



AUTOMATING AUDIO LABELING WORKFLOW USING DEEP LEARNING FOR VOICE ACTIVITY DETECTION

RAMAKRISHNAN RAMAN

FELLOW

HONEYWELL TECHNOLOGY SOLUTIONS

VASANTHA SELVI PAULRAJ

LEAD EMBEDDED ENGINEER

HONEYWELL TECHNOLOGY SOLUTIONS

AUTOMATING AUDIO LABELING WORKFLOW USING DEEP LEARNING FOR VOICE ACTIVITY DETECTION

- This presentation discuss our experience in using a pre-trained Deep Learning model in a labeling algorithm and demonstrates how we have adopted automation of Audio Labeling workflow towards audio Voice Activity Detection(VAD) using MATLAB.

HONEYWELL AEROSPACE - [HTTPS://AEROSPACE.HONEYWELL.COM/](https://aerospace.honeywell.com/)

Honeywell

AEROSPACE

Learn

Purchase

Services & Support

MyAerospace

Shaping the future of Aerospace: Get inspired by our latest customer success stories



Honeywell Partnership Supports CIAC's Market Expansion

Engineers and end customers benefit from Honeywell Authorized Service Center status. Find out more!

[MATLAB EXPO 2022](#)



Honeywell Inertial Navigation System Delivers Precise Location Accuracy

Honeywell HGuide n580 INS/GNSS navigator delivers precise navigation information for UAV hardware and service provider. Read more!



Lockheed Martin's Mars Lander Counts on Honeywell MIMU to Study Red Planet

Honeywell's Miniature Inertial Measurement Unit (MIMU) has been a part of Lockheed Martin's Mars satellites and landers since 1998. Read more!

Honeywell Aerospace products and services are found on virtually every commercial, defense and space aircraft in the world. With an unmatched heritage of innovation that spans more than a century, our aim is to solve the greatest challenges CEOs, pilots, operators, passengers, finance, maintenance and cabin crews face -- and transform the way we all fly. We deliver improved fuel-efficiency, more direct and on-time flights, safer and more comfortable travel and better flight planning and traffic management. We do this through one of the industry's broadest and most advanced portfolios including world-class engines, cockpits, cabin design, wireless connectivity and enterprise performance management services -- even for emerging market segments like flying taxis, while providing world-class customer service, repairs and technical support.

HONEYWELL: SPEECH TECHNOLOGIES



Speech TECHNOLOGY

Speech Recognition Customer Self Service Virtual Assistants Analytics Artificial Intelligence More Topics

HOME SUBSCRIBE NEWS IN DEPTH WHITE PAPERS WEBINARS RESOURCES CONFERENCES ABOUT

Honeywell Tests Voice Recognition in Plane Cockpits



Honeywell Aerospace has begun testing voice recognition inside aircraft cockpits.

The company is currently using its Honeywell Innovative Prototyping Environment (HIPE) to flight-test voice recognition in an Embraer ERJ170 aircraft and is working with pilots and customers to assess its usability, safety, and efficiency in real airborne scenarios.

MATLAB EXPO 2022

Speech
Mag

Method and system for remotely training and commanding the speech recognition system on a cockpit via a carry-on device in a connected aircraft

Patent number: 10672385

Date of Patent: June 2, 2020

Assignee: Honeywell International Inc.

Inventors: Bharath Sundararajan, Vasantha Selvi Paulraj, Rengarajan Mohan

Audio processing system and method using push to talk (PTT) audio attributes

Patent number: 10278034

Date of Patent: April 30, 2019

Assignee: HONEYWELL INTERNATIONAL INC.

Inventors: Sivaraman Saptharishi, Vasantha Selvi Paulraj

Methods and apparatus for post-processing speech recognition results of received radio voice messages onboard an aircraft

Patent number: 10157616

Date of Patent: December 18, 2018

Assignee: HONEYWELL INTERNATIONAL INC.

Inventors: Jitender Kumar Agarwal, Vasantha Selvi Paulraj, Kiran Gopala Krishna, Chaya Garg, Roger W. Burgin, Robert E De Mers

