The project will focus on reproducing and improving a research paper.

The paper should focus on the **task of time series classification or extrinsic regression**. It can be selected from the provided list below, or from any of the top A/A* AIML conferences or journals (eg NEURIPS, ICLR, KDD, ICML, ECML, AAAI, IJCAI-ECAI, CIKM; Data Mining and Knowledge Discovery Journal, Machine Learning Journal, JMLR, etc). The paper should provide data and code as open source. **If you select a paper from outside the provided list, you need to get approval from the TA and lecturer before the project starts.**

Before proceeding with a paper, you should verify that you have access to the data and the code, and that you can run the code.

The focus of the project is to use the data and code from the published paper, check if the published results stand up to scrutiny, report on lessons learned, as well as propose a few ways to improve the results.

Deliverable: A written report to be submitted via Brightspace (6 pages of content + 1-2 pages for references).
The report should take the form of a short paper, max. 6 pages double-column in IEEE format (Overleaf template). Please make your own Overleaf/Latex repo by downloading this template and customising it. At the end of the group report, 1 extra page should be used to describe: the project work plan, work division among members and the final contribution of each member (so 6 pages of report content + 1 page for work plan and work division + 1-2 pages for references). Once complete, export to PDF and submit your report to Brightspace.

Evaluation criteria for grading the report:

1. [10/100] Motivation for selecting the paper and impact of the work on the AIML community.
2. [10/100] Approach taken to reproduce the results.
3. [30/100] Lessons learned from attempting to reproduce the results (e.g. what works as published, what doesn't, any other interesting findings).
4. [30/100] Proposed improvements and evaluation of improvements (e.g., show experiments to test how effective are the proposed improvements and discuss).
5. [20/100] Open source github repo with a readme providing instructions for reproducing the results in your project. Include the link to your repo in the report.

Recommended reading:
Several AIML top conferences have reproducibility tracks. You can read more here about the importance of reproducibility:
https://onlinelibrary.wiley.com/doi/10.1002/aaai.70002
https://reproml.org/call_for_papers/
https://ecir2025.eu/call-for-reproducibility-papers/#:~:text=Reproducibility%20is%20key%20for%20establishing,as%20part%20of%20the%20submission.
https://sigir2026.org/en-AU/pages/submissions/reproducibility-track

The importance of proper baselining in ML research:
https://www.researchgate.net/publication/392760978_An_Empirical_Evaluation_of_Foundation_Models_for_Multivariate_Time_Series_Classification

Empirical evaluation papers for TSC and TSER (to give an idea of methods that work well on each task and the current state-of-the-art):
Benchmarking for TSC: Bake off redux: a review and experimental evaluation of recent time series classification algorithms
Benchmarking for TSER: Automatic Feature Engineering for Time Series Extrinsic Regression: a Comparative Study of Signal Processing Libraries

Suggested list of papers for the project:

Dictionary & Kernel-Based Methods (The ROCKET Family)

1. **ROCKET**
   - *ROCKET: Exceptionally Fast and Accurate Time Series Classification Using Random Convolutional Kernels* (Data Mining and Knowledge Discovery 2020)
2. **MiniRocket**
   - *MiniRocket: A Very Fast (Almost) Deterministic Transform for Time Series Classification* (KDD 2021)
3. **MultiROCKET**
   - *MultiRocket: Multiple Pooling Operators and Transformations for Fast and Effective Time Series Classification* (Data Mining and Knowledge Discovery 2022)
4. **Hydra**
   - *Hydra: Competing Convolutional Kernels for Fast and Accurate Time Series Classification* (Data Mining and Knowledge Discovery 2023)

Interval, Symbolic, & Feature-Based Methods

1. **Quant**
   - *QUANT: A Minimalist Interval Method for Time Series Classification* (Data Mining and Knowledge Discovery 2024)
2. **MrSQM**
   - *Fast Time Series Classification with Random Symbolic Subsequences*
3. **Catch22**
   - *catch22: CAnonical Time-series CHaracteristics* (Data Mining and Knowledge Discovery 2019)
4. **DrCIF**
   - Unsupervised Feature Based Algorithms for Time Series Extrinsic Regression

Convolutional & Deep Learning Architectures

1. **FCN (Fully Convolutional Network)**
   - *Time Series Classification from Scratch with Deep Neural Networks: A Strong Baseline* (IJCNN 2017)
2. **InceptionTime**
   - *InceptionTime: Finding AlexNet for Time Series Classification* (Data Mining and Knowledge Discovery 2020)

3. **LITETime**
    - *Look Into the LITE in Deep Learning for Time Series Classification* (arXiv 2024)
4. **ConvTran**
    - *Improving Position Encoding of Transformers for Multivariate Time Series Classification* (2023)