**SE 4485: Software Engineering Projects**

Spring 2024

Project Management Plan

| Group Number | **8** |
| --- | --- |
| Project Title | **Crop Yield Prediction Model** |
| Sponsoring Company |  |
| Sponsor(s) | **Marc J. Perna (*Main Sponsor*)**  **Daryl P. Nelson (*Former, as of 03/14/24*)** |
| Students | **1. Ryan Haven (*Team Leader*)**  **2. Cameron Sutton**  **3. Melvin Sajeev**  **4. Ibrahim Barney**  **5. Nisai Sun** |

# 

# Abstract

This abstract provides an overview of a crop yield predictive analysis application sponsored by Raytheon. It outlines the project's objectives, associated risks and opportunities, chosen models, and requirements. Additionally, it addresses key considerations, including monitoring reports, standards, conduct, and constraints.

# 

# Table of Contents

[Abstract 2](#_ik4n26sfekj7)

[Table of Contents 3](#_te9z38ucfz4n)

[Project Organization / Process 5](#_fhmvzj7u0wea)

[Lifecycle Model 6](#_eohskn6xyrgk)

[Risk Mitigation 7](#_jlxtmojrqb12)

[Software / Hardware Requirements 9](#_wr4lzpxizoye)

[Deliverable Schedule 10](#_at48j4c5dpjo)

[Monitor / Control Mechanisms 13](#_vay9z7icbnwl)

[Professional Standards 14](#_tj07wlopblgj)

[Evidence of Configuration Management 15](#_b102z8ua8943)

[Configuration Management Table 1](#_3md9zrnuylqm)7

[Engineering Standards / Constraints 1](#_fzensla9gs0y)8

[References 1](#_qh53swmuy08f)9

Introduction

In today's rapidly changing world where climate change poses a looming threat to our agricultural systems and food security, our project aims to develop a robust crop yield prediction model. We will be using modern cloud and data science tools to leverage datasets from data.gov that encompass historical weather patterns and crop yield data. We endeavor to address the pressing need for understanding the trends in crop yield. The prediction model will help in informing crop rotation, watering, and other useful information to optimize crop return.

# 

# Project Organization / Process

Ryan will be the team leader for this project. The team will prioritize tasks and assign collaboratively. Ryan will make sure everything is allocated equitable and will break any ties.

Process risks:

* Out Of Date Documentation:
  + **Description**: Server/main host documentation is out of date.
  + **Process**: All documentation will be updated the day of the application update implementation.
* Team does not have access to source code locally in offline situation:
  + **Description:** team member forgets to clone/update local repository, does not have access to most up to date version.
  + **Process:** Provide a backup of the most up to date source code.
* Team member needs to be caught up to speed:
  + **Description:** Team member either misses a meeting or important update.
  + **Process:** all updates will be posted /provided in discord for viewing.
* Problems with communicating with sponsors:
  + **Description:** Problem communicating with a sponsor in a specified week
  + **Process:** Professors notified. Continue with production and prepare questions / conversation topics for the following week

# Lifecycle Model

Our project will be developed using an Iterative model. Our project can be split up into several well-defined steps (Collection, Preprocessing, Processing, Visualization, etc.) and an iterative lifecycle model will let us develop our project most efficiently.

Since we have a broad knowledge of each step but not much detailed information on the best way to implement each step, an iterative process will let us take on each step one at a time and iterate/improve on it until we have something that works smoothly and robustly. By doing this for each step of our project, we will be left with a well-polished product that should work very smoothly.

We first considered a Waterfall lifecycle but decided against it since it offers very little support for going back and changing something after it has been created. We would have been locked into our first design choice which isn’t ideal since we wouldn’t be able to effectively incorporate feedback from the sponsors.

We also considered an Agile system but realized the quick prototyping nature is not well suited for our project. A barebones prototype of our project would be a glorified “hello world” program that processes hardcoded data which didn’t seem ideal. This was our rationale that led us to the Iterative lifecycle model and we think it will be very effective for us when creating our project.

Communication between sponsors will be through email and zoom meetings. Team members will communicate with one another through discord / email / in person meetings. Team members will communicate with UTD professors via email.