Sample
Patents

PROBLEM CONTEXT

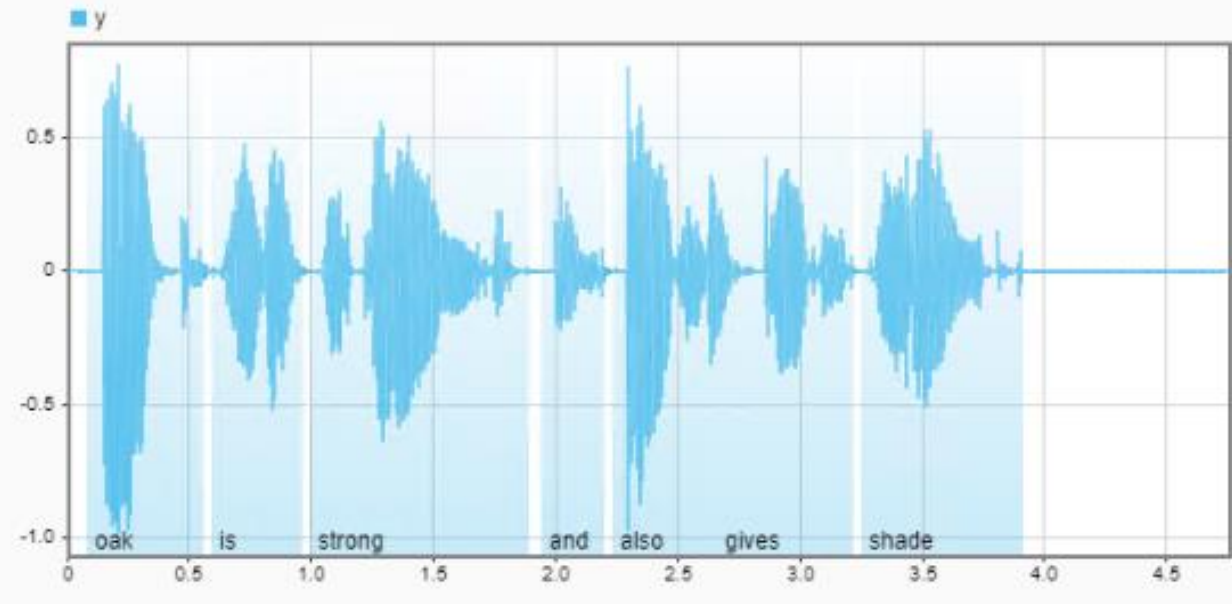
- Speech technologies for Aviation domain has a unique challenge as it involves heavy domain specific vocabulary.
- There is need for large amount of domain specific labeled corpus for each accent and language
- Larger vocabulary use cases require huge training data as well to be labeled

CHALLENGES IN LABELING OF AUDIO DATA

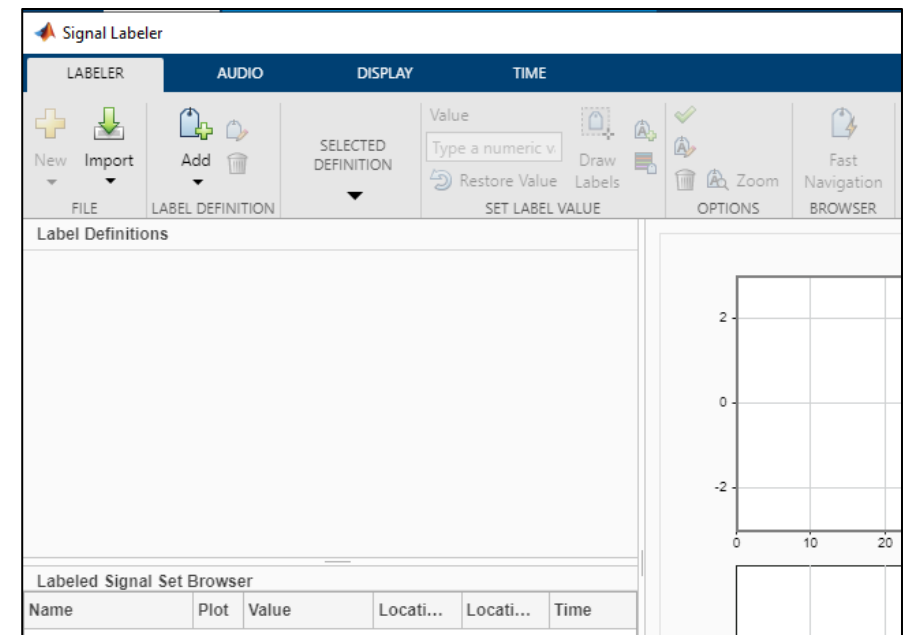
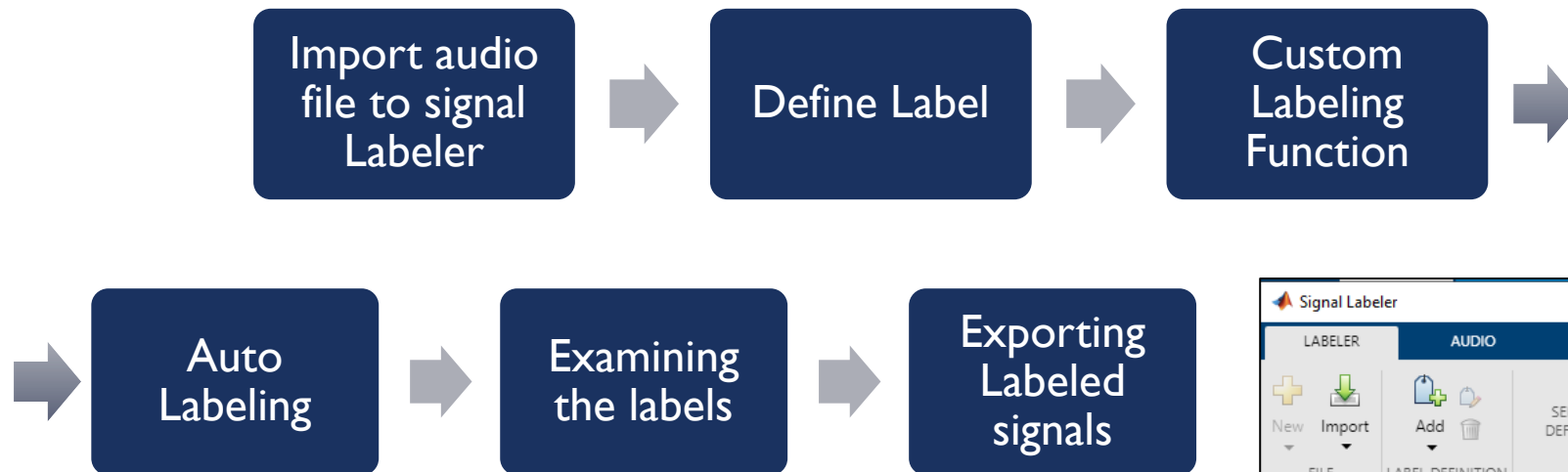
- For audio files, labeling involves analyzing segments of audio and listening to them and manually assigning appropriate labels for specific time slots in the audio files.
- However, such an Audio Labeling workflow is a labor-intensive process and slows the development cycle time
- For instance, if we are planning to train a Deep Learning Model to classify segments into speech or noise (for VAD - Voice Activity Detection purposes), we need to label the speech and noise segments in the audio files.
- Manual labeling of audio data is a resource intensive and time-consuming process

