

Tabitha O'Malley

Sprint 1



SDD

Tabitha, Milan, David, Max, Tisha, Adam	09/29/2023	Start Document	V1.0
---	------------	----------------	------

SRS

Tabitha, Milan, Tisha, David, Adam, Max	09/29/23	Starting the document	V1.0
---	----------	-----------------------	------

9/21

Using both the frequency and time domain for the project

Frequency is better because it has a smaller data size and it works the same as the human ear

Challenge in speech include : length, pause, intonation, pitch, stress

Recording is not perfect - background noise and channel responses

Log operation is used to shrink the range of change in the frequency

Probability theory

Random variable (RV)

PFM for discrete RV(tossing a die) (probability math function)

PDF for continuous RV (measuring body temperature) (probability density function)

GMM (Gaussian Mixture Model) - weight of cats

Multiple RVs for a certain model

HMM (Hidden Markov model) - state driven, can not go backwards, cannot observe the state but can see RV

Condition probability the chance that something will be said based on the past words

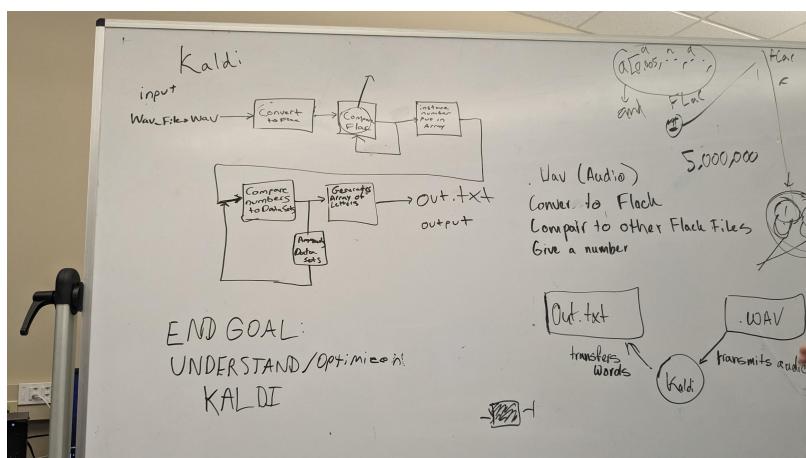
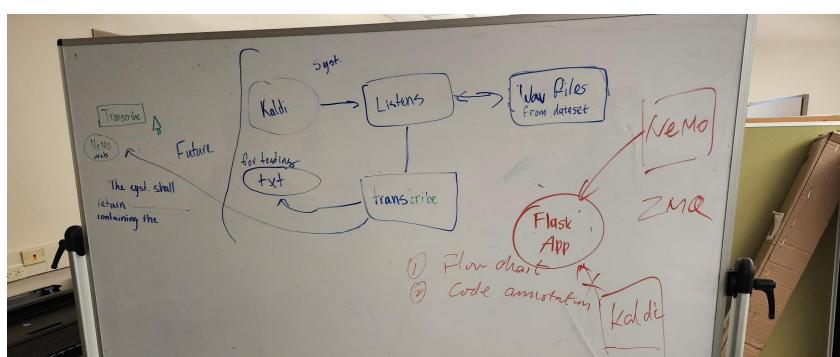
Markov model every prediction is based off only the previous word and no more

