The following is a scheme for weighting variables in a mixed-type dataset to convert into a Gower matrix.

x is a column of matrix X of shape (n, d).

If x is a numeric variable, then $w \leftarrow 1$.

Otherwise, if x is a categorical variable, then we can use the following equation to calculate the entry-wise unbiased sample kurtosis of H, the one-hot encoding of x, where h_c is the column of H corresponding to the c^{th} of C classes in x:

$$\widehat{\text{Kurtosis}} = \frac{(C^2n^2 - 1)(C + \frac{1}{(C-1)} - 2) - 3(Cn - 1)^2}{(Cn - 2)(Cn - 3)} + 3 \tag{1}$$

Then, we can calculate the weight of x as follows:

$$w \leftarrow \tanh \frac{\left(\sum_c^C \widehat{\text{Var}}(h_c)\right)\left(1 - \sum_c^C \widehat{\text{Var}}(h_c)\right)\left(\widehat{\text{Kurtosis}}\right)}{\log n} \tag{2}$$

The first term in parentheses is zero when all the values of the variable are the same, the second term is zero when they're all different, and the third term increases with the number of classes and decreases with the number of observations in each class.