|  |  |
| --- | --- |
| Capstone Project Proposal  Detect Emotions from voice |  |

*Fatimah Aljuhani*

**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? | Problem: Devices do not understand feelings, although most of our communication in work and life is through devices  The aim is to build an application that can identify feelings from voice  This application can be used in different fields,  Such as:   * Teachers can use it in online classes if the sound of student is sad, happy, confident …ext. * Can be used in marketing * Someone who wants to know the feelings of anyone who talking with * Customer service   I will use ML/AI to build a model with google AutoML and use this model through google API |
| **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. | Machines have cognitive intelligence, but what if machines could understand people's feelings through their voice?  The output or answers would be different if these devices could understand the speaker's feelings  For example, in Siri, the user will not be answered as an robotic response, but the program will be able to speak in an angry or playful way, depending on how he understands the person he is talking to  This application can be sold recurring payment  I think this application will be unique there is no such application in the market and many people can be happy with it  this application can be develop every time receives new voices to translate it, adds these new information to dataset, and therefore the more data it has, the better its detect of feelings becomes.  Also can be expanded in many different languages in the future |
| **Application of ML/AI**  What precise task will you use ML/AI to accomplish? What business outcome or objective will you achieve? | I will collect my dataset from kaggel website and it is free  [https://www.kaggle.com/uwrfkaggler/ravdess-emotional-speech-audio](https://www.kaggle.com/uwrfkaggler/ravdess-emotional-speech-audio%20)  there are1440 files and it is free  I will use AutoML to build a model and train it  For this model will use a sound classifier  I should start by creating training data (training data is what allows the model to learn to classify data)  Using 4 labels putting them in folders: Happy, sad, usual, angry  This set will make a good model |

**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. | Success metrics are number of application downloads vs active installation  And also can be in positive reviews from users |

**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 source is kaggel website, and training is done using AutiML to create the model  As Data source cost is free no cost for data collection.  Model is deployed to be used online with google API  So the Cost is about 300$ to deploy the model on google cloud  Data sensitivity: the is no sensitive issues or information in the app  To increase accuracy, we may delete some sounds, if I found on the reviews some sounds have a confusion we can delete those sounds  Or even we can add more sounds, there are different extra datasets on kaggel website and it is free too, but retrain the model costs about 100$. so, we can do this process each 6 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? | The size of data is 590.35MB, 1440 files.  Quality may produce some bias the dataset has only 24 actors, To build a good model I think we need more than that, we need files of men, women, kids, in different ages  Almost all the data we have almost speak the same accent , So we may need to add more diversity of accent |
| **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 are   * Happy * Sad * Normal * Angry   I chooses this labels as it is the target of the application to recognize the feelings by input voices |

**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 use google AtoML to build the application model. I will deploy the model on google cloud and access the model online by google API’s  Because if I choose offline model then I cannot retrain the data  But in online AutoMl model I can access the data and retrain it, or add new data whenever I want |
| **Evaluating Results**  Which model performance metrics are appropriate to measure the success of your model? What level of performance is required? | to measure performance of model I will use confusion matrix  the level of performance required increase  precision by 90% |

**Minimum Viable Product (MVP)**

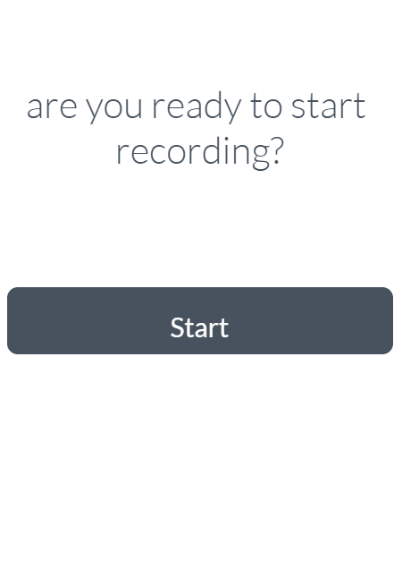
|  |  |
| --- | --- |
| **Design**  What does your minimum viable product look like? Include sketches of your product. | Minimum viable product has 3 faces on the app  It has simple interface first asking the user if ready to start recording  Then ask the user to enter the sound or start recording  Finally show the result if the sound is sad, Happy, angry, normal |
| **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? | I am designing this application for people who need to know the feeling from sound  The customer use case is:   * Request to add voices * Request to record voices   The administrator use case includes:   * Request model retain from AitoML * Request add more sounds to dataset * Request model performance metrics |
| **Roll-out**  How will this be adopted? What does the go-to-market plan look like? | Application will be available on apple store and play store  The plane includes:   1. Testing: we will need partnership with background of computer science using the app to showcase the effectiveness and time saving ability 2. Before launch: once the testing is over the app will deployed for 9$ at the partner companies for collecting data on savings and for demos for other companies during launch. 3. Launch: finally deployed the app on for iphone and android for 29$ for 3 months |

**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 the application by adding sounds with different accents or even different languages also I can add more labels like afraid, crying…ext    Real world data might have new types and different quality than training data  My application does not need A/B test |
| **Monitor Bias**  How do you plan to monitor or mitigate unwanted bias in your model? | I plan to review model statistics through confution matrix  If any bias is detected I will retrain data or add more data or delete the confused data |

Appendix

1. Phase 1



1. Phase 2



1. Phase 3

