Exercise 3b

Advantages:

- it is often faster and simpler to obtain sparse representations via autoencoders
- huge reduction in parameters(example: in case of natural images)
- sparsifying non-linearity
- the estimate of the expectation E $[h_j(x; W, b)]$ is very noisy in direct optimization but autoencoder denoises the data
- easier to implement
- faster to optimize (no need to keep track of source codes)
- good initial guess for si, optimize from there --> save iterations of source optimization
- can be trained by backpropagation

Disadvantages:

- 2 layers to train
- (more parameters)
- no control of regularization
- bad encoder --> bad decoder