

### **Project Proposal: Skill Combination Predictor for Student Success**

#### **Team Members**

CAP4910 – Capstone Project

* **Ramon Castro**
* **Krister Martinez**

**Jira Link:**

[**https://capstonemdc.atlassian.net/jira/software/projects/KAN/boards/1?atlOrigin=eyJpIjoiNjY2YzQyOGZhZWQ3NGY4ZWFlMGMwYmIxNjM0NjNmNjciLCJwIjoiaiJ9**](https://capstonemdc.atlassian.net/jira/software/projects/KAN/boards/1?atlOrigin=eyJpIjoiNjY2YzQyOGZhZWQ3NGY4ZWFlMGMwYmIxNjM0NjNmNjciLCJwIjoiaiJ9)

**Github Link:**

[**https://github.com/MiamiCrypto/Capstone-Project-**](https://github.com/MiamiCrypto/Capstone-Project-)

#### **Chosen Project**

**Project Title:** Skill Combination Predictor for Student Success

**Objective:** Develop a predictive analytics system that utilizes detailed student resume data, along with academic and extracurricular datasets, to forecast successful skill combinations for students. This system aims to guide students in selecting courses and activities that align with successful outcomes in their desired career paths.

#### **Business Value**

* **Guided Academic Choices:** Provide students with data-driven recommendations on skills and courses that enhance employability and academic achievement.
* **Curriculum Development:** Assist educational institutions in designing curricula that align more closely with market demands and successful student profiles.
* **Enhanced Employer Recruitment:** Enable employers to identify potential candidates who possess the desired combinations of skills for specific roles.
* **Strategic Resource Allocation:** Facilitate better resource planning within schools based on predictive insights into popular and successful skill sets.

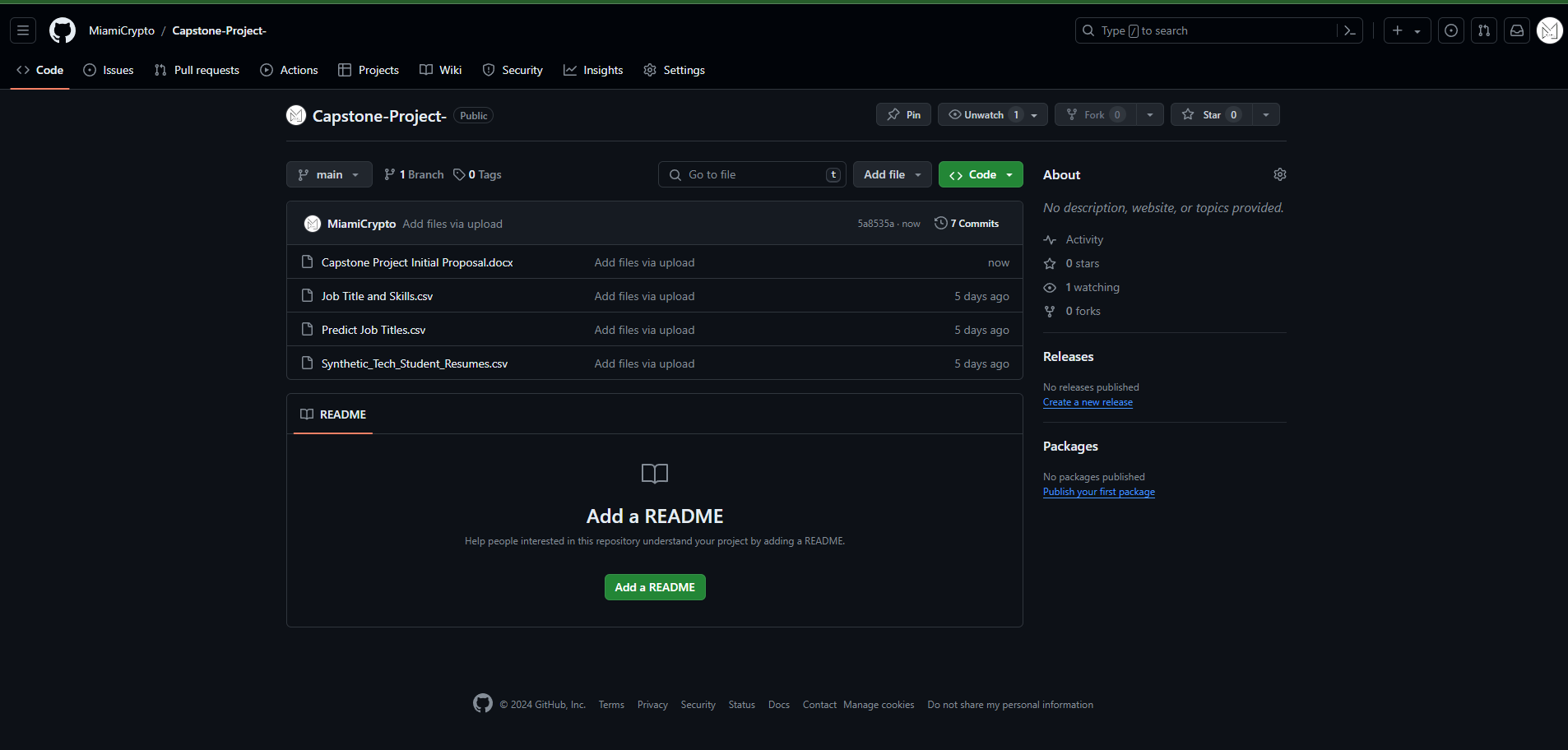
#### **Technical Architecture**

* **Data Layer:** Gather student and employer datasets from online resources such as Kaggle and transform the data with the help of ChatGPT. Expand the datasets using resources such as Synthetic Data Vault. Utilize secure cloud storage solutions such as AWS or other services for handling synthetic and real student datasets, ensuring privacy and compliance.
* **Processing Layer:** Implement data processing and machine learning model development using Python, with libraries such as scikit-learn for predictive modeling.
* **Application Layer:** Develop a front-end interface for model interaction with Streamlit or React for user-friendly access.
* **Visualization and Reporting:** Use PowerBI or Tableau for dynamic reporting and visualization of predictive results and student data insights.

