sklearn api for linear regression, logistic regression, ridge and lasso

LinearRegression fits a linear model with coefficients w = (w1, …, wp) to minimize the residual sum of squares between the observed targets in the dataset, and the targets predicted by the linear approximation.

class sklearn.linear_model.LinearRegression(*, *fit_intercept=True*, *normalize=False*, *copy_X=True*, *n_jobs=None*, *positive=False*)

Parameters

- **fit_intercept**bool, default=True: Whether to calculate the intercept for this model. If set to False, no intercept will be used in calculations (i.e. data is expected to be centered).
- **normalize**bool, default=False: This parameter is ignored when fit_intercept is set to False. If True, the regressors X will be normalized before regression by subtracting the mean and dividing by the l2-norm.
- **copy_X**bool, default=True: If True, X will be copied; else, it may be overwritten.
- **n_jobs**int, default=None: The number of jobs to use for the computation.
- **positive**bool, default=False: When set to True, forces the coefficients to be positive.

Attributes

- **coef_**array of shape (n_features,) or (n_targets, n_features): Estimated coefficients for the linear regression problem.
- **rank_**int: Rank of matrix X..
- **singular_**array of shape (min(X, y),): Singular values of X.
- **intercept_**float or array of shape (n_targets,): Independent term in the linear model. Set to 0.0 if fit_intercept = False

Logistic Regression classifier.

class sklearn.linear_model.LogisticRegression(*penalty='l2', *, dual=False, tol=0.0001, C=1.0, fit_intercept=True, intercept_scaling=1, class_weight=None, random_state=None, solver='lbfgs', max_iter=100, multi_class='auto', verbose=0, warm_start=False, n_jobs=None, l1_ratio=None*)

Parameters

- **penalty**{'l1', 'l2', 'elasticnet', 'none'}, default='l2': Used to specify the norm used in the penalization.
- **dual** bool, default=False: Dual or primal formulation.
- **C** float, default=1.0: Inverse of regularization strength; must be a positive float.
- **fit_intercept**bool, default=True: Specifies if a constant (a.k.a. bias or intercept) should be added to the decision function.
- **intercept_scaling**float, default=1: Useful only when the solver 'liblinear' is used and self.fit_intercept is set to True.
- **random_state**int, RandomState instance, default=None: Used to shuffle the data..
- **max_iter**int, default=100: Maximum number of iterations taken for the solvers to converge.
- **multi_class**{'auto', 'ovr', 'multinomial'}, default='auto': If the option chosen is 'ovr', then a binary problem is fit for each label.
- **verbose** int, default=0: For the liblinear and lbfgs solvers set verbose to any positive number for verbosity.
- **n_jobs** int, default=None: Number of CPU cores used when parallelizing over classes if multi_class='ovr'".

Attributes

- **classes_**ndarray of shape (n_classes,): A list of class labels known to the classifier.
- **coef_**ndarray of shape (1, n_features) or (n_classes, n_features): Coefficient of the features in the decision function.
- **intercept_**ndarray of shape (1,) or (n_classes,): Intercept (a.k.a. bias) added to the decision function.
- **n_iter_**ndarray of shape (n_classes,) or (1,): Actual number of iterations for all classes.

Lasso

class sklearn.linear_model.Lasso(*alpha=1.0, *, fit_intercept=True, normalize=False, precompute=False, copy_X=True, max_iter=1000, tol=0.0001, warm_start=False, positive=False, random_state=None, selection='cyclic'*)

Parameters

- **alpha** float, default=1.0: Constant that multiplies the L1 term.
- **fit_intercept** bool, default=True: Whether to calculate the intercept for this model.
- **normalize** bool, default=False: This parameter is ignored when fit_intercept is set to False. If True, the regressors X will be normalized before regression by subtracting the mean and dividing by the l2-norm.
- **precompute**bool or array-like of shape (n_features, n_features), default=False: Whether to use a precomputed Gram matrix to speed up calculations.
- **copy_X**bool, default=True: If True, X will be copied; else, it may be overwritten.
- **max_iter**int, default=1000: The maximum number of iterations.
- **positive**bool, default=False: When set to True, forces the coefficients to be positive.
- **random_state** int, RandomState instance, default=None: The seed of the pseudo random number generator that selects a random feature to update.
- **selection**{'cyclic', 'random'}, default='cyclic': If set to 'random', a random coefficient is updated every iteration rather than looping over features sequentially by default.

Attributes

- **coef_**ndarray of shape (n_features,) or (n_targets, n_features): Parameter vector
- **dual_gap_**float or ndarray of shape (n_targets,): Given param alpha, the dual gaps at the end of the optimization, same shape as each observation of y.
- **intercept_**float or ndarray of shape (n_targets,): Independent term in decision function.
- **n_iter_**int or list of int: Number of iterations run by the coordinate descent solver to reach the specified tolerance.

Ridge

class sklearn.linear_model.Ridge(*alpha=1.0, *, fit_intercept=True, normalize=False, copy_X=True, max_iter=None, tol=0.001, solver='auto', random_state=None*)

Parameters

- **alpha**{float, ndarray of shape (n_targets,)}, default=1.0: Regularization strength; must be a positive float.
- **fit_intercept**bool, default=True: Whether to fit the intercept for this model.
- **normalize**bool, default=False: This parameter is ignored when fit_intercept is set to False. If True, the regressors X will be normalized before regression by subtracting the mean and dividing by the l2-norm.
- **copy_X**bool, default=True: If True, X will be copied; else, it may be overwritten.
- **max_iter**int, default=None: Maximum number of iterations for conjugate gradient solver.
- **solver**{‘auto’, ‘svd’, ‘cholesky’, ‘lsqr’, ‘sparse_cg’, ‘sag’, ‘saga’}, default=’auto’
- **random_state**int, RandomState instance, default=None

Attributes

- **coef_**ndarray of shape (n_features,) or (n_targets, n_features): Weight vector(s).
- **intercept_**float or ndarray of shape (n_targets,): Independent term in decision function. Set to 0.0 if fit_intercept = False.
- **n_iter_**None or ndarray of shape (n_targets,): Actual number of iterations for each target. Available only for sag and lsqr solvers. Other solvers will return None.