

The main reason the features are as they are is that numer.ai isn't allowed to give out the financial datasets that they are subscribed to.

By hiding the name of the stock, and by anonymizing the features, they can still share the data with us.

The plus side is that you have little (if any!) data cleaning to do. All features are \sim normal distributions centered at 0.5. So maybe you want to subtract 0.5 from them when you feed them to a neural network, but that is about it.