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. Credit card fraud detection with logistic regression

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

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. Image detection with CNN networks on the Fashion-MNIST dataset

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

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

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

8. Predict heart disease based on health measurements using linear regression

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

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 (csv file)

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

12. Detect diabetes from diagnostic measurements with MLP networks (csv file)

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

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

MedMNIST datasets:

Available here: https://medmnist.github.io/, or see paper here: https://arxiv.org/pdf/2010.14925.pdf

- 15. Cancer patology detection from the PathMNIST dataset with MLP networks
- 16. Chest Xray classification on ChestMNIST with MLP networks
- 17. Skin lesion classification on DermaMNIST with MLP networks
- 18. Retina disease classification on OCTMNIST with MLP networks
- 19. Pneumonia detection on PneumoniaMNIST using MLP networks (or logistic regresison)
- 20. Diabetic retinopathy severity classification on RetinaMNIST with MLP networks
- 21. Breast cancer detection on BreastMNIST with MLP networks (or logistic regresison)
- 22. Organ identification from CT image slices (OrganMNIST_Axial) with MLP networks
- 23. Organ identification from CT image slices (OragnMNIST_Coronal) with MLP networks
- 24. Organ identification from CT image slices (OrganMNIST_Sagittal) 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.com
- https://archive.ics.uci.edu/ml/datasets.php