A person in a uniform

Description automatically generated with low confidence

**YiLong 仪陇**

**Package for While- and Post-Tuning Analysis**

*10/01/2023*

**Background**

YiLong, the native city of Chinese Revolutionary and Leader Zhu De (朱德), is what this While Tuning Analysis Package is named after. The purpose of this package is to cooperate with tuning packages such as JiXi and YangZhou to provide real-time analysis of tuning results, as well as post-tuning analysis.

The naming was particularly chosen as Zhu was a faithful deputy wielding the military arm to Paramount Leader Mao Zedong, whose native city alongside those of other paramount leaders are used to perform tuning.

The package performs two main functions in viewing the tuning data sorted in descending order by the validation score, and also viewing the means of evaluation metrics for combinations containing each individual value of a hyperparameter – for each hyperparameter.

**Functions**

|  |  |
| --- | --- |
| Function | Purpose |
| combine\_tuning\_results(tuning\_results, output\_address) | Takes in multiple tuning output DataFrames, merging and exporting it into one file  Parameters:  tuning\_results – list of DataFrames – at least length 2, all DataFrames must have same column name  output\_address – str – does not need to contain ‘.csv’ |

**Class**

|  |  |
| --- | --- |
| Class | Purpose |
| YiLong | Object that reads in a tuning result DataFrame, and can display:   1. Combinations sorted by validation score 2. Means of train, validation, test scores for combinations containing each individual value of a hyperparameter – for each hyperparameter. 3. Groups of train, validation, test scores grouped by combinations containing each individual value of a hyperparameter – for each hyperparameter |

**Methods:**

|  |  |
| --- | --- |
| Methods | Purpose |
| *YiLong(type)* | Initialisation must input type  Parameters:  type – str - either 'Classification'  or 'Regression' |
| read\_tuning\_result(address, extra\_to\_discard\_columns = None) | Reads in Tuning Result DataFrame using address as guide to csv  Parameters:  address – str – include ‘.csv’  extra\_to\_discard\_columns – list – optional, default None. List of column names which are not useful hyperparameters nor evaluation metrics |
| read\_sorted\_full\_df(interested\_statistic = None, ascending = False) | View the Tuning Result DataFrame in sorted evaluation metric (default Validation Score, decreasing) order.  Displays top 60 and bottom 60 (if len(DataFrame) <= 120 than may be overlap; if len(DataFrame) <= 60 then displays top and bottom len(DataFrame) combinations, where top and bottom dataframe is exactly the same  Also displays and returns the best combination according to the evaluation metric  Parameters:  interested\_statistic – str – must be a valid evaluation metric of the model type  ascending – bool – default False |
| read\_mean\_val\_scores() | View the means of evaluation metrics for combinations containing each individual value of a hyperparameter – for each hyperparameter. |
| read\_grouped\_scores() | View all evaluation metrics for combinations grouped by containing each individual value of a hyperparameter – for each hyperparameter  If any of the individual values of a hyperparameter exceeds 60, then sample down to 60 without replacement, using seed 19861201 |

**Objects:**

|  |  |
| --- | --- |
| Objects | Purpose |
| clf\_type | Str – either ‘Regression’ or ‘Classification’ |
| tuning\_result | DataFrame |
| hyperparameters | List |
| regression\_extra\_output\_columns  &  classification\_extra\_output\_columns  &  GLM\_Regression\_extra\_output\_columns | Lists containing column names that are used internally by YiLong |
| discard\_columns | List |

**Dependencies**

pandas