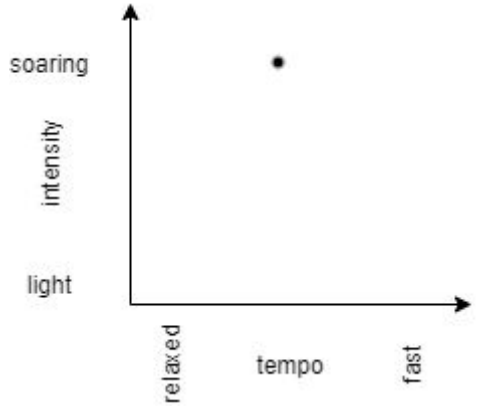
Naive Bayes

Supervised Classification: is the technique most often used for the quantitative analysis of remote sensing image data. At its core is the concept of segmenting the spectral domain into regions that can be associated with the ground cover classes of interest to a particular application. For Example:

1) From a album of tagged photos, recognize someone in a picture
2) Given someones music choices and a bunch of features of that music (tempo, genre etc) recommend a new song

Features and Labels: Suppose you like a song. Based on that song it is possible to figure out other songs you may like. In order to do that The Features of that song use as dimension. And the variety of the features known as Labels. Suppose you like "Let it go". The song has 2 features and both features has 2 labels. So the graph look like below



Decision Surface: In a statistical-classification problem with two classes, a decision boundary or decision surface is a hypersurface that partitions the underlying vector space into two sets, one for each class.

- ❖ When a decision surface is straight line it is called "Linear".

Bayes Rule:

$$P(A|B) = P(B|A).P(A)/P(B)$$

A, B = Events

$P(A|B)$ = Probability of A givren B is True

$P(B|A)$ = Probability of B givren A is True

$P(A),\ P(B)$ = the independent probabilities of A and B

Strengths and Weaknesses of Naive Bayes

Strengths: Easy and quick way to predict classes, both in binary and multiclass classification problems. In the cases that the independence assumption fits, the algorithm performs better compared to other classification models, even with less training data.

Weaknesses: The greatest weakness of the naïve Bayes classifier is that it relies on an often-faulty assumption of equally important and independent features which results in biased posterior probabilities.