

Machine Learning Toolbox: Program Create Model

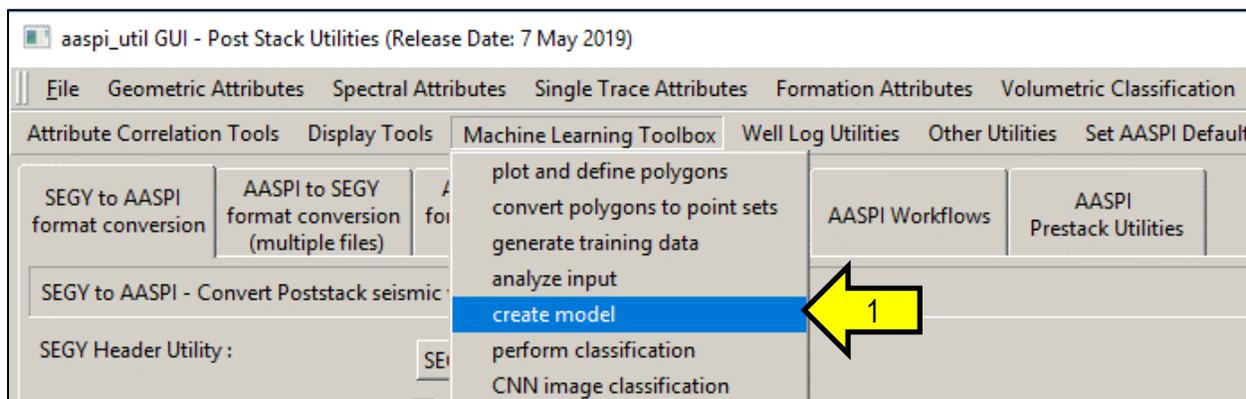
MACHINE LEARNING TOOLBOX – Create Model

Contents

Running aaspi_machine_learning_create_model	1
Output file naming convention	3

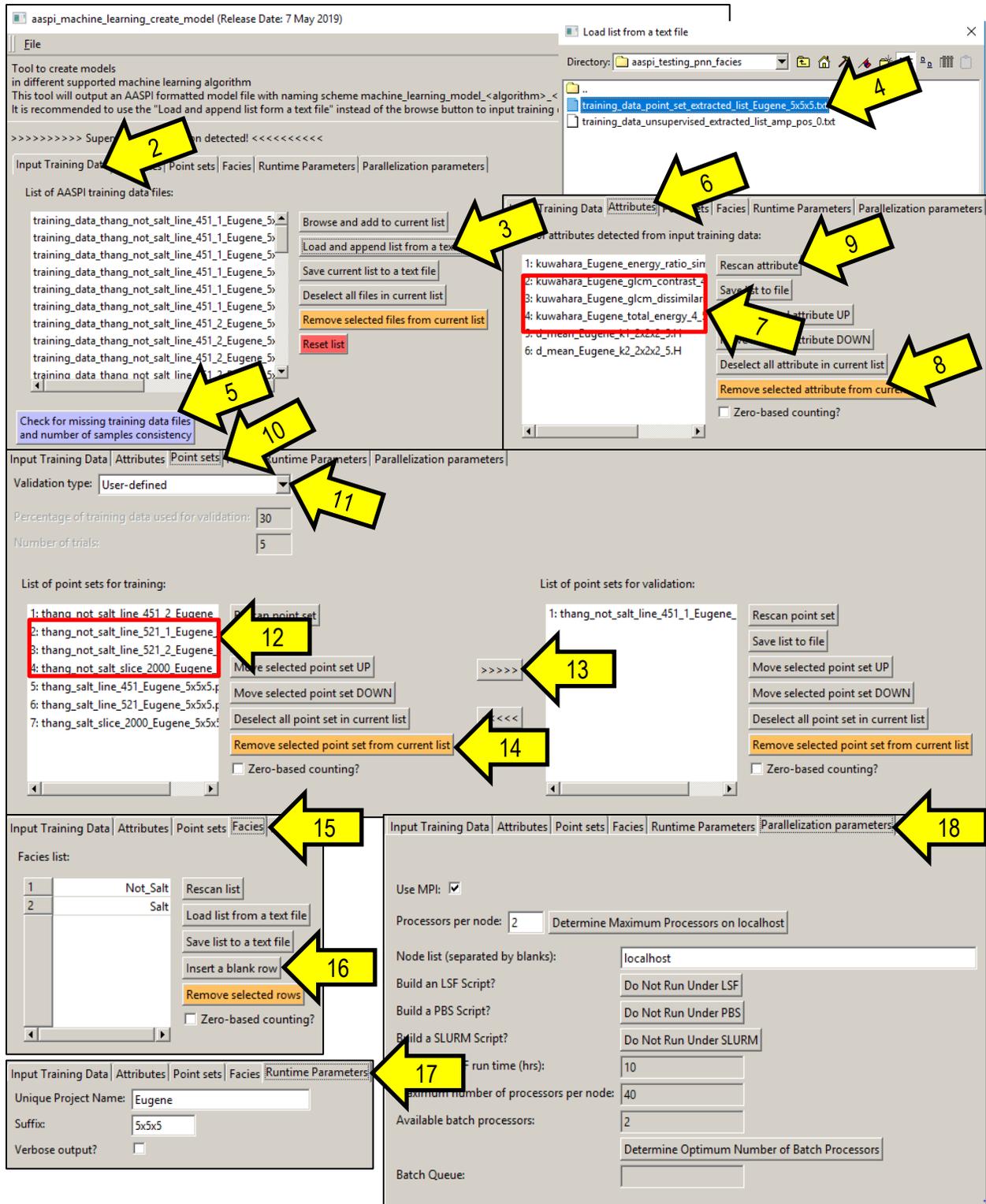
Running aaspi_machine_learning_create_model

