

**SCHOOL OF COMPUTER TECHNOLOGY**

AASD 4005 Adv. Mathematical Concepts for Machine Learning

**Project**

**Presentation: Week of Nov 21st**

**Due date to submit the full work: 30th of Nov**

**Students to form groups of 5 or 6, this is a group project.**

**Project: Part 1**

**Building an Image Enhancer using Fourier Transform**

In this project, students are expected to explore the Fourier Transform to enhance the image such as blurring, sharpening, edge detection, noise suppression, etc.

Students are required to build the enhancer module using Fourier Transform with the following functionality:

1. Blurring using low-pass filter
2. Sharpening using high pass filter
3. Edge Detection
4. Noise suppression

Students are required to create a python class, and can implement these functionalities as different methods.

Steps that can be followed ( Please note these are only recommended steps, students are free to use other steps) :

1. Convert Image from spatial domain to frequency domain
2. Apply appropriate filter/processing on frequency domain to achieve the task.
3. Convert back to the spatial domain from the frequency domain.
4. Apply the required post-processing to improve the results.

**Dataset:**

No specific requirement on the database, students can use 5 different images to enhance.

**Part 2: Building a classifier**

In this project students are expected to build **one of the 2** (your choice which one to do) classification models:

1. A classifier that classify text samples into different categories (NLP)
2. A classifier that classify images into different categories

Students are required to build the classifier using 2 of the following models

1. Logistic regression
2. Random Forest Classifier (and/or SVM)
3. Other models but not Neural Networks

For the NLP cases, here are few examples/suggestions, building a model that

1. classify emails (or text messages) into spam and not spam
2. classify news into different categories
3. classify IT tickets into different categories (request types)
4. Assign tasks (requests) to different personal based on expertise
5. …

Note: Students are free to choose their dataset and decide what you want to work on.

If there are groups that like to work on other type of data than text or images (e.g. time series, signals,…), that also would be possible, please talk to me.

**Database:**

No specific requirement on the database, of-course it should have at least 2 categories.

These parts are common for both parts:

**Report:**

Put your results in a report, you are required to use the template I have provided, but of course you can add new sections. Your report needs to include the following sections (not limited to):

* Names of the group members, date, title (if you have any),
* A table of responsibilities and tasks each member had and performed (need to be clear who did what)
* The Problem statement,
* The Database,
* The model you picked to solve the problem,
* Results, the model performance (test, valid), the loss, predictions…

(like use of confusion matrices etc…)

* Conclusions

In your results please comment and discuss the followings:

1. Evaluate the model, how?
2. Tuning your models, for example how your model change when the number of estimators (decision trees, if you used this method) changes?
3. Compare the performance of the different model you have trained (e.g. the random forest vs SVM etc… if you trained these)
4. Are any of your models are overfitted? Why?

Important Note:

1. Make sure to document your findings and back up your conclusions using the results, use visualization.
2. Note that you have to present your progress/results during the last week of class, (Note that this date can-not be moved, department need the grade by end of that week), This is a **10-15 min** long presentation.
3. There will be a draw on whom to present on which day.
4. Only one person from each group to submit (report to include all group members),
5. Your submission need to include: your **data, code, report and your slides**
6. Your codes should run without error, If I test them,
7. Compress all of your files into a zip (or similar) files and only submit that single file (one zip file)
8. You can submit few times, but I only consider the latest submission.
9. Note: You will be marked on the followings:
   * Your general quality of the report you submitted
   * Your results/conclusions
   * Your presentation quality and kills, (smoothness and flow (problem>>methodology>>results>>conclusion), clarity, slides, time…)