- Kaggle
- Img 2 Biomass Contest

--- background & dataset: img of pasture taken by phone->estimation of the amount of green and withered grass

Model1： Img ->feature ->XGBoost

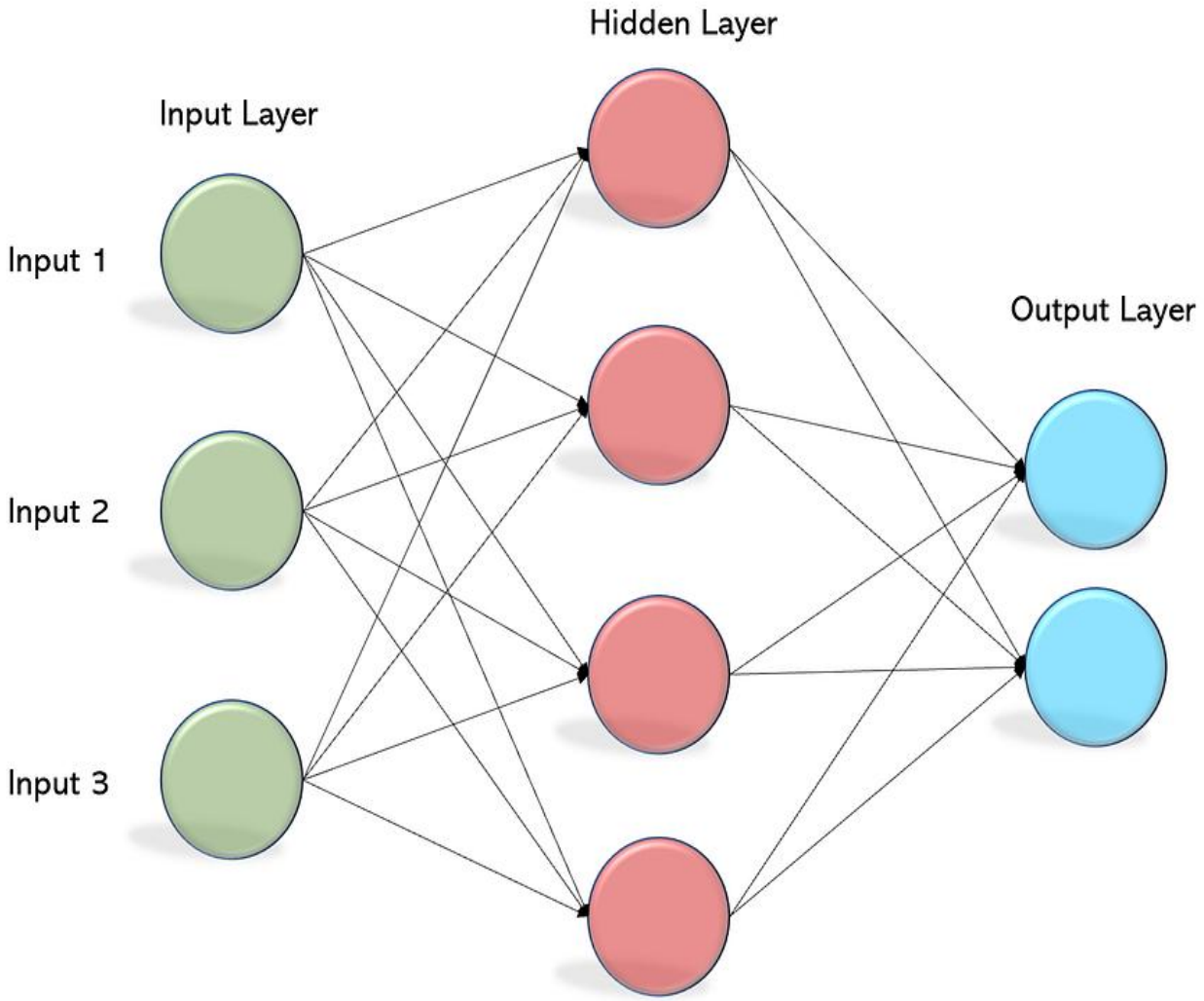
XGBoost is a decision-tree algorithm and a frequent winner in data science competitions like Kaggle.

Think of it as a "team of top students." The first student solves a problem. The second student focuses on fixing the first one's mistakes. The third student corrects the errors made by the previous two, and so on. Finally, the team combines their insights to produce a highly accurate answer.

