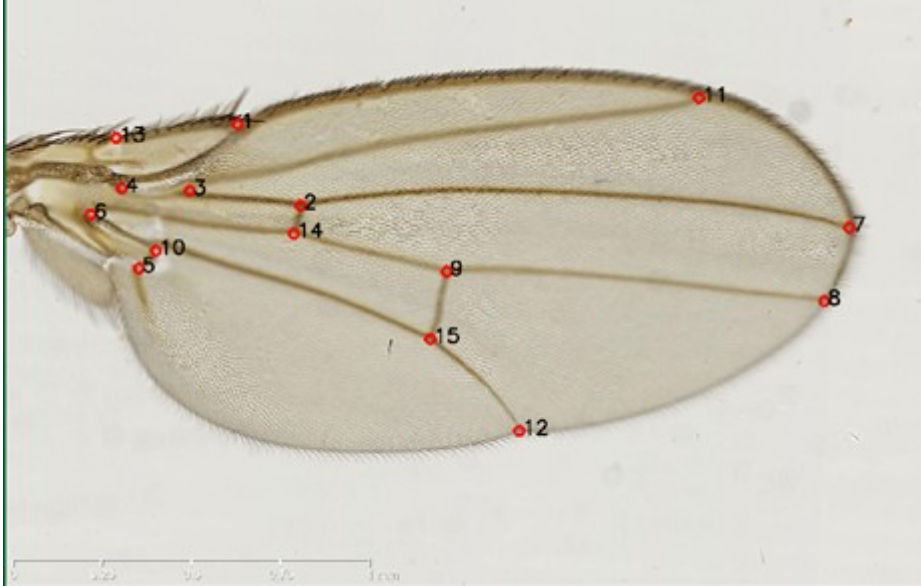


1. Self-study naive bayes classifier
2. Follow the tutorial on using logistic regression with scikit learn
<https://towardsdatascience.com/logistic-regression-using-python-sklearn-numpy-mnist-handwriting-recognition-matplotlib-a6b31e2b166a>
3. Apply Naïve Bayes classifier and artificial neural network (MLPClassifier) with mnist dataset
4. Using various classifier to solve the landmark prediction problem. The problem is described as followed:
Given a data set of images of drosophila wings. Predict the location of the landmark point in the image.



An example of landmark point in the wing images.

The framework included of two stages:

Training stage:

The goal of the training stage is to build a classifier which can predict if a point in the image is a landmark point (positive) or not (negative)

- o To select the landmark points (positive) and non-landmark points (negative). The positive points are selected as the point which are very close to the true landmark. The negative points are selected as the point which are far away from the landmark.
- o Use feature extraction to extract the local feature of each point (both negative and positive).
- o Train a classifier, taking the input as the local feature vector of one point. Predict if a point is landmark or non-landmark

Prediction stage:

- o Determine the region where the landmark should belong to
- o Randomly select 500 random points in the region. Those 500 points are called the candidate points
- o Extract local feature for those 500 candidate points
- o Use the classifier trained in the training stage to predict if a point is positive or negative
- o The predicted landmark is computed as the center point of all positive points

More information can be found in the paper :” Landmark detection in 2D bioimages for geometric morphometrics: a multi-resolution tree-based approach”

In this exercise we focus on the landmark number 9 only

Your materials:

RAW feature vectors of some positive and negative points (mark as _pos and _neg in the file name)

Some images with their true landmark

A function to extract RAW feature vector

Your task:

- Load the RAW feature vector, train the classifier to classify if an image is negative or positive
- Determine the region in the image where the landmark 9 should belong to
- Randomly select 500 candidate points in the region
- Use the classifier trained in the training stage to predict if a point is positive or negative. The predicted landmark is computed as the center point of all positive points
- Calculate the mean absolute error over all the images.

