

# Milestone 2: Vision and Scope, User Stories, Iteration Plan

Group 21 - Interactive Visualization for AI Education

# Table of Contents

<b>Table of Contents</b>	<b>2</b>
<b>Introduction</b>	<b>3</b>
<b>Background</b>	<b>3</b>
Problems	3
Current Solutions	3
Project History	4
<b>Vision</b>	<b>5</b>
Hypotheses	5
Growth hypothesis	5
Value hypothesis	5
Requirements	6
Functional requirements	6
Non-functional requirements	6
<b>Prioritized Project Constraints</b>	<b>7</b>
Time	7
Resources	7
Scope	7
<b>Scope</b>	<b>9</b>
Process Flows	9
User Stories	10
<b>Iteration Plan and Estimate</b>	<b>12</b>

# Introduction

Milestone 2 will lay out a holistic view of the Interaction Visualization for AI Education project. Group 21 is the first team to work on a project of this nature allowing the team to build a new fun and educative tool from the ground up. The overall goal is to create an online learning tool that guides users from start to finish through creating a model to classify the sentiment in a given sentence and empower them to find their own datasets and build a subsequent classification model on their own.

## Background

### Project Introduction

This project aims to be an interactive visualization tool to teach beginners about artificial intelligence (AI) and machine learning (ML). This tool will allow users to interact and actively learn about complex computer science topics. Learning about artificial intelligence and machine learning can be difficult given that these topics are a complicated and niche sector of computer science. There lacks fun, interactive, informative education tools to teach beginners about AI and ML and so this interactive web application project aims to serve as an educational tool that will help novices become more acquainted with AI and ML. Users that will be included in testing this tool including students aspiring to learn about AI, students who have some background or experience with AI, and potentially faculty or industry professionals who are interested in AI visualization tools.

Oregon State University professor Dr. Minsuk Kahng had proposed this project to create a tool to help users learn about AI and how AI works. The current team members include: Laura Jiang, Junhyeok Jeong, Owen Markley, and Thuy-Vy Nguyen.

### Problems

Machine learning and artificial intelligence can be difficult to pick up because of how overwhelming the field is. From understanding the terminology in artificial intelligence (AI) to the actual math underlying AI, this is a difficult topic to both attempt to learn on your own and to teach to other people. While there is currently a wealth of information on languages, frameworks, tools, and resources, it can feel overwhelming especially for beginners. What resource is best to start out with? How can a web application be useful to help both unfamiliar students and professionals?

### Current Solutions

There are a few different methods that people currently use to help them learn about machine learning and artificial intelligence topics. Some of these websites include Coursera, EdX, and other universities which offer online courses on a variety of topics.

There are also a few sites out there whose goal is to abstract away programming and move towards visualizing machine learning to make the topic more accessible. Google in particular has a large collection of interactive tools and courses through their Google AI Education site. In particular, two of their popular tools include Teachable Machine and Facets. Facebook also has their own collection of tools, linked below.

Source	Value
<a href="#">Google AI Education</a>	Provides a mass collection of Google-backed resources, including visualizations, guides, and crash courses.
<a href="#">Teachable Machine</a>	Also Google-backed. Easy method for on-the-fly creation of image classification to learn about epochs and learning rate.
<a href="#">Facets</a>	A complex deep-dive into a particular dataset. One set is already provided on load, but the user can also choose to upload their own data.
<a href="#">Facebook AI</a>	Provides a collection of resources related to implementation assets for programmers and hobbyists, such as languages, libraries, etc.

The distinguishing factor of the capstone project from these current solutions is that they do not give the user the full end-to-end experience of both learning and applying concepts at the same time. The Google AI and Facebook AI tools are both collections of resources but still lead to the issue of the user feeling overwhelmed and uncertain where to start. While Teachable Machine and Facets focus on visualization, they do not grant users any opportunity to learn what the programming aspect of machine learning looks like. The capstone project will remedy the issues from both groups by both providing an entry-level explanation for beginners and an opportunity to access and play around with code if desired.