Model2： Img ->feature ->MLP

A multilayer perceptron (MLP) neural network model was employed in this study. It is one of the simplest forms of an ANN.

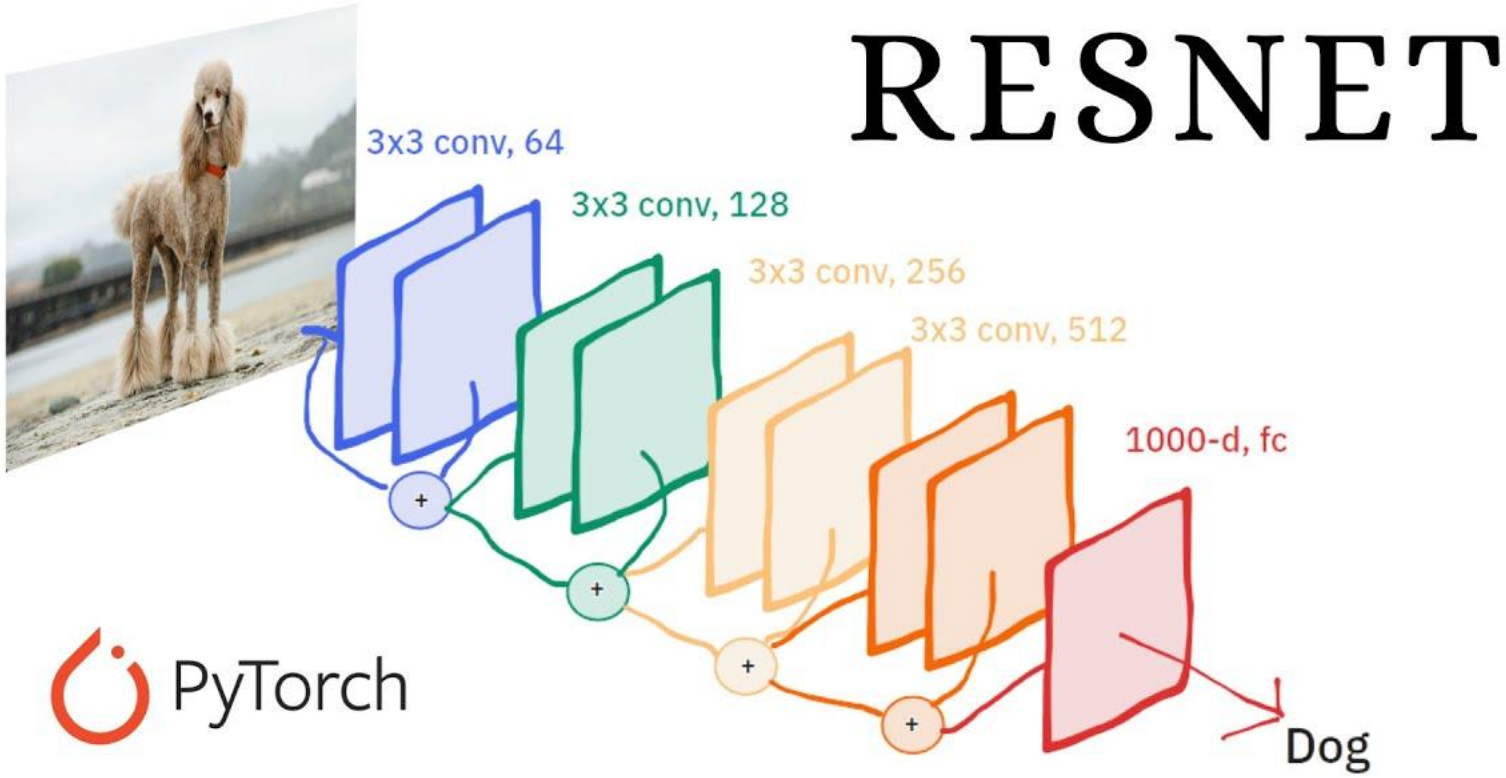
- two hidden layers (64 nodes each) and one output
- activation function - rectified linear unit and the
- optimisation algorithm - Adam
- learning rate of 0.001
- 5 epochs.



source: https://becominghuman.ai/multi-layer-perceptron-mlp-models-on-real-world-banking-data-f6dd3d7e998f

