

## intro[Objective]

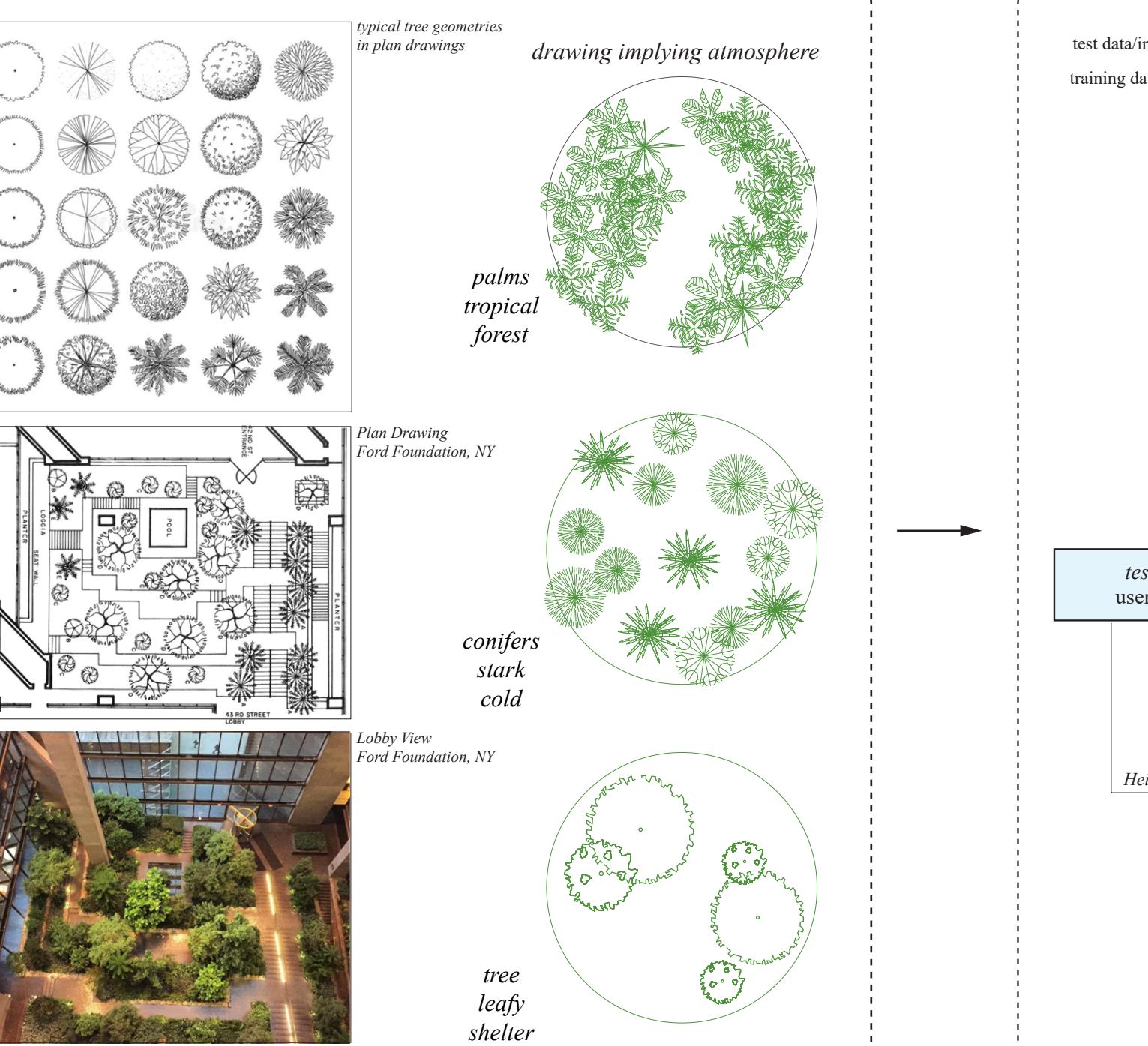
### Suggestive Trees

A tree species recommendation system based on vector geometry and taxonomical characteristics using Lunchbox neural networks

