Deforestation Risk Assessment

Submitted By: Sushil Jangra
Submitted to: Dr. Sujit Ghosh

1. Introduction & Problem Statement

Examining deforestation in a spatial context aids in pinpointing pivotal factors that affect deforestation risks and gauging their relative impacts. Foreseeing alterations in forest coverage is imperative as it allows for anticipation of deforestation consequences, including carbon emissions or biodiversity loss.

Our objective is to create deforestation risk maps for a specified area (Para, Brazil). The primary goal of this endeavor is to generate maps indicating the likelihood of deforestation based on various factors such as population density, road networks, land usage, and topographical features.

2. Data Description

We have been provided with a file containing four classes representing forest cover transitions from 2015 to 2021. The value 1 indicates areas deforested during the 2015-2018 period, while the value 2 represents areas deforested during the 2018-2021 period. Value 3 signifies stable forest cover throughout the 2015-2021 period, and value 0 denotes stable non-forest cover during the same period. Our objective is to extract relevant features for predicting the risk of future deforestation. To achieve this, data from Google Earth Engine and DIVA-GIS for the specified region of interest has been collected, including:

- Land Use
- Population Density
- Elevation
- Distance from the nearest Road

3. Methodology

Our approach involves the following key steps:

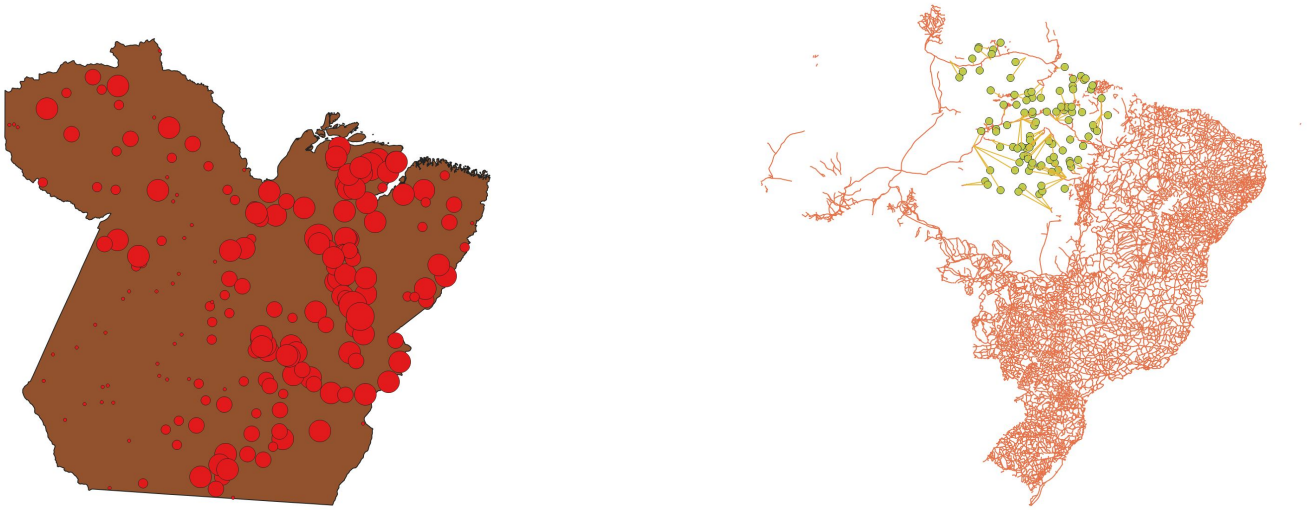
- **Data Preprocessing**: All these datasets have been unified into the same coordinate system, namely EPSG 32722.
- **Data Sampling**: We commenced by collecting a balanced dataset consisting of 1000 samples, equally divided into 250 sample points belonging to each of the four classes i.e., 0, 1, 2, 3.
- **Feature Engineering**: Values corresponding to each sampled point were extracted for every feature. Subsequently, these values underwent standardization and encoding for consistency while model training.
- **Model Training**: The above described features were used to train various ML models. Before that class was assigned to each of the data point based upon deforested region or stable region. Various models like Random Forest, AdaBoost etc. were tried but Artificial Neural Network (ANN) provided the best result.

Serial No.	Feature Name	Description
1	Elevation	Elevation at sampled point
2	Population Density	Population Density at sampled point
3	Land Use	Type of land use at sampled point
4	Distance from Road	Distance between sampled point and nearest road

Table 1: List of Features and Descriptions

4. Result and Discussion

To assess the risk of deforestation, we collected new data points from the region of interest. By applying the trained model, we derived the probability of deforestation for each point, and visualized the results as follows: The size of the dots on the plot corresponds to the level of deforestation risk, with larger dots indicating a higher probability of deforestation.



(a) Risk of Deforestation

(b) showing nearest road from the sampled points

Figure 1: Images a. showing Deforestaion risk in the Para (Brazil region) and b. road network in Brazil and few sampled points in the Para region