

Capstone Project Proposal



Rosemary Nwosu-Ihueze

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?	<p>Threat detection is a huge and evolving field. Threat detection can come in form of cyber protection mechanism, or software in perceived important places. What of personal safety? How does one detect danger to themselves?</p> <p>This proposed product is a threat detection mechanism for mobile phones that can detect danger using energy emissions from weapons.</p> <p>Data is collected using heat wave detection software and labelled in accordance with weapon and fed to the model. Classification will be for segregating and regression for prediction.</p>
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.	<p>This is important as it not only improves personal safety of users, it also reduces money spent by the government in threat detection.</p> <p>This project is proposed to generate revenue for our team in the following ways:</p> <ol style="list-style-type: none">1) Private company partnership like the telecommunication companies wanting this feature incorporated into their sim cards and phones while software companies will try to out compete each other in funding to make a better detection model.2) The more consumers use this feature, the more traction the company gains through traffic.

	<p>3) The government would also be interested in funding the project as it reduces the expenditure and a faster and convenient way to detect threat early before it happens.</p>
<p>Application of ML/AI</p> <p>What precise task will you use ML/AI to accomplish? What business outcome or objective will you achieve?</p>	<p>My proposed project will use ML/AI to detect and differentiate heat or energy from the human body and various weapons, to enable threat detection.</p> <p>The models will utilize neural networks to automatically</p> <p>1) Detect type of weapon in the vicinity by assailants</p> <p>2) Differentiate human heat emission from weapon heat emission for accurate recognition.</p> <p>3) Detect human emotions such as aggressiveness.</p> <p>The outcome to be achieved is improved security in neighborhood and early alert to potential threats to the security agents.</p>

Success Metrics

<p>Success Metrics</p> <p>What business metrics will you apply to determine the success of your product? Good metrics are clearly defined and easily measurable. Specify how you</p>	<p>Success metrics will be tested not only from the performance of the model from real world data, but also from the performance on MVP testing.</p> <p>Success metrics will be measured by the decrease in customer acquisition cost over time as our target market will be properly defined and email segmentation.</p>
---	---

will establish a baseline value to provide a point of comparison.	<p>It would also be measured by the traffic to the website through SEO improvement and faster load time.</p> <p>We will also measure through our baseline performance traffic which is 10k users within 3 months post-launch.</p>
---	---

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?	<p>The dataset will be sourced from LWIR segment of an electromagnetic camera and military weapons detection which will be compared and isolated to get a comprehensive dataset of heat emission from varying hand weapons and the human body.</p> <p>It would cost a lot to do this and since there would be military detection dataset, there would need to be cleared to undertake such task.</p> <p>There won't be data sensitivity issues as what is being extracted is thermal or heat energy.</p> <p>A large amount of data will be needed to start the process with continuous refreshment as more data is sourced and added.</p>
Data Source Consider the size and source of your data; what biases are built into the data and how might the data be improved?	<p>The bias will be from labelling, comparison, and isolation of dataset.</p> <p>This data will be improved by engaging with domain experts from the inception of the project to minimize anomalies to the barest minimum.</p>
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?	<p>The labels I would be working with is dependent on the range of weapons chosen and the emotions to be considered.</p> <p>The MVP will contain labels for: Guns, hand grenades, knives, emotions like anger, sadness, depression.</p> <p>The strength of the labelling scheme is that data collection would be well defined and annotators having a</p>

	<p>clear picture of what to expect.</p> <p>The downside to this scheme is that different guns give off different heat signal and may cause bias in model performance.</p>
--	---

