Improving the Accuracy of Voice Detection for SocioApp with Deep Learning, Spring 2023

Presented by: Winfred Afeaneku

Advisors: Dr. Jaisinghani, Dr. Diesburg, Dr. Berns

Date: May 12, 2023

UNI / University of Northern Iowa.

Problem Statement

Problem: The accuracy of the Neural Network when deciphering between 2 individual had a successful prediction of around ~50 which is merely a coin toss.

Hypothesis: The accuracy of neural networks in distinguishing between two individuals initially is often only around 50%, equivalent to a coin toss. Furthermore, increasing the amount of data given to neural networks may improve their accuracy in prediction.

Aim: Devise a mechanism to passively detect social status of a user and propose prospective remedies.

Current Goal: Train a Neural Network to decipher between 2 individuals accurately

Related Work

"StudentLife: Using Smartphones to
Assess Mental Health
and Academic Performance of
College Students." Mobile Health,
2017
Authors: Wang, Rui et al.

- Results from the StudentLife
  study show a number of
  significant correlations between
  the automatic objective sensor
  data from smartphones and
  mental health and educational
  outcomes of the student body.

"How Does Social Behavior
Relate to Both Grades
and Achievement Scores?" Front.
Psychol., 4 June 2018,
Authors: DeVries, Jeffrey M, et al.

- Results showed that grades
  correlate more strongly to
  social behavior than test
  scores do at younger ages,
  and that specifically peer
  problems have a stronger
  relationship to academic
  performance

UNI / University of
Northern Iowa.

Background and Motivation

Background
- A student's performance in school is correlated with their mental health
- Social status of a student refers to the rank of a student in terms of interaction with others – professors or peers
- Does not measure sociability through calls and social media
- We are interested in physical interactions

Motivation
- AntiSocial-Behavior has negative effects, if prolonged
- Smartphones today are packed with sensors that we could use
- Students who discuss their challenges with others would be more likely to graduate with better grades

Advancements since earlier work

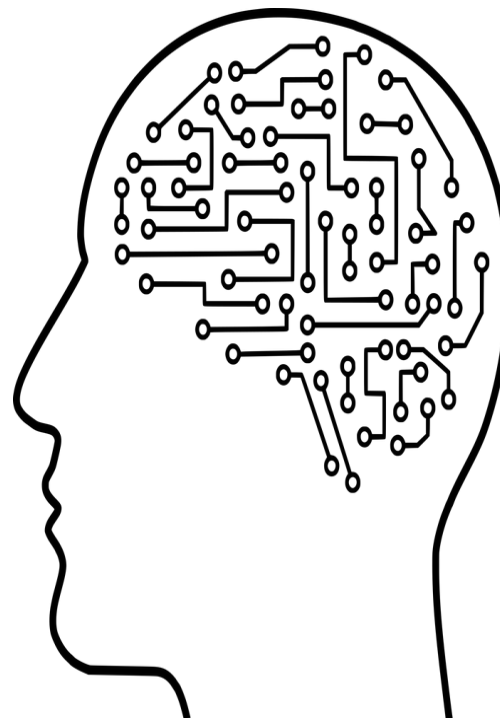**Winfred Afeaneku Fall 2022**
- Audio_Cutter.py
    - Takes in Audio
    - Outputs cut up pieces of audio
- Ffmpeg_convert_wav_to_flac
    - Takes in wav files
    - Outputs flac file with desired frequency and bit rate
- Collected 3 hours worth of Audio data from 2 participants
- Prediction accuracy of model
    - ~50%

Winfred Afeaneku Spring 2023
- ExtractData.py
    - Cuts Audio data
    - Turn Audio data into flac with desired output
    - Turns flac data into features
- Collected 22 hours worth of Audio data
- Prediction accuracy of model
    - Varies depending on constraints

UNI / University of Northern Iowa.

Implementation Details

- Pixel 5a
  - Microphone sensor used to collect audio from participants
- Tensorflow
  - Open source machine learning library
  - Like the Big Engine
- scikit-learn
  - Open source machine learning algorithm for python
- Keras
  - High level deep learning API
  - Provides a friendly interface to use Tensorflow

Dataset Details

- Recordings were done in both a Quiet and Loud environment
- Participants given a pixel 5A
- Duration: 30-60 second clips of them speaking into the phone
- Speakers: Winfred Afeaneku and Aaron Walker
- Genders: Male
- Size of dataset: 3.02MB 1,902 sample file
    - Fall 2022 dataset: 1.28 MB, 814 sample files
- 2 participants
    - Fall 2022 I had 2 participants
- Script 1: 72 Harvard Sentences
    - Fall 2022 I used the same sentences
- Script 2-4: The Eye of the World
    - Fall 2022 I used only script 2

UNI / University of Northern Iowa.