After input training data are analyzed and users know the exact selection of attributes and point sets they want to use for classification, the next step is to build a model for machine learning. Go to *Machine Learning Toolbox* → *create model* (1).



Click on Input Training Data tab (2). To quickly browse all training data generated by **aaspi_training_data**, click “Load and append a list from a text file” button (3), then select the extracted list text file (4). This list contains all extracted training data files from generate training data step. If you decide to remove some training data files from the list, make sure to click “Check for missing training data files and number of samples consistency” button (5). If there is no error, you can proceed. It is not recommended to remove training data files directly this way, but rather removing all training data belonging to a specific attribute or point set. More on this later.

Machine Learning Toolbox: Program Create Model



Click on “Attributes” tab (6) to view the automatically detected attribute list associated with the input training data. You can move attribute up/down within the list. To remove some attributes and all of their associated training data files, first highlight them in the list (7) by clicking on

Machine Learning Toolbox: Program **Create Model**

them, and then click the “*Remove selected attributes from current list*” button (8). If you want to reset to the default attribute list, click on “*Rescan Attribute*” button (9).

Specify transformation parameters for your input attributes. Refer to Machine Learning Toolbox Program Analyze Input documentation for more detail about transformation parameters.

For supervised classification, click on “*Point sets*” tab (10) to view and define training point sets and validation point sets. For model creation, currently AASPI only supports user-defined training/validation data distribution (11): the user will decide which point sets to be used for training and which for validation. To transfer some training point sets to the validation list, first select the training point sets you want to move by clicking on them (12), then click on the transfer button (>>>>) (13) to swap them to the validation list (and vice versa). A model is created just on that particular training-validation setting. For PNN algorithm, model creation only depends on training point sets.

To remove a point set and all associated training data files, highlight them and click on “*Remove selected point sets from current list*” button (14).

For supervised classification, Click on “*Facies*” tab (15) to view and edit the facies list associated with the input training data. **CAUTION!!! DO NOT modify facies names, unless you are trying to merge multiple training data sets with different facies lists. The model creation process depends on facies names to match a training data file to a facies, and if it cannot find a matching facies name, it will discard the training data file from the model! In case you do need to modify the facies list, you can insert a blank row (16) and then double click on the blank row to define the new facies name. DO NOT modify existing facies!**

Click on “*Run time parameters*” tab (17) to define unique project name and suffix. These should be automatically loaded after user browse input training data.

Finally, click on “*Parallelization parameters*” tab (18) to define MPI parameters (i.e. number of parallel processors).

In the lower section, click on the desired algorithm based on which you want to create a model. Please refer to each algorithm’s documentation to see how to set up the parameters for each of them. The documentation file name of individual algorithm is as follow:

Machine_Learning_Toolbox-create_model-< algorithm_name>.pdf

Output file naming convention

Program `aaspi_machine_learning_create_model` will always generate the following output files:

Output file description	File name syntax
program log information	machine_learning_create_model_algorithm_name_unique_project_name_suffix.log

Machine Learning Toolbox: Program **Create Model**

program error/completion information	machine_learning_create_model_ <i>algorithm_name_unique_project_n</i> <i>ame_suffix</i> .err
--	---

where the values in red are defined by the program GUI. The errors we anticipated will be written to the **.err* file and be displayed in a pop-up window upon program termination. These errors, much of the input information, a description of intermediate variables, and any software trace-back errors will be contained in the **.log* file.