# FourthBrain Capstone Projects

Welcome to Capstone Project Time!

An end-to-end project drives you to think strategically about approaching a business problem at data-driven companies, to strengthen your knowledge and showcase your skills, while enhancing teamwork and communication with your fellow engineers. Capstone projects are meant to synthesize various skills that you have learned and developed through the course of this cohort and apply them to build a new product that drives business value.

You’ll be working in a group of three engineers to complete an end to end product. The final **deliverables** include

* Deployed Demo
* 10-minute Presentation
* GitHub Repo ( description in README + code )

You can see sample capstone projects on our public [Capstone Projects Gallery](https://www.fourthbrain.ai/capstone-projects).

## Project Timeline

**Week 1–4: Overview and Team Formation**

**Week 1 Live Session**: Capstone Process Overview. Review all the industry designed projects and (Optional) develop your project pitch. Team search has started!

**Week 2 Live Session:** Initial Project Selection and Student Project Pitches.

**Week 3 Live Session:** Teaming and data/model scoping. The teams after Week 3 should be relatively stable; inform a staff member if you don’t have a team by the end of Week 3.

**Week 4 Live Session**: Final Team / Project formation. Before the Week 4 Live Session, continue discussing with your teammates and conduct any research needed. Make sure you meet with your team at least once to confirm this is a good match.

**Week 5 Live Session**: Project Proposal Presentations (2–3 mins)

**Week 6–14**: Project Work

**Week 9 Live Session**: Project Updates (3–5 minutes)

**Week 14 Live Session**: Capstone Practice Presentations (5 minutes)

**Week 15**: Final Presentation / Demo Day

## Selecting a Capstone Project

We have a library of [Industry Projects](https://drive.google.com/drive/folders/1H-I0wGcOcnXotu_5pe16PtMyhYBKYNU1?usp=sharing) co-designed with some of our partners. You may select one of these projects, or work on a Self-Designed project where you have the option to pitch your own project during the Week 2 live session.

When you are discussing project ideas and specifics with your classmates, you should consider the following:

1. Start with **Why**. Why is this problem worth solving? Why is machine learning necessary to solve the problem? Why would stakeholders fund you (explainability, reliability, ROI, scalability, fairness, regulations)? Who are the potential customers and why should they choose your product?
2. **How** to frame the problem (regression, multiclass, multilabel, time series)? How will you evaluate the product, from a machine learning perspective, and/or from a business’s standpoint? How will you scope the work? Do you need to collect the data, and if yes how? Are there any privacy concerns related to the data?
3. **What** are the time and budget **constraints**? Can we leverage existing paper/models/code? Either way, can we establish a baseline model, a working system by week 8, and then iterate it each week afterwards? What are the system leve**l** considerations that you have to make (AWS, GCP, platform, software, tools etc.)?

## Industry-Designed Projects

You will have the choice to work on an Industry Sponsored Project that is already a pre-defined problem, often with a code base and data sets to work on. The difference between an industry-sponsored project and a personal project is that the industry-sponsored project is already polished with prior documentation for a smooth start. Some companies may even be available for asynchronous input throughout the project!

If more than three of you want to work on the same project, we will split the group into multiple teams working on the same project. The teams will align at the start to make sure each team is taking a different approach so the final products are sufficiently different.

## Selecting teammates

Picking teammates you can work well with is important!

1. **Alignment:** Make sure you are aligned on the expectations for the project, that you are **available at similar times** to plan work sessions/meetings. Be realistic about how much time you can dedicate to the project.
2. **Team Experience**: Some prior experience on the specific use-case (NLP, Computer Vision, Audio processing, signal processing, biomedical engineering) will help steer you in the right direction rather than picking a brand new topic of interest. Teams with complementary skills are found to be most efficient.
3. **Project Team Size**: Based on our experience 3 person project teams are most effective and efficient. However, in some cases 2 person project teams may also be formed. ALL team members are to contribute equally to CODE development and GitHub commits. Project teams should be formed not just based on common topic interests but also on common working hours since you will have to work together throughout the week.

## Things to Consider for your Capstone Project Proposal

Envision your project as serving one of two goals. First, you could be called on to develop a new prototype for your company and demonstrate a deployed end-to-end solution that scales. Second, you could belong to a company that has an established data processing pipeline, but there is a change in either the data or pipeline and that needs careful evaluation regarding the benefits or limitations of the envisioned change. In this case you are required to provide a summary of your findings and recommendations.

Thus, the Final Project (on Demo Day) could either be a deployed Machine Learning Solution on a cloud platform, or a complete analysis of a data and system pipeline assessed on at least 2 data sets. One data set can be small (<1GB) while the other data sets should be real sized (>1GB). The goal here is to validate if your model transfers learning across data sets with minimum platform dependencies or not.

From a Process and System Design perspective, think of your project as a proof-of-concept prototype that needs to be deployed in versions. You will need to identify who your customers are and think about the data, data treatment, models and processes, output, recommendations and feedback and communicate your findings to your team and advisors on a weekly basis. This project will give you a real gauge into the day-to-day working of a machine learning engineer to navigate through hurdles and deliver your project by meeting targets on time!

## How to Prepare a Project Pitch

During Week 2, any of you who are interested in working on a project topic of your own can pitch their idea in class. The goal is to get other engineers interested in the project to work on it. If you have a project idea and you are enthusiastic about it, please consider attending an instructor’s Office Hours to help refine and scope your idea. Considerations of the project pitches are as follows:

1. The project idea MUST be well developed. This implies that the idea should have

* a NEED
* public data sets available to work on

1. The topic should be interesting enough to attract more engineers to work on it. The capstone project cannot be completed solo so you need buy-ins from other engineers.
2. At the end of the Week 2 Live class, a project owner may choose to leave their own and work on another project. So pitching a project does not tie you down to work on that idea. You are free to choose another project if that interests you more.

References:

1. See the [project template](https://docs.google.com/document/d/1xnpANat3qM6WvWEIH5GB1L3-qUNjgvDl/edit) for guidance on a fine tuned project pitch.
2. You can also see some sample pitches here:

* [Edge2Pic Pitch](https://docs.google.com/document/d/15ONDXIt0MSTYynKlp2iA0QdfWVyQZWac2oF1Ydm4f30/edit)
* [Visual Art Gallery Pitch](https://docs.google.com/document/d/1cN6VG8D4bTIFLDLUz7AS4yozA8wizDsj/edit)