



# NHẬN DẠNG NGUYÊN ÂM KHÔNG PHỤ THUỘC NGƯỜI NÓI

GVHD: Ninh Khánh Duy



Nhóm 9

Tên thành viên:

- Nguyễn Thùy Linh
- Nguyễn Đỗ Vương
- Vũ Xuân Nguyên
- Dương Phú Ninh



# Nội dung trình bày

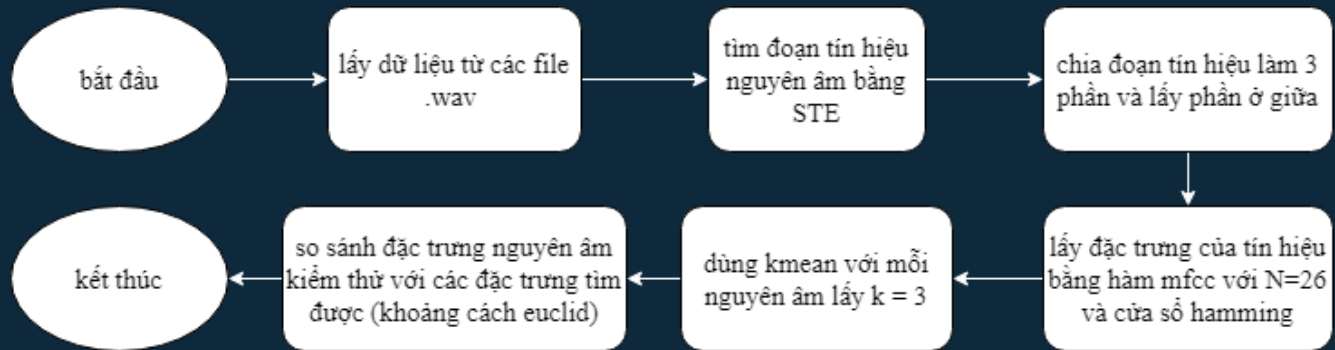
1. Thuật toán nhận dạng nguyên âm
2. So sánh hiệu quả của 2 loại vector đặc trưng FFT và MFCC
3. Đánh giá hiệu quả của K-means khi K thay đổi và nMFCC thay đổi
4. Báo cáo kết quả

A series of hexagonal icons in various shades of blue and cyan are arranged along the left edge of the slide. The icons include a lightbulb, a thumbs-up, a network node, a smartphone, a magnifying glass, a gear, and a speech bubble.

