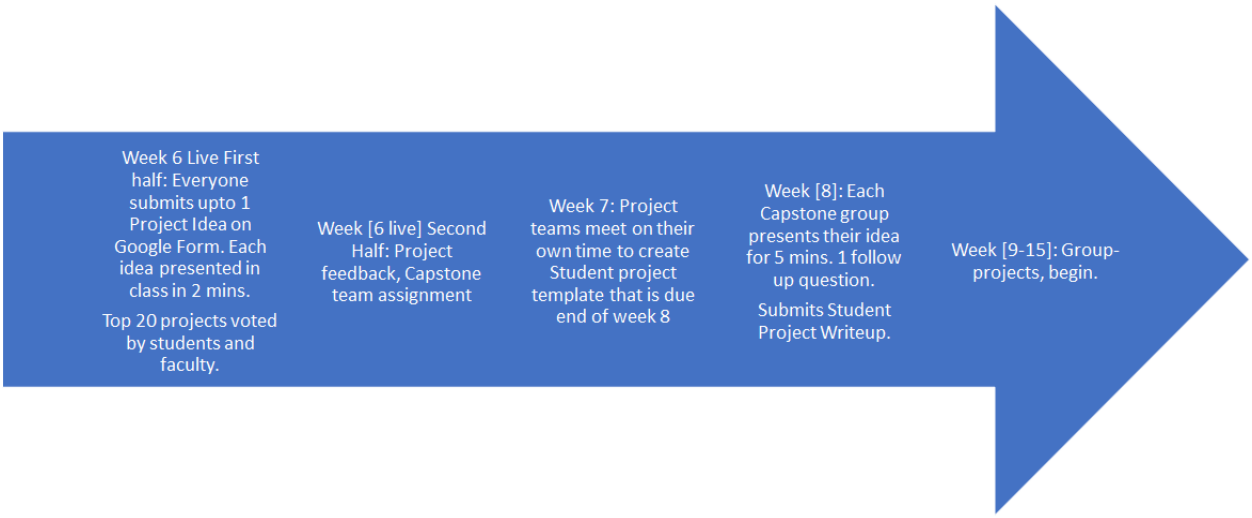
Group Project Selection Guidelines

1. Project Timeplan



Major Timelines
Week 6 Live class: Individual Project Pitches
Week 8 Live Class:Project Proposals
Week 9-14: Project Work
Week 15: Demo Day.

2. Things to consider for deciding a Capstone Project topic

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 while working with team members and other teams that help you with data and processes. 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 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 delivering your project by meeting targets on time!

Questions to ask yourself before submitting a project idea (on Google form) and pitching it on Week 6 are as follows:

Q1. **Why** is the problem important? Has there been enough work done in this area or is this idea groundbreaking? If work has already been done on the topic then what is the novelty in your proof of concept (performance, scalability, speed)? If the idea is groundbreaking, how will you convince stakeholders to fund the project (explainability, benchmarking on public data sets)?
Q2. **What** is the delivery process going to be? How many iterations will you require? Is there some baseline paper/code you should replicate first to benchmark? What modifications can you then make to the benchmark? What are the **System level** considerations that you have to make (AWS, GCP, platform, software, tools etc.)?
Q3. **How** are you going to execute the project? What will be your time plan? Do you have plans to collect the required data sets (with labels) and develop processing data models. How will you **pre-process**/ wrangle the data set?
Q4. Is this project going to be something that attracts employer/stakeholder attention to fund in future (scope of extendability)? What aspects should you consider to ensure the project is kept alive/ be appealing in future.
Q5. Is this work doable in 6 weeks by a team of 3 people spending 10-15 hrs a week for this work? Scoping the work is most important.
Q6. Is there existing public domain data sets and methods/models/code/papers that can be useful for the project?

3. **Expectations for Capstone Project Topic**

   One major requirement to complete the FourthBrain MLE program is the 6-week Capstone Project. The expectations for the final project deliverable are:

   1. **Problem Positioning:** There should be a need for the final deliverable.
   2. **ML Demo:** There should be a final deployed ML model demo.
   3. **Processing restrictions:** To work on your projects you will have access to AWS and GCP resources, but the resources have limitations. Each team will have a strict budget that will allow for upto 12 days of Ec2 processing using P2 spot instances. This means highly compute intensive projects (for instance hi fidelity GAN models) will NOT be a good project fit or they have to be scoped down with limited data and processing constraints. The cloud compute resources are shared for the whole batch and MUST BE USED RESPONSIBLY.
   4. Qualitative and Quantitative **Reporting**: The final deliverable includes a 10 minute verbal presentation and a detailed Technical Report. Focus on data representation, visualization, explainability, Tables and metrics is key for a successful project.

You may choose to extend your Individual Kaggle Project idea for a Capstone Project as well by expanding in scope. For more ideas on Capstone projects, look at the prior projects https://blog.fourthbrain.ai/check-out-our-graduates-final-projects

4. **Expectations for Capstone Project Team**
   For your Capstone project, you should work on a project that aligns best with your career interests after the course completion. Some guidelines for team formation are as follows:
   1. **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.
   2. **Industry-Sponsored projects**: You will have the choice to work on an Industry Sponsored Project that is already a pre-defined problem 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, and aligns with skills that the partner company is interested in. Companies will not be providing input or mentorship through the project. Updates on projects can be communicated to the Industry sponsors WITHOUT expectations on feedback.
   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.

5. **Expectations for Preparing a Project Pitch**
   On Pitch day (Week 6) all students who are interested in working on a project topic of their own can pitch their idea in class for 2 mins each. You may or may not use slides. 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, consider refining the idea before pitching it in class. Please consider attending the Office hours of the SME in Week 6 to help refine 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
      i) a NEED,
      ii) some existing code
      iii) public data sets available to work on.
      See the project template for guidance on a fine tuned project pitch. Project topics that don't satisfy the 3 constraints given above may not be considered for final project ideas.
   2. 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 here.

3. If you pitch a topic, you become the topic owner for week 6 and are responsible to receive feedback and fine tune the project idea even further in class.
4. At the end of the feedback session on Week 6 Live class, a topic owner may choose to leave their own topic and work on another project idea. So pitching a topic does not tie you down to work on that project idea only. You are free to choose another topic if that interests you more.

6. Capstone Project Evaluation Criteria

At the end of week 6 Live class, each candidate will have chosen their Capstone Project and team. The student teams have until week 8 to make changes to their project topic or team. In the Week 8 Live session, all Capstone project teams will make 10 minute presentations regarding the Why, What and How, System Design and Ethical Considerations (more details to follow on presentation materials). Based on the presentations, candidates and a panel of judges from FourthBrain will score each project on a scale of 1-5 for the following 3 crateria:

1. Overall Project Plan: Importance on final deliverables and execution plan. Time plan, system and ethical considerations are key here.
2. Project Feasibility: Is the project doable in 6 weeks?
3. Project Usefulness: Is the project well positioned. Will it be useful and extendable?

The teams that are awarded the highest scores out of the weighted combination from judges and candidate scores will receive “People’s Choice Award for Project Pitches”