Capstone Project Proposal

<Rahul Sarkar>

Business Goals

Project Overview and Goal 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?	The business problem that I am trying to solve is how to enable schools to take online examinations by ensuring academic integrity. Due to COVID-19, all schools around the world have been forced to delay their examinations and conduct online classes. School authorities are worried about scheduling the pending exams due to lack of proctored environment which can prevent students from practicing aberrant behaviors. Thus, as an ed-tech startup, the idea is to provide a platform for conducting both online classes and online examinations using an **Automated Proctoring System**. AI can play a crucial role to track the student's behavior on real-time basis during the exam and flagging all the aberrant behaviors, thus helping to identify those students who are trying to cheat during the exam.
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.	As a startup providing online platform to schools for conducting online classes, there has been an increasing demand for automated proctoring system for conducting online exams from existing clients. Designing such a system has following advantages for the startup. 1. This can prove to be a unique offering in the Indian education market as other ed-tech players don't have such an offering currently. Hence, this can provide first-mover advantage to the startup. 2. This system can be sold both independently and by bundling with the existing service packages, thus bringing more revenue for the company. 3. Existing schools will be happy with the service and thus act as a referral source to capture more leads, thus helping to increase the market share.

| **Application of ML/AI**

What precise task will you use ML/AI to accomplish? What business outcome or objective will you achieve? | AI will help us to flag the below-mentioned aberrant behaviors and thus build a robust proctoring system:

• With the help of **face verification**, we can verify the identity of the test taker.
• With the help of **object detection** technology, we can detect mobile phones, books, calculators etc.
• We can detect whether any collaborators are helping the test taker during the test by identifying more than 1 person in front of the camera.
• We can also detect if the test taker leaves during the middle of the test.

Thus, AI can help us to create the Automated Proctoring System, which can be a powerful offering for our company, thus leading to client satisfaction, increased revenue and high market share. |

Success Metrics

| **Success Metrics**

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. | **Acquisition:** How many clients registered for the automated proctoring system
Activation: How many clients have started using this service out of the total registered clients
User Engagement:
• How many tests are being scheduled by the clients on monthly/yearly basis
• How many wrong predictions were flagged by the clients
Retention:
• What is the retention rate of clients
• NPS and CSAT
• Reviews/Comments from customers
• Number of complaints filed by clients
Revenue:
• Revenue earned on monthly/yearly basis
• Market Share on yearly basis
• Cost incurred in sales and marketing
• Net profit of the company
Referral:
Testimonials and Word of Mouth

For baseline value, we can compare with other companies if public data is available. Otherwise, we can fix goals and try to improve our metrics on monthly/yearly basis. |

Data

Data Acquisition	For face verification model, we can use pre-trained models from existing open source tools like OpenCV and test it on some of our school students. In case the accuracy is low, we will have to collect the student identity details from our clients (schools) and our own internal database, annotate them (either in-house or use platform like Appen), train the model and launch it with better accuracy.
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?	And for object detection model, we can use the pre-trained single shot detectors like YOLO etc. Since objects like mobiles, study books/notes, calculator etc. are same throughout the world, we can use the pre-trained model only for our project. In case of training our own model, we have to depend on our own database, open source datasets and third-party annotators for data collection purpose.
	Once we are ready with the proctored system, we can start collecting our own data during the online examinations and use them for re-training the model in future.
	The entire cost for data acquisition will be free in future as we will store the video recordings of students on our servers and let our clients know about it beforehand only.
	For ensuring data sensitivity, we will make sure to protect the data from hackers by installing firewalls and always ask opt-in from our clients before conducting proctored tests.
	Since live data will be recorded on every test, new data will be collected on an ongoing basis.
Data Source	If we need to train our models for better accuracy through transfer learning, then the approach will be as follows:
Consider the size and source of your data; what biases are built into the data and how might the data be improved?	For face verification model, we need 2 models: one to identify the face with a bounding box and another to convert the faces into a feature vector of n dimension for finding similarity between faces. For the first model, we can use our object detection model only.
	For the second model, we can use around 1200 images for building the model. (*details are given below*)
	For the object detection task, we will define categories like human faces, mobiles, books and calculator. For each of these categories, we will collect around 250 annotated images, thus keeping the data balanced with a total of 1000 (250*4) images.

