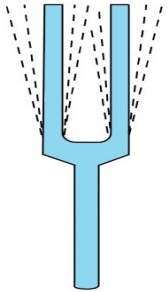
Introduction to Speech Command Identification

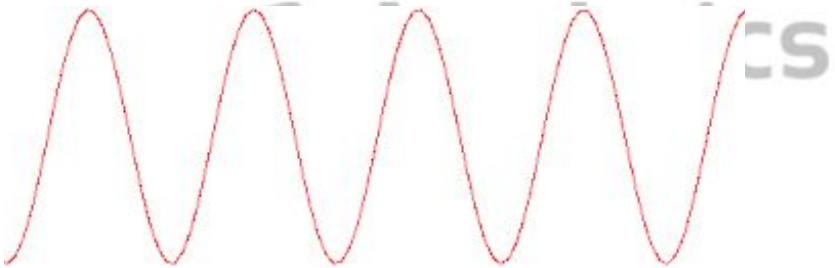
Analytics Vidhya

Recap: What is an Audio Signal?

Recap: What is an Audio Signal?

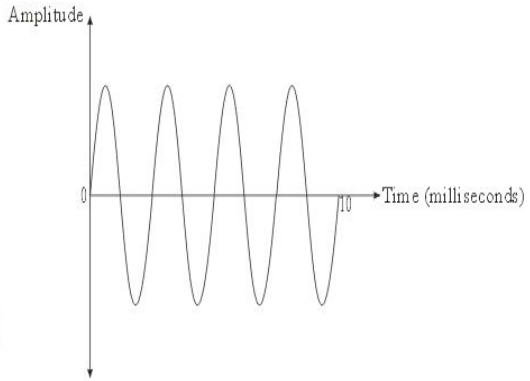
- Any vibrating body produces a sound wave





Recap: Audio Representation

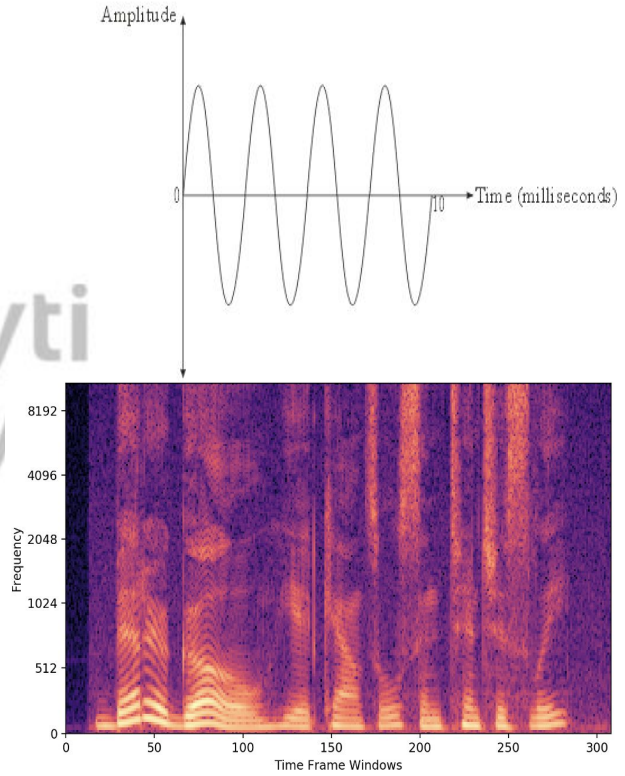
Recap: Audio Representation



- **Time Domain:** Amplitude vs Time

Recap: Audio Representation

- **Time Domain:** Amplitude vs Time
- **Spectrogram:** Every point represents an amplitude of a frequency at particular time