## Project History

This particular project does not have any prior history - this will be the first iteration of the project.

# Vision

There are two main user segments that this capstone project is aimed at: students and professionals in the technology industry with an interest in learning about AI. The project will help solve two large pain points experienced by these two user groups.

The first pain point is an overwhelming amount of resources without a string between those resources, such as the resources provided by Google and Facebook AI mentioned above. By tying everything together in a singular application through this capstone project, the user can work in the same tab. This simplifies their workflow and grants them a unified knowledge base. Next, the second pain point is minimal instruction. Many existing interactive visualizations are complex and do not feel incredibly user friendly, such as Facets which is a deep-dive into a provided dataset.

Comparative to other existing visualizations, this project is aimed to be beginner friendly and help develop skills for users new to machine learning. This tool will be beginner friendly in that it is an interactive and visual educational tool to teach users about introductory artificial intelligence concepts.

The minimum viable product (MVP) is a multi-page web application that will guide users through creating a sentiment classification model from start to finish. Users will be introduced to concepts like how to train a model, what kind of data is suited for a classification model, loss functions, and building a classification model. Code for these elements can be optionally accessed by the users.

## Hypotheses

### Growth hypothesis

Within the resource space for educational machine learning tools, the team anticipates that this product will see particular adoption because it stands apart from current solutions; users will naturally gravitate towards this product for its visual and hands-on learning experiences (visualization and programming access). Another hope for adoption is that this product will be directly helpful in a formalized education setting; educators can showcase and leverage this product as an instructive tool and students will be inclined to review and understand the concepts presented on the site.

### Value hypothesis

It is hoped that customers will use this tool because they want an engaged and enriched learning experience. Customers who use this tool alongside other artificial intelligence or machine learning resources may have a deeper understanding of the content and be able to explain these concepts more easily due to the in-depth explanations that the site provides them. Compared to existing solutions, this product is designed to educate users from different levels of expertise and does not expect any pre-requisite knowledge. Additionally, this product will pair

direct code snippets with educational text which is also a key distinguishing factor from the status quo.

## Requirements

The larger goal is to build a web application that is able to both teach the user about machine learning concepts and grant them a platform where they can start playing around with a machine learning system.

Below is a more detailed description of the functional and non-functional requirements.

### Functional requirements

- The system will contain a walkthrough of how to build a sentiment classification machine learning model.
- The system will show the user's progress as they move through the site's topics.
- The system will include an interface that allows users to work through the website topics based on the user's learning interests.
- The system will then provide visualization and/or explanation of each step within the model based on the chosen dataset to educate the user about the ML model process.
- The system will feature code snippets that correspond to the topics being covered on that particular page.
- **(STRETCH GOAL)** The system will walk users through the process for simple image classification.
- **(STRETCH GOAL)** The system should allow for user modification of that code and be able to run and display any changes.
- **(STRETCH GOAL)** The system will allow the user to upload a custom dataset and build a model around their dataset.

### Non-functional requirements

- The system will display changes made by the user within 10 seconds.
- The system should be able to load information on different pages or screens within 5 seconds.
- The system should be functional on different sizes of screens and browser windows.
- The system should be open source to encourage maintainability of the site and additional future updates.

# Prioritized Project Constraints

## Time

One obvious time constraint is that the Capstone timeline is to complete a project within the school year (9 months). By the end of fall term, the team hopes to complete the first prototype of a web application that can support some form of user interaction related to machine learning and demonstrate as a prototype for the design gallery. By the middle of winter term, the team would like a second or third prototype, iterating upon user research conducted in the first weeks of winter term.

Members of the team will allocate time for the Capstone project as appropriate. Each sprint will have features assigned to a particular team member. Although there is no particular time requirement, they will be expected to complement that task within the sprint. An estimate would be around 2 - 7 additional hours a week (not including the time required for team stand-ups or project partner meetings), depending on the task and the phase in the project.

