**OVERVIEW**

* I participated in the Hack The Carbon hackathon, sponsored by InstaDeep at Deep Learning Indaba (DLI) 2025 in Kigali, Rwanda — Africa’s largest annual AI and Machine Learning event. In 2025, over 1,200 carefully selected students, academics, and researchers from across the continent gathered to exchange ideas, attend workshops, and engage in competitions like this hackathon.
* Out of 260 entrants, 87 participants, 54 teams, and 435 submissions, ***I achieved 1st place,*** demonstrating strong innovation and technical skill in developing machine learning solutions for forest biomass estimation across Africa.

**MODEL ARCHITECTURE & TRAINING SUMMARY**

* **Model:** U-Net with an EfficientNet-B3 encoder backbone, pre-trained on ImageNet.
* **Input Data:** The given training test and validation time-series satellite chips (3 timesteps × 6 bands = 18 features/channels) were used to engineer 4 vegetation indices per timestep (NDVI, NDWI, EVI, NBR) per timestep, expanding inputs to 30 channels. Thie enriched the spectral information and improved the model performance
* **Data Splitting:** A **random split** was employed instead of a geographical/temporal split. This ensures the model learns from a well-mixed distribution of features across the entire dataset but may overestimate real-world generalization performance.
* **Training:** The model was trained for 15 epochs, with performance tracked using R², RMSE, and MSE.
* **Results:** The training process showed strong and consistent convergence. The error metrics decreased sharply and began to plateau around epoch **15**, achieving a robust **R² score of ~0.81**. This indicates the model explains a very high proportion of the variance in the target variable and has learned the training distribution effectively.

**Key Takeaways:**

* The combination of a powerful U-Net architecture with a pre-trained EfficientNet backbone proved highly effective.
* The custom vegetation indices were successfully integrated and contributed to the model's strong performance.
* The rapid convergence and high R² score suggest the model is well-fit to the data.

**Next Steps:**

* Evaluate on a held-out test set with a **geographical or temporal split** to better assess real-world generalization.
* Experiment with learning rate schedules or slight architectural tweaks to potentially push performance beyond the plateau.
* Perform error analysis to understand what types of samples the model still struggles with.