

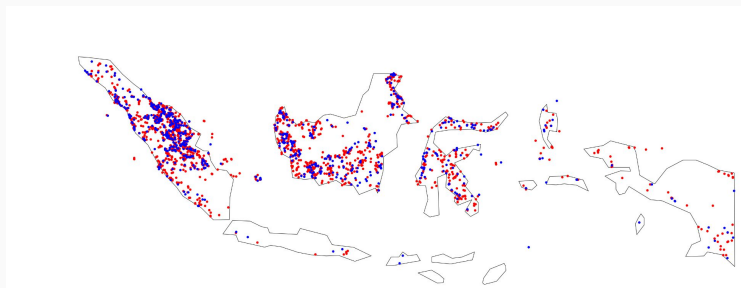
Zero Deforestation Challenge

Solution by Team “conquistador”



Data preparation

- A 90% of the data was reserved randomly to train split and 10% for validation split for model validation. **Why: It was observed that the data was distributed similarly across Indonesia (train=red, test=blue).**
- Due to time restrictions, only images were used. And only the data preprocessing that was better suited for the pre-trained models was used (subtract the mean of Imagenet dataset and divide by its standard deviation.)
- External data was allowed. Therefore, a dataset of the same domain ([ForestNet](#)) was used to increase the data. **Why: This was done to increase the model performance, given the limited samples available for the challenge.**



Model training

- An EfficientNet B0 pretrained on Imagenet was used. **Why: It was used because of its relatively high throughput and good results on a general purpose benchmark (Imagenet).** This model already suffered of overfitting, so bigger models could have increased it.
- The model was validated on the train/val split proposed on the previous slide. Results can be observed below.
- Once validated, train and validation splits were concatenated, and a final model (with the same architecture) was trained.