LABELING OF DATA

- Deep Learning models require labelled data for training purposes. Labeling of the data is a crucial and important step.
- To overcome the challenges, feasibility of automated labeling of the data was explored, using MATLAB Signal Labeler



LABELING WORKFLOW



STEP 1: LOADING AUDIO FILE & IMPORT TO SIGNAL LABELER

- Open Signal Labeler. On the Labeler tab, click Import and select file (use 'select from workspace' in the Members list, if loaded prior).
- In the dialog box, select the signal, *y*.
- Add time information: Select Time from the drop-down list and specify fs as the sample rate, which is measured in Hz.
- Close the dialog box ("Import & close"). The signal appears in the Labeled Signal Set Browser.

The screenshot shows the Signal Labeler interface. The 'Import Members from Workspace' dialog box is open, displaying a 'Workspace Browser' table with one entry: *y* (9600000x1 double). Below the table, the 'Working in:' dropdown is set to 'Time', and the 'Sample Rate' is set to 8000 Hz. The 'Labeled Signal Set Browser' at the bottom shows the signal *y* with a sample rate of 8 kHz.

NAME	SIZE	CLASS
<i>y</i>	9600000x1	double

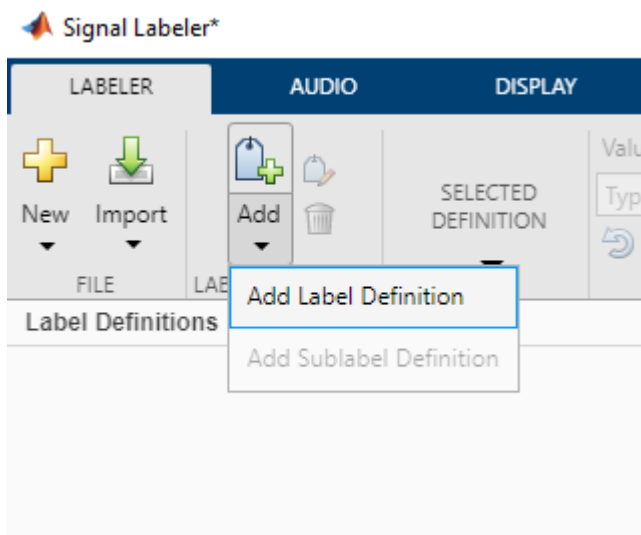
Name	Plot	Value	Locatio...	Lo...	Time
<i>y</i>	<input type="checkbox"/>				Fs: 8 k...

To load audio file in workspace:

