

CSCE 633

Classify three types of flowers from the UCI IRIS dataset

1. Data exploration: compute the number of samples belonging to each class, plot the histogram of each feature, plot scatter plots of all pairs of features
2. Implement a **KNN** classifier from scratch using the Euclidean distance as a distance measure to classify between the three classes
3. Plot the accuracy of the model against with the different K values, choose the best K^*

Machine learning with Pokemon GO

1. Data exploration: visualize the scatter plots and compute the Pearson's correlation coefficient between the numerical attributes themselves and the combat points
2. Pre-processed the categorical attributes with the one hot encoding
3. Implement a **Linear Regression** model to predict the combat points, use the 5-fold cross-validation to evaluate the model.

Machine learning for facial recognition

1. Visualize the grayscale images for each emotion
2. Use the **FNN** and **CNN** to perform the emotion classification task respectively

Classify benign vs malignant tumors

1. Data exploration: plot the histograms of the class outcome and each feature, compute the number of samples belonging to the benign and malignant case
2. Implement a function that computes the conditional entropy of each feature
3. Use a **Decision Tree** and **Random Forest** classifier to classify between benign and malignant tumor based on the features provided

Perform COVID-19 diagnosis from chest Xray images

1. Crop and resize the images, extract HoG features
2. Train **SVM** and **Random Forest** classifier with that features
3. Improve the model by VGG-19 neural-network under the guidance of the research papers.