Script Table

This table shows what script and pages will be read on each day over the course of a week.

indicates the pages are subject to change
indicate the new pages for each script 3/20

Days	Scripts/Pages
Day 1 sc1	Harvard Sent.
Day 2 sc2	Pg. 9-22
Day 3 sc3	Pg. 23-35
Day 4 sc4	Pg. 36-59* 36 - 51**
Day 5 sc5	Pg. 60-73* 52-64**
Day 6 sc6	Pg. 74-87* 65-75**
Day 7 sc7	Pg. 88-101* 76-87**

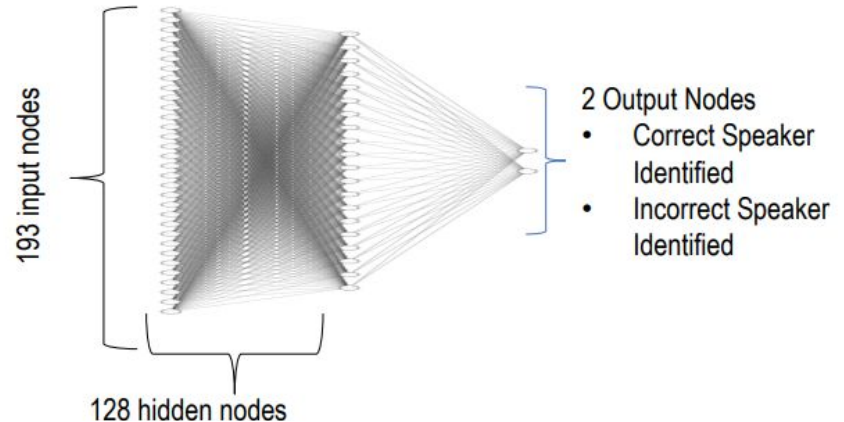**UNI** / **University of Northern Iowa.**

Dataset Details

- Harvard Sentences examples
    - The birch canoe slid on the smooth planks
    - Glue the sheet to the dark blue background.

- The Eye of the World examples
    - The Wheel of Time turns, and Ages come and pass leaving memories that become legend.
    - Legend fades to myth, and even myth is long forgotten when the Age that gave it birth comes again.

University of Northern Iowa.

Neural Network

Machine learning Algorithm
- Data Samples
    - Large Vector for each sample
- Large Vector
    - Consisted of 193 nodes (features)
    - Clips transformed into features in single data sample
    - Datasample is train on a machine learning algorithm
    - Condense 193 nodes into 2 output nodes
- 2 output nodes
    - 1 for correct speaker identified
   - 1 for incorrect speaker identified

193 input nodes

128 hidden nodes

2 Output Nodes
- Correct Speaker Identified
- Incorrect Speaker Identified

UNI / University of Northern Iowa.

Quiet and Alone Environment

Batch	Epochs	Accuracy
1	10	65%
2	2	65%
3	5, 6, 8	60%
4	1, 2, 3, 9, 10	50%
6	2, 5, 8	55%
7	7	60%
9	3, 5, 7	55%
10	6, 9	60%

Batchsizes over Epochs in Quiet Env and Alone
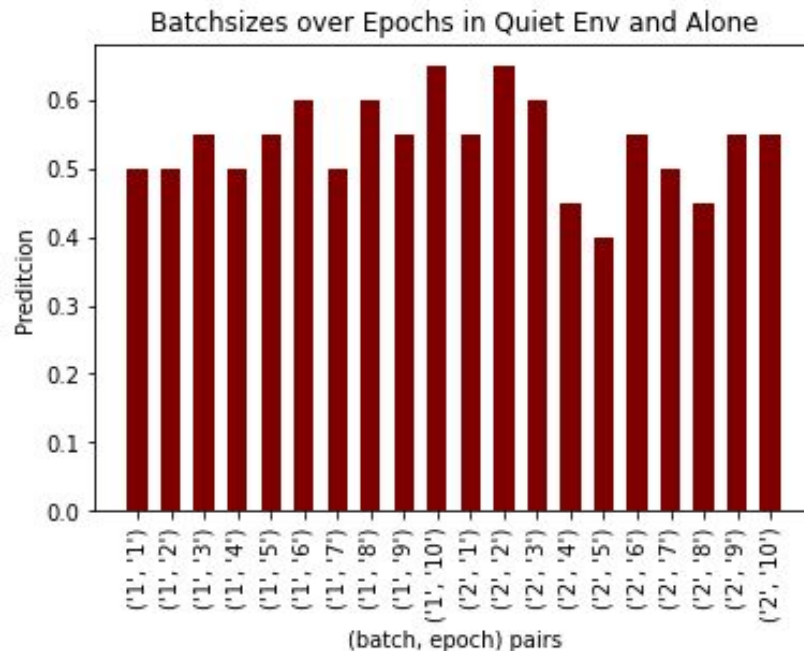
Loud and alone Environment

Batch	Epochs	Accuracy
1	9	55%
4	7	60%
4	3, 7, 8	60%
5	10	65%
6	7	70%
8	7	60%
9	4	65%
10	4	70%