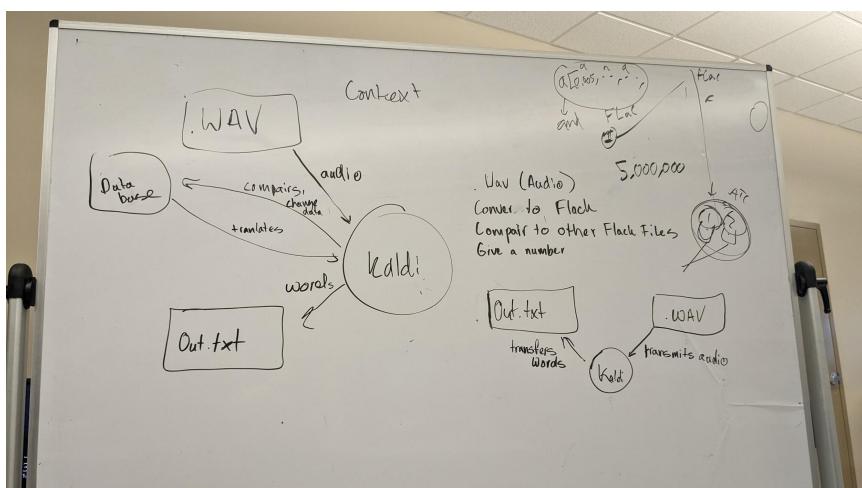
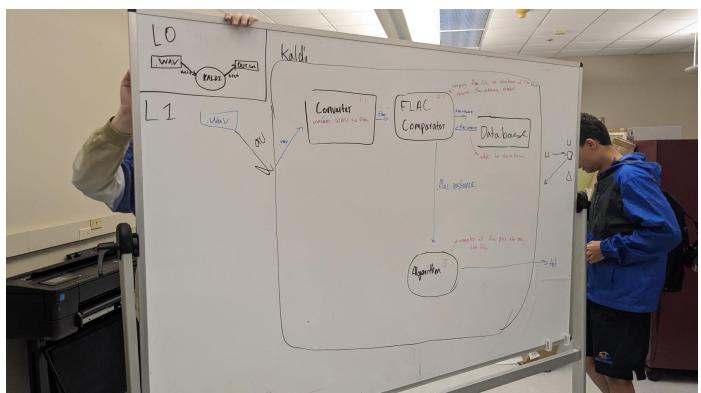
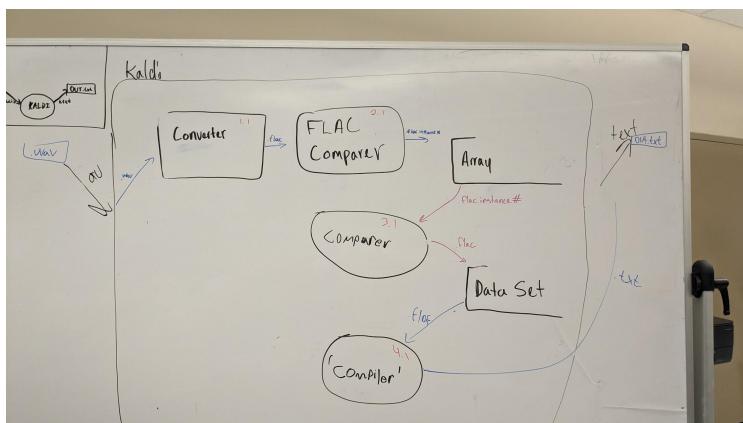
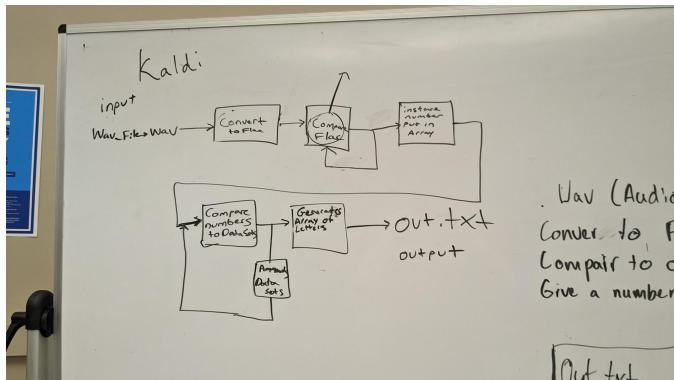
Three states for normal sound, five states for silence and stutter

Phone used to model the pronunciation of words (Lexicon Dictionary)