## Resources

User studies will be conducted in order to most effectively design the product. The user studies will be conducted to understand motivations behind the project and how to design an intuitive website. The study process will be shortened to informal, short interviews or user walkthroughs to receive quick feedback.

Other resources include hosting the final polished project onto a website so that members outside of the team will be able to use it. The team has not decided where the site will be hosted but there is a high likelihood that the site will be hosted on a platform like AWS. The project partner will be available to host the site indefinitely as long as the project is seeing usage and occasional maintenance.

In addition, our project owner, Dr. Minsuk Kahng, can provide advice based on various data visualization and AI project experiences like ActiVis, GAN Lab, and ETable. These projects are related to this capstone project and will use similar tools for visualization.

## Scope

The team has decided on a set of minimal features and functionality that should be completed by the end of the school year. This consists of the minimal interactions the user will need to have in order to effectively use this tool: providing training and test data for users, permitting users to modify certain parameters of the machine learning (ML) model, and an instructional guide to explain each step of the sentiment classification process to the users.

One important aspect of this capstone project is user studies and user research. The inclusion of user feedback may slow down parts of the project, limiting the scope of the project as well. There is no clear projection on what this will look like now.

Nevertheless, features may be altered or additional features may be added depending on time constraints, results of user studies, or discussion during team meetings.



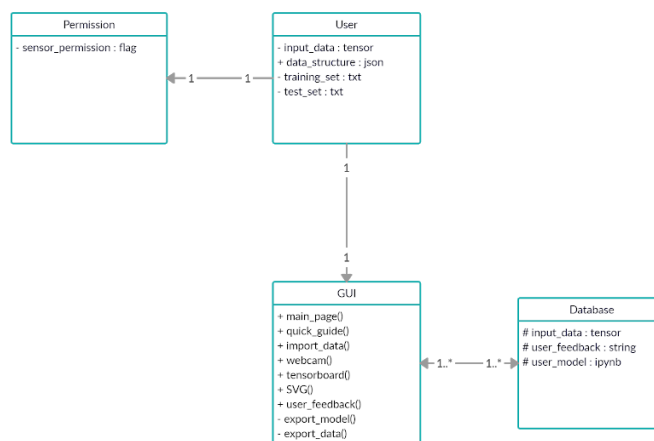
# Scope

One central tenet of this project is designing for the user and understanding their needs. With a student or professional beginning to learn about machine learning, general concepts can take time to process and understand how to apply them. The web application should be easy to pick up and also take users through an engaging introduction of fundamental concepts in machine learning.

For the team's workflow, as mentioned, informal user testing will be central to improving the application. Accessibility and usability are the foundation of turning this into a highly successful project.