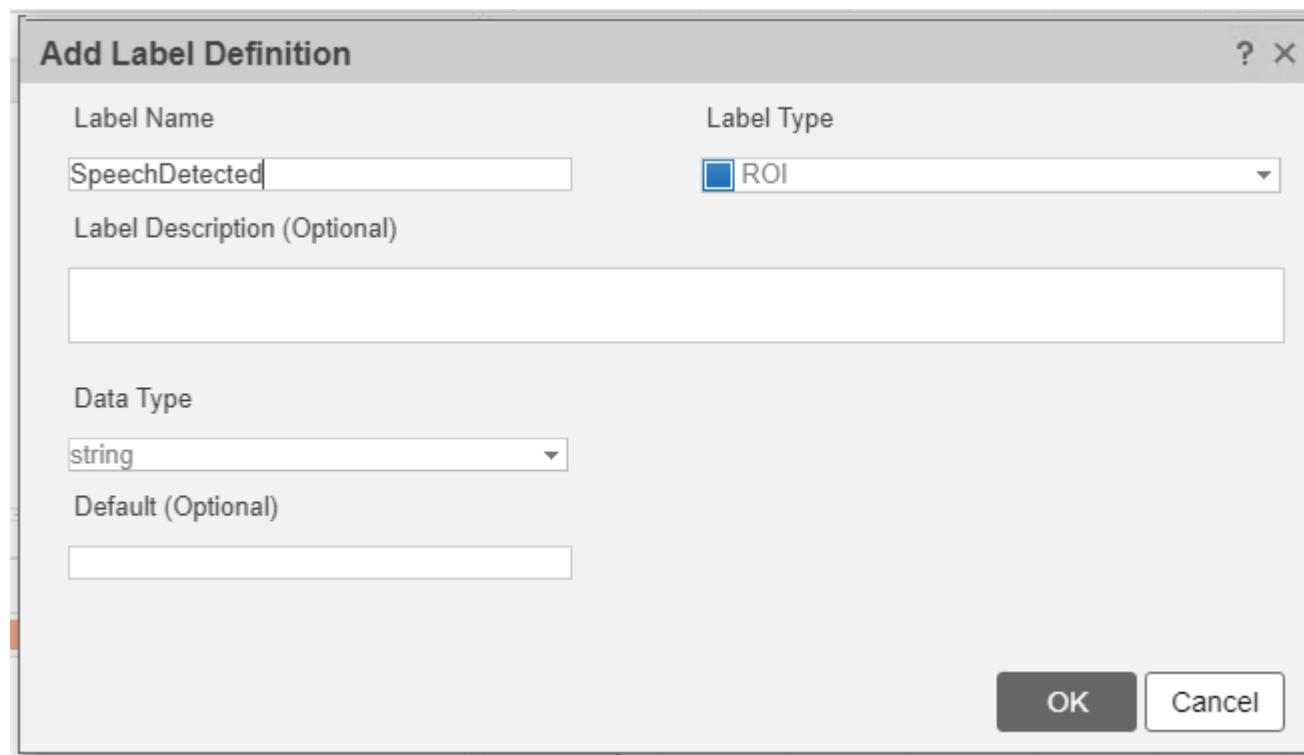
- `[y,fs] = audioread('myAudioFile.wav');`
- % To hear, type `soundsc(y,fs)`

Name	Value
<i>fs</i>	8000
<i>y</i>	9600000x1 double

STEP 2: DEFINE LABEL



- Click Add Definition on the Labeler tab.
- Specify the Label Name as “SpeechDetected”, select a Label Type of ROI, and enter the Data Type as string.



The 'Add Label Definition' dialog box is shown. It contains the following fields and options:

- Label Name:** A text input field containing 'SpeechDetected'.
- Label Type:** A dropdown menu with 'ROI' selected.
- Label Description (Optional):** An empty text input field.
- Data Type:** A dropdown menu with 'string' selected.
- Default (Optional):** An empty text input field.

At the bottom right of the dialog box, there are two buttons: 'OK' and 'Cancel'.

STEP 3: CUSTOM LABELING FUNCTION

The screenshot shows the Signal Labeler software interface. The main window is titled "Signal Labeler*" and has a menu bar with "LABELER", "AUDIO", "DISPLAY", and "TIME". Below the menu bar is a toolbar with various icons for file operations, label definitions, and analysis. The "DRAW LABELS" button is highlighted. The main area is divided into two panes. The left pane, titled "Label Definitions", shows a list of labels with "SpeechDetected" selected. The right pane, titled "Labeled Signal Set Browser", shows a table of labeled signals.

Name	Plot	Value	Locatio...	Lo...	Time
y	<input checked="" type="checkbox"/>				Fs: 8 k...

The waveform plot shows a signal labeled "y" over a time range of 0 to 18 minutes. The signal is a complex, periodic waveform with a peak amplitude of 1 and a trough amplitude of -1. The plot is titled "SpeechDetected" and has a y-axis labeled "y" and an x-axis labeled "Time (minutes)".

The screenshot shows the "ROI FUNCTIONS" and "POINT FUNCTIONS" panels in the Signal Labeler software. The "ROI FUNCTIONS" panel has a search bar and a list of functions, including "LSTM_sp..." and "LSTM_speech_detn". The "POINT FUNCTIONS" panel has a list of functions, including "Peak Labeler".