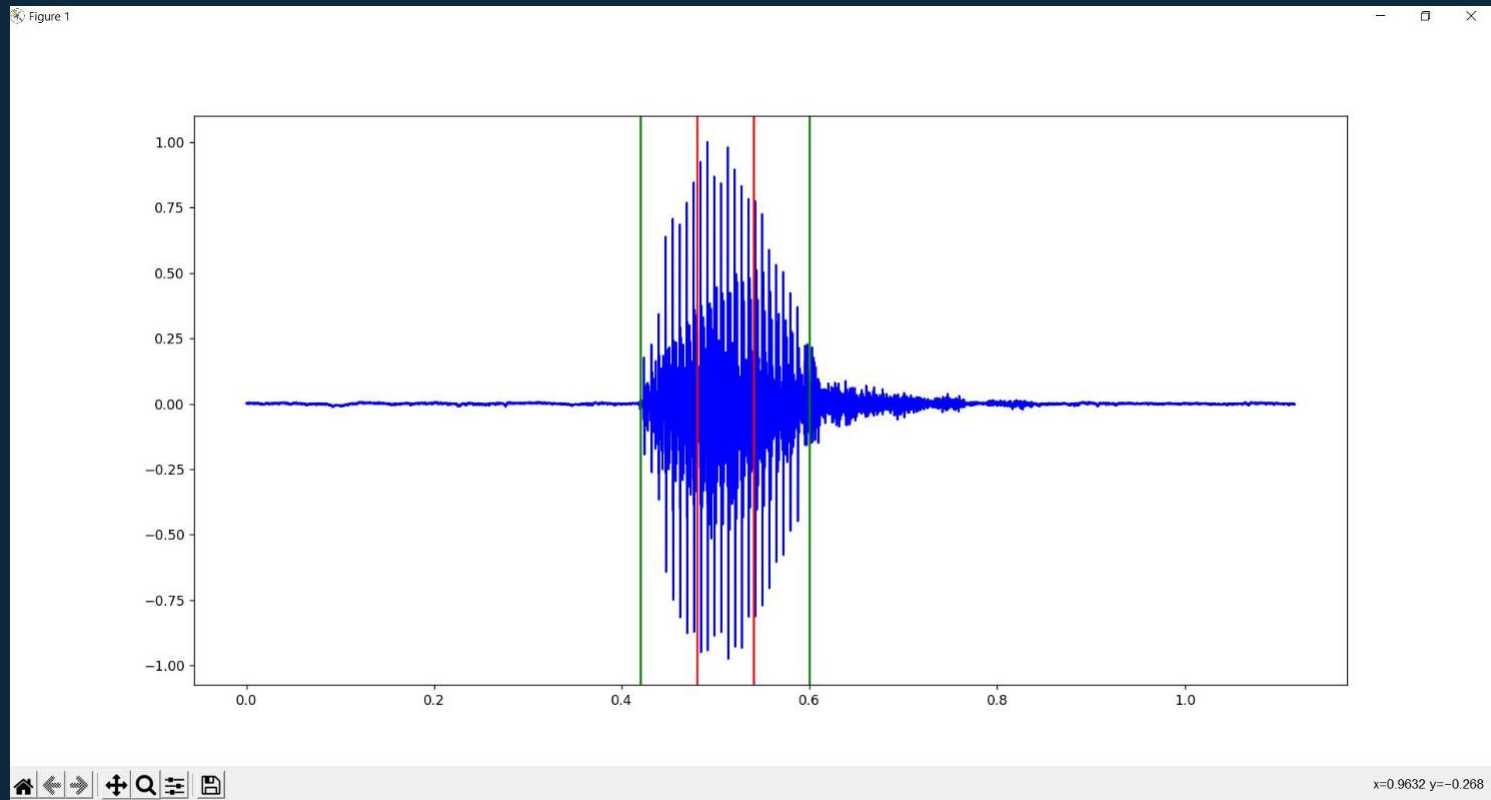
1

# Thuật toán nhận dạng nguyên âm

