**1. Project Overview**

The project plan must include the following:

**· A Project Title.**

Predictive Analytics on the Recent California's Wildfire Risk Assessment and Damage Estimation Using Climate and Environmental Data

**· A short summary of the project topic and background.**

There have been increasingly severe wildfires in California leading to casualties, property damage and an impact on the economy. For example, they caused at least29 deaths and over 16,000structure losses, and the total property damage from these fires is estimated to have reached up to $275billion in 2025 Southern California wildfires. These events underpin the importance of disaster risk predictability and effective predictive analytics to facilitate estimation of such damage. The goal of this project is to blend the data regarding climate and environment to build predictive models that would provide info related to wildfire risk and the corresponding damages in California. The project will analyse datasets like the CAL FIRE Damage Inspection (DINS) data or state expenditure records to understand which aspects have a strong relation to the wildfire occurrences and the impact of them.

**· A Research Question.**

How do machine learning models use climate, environmental, and historical wildfire data to accurately predict wildfire risk and estimate potential damage in high-risk Californian regions?

**· The Project Objectives.**

* To create machine learning models utilizing climate and environmental data to predict areas in California at heightened risk for wildfires.
* To utilize predictive models to estimate potential damages, including property loss and economic impact, associated with predicted wildfire events.
* To provide insights to policymakers and emergency management agencies to inform resource allocation and develop effective wildfire mitigation strategies.

**· Reference List**

Qiu, L., Chen, J., Fan, L., Sun, L. and Zheng, C., 2022. High-resolution mapping of wildfire drivers in California based on machine learning. *Science of The Total Environment*, *833*, p.155155.

Hernandez, K. and Hoskins, A.B., 2024. Machine learning algorithms applied to wildfire data in California's central valley. *Trees, Forests and People*, *15*, p.100516.

Jain, P., Coogan, S.C., Subramanian, S.G., Crowley, M., Taylor, S. and Flannigan, M.D., 2020. A review of machine learning applications in wildfire science and management. *Environmental Reviews*, *28*(4), pp.478-505.

**2. Project Plan: Task List and/or Project Timeline**

| **Task Number** | **Task Description** | **Start Date** | **End Date** | **Notes** |
| --- | --- | --- | --- | --- |
| 1 | Conduct Literature Review | Feb 10, 2025 | Mar 5, 2025 | Review existing research on wildfire predictive analytics and relevant datasets. |
| 2 | Data Collection and Preprocessing | Feb 15, 2025 | Mar 10, 2025 | Collect and preprocess CAL FIRE DINS data and state expenditure records. |
| 3 | Exploratory Data Analysis (EDA) | Mar 1, 2025 | Mar 20, 2025 | Perform EDA to understand data distributions and identify key features. |
| 4 | Model Development | Mar 10, 2025 | Apr 20, 2025 | Develop predictive models using machine learning techniques. |
| 5 | Model Evaluation and Validation | Apr 10, 2025 | Apr 30, 2025 | Evaluate model performance and validate using appropriate metrics. |
| 6 | Results Interpretation and Documentation | Apr 20, 2025 | May 5, 2025 | Interpret results and document findings comprehensively. |
| 7 | Report Writing and Presentation Preparation | Apr 25, 2025 | May 5, 2025 | Prepare the final report and presentation materials. |
| 8 | Final Assessment Submission | May 5, 2025 | May 5, 2025 | Submit the final report and present findings to stakeholders. |

3. Data Management Plan

*Overview of the Dataset*

A very detailed data that talks of features of structures damaged or destroyed in California wildfires is offered. It is used to assess wildfire impacts and is collected by the California Department of Forestry and Fire Protection (CAL FIRE). These records are available through the Open FI$Cal platform, and they detail California’s departmental spending transactions which gives context to the allocation of funds on wildfire management and recovery.

*Data Collection*

* **CAL FIRE DINS Data:** Accessible at [CAL FIRE DINS Data](https://data.ca.gov/dataset/cal-fire-damage-inspection-dins-data/resource/b8aeb030-140d-43d2-aa29-1a80862e3d62).
* **State Expenditure Records:** Available at [Open FI$Cal Department Spending Transaction Files](https://open.fiscal.ca.gov/dept_spending_transaction.html).

*Metadata*

It provided CSV format, which contained records of inspected structures including, damage level, structure type and geographic location. The dataset size is about 50 MBs. Monthly spending transactions by department is available in CSV format. The dataset size roughly is about 200 MB.

*Document Control*

Version control and code management will happen through a dedicated GitHub repository for this. The repository will be committed with code and documentation updates per week.

*ReadMe File*

* **Project Overview:** A brief description of the project's purpose and objectives.
* **Dataset Information:** Details on the datasets used, including sources and descriptions.
* **Setup Instructions:** Guidelines on how to set up the project environment and dependencies.
* **Usage Instructions:** Examples of how to run the code and interpret the outputs.
* **Contact Information:** Details for further inquiries or contributions.

*Security and Storage*

* **Backup Frequency:** Data and code will be backed whenever the site updates their record.

Ethical requirements: You must address each of the following issues and state how your specific dataset meets these requirements, give evidence when possible (e.g. screenshots or references):

1. Does the data come under GDPR requirements?

The European Union regulations regarding the processing of personal data of individuals living within the EU are collectively referred to as the General Data Protection Regulation, or the GDPR. Both the CAL FIRE Damage Inspection (DINS) data and California’s state expenditure records are drawn from U.S. entities and relate to events occurring within California. Such project does not concern personal data of EU residents to which GDPR relates to, so these datasets do not contain such.

2. Does the project conform to UH ethical policies?

According to University of Hertfordshire’s policy, all researches involving human participants or their data or biological material must achieve ethical approval before commencement of such researches. Although the project only makes use of secondary data that is publicly available and does not use human participants or personal data. Research that only involves use of published secondary sources does not require approval but may require a declaration of secondary research according to UH guidelines. The researcher will refer to their supervisor in order to be compliant with UH ethical policies.

3. Do you have permission to use the data for your proposed research project?

The datasets in question are publicly accessible:

* **CAL FIRE DINS Data:** Available at [CAL FIRE DINS Data](https://data.ca.gov/dataset/cal-fire-damage-inspection-dins-data/resource/b8aeb030-140d-43d2-aa29-1a80862e3d62).
* **State Expenditure Records:** Accessible via [Open FI$Cal Department Spending Transaction Files](https://open.fiscal.ca.gov/dept_spending_transaction.html).

These platforms provide such data for public use, it is not forbidden to exploit them for research purpose. The researcher will cheque whether there are any associated terms of use, and if they are, he or she will make sure compliance with all specified conditions.

4. Are you assured that the data was collected ethical (i.e. by the original people who gathered/collected/ collated/made the data)?

Parametric and derived wildfire indicators are compiled from and disseminated by reputable government agencies based on datasets collected by the California Department of Forestry and Fire Protection (CAL FIRE) to impact wildfire assessment. It is discharged by the State of California to advance transparency in governmental profitable movement. Because these agencies have ethical guidelines to adhere to when collecting data, they ensure that the data they collect is done responsibly and ethically.