Batchsizes over Epochs in Loud Env and Alone

UNI / University of Northern Iowa.
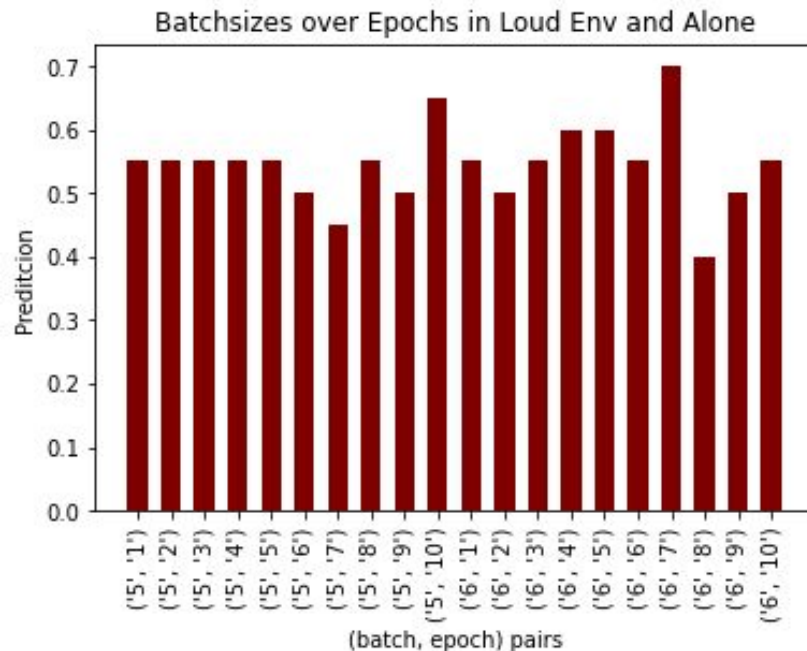
Quiet with Peer Environment

Batch	Epochs	Accuracy
1	3 & 8	65%
4	1	60%
5	3,5,8,9	55%
6	8	60%
7	9	50%
8	2,3,4,5,7, 10	50%
9	1 & 2	65%
10	10	60%

Batchsizes over Epochs in Quiet Env with Peer
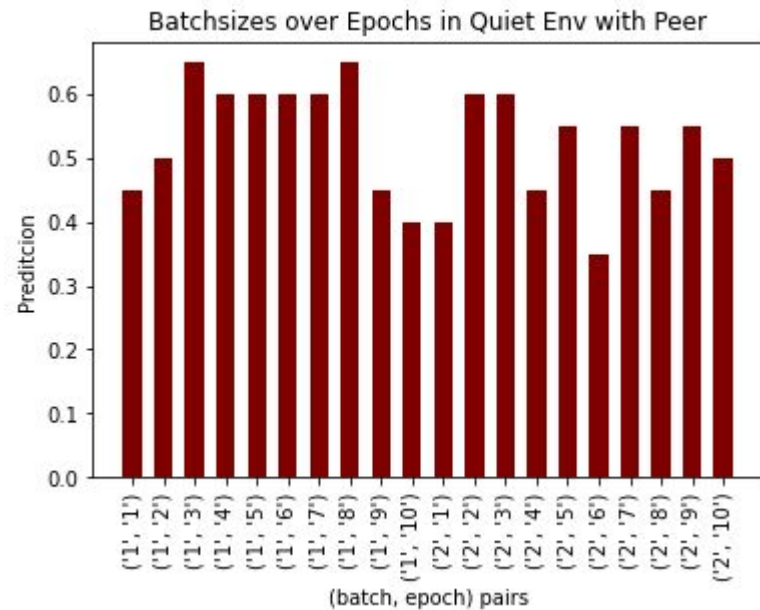
Loud with Peer Environment

Batch	Epochs	Accuracy
1	6	60%
3	2	55%
5	9	55%
6	2, 3	50%
7	9	50%
8	2, 3, 7	50%
9	7, 8	50%
10	3, 4, 7, 8, 9 10	45%