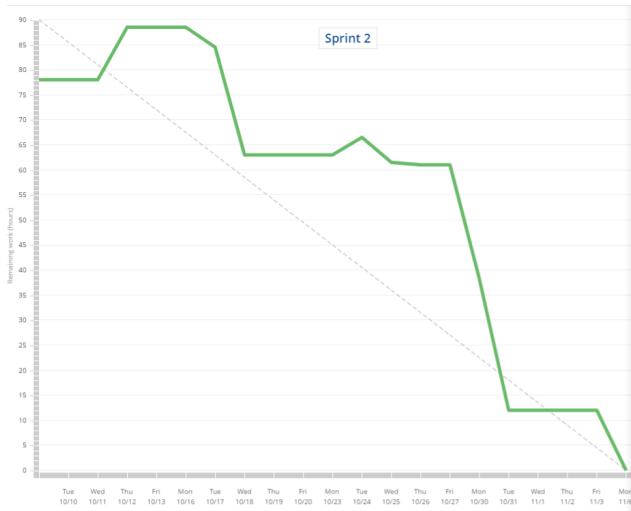
Phone and triphones cannot be observed

Decision trees significantly shrink the size of useful tri-phone





Sprint 2



SDD

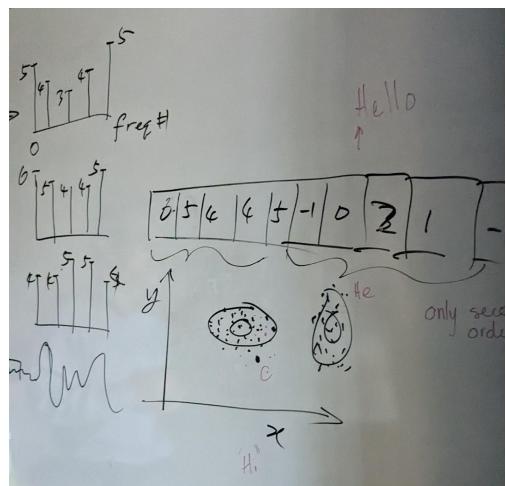
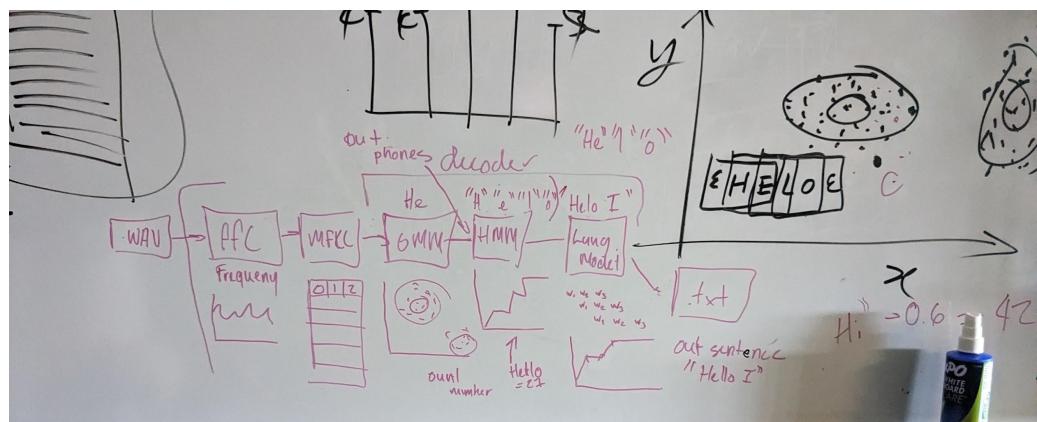
Tabitha, Milan, David, Max, Tisha, Adam	09/29/2023		V1.0
Tabitha	10/24/2023	Writing Section: 1.2.1	V2.1
Tabitha, Tisha, Milan	10/24/2023	Writing the section, Rewriting, and editing: 1.2	V2.3
Tisha, Tabitha, Milan	10/24/2023	Rewriting the section: 2.2	V2.2
Tabitha	10/25/2023	Writing Sections: 2.1, 1.5	V2.4
Tabitha	10/26/2023	Writing Sections: 1.1, 1.2.2, 1.3	V2.5
Tabitha	10/29/2023	Writing/Rewriting Sections: : 1.2.1, 2.1, 2.2, 3.1, 3.2, 4, 4.1, 4.2, 5.2	V2.10
Tabitha	10/30/2023	Rewriting Section: 2.1	V2.15
Tabitha	10/31/2023	Updating Models Editing Sections Writing Section: 4.2	V2.16