Model3： Img ->ResNet

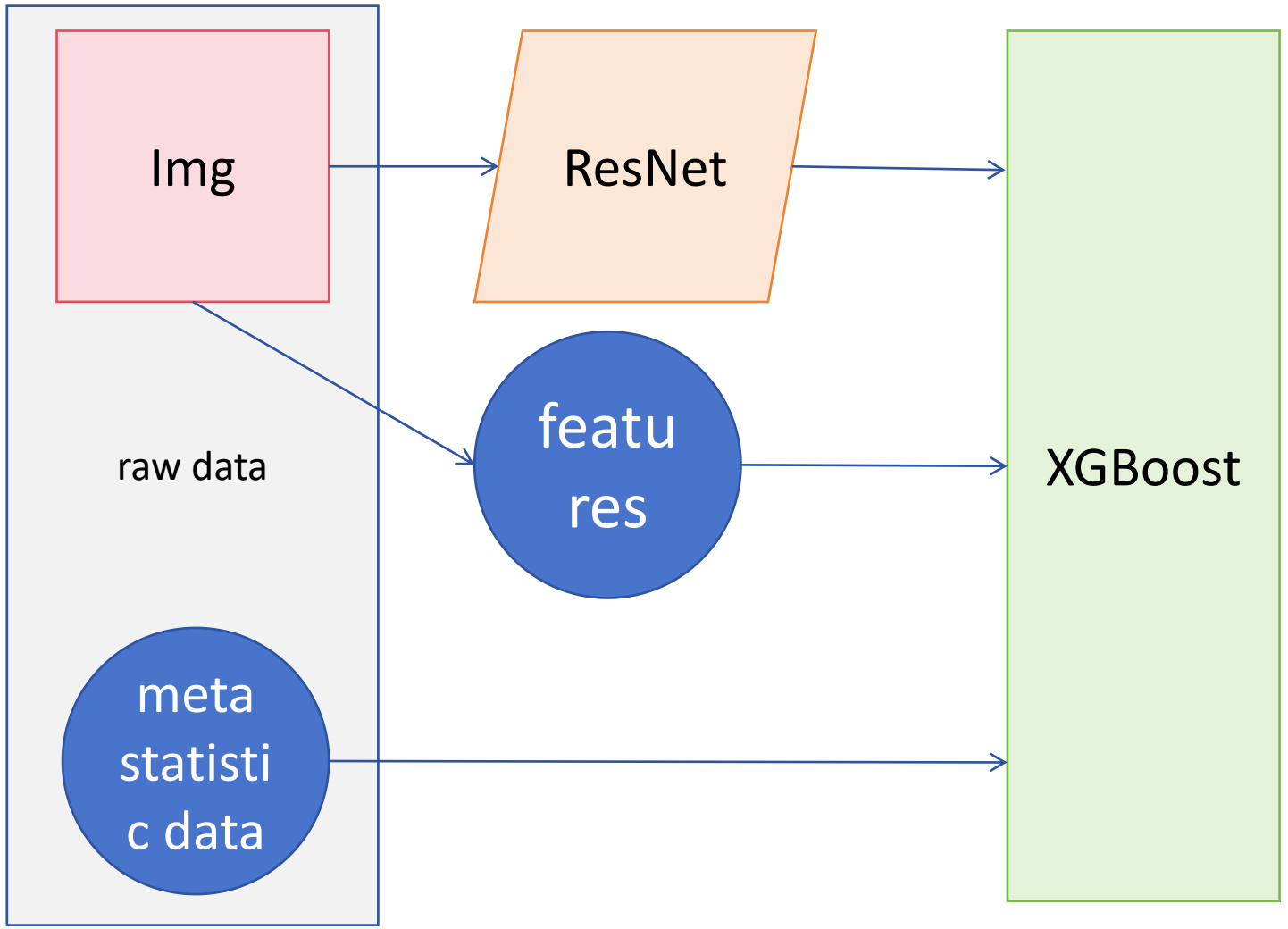
Images are processed by ResNet, where the output of the final convolutional layer serves as an embedding to model pasture biomass.