Batchsizes over Epochs in Loud Env and Alone
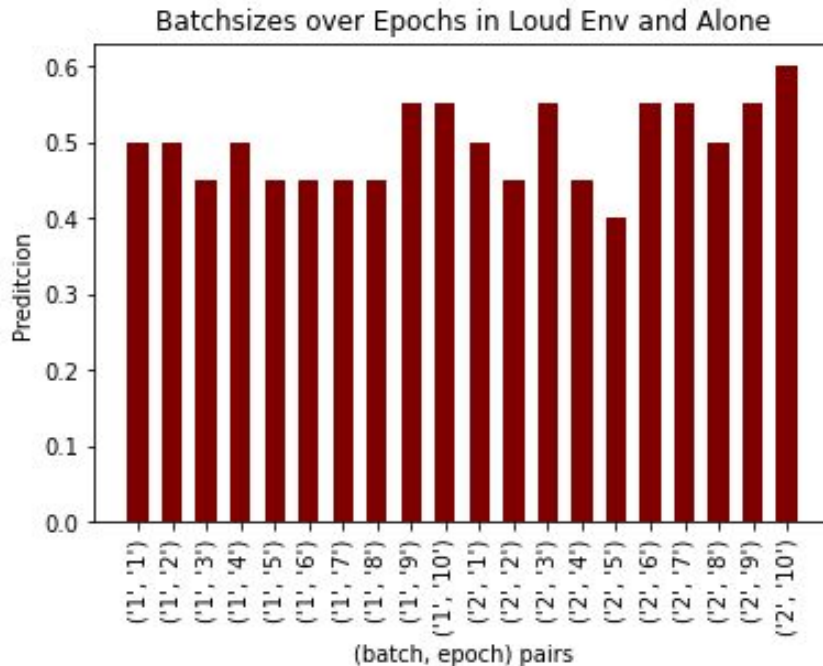
Results of Each Graph

- Quiet and alone
  - (batch= 1, epoch 10)
  - (batch= 2, epoch 2)
  - 65%
- Quiet with Peer
  - (batch= 9, epoch 1)
  - batch= 9, epoch 2)
  - 65%
- Loud and Alone
  - (batch= 6, epoch 7)
  - (batch= 10, epoch 4)
  - 70%
- Loud with Peer
  - (batch= 1, epoch 6)
  - 60%

Drawing Meaning
- Environment plays a role in accuracy
- More data increased the accuracy of the model

Challenges/Limitations

- TIME
- Not enough data samples
    - Meaning not enough data that can get a accuracy around 90% or better
- Small pool of participants
- Getting consistent audio recording each week from participants

University of Northern Iowa

What did I got out of this research?

- Applied ML with Python (Coursera course)
- Insight on Machine Learning
    - Environment Set Up
    - Learned new Algorithms
- Learnt planning project execution with deadlines - I was the BOSS
- Presented research work to professionals from academia and industry
    - IINSPIRE LSAMP Conference in Chicago, October 2022
    - IINSPIRE LSAMP Conference in Iowa City February 10, 2023
- Won the Most like poster award at the Iowa City LSAMP Conference
- Learnings from the Course Intermediate Computing helped me understand the implementation of SocioApp as an Android Application and the Course intro to Computing helped me with Python
- I was able to bring classroom learnings to the real world application

UNI / **University of Northern Iowa.**

Wrap up/Timeline of project

- In the summer of 2022, I began the SocioApp research
    - Summer 2022, learn about machine learning
    - Fall 2022, record audio data and test
    - Spring 2023, record and analyze models accuracy
- I learned how to manage my time effectively, and work efficiently
- I developed the ability to simplify complex ideas and communicate them clearly and understandably
- I became more open to trying out new ideas and approaches
- I will be passing the project off to another student
- I will be researching a new topic in the Summer.

UNI / University of Northern Iowa.

Thank you

Future Work

- Collect more data in different locations
- Collect Large amount of data from many people from diverse backgrounds
  - Age, Ethnicity, Male, and female
- Improve the Neural Network for identifying speakers precisely
  - Look into what vectors neural networks use the most
  - Find out what shortcuts the machine learning algorithm takes
  - Find the best epoch number that reduces error of accuracy close or to 0
- Develop models for detecting and correlating human activities conducted in proximity of others
- Create a metric to measure Sociability

SocioApp: Detecting Your Sociability Status with Your Smartphone

Presented by: Winfred Afeaneku

Advisors: Dr. Jaisinghani, Dr. Diesburg, Dr. Berns

Date: December 16, 2022

UNI / University of Northern Iowa.

Problem Statement

Problem: A lot of students are not getting the most out of their education and factors such as interacting with others, and not seeking help when needed are blockaides that hinders students full potential

Hypothesis: If students were informed on their social status they would be more inclined to address it, making necessary changes that would improve their academic performance and social lives

Aim: Devise a mechanism to passively detect social status of a user and propose prospective remedies

Background and Motivation

Background
- A student's performance in school is correlated with their mental health
- Social status of a student refers to the rank of a student in terms of interaction with others – professors or peers
- Does not measure sociability through calls and social media
- We are interested in physical interactions

Motivation
- AntiSocial-Behavior has negative effects, if prolonged
- Smartphones today are packed with sensors that we could use
- Students who discuss their challenges with others would be more likely to graduate with better grades

UNI / University of Northern Iowa.

Related Work

"StudentLife: Using Smartphones to Assess Mental Health and Academic Performance of College Students." Mobile Health, 2017
Authors: Wang, Rui et al.

- Results from the StudentLife study show a number of significant correlations between the automatic objective sensor data from smartphones and mental health and educational outcomes of the student body.

"How Does Social Behavior Relate to Both Grades and Achievement Scores?" Front. Psychol., 4 June 2018,
Authors: DeVries, Jeffrey M, et al.