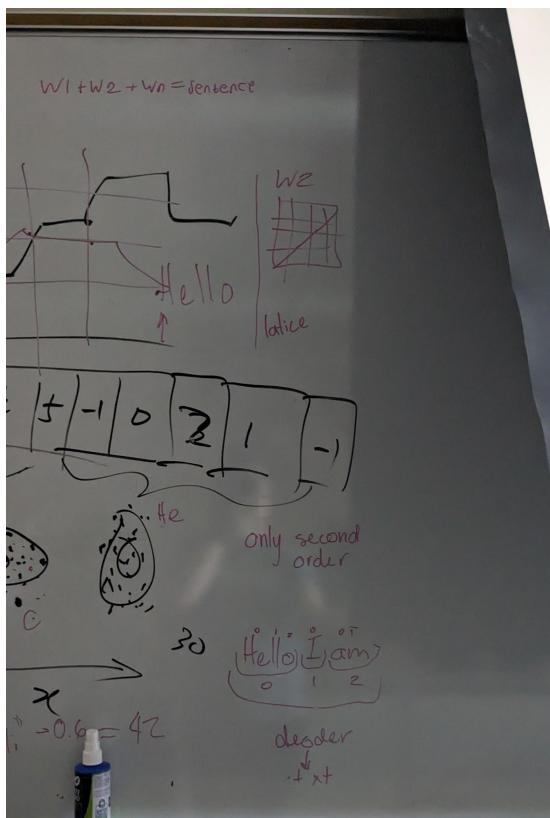
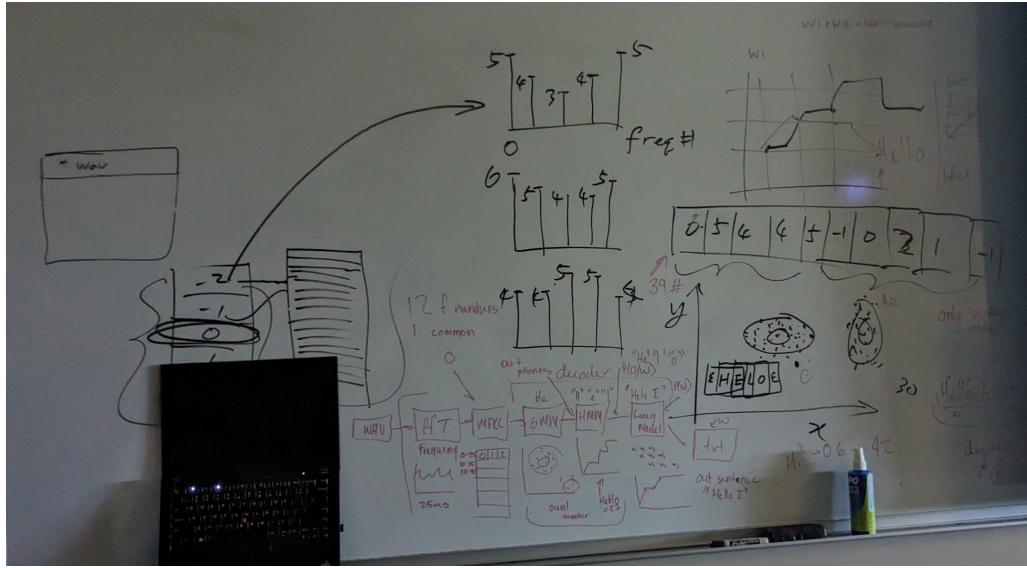
SRS

Tabitha	10/25/23	Formatting Revision History Editing/Formatting Appendix Writing Section: 1.5	V2.1
---------	----------	--	------

Tabitha	10/26/23	Writing Sections: 1.2, 2.2, 2.5, 3.1	V2.2
Tabitha	10/27/23	Writing Requirements: 3.1	V2.3
Tabitha	10/29/23	Writing Section: 2.3, 2.4, 2.5	V2.6
Tabitha	10/30/23	Writing Section: 4, 4.1, 4.2, 4.3, 4.4, 5.1, 5.2, 3	V2.8
Tabitha	10/31/23	Update Model Editing Sections	V2.11

