Project List

This list contains only **suggested** datasets and tasks. You can use a custom dataset (for safety, check with me first). The only requirement is that it is interesting.

1. Brain tumor classification from MRI images with CNN networks

   https://www.kaggle.com/sartajbhuvaji/brain-tumor-classification-mri

2. Credit card fraud detection with MLP networks

   https://www.kaggle.com/mlg-ulb/creditcardfraud

3. Fruit and vegetables image detection with CNN networks

   https://www.kaggle.com/moltean/fruits

4. Image detection with MLP networks on the Fashion-MNIST dataset

   https://www.kaggle.com/zalando-research/fashionmnist

5. Recognize dog vs cat image with CNN networks

   https://www.kaggle.com/vaishnavkapil/feature-detection-cnn

6. Detect handwritten letters with MLP networks

   https://www.kaggle.com/sachinpatel21/az-handwritten-alphabets-in-csv-format

7. Detect heart disease based on health measurements with logistic regression or MLP networks

   https://www.kaggle.com/ronitf/heart-disease-uci

8. Gastrointestinal Cancer MSI MSS Prediction with CNN networks

   https://www.kaggle.com/linjustin/train-val-test-tcga-coad-msi-mss

9. Predict mushroom toxicity with linear regression

   https://www.kaggle.com/uciml/mushroom-classification

10. Detect pneumonia from chest X-Rays using CNN networks

    https://www.kaggle.com/paultimothymooney/chest-xray-pneumonia

11. Detect diabetes from diagnostic measurements with logistic regression or MLP (csv file)

    https://www.kaggle.com/uciml/pima-indians-diabetes-database

12. Detect dog breed from images using CNN networks

    https://www.kaggle.com/venktesh/person-images (only the dog breed part of the dataset)

13. Recognize sign language using MLP networks on the Sign Language MNIST

    https://www.kaggle.com/datamunge/sign-language-mnist

14. Recognize Chinese handwritten numbers using MLP networks

    https://www.kaggle.com/gpreda/chinese-mnist

15. Bird species classification with CNN networks

    https://www.kaggle.com/gpiosenka/100-bird-species

16. Bus vs car image detection with CNN networks

    https://www.kaggle.com/positivepc/object-detection

MedMNIST datasets:

All data for the next projects are available here: https://medmnist.github.io/, or see paper here: https://arxiv.org/pdf/2010.14925.pdf

- 17. Cancer patology detection from the PathMNIST dataset with MLP networks
- 18. Chest Xray classification on ChestMNIST with MLP networks
- 19. Skin lesion classification on DermaMNIST with MLP networks
- 20. Retina disease classification on OCTMNIST with MLP networks
- 21. Pneumonia detection on PneumoniaMNIST using MLP networks (or logistic regresison)
- 22. Diabetic retinopathy severity classification on RetinaMNIST with MLP networks
- 23. Breast cancer detection on BreastMNIST with MLP networks (or logistic regresison)
- 24. Organ identification from CT image slices (OrganMNIST_Axial or Coronal or Saggital) with MLP networks
- **Notes**:
  - data file have extension *.npz, can be unzipped like a normal zip archive
  - inside there are multiple arrays saved as *.npy files. These are originally Python files, but they can be read in Matlab with the functions from here: https://github.com/kwikteam/npy-matlab
  - if you have problems, contact me, I can convert them for you
  - most images are resized to 28x28 from larger images. If you get the original images, you can use CNN networks instead.

25. **Or any other** interesting dataset found on the Internet

    Popular dataset sources:

    - https://www.kaggle.com/datasets
    - www.zenodo.org
    - https://archive.ics.uci.edu/ml/datasets.php