- Results showed that grades correlate more strongly to social behavior than test scores do at younger ages, and that specifically peer problems have a stronger relationship to academic performance

# Advancements since earlier work

Aaron Walker
- Developed SocioApp App
  - Records audio from users
  - Transforms audio file into flac
  - Transforms flac file into sample file that contains features
- Remote Server
  - Machine learning Algorithm
  - Features are used to train the server
- Prediction accuracy of model
 - ~30%

Winfred Afeaneku
- Audio Cutter
  - Takes in Audio
  - Outputs cut up pieces of audio
- Ffmpeg convert wav to flac
  - Takes in wav files
  - Outputs flac file with desired frequency and bit rate
- Collected 3 hours worth of Audio data from 2 participants
- Prediction accuracy of model
  - ~50%

UNI / University of Northern Iowa.

SocioApp: System Overview
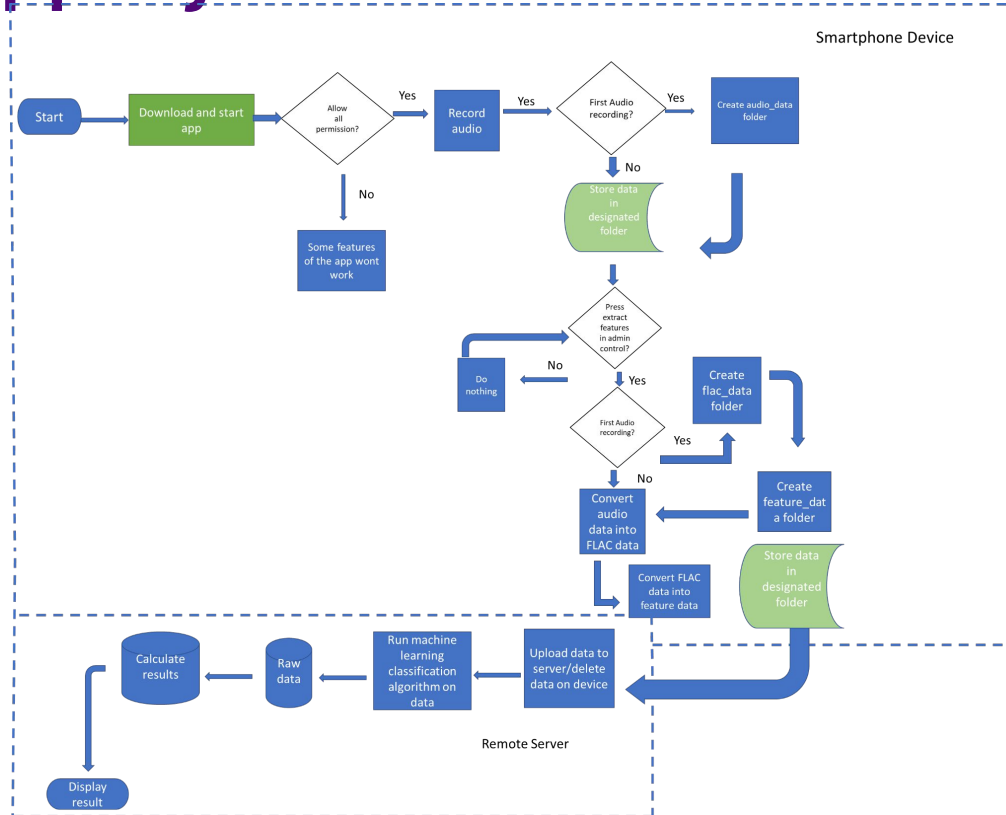
Implementation Details

- Pixel 5a
  - Microphone sensor used to collect audio from participants
- Tensorflow
  - Open source machine learning library
  - Like the Big Engine
- scikit-learn
  - Open source machine learning algorithm for python
- Keras
  - High level deep learning framework
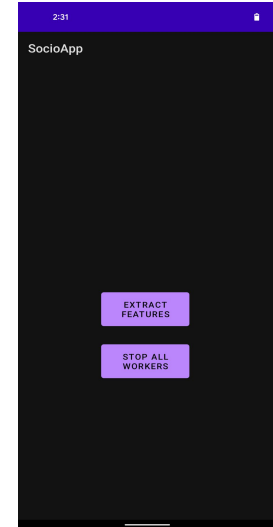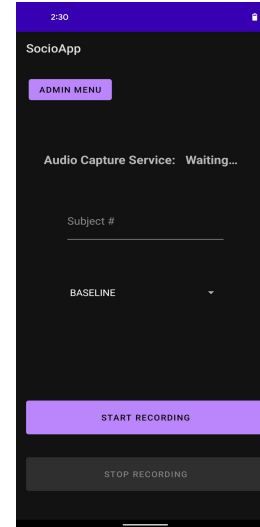
Dataset Details

