# Deep Learning - Tree Survival Prediction

Recommended:

Kaggle is the world's largest data science community with powerful tools and resources to help you achieve your data science goals.

Register via <https://www.kaggle.com>

Tree seedling functional traits mediate plant-soil feedback survival responses across a gradient of light availability.

**Methodology:**

(the following information provided by the authors of the experiment)

We conducted a factorial blocked design field experiment, consisting of four tree species, seven soil sources (sterilized conspecific, live conspecific, and five heterospecific), and a gradient of forest understory light levels (low, medium, and high), for a total of 3,024 seedlings. We monitored seedling survival twice per week over one growing season, and we randomly selected subsets of seedlings to measure mycorrhizal colonization and phenolics, lignin, and NSC measurements at three weeks. We used Cox proportional hazards survival models to evaluate survival and linear mixed effects models to test how light availability and soil source influence traits.

**Detailed information about each column follows:**

* **No**: Seedling unique ID number.
* **Plot**: Number of the field plot the seedling was planted in. (1-18)
* **Subplot**: Subplot within the main plot the seedling was planted in. Broken into 5 subplots (1 per corner, plus 1 in the middle). (A-E)
* **Species**: Includes Acer saccharum, Prunus serotina, Quercus alba, and Quercus rubra.
* **Light ISF**: Light level quantified with HemiView software. Represents the amount of light reaching each subplot at a height of 1m.
* **Light Cat**: Categorical light level created by splitting the range of Light\_ISF values into three bins (low, med, high).
* **Core**: Year the soil core was removed from the field.
* **Soil**: Species from which the soil core was taken. Includes all species, plus Acer rubrum, Populus grandidentata, and a sterilized conspecific for each species.
* **Adult**: Individual tree that soil was taken from. Up to 6 adults per species. Used as a random effect in analyses.
* **Sterile**: Whether the soil was sterilized or not.
* **Conspecific**: Whether the soil was conspecific, heterospecific, or sterilized conspecific.
* **Myco**: Mycorrhizal type of the seedling species (AMF or EMF).
* **SoilMyco**: Mycorrhizal type of the species culturing the soil (AMF or EMF).
* **PlantDate**: The date that seedlings were planted in the field pots.
* **AMF**: Percent arbuscular mycorrhizal fungi colonization on the fine roots of harvested seedlings.
* **EMF**: Percent ectomycorrhizal fungi colonization on the root tips of harvested seedlings.
* **Phenolics**: Calculated as nmol Gallic acid equivalents per mg dry extract (see manuscript for detailed methods)
* **NSC**: Calculated as percent dry mass nonstructural carbohydrates (see manuscript for detailed methods)
* **Lignin**: Calculated as percent dry mass lignin (see manuscript for detailed methods)
* **Census**: The census number at which time the seedling died or was harvested.
* **Time**: The number of days at which time the seedling died or was harvested.
* **Event**: Used for survival analysis to indicate status of each individual seedling at a given time (above)
  + 0 = harvested or experiment ended
  + 1 = dead
* **Harvest**: Indicates whether the seedling was harvested for trait measurement.
* **Alive**: Indicates if the seedling was alive at the end of the second growing season. "X" in this field indicates alive status.

Missing data is coded as NA.

**Acknowledgements:**

All data was collected from single experiment and is presented in the associated manuscript: Wood, Katherine; Kobe, Richard; Ibáñez, Inés; McCarthy-Neumann, Sarah (2023). Tree seedling functional traits mediate plant-soil feedback survival responses across a gradient of light availability.

If you use this dataset in your research, please credit the original authors.  
<https://doi.org/10.5061/dryad.xd2547dpw>