Configuration Options for MDITRE

July 9, 2021

1 General options

- **data**: path to the pickle file containing the preprocessed dataset
- **data_name**: Name of the dataset, used to create a directory to store the model output
- **save_as_csv**: Save all the model output as CSV files
- **verbose**: Print training output logs to the console

2 Training options

- **workers**: Number of cpu threads to use for pytorch data loading
- **epochs**: Number of training iterations to run the model
- **batch-size**: Number of training samples in a single batch. For full batch training it is set to the dataset size
- **deterministic**: Deterministic training for reproducibility
- **seed**: Random seed for reproducibility
- **cv_type**: Cross-Validation procedure to be used. Options are "loo" (leave-one-out), "kfold" and "None"
- **kfolds**: Number of folds to use if cv_type chosen is "kfold"
- **distributed**: Use multiprocessing for training
- **local_rank**: Rank of the current process if using distributed training

3 Optimization options

- **lr_fc**: Learning rate for the regression coefficients. Value of 0.001 works well in practice.
- **lr_bias**: Learning rate for the regression bias term. Value of 0.001 works well in practice.
- **lr_alpha**: Learning rate on detector selector parameters. Value of 0.001 works well in practice. Using a very high value may result in model not training.
- **lr_beta**: Learning rate on rule selector parameters. Value of 0.001 works well in practice. Using a very high value may result in model not training.
- **lr_thresh**: Learning rate on abundance threshold parameters. This value is dependant on the scale of the abundances. We found that value of 0.001 works well for all our datasets.
- **lr_slope**: Learning rate on slope threshold parameters. This value is dependent on the scale of rate of change of abundances. We found that a value of 0.0001 works well for all our datasets.
- **lr_time**: Learning rate on time window length parameters. Value of 0.01 works well in practice since the time windows are generally on a higher scale (days, months etc.)
- **lr_mu**: Learning rate on time window center parameters. Value of 0.01 works well in practice since the time window centers are generally on a higher scale (day, month etc.)
- **lr_kappa**: Learning rate on phylogenetic radius parameters. Value of 0.001 works well in practice.
- **lr_eta**: Learning rate on phylogenetic embedding parameters. Value of 0.001 works well in practice.

4 Model options

- **min_k_bc**: Initial temperature (sharpening factor) before annealing on the rule and detector selectors.
- **max_k_bc**: Final temperature (sharpening factor) after annealing on the rule and detector selectors.
- **min_k_thresh**: Initial temperature (sharpening factor) before annealing on the threshold detector response.
- **max_k_thresh**: Final temperature (sharpening factor) after annealing on the threshold detector response.

- **min_k_slope**: Initial temperature (sharpening factor) before annealing on the slope detector response.
- **max_k_slope**: Final temperature (sharpening factor) after annealing on the slope detector response.
- **min_k_time**: Initial temperature (sharpening factor) before annealing on the temporal focus response.
- **max_k_time**: Final temperature (sharpening factor) after annealing on the temporal focus response.
- **min_k_otu**: Initial temperature (sharpening factor) before annealing on the phylogenetic focus response.
- **max_k_otu**: Final temperature (sharpening factor) after annealing on the phylogenetic focus response.
- **z_mean**: Mean of the Negative Binomial prior on the detector selectors.
- **z_var**: Variance of the Negative Binomial prior on the detector selectors.
- **z_r_mean**: Mean of the Negative Binomial prior on the rule selectors.
- **z_r_var**: Variance of the Negative Binomial prior on the rule selectors.
- **w_var**: Variance of the Normal prior on the regression coefficients.