- All recordings were done in a quiet place, which place
- Participants given a pixel 5A
- Duration: 30-60 second clips of them speaking into the phone
- Speakers: Winfred Afeaneku and Aaron Walker
- Genders: Male
- Size of dataset: 1.28 MB, 814 sample files
    - Aaron dataset: 228 sample files
- 2 participants
    - Aaron had 4 participants
- Script1: 72 Harvard Sentences
    - Aaron used this script
- Script2: The Eye of the World
    - Aaron did not use this script
- 3 hours worth of recorded data in quiet environments turned into 30 -60 second clips
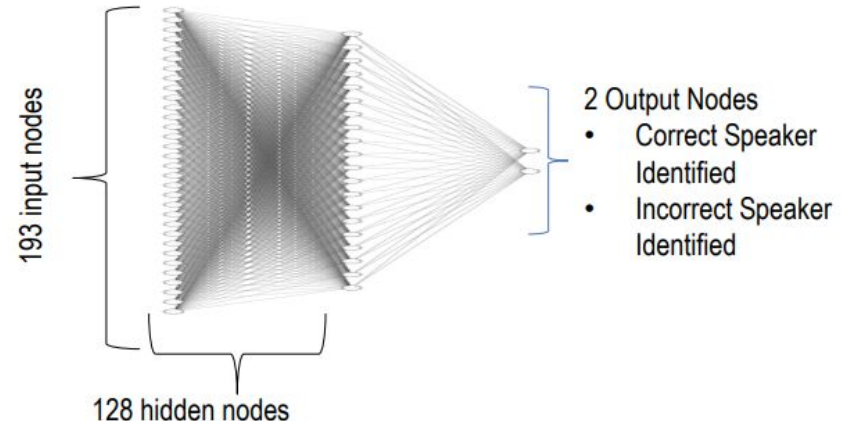- Researcher collects the phone each week and extract features

Dataset Details

- Harvard Sentences examples
    - The birch canoe slid on the smooth planks
    - Glue the sheet to the dark blue background.

- The Eye of the World examples
    - The Wheel of Time turns, and Ages come and pass leaving memories that become legend.
    - Legend fades to myth, and even myth is long forgotten when the Age that gave it birth comes again.

**UNI** / University of Northern Iowa

Neural Network

Machine learning Algorithm
- Data Samples
  - Large Vector for each sample
- Large Vector
  - Consisted of 193 nodes (features)
  - Clips transformed into features in single data sample
  - Datasample is train on a machine learning algorithm
  - Condense 193 nodes into 2 output nodes
- 2 output nodes
  - 1 for correct speaker identified
  - 1 for incorrect speaker identified

193 input nodes

128 hidden nodes

2 Output Nodes
- Correct Speaker Identified
- Incorrect Speaker Identified
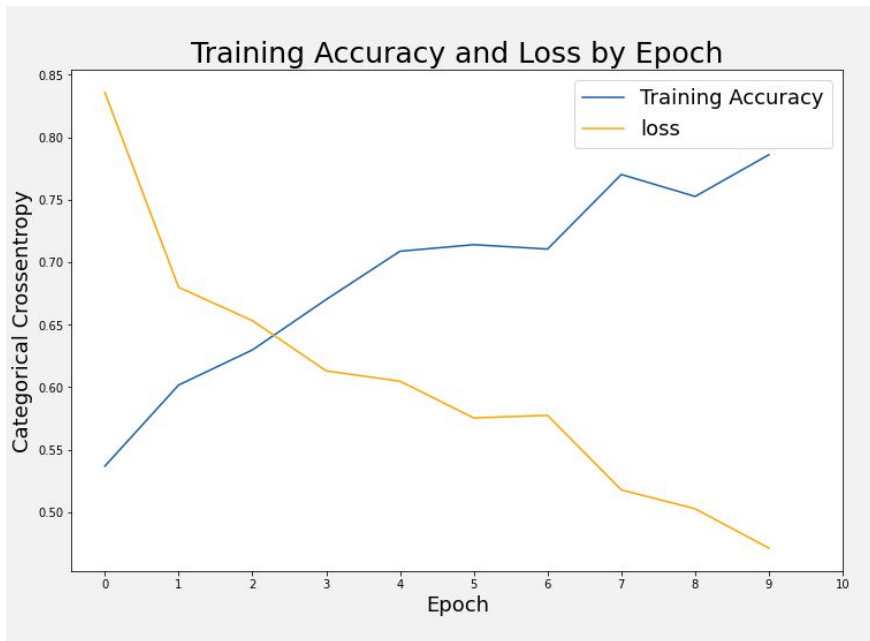
UNI / University of Northern Iowa.