	Some biases which might be present in the data are: • First, the data that we would collect will be country-specific. e.g. if we are operating in India, we can collect data of Indian students only and retrain our models. But it might not work accurately for detecting faces from countries like China, America etc. as the training data is not representative of the entire sample. • Secondly, the camera quality, lighting and ambience of the student will affect both the output/prediction and the data quality. For the first case, we have to ensure that before entering any new market, we test the accuracy of our existing system on some new test data. In case its low, we will have to purchase annotated data from some 3rd party for training our model and go for transfer learning. For the second case, we can provide guidelines to students before their exam starts and ask them to arrange for laptop/mobile with better camera and select a better lighting environment during the test.
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?	For the face verification, we need to build the model for converting the face into feature vector of n dimension. For this we will train the model by comparing 3 images at a time (1st and 2nd image from the same person and 3rd from another person) and the objective will be to minimize the loss function: $d(i1,i2) <= d(i1,i3)$ i.e. the distance between image 1 and 2 will be less than the distance 1 and 3. Thus, we will keep performing the same task on new images (3 images per iteration) and adjust the model weights accordingly using the backpropagation method. Later, during the model evaluation, we will define 2 output labels: "match" and "mis-match" for testing our model. And for the object detection model, the output label will be the final category of the object with the bounding box on the image. Output labels will be: human face, mobile, study books and calculator. We will keep tracking more data on real time basis from proctored tests, that can be used for re-training our model periodically.

Model

Model Building	As lot of open source pre-trained models are already

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?	available on face detection and object detection, I would like to go for in-house development by our internal data science team for this project. Initially, we will be testing the performance of pre-trained open source models for both face detection and object detection on a small dataset collected from schools and our own database. If the performance is poor, we will collect more data from our school and internal database, annotate them (either in-house or third party) and go for transfer learning for building the final model using our internal data. We will ensure all the data collected in our platform is safe by implementing all secure measures and firewalls. Also, before collecting the sensitive data like student identity details and video recordings, we will ask for the opt-in of our school clients, plus notify the student before the examination that the video recordings will be used for training and development purpose.
Evaluating Results Which model performance metrics are appropriate to measure the success of your model? What level of performance is required?	We will divide our dataset into 80% training, 10% validation and 10% testing. Then we will analyze the model performance using 4 metrics: precision, recall, F1 score and accuracy. In case of the model used to convert faces into feature vector, we will create a confusion matrix of 2 labels: match and mis-match. All the above 4 metrics will be verified in this case and ideal score for all 4 metrics would be >=90%. For the object detection model, we will create the confusion matrix with 4 labels (faces, mobiles, calculator and books). For accuracy of bounding box, we will measure IOU (Intersection over Union): IOU = Area of overlap/Area of Union Here also, we will verify all the above 4 metrics in addition to IOU and ideal value for each will be >=90%.

Minimum Viable Product (MVP)

|	
Design What does your minimum viable product look like? Include sketches of your product.	Added in the last page

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 prime users of our product are students and teachers of schools. The main use cases of our proctoring system are as follows: • Face Verification: Identifying whether the test taker's face matches with the image present in the identity card. • Check whether the test taker is performing any aberrant behavior like using items like mobile, books, calculator during the test. • Check whether the test taker is receiving any external help from friends/peers. • Check if the test taker is absent in the middle of the test for some reason. Students/Teachers will have to create their account while sign up process. Once logged in successfully, they will be able to use all the services hassle-free.
Roll-out How will this be adopted? What does the go-to-market plan look like?	Go-To-Market plan for this product: **Pre-Launch Strategy** First, we will position our product using 4Ps framework: Product: Our automated proctoring system helps to flag the aberrant activities of students during the test without needing any manual intervention, thus saving both time and money for schools. Price: We will target tier 1 and tier 2 schools, and thus price our product on medium price range. We will provide discount if schools purchase both online platform and proctoring service. We will maintain 4 packages (monthly, quarterly, half-yearly and annual) on a subscription basis Place: Since we are into B2B business, we will focus on our internal sales team for building long term partnership with new schools. For existing schools, we will try to cross sell our new offering by giving attractive discounts to them. Promotion: For attracting schools, we will provide discounts and 1-week free trial so that schools can try out our product before purchasing it. During the product launch, we will host public events to create awareness in the market. We will focus on digital marketing through channels like newspaper, google ads and social media, but larger focus will be on sales team. **Post-Launch Strategy** We will track our engagement and retention metrics very closely to find out whether our clients are facing any problem and improve all the pain points. For the satisfied customers, we will try to further develop our product offerings. We will check product-market fit and focus mainly on the needs of the satisfied customers for product development.

