Project Plan

DATA11004 Data Science Project I

Please use this template to make your project plan. Each of the sections of the plan should be brief (1-2 paragraphs). The idea is to plan and document all relevant issues, make it easy to manage the project, and make any changes later, if necessary. Please keep the heading structure. During your first meeting, please fill in the items in [red]! You should complete the project plan after meeting with the client (task 2). You can update the plan during the project plan if a need arises (task 1); however, please save copies of the earlier versions of the project plan. You can remove these boxed instructions when done.

Summarising, during your first meeting, please fill out the following items to the extent possible (you can revise them afterwards):

- Section “0 General”

- Section “4 Risks”: initial risk table

- Section “5 Analysis of the project environment”

- Section “7 Rough schedule and mid-term results”: indicate at least the dates of your presentations.

- The following tasks: “Task 1: Management” and “Task 2: Agreeing about the project scope and outputs with the client”

- Section “9 Project organisation”

- Section “10 Workload and costs”: at least fill in the resourcing table for the project duration of all project participants so that you allocate everyone 115 hours.

- Section “11 Project management description”

Please make a complete project plan after you have met with the client.

## 0 General

[fill this during the first meeting]

**Project name:** X

**Project duration:** X.X.202X - X.X.202X

**Date of the plan:** X.X.202X

**Version number:** 1

# Summary

## 1 Background

Earlier development and status. Briefly describe the need (why is the project helpful?) and the current state-of-the-art solution (that you will be developing).

## 2 Project objective and results

Please describe here the project objective and the planned main result. Please provide a clear and concrete criterion for success. After reading this section and seeing your project results, you should be able to decide if you failed or succeeded! The length of this section should be one paragraph of text.

# Project strategy

## 3 Scope

Please describe here the scope of your project. i.e., what you will do and what you leave for future work.

## 4 Risks

Please list the most critical risks in the table below and how you have addressed them. Classify their P=probabilities and I=impacts (effect on the project if the threat materialises) as L=low, M=medium, or H=high. You do not need to list risks with low probability and impact. You should track the risk table during the project (has the risk materialised, has the likelihood, or expected impact changed, are there new risks, how you address the risks and so forth). In your first meeting, you should minimally add one specific risk for your project topic; please keep the risk table updated as part of your project management process.

[fill this during the first meeting]

|  |  |  |  |
| --- | --- | --- | --- |
| **Risk** | **P** | **I** | **How addressed** |
| A group member ends the project prematurely | L | M | All group members commit to the course and reserve time for the coursework in their spring schedule. The schedule is flexible, so other group members can compensate if one member leaves. |
| We do not reach all project objectives | L | H | We will first focus on making a “minimal viable product” that satisfies the minimal requirement. After completing the minimal viable product, we will start working on the more advanced features. |
| Some data science tool components will not be finished in time, which delays the project. | M | H | We will first focus on making a “minimal viable product”, which is easiest to make. During development, we will replace uncompleted parts with dummy parts (e.g., we will use a dummy classifier instead of the final classifier). Dummy parts allow us to develop and test the whole system and the other components as planned, even if one part is delayed. |
|  |  |  |  |

## 5 Analysis of the project environment

Please list your group members’ names and email addresses (Alice, Bob, Carol, and Dave in this example template) and the client. Describe any other relevant stuff about your project environment (the essential tools and infrastructure you plan to use, including collaboration tools, version control systems, programming languages, machine learning libraries, platforms etc.).

[fill this during the first meeting]

## 6 Description of the project result

You can describe the project outcomes in more detail, including technical stuff.

## 7 Rough schedule and mid-term results

Indicate here the course schedule (start of the project, the dates of your mid-term and final presentations, and what will be ready in mid-term).

[fill this during the first meeting – at least indicate the significant deadlines for your work]

# Detailed plan

## 8 Detailed timetables

List of tasks and task dependencies

Break down tasks in a Work Breakdown Structure (WBS). 100% of the project scope and deliverables should be in WBSs! At simplest, WBSs are a flat list of tasks; larger projects should have a more hierarchical tree-like structure where the whole project would be the root WBS. Each of the tasks should satisfy the following properties:

- measurable

- tangible results & short duration

- easily understood

- clear/single owner

- clear deliverables defined (e.g., not just “data exploration”; also explain *why* you explore the data and what are the *concrete outputs* of the work – see the examples below)

You can also use Gantt charts to help schedule the tasks and their dependencies.

You can find an example of the initial tasks below. Please change tasks or add more tasks depending on your project.

### Task 1: Management

[fill this during the first meeting]

**Duration:** X.X.202X - X.X.202X (whole project)

**Description:** Project management, meetings, tracking deliverables, quality control, managing risks, and any changes. Most of the work is done in weekly project meetings, as invited by calendar invitations sent by Alice at the beginning of the project.

**Owner:** Alice

**Resourcing:** Alice (10 h), Bob (4 h), Carol (4 h), Dave (4 h)

**Deliverables:** (i) Updated project plan (when needed), (ii) project meeting notes.

### Task 2: Agreeing about the project scope and outputs with the client

[fill this during the first meeting]

**Duration:** X.X.202X - X.X.202X

**Description:** We will have two workshops with the client where we agree on the project scope and deliverables and finish the project plan.