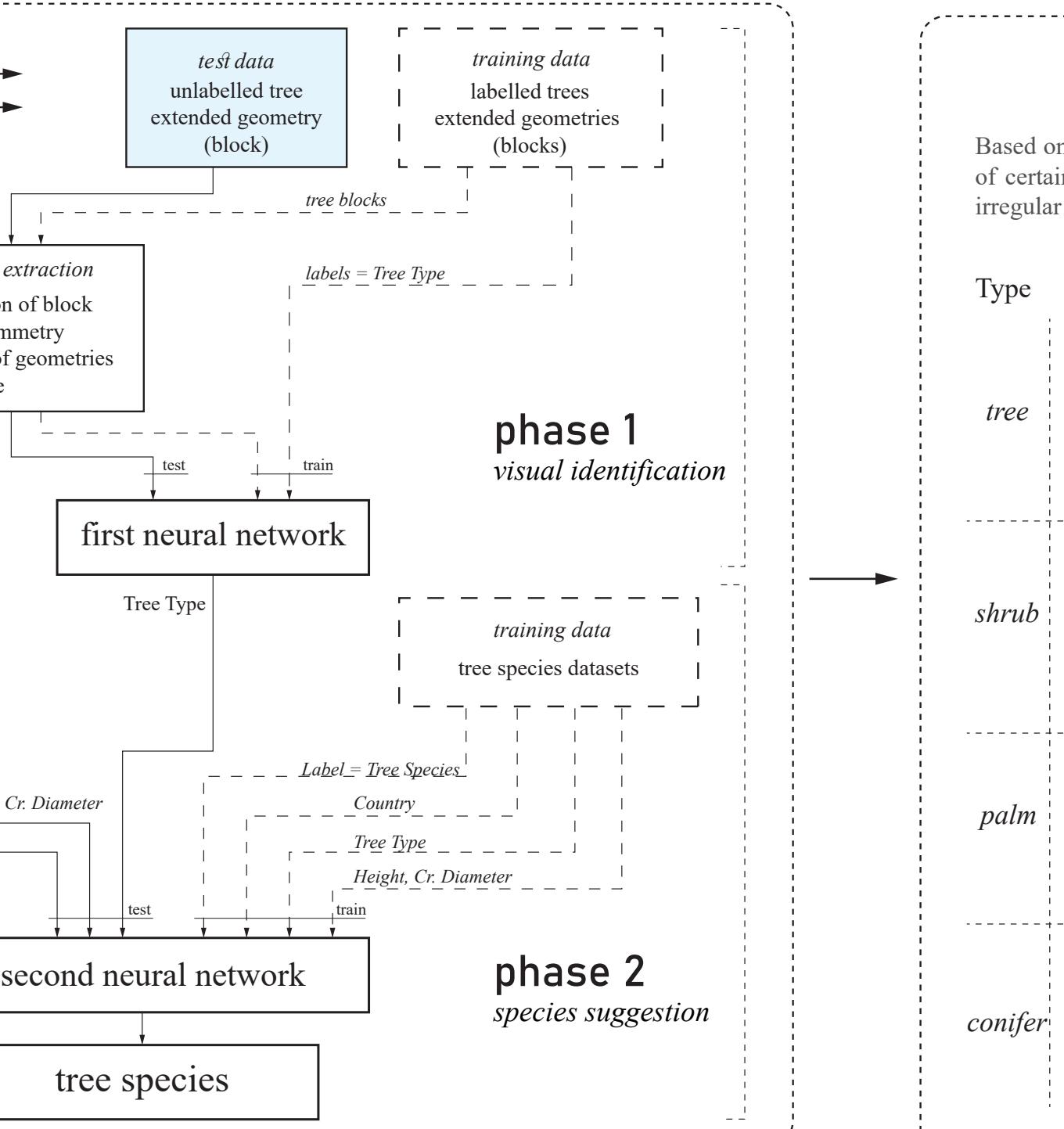
Apart from polylines that denote the built environment of a project, architectural CAD drawings contain several non-standard representations of objects, that a human being is usually able to interpret by their top view, such as furniture, building components like elevators, and landscape features.

Out of these, landscape features such as trees are of special interest as non-standard representations. The features of these vector geometries, usually propagated in libraries over the internet and used in the form of CAD blocks, or extended geometries, can be interpreted as humans to understand the ‘green’ nature of a space, by certain visual cues. These 2d representations of trees can potentially provide information on the diameter of the tree crown, density of greenery, ‘bushiness’, or other features that communicate the atmosphere of the space to the reader of the plan.

This model was built for the purpose of perceiving input vector geometries, 2d representation of trees, and using user-input and implicit data to suggest a tree species to the user.