# 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?	For future improvement, we would focus mainly on our engagement and retention metrics initially. We would interact with our customers closely on regular basis to find what are the areas where they are facing problems and improve them on phase wise manner. For product enhancement, we would focus on how the existing offerings can be further improved and also ask our satisfied customers for their suggestions. Real data might contain noise due to external factors like user ambience, camera quality etc. Also, the training data might not represent the entire student population, thus leading to false positive and false negative cases. For re-training the model with new data, we will first need to break the videos into images and then annotate them before training our model. Also, we have to ensure that we are using balanced dataset for training purpose. A/B Testing: After deploying our new model, we will distribute the traffic (60%-40%) between our old and new model and evaluate the performance metrics before permanently moving to the new model.
Monitor Bias How do you plan to monitor or mitigate unwanted bias in your model?	Following strategy can be used to avoid bias: - For data bias, we need to ensure that our training dataset is representative of our actual users. Also, confusing images should be avoided and images with better quality should be used. - For annotating bias, we need to ensure that the images are being annotated correctly. If in-house is proving to be difficult, we should go for third party like Appen, because if we feed wrong data to our models, we will get wrong results as *garbage in will lead to garbage out*. - For model bias, we need to focus on the performance metrics of the model very closely.

Product Screenshot
Sign In/Sign Up
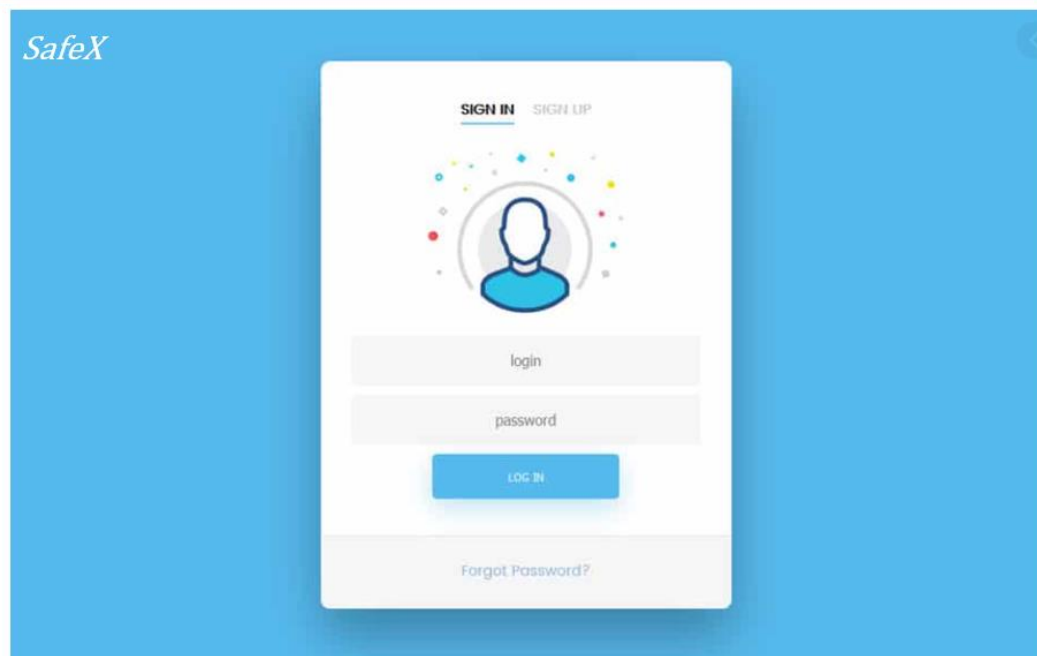
Student Dashboard
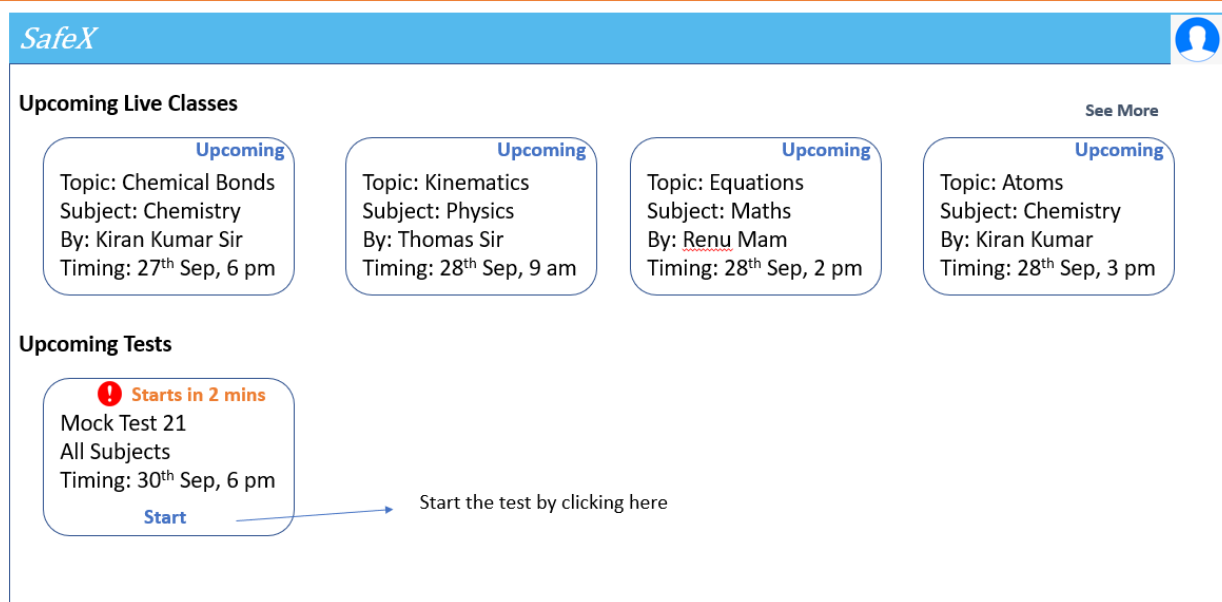
Student Screen