# Risk Mitigation

Possible Risks

* Online data access being taken offline on the API provider’s end
  + **Action**: If data becomes unavailable by the provider, backups to the data will be created to allow access to the
  + **Mitigation**: Try to obtain and keep an offline copy of the live data as much as possible, within our budget, scope, and resources.
  + **Retirement:** Data fully backed up; ready to access and use.
* Team member is out sick
  + **Action:** If a team member is out of commission, other team members will be notified as soon as possible to allow leeway to make urgent and critical decisions
  + **Mitigation:** Fallback team members assigned to tasks, in the event that one of the team members are unavailable or unable to fulfill tasks, whether it is emergencies or events out of any of the members’ control.
* Github down for maintenance
  + **Action**: If Github goes down for maintenance team members will be notified to ensure that there is not a disruption in version control.
  + **Mitigation**: Multiple ways to reduce risk is to mirror the git repo to other platforms or a self-hosted git repository that is limited to use and viewing by team members, and possibly sponsors upon request and approval by the team.
* Lack of technical knowledge
  + **Action:** If there is a certain skill set that the team is lacking that is required for completion of the project. Then a task will be delegated out to a team member for research.
  + **Mitigation:** Planning ahead of time to foresee a possible area where there would be a lack of expertise will help with preventing a slowdown.
* Unable to display results
  + **Action:** If the team is unable to display results to sponsors in an appealing way, steps will be taken to ensure a presentation presenting the predictions concluded by the application
  + **Mitigation:** Ensure enough time to tackle the time intensive task of creating proper visualization
* Unable to create correlations
  + **Action:** If the application is unable to produce the appropriate results, examination of the data / algorithms being used will be applied
  + **Mitigation:** Proper prep work and question asking in terms of what the data can answer will be established to prevent possible issues

Possible Opportunities

* The satisfaction of having learned about and created a tool that could help with decisions related pertaining to food insecurity and crop yields.
* Farmers having a more positive outlook on AI and machine learning technology to help plan and chart out their future.

# Software / Hardware Requirements

We will be using Python, It is widely used in data science and machine learning. Python has extensive libraries and frameworks such as NumPy, and TensorFlow that are crucial for data analysis and modeling. Version Control will be Git and it is essential for tracking changes in your codebase, collaborating with others, and maintaining a version history. Geospatial Analysis (Optional), If the analysis involves geospatial data, these libraries will be helpful for reading, writing, and processing spatial data.For the Hardware Requirements, it will be a Processor (CPU), A multi-core processor (Intel Core i5 or i7) for faster data processing. Random Access Memory At least 16 GB of RAM is recommended, especially for handling large datasets. Also, sufficient disk space to store datasets, code, and model files.

*Rationale*: Python: Chosen for its readability, extensive libraries, and popularity in the data science community. Database Management: PostGreSQL. Geospatial Analysis (Optional): Useful if the analysis involves geographical aspects of crop yield. Version Control (Git): Facilitates collaboration, code tracking, and easy rollback to previous versions. GPU (Optional): Accelerates model training, especially for deep learning, significantly reducing processing time. RAM and Storage: Essential for handling large datasets and ensuring smooth execution of resource-intensive tasks.

