|  |  |
| --- | --- |
| Capstone Project Proposal |  |

*Ayeni Trust Olamilekan*

**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? | * I want to eliminate race and gender biases in voice and speech recognition systems, these aspects of AI now plays an important role in human daily activities including immigration decisions, job hiring, especially in tech gadgets (cars, bikes, phones, airplanes) among many other things. Thus, this means that speech/voice recognition accuracy - - or lack thereof – could prevent you from migrating to a country, getting a job, getting some IOT devices or getting a new car, airplane and so on. The reason being that databases have lots of with male data, and less data on female and minority voices (for example what about African, Asian, north American, and other continents voice data).   These factors reduces customers satisfaction or purchasing decisions.   * **ML/AI** should be used in solving this task as it requires lots and lots of voice data probably **Gigabytes** or **Terabytes** of data. Hence data is collected from various countries would prevent AI setup from failing and making the model perform well with data it sees infrequently, such as female and more diverse voices making a huge milestone in the purchase or use of voice recognition systems/devices in companies, offices and various homes.   Therefore, only ML/AI algorithms are suitable when dealing with this spectrum of data. |
| **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. | * Voice AI is becoming ubiquitous and powerful. Forecasts suggests that voice commerce will be an $80 billion business by 2023.   Recently, the Algorithmic Justice League’s voice erasure project found that speech recognition systems from Apple, Amazon, Google, IBM and Microsoft collectively achieve word error rates of 35% for African American voices versus 19% for White voices.  In African countries, as a case study various companies including tech companies are considering tapping into the power of AI to make life easier by making purchases online, front desk automation, voice-based digital assistants, and more.  But due to the bias in VR systems many Africans prefer lower budget technologies/ IOT devices, not that they don’t have the money for purchase but the mentality that “I can’t buy what cannot understand what I’m saying” is what makes Africans ignore those techs.   * It is a good business sense to build this product since:  1. VR systems misidentify words about 19% of the time with white people. With the blacks, mistakes jumped to 35%. 2. About 2% of audio snippets from white people were considered unreadable by these systems while it rose to 20% with black people.   Thus, if this loophole is solved:  The VR model can be implemented into   1. In-car systems 2. Health care systems (medical documentation, therapeutic use) 3. Military (high performance fighter aircrafts, helicopters, training air traffic controllers) 4. Telephony systems (computer gaming, smartphones) 5. Education and daily life (useful in learning a second language) 6. Robotics 7. Real time assistants 8. Aerospace 9. Home automation systems  * Hence, it would be a revolutionary breakthrough in the AI market if these biases are neutralized, as it would bring in substantial revenues from the large footprints of customers and innovators who supports it. |
| **Application of ML/AI**  What precise task will you use ML/AI to accomplish? What business outcome or objective will you achieve? | * I will use AI to train this large corpus of data and fine-tune this multi-class classification model on the voices collected from different continents (countries) and then deploy the model into the cloud and mobile devices through a conversational app. This product will also support online inferences for 3rd party integrated tools, through REST-APIs. * Conversely, anyone who wants to use my product can access it through the cloud or mobile devices, implementing this model to support their intended purpose. |

**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. | **I will use the success metrics to measure target audience awareness such as:**   * Referral traffic   Goal: To aim for more than 50% of traffic. Exact number will depend on the size of my target audience.   * Return visitors   Goal: To aim for 25-50% return visitors  **Success metrics to measure product recommendation:**   * Net promoter score   Goal: To aim for 90-100%  The NPS measures how likely someone is to recommend my product, for this I would use quantitative data than qualitative data.  If my NPS score is low I would find out why?’,  improve on that information and regularly update my model.  I can establish a baseline value by getting data about the metrics listed above from a substantial sample size of users.  The higher the first metric, the baseline for this should be at least 80% when starting off, else there is a huge risk of losing users.  While in the long run, the goal should be to get to 100%. |

**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? | **Data sources**   * Questionnaires: Survey monkey, Google forms. * Forming partnership with various institutions and companies who will be the customers for the product. * Internet   Getting these data would require lots of money, time and energy. Money to purchase/download voice data from the internet, to set up partnership with institutions and companies interested in the model. I would also travel to different countries to collect my own data (primary data).  **PII (personally identifying information)**   * No, I will not be using any sensitive or personally identifying data, but using voice data related to a particular model class.   My product would use a large amount of data that will be refreshed every 2 months. |
| **Data Source**  Consider the size and source of your data; what biases are built into the data and how might the data be improved? | **Biases**  Bias such as race and gender bias can still occur if the product becomes widely acceptable, many users in countries/continents whose voice data have not been collected also, users with different intonation, pitch, language can be at greater loss.  All of this points to the fact that our data, to a certain degree, might not be inclusive and diverse enough to represent remote people thus skewing the model’s learning and inference to favor a particular class than another.   * Measurement bias:   Differences in the data collected for training would differ from the data collected during production.   * Algorithm bias:   This bias is caused by how the model was developed or how the model was trained that results in unfair outcomes.   * Exclusion bias:   I might add/remove features that I think are relevant or not without bringing in a domain knowledge expert to conduct feature engineering.  **Improvement**  The data can be improved through:   * Data profiling: this is used to examine the data defects analyzing the correctness and usefulness of data. * Data normalization: My model definitely contains data from different sources and includes a variety of spelling options. Normalization is the process of reorganizing data so that it can be easily accessed. * Quality control on the data being collected. * A/B Testing on a few ML models to check for biases. * Bringing in a domain expert to work out features to be included or not. |
| **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? | **Labels**   * **Male**: This label contains data of female voices collected from different countries. * **Female:** This data contains data of male voices collected from different countries. * **Country:** This label consists of various countries data was collected from and different male and female voice data tagged with it.   I am using these labels because it precisely arranges my data making it easier for me to normalize the data-set. |