SafeX

Complete details and start test

Step 1. Enter personal info   Step 2. Upload your identity proof   Step 3. Click your picture

Name	Enter your name
Class	Enter class
Roll No.	Enter roll no.
School	Enter school

Next

Student Screen

SafeX

Complete details and start test

Step 1. Enter personal info   Step 2. Upload your identity proof   Step 3. Click your picture

Click or Upload Proof

Next

Student Screen

Student Test Screen

Instructions

1. Total duration of this test is **180 minutes**.

2. There will be **90 Single Option Correct MCQ questions.**

3. There are **3 sections** in the question paper consisting of Physics, Chemistry, Mathematics, having **30 (Physics), 30 (Chemistry), 30 (Mathematics) questions** respectively.

4. Each question is allotted **4 (Four)** marks for correct response.

5. **1 (One) mark will be deducted** for indicating incorrect response for each question.

6. **No deduction** from the total score will be made if no response is indicated.

7. **To activate RankUp - Guaranteed improvement plan, you must spend atleast 30% of the time allocated to the test.**

8. Please use ONE device, browser and tab while taking this test. As an anti-cheating measure, switching between tabs, browsers or devices will result in loss of your test data.

9. To get access to your Advanced Test Feedback, ensure that you have:
   - spent at least 10% of the total test time and
   - attempted 10% of the total number of questions in the test.

☑ I have read and understood the instructions.

Start Test

Student Test Screen

Q 1 - Single Choice Question   🔖 Mark For Review

A massless rod BD of length l is suspended by two identical massless strings AB and CD of equal lengths. A block of mass m is suspended from point P such that BP is equal to x. If the fundamental frequency of the left wire is twice the fundamental frequency of right wire, then the value of x is

PHYSICS (30)

1 2 3 4 5
6 7 8 9 10
11 12 13 14 15

< PREVIOUS   CLEAR SELECTION   NEXT >
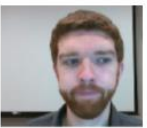
Teacher Screen
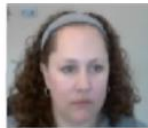
SafeX

Mock Test 21
20 test takers

Search student by name 🔍
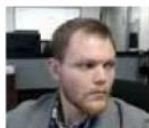
5 Flags violated

3 Flags violated

3 Flags violated   Click on student to see detail info

1 Flags violated

0 Flags violated

0 Flags violated

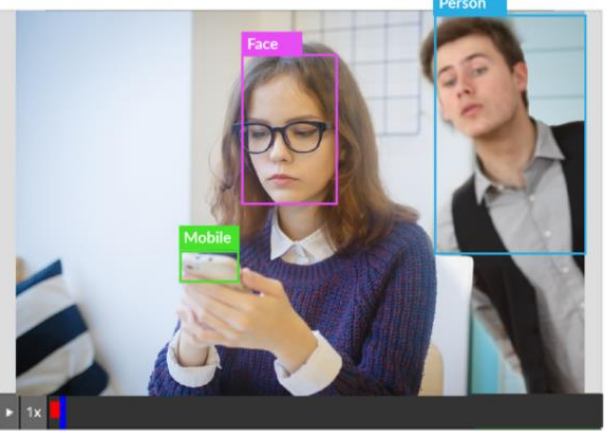Chat(24)   ✕

Margaret

Amy

Steve

Robert

Timothy

Teacher Screen

SafeX

Mock Test 21   Student Name: Margaret   Pause Test   End Test

Person

Face

Mobile

► 1x

Student Log

Face Verified: ✓   Object Detection in action

Total Flags Violated: 2

Additional Faces: Yes

Mobile Detected: Yes

User Absent: No

Calculator Detected: No

Books/Notes Detected: No

Chat with the student