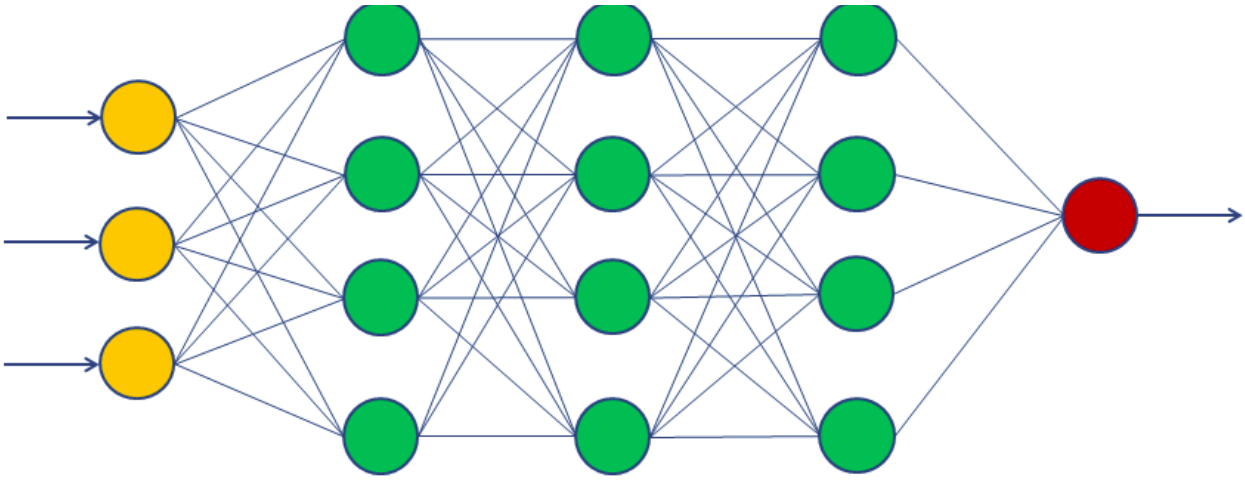


Code Overview and Models

Feature Engineering

Landsat7 data for the Indian shape was obtained from Google Earth Engine. The image downloaded was at 500m resolution and was the average of images from the start of a year to its end. Necessary procedures were followed to reduce the cloud cover of the image. The data consisted of 9 primary band and we added three derived bands to the data set.

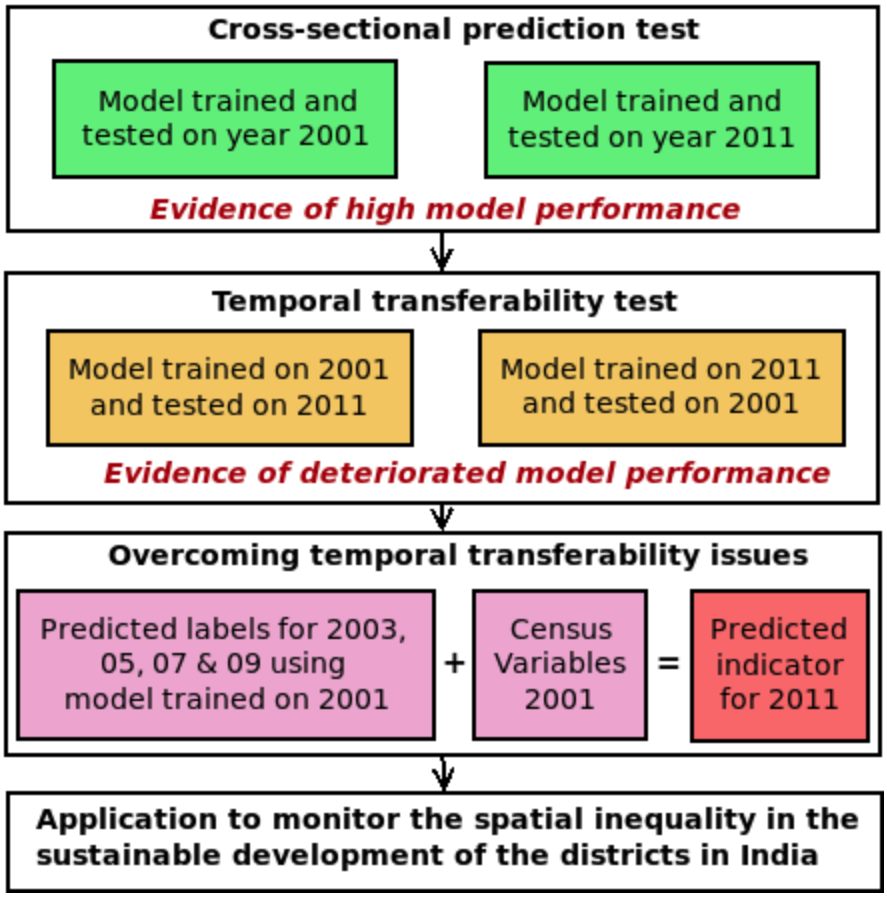
Band	Type	Resolution
B1	Blue	30m
B2	Green	30m
B3	Red	30m
B4	Near Infrared	30m
B5	Shortwave Infrared 1	30m
B6_VCID_1	Low-gain Thermal Infrared	30m
B6_VCID_2	High-gain Thermal Infrared	30m
B7	Shortwave Infrared 2	30m
B8	Panchromatic	15m
(B4-B3)/(B4+B3)	Normalized Difference Vegetation Index (derived)	30m
(B2-B5)/(B2+B5)	Modified Normalized Difference Water Index (derived)	30m
(B5-B4)/(B5+B4)	Normalized Difference Built Index (derived)	30m

