

**SCHOOL OF COMPUTER TECHNOLOGY**

AASD 4010 DL I

**Project: Building a NN model**

**Presentation: Last week of class on** Wed, or Fri

**Due date to submit the full work:** 1 week after the last day of class (Friday end of the day)

In this project students are expected to design, develop, train and evaluate a Deep Learning model.

It can be one of the following models:

1. A predictive model
2. A classifier
3. Others? (sentiment analysis, text prediction, other…please discuss with me)

The model will be a RNN, ANN, combination of these, NLP or use of transfer earning.

Few examples: Building a model that

1. The price for this house
2. Detect fraud activity (classification)
3. Predict the price of oil for the next year
4. Predict revenue per day
5. Classify text (NLP)
6. Sentiment analysis
7. …

**Group, roles and responsibilities:**

This is a group project; it is expected that you form your groups in the first week of class (by Wednesday). Every group need to identify a Project Manager (students can decide among themselves or vote for this position). Group size should be minimum of 4 students (Please discuss it with me if you need a different arrangement).

Some of Project Managers responsibilities include coordination and management of the project, communications, monitor timelines, running meetings and take minutes of meeting for each meeting (team need to meet at least 3 times a week). It is expected that project managers to communicate issues within the team to your instructors and find solutions in a timely manner.

**Database:**

No specific requirement on the database, of-course it should have at least 2 categories (for classification cases).

You are free to choose your 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 time series, please talk to me, to make sure it is possible before start the work.

Mainly the dataset for this project can include:

1. Time series (example price of a product over time)
2. Text (e.g. product reviews)

**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):

* Background and problem statement,
* Plan of attack, what is your approach toward this?
* 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…)

Also include the steps you took to tune your hyperparameter, architecture etc… (tell the whole story along the way)

* 1. Did you have a benchmark model?
  2. Did you plan of attack made sense, now you have a better idea?
  3. Do you do differently next time?
* Conclusions

In your results please comment and discuss the followings:

1. Basic data analysis, e.g. data imbalance, outliers, how to deal with them?
2. Evaluate the model, how?
3. Tuning your models, for example how your model change when the model structure, architecture, hyperparameters are changes?
4. Are any of your models are overfitted? Why?

What will be in your final submission:

1. Your report

* Background and problem statement,
* Plan of attack, what is your approach toward this problem?
* The Database (investigate the dataset, present findings),
* The model you picked to solve the problem,
* Results, the model performance (test, valid), the loss, predictions…

(like use of confusion matrices etc…)

Also include the steps you took to tune your hyperparameter, architecture etc… (tell the whole story along the way)

* 1. Did you plan of attack made sense, now you have a better idea?
  2. Do you do differently next time?
* Conclusions

1. All of your codes
2. Example input data with model predictions on them (only few)

Presentation:

Note that you have to present your progress/results on date specified, (you may not have completed all the work by presentation date, the final due date for delivery of report is later). Note that this date can-not be moved, department need the grade by end of the week, This is a **15 min** long presentation.

(During presentation you don’t need to have everything ready, you present what you have achieved until then)

There will be a draw on whom to present on which day.

Important Note:

1. Make sure to document your findings and back up your conclusions using the results, use visualization (this is important).
2. Only one person from each group to submit on BB (report to include all group members),

If D2L give you problems, due to size etc…, you can share it through an online drive like google.

1. Your submission need to include: your **data, code, report and your slides.**
2. If we test them, your codes should run without error.
3. **Important:** Compress all of your files into a zip (or similar) files and only submit that single file.
4. You can submit few times, but I only consider the latest submission.

Evaluation:

You will be marked on two different component, individual contribution and group work. Both of these component will be considered in: 1. The actual work, 2. Report, 3. Presentation, 4. Team work, communications and other…

Team will divide the work in a reasonable way, work will commence, and will be monitored by managers but also everyone on the team.

It is completely fine that team members do the same work, using different approaches or techniques (e.x. two or more people may train a model but with different architecture or parameters, another example: several people may contribute in writing a conclusion in the report)

Work will be mark on these items:

* + Your general quality of the work
  + Your approach and methodologies
  + Your results/conclusions
  + Your general quality of the report you submitted
  + Your presentation quality and skills, (smoothness, clarity, slides, being on time and flow:

e.g.: Background and Problem>>Data>>task and responsibilities>>methodology>>results>>conclusion

(I understand you may not have the final results by time of submission)