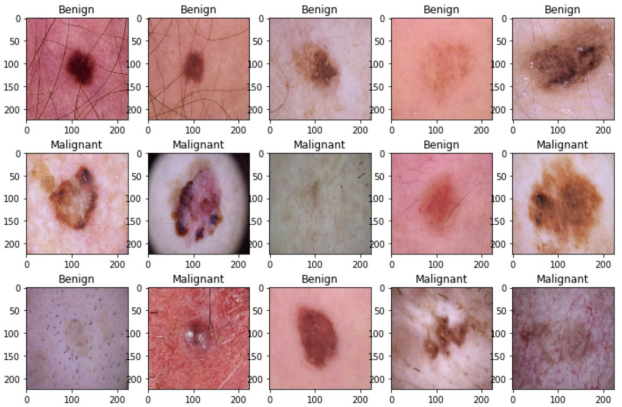
Ai4All Class 10

Project Ideas

Medical Domain (Idea 1)

Skin Cancer Detection (Supervised Learning)

1. Dataset - The data consists of two folders with each 1800 pictures (224x244) of the two types of moles.
2. Aim - To classify images as Malignant / Benign Skin Cancer types
3. Image Classification Problem
4. Skills - Deep Learning (Pytorch / Deep Learning Architectures)
5. To read :
    a. CNNs:
        i. https://towardsdatascience.com/convolutional-neural-networks-explained-9cc5188c4939
        ii. https://medium.com/@RaghavPrabhu/understanding-of-convolutional-neural-network-cnn-deep-learning-99760835f148
        iii. https://ai.plainenglish.io/deep-dive-into-the-world-of-cnns-8cf22cd84e7
    b. Metrics:
        i. https://www.analyticsvidhya.com/blog/2021/07/metrics-to-evaluate-your-classification-model-to-take-the-right-decisions/



Medical Domain (Idea 2)

Brain Tumor (Supervised Learning)

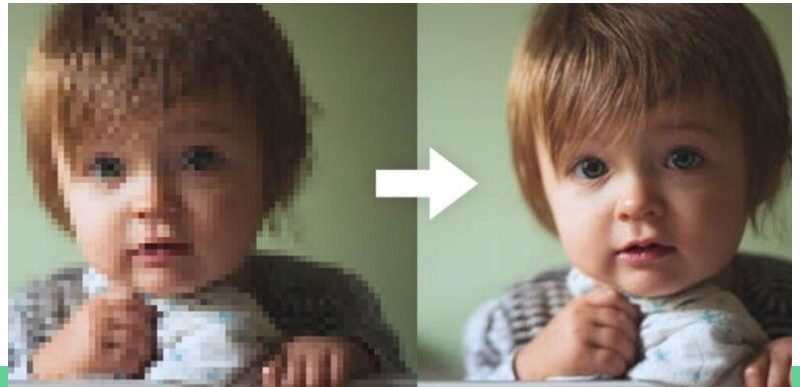
1. Dataset - The dataset contains 7465 columns and 2 labels
2. Aim - Detect if a patient has brain tumor / not
3. Numerical Classification Problem
4. Skills -
    a. We can use one of the classification algorithms
        i. Logistic Regression , KNN(K Neighbors) , Decision Tree , Random Forests , SVM , GBM (Gradient Boosting Machine), AdaBoost, Bagging
    b. Exploratory Data Analysis
    c. Metrics - Confusion matrix

#	# X53416	# M83670	# X90908	# M97496	# X90908.1
0	70	-81	25	10	22
1	108	-30	-7	60	0
2	75	-1	5	48	6
3	871	4	14	78	-6
4	-92	-34	14	19	11
5	21	-13	5	11	-18
6	225	118	-5	175	-40
7	-346	-35	37	42	39
8	-378	31	-29	105	-54

Image Generation

Improve the quality of low resolution images by using Generative Modelling (Unsupervised Learning)

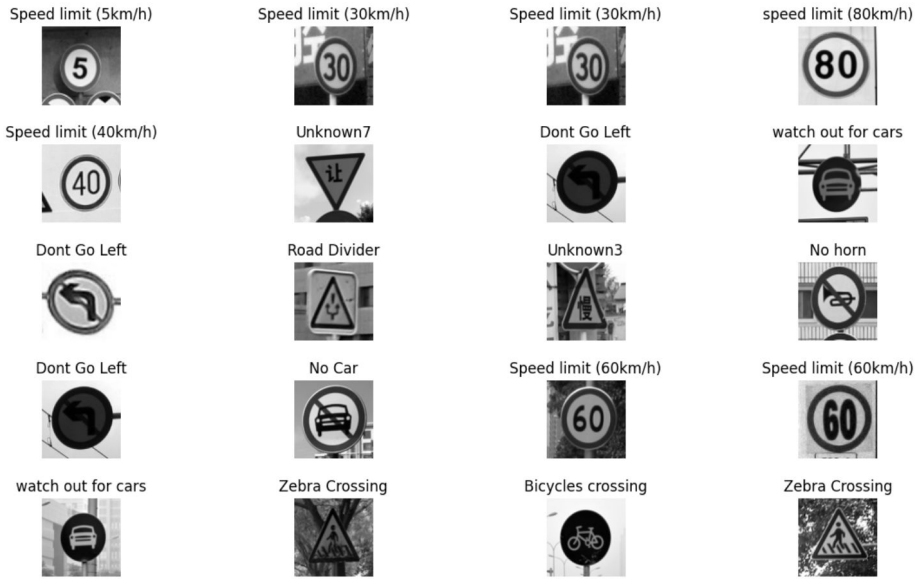
1. Dataset - The dataset contains 5000 low and high resolution images
2. Aim - To convert low resolution images to high resolution images using Autoencoders (Generative AI)
3. Skills -
   a. https://towardsdatascience.com/auto-encoder-what-is-it-and-what-is-it-used-for-part-1-3e5c6f017726
   b. https://www.analyticsvidhya.com/blog/2021/06/autoencoders-a-gentle-introduction/#:~:text=FAQs-,What%20are%20Autoencoders%20%3F,are%20called%20self%2Dsupervised%20models.



Traffic Congestion (Idea 1)

(Supervised Learning)

1. Dataset - There are around 58 classes and each class has around 120 images
2. Aim - Classify Traffic Signs into multiple classes
3. Image Classification Problem
4. Skills - As mentioned in the Medical Idea 1 slide



Traffic Congestion (Idea 2)

Addressing False Positives in Pedestrian Detection (Supervised Learning)

1. Dataset - contains 1338 images of person / person-like images
2. Aim - Detect Person / Person-like Classes
3. Image Classification Problem
4. Skills - As mentioned in the Medical Idea 1 slide