**Owner:** Carol

**Resourcing:** Alice (4 h), Bob (4 h), Carol (4 h), Dave (4 h)

**Deliverables:** Complete the project plan (this file).

### Task 3: Data cleaning and preparation (for example, change to your needs!)

**Duration:** X.X.202X - X.X.202X

**Description:** We will transform the existing student into the CSV table format and prepare scripts to transform the data into any form used by MiniZinc. We will also do a quality check of the data and address any discrepancies or other quality issues (missing values, outliers etc.).

**Owner:** Dave

**Resourcing:** Alice (0 h), Bob (0 h), Carol (0 h), Dave (10 h)

**Deliverables:** (i) Data table in CSV format, (ii) data processing scripts, (iii) report on data preprocessing (to be used later in the final report).

### Task 4: Data exploration and feature selection (for example, change to your needs!)

**Duration:** X.X.202X - X.X.202X

**Description:** We will study and visualise the data. We will test different logic solvers for student group allocation on the data and recommend how to select the solver and constraints.

**Owner:** Bob

**Resourcing:** Alice (10 h), Bob (10 h), Carol (0 h), Dave (0 h)

**Deliverables:** (i) Recommendation for the choice of a solver and constraints, (ii) report on data and its exploration and feasible constraints (to be used later in the final report).

### Task X: X

**Duration:** X.X.202X - X.X.202X

**Description:**

**Owner:**

**Resourcing:**

**Deliverables:**

### Milestones

Please list milestones with dates; see some suggestions below. Milestones are reference points in the project that you can use to measure the progress towards the goal.

* Project plan presentation X.X.202X
* Mid-term presentation X.X.202X
* Final presentation X.X.202X
* Final project report X.X.202X
* Supplying the result to the client X.X.202X

## 9 Project organisation

Please explain here the tasks and responsibilities of the team members.

[fill this during the first meeting]

Alice works as a project manager. All members contribute to the project as detailed above in the task. We can change the plan later; this work is included in task 1, “management”.

## 10 Workloads and costs

Each student will distribute 135 hours of work to the course (the course is worth 5 ECTS credits, one credit corresponding to 27 hours of work). The five two-hour lectures take 10 hours total, and if we count 10 hours for any other individual stuff (such as the learning diary), this leaves 115 hours per student or about 16 full-time workdays to be distributed in this plan (a typical workday is 7¼ hours of work). The hours allocated for each task are detailed above with the task descriptions. The table below shows the workload for each group member (H1 and H2 indicate the first and second half of the month, respectively – please feel free to change this!). Please make this resourcing table compatible with the task resourcing above. You should also record substantial other commitments for team members (work, other studies, exams, travel etc.) that may affect the project. You may not allocate more than 40 hours per week per student (this corresponds to full-time work with no additional work or studies!).

[fill this during the first meeting - at least fill in the resourcing table for the project duration of all project participants so that you allocate everyone 115 hours. Make sure that all group members have the hours available. You can change the hours later as part of your project management process.]

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | Jan H2 | Feb H1 | Feb H2 | Mar H1 | Mar H2 | Apr H1 | Apr H2 | May H1 | Total hours |
| Alice | 10 | 5 | 5 | 40 | 40 | 5 | 5 | 5 | 115 |
| Bob | 10 | 5 | 5 | 40 | 40 | 5 | 5 | 5 | 115 |
| Carol | 10 | 5 | 5 | 40 | 40 | 5 | 5 | 5 | 115 |
| Dave | 10 | 40 | 40 | 5 | 5 | 5 | 5 | 5 | 115 |
| Total hours | 40 | 55 | 55 | 125 | 125 | 20 | 20 | 20 | 460 |

## 11 Project management description

Steering practices (leadership), risk management, quality monitoring and change control, methods, and tools.

Here you can, e.g., agree on weekly project meetings where you go through the status of the WBSs and current risk table (i.e., has any of the risks realised, is the probability higher, are there new risks). You can use a traffic light system where you show the status of the old and new risks with green, yellow, or red. You can update the risk table at each project meeting, look at the project results (are they ok, quality control), and any changes to the project plan. In the example above, this work is in “task 1”.

[fill this during the first meeting]

## 12 Other issues

Please write here any other matters of relevance.

Notice that machine learning projects differ from “regular” software projects. You may want to take these differences into account in your planning. Please read Sculley et al. (2015) for some insights!

Sculley, D., Holt, G., Golovin, D., Davydov, E., Phillips, T., Ebner, D., Chaudhary, V., Young, M., Crespo, J.-F., & Dennison, D. (2015). Hidden technical debt in machine learning systems. In C. Cortes, N. Lawrence, D. Lee, M. Sugiyama, & R. Garnett (Eds.), *Advances in neural information processing systems* (Vol. 28). Curran Associates, Inc. <https://proceedings.neurips.cc/paper/2015/file/86df7dcfd896fcaf2674f757a2463eba-Paper.pdf>