Recap: Data Science Problem - Audio Classification

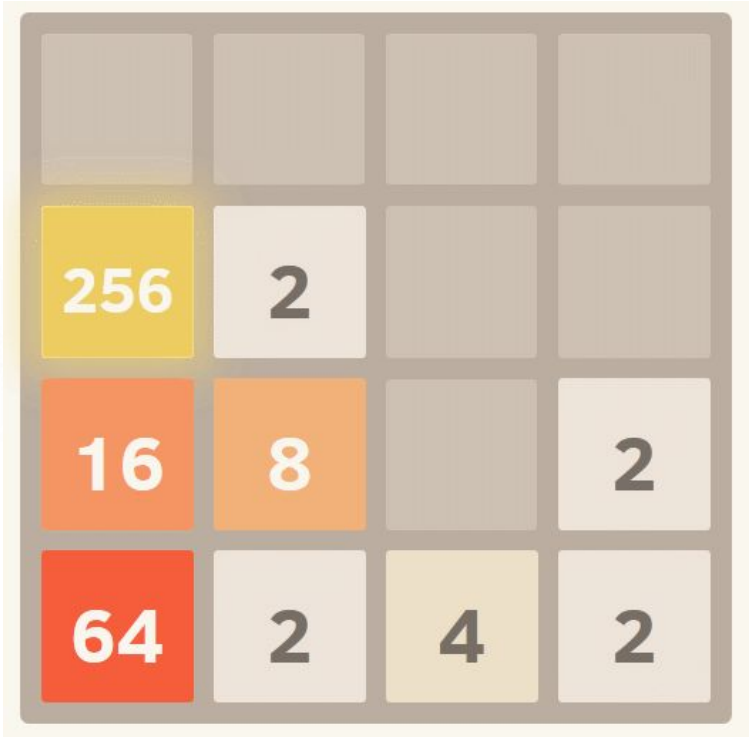
Objective: To build a model that classifies the audio into emergency or non emergency

Emergency Vehicle

Non Emergency Vehicle

Overview of Problem Statement

- *Play the game “2048” using speech commands*

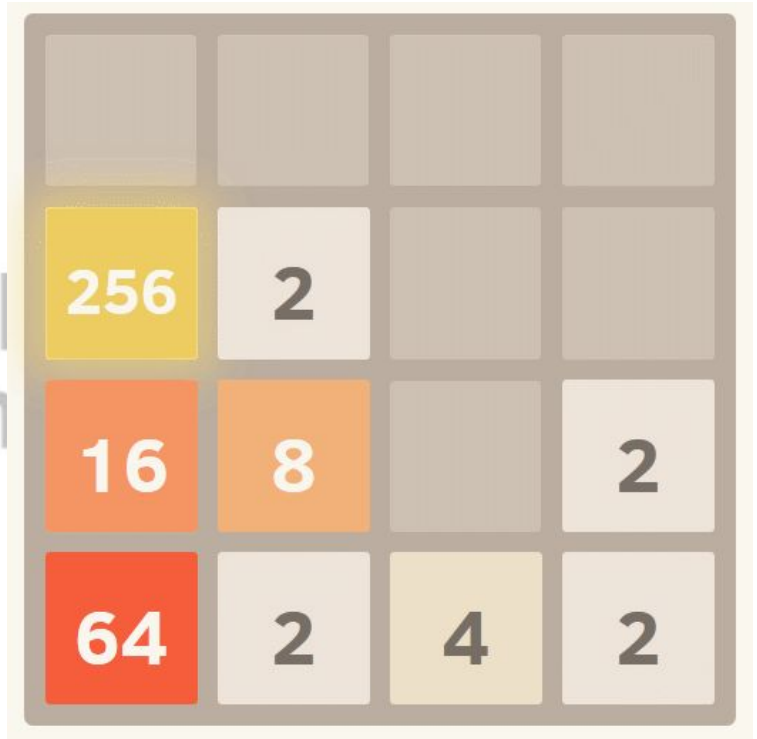


Analytics Vidhya

How to play the game “2048”

How to play the game “2048”

