非任务导向的时序数据探索 Review

|  |  |  |  |
| --- | --- | --- | --- |
| Function name | Description | Status | Follow up |
| TimeSeriesAnalyzer | 初始化 | Completed. | 目前仅支持一种数据格式，参见[data\_formats](https://tsfresh.readthedocs.io/en/latest/text/data_formats.html) ，暂时选定Option 1作为目前的格式。  可支持多种数据格式。 |
| \_remove\_outliers | 去除离群点 | Completed. | 目前仅支持Target为数值型的sample级别的离群点去除。  可加入Feature级别的集群点识别。 |
| get\_problem\_category | category 取值：“1” 或者 “2”  “1” 代表分类问题  “2” 代表回归问题 | Completed. | 无 |
| get\_missing\_data | 返回当前数据集中数据缺失的信息：每一feature的数据缺失比。 | Completed. | 平滑算法填充缺失值 |
| get\_problem\_desc | 返回对于研究目标的描述：  离散型：   1. 类别占比   连续性   1. 统计信息 2. 取值分布 3. 偏度与峰度 | Completed. | 暂无 |
| get\_feature\_desc | 返回对于某一feature的描述：   1. 统计信息 2. 取值分布 3. 偏度与峰度 4. 移动统计量 5. 时间序列分解：趋势，季节性和残差 | Completed. | 暂无 |
| get\_features\_rel | 返回feature之间两两的相关度：   1. 特征相关系数矩阵（协方差） | Completed. | 升级算法 |
| get\_features\_target\_rel | 返回feature与target之间的相关性：  Target为连续型：   * 特征相关系数矩阵（协方差）   Target为离散型：   * Feature降维后可视化 | Completed. | 升级算法 |