Capstone Project Proposal

Ashkan Yousefi

Business Goals

Project Overview and Goal	Globally, the number of overweight children under the age of five is estimated to be over 41 million. Overweight children are likely to stay obese into adulthood and develop noncommunicable diseases. Societies that are transitioning to westernized lifestyles are experiencing substantial increases in its prevalence. You can't tackle hunger and the paradox of the obesity crisis among hungry children without educating children on healthy eating. As proven through our academic validations funded by the National Science Foundation, we know that educating children on these topics at an early age, is crucial in contributing towards a solution. It is the parent's responsibility to teach children healthy eating habits. But the demographics that suffer most from obesity are often children of working-class immigrants, whose parents are working more than 2 jobs. Teachers understand the need for teaching nutrition to children and want to step in, but their time and budget in the classroom are limited, and their priority as an educator is to ensure children are succeeding in school subjects. By aligning nutrition and sustainability subjects (which are not mandatory in elementary schools) with math and science learning (which is mandatory), we have presented a way for teachers to obtain the budget and approval needed from Districts and Principals to fund our project and make time in the classroom for teaching these valuable lessons. We have also proven that by introducing cooking in the classroom, we have succeeded in improving math and science test
What is the industry problem you are trying to solve? Why use ML/AI in solving this task? Be as specific as you can when describing how ML/AI can provide value. For example, if you're labeling images, how will this help the business?	

	scores by 50%. A healthy school community sends children the same message in the home. Children who attended our gamified cooking lessons have made informed, healthy decisions that affected their own lives and the lives of their families. The use of AI/ML can provide custom crafted recipes and adjust the lessons based on each individual student. In addition, integration of chatbots can encourage students to be involve more heavily in the process and help the app to get considerable success.
Business Case Why is this an important problem to solve? Make a case for building this product in terms of its impact on recurring revenue, market share, customer happiness and/or other drivers of business success.	Our market fits into several categories within the mobile learning and elementary classroom technology markets. The global market size of the mobile learning product category was approximately USD $12.2 billion in 2017 and is expected to grow to USD $37.8 billion by 2020. Within the broad mobile learning category, the niche market of educational games alone is expected to reach USD $13 billion by 2020 (McKinsey, 2013). The global market size of the mobile learning product category was approximately USD $12.2 billion in 2017 and is expected to grow to USD $37.8 billion by 2020. Within the broad mobile learning category, the niche market of educational games alone is expected to reach USD $13 billion by 2020 (McKinsey, 2013). The Education Market segment, which includes educational gamification in the form of simulations and digital learning platforms, is expected to grow at a high compound annual growth rate (CAGR) of 24.85% in the US with a projected global growth of 68% by 2020. An increasing focus on experiential and inquiry-based learning to operationalize theoretical concepts using interactive technologies is a primary driver of his growth. Thus far customers have been finding us organically and through the promotion of the US Department of Education who finds our pedagogy and approach to education refreshing and innovative. Thanks to our successful pilot studies that were supported by the National Science Foundation and evaluated by advisors from Stanford Research Institute, we have proven that we make a true impact on student learning, thus gaining the trust of customers has been easy. We

	plan to leverage standard customer acquisition techniques in scaling our product -market fit, such as presence in conferences, email campaigns, and partnerships with thought leaders in the education space.

Application of ML/AI
 	
What precise task will you use ML/AI to accomplish? What business outcome or objective will you achieve?
	In regard to AI we are definitely planning on building adaptive feedback into the Q & A interactions that are sensitive to a learner's performance (within a session). And we are also planning on tracking many in-app behaviors (including timing) to provide data for behavioral analysis leading into more AI based adaptive content coverage from session to session (beyond in-session adaptivity and cross-session leveling). Once data is collected from about 1-2 thousand students, we should be ready to consider how AI can aid our understanding of student's learning, engagement, and could inform future features, like
	
	    o Automated norming of our test questions year to year, and region to region
	    o Identifying students with gifted or special needs so that adaptivity can better and more quickly adjust difficulty (and content) to best help and challenge them at appropriate levels
	    o Reporting to teachers when students' show signs of dyslexia, color blindness, dyscalculia, sensory sensitivities, or other patterns of learning disabilities
	
	AI assisted search and contextual assistance could be a future upgrade, as could in-app AI simulations of cooking, nutritious recipe creation, farming or healthy development.