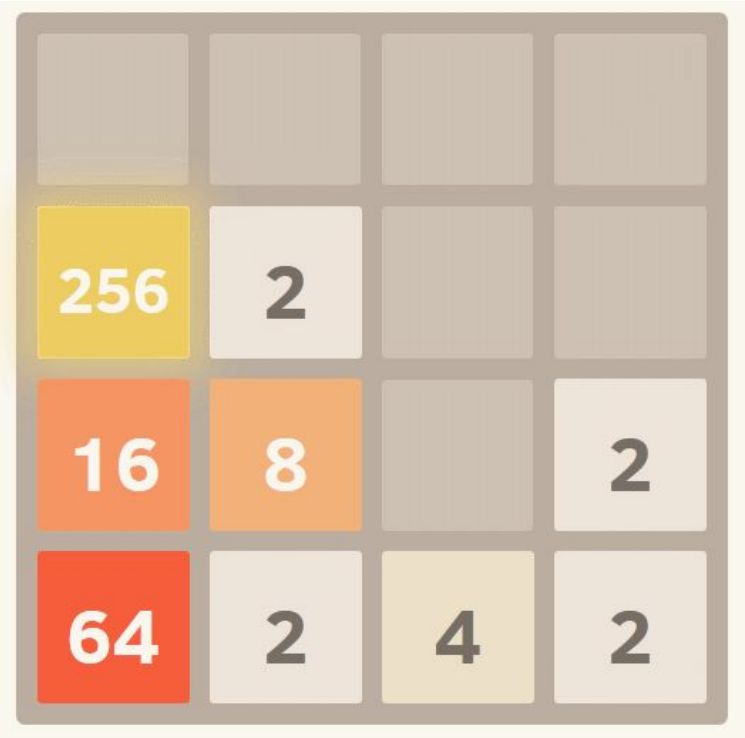
- Four ways to move the numbers
  - Up
  - Down
  - Left
  - Right



How to play the game “2048”

- Four ways to move the numbers
  - Up
  - Down
  - Left
  - Right

- If two tiles of the same number touch - they merge

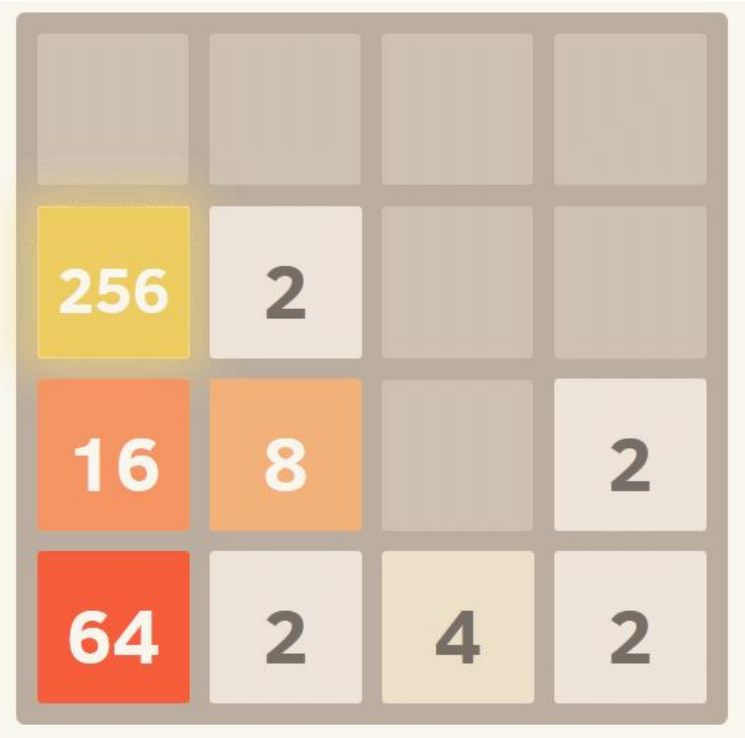


How to play the game “2048”

