Sklearn Regression

Title Sklearn Regression

Description

This is a largely experimental scripting tool intended to provide general access to various Scikit-Learn regression models that can be accessed through a tool interface. The tool takes an input feature class,a linear model choice, a dependent field, independent variable fields, and a few choice parameters of some of the models, if they support them.

Source: http://scikit-learn.org/stable/modules/classes.html#module-sklearn.linear_model

Model Sources: http://scikit-learn.org/stable/modules/linear_model.html#linear-model

Usage

This is a largely experimental scripting tool intended to provide general access to various Scikit-Learn regression models that can be accessed through a tool interface. The tool takes an input feature class,a linear model choice, a dependent field, independent variable fields, and a few choice parameters of some of the models, if they support them.

Source: http://scikit-learn.org/stable/modules/classes.html#module-sklearn.linear_model

Generalized linear models: http://scikit-learn.org/stable/modules/linear_model.html#linear-model

Syntax

SklearnRegression (Input_Feature_Class, Regression_Model, Dependent_Variables, Independent_Variables, {Chosen_Alpha}, {Normalize_Data}, {Output_Directory})

Parameter	Explanation	Data Type
Input_Feature_Class	Dialog Reference This is the input feature class that has both the dependent (y) and independent variable (X) fields to be fit to an abitrary regression model. Predicted values from the model will extended to this feature class in a new field. There is no python reference for this parameter.	Feature Layer
Regression_Model	Dialog Reference Select a regression model from the linear model libary in scikit-learn. http://scikit-learn.org/stable/modules/classes.html#module-sklearn.linear_model Python Reference You can alter the code or the script tool parameters to add or support more aspects of a chosen model.	String

Dependent_Variables	Dialog Reference This is the feature (y) that you are trying to predict based on the chosen independent variables. Only one field is chose. Null values are interpreted as zeros. There is no python reference for this parameter.	Field
Independent_Variables	Dialog Reference These are the features (X) that are used attempt to predict a chosen dependent (Y). Null values are interpreted as zeros. There is no python reference for this parameter.	Multiple Value
Chosen_Alpha (Optional)	Dialog Reference Some regression types such as ridge and lasso have a regularization alpha. Constant that multiplies the L1 term. Defaults to 1.0. alpha = 0 is equivalent to an ordinary least square, solved by the LinearRegression object. ——————— Python Reference If the chosen model does not have an alpha parameter it won't be set.	Double
Normalize_Data (Optional)	Dialog Reference Many models have a normalize data parameter. If True, the regressors X will be normalized before regression. This parameter is ignored when fit_intercept is set to False. When the regressors are normalized, note that this makes the hyperparameters learnt more robust and almost independent of the number of samples. The same property is not valid for standardized data. However, if you wish to standardize, please use preprocessing.StandardScaler before calling fit on an estimator with normalize=False. ——————— Python Reference If the chosen model does not have an normalize parameter it won't be set.	Boolean
Output_Directory (Optional)	Dialog Reference This output workspace is where the pickeled model, and output report text file will be output. See Docs on model persistence: http://scikit-learn.org/stable/modules/model_persistence.html There is no python reference for this parameter.	Folder

Code Samples

There are no code samples for this tool.

Tags

Regression, Linear Models, Scikit-Learn

Credits

David Wasserman

Use limitations

There are no access and use limitations for this item.

You are currently using the Item Description metadata style. Change your metadata style in the Options dialog box to see additional metadata content.