## intro[Process Flowchart]



### phase 1 visual identification

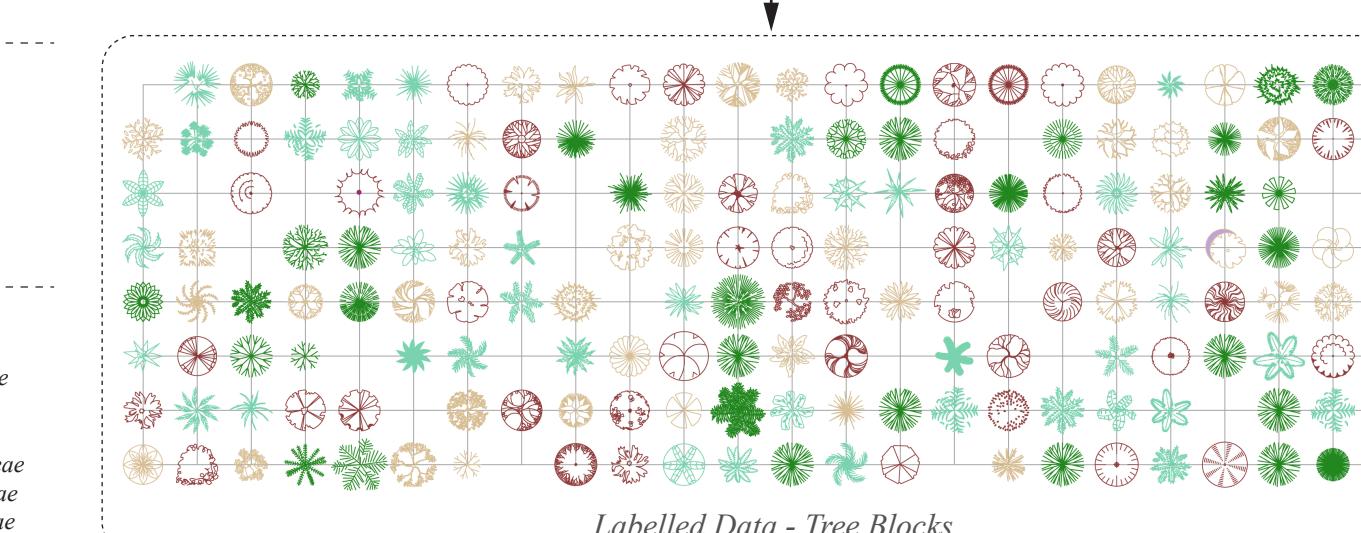
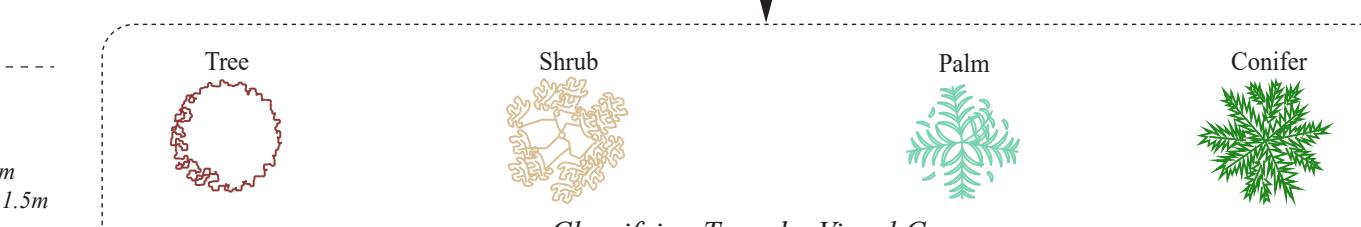
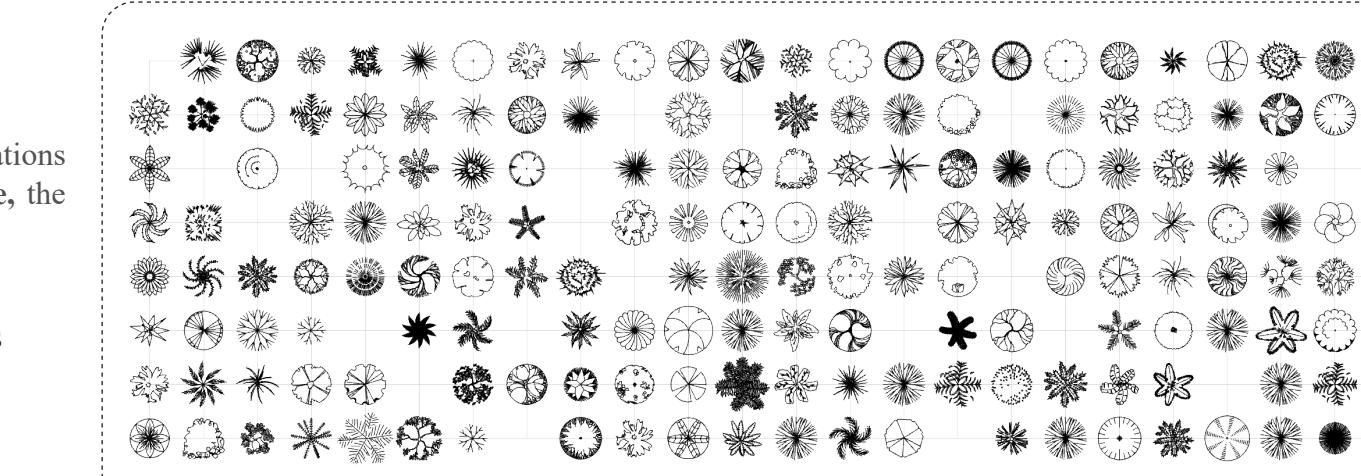
### phase 2 species suggestion

