PROBABILITY CALCULATION EXAMPLE

Consider the material As_2Ta_1 which was predicted to be topological by our algorithm. The R-number values (denoted by R_e for elements and R_p for pairs) and corresponding compositions (denoted by c) for the Complete Dataset are:

$$R_e(As) = 0.464, \quad c(As) = 2/3 = 0.667$$
$$R_e(Ta) = 0.918, \quad c(Ta) = 1/3 = 0.333$$
$$R_p(As, Ta) = 1.75, \quad c(As, Ta) = 3/3 = 1$$

If an element not present in the dataset is encountered, it can simply be neglected. The calculated feature for As_2Ta_1 values using parameters from the Complete Dataset results are:

$$F_1 = (0.667 * 0.464 + 0.333 * 0.918) = 0.6152$$
$$F_2 = 1 * 1.75 = 1.75$$
$$F_3 = 0.667 * (1/0.464) + 0.333 * (1/0.918) = 1.8002$$
$$F_4 = 1 * (1/1.75) = 0.5714$$

The probability of this material to be topological comes out to be 0.7222 as shown below. As it is more than 0.5, so it is most likely a topological material.

$$t = w_1F_1 + w_2F_2 + w_3F_3 + w_4F_4 + b = 0.9552$$
$$\text{Probability} = \frac{1}{1+\exp(-0.9552)} = 0.7222$$

NOTE:

It is best to use the R-number and formula parameter values from Complete Dataset results. This is because the results from this dataset are more acceptable than from other datasets, as it involves training over all the materials. The total number of predictions from this dataset is also low and not very high like that from the Spaced dataset.