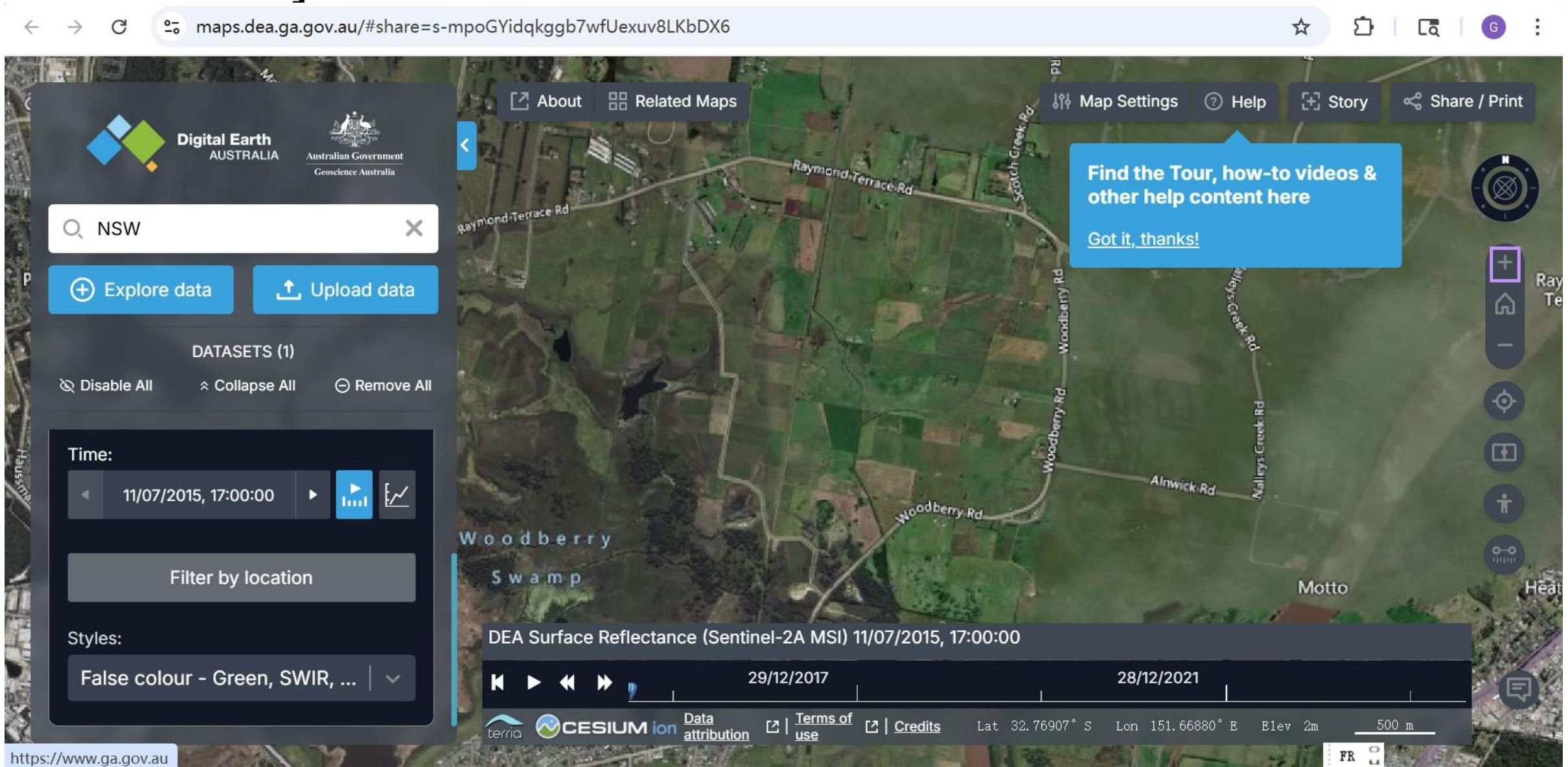
source: https://www.youtube.com/watch?v=woEs7UCaITo

Model4： Fusion model

Handcrafted features, metadata, and ResNet embeddings (extracted from the final conv layer) are concatenated and input into XGBoost.