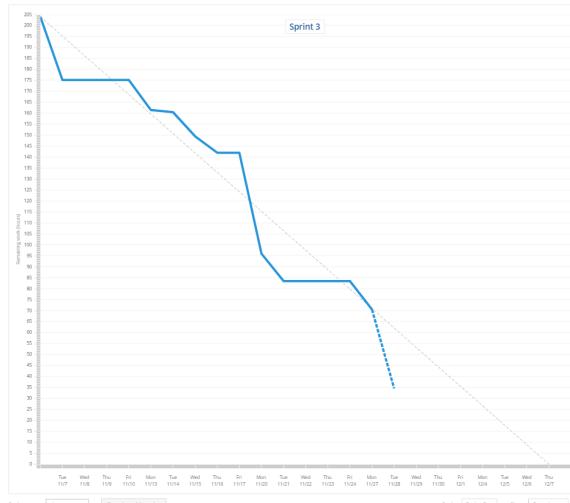
10/26





11/05

Sprint 3



SDD

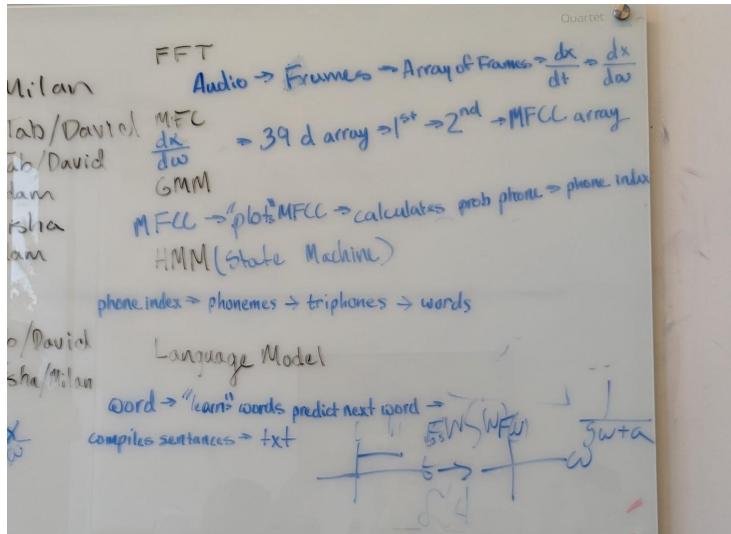
Tabitha	11/05/2023	Class Diagram: 2.1	V3.1
Tabitha	11/11/2023	Editing and Rewriting: 4.2	V3.8
Tabitha	11/15/2023	Update all DFD Models Updating/Rewriting Section; 4.2 Editing/Adding: 5.1	V3.13
Tabitha	11/16/2023	Reading and Commenting Sections : 2-5 For accuracy to the current requirements Update Classes Diagram Update/Rewrite Section: 2.1	V3.15
Tabitha	11/18/2023	Editing and rewriting Use Cases Remaking Use Case Diagram V2.1.1 Making Use Case Diagram V2.2.2 Editing DFD V3.1.3, V3.2.2, V3.3.1 Rewriting and editing 4.1	V3.17
Tabitha	11/19/2023	Editing: 5.1, 5.2	V3.21

SRS

Tabitha	11/05/23	Rewrite: 2.3 Add to: 2.2	V3.3
Tabitha	11/07/23	Adding/Editing: 3	V3.6
Tabitha	11/11/2023	Editing: 3	V3.10
Tabitha	11/14/2023	Update Requirements	V3.13

Tabitha	11/15/2023	Editing: 1.5, 2.3, 4.1, 5.1	V3.16
Tabitha	11/16/2023	Review/Edition/Commenting Section: 4 Update Class Diagram Update/Review Section: 2.3	V3.17
Tabitha	11/19/2023	Adding/Editing/Rewriting: 4, 5.1, 5.2 Appendix A, B	V3.22

