




Final Project

 Discover & Learn


Important Information

	<p>School of Engineering Technology and Applied Science</p> <p><i>Information and Communication Engineering Technology</i></p> <p>Unsupervised and Reinforcement Learning (COMP257)</p> <ul style="list-style-type: none">• Final Project (25%)• Due Date: Wednesday of Week 14 by 11:59 pm EST (late penalty at 10 points per day)• Upload your final project here: Final Project (d2l/common/dialogs/quickLink/quickLink.d2l?ou=694716&type=dropbox&rcode=CENCOL-3439547)
---	---

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">• SYNOPSIS ¹ <p>Your team has been approached by a law enforcement security company ABC that would like to incorporate facial recognition and identification into their system. ABC has provided your team with the umist_cropped.mat (https://www.google.com/url?q=https://cs.nyu.edu/~roweis/data/umist_cropped.mat&sa=D&source=editors&ust=1619725622992000&usg=AOvVaw0_BujtSzMDSbT) dataset to train your system.</p>
<p>Requirements</p> <p>[100 points]</p>	<ol style="list-style-type: none">1. Split the training set, a validation set, and a test set using stratified sampling to ensure that there are the same number of images per person in each set. Provide your rationale for the split ratio [5 points]2. Discuss the rationale behind how your team preprocess the data. Include the mathematical equations used and any dimensionality reduction applied to the instanced and discuss its relevance to the problem at hand. <i>Note that your team will receive more points if you perform data preprocessing that help improve the eventual training process.</i> [20 points]3. Select a clustering technique taught in this course and apply it on the training instances. Provide the rationale behind your team's choice of clustering technique and how your team tuned the parameters for the technique implemented. [30 points]4. Clearly discuss - with illustrations - the architecture your team has selected for training and predicting the test instances. Discuss the rationale behind your team's choice of activation functions, loss function, and how you tuned the hyperparameters of the network model. [30 points]5. Discuss the results of the trained system. [5 points]6. Submit a video that demonstrates your system. Discuss the decisions your team made, the challenges your team encountered, how your team resolved the problems, and the results. [10 points]

¹ Note that this is just a hypothetical synopsis for academic purposes.