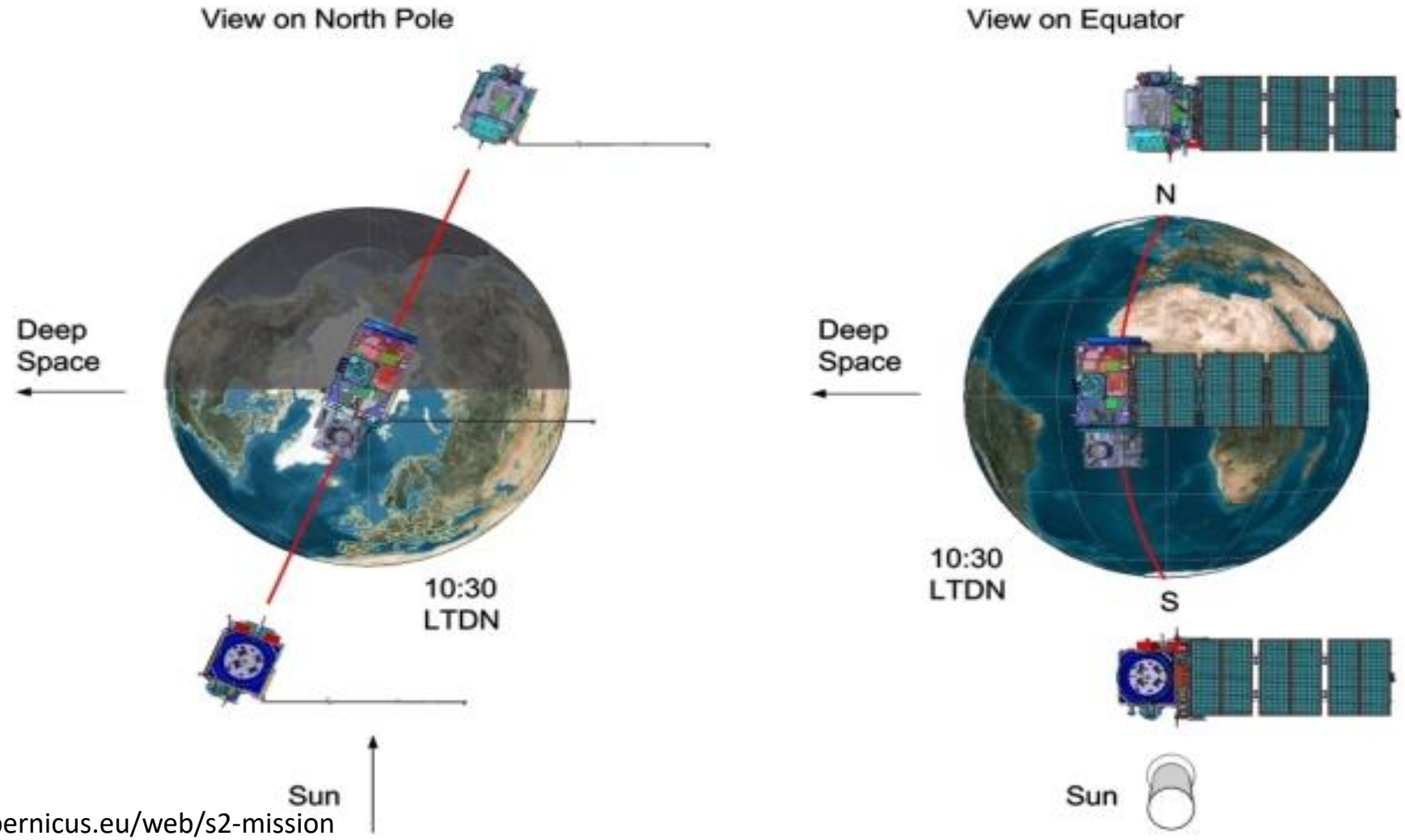


Model5： [Img ->feature + Remote Sensor->feature]->MLP



Model5：[Img ->feature + Remote Sensor->feature]->MLP

◆ Remote sensor:How it works

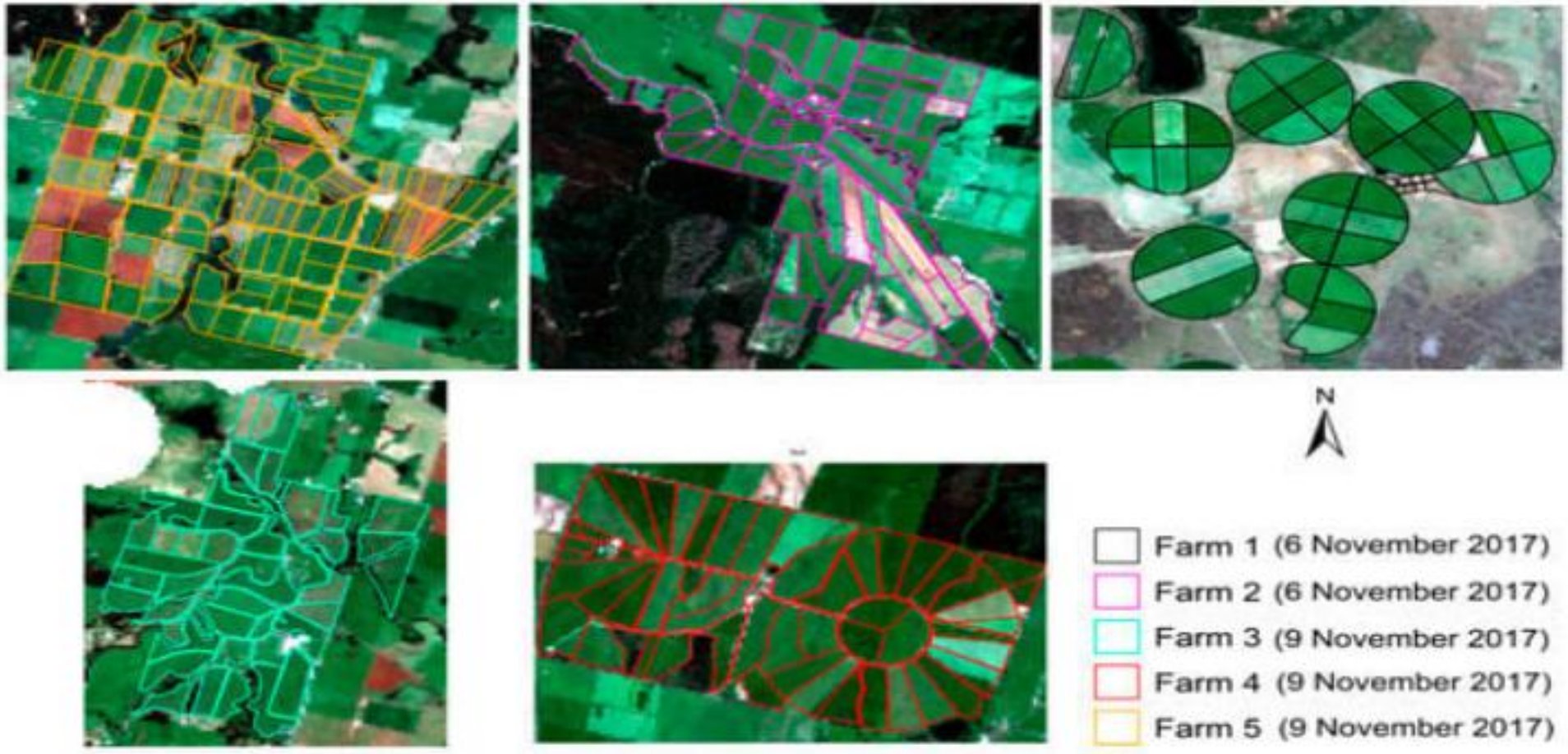


source: https://sentiwiki.copernicus.eu/web/s2-mission

Model5： [Img ->feature + Remote Sensor->feature]->MLP

◆ Data: How to use

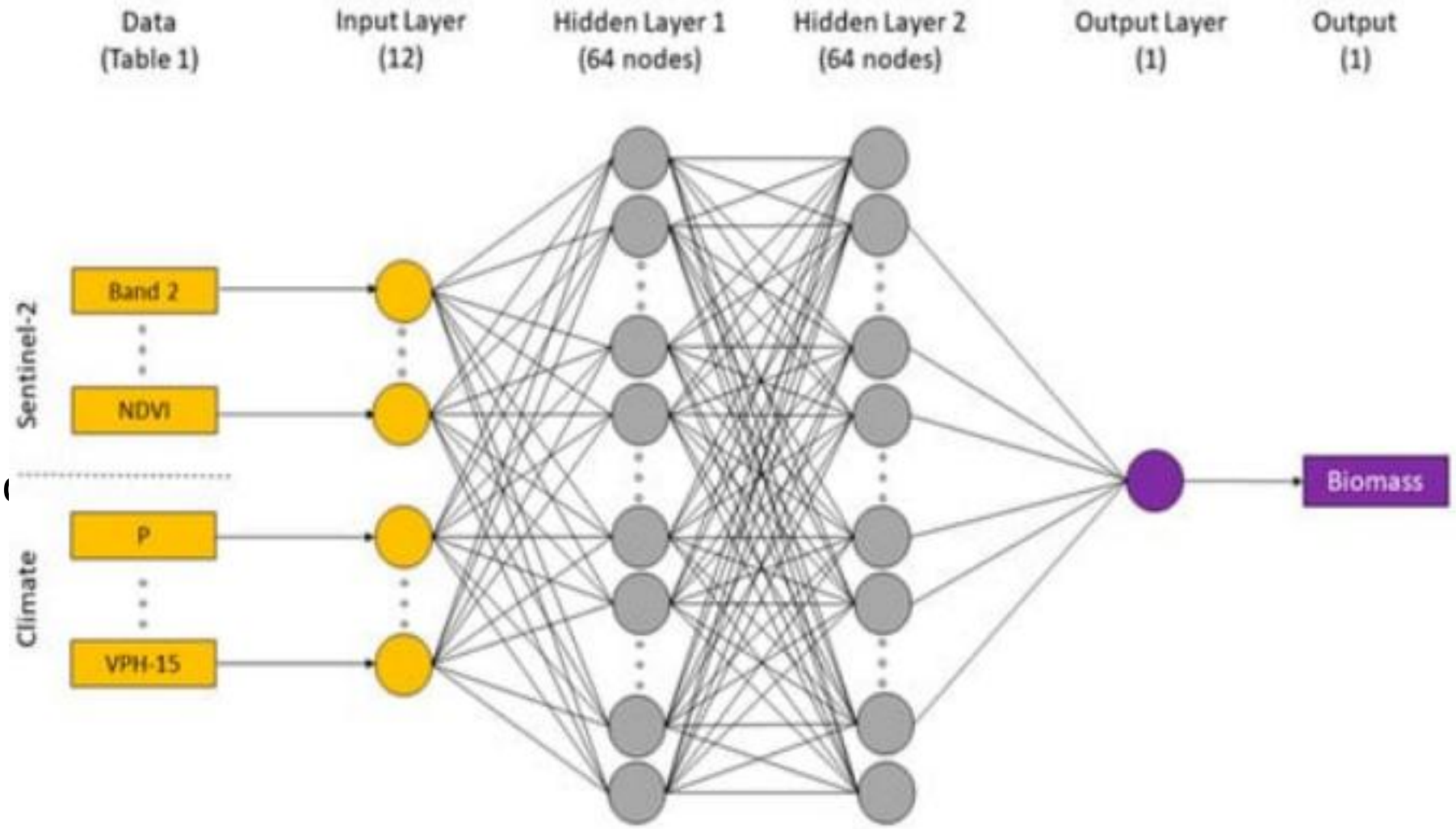
coordinate (shape)
time
radio range
resolution
satellite
cloud cover



source: https://www.mdpi.com/2072-4292/13/4/603#Appendix_A_Cloud_Coverage_Analysis_of_the_Sentinel2_Imagery

Model5：[Img ->feature + Remote Sensor->feature]->MLP