- Four ways to move the numbers
  - Up
  - Down
  - Left
  - Right

- If two tiles of the same number touch - they merge

- At each move, one tile randomly occurs in an empty position



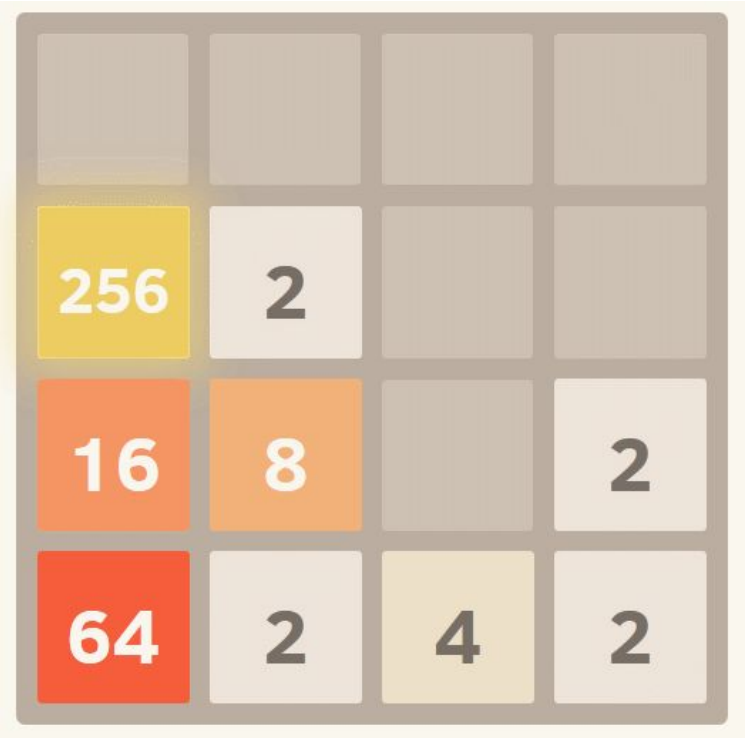
How to play the game “2048”

- Four ways to move the numbers
  - Up
  - Down
  - Left
  - Right

- If two tiles of the same number touch - they merge

- At each move, one tile randomly occurs in an empty position

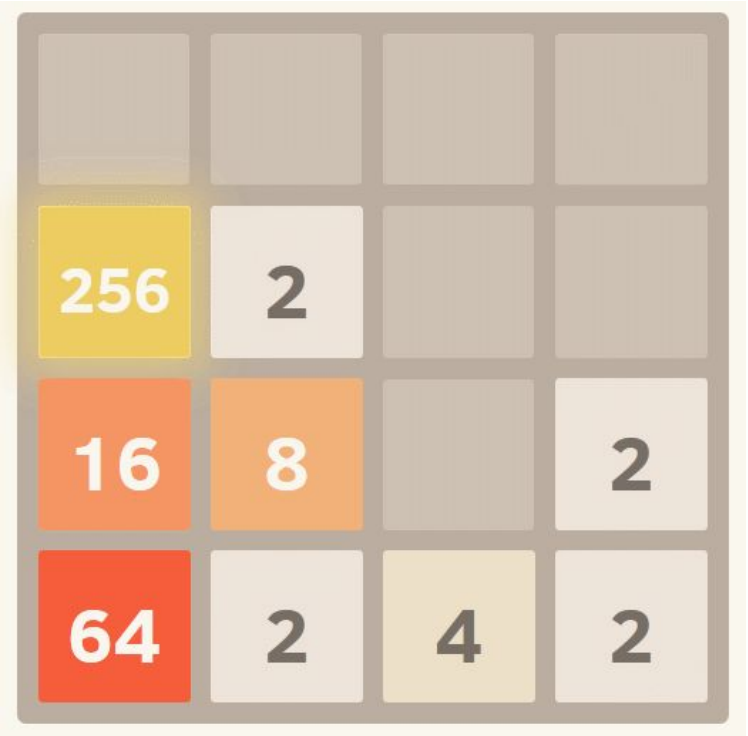
- The game ends when there’s no move available



