

Capstone Project Plan Template

Last updated: November 13, 2023

**Instructions**: Each capstone team can use this template to capture and summarize information about the project. This can be shared with the sponsor and mentor. When submitting the plan during the course, a PDF file is preferable.

**Stakeholder Names and Roles**

*In the table below, enter information for each team member, mentor, and sponsor. On the sponsor side, there may be several stakeholders including project manager, data contact, etc.*

|  |  |
| --- | --- |
| **Stakeholder** | **Role** |
| *Name 1* | *Team member / mentor / sponsor role* |

**Project Title:** [enter title]

**Abstract**

*Enter abstract*

**Outline of the Project**

*This should discuss:*

*- the business purpose of the project*  
*- why the project is important*  
*- who are the stakeholders*  
*- important assumptions*  
*- what is in scope / out of scope (as needed)*

*Example: Legal assistants spend a lot of time reviewing documents related to court cases. This project will use an ML system that will take legal documents as input and output a summary and helpful metadata. This will save legal assistants time, reduce burnout, and allow them to focus on helping their clients. The assumption is that all input documents are legal documents.*

**Success Criteria**

*Success criteria define the needs for a successful project. These should be identified with help from the sponsor. It is important to revisit these criteria with the sponsor as the project progresses to ensure alignment.*

|  |  |
| --- | --- |
| *SC1* | *description1* |
| *SC2* | *description2* |

**Assumptions and Limitations**

*For any project, there may be assumptions* [A] *and limitations* [L] *on the data and the modeling approach. These can be documented here. Example:* [L] *Ideally, the dataset would include variable X, but we did not have access to this data, which was a limitation.*

|  |  |
| --- | --- |
| **Identifier** | **Description** |
|  |  |

**Potential Background Literature and Resources**

*List readings / repositories / etc. that might help with the project.*

**Brief Outline of the Data**

*This should discuss as much as possible:*

*- the history of the data and how it was sourced*  
*- the target variable*  
*- the raw data*  
*- potential predictor variables, if known*  
*- data aspects which may present a challenge: missing data, bad data, etc.*  
  
  
**Brief Plan of How the Data will be Modeled and Processed**

*For example, you might discuss:*

*- Will you need to build a database? How will you construct it?*  
*- Will you need to supplement with additional data? How will you find it?*  
*- Thoughts on processing raw data in refined data useful for modeling*

**Brief Plan of Modeling Approaches**

*For example, you might discuss:*

*- Specialized techniques that may be used, such as NLP*

*- Models you might try, such as gradient boosted trees*

*- Metrics that are important (e.g., F1 score / recall / precision), and explanation of why they are important.*

**Potential Concerns [C] and Blockers [B]**

*Based on what you know about the project and data, are there things that might present a challenge in moving forward?*

*Are there things you can do to clear these challenges? For example, are there resources you might need? What are they?*

|  |  |
| --- | --- |
| **Identifier** | **Description** |
|  |  |