For each sample, the median value of all Sentinel‑2 bands is combined with the engineered features and metadata, then passed into an MLP.



source: https://www.mdpi.com/2072-4292/13/4/603#Appendix_A_Cloud_Coverage_Analysis_of_the_Sentinel2_Imagery

Result

Model Input + Algorithm	Target: GDM_g (R2) (Train / Val)	Target: Dry_Total_g (R2) (Train / Val)	Training time (sec)
Images + ResNet	-0.4491 / -1.2452	-0.6701 / -1.2970	267.15 (Slowest)
Hand-crafted + Meta + XGBoost	0.9997 / **0.7140** 👑	0.9998 / **0.6618** 👑	251.00
Hand-crafted + Meta + MLP	0.5519 / 0.3956	0.4625 / 0.1569	**112.91** (Fastest)
Sentinel-2 + Features + MLP	0.3524 / 0.5116	0.3898 / 0.2550	126.72
Fusion (ResNet + XGBoost)	0.9511 / 0.6518	0.9999 / 0.5423	182.58

Result

- Best Model (Model 2): By achieving the highest validation R^2 on both GDM_g and Dry_Total_g, Model 2 demonstrates that Structured Features paired with XGBoost is the optimal approach for this data.

- Value of Satellite Data: Comparing the MLP models, adding Sentinel-2 data (Model 4) raised the GDM_g score from 0.3956 to 0.5116 over the baseline (Model 3), validating the utility of satellite inputs.

- Image Data Issues: Model 1 (ResNet) failed to converge, resulting in negative scores. Consequently, the fusion model (Model 5) underperformed Model 2, as the image data effectively acted as noise.

Further Work: other remote sensor data, network(DCN)