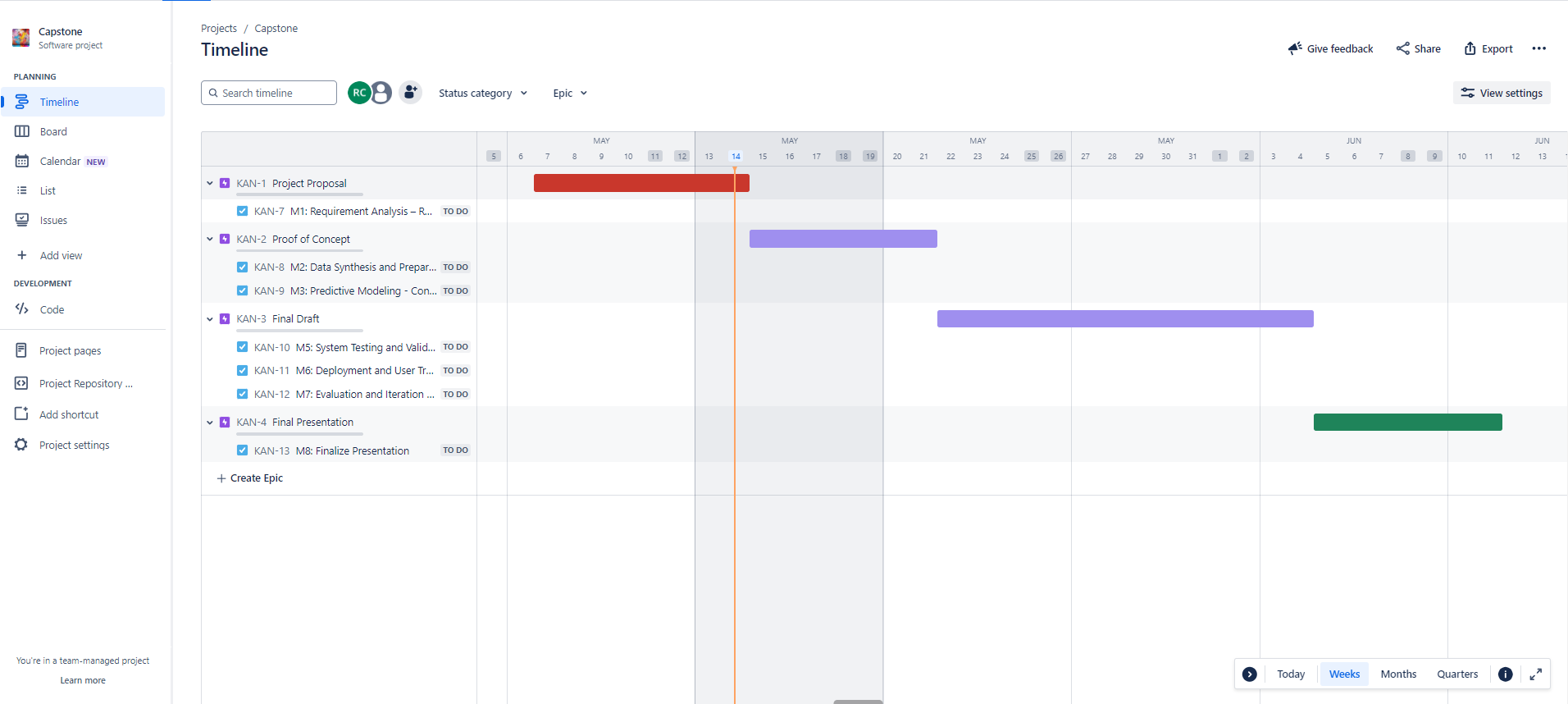
#### **Initial Setup of Jira Board (Kanban)**

* **Milestones:**
* **M1: Requirement Analysis** – Research and understand needs and finalize project scope.
* **M2: Data Synthesis and Preparation** - Create a comprehensive synthetic dataset that mirrors real student data, emphasizing skill diversity.
* **M3: Predictive Modeling** - Construct and train models to predict successful skill combinations based on historical data.
* **M4: Integration and Interface Development** - Integrate backend models with a front-end interface, ensuring seamless user interaction.
* **M5: System Testing and Validation** - Rigorous testing of the system’s predictive accuracy and user interface, followed by stakeholder validation.
* **M6: Deployment and User Training** - Deploy the system into a production environment and train users in its functionalities.
* **M7: Evaluation and Iteration** - Evaluate system performance and user feedback, iteratively refine the tool.

**Project Github**



**Jira Timeline**



**Proposed Synthetic Dataset**



Skills Dataset