## Process Flows

UML Class Diagram



UML Activity Diagram

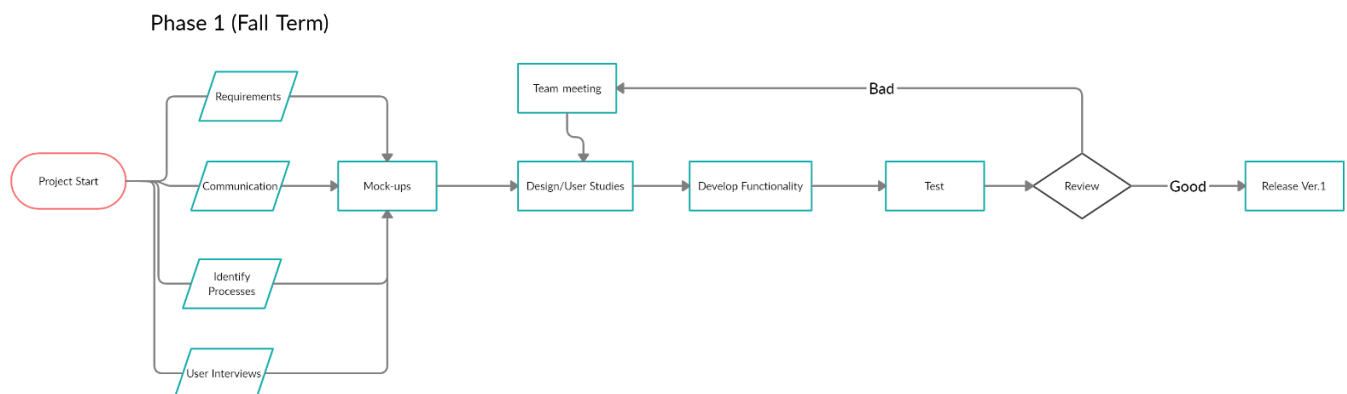


Fig 1. UML Class Diagram



Fig 2. UML Diagram for User Flow

## User Stories

1. As a college student trying to learn about how to build my first classification model I need an inclusive step-by-step tutorial so that I do not need to jump between different resources to learn any prerequisite concepts.
2. As a hobbyist programmer I need to understand how changing the learning rate affects the accuracy of my classification model so that I can learn what learning rates are more effective when building my own models.

3. As a user I need to understand what a training set and learning set is so that I know what a classification model's prediction accuracy is founded upon and how I can improve the model's accuracy.
4. As a high school student interested in AI, I need a singular tool for learning AI so that I am not overwhelmed with information or background knowledge that I am expected to know and I can move at my own pace.
5. As a technical product manager, I need to understand the technical underpinnings of what a classification model is and precisely how it can classify an incoming message so that I can communicate between the engineers and our customers to explain how my company's spam filter works.
6. As a college student working with machine learning models for the first time I need a way to build upon my knowledge of machine learning so that I can build confidence and comfort in this subject.
7. As a technical product manager I need to understand different basic machine learning concepts so that I can better communicate with my engineering team and better understand the products that my company is building.
8. As a software engineer, I need to spin up an accessible demonstration of a machine learning model for non-technical people so that I can demonstrate machine learning to other teams.
9. As a college student, I need a step-by-step guide on what an interactive visualization does so that I can effectively use that tool to supplement my learning.

# Iteration Plan and Estimate

The current expectation is that two sprints be completed by the end of the fall term. The team's sprints are broken up into 2 week periods. The first of these sprints will focus primarily on gaining a deeper understanding of the needs of the different users that are expected to be using this tool, as well as developing a basic prototype that will represent the user experience. Information from interviews and user research will be incorporated into this first iteration. Subsequent interviews and interactions with users, as well as feedback from Dr. Kahng will determine the form of the functional and nonfunctional requirements initially provided by Dr. Kanhg.

While working on this first sprint, the users' needs will always be in focus. In depth diagrams and supplementary material will also be updated as the prototype progresses. Supplementary material should map directly to the users' needs. During this phase it will also be determined as to which frameworks, databases, libraries will be used. By the end of this first sprint, there should be some form of deliverable to show our progress, and more importantly, a clearer vision for a development.

The second sprint is dedicated to further developing the first iteration of the user interface, and having a fully functional backend built for at least one type of model. Ideally, this iteration would be able to work with sample data sets to build out the sentiment classification model. Once a basic backend and user interface has been created and is usable, the team may then seek another round of user feedback. At this point time and resources can be allocated towards either tracking and adjusting existing functionality to better suit the needs of the user or working to create a rough prototype of what the incorporation of code would look like on the user interface side.

Lastly, any remaining time is dedicated to ensuring that there are no bugs or issues with the functional prototype that would render it unusable. It should be ensured that every single one of the predetermined goals for the term is met, and that functional requirements are being implemented. This portion may also be partially dedicated towards the testing of a working prototype. Once the final sprint is completed, a plan for the subsequent Winter 2020 term may be created by the team.