ROI FUNCTIONS

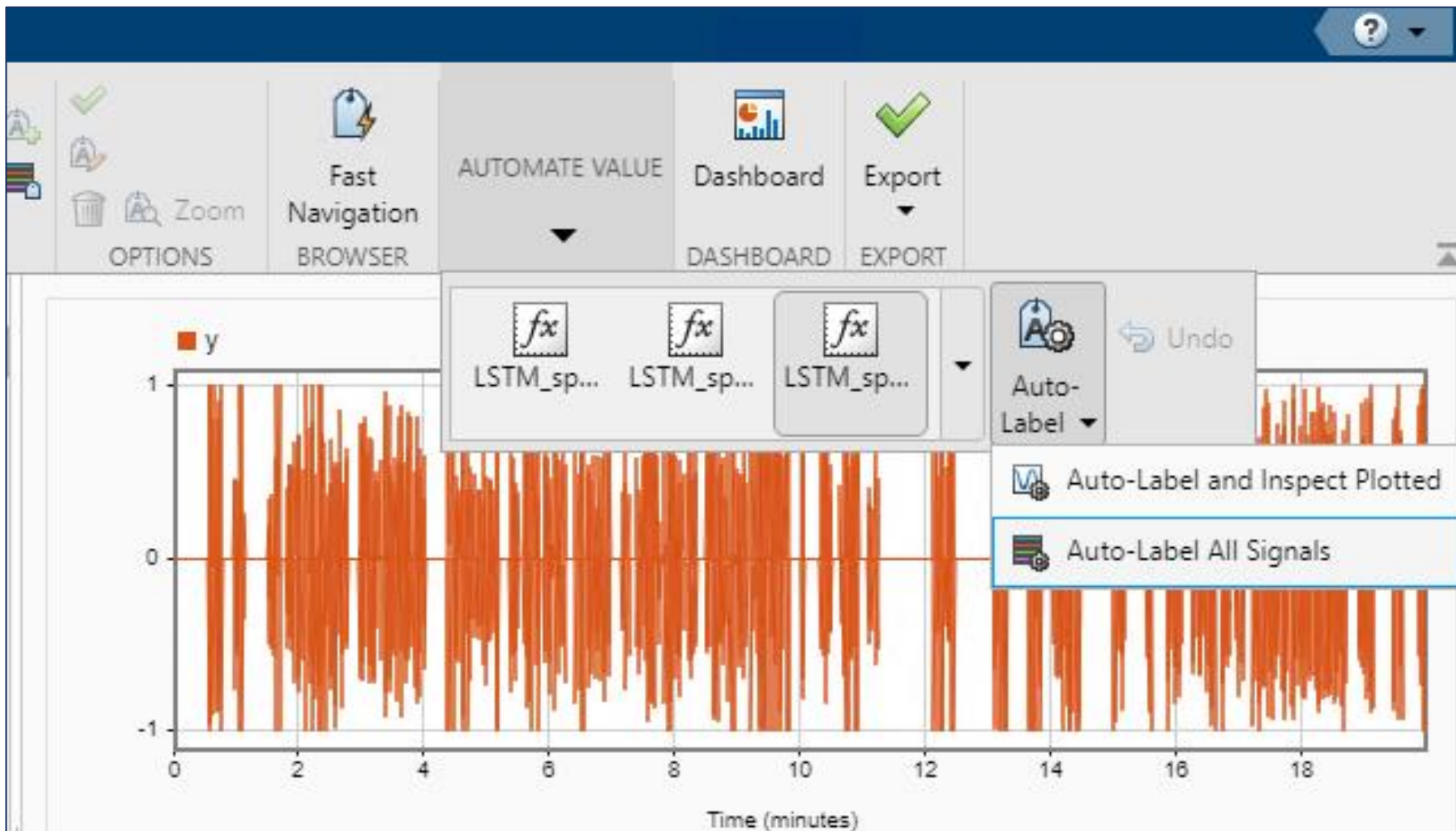
- LSTM_sp...
- LSTM_sp...
- LSTM_sp...

