* Earth Data Open Access for Open Science. <https://earthdata.nasa.gov/learn/backgrounders/remote-sensing>

NOAA. (2021, 2 26). *What is remote sensing?* Ocean Facts. <https://oceanservice.noaa.gov/facts/remotesensing.html>

Remote sensing. (2022, January 25). In *Wikipedia*. <https://en.wikipedia.org/wiki/Remote_sensing>