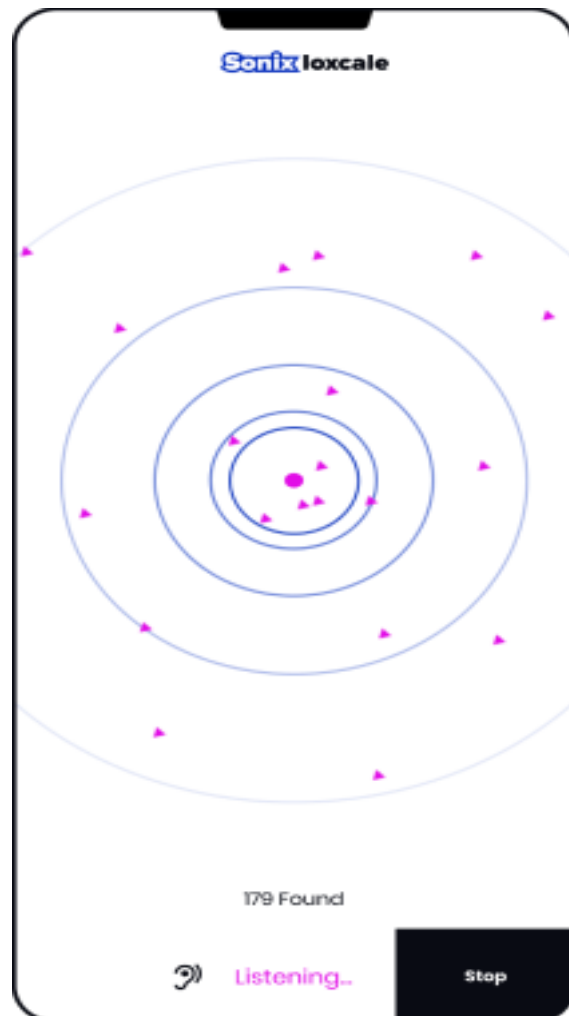
Model

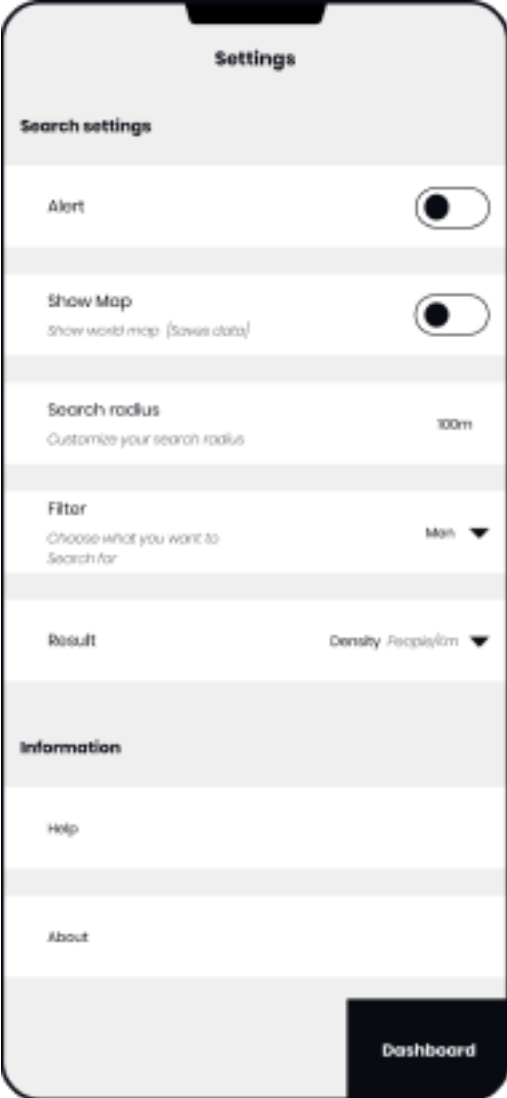
<p>Model Building</p> <p>How will you resource building the model that you need? Will you outsource model training and/or hosting to an external platform, or will you build the model using an in-house team, and why?</p>	<p>The model will be built using the in-house team as the team is comprised of domain experts. The storage will be outsourced as large amount of data will be used.</p>
<p>Evaluating Results</p> <p>Which model performance metrics are appropriate to measure the success of your model? What level of performance is required?</p>	<p>Metrics to be used include the ROC AUC curve, precision, and recall. The performance level required is at least at 85% as this is a proposed security apparatus and needs as little FPs and FNs as possible.</p>

Minimum Viable Product (MVP)

Design

What does your minimum viable product look like? Include sketches of your product.



	
<p>Use Cases</p> <p>What persona are you designing for? Can you describe the major epic-level use cases your product addresses? How will users access this product?</p>	<p>I'm designing for night walkers who enjoy a bit of the outdoors.</p> <p>Also, for anyone who feels unsafe walking around in a dangerous neighborhood especially women.</p> <p>Major case: A young lady leaves work very late and takes a train home. On her way home, she feels uneasy as though she's being followed. The product detects unusual energy such as aggressive emotions or a weapon close by, it alerts the authority first and then signals the lady using the secret tone setup by the user.</p> <p>This product can be accessed by being integrated in the phone, by phone companies and sim cards by the</p>

	telecommunication companies which means the product is in-built.
Roll-out How will this be adopted? What does the go-to-market plan look like?	<p>Like I mentioned earlier, it would be integrated into phones and sim cards with software update incorporating it into existing user's phones. No need for go-to market plan.</p> <p>The roll-out will include the following milestones:</p> <ol style="list-style-type: none"> 1) Hire team comprising of: Product owner, product designer, data scientists, ML engineers, QA, weapons specialists and 2 software engineers. This team is expected to be assembled in two months. 2) Generate labeled dataset: We will need to hire an external data annotation company to label the dataset with domain experts vetting this labelling process. We expect this to be done in a month. 3) Build and train model: We expect model building and training to take about 2-3 weeks for a first iteration. 4) Test model, and iterate until precision meets specification: We expect the testing/iteration phase to last around 2 weeks before launching the MVP. 5) Launch: Invariably, we think that we can get to launch in about 6-8 months. 6) Iteratively add new video/image content: We plan to inject a new set of data into the model and re-train on approximately a monthly cadence, and we'll need the data scientist and ML engineer to set aside about a day each time this needs to be done.

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?	<p>As detections are being confirmed, it acts as the feedback to the model. With time more labels would be added, the model retrained on these updated labels and tested to ensure efficiency.</p> <p>There won't be difference between training data and real world as the data used for training are real world data curated by the team.</p> <p>After new model addition, 30% would be used on the updated model and 70% on the current model and performance compared.</p>
Monitor Bias How do you plan to monitor or mitigate unwanted bias in your model?	<p>User feedback will mostly be used to mitigate unwanted bias.</p>