Table 1: Landsat 7 bands

The data obtained from the GEE was continuous in nature as it contained the values of individual pixels for all 12 bands. **Binning** was applied to the data set to perform **discretization** of the database. **Ten bins** were chosen for each of the twelve bands. **Quantile binning** strategy was used to ensure each bin gets an equal number of data points. After obtaining the bin edges the features were made by the **ordinal encoding** of each district tiff file. The ground truths were obtained from the census data of 2011 and 2019. The labels present in the ground truth were asset ownership, bathroom facility, fuel for cooking, condition of household, main source of light, main source of water, literacy, and formal employment.

Model Training and Predictions.

Various training and testing iterations were done to identify the temporal transferability of the satellite data. The complete gist of models is depicted in the picture given below.



Cross-Sectional Models:

These models were trained and tested in the same year. This was done for both years 2001 and 2011, The accuracies were estimated using **Weighted F1 Score** and by performing a **training-testing split** on the database of the same year. The algorithms used were **SVCs, Random Forrest, and XGBoost.** XGBoost had the highest accuracies among all the models. And the overall accuracies for these models were decent enough to indicate that satellite data generalize well over the same time period.

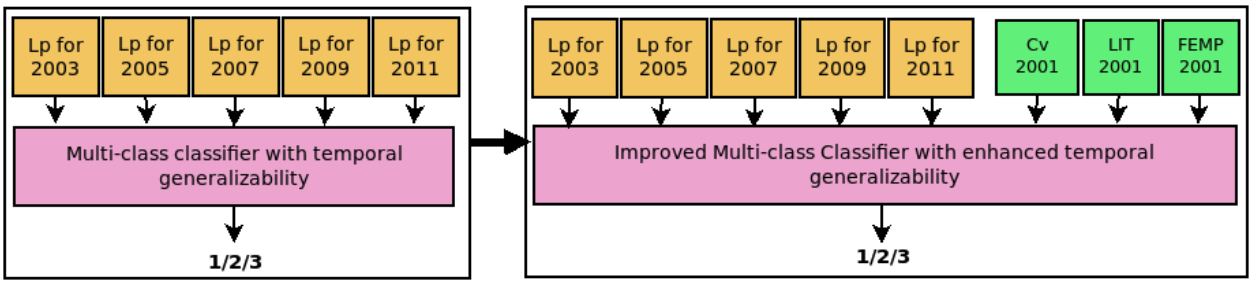
Single Step Temporal Transferability :

Now, to check the temporal transferability of the satellite data we trained models on the year 2001 and tested them on the year 2011 (Direct Forward Classification) and vise-versa. Since XGBoost was used for this purpose because it provided the

best accuracies in Cross-Sectional classification. The accuracies obtained were a bit low indicating the deteriorated performance of the model.

Two Step Temporal Transferability :

To improve the model for temporal transferability a technique similar to **Bagging Ensemble Classifier** was applied. This method applies the basis direct model to predict labels for several intermediate years between the base year (whose ground truth is available) and the target year (whose labels are finally to be predicted) and uses these as features for a second classifier to make the final prediction. We call this second classifier the **forward classifier**. This model relies on the technique of **ensembling weak learners to make strong learners**.



This improved two-step classification is further complemented by including some census variables of the base year as features in the forward classifier, which are likely to affect the socio-economic change that would be taking place in a district.