Training and Validation Accuracy by Epoch
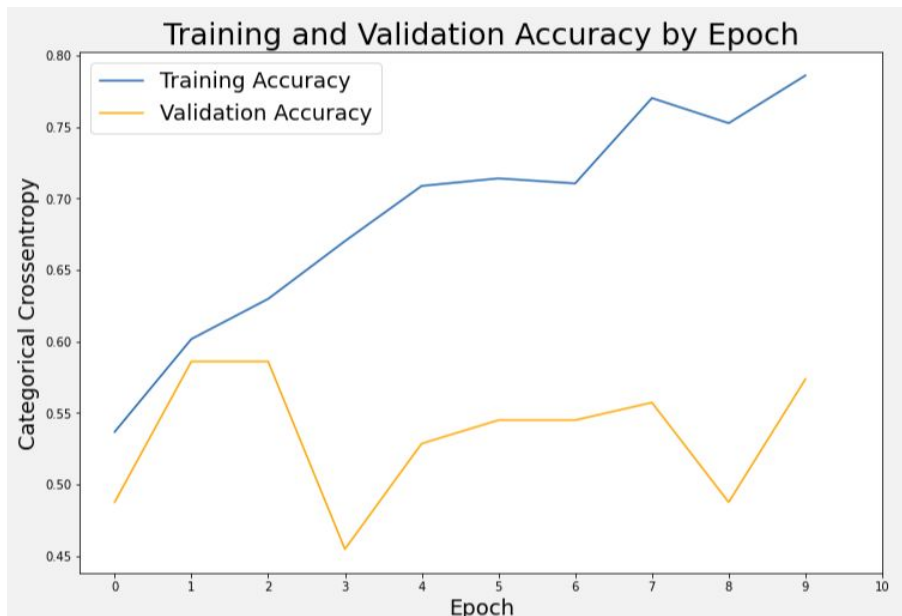
- Multi-Class Classification Neural Network
- Current Accuracy ~50%
- Epochs 10
- Validation checks how accurate the training model identify different speakers

Training Accuracy and Loss by Epoch

Training Accuracy and Loss by Epoch

- Relationships shows that as Training accuracy goes through more epochs the loss decreases
- Closer loss is to 1 means more errors occurred
- Error is minimized when the model goes through more epochs
- Epochs 10

Training Val_loss and Val_accuracy by Epoch

Training Val_loss and val_acuracy by Epoch
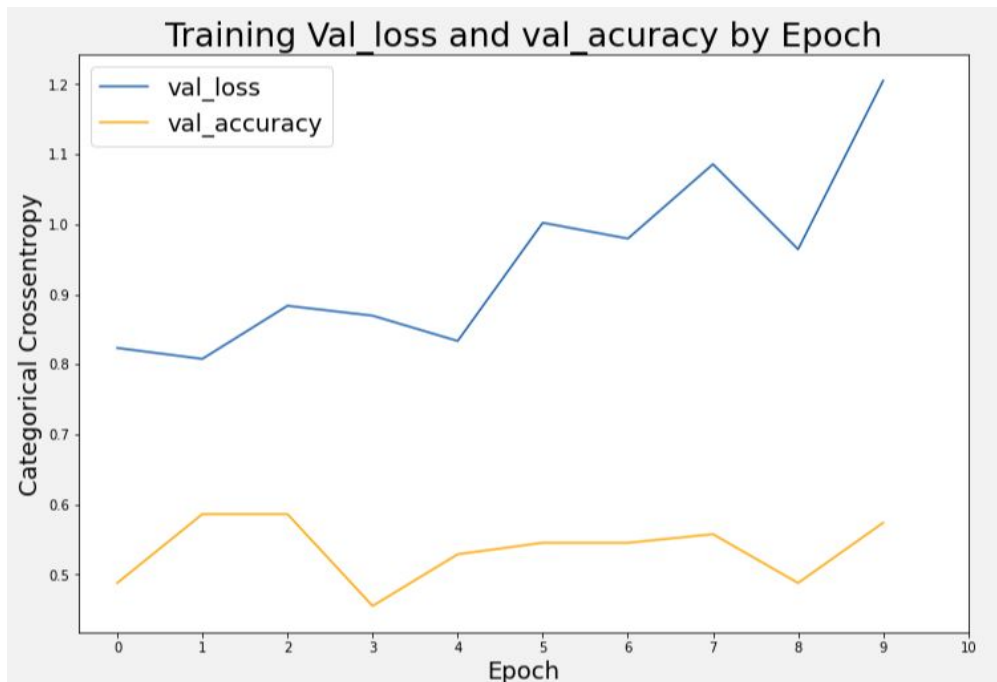
- Relationship is as we go through more epochs Val_acuracy stays relatively the same but Val_loss increase as we go through more epoch
- Graph indicates poor performance
- Epochs 10

Challenges

- Insufficient recordings from participants
- Not enough data samples
- Small pool of participants
- Getting consistent audio recording each week from participants

How this research helped me?

