Complement Coding Example

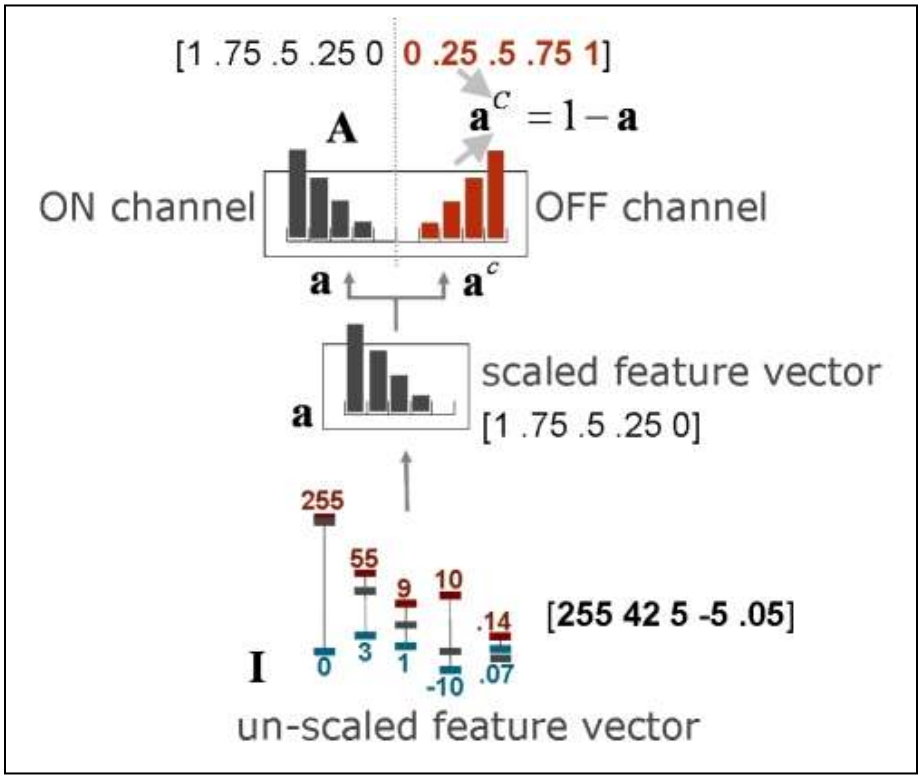
***Usage*: Complement_Code(feat_vect,min_vect,max_vect) :**

feat_vect is an input feature vector of size 1 X M where M is the number of features.

min_vect is a row vector of size 1 X M indicating the *minimum* allowable feature values.

max_vect is a row vector of size 1 X M indicating the *maximum* allowable feature values.

The following example is depicted in this complement coding icon



Consider an un-scaled, five-dimensional input feature vector I = [255 42 5 -5 0.05], where each feature has the following pair of *minimum* and *maximum* allowable values:

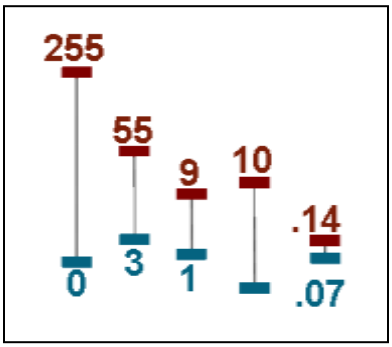
$I_1^{min} = 0$ $I_1^{max} = 255$

$I_2^{min} = 3$ $I_2^{max} = 55$

$I_3^{min} = 1$ $I_3^{max} = 9$

$I_4^{min} = -10$ $I_4^{max} = 10$

$I_5^{min} = 0.07$ $I_5^{max} = 0.14$



The minimum and maximum values are depicted as shown above.

The **min_vect** and **max_vect** feature values, based on the above values, are:

min_vect = [0 3 1 -10 0.07]

max_vect = [255 55 8 10 0.14]

The complement coded feature vector is obtained by running:

Complement_Code(feat_vect,min_vect,max_vect)

≡ **Complement_Code(**[255 42 5 -5 0.05]**,** [0 3 1 -10 0.07]**,** [255 55 9 10 0.14]**)**

To give the 10-dimensional complement coded output:

[1.0 .75 .50 .25 0 **0 .25 .50 .75 1.0**]

Each feature value is first scaled into the [0 1] range by using the minimum and maximum feature values, and all values that exceed the maximum or are below the minimum are truncated to 1 and 0 respectively.

Therefore, $a_2 = (I_2 - minimum) / (maximum - minimum) \equiv (42 - 3) / (55 - 5) = 0.75$

In the case of feature I_5, *the feature value is less than the allowable minimum* $I_5^{min} = 0.07$.

As a result, a_5 is truncated to 0.

Also, each feature and its complement in the complement coded vector add up to 1.

$a_1 + a_6 \equiv 1 + 0 = 1$

$a_2 + a_7 \equiv 0.75 + 0.25 = 1$

$a_3 + a_8 \equiv 0.50 + 0.50 = 1$

$a_4 + a_9 \equiv 0.25 + 0.75 = 1$

$a_5 + a_{10} \equiv 0 + 1 = 1$

Batch mode:

In the batch mode, multiple feature vectors, *I*, are provided in a single array with each row being one un-scaled feature vector.

To complement code un-scaled feature vectors I^a , I^b , and I^c , where the vectors are:

$\mathbf{I^a}$ = [255 42 5 -5 0.05]

$\mathbf{I^b}$ = [102 29 7 -4 0.14]

Using the same minimum and maximum feature values as above, the complement coded feature vectors can be obtained in an array by running

Complement_Code(data , min_vect, max_vect) ≡

Complement_Code ([255 42 5 -5 0.05; 102 29 7 -4 0.14],

[0 3 1 -10 .07]**,**

[255 55 9 10 0.14]**)**

The output is as follows:

[1.0 .75 .50 .25 0.0 **0.0 .25 .50 .75 1.0**

0.4 .50 .75 .30 1.0 **0.6 .50 .25 .70 0.0**]

The rows in the output array correspond to the complement coded vector of $\mathbf{I^a}$ and $\mathbf{I^b}$.

When running the code with the same input data, but without the *minimum* and *maximum* vectors, the *minimum* and *maximum* values are imputed from the data.