POINT FUNCTIONS

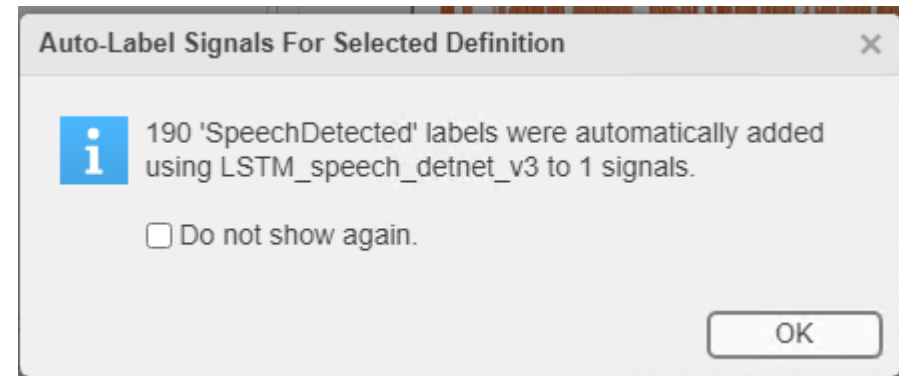
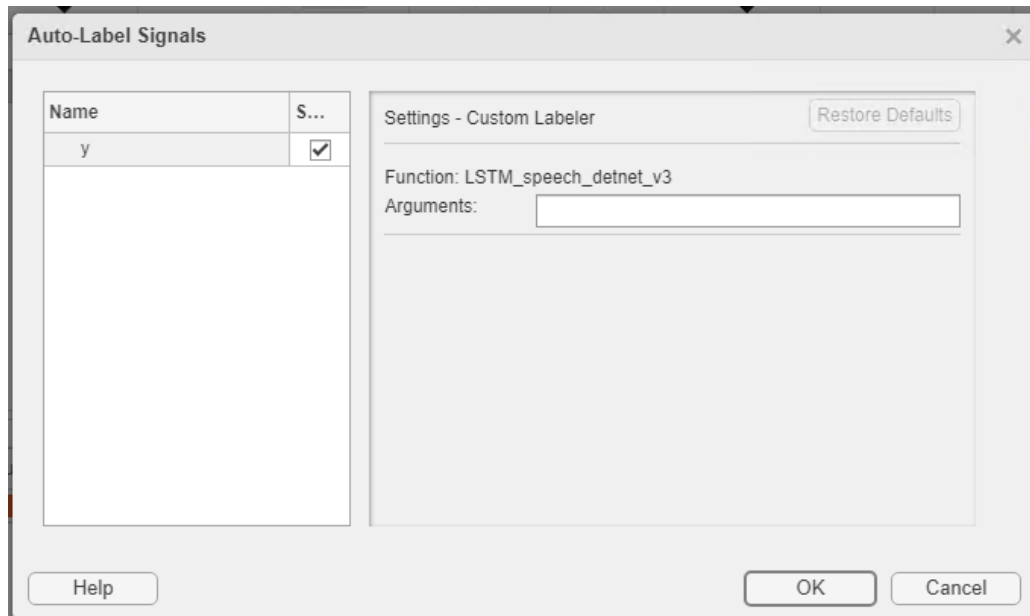
- Peak Labeler