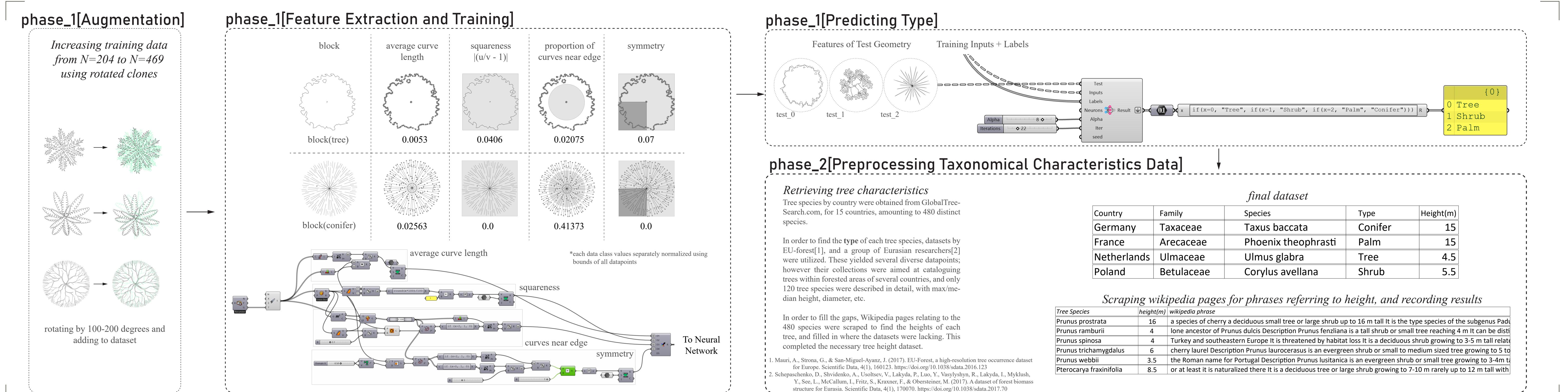
## phase\_1[Preprocessing Vector Graphics]

### Labelling Training Data

Based on the visual cues of the dataset, as well as typical representations of certain species, 4 distinct types can be identified: the leafy tree, the irregular shrub, the conifer, and palm.