**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? | I will resource building the model by creating models that test different costings, budgets and outcomes. Figuring out cost and how long it will cost, how long it will take. Comparing different models to find out the most efficient and cost-effective approach to completing such project.  I also intend to build the model using an external platform which, will then expand to a robust ML team and product team. |
| **Evaluating Results**  Which model performance metrics are appropriate to measure the success of your model? What level of performance is required? | I would use **precision** as the main metric since it is an  Here, I need to aim for high precision at a reasonable threshold.  A precision level **> 80%** is what I should aim for. |

**Minimum Viable Product (MVP)**

|  |  |
| --- | --- |
| **Design**  What does your minimum viable product look like? Include sketches of your product. | The MVP will be a cross-platform speech recognition app and a REST-API. The app would first display a conversational user interface that collects voice data from the user and his/country, adding this data to the model’s corpus of knowledge (pre-trained model).    The app is then able to recognize spoken words from anyone in any country.  Furthermore, the REST API allows for third party applications to use the hosted online model for inference.  The sketches of the product interface are below: |
| **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 main personas:**   * Students * Tech Companies planning to use the product in their services * Individuals * Military   **Use cases:**   * **In the educational sector**, students who are blind or have very low vision can benefit from using this technology to convey words and then hear the computer recite them, as well as using a computer by commanding it with their voice, instead of having looking at the screen or keyboard. * **Tech companies: (automobiles)** simple voice commands can be used to initiate phone calls, select radio stations, play music from a compatible smart phone, MP3 player or flash drive.   Voice recognition capabilities vary between car make and model. Some of the most recent models offer Natural Language speech recognition in place of a fixed set of commands, allowing the driver to use full sentences and common phrases with no need of memorizing a set of fixed command words.   * **Military: In fighter air crafts**, Speech recognition is used for a wide range of cockpit functions.   Voice commands are confirmed by visual and/or aural feedback. The system reduces the pilot workload and even allow the pilot to assign targets to his aircraft with simple voice commands.   * **Medical sectors:** The model can be implemented in front-end or back-end of the medical documentation process. A large part of the clinician’s interaction with the **EHR** (Electronic Health Record) involves navigation through the user interface using menus, and tab/button clicks which is heavily dependent on keyboard and mouse. **Thus, voice based navigation provides only modest ergonomic benefits.**   Finally, users can access the product through the cloud APIs or on IOT devices. |
| **Roll-out**  How will this be adopted? What does the go-to-market plan look like? | Go-to-market plan   1. **Planning phase**(pre-launch): Securing partnerships with educational institutions and tech companies where the system can be tested, evaluated and proven to be effective. 2. **Launch phase**: When tested to be effective, the product can now be made available to other institutions and tech companies for data collection and analysis through university boards, social media pages of other tech companies. 3. **Post-launch**: The product is on solid footing to succeed in the IT market through recommendations from supported users and a strong web presence to quickly deploy and close sales.   Also injecting a new set of audio content into the model and re-training on monthly basis would increase the model’s accuracy. |

**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? | I can improve my products longevity by   * **Identifying customers needs and requirements:** Flexibility is required when balancing the needs and requirements of customers with respect to price and performance of the product. **(Focusing on outcome rather that output)** * **Forming good communication skills:**   The gap between market research and production should be minimum, the more layers of bureaucracy added between, the more likely it is that the design put forward by the research team, and the final product will have several gaping holes.   * **Continuous learning of AI model:** by using smart selection of training data to improve accuracy and discarding any irrelevant data.   The training set won’t know how to deal with value that are not in similar range, due to the fact that data distributions changes over time.  My product will learn from new data over time by displaying an interface that collects voice data from new users in any country and adding the data to its corpus of knowledge. Once, this is done I can train the newly collected data and test it with its older data-set thereby improving the accuracy of the model.  **A/B testing** would prove helpful in improving my product by letting me know what model performs better.  I can implement A/B testing by splitting the collected data into two versions (tested model and not yet tested model),  Training the data to see how it performs with another newly collected data. This method further reinforces my model’s confidence level to perform better when feed with new data. |
| **Monitor Bias**  How do you plan to monitor or mitigate unwanted bias in your model? | **Steps taken to monitor/ mitigate unwanted bias**   * **Narrowing the business problem**:   I want to avoid solving for too many scenarios which would mean needing tons of labels across an unmanageable number of classes.   * **Structuring data gathering that allows for different opinions**. * **Understanding my training data:**   The more I understand my data, the less likely I am to be surprised by objectionable labels   * **Gathering a diverse ML-team that asks different questions:**   People from diverse backgrounds will inherently ask different questions and this can help me catch problems before the model is in production.   * **Thinking about all end-users:**   I can also avoid AI bias by learning to anticipate how people will interact with the technology and what problems might arise in doing so.   * **Annotate with diversity:**   An option is to source from a global crowd of annotators, who can not only provide different perspectives, but also support a variety of languages, dialects, and geographically-specific content.   * **Testing and deploying with feedback in mind:**   By opening a forum and discussion for feedback, will continue to ensure that my model is maintaining optimal performance levels for everyone.   * **Having a concrete plan to improve my model with the feedback:** Receiving feedbacks and giving my model an improved feedback of my own would constantly iterate my model towards higher accuracy. |