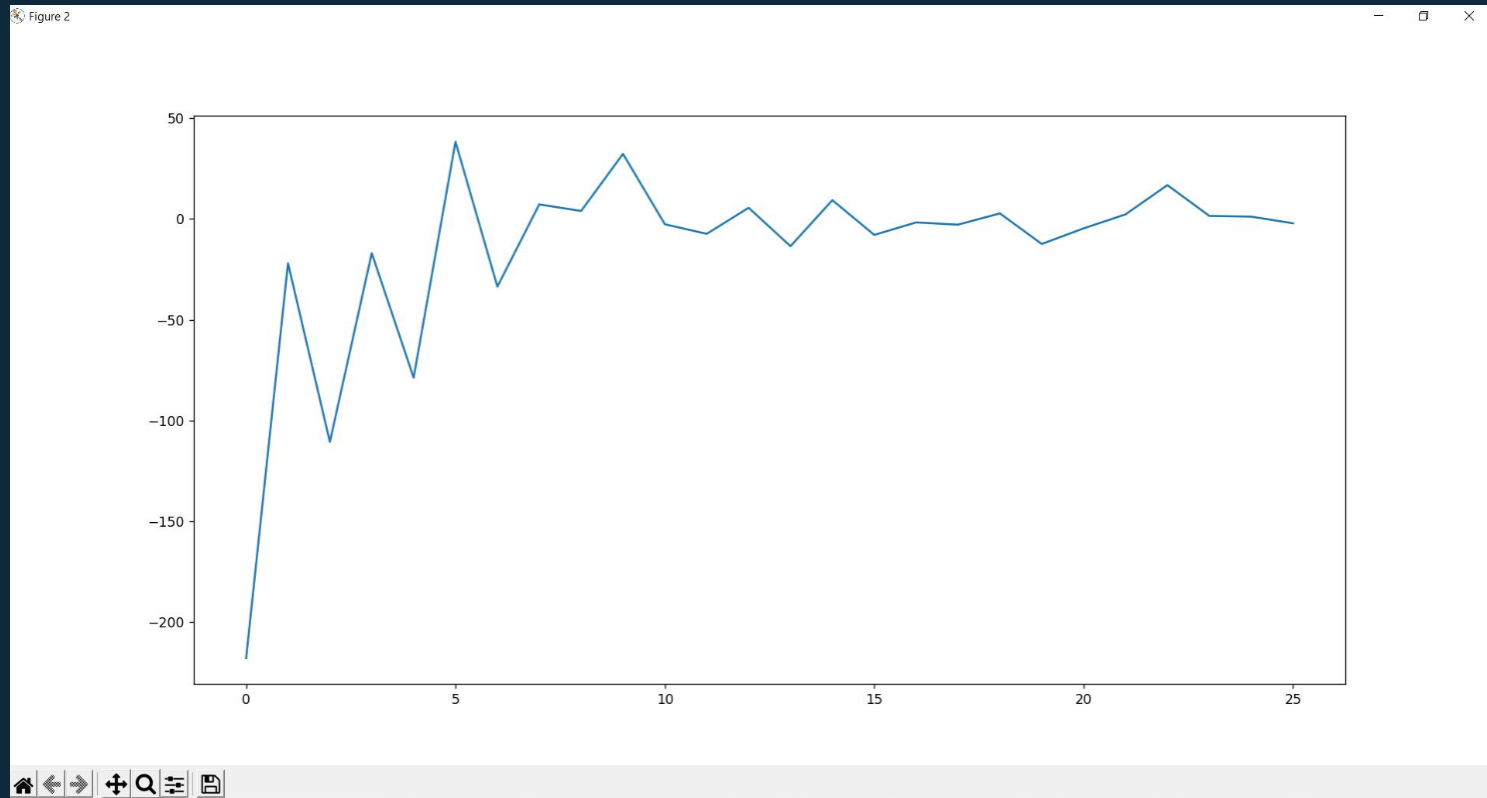
# Nhận dạng nguyên âm bằng vector đặc trưng MFCC