Overview of Problem Statement

- *Play the game “2048” using speech commands*

Step I - Train a Deep Learning model to identify speech commands



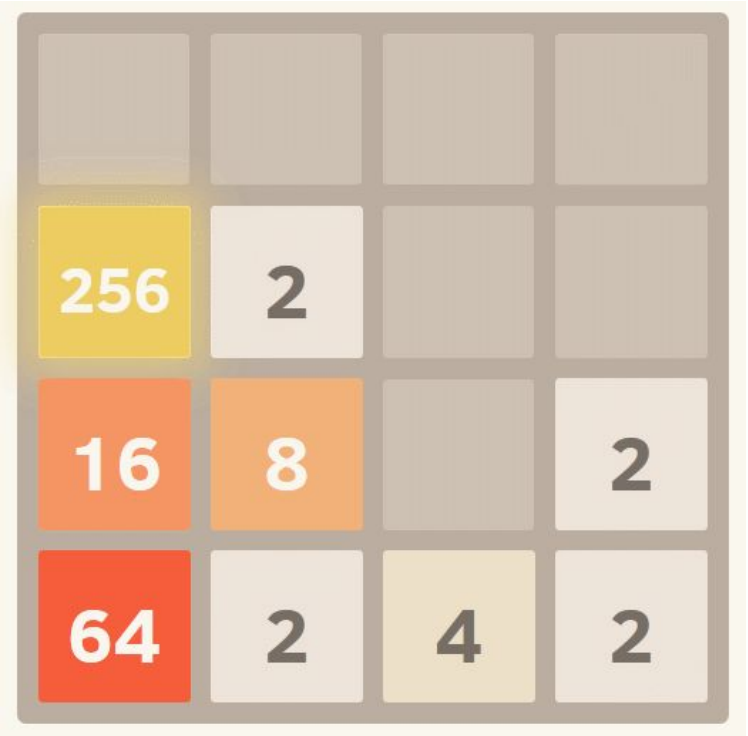
Analytics Vidhya

Overview of Problem Statement

- *Play the game “2048” using speech commands*

Step I - Train a Deep Learning model to identify speech commands

Step II - Identify the speech commands in real time to play the game

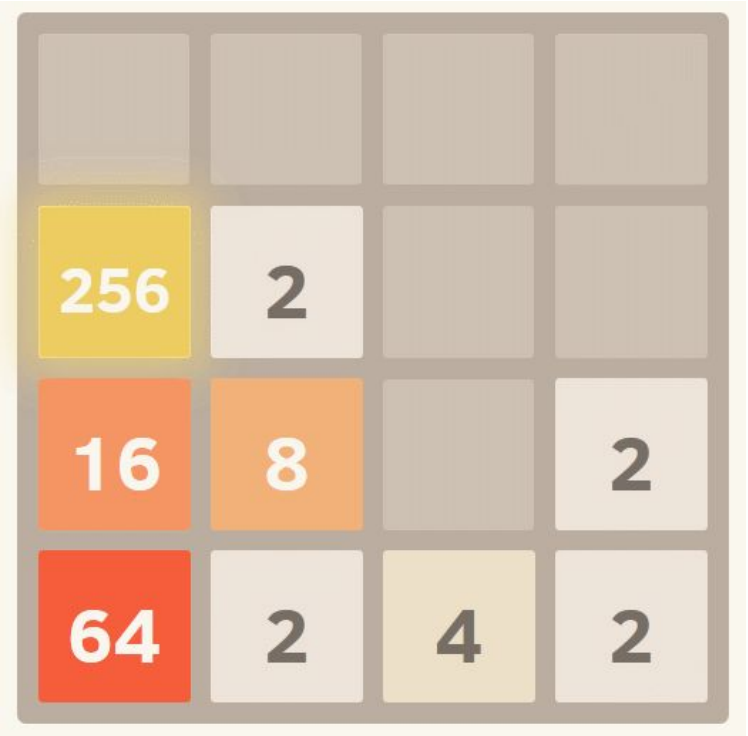


Analytics Vidhya

Overview of Problem Statement

- *Play the game “2048” using speech commands*

Step I - Train a Deep Learning model to identify speech commands



Analytics Vidhya

About the Dataset - Speech Commands

Speech Commands: A Dataset for Limited-Vocabulary Speech Recognition

Pete Warden
Google Brain
Mountain View, California
petewarden@google.com

April 2018

1 Abstract

Describes an audio dataset[1] of spoken words designed to help train and evaluate keyword spotting systems. Discusses why this task is an interesting challenge, and why it requires a specialized dataset that's different from conventional datasets used for automatic speech recognition of full sentences. Suggests a methodology for reproducible and comparable accuracy metrics for this task. Describes how datasets encourages collaborations across groups and enables apples-for-apples comparisons between different approaches, helping the whole field move forward.

