**The goal:**

1. we look into whether the sbs signatures could be used to determine the cancer types
2. we then find the most powerful sbs signatures for determine such cancer (eg. Lung cancer), then we set those sbs signatures as input and then set the gene mutation status as labels and classify the gene mutation status.
3. We then could find out which of the mutated gene could be used to determine the occurrence of that cancer. Thus，we could predict which of the gene are mutated in different cancer types.

**The steps:**

Classification of cancer types based on softmax function.

1. Building cancer label as one-hot.
2. Using SBS features (represented by an n-dimensional vector , ) as input.
3. The model is represented by the following formula:

Where is a matrix as the model parameter, is an m-dimensional vector as the model output and the is function of softmax.

1. Then we get the .

Classification of gene types based on softmax function.

1. Building gene label as one-hot.
2. Using SBS features (represented by an n-dimensional vector , ) as input.
3. The model is represented by the following formula:

Where is a matrix as the model parameter, is an m-dimensional vector as the model output and the is function of softmax.

1. Then we get the .