There is validation data in the tournament dataset which includes target values, thus there being a target column. Target data for data_types “test” and “live” are filled with “x”

The obfuscated values represent movements within those signals. When ran through a machine learning algorithm they are taken as continuous variables. These algorithms average out values when they are continuous which is why you get values really close to 0.5 (you can expect to see an average of 0.5 when you randomly sample a uniform distribution between 0 and 1, which is kinda what is happening).

NumerAI gauges your performance based on correlation, so even if your predictions are really close to 0.5, the relative distance from 0.5 of your predictions is used to calculate your correlation score.