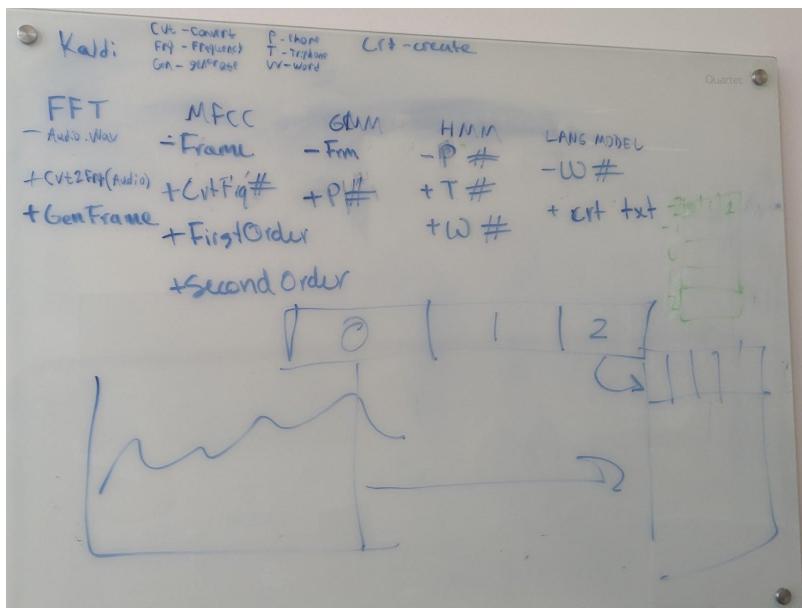
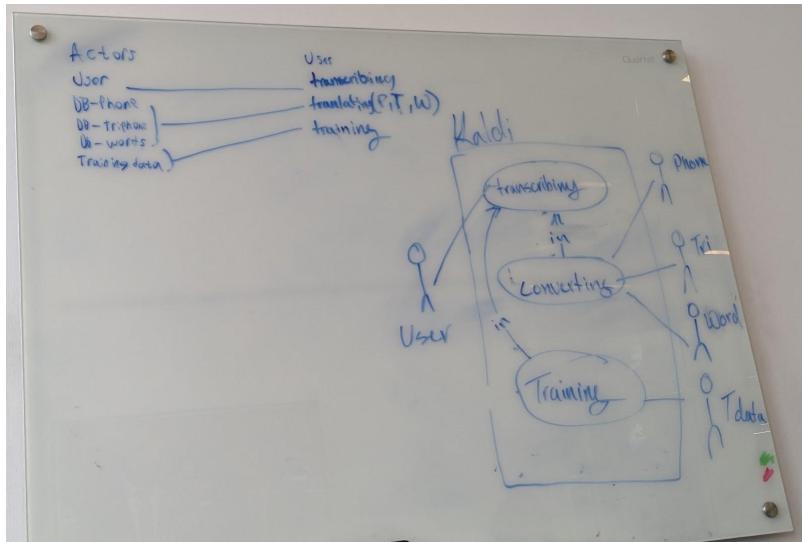
Name	Date	Reason For Changes	Version
All	10/18/2023	Write wrong information	V1.0
Tabitha	11/20/2023	Added in comments	V2.1
Tabitha	11/27/2023	Sections 3.3, 3.3.1, 3.3.2, 7, 8	V2.3
Tabitha	11/28/2023	Section 1.1 Editing/Reading All Sections Editing 3.3.2, 4, 5	V2.7



Clause 0 ???

if iteration n has a worse (i.e., lower) transcription accuracy than iteration n-1, iteration n shall cease to exist.

if iteration n has a greater or equal transcription accuracy than iteration n-1, the system shall overwrite iteration n-1 with iteration n



MFC Converter

- frame[]
- MFCC [] \Rightarrow vector
- + createMFCArray (frame[], firstder[], second
firstder (frame[]))
- + Frame[] firstder (frame[])

FFT

- Audio audiofile
 - Frame frame
 - Frame[] frames
- + Frame cvt2frm(Audio)
- + Frame[] framearr(Frame)
- + Frame CVt2freq(Frame[])

GMM

- MFC [...] ... [...] vector
 - int MFC state
 - int HMM state
 - double distance
- + find MFC state (MFC[]) vector
- + calculate distance between MFC and HMM states

HMM

- int MFC state
 - int HMM state
- + int compare MFC state to phoneme states
(MFC state, phonemes)
- int phonemes
- + int get phonemes (File read)

WFST

- int [] phoneme chain
 - string sentence
 - int[3] triphones
 - int monophone
 - int word
- + int [] HMM (phoneme chain)
- + int [] get triphone (phoneme chain, File read phones)
- + int convert from tri to mono (triphone)
- int [] Monophones
- + int[] create array of mono phones (monophones)
- + int convert from mono to word (Monophones)
- + string compile sentence (word)

Viterbi Decoder

- double distance
 - int[] phoneme chain
- + int[] calculate phoneme chain (distance)