Type	Top View	Example	Descriptors
tree			* base scenario * trunk thicker than 12 cm * crown base higher than 1.5m * height higher than 3m
shrub			* trunk thinner than 12 cm * crown base lower than 1.5m * height lower than 3m
palm			* belonging to family: Arecaceae
conifer			* belonging to families: Pinaceae, Podocarpaceae, Cupressaceae, Araucariaceae, Cephalotaxaceae, Phyllotladaceae, Sciadopityaceae





1. Mauri, A., Strona, G., & San-Miguel-Ayanz, J. (2017). EU-Forest, a high-resolution tree occurrence dataset for Europe. *Scientific Data*, 4(1), 160123. <https://doi.org/10.1038/sdata.2016.123>
2. Schepaschenko, D., Shvidenko, A., Usołtsev, V., Lakyda, P., Luo, Y., Vasylshyn, R., Lakyda, I., Myklush, Y., See, L., McCallum, I., Fritz, S., Kraxner, F., & Obersteiner, M. (2017). A dataset of forest biomass structure for Eurasia. *Scientific Data*, 4(1), 170070. <https://doi.org/10.1038/sdata.2017.70>

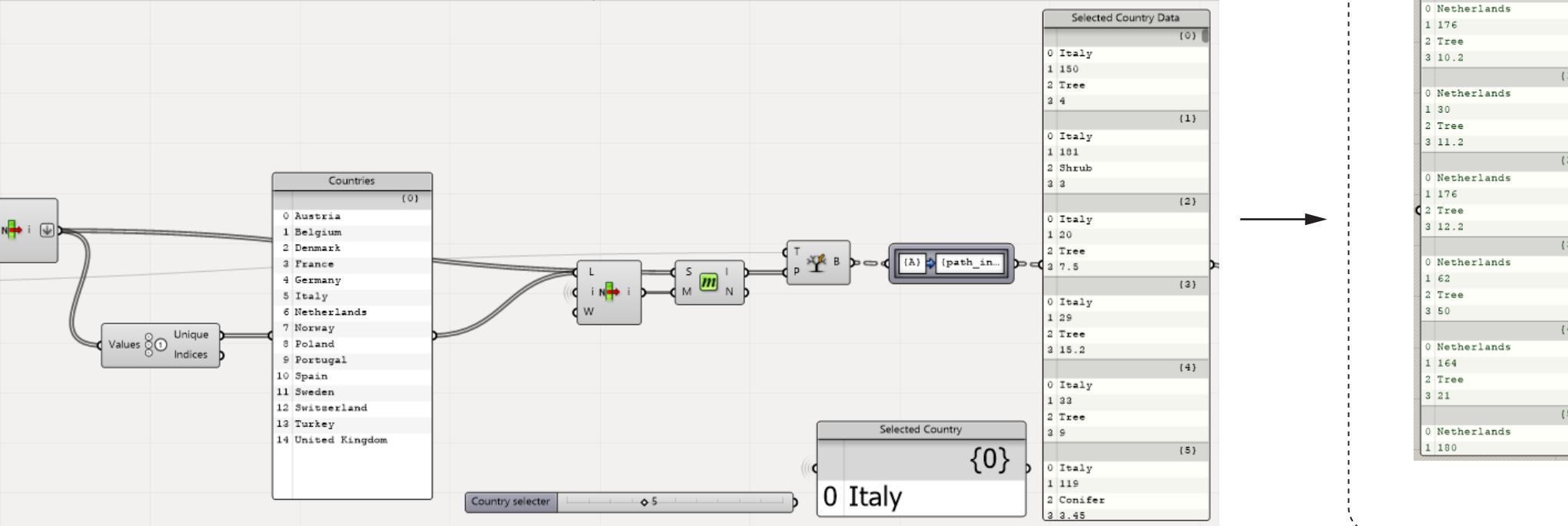
# Phase\_2[Augmentation]

*Increasing training data  
N=1068 to N=4272 by creating variations  
in possible heights*

Poland	Betulaceae	<i>Corylus avellana</i>	Shrub	6.5
Poland	Betulaceae	<i>Corylus avellana</i>	Shrub	5.5
Poland	Betulaceae	<i>Corylus avellana</i>	Shrub	4.5

Phase\_2[Feature extraction] ↓

using separate country-specific datasets to avoid geographic errors



sing]

—  
—  
—

In



*datasets*  
*ntry's data*

The diagram illustrates a neural network architecture for species classification. It consists of three main layers:

- Input Layer:** Represented by a dashed vertical line on the left. A horizontal line labeled "normalized heights" originates from this layer.
- Hidden Layer:** Represented by a dashed vertical line in the center. A horizontal line labeled "one-hot encoded vectors denoting type" originates from this layer.
- Output Layer:** Represented by a dashed vertical line on the right. A horizontal line labeled "species re-labelled for frequency within country set" originates from this layer.

A large black arrow points horizontally from the Input Layer towards the Output Layer, indicating the flow of data through the network.

## phase\_2[P]

