برای انجام عملیات اصلی ( پیش بینی خطای نرم افزار)، داخل فایل پیکربندی به نام (WPDP)، فیلد granularity را به عدد 1 تغییر دهید، در غیر اینصورت و برای انجام عملیات پیش پردازش و عملیات متفرقه، عدد 3 را به granularity منتسب کنید.

برای انجام امور متفرقه باید فایل utils را اجرا نمایید. داخل این فایل تعداد تابع وجود دارند که عملکرد آن ها به شرح زیر است:

Find\_nonoverlapping: این تابع مسئولیت یک سان سازی تعداد رکورد های دیتا ست های مختلف جمع آوری شده از ابزار های مختلف را بر عهده دارد. این تابع دو ورود دارد. ورودی اول دیتا ست اول و ورودی دوم دیتا ست دوم جهت مقایسه می باشد. اگر دیتا ستی که میخواهید تعداد نمونه های آن را با توجه به یک دیتا ست دیگر کاهش یا افزایش دهید، باید به عنوان ورودی اول قرار بگیرد و دیتا ست دوم همان دیتا ست محوری یا لولا pivot شماست که مقایسه نسبت به آن انجام می شود. توجه داشته باشید که برای لود کردن دیتا ست ها توسط load\_datasets از کلاس پیش پردازش از فایل data\_handler، پارامتر این Drop\_unused\_column باید برابر با مقداری غیر از old باشد. همینطور باید داخل تابع load\_datasets نباید ستون اول که نام فایل ها می باشد حذف شود.

آدرسی که میخواهید دیتاست همسان سازی شده شما در آنجا کپی شود، باید داخل آخرین خط این تابع تغییر کند. ورودی دوم تابع to\_csv داخل تابع find\_nonoverlapping.

Copy\_status\_to\_new\_dataset: این تابع مسئولیت کپی کردن وضعیت buggy و clean بودن فایل های متناظر در دیتا ست اصلی به دیتا ست تولید شده توسط ابزار ها را بر عهده دارد. ورودی اول آن دیتا ست جدیدی است که تولید شده است و فاقد هر گونه متغیر کلاسی است. و دیتا ست دوم همان دیتا ست اصلی است. برای تغییر آدرس ها، بالای فایل در خطوط اول این قطعه کد می توانید تغییر دهید.