Buttons: Add Custom Function, Manage Custom Functions

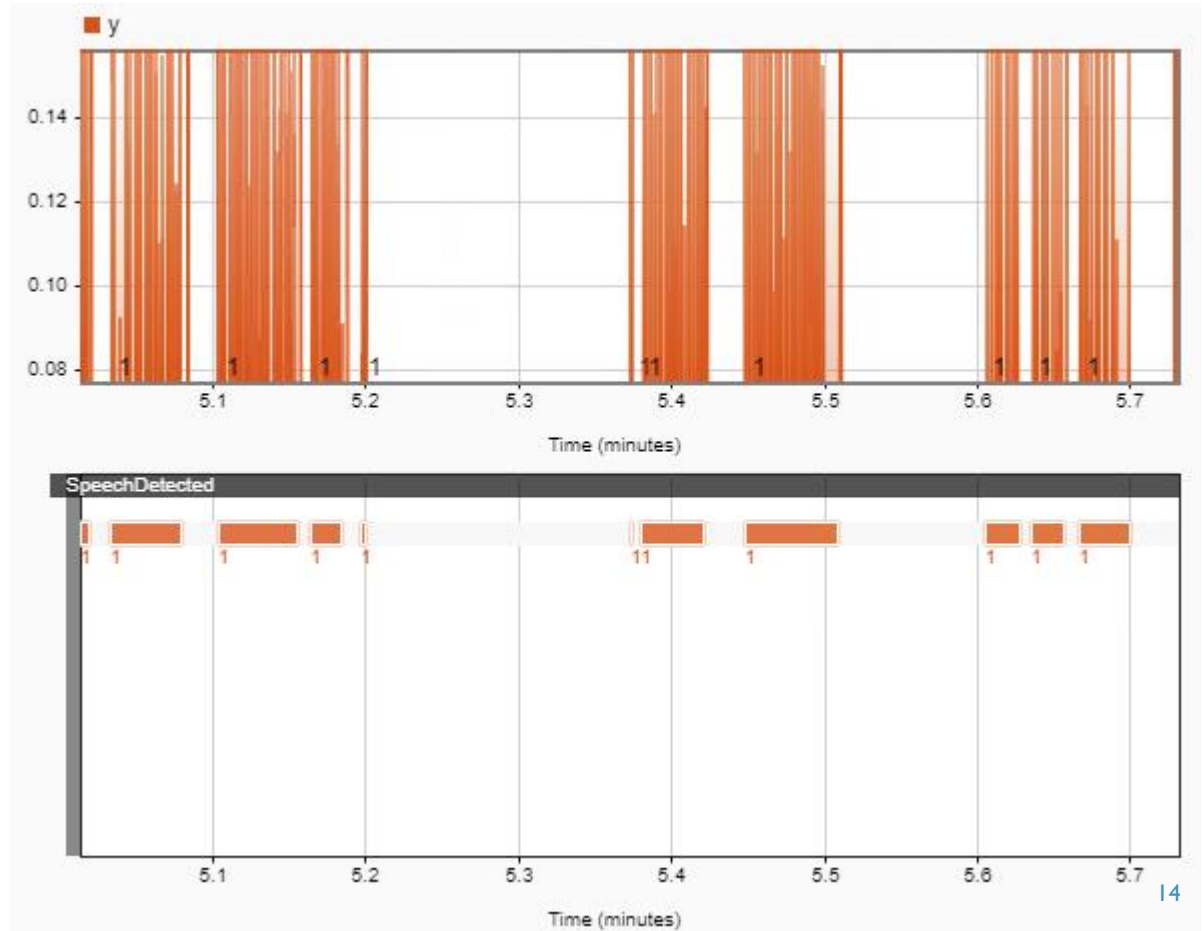
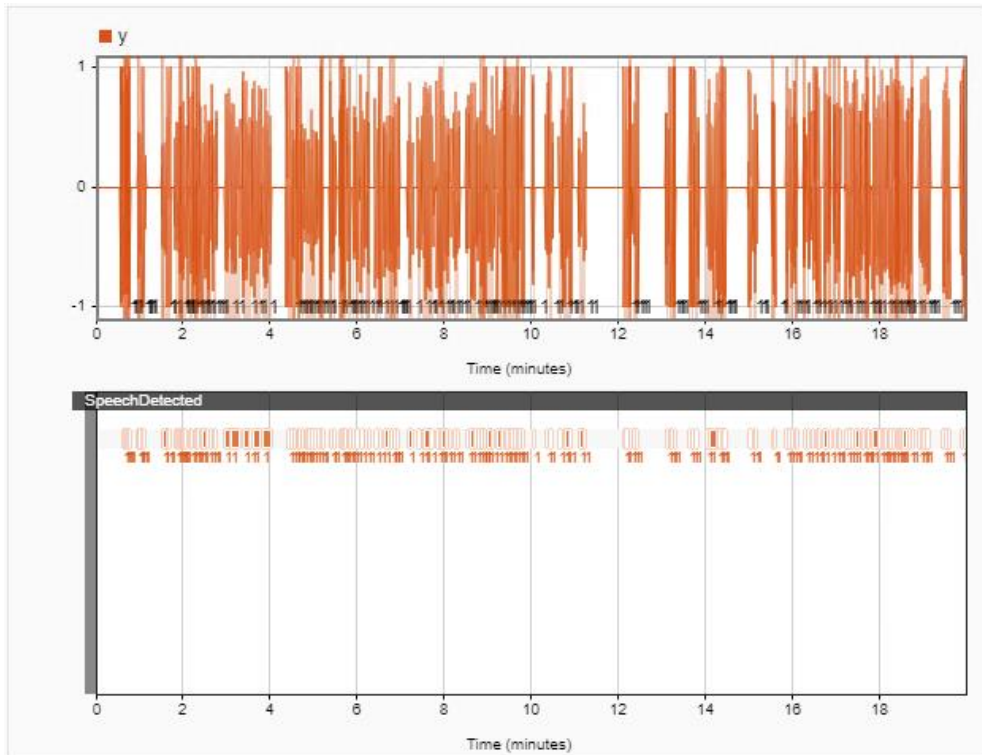
STEP 4A: AUTO LABELING