# 

# Deliverable Schedule

| **Date/Time** | **Participants** | **Activity** | **Rationale** | **Dependencies** |
| --- | --- | --- | --- | --- |
| February 3rd, 2024 | *Group Team 8* | Git Repository Overview, Setup, and Familiarization | Make sure the whole team is on the same page on how code revisions will be processed and published | Git, GitHub, in-person meeting, computers |
| February 10th, 2024 | *Group Team 8* | Select data sets | The data sets are what make this project work, so selecting appropriate data is crucial | Related data repositories, backup servers, |
| February 17th, 2024 | *Group Team 8* | *First iterative stage of the development process* | This is the first iteration of the project, so starting from scratch with all the material and software tools at our disposal. | In-person meeting, sponsors feedback noted, access to data and code, computers |
| February 24th, 2024 | *Group Team 8* | *Second iterative stage of the development process* | After getting feedback from the first iteration demonstrated to the sponsors, changes (additions, removals, etc) will be made. | In-person meeting, sponsors feedback noted, access to data and code, computers |
| March 2nd, 2024 | *Group Team 8* | *Third iterative stage of the development process* | Again, further development made based on what was accomplish, what needs to be added, and what changes must be made based on feedback on the previous meeting with sponsors | In-person meeting, sponsors feedback noted, access to data and code, computers |
| March 8th, 2024 | *Group Team 8* | *Fourth iterative stage of the development process* | Continue with the iterative development process and prepare for the in person meeting based on criteria set by sponsors the coming week | In-person meeting, sponsors feedback noted, access to data and code, computers |
| March 22nd, 2024 | *Sponsors + Group Team 8* | In-person meeting at the Richardson campus of Raytheon | Make sure to bring everything agreed upon by sponsors and make notes and use opportunities to meet face-to-face with deep-in-the-thick-of-things developers | In-person meetup, notepads, transportation, sponsors being present (for security reasons) |
| March 29th, 2024 | *Group Team 8* | *Fifth iterative stage of the development process* | Continuing the iterative process | In-person meeting, sponsors feedback noted, access to data and code, computers |
| April 5th, 2024 | *Group Team 8* | *Sixth iterative stage of the development process* | Continuing the iterative process | In-person meeting, sponsors feedback noted, access to data and code, computers |
| April 12th, 2024 | *Group Team 8* | *Seventh iterative stage of the development process* | Continuing the iterative process | In-person meeting, sponsors feedback noted, access to data and code, computers |
| April 19th, 2024 | *Group Team 8* | *Eighth iterative stage of the development process* | Continuing the iterative process | In-person meeting, sponsors feedback noted, access to data and code, computers |
| April 26th, 2024 | *Group Team 8* | *Ninth stage of the development process* | Continuing the iterative process | In-person meeting, sponsors feedback noted, access to data and code, computers |
| May 3rd, 2024 | *Group Team 8* | *Final stage of the development process* | Likely polishing, in terms of minor bug fixes, polishing the UI/UX, and addressing any final feedback given by the sponsors of the project | In-person meeting, sponsors feedback noted, access to data and code, computers |

# 

# Monitor / Control Mechanisms

Monitoring Mechanisms:

* Weekly Progress Report
  + *Description*: Weekly report that will consist of all work completed in the current week. The report will also include what the plan for the following week will be.
  + *Rationale*: This will give a clear picture as to what work has been completed, and give an established goal as to what needs to be accomplished in the coming week.
* Attendance Report
  + *Description*: Weekly Attendance Report to be sent to instructor and TA’s. This will keep track of each team member's participation with our sponsors and within our group.
  + *Rationale*: Required reporting for senior design class. This will help keep in compliance with the requirements needed by the class.

Control Mechanisms:

* Version Control
  + *Description*: The project will be version controlled through the use of Git.
  + *Rationale*: To ensure consistency and protection of the code base. As the project continues to grow, this will help with ensuring protection of the overall production cycle of the project, reducing the risk of needing to start over.

# Professional Standards

Ensure compliance with data protection laws and regulations. Implement secure storage and processing practices to safeguard sensitive information. Collect and use data ethically, respecting the rights and privacy of individuals. Communicate the purpose of data collection and obtain appropriate consent. Document your methods, assumptions, and limitations. Transparency is essential for others to understand and trust your analysis. Structure your code and analysis to allow others to reproduce your results. This promotes transparency and helps in the validation of findings. Be aware of potential biases in your data and algorithms. Strive to address and mitigate biases to ensure fair and equitable predictions. Maintain comprehensive documentation for your project, including data sources, preprocessing steps, model architecture, and evaluation results. Encourage peer review of your work. Collaborate with domain experts and stakeholders to ensure your analysis aligns with domain knowledge and addresses relevant issues. Stay updated on the latest data science, machine learning, and agricultural advancements. Continual learning ensures that your analysis incorporates the most recent techniques and knowledge. Consider the environmental impact of your analysis. Optimize code and algorithms to minimize resource consumption, especially when deploying models in resource-constrained environments.

