



Assignment 5: Autoencoders

Discover & Learn


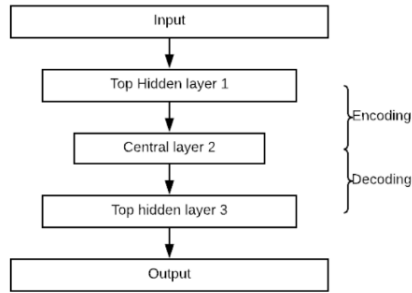
Important Information

	<p>School of Engineering Technology and Applied Science</p> <p>Information and Communication Engineering Technology</p> <p>Unsupervised and Reinforcement Learning (COMP257)</p> <ul style="list-style-type: none">• Autoencoders (15%)• Due Date: Friday of Week 8 by 11:59 pm EST (late penalty at 10 points per day)• Upload your assignment here: Assignment 5: Autoencoders (d2l/common/dialogs/quickLink/quickLink.d2l?ou=694716&type=dropbox&rdoc=CENCOL-3439545)
---	--

Instructions

	<ul style="list-style-type: none">• This assignment requires students to work in teams of two. In an odd head count in the class size, there will be one team with three members.• Your team is free to choose any toolkits to solve the problems at hand (e.g., TensorFlow, Sci-Learn, etc.)• All written reports and codes are to be maintained on a repository of your choice such as Github. The course instructor will discuss and exchange with you information to get access to your code.• IMPORTANT NOTES:<ul style="list-style-type: none">◦ 1 point will be deducted for each incident that does not conform to the requirements (e.g., code not properly formatted, comments not relevant to support documentation of code, missing code documentation, etc.).◦ All points will be deducted for submission of <i>nonsensical code</i> (i.e., code that doesn't contribute to the relevancy of the task at hand). This is question-specific.
---	--

Questions

	<ul style="list-style-type: none">• Read the question below carefully.
<p>Question 1</p> <p>[100 points]</p>	<p>In this assignment, you will implement an autoencoder that encodes and decodes data using the Olivetti faces dataset.</p> <ol style="list-style-type: none">1. Use the training set, validation set, and test set from Assignment 3 (Hierarchical Clustering) for this Assignment. [0 points]2. Use PCA preserving 99% of the variance to reduce the dataset's dimensionality as in Assignment 4 (Gaussian Mixture Models) and use it to train the autoencoder [0 points]3. Define an autoencoder with the following architecture: <div data-bbox="587 1301 999 1592"></div> <ol style="list-style-type: none">a) Use k-fold cross validation to fine tune the model's learning rate and hyperparameter of the regularizer. Due to the long training requirements, for the number of hidden units, try two or three different values for each hidden layer. [75 points]b) Discuss the rationale with respect to the activation functions and loss function used in your model. [10 points] <ol style="list-style-type: none">4. Run the best model with the test set and display the original image and the reconstructed image. [15 points]