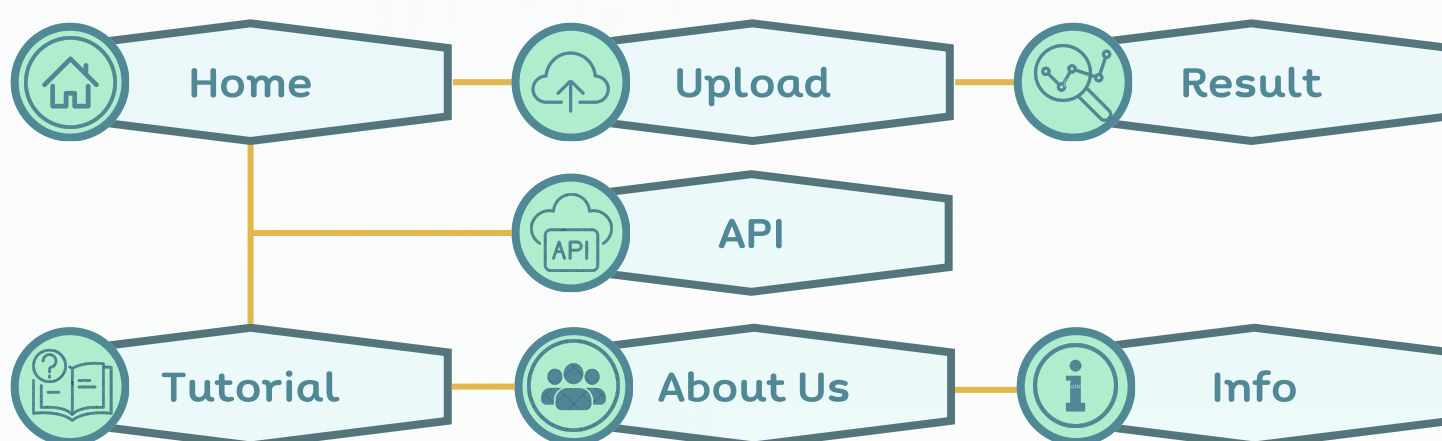







### Introduction

โปรแกรม Vocal Sense เป็นเครื่องมือที่พัฒนาขึ้นเพื่อช่วยตรวจสอบเสียงที่ถูกสังเคราะห์โดยใช้เทคโนโลยี Deep Learning และ Machine Learning มีจุดประสงค์เพื่อป้องกัน และเสริมสร้างความปลอดภัยให้แก่บุคคลที่อาจถูกโจมตีหรือถูกแอบอ้างผ่านการใช้เสียงสังเคราะห์ Vocal Sense ให้บริการในรูปแบบของ Web Service ที่สามารถเข้าถึงได้ง่ายและรวดเร็วผ่านเบราว์เซอร์ โดยไม่จำเป็นต้องติดตั้งโปรแกรมใด ๆ ระบบถูกออกแบบให้ใช้งานง่ายและให้ผลลัพธ์ที่รวดเร็ว

### Application



### Target group

-  **Entertainment Industry**  
(Social Media Voice Verification)
-  **Music Industry**  
(AI-Generated Vocal Verification)
-  **General Public**  
(Call Center Voice Verification)



# VOCALSENSE

### Results

System	Train/Test on	Result	
		EER (%)	min t-CDF
RawNet2	Eng	6.61	0.264
	Thai	16.01	0.510
Random forest + LFCC	Eng	13.76	0.323
	Thai	5.00	0.200
ANN + LFCC	Eng	22.68	0.691
	Thai	10.06	0.366
KAN + LFCC	Eng	17.67	0.521
	Thai	9.23	0.340

ผลการทดลองของแต่ละแบบจำลองโดยใช้ EER (Equal Error Rate) และ min t-DCF (Minimum Tandem Detection Cost Function) เป็นเกณฑ์ในการวัดทั้งภาษาไทยและอังกฤษ \*EER, min t-DCF ยิ่งน้อยยิ่งดี\*

### Methodology

#### 1. Data

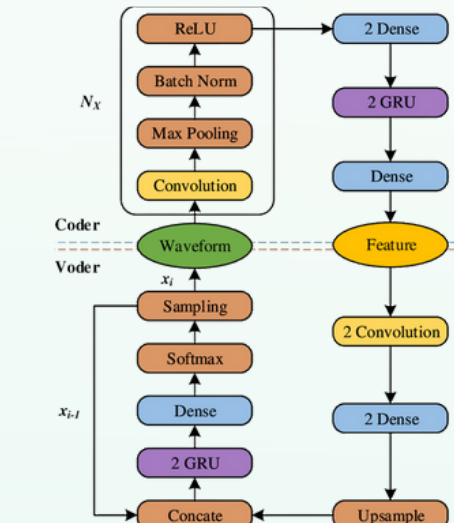


ข้อมูลสำหรับการฝึกสอนโมเดลมาจากหลายแหล่ง ข้อมูลเช่น ASVspoof2021, AI FOR THAI, Mozilla เพื่อให้ครอบคลุมการตรวจจับเสียงปลอมอย่างทั่วถึงรองรับทั้งภาษาไทยและอังกฤษ

#### 3. Functional Specification



#### 2. Model



จากการทดลอง โมเดลที่ดีที่สุดมีชื่อว่า RawNet2 สถาปัตยกรรมของโมเดลมีการใช้ทั้ง CNN และ LSTM รวมถึงยังถูกออกแบบให้ใช้สำหรับ Raw Waveform โดยไม่จำเป็นต้องใช้ Feature

เครื่องมือที่ใช้ในการพัฒนา :