# Evidence of Configuration Management

1. **Name of the CM tool used by your team**: Google Docs
2. **Version number of each document after it is checked in**: Google Docs does not have version numbers, but has versions based on timestamps for each session. We can view these in the version history in *File > Version History*.
3. **Version number of each document before it is checked out**: Google Docs does not have check out, or check in, the team works on the latest time stamped version of the document.
4. **Difference between two consecutive versions**: Users can view the revision history through Google docs to see the different versions of the document.
5. **Review of each change**:
   1. After the document is finalized, we have the team double check it to make sure all the input is correct.
6. **Other information that helps the understanding of each change**:
   1. Different team members are assigned different tasks and put this information in the table below.
   2. Every team member is added with their gmail accounts so we know who made each change.

# 

# Configuration Management Table

| **Changes Made** | **Update Description** | **Submitted by** | **Date Updated** |
| --- | --- | --- | --- |
| Abstract, Table of Contents, Monitoring / Reporting / Controlling Mechanism | Created original documentation | Ryan | 01/30/24 |
| Project Organization, Lifecycle Model | Created original documentation | Cameron | 01/31/24 |
| Risk Analysis / Deliverable Schedule | Created original documentation | Nisai Sun | 02/02/24 |
| Introduction / Engineering Standards | Created original documentation | Ibrahim | 02/02/24 |
| Standards / Constraints | Updated Standard / Constraints with Professor's email | Ryan | 02/14/24 |
| References | Updated References with professors email | Ryan | 02/14/24 |
| Cover Page | Updated Cover Page orientation | Nisai | 03/09/24 |
| Evidence of Configuration Management | Added Evidence of Configuration Management Page / Table | Ryan / Ibrahim | 03/09/24 |
| Engineering Standards and Constraints | Updated Standards and Constraints to be in Line with Project Management | Ryan | 04/08/24 |
| Engineering Standards and Constraints | Updated links for documents to working links | Ryan | 04/08/24 |
| Additional References | Added additional references | Ryan | 04/08/24 |

# Engineering Standards and Multiple Constraints

* IEEE Std 1058-1998: Software Project Management Plans [[pdf](https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=25325)]
* PMBOK® Guide: Project Management Body of Knowledge [[pdf](https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=1302773)]
* IEEE Std 12207: Software Life Cycle Processes [[pdf](https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=8742773)]
* IEEE Std 15939: Measurement Process [[pdf](https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=4775910)]
* ISO/IEC/IEEE Std 29148-2018: Systems and Software Engineering
  + Life Cycle Processes
  + Requirements Engineering [[pdf](https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=8559686)]

## Additional standards suggested by the sponsor

* IEEE 1490: Standard for Software Engineering Management Plans [[pdf](https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=1302773)]
* IEEE 828: Standard for Software Configuration Management Plans [[pdf](https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=6044675)]

# 

# Additional References

* *Software engineering: Project Monitoring and control - javatpoint*. www.javatpoint.com. (n.d.). <https://www.javatpoint.com/project-monitoring-and-control>
* *Geospatial Data Analytics: What it is, benefits, and top use cases*. SafeGraph. (n.d.). <https://www.safegraph.com/guides/geospatial-data-analytics#:~:text=Geospatial%20data%20analysis%20involves%20collecting,in%20the%20relationships%20between%20places>
* Larson, E. and Gray, C., 2014. *Project Management: The Managerial Process.* McGraw Hill
* Humphrey, W.S. and Thomas, W.R., 2010. *Reflections on Management: How to Manage  
  Your Software Projects, Your Teams, Your Boss, and Yourself.* Pearson Education
* M. T, K. Makkithaya and N. V. G, "A Federated Learning-Based Crop Yield Prediction for Agricultural Production Risk Management," *2022 IEEE Delhi Section Conference (DELCON)*, New Delhi, India, 2022, pp. 1-7, doi: 10.1109/DELCON54057.2022.9752836.
* K. Styk, J. Liszcz and K. Drobek, "Basic Project Management Documentation Based on the Example of the Student Project AGH Lean Line," 2019 8th International Conference on Industrial Technology and Management (ICITM), Cambridge, UK, 2019, pp. 45-49, doi: 10.1109/ICITM.2019.8710717.