

## Mid-Term Project Requirement and Datasets

The assignment is worth 15% of your final grade.

### Why?

This project as a warm-up aims to explore an interesting but small machine learning problem of your choice in the context of a real-world data set. Your class project must be about new things you have done this semester; you can't use results you have developed in previous semesters.

We hope you will learn how to explore dataset in the wild and apply machine learning ideas to the dataset for insights. It works like this:

1. You find your dataset and machine learning ideas.
2. In phase 1, you put your idea in a one paragraph (no more than one page) and submit for review.
3. In phase 2, you will implement the idea with our feedback. In the worst case, your idea may be trimmed to be practical or changed to fit the course scope.

### Requirement

- No deep learning. I know it is fun, but you shall learn to walk first.
- The mid-term project must be done individually. Below, you will find some project ideas.
- If you are having trouble writing a proposal, feel free to consult with a TA or the instructor. Once you submit your proposal, we will give you feedback. Of course, the final responsibility to define and execute an interesting piece of work is yours.

**NOTE: your final project will be a team project and you may want to bring what you learned in this mid-term project to your team.**

### The Problems Given to You

| Easy Idea  | Intermediate Ideas                                     |
|--|--|
| 1. Cartoonify Image with Machine Learning              | 1. Music Genre Classification Machine Learning Project |
| 2. Iris Flowers Classification Project                 | 2. Bitcoin Price Predictor Project                     |
| 3. Emojify – Create your own emoji with Python         | 3. Uber Data Analysis Project                          |
| 4. Loan Prediction using Machine Learning              | 4. Personality Prediction Project                      |
| 5. Housing Prices Prediction Project                   | 5. Handwritten Character Recognition                   |
| 6. MNIST Digit Classification Machine Learning Project | 6. Xbox Game Prediction Project                        |
| 7. Stock Price Prediction using Machine Learning       | 7. Credit Card Fraud Detection Project                 |
| 8. Titanic Survival Project                            | 8. Sign Language Recognition                           |
| 9. Wine Quality Test Project                           | 9. Barbie with Brains Project                          |
| 10. Fake News Detection Project                        | 10. Customer Segmentation                              |

See details at <https://data-flair.training/blogs/machine-learning-project-ideas/>

You can find other ideas and dataset by web search:

<https://www.google.com/search?q=machine+learning+research+project+for+beginners>

## What you need to submit

### Phase 1: Project Proposal (Due Oct 2 11:59pm ET)

You are encouraged to use one of the suggested data sets, because we know that they have been successfully used for machine learning in the past. If you prefer to use a different data set, we will consider your proposal, but you must have access to this data already, and present a clear proposal for what you would do with it.

**Page limit:** Proposals should be at least one paragraph (one page maximum).

Include the following information:

- Project title
- Data set
- Project idea. This should be approximately one paragraph.
- Software you will need to write.
- Articles to read. Include 1-3 relevant articles including scientific journal papers, weblogs or Kaggle reports. You will probably want to read at least one of them before submitting your proposal.

### Phase 2: Mid-term Report and Code (Due Oct 23 11:59pm ET)

This should be a 2-3 pages short report in the form of a [NeurIPS](#) paper. You can also use the word document provided by instructor. It should consist of background, method, experiment, conclusion and references. The sections on the experiments and conclusions will have whatever results you have obtained. If you expect to work on it in future (Final project), you can have a discussion section which covers you plan or results you hope to obtain.

**Page limit:** mid-term report should be 2-3 pages (4 pages maximum).

Grading scheme for the project report:

- 30% for Data acquisition and Exploratory Data Analysis.
- 40% for proposed method and experiments conducted to obtain insights from the data.
- 20% on Data Visualization including graphs and charts.
- 10% on code. Is it runnable? Can it produce the same charts as shown in the report?

## Appendix: Report Sections

- Introduction, including a summary of the project, related work/methods, and the results of this project.
- Problem description, including a detailed description of the problem you address in this project.
- Methodology, including a detailed description on the methods you developed or used.
- Experiments, including, experiment setting, results, and your observations.
- Conclusion and future work, including a summary of the main contributions of this project and potential future directions.