Success Metrics

Success Metrics	The ELA/ELD as well as the STEAM scores can be a good index to emphasis the impact of AI powered educational app for the students.
What business metrics will you apply to determine the success of your product? Good metrics are clearly defined and easily measurable. Specify how you will establish a baseline value to provide a point of comparison.	Conducting a repeated measures analysis of variance (ANOVA), a test to detect any overall differences between related means. With the time factor (i.e., pre- versus post-test) as the within-subjects factor, and condition (i.e., control versus treatment) as the between-subjects factor, we observed that in the control group, there was a slight, statistically significant increase in ELA/ELD scores (13%), and almost no increase in the STEAM scores (3%) (Table 2). In contrast, in the treatment group, there was a statistically significant and relatively large improvement in both ELA/ELD scores (55%) and STEAM scores (51%). While students in the control group improved from 51.83% in the pre-test to 56.09% in the post-test, which represents a statistically small effect, students in the treatment group improved from 47.10% in the pre-test to 72.05% in the post-test, which corresponds to a 53% improvement, and represents a statistically large effect.

Data

Data Acquisition Where will you source your data from? What is the cost to acquire these data? Are there any personally identifying information (PII) or data sensitivity issues you will need to overcome? Will data become available on an ongoing basis, or will you acquire a large batch of data that will need to be refreshed?	Our data will be sourced through implemented pilots in the schools. There is a cost related to the data acquisition which is basically the cost of pilot implementation. The data contains the PII which need to be managed before processing for the AI implementation. As a result of pilot implementation, the ongoing data can be available on a daily basis and the implemented AI can be improve by feeding the new sets of data.
Data Source Consider the size and source of your data; what biases are built into the data and how might the data be improved?	Considering the different demography's of the kids in the school from state to state and also in various geographic locations in the world, the collected data is biased and the suggested food and also the metrics for the success could be different depending on which data types used for the AI training.
Choice of Data Labels What labels did you decide to add to your data? And why did you decide on these labels versus any other option?	The labels which we are going to use is related to the average mark that students can get from the quiz following the cooking of the specific recipe. The main point of the planned educational app is to train the students in a fun way to get the better performance in their education including the math. The quiz mark followed by the cooking a specific recipe could be a very effective approach on determining which recipe is more effective in training the students.

Model

Model Building How will you resource building the model that you need? Will you outsource model training and/or hosting to an external platform, or will you build the model using an in-house team, and why?	As we are a small team, building a scalable model could be a challenge and we do not have the inhouse knowledge and skills to build and scalable and reliable model. Our preference is to build the model using one of the available platforms.
Evaluating Results Which model performance metrics are appropriate to measure the success of your model? What level of performance is required?	The accuracy of the model can be around 80% and it is enough for the purpose of the education. This is not a mission critical application so high level of accuracy and recall are not necessary.

Minimum Viable Product (MVP)

Design What does your minimum viable product look like? Include sketches of your product.	 It is a AI powered app which helps the children around five years old to learn about healthy cooking habits as

	well as math skills.
Use Cases What persona are you designing for? Can you describe the major epic-level use cases your product addresses? How will users access this product?	The product is designed for the children under 6 years old and the product will be an ios and android app as well the web application. The users can access the web application in the school and the ios and the android app through app store.
Roll-out How will this be adopted? What does the go-to-market plan look like?	The go to market plan includes attending the major conferences and exhibitions, demonstration of the product in the local schools in Northern California and the email marketing.

Post-MVP-Deployment

Designing for Longevity How might you improve your product in the long-term? How might real-world data be different from the training data? How will your product learn from new data? How might you employ A/B testing to improve your product?	The following tasks and implementations are required for the post MVP deployment. Performance (minimizing writes & db calls, caching, strict content sizing etc.) Scalability (via modularity, tiered abstraction, externalizing content for dynamic updates, etc.) Internationalization (expanding to Spanish) Easy, Consistent UX (simplified and adhering to style standards) Age Appropriate Content (with versions of lessons for adaptive difficulty and reuse across grades) Two Modes of Use (directed teaching vs exploratory free-play)

	Expanded Teacher Control and Reporting (sharable with parents)
	The real data might be different as a result of different demographic's and also the teaching style in different countries.
	Regarding the A/B testing, the product can be divided into two groups and one group use one additional feature whereas the second team do not have access to the specific feature under study and afterwards the analytics results can be used to make a final decision. One example for the A/B testing could be chatbot option or another option could be possibility of using the amazon fresh single button order.
Monitor Bias How do you plan to monitor or mitigate unwanted bias in your model?	One approach which we are planning is to use different versions of the model for our range of clients as an example specific model can be used for the northern California and different model for the Mexico which the geography and the demography of the people are greatly different.