KF5012 Software Engineering Practice

AI path project brief

Introduction.

This is the brief for the group project assessment of the Artificial Intelligence (AI) stream of KF5012 Software Engineering Practice. To fully understand this brief you will need to also read the Assessment Specification (also known as the “umbrella assessment”). That document lays out general principles, how the group-work works, and the marking schemes for the different types of mission. This document focuses on defining the technology to be used, the basic parameters of the project, and specifics of the “missions” associated with this brief. When you create your team’s project proposal, you will need to choose which missions your team will be attempting and identify which team member is responsible for which mission(s).

The Technology

You are required to deliver a project using machine learning methods and tools to tackle an AI problem and report on your experimental investigation.

The AI problem

The AI problem could be (but not limited to):

- Object detection in images.
- Handwriting Recognition.
- Classification of text, image, audio etc.
- Separation of audio data.
- Forecasting a time-series.
- Diagnosis of (publicly available) medical data/images.
- Voice recognition.
- Chatbot for businesses.
- Personality recognition through images.
- Facial recognition.

You may come up with your own AI project idea, but you need to discuss it with the tutor, and have it approved before you start.

The Missions

This AI module aims to give students experience working with machine learning methods and their application to the development of machine learning solutions to

solve artificial intelligence problems. The module will also prepare students for further research projects in the coming 3rd year.

There are several **main missions** (required missions) and several **side missions** (optional missions). There is one double mission which the whole team will work on, which counts as two normal (single) missions. Please see the Assessment Specification for information on how many missions you need, and how these should be distributed between individual missions and group missions.

Please note, the specialist teaching for the AI path focuses on the AI methodologies and their implementations. It will give you enough to get started on most topics, but you will need to go beyond this to gain high marks. If you choose these missions, it is assumed you either already have the skills, or are willing to spend some time developing those skills. In all missions, going beyond expectations, will probably involve extra study.

The six main missions are:

- Project Ideation
- Solution Design
- Baseline Implementation
- Iterative Development (double)
- Solution Testing
- Project Management

The 5 side mission are:

- Additional Data
- Mathematical Review
- Solution Review
- User Interface
- Ethical Evaluation

Details of all missions can be found below.

Main Mission	Single Mission	Individual Responsibility

Project Ideation

To start the project, you will need to clearly state the problem the project aims to tackle, why the problem is important, who in the world has tackled the problem before, and what approaches did they take, and solutions did they try. The project ideation will also include a project statement, outlining what you plan to do.

The output of this mission will be a project ideation document including the following **elements**:

- *Problem Statement* - statement of the problem you have chosen.
- *Problem Motivation* - justification for why the problem is important.
- *Literature Review* - discussion of existing solutions and related work.
- *Project Statement* - outline of what you plan to do.

The project ideation document should exhibit the following **qualities**:

- *Soundness* - problem motivation uses valid reasoning and good judgement.
- *Coverage* - literature review covers a suitable amount of related work.
- *Depth* - related work is discussed to a suitable level of detail.
- *Communication* - ideas are communicated effectively.
- *Clarity* - clear use of error free English.

To be **outstanding,** you may extensively review existing solutions and related work, summarise their weaknesses and strengths, and justify the plan in your project statement.

To be **exceptional**, either your problem should be novel with sound justification for why it is a worthwhile problem, or the plan in your project statement should include a novel approach / solution with sound justification for why it is worth exploring.

Main Mission	Single Mission	Individual Responsibility

Solution Design

To start development of a solution, you will need to work the project statement into a specification of an initial solution with clearly stated requirements for the solution and related data. Your design choices will need to be well justified and requirements will need to be communicated clearly enough to be implemented.

The output of this mission will be a solution design document including the following **elements**:

- *Solution Diagram* - diagram of proposed solution.
- *Solution Description* - description of proposed solution.
- *Data Description* - description of the dataset required.
- *Solution Motivation* - reasoning behind the proposed solution.

The solution design document should exhibit the following **qualities**:

- *Understanding* - proposed solution and data are suitable for the problem.
- *Technical* - technical details are communicated effectively.
- *Soundness* - reasoning for design choices is sound.

To be **outstanding,** you may provide evidence of a deep understanding of the chosen problem and existing solutions and clear justification for your design choices.

To be **exceptional** your solution design will include ideas beyond the related work, or your design will involve the application of existing methods in a novel way, with clear justification for both.

Main Mission	Single Mission	Individual Responsibility

Baseline Implementation

To lay the foundation for iterative development, the first solution you will implement will be a baseline solution which you will use to test your infrastructure and to set a benchmark for performance you can compare improved solutions to. It will be beneficial to get your environment and tooling right, and not focus too much on the machine learning element of the system.

The output of this mission will be a download of an early version of your software repository including the following **elements**:

- *Baseline Pipeline* - preparation of initial version of the dataset.
- *Baseline Solution* - implementation of an initial solution.
- *Baseline Evaluation* - evaluation of the baseline solution.
- *Environment Setup* - everything required to run the code.