- Applied ML with Python (Coursera course)
- Machine Learning - Environment Setting, Learning new Algorithms, Exploring Algorithms
- Learnt planning project execution with deadlines - I was the BOSS
- Opportunity to present the research work in front of various personnel from academia and industry during LSAMP Conference in Chicago, October 2022
- Learnings from the Course Intermediate Computing helped me understand the implementation of SocioApp as an Android Application and the Course intro to Computing helped me with Python
- I was able to bring classroom learnings to the real world application

UNI / University of Northern Iowa

Thank You

Future Work

- Collect more data in different locations
- Collect Large amount of data from many people from diverse backgrounds
  - Age, Ethnicity, Male, and female
- Improve the Neural Network for identifying speakers precisely
  - Look into what vectors neural networks use the most
  - Find out what shortcuts the machine learning algorithm takes
  - Find the best epoch number that reduces error of accuracy close or to 0
- Develop models for detecting and correlating human activities conducted in proximity of others
- Create a metric to measure Sociability

Spring 2023 Data Collection Plan

(1 hour daily) Expected weeks= 4	Quite Env. Self	Loud Env. Self	Quite Env. w Peer	Loud Env. w Peer
Day 1	2/1/2022 1hr -hv	2/16 10min hv	3/22/23 35min hv 3/22/23 25min hv	3/29 30min hv
Day 2	2/2/2022 1hr -hv	3/14 30 min hv	3/24 18 min sc1 3/24 31 min sc1	3/29 30 min hv 4/4/2023
Day 3	2/3/2022 1hr -hv	3/20 20min hv	3/25 11min sc1 3/25 30 min sc2 3/25 30min sc2	4/4/ 30 min sc1 4/4/ 30 min sc1
Day 4	2/3/2022 30m -hv	3/20 25min s1	3/26 30 min sc3	4/8 1hr sc2
Day 5	2/7/22 1 hr bk s1	3/21 25min s1 3/21 15min sc1	3/27 30 min sc3 3/27 30 min sc4	4/9 1hr sc3 4/9 1hr sc4
Day 6	2/14/23 1hr bk s2	3/21 9 min sc2 3/21 36 min sc2 3/21 16min sc2	3/28 30 min sc4	

Next time

- 4 models of every scenario
- The same code run for 4 different scenarios
- Done collected data
- Ask LSMAP if they will sponsored me or if not talk to Professor about Surp

Spring 2023 Data Collection Plan

(1 hour daily) Expected weeks= 4	Quite Env. Self	Loud Env. Self	Quite Env. w Peer	Loud Env. w Peer
Day 8	2/18 1hr bk s3	3/22 34min sc3 3/22 25 min sc4		
Day 9	2/19/23 1hr bk s4	3/24 31min sc4 3/24 5min sc4		
Day 10	2/20/22 1hr bk s5			
Day 11	2/21/22 15min bk s6			
Day 13				
Day 14				

Script Table

This table shows what script and pages will be read on each day over the course of a week.

indicates the pages are subject to change
*** indicate the new pages for each script 3/20*

Days	Scripts/Pages
Day 1	Harvard Sent.
Day 2 sc1	Pg. 9-22
Day 3 sc2	Pg. 23-35
Day 4 sc3	Pg. 36-59* 36 - 51**
Day 5 sc4	Pg. 60-73* 52-64**
Day 6 sc5	Pg. 74-87* 65-75**
Day 7	Pg. 88-101* 76-87**

When I extracted from phone

Days	Dates extract from phone
Day 1	2022-02-12 hv 2022-02-13 hv
Day 2	2022-02-14 hv 2022-02-15 hv
Day 3	2022-02-18 bk sc1 2023-02-15 bk sc2 2023–02-18 bk sc3
Day 4	2023-02-19 bk bk sc4 2023-02-20 bk sc5
Day 5	2023-02-21 bk sc6
Day 7	

When I extracted from phone

Days	Dates extract from phone
Day 1	3/14 hv alone loud 3/14 hv alone loud 3/20 hv alone loud
Day 2	3/20 sc1 alone loud 3/21 sc 1 alone loud 3/21 sc1 alone loud
Day 3	3/21 x5 sc2 alone loud
Day 4	3/21 sc3 alone loud 3/22 sc3 alone loud
Day 5	3/22 x4 sc4 alone loud
Day 6	

UNI / University of Northern Iowa.

When I extracted from phone

Days	Dates extract from phone
Day 1	3/25 _PQ_sc3 3/26 _PQ_sc3
Day 2	3/26 PQ sc3 3/27 PQ sc3
Day 3	3/27 PQ sc4 3/28 x3 PQ sc4
Day 4	3/29 PL HV 3/30 PL HV
Day 5	3/30 PL sc1 4/5x2 PL sc1 pt2 3/31 PL sc1
Day 6	4/09 PL sc3, 4/09 sc4

UNI / University of Northern Iowa.

Updates I want to make to the app

- After I press extract features I want the app to upload the audio data to a google folder
- I want to incorporate my format adjuster to the app so frequency and bit rate is the same, meaning I dont have to pull the data from the phone and run it another program I created
- I want to a incorporate my 30 seconds slicer on the app itself so i can talk for as long as I want and have the app do split up the audio in 30 sec increments, I have this program wrote already