Project Name: - Lice classification using transfer learning models

Name: - Aarti Darji

1. Introduction  
   My project focuses on image classification using transfer learning models. A transfer learning model is when we use a pre-trained model rather than a new model that is trained on small dataset. By using a transfer learning model, we can achieve higher performance.
2. Dataset  
   The dataset I’m using is the lice data set provided by Dr. Rostami. The lice are classified by 8 species depending on their size and morphology. It is of JPEG format.
3. Data Preparation  
   For the data preparation part of my project, I will first use data augmentation to get multiple data points of the lice images with different alterations, this reduces the overfitting problem. Then, we reshape the images to fit to the model. After increasing the size of the dataset, we can split it into training dataset and test dataset.
4. Exploratory Data Analysis (EDA)  
   Exploratory data analysis is done to understand the relationship between the 8 different classes of lice and by visualizing the datasets by plotting the distribution of different classes of lice. Furthermore, I will get an insight on the image metrics, such as the dimensions, mean/median, timestamp, and image formats.
5. Modelling  
   I’ll be using pre-trained neural network models (transfer learning models) for faster results and maximum optimizations. The ones I’ll be using are AlexNet, VGG and ResNet. AlexNet is an 8-layer deep convolutional neural network that classify regular objects such a mouse, keyboard and a bicycle. VGG is deep convolutional neural network that can support either 16 or 19 layers and can have up to 1000 categories/classes. It is one of the most popular model used for image recognition. ResNet is also a convolutional neural network used for deep learning. It is different in the sense that it has something called ‘skip connections’ which serve as a solution to the vanishing gradient problem.
6. Evaluation  
   For evaluating the model, I can plot the ROC/AUC curve to determine its accuracy and precision. I can also plot the confusion matrix to visualize how its accuracy. Based on the error rate and the accuracy, we can determine the performance of the model.