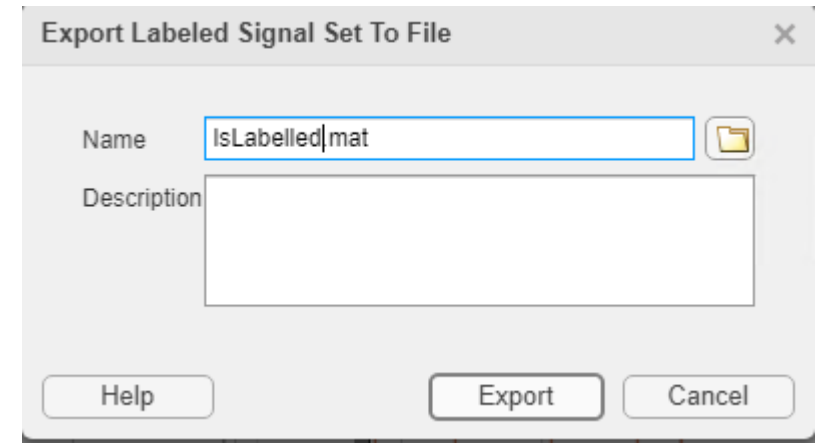
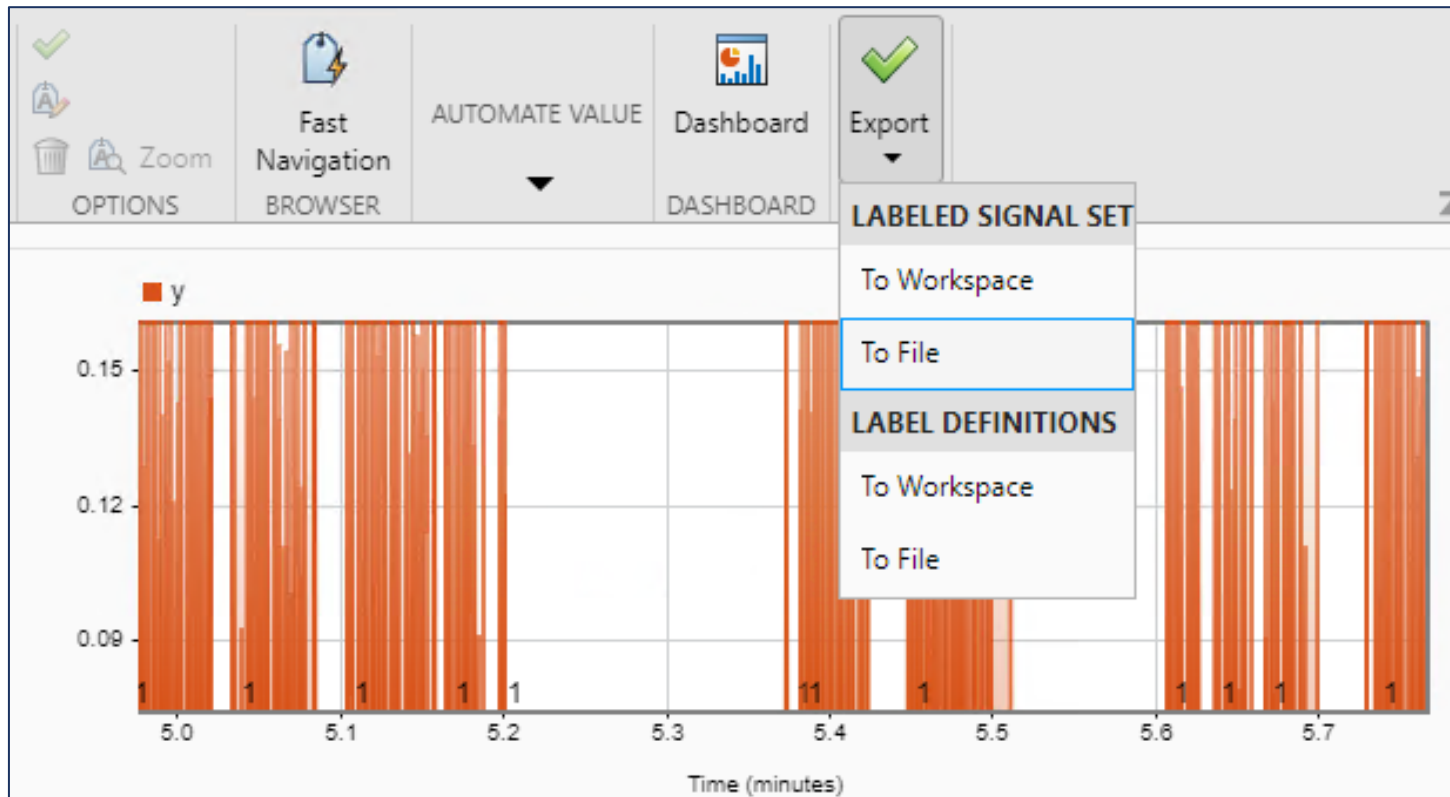
STEP 4B: LABELING COMPLETE



STEP 5: EXAMINING THE LABELING



STEP 6: EXPORTING LABELED SIGNAL SET



LABEL SET GENERATED

1x1 labeledSignalSet	
Property ▲	Value
SampleRate	
SampleTime	
TimeValues	
TimeInformation	"inherent"
Description	""
NumMembers	1
Source	1x1 cell
Labels	1x1 table

Is.Labels.SpeechDetected{1,1}		
	1 ROILimits	2 Value
	36.1760	38.6880 "1"
	39.3760	40.2560 "1"
	40.5280	41.9520 "1"
	42.6720	44.7520 "1"
	45.0240	46.6880 "1"
	57.6480	58.7360 "1"
	58.8320	59.3440 "1"
	59.5040	62.8160 "1"
	63.5040	65.4720 "1"
	65.7760	67.9520 "1"
	91.5520	92.6400 "1"
	92.7360	98.0160 "1"
	98.7520	101.0400 "1"
	101.3280	102.1600 "1"
	108.7040	111.2000 "1"
	111.7920	112.7840 "1"
	113.4880	115.2000 "1"

CONCLUSIONS

- Speech recognition systems requires large amount of domain specific labeled corpus
- Automation of Labeling can provide immense benefits in significantly reducing the cycle time
- This presentation illustrated automation of Audio Labeling workflow towards audio Voice Activity Detection(VAD)

ACKNOWLEDGEMENT

- Presenters gratefully acknowledge the help & support received from Jayanth Balaji A., Application Engineer, MathWorks



THANK YOU