It is understood that your iterative development will almost certainly be built from the foundation of the baseline implementation. To allow your baseline implementation to be marked, you **must** keep a snapshot of your project at the point where this is complete.

The baseline implementation should exhibit the following **qualities**:

- *Suitability* - should be enough to test the pipeline diagram works.
- *Benchmark* - should produce a reasonable benchmark for performance.
- *Readability* - code should be clear and understandable.
- *Error-free* - should run without errors.

To be **outstanding**, you will implement the baseline in a way so that the learning parts of the system are encapsulated and you can test everything around it.

To be **exceptional**, you will implement and evaluate the baseline in a way that enables rapid iterative development of solutions beyond the baseline.

Main Mission	Double Mission	Team Responsibility

Iterative Development

To move the solution beyond the baseline solution, the whole team will be involved in the iterative development of solutions. This will give everyone in the team hands-on experience with machine learning development, and you should explore a variety of data pre-processing techniques, machine learning models, and parameter tuning methods towards improved predictive performance. It is important to document the development process and to have clear justifications for the decisions made.

The output of this mission will be the final version of your software repository and a solution development report including the following **elements**:

- *Pipeline Evaluations* - evaluation of different data pre-processing techniques.
- *Model Evaluations* - evaluations and comparisons of different models.
- *Parameter Justifications* - explanations for parameters and tuning methods.
- *Reproducible Code* - code for reproducing all experiments and results.

The solution design document should exhibit the following **qualities**:

- *Progress* - progress was made towards solutions with improved performance.
- *Creativity* - problem solving was creative during development.
- *Rigour* - process was rigorous and design choices were well justified.
- *Communication* - work is well documented and results well organised.

To be **outstanding,** you may conduct extensive investigation into model performance, exhibit outstanding creativity in problem solving, and exhibit outstanding rigor in the justification of your design choices.

To be **exceptional** you will identify and develop novel solutions leading to improved performance.

Main Mission	Single Mission	Individual Responsibility

Solution Testing

To ensure your solution is robust, you will need to work to identify potential weaknesses in the system and design test case intended to expose those weaknesses. This might involve trying to break the system or confuse the model with a variety of data inputs, and should ultimately lead to a more robust system and model.

The output of this mission will be a solution testing plan and report including the following **elements**:

- *Test Plan* - state what is going to be tested and why.
- *Test Cases* - specify test cases for identifying flaws in the solution.
- *Test Report* - perform tests and report results.
- *Recommendations* - create a set of recommendations based on results.

The solution testing plan and report should exhibit the following **qualities**:

- *Justification* - the test plan is well justified and worthwhile.
- *Fitness* - test cases are suitable for the intended purpose.
- *Insight* - recommendations are useful and improve the system.
- *Clarity* - reasoning is clear and results are well analysed.

To be **outstanding,** you may perform extensive tests of the system and identify multiple weaknesses leading to insightful recommendations for making the solution more robust.

To be **exceptional**, the test cases should include attempts to confuse the underlying model and lead to insights that improve model robustness.

Main Mission	Single Mission	Individual Responsibility

Project Management

For the project to run smoothly, someone will need to take responsibility for the project management aspects of the project. They will need to prepare the project management documentation and adapt and update the documentation as the project progresses. The manager needs to keep a record of meetings, decisions and agreements made, and have a clear picture of progress; and will be responsible for settling disputes, handling contingency for unforeseen problems and anything else that affects the effective working of the team.

The output of this mission will be project management documentation including the following **elements**:

- *Code of Conduct* - code of conduct for the team
- *Skills Audit* - skills audit of the team
- *Risk Analysis* - summary of the risks involved in the project
- *Task Management* - task breakdown, prioritisation and allocation
- *Master Schedule* - timeline with deliverables and milestones
- *Progress Report* - logs of progress and records of communication

The project management documentation should exhibit the following **qualities**:

- *Clarity* - clear and well organised documentation.
- *Completeness* - elements completed as required.
- *Communication* - clear and effective communication.
- *Ethical* - fair and ethical treatment of team mates.

To be **outstanding,** you will deliver elements to a professionally acceptable standard, using and referencing established project management practices.

To be **exceptional**, you will be creative in the adaptation of project management practices to this particular project, identifying novel processes with justification for their use.

Side Mission	Single Mission	Individual Responsibility

Additional Data

To improve performance or make your solution more robust, you might consider including additional datasets beyond the data specified in the design stage of the project. The extra datasets are expected to be either collected by yourself or downloaded from any public web sources with clear justification for why.

The output of this mission will be any additional datasets and a data preparation and integration report including the following **elements**:

- *Data Motivation* - reasons for introducing additional datasets.
- *Data Preparation* - methods used to prepare and integrate additional data.
- *Impact Report* - analysis and discussion of the impact of additional data.