The Speech Commands dataset is an attempt to build a standard training and evaluation dataset for a class of simple speech recognition tasks. Its primary goal is to provide a way to build and test small models that detect when a single word is spoken, from a set of ten or fewer target words, with as few false positives as possible from background noise or unre-

Analytics Vidhya

About the Dataset - Data Distribution

Speech Command	Audio Sample	Number of audio files
left		3801
right		3778
up		3723
down		3917

Steps to solve Speech Command Identification

Steps to solve Speech Command Identification

1. Data Loading and Preprocessing

Steps to solve Speech Command Identification

1. Data Loading and Preprocessing

1.1 Install important libraries

Steps to solve Speech Command Identification

1. Data Loading and Preprocessing

1.1 Install important libraries

1.2 Data Exploration

Analytics Vidhya

Steps to solve Speech Command Identification

1. Data Loading and Preprocessing

1.1 Install important libraries

1.2 Data Exploration

1.3 Load data using custom Data Loader

Analytics Vidhya

Steps to solve Speech Command Identification

1. Data Loading and Preprocessing

1.1 Install important libraries

1.2 Data Exploration

1.3 Load data using custom Data Loader

2. Speech Command Identification using CNN

Steps to solve Speech Command Identification

1. Data Loading and Preprocessing

1.1 Install important libraries

1.2 Data Exploration

1.3 Load data using custom Data Loader

2. Speech Command Identification using CNN

2.1 Define model architecture

Steps to solve Speech Command Identification

1. Data Loading and Preprocessing

1.1 Install important libraries

1.2 Data Exploration

1.3 Load data using custom Data Loader

2. Speech Command Identification using CNN

2.1 Define model architecture

2.2 Train the model

Steps to solve Speech Command Identification

1. Data Loading and Preprocessing

1.1 Install important libraries

1.2 Data Exploration

1.3 Load data using custom Data Loader

2. Speech Command Identification using CNN

2.1 Define model architecture

2.2 Train the model

2.3 Get Inference

Thank You

Analytics
Vidhya

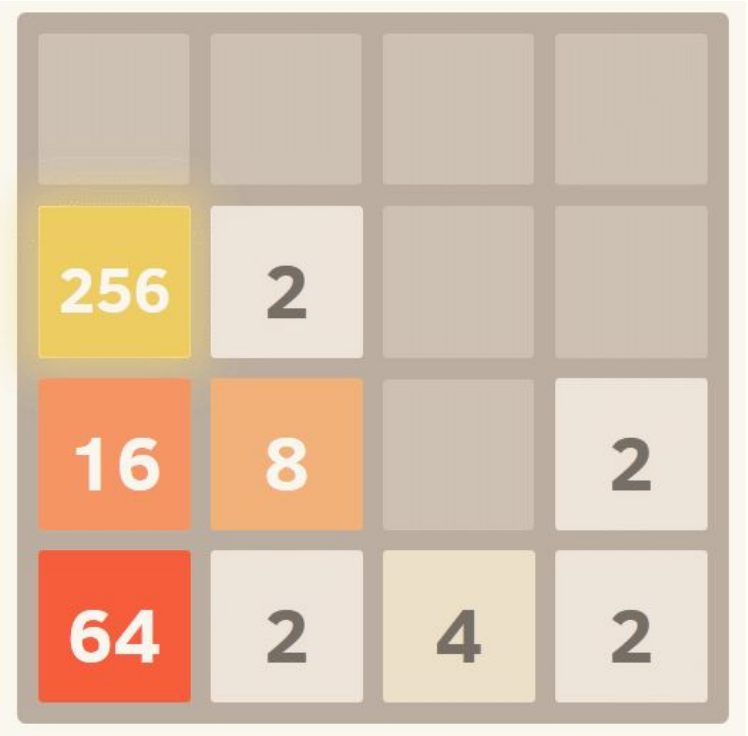
3. Understanding Real-time Audio Capture

Overview of Problem Statement

- *Play the game “2048” using speech commands*

Step I - Train a Deep Learning model to identify speech commands

Step II - Identify the speech commands in real time to play the game



Analytics Vidhya

4. Evaluation of Real-time Speech Commands

5. Play 2048 game using Real-Time Speech Command Identification