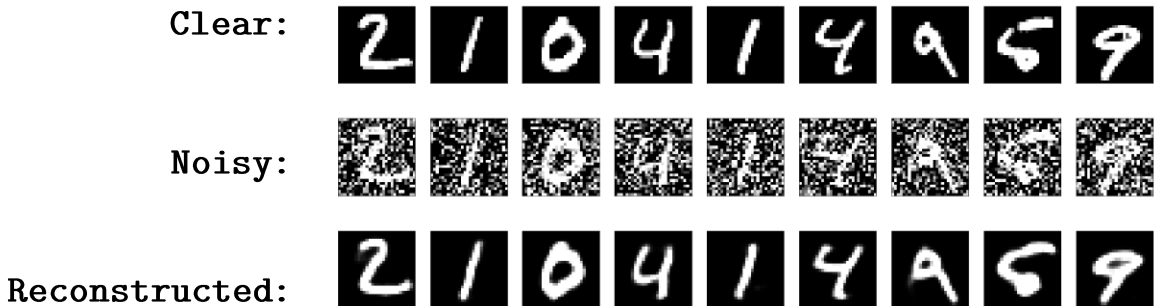# MIW (eng) – mini-project 8 (20.05.20)

Fraud investigations: reconstructing noisy images with an autoencoder

1. **Backstory** 🧑‍🦰🙍‍♀️

   FBI has been investigating the biggest fraud that has ever happened in history. Everything was engaged from satellite images to spying on people's browser history and a suspicious website was found. But it had a password. A detective Angela took a picture of some numbers and she was convinced that was indeed a password, but her phone was old to make good pictures. Now FBI is asking PJATK students to help them clear out the pictures.

2. **Get and prepare data to help FBI**

   Load the MNIST dataset of handwritten digits. You will be using only X subsets (feature subsets). Normalize them to be in the range from 0 to 1, reshape to fit the input layer of the convolutional neural network, add noise to those images and make sure the values are still in the range 0 to 1.

3. **Build, train and test:**

   Build a convolutional neural network according to this structure:

Layer (type)	Output Shape
InputLayer2D (shape)	(None, 28, 28, 1)
Conv2D (32 filters, kernel (3, 3), relu, padding same)	(None, 28, 28, 32)
MaxPooling2D (pool (2, 2), padding same)	(None, 14, 14, 32)
Conv2D (32 filters, kernel (3, 3), relu, padding same)	(None, 14, 14, 32)
MaxPooling2D (pool (2, 2), padding same)	(None, 7, 7, 32)

```
Conv2D (32 filters, kernel (3, 3), relu, padding same)        (None, 7, 7, 32)
UpSampling2D (pool (2, 2), padding same)                      (None, 14, 14, 32)
Conv2D (32 filters, kernel (3, 3), relu, padding same)        (None, 14, 14, 32)
UpSampling2D (pool (2, 2), padding same)                      (None, 28, 28, 32)
Conv2D  (1 filter, kernel (3, 3), sigmoid, padding same)      (None, 28, 28, 1)
=============================================================================
```

Compile with **adadelta** optimizer and **binary_crossentropy** loss function. Train with noisy images as a feature dataset and original images as a label dataset. Make predictions and visualize original and reconstructed images. Don't forget to visualize, FBI wants numbers. It's important. Good luck!

Submission and grading:

1. Visualization of original vs. reconstructed images will count as 3 points. The rest is for 7 points (2 and 3 points for the 3rd deadline accordingly).
2. Send a link to your notebook on https://colab.research.google.com/ to my email ihalych@pja.edu.pl.

Deadline rules:

1. By midnight 20.05.20 – max amount of points is 10
2. By midnight 27.05.20 – max amount of points is 10
3. By midnight 3.06.20 – max amount of points is 5
4. Later – 0 points

Documentation: https://keras.io/api/