The data preparation and integration report should exhibit the following **qualities**:

- *Justified* - reasons for introducing additional data are well justified.
- *Suitable* - choice and preparation of data are suitable for the problem.
- *Rigour* - evaluation of the impact uses sound and relevant experimentation.

To be **outstanding,** you will provide strong and clearly communicated justifications for introducing additional data, the data will be prepared in a way that maximises its potential to improve performance, and analysis of the impact of the additional data using well established machine learning methods.

To be **exceptional**, you will be creative in your use, integration, and/or evaluation of additional data towards improved performance of your solution.

Side Mission	Single Mission	Individual Responsibility

Mathematical Review

For you to place your work within the context of existing theory, you will need to formalise the problem, potential solutions, and choices of model parameters. This mission will be a test of your mathematical understanding and your ability to derive solutions or make design decisions based on mathematical / theoretical results in addition to experimental results.

The output of this mission will be a mathematical review document including the following **elements**:

- *Problem Formulation* - mathematical formulation of the problem.
- *Solution Justification* - mathematical justification for proposed solutions.
- *Parameter Justification* - mathematical justification for parameter choices.

The solution design document should exhibit the following **qualities**:

- *Correctness* - correct use of mathematics and consistency throughout.
- *Depth* - depth of understanding and references to relevant work.
- *Communication* - clarity and effectiveness of communication.

To be **outstanding,** you may provide evidence of a deep understanding of the mathematical formulation of the problem, the mathematical derivation of solutions, and mathematical justification of parameter choice where possible.

To be **exceptional** you will provide mathematical formulation for a novel problem, derive a novel solution, or provide novel justification for a choice of parameter. You will need to make it clear why it is novel.

Side Mission	Single Mission	Individual Responsibility

Solution Review

To consider how your solution might generalise to new problems and environments, you will need to compare your results to other results in the existing literature. This mission will require you to identify suitable comparisons and explain why they are fair or not.

The output of this mission will be an solution review document including the following **elements**:

- *Comparison* - comparison of your results to others in the literature.
- *Discussion* - justification for and fairness of comparisons.
- *Recommendations* - recommendations based on comparisons.

The solution review document should exhibit the following **qualities**:

- *Suitability* - the articles for comparison were suitable and well justified.
- *Depth* - the comparisons were thorough and fairness was considered.
- *Insight* - the recommendations were insightful.
- *Clarity* - comparisons were presented and explained clearly.

To be **outstanding,** you will choose a number of suitable articles for comparison with well justified reasons, the comparisons will cover key similarities and differences, and have strong arguments for whether the comparisons are fair or not.

To be **exceptional** your comparisons will lead to novel insights and recommendations that go above and beyond what already exists in the literature.

Side Mission	Single Mission	Individual Responsibility

User Interface

To deploy your solution in the real world, you will need to develop an interface so users and user applications can interact with your solution in a standardised way. This mission will focus on making your solution accessible to the world.

The output of this mission will be the code required to run a user interface and documentation including **elements**:

- *API Design* - design of the API with justifications.
- *UI Design* - design of the user interface with justifications.
- *Solution Integration* - explanation for how your solution is integrated.
- *User Testing* - log of tests of the API/UI with justifications and discussion.

The API/UI and documentation should exhibit the following **qualities**:

- *Usable* - the design is usable.
- *Error-free* - works without errors.
- *Robust* - is robust to a range of inputs.
- *Efficient* - is efficient using good software engineering practices.

To be **outstanding,** you may develop an API/UI that is highly performant, robust and developed using clear code and good software engineering and design practices.

To be **exceptional**, your API/UI will include novel design choices, implementation details, or an extension/use of the API in a novel way.

Side Mission	Single Mission	Individual Responsibility

Ethical Evaluation

For your solution to see real-world use, all possible ethical impacts must be considered. In this mission, you may discuss the potential societal impact that your project could have, analyse and account for sources of bias, or develop an interpretable mechanism to understand how your model arrives at a classification.

The output of this mission will be a report on the ethical considerations including elements:

- *Bias Analysis* - detail bias found in the input data set and/or output model.
- *Ethical Discussion* - report other potential ethical issues of the project.
- *Technical Solution* - solve identified problems (where possible).
- *Solution Verification* - present comparison of results.

The ethical report should exhibit the following **qualities**:

- *Sensitivity* - problems are highlighted in a professional manner.
- *Clarity* - the discussion is unambiguous.
- *Thorough* - a wide range of potential problems are considered.

To be **outstanding**, you may incorporate state-of-the-art third-party tools into your system to correct model bias during training. You will write a detailed report on your reasons for choosing the selected solution for your particular task and include a concrete discussion of how the solution works.

To be **exceptional** you will design a novel technical approach to deal with bias that has been identified within your system. You will provide a thorough mathematical discussion of how your techniques work and compare it with other possible solutions.