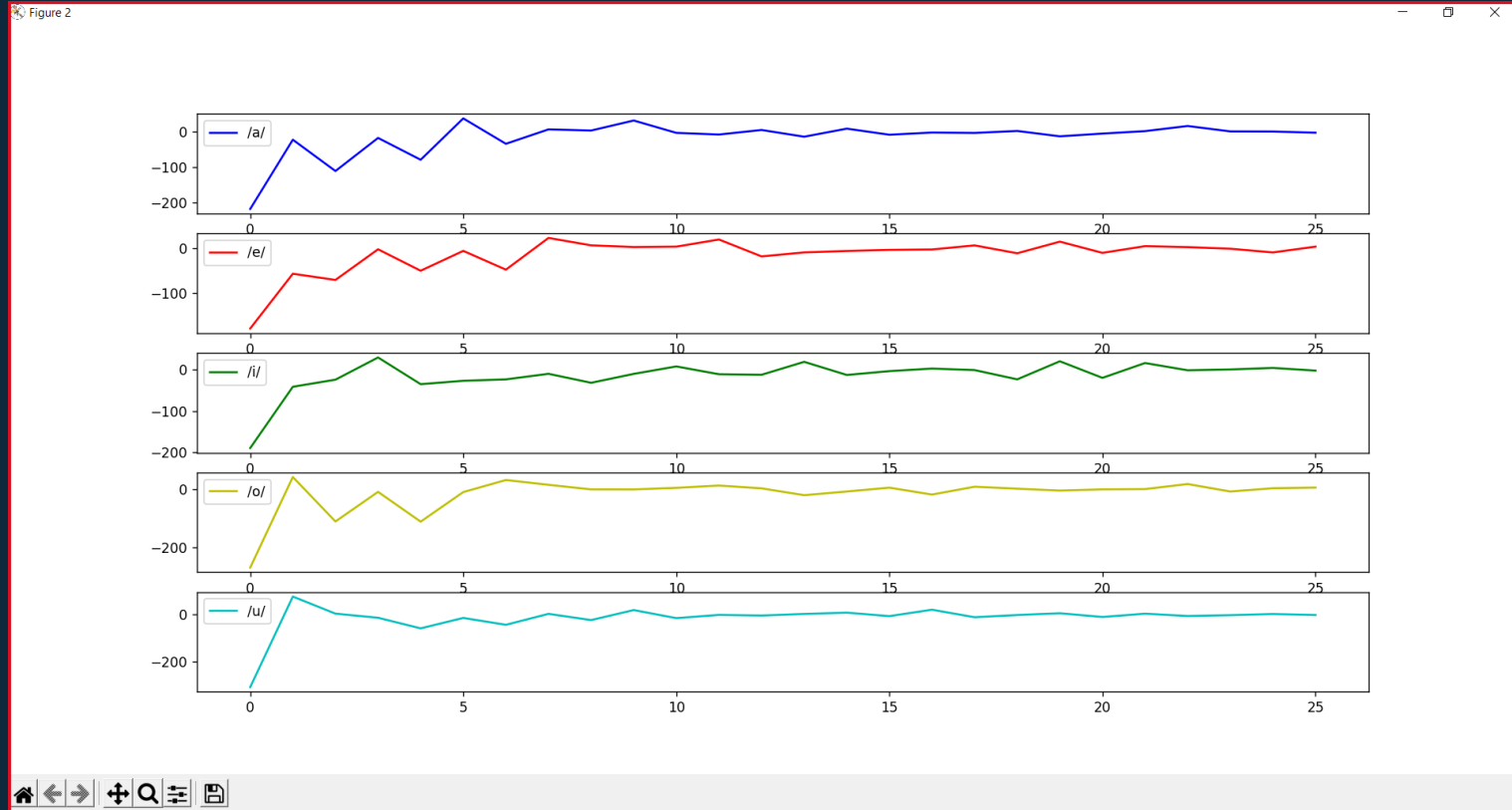
◇ Xác định vùng có đặc trưng phổ ổn định để trích xuất vectơ đặc trưng



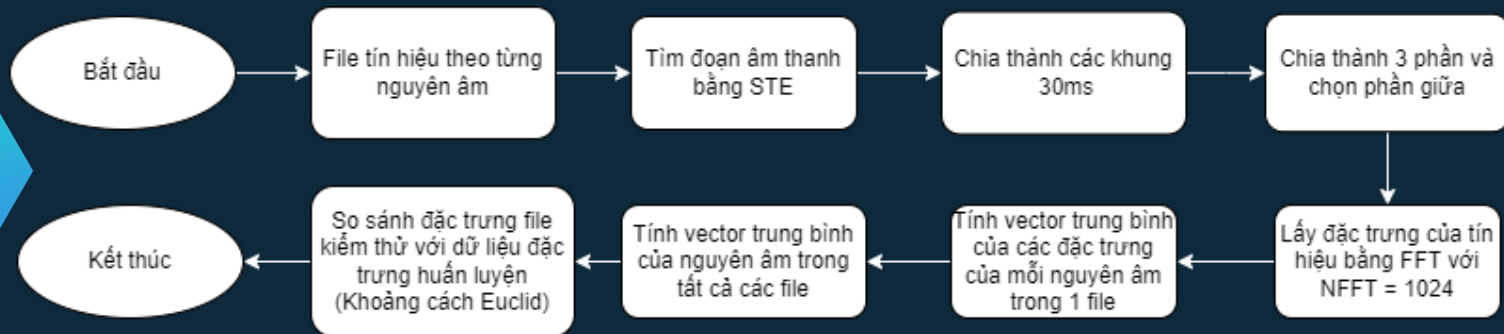
# Xác định vector MFCC từ các khung đã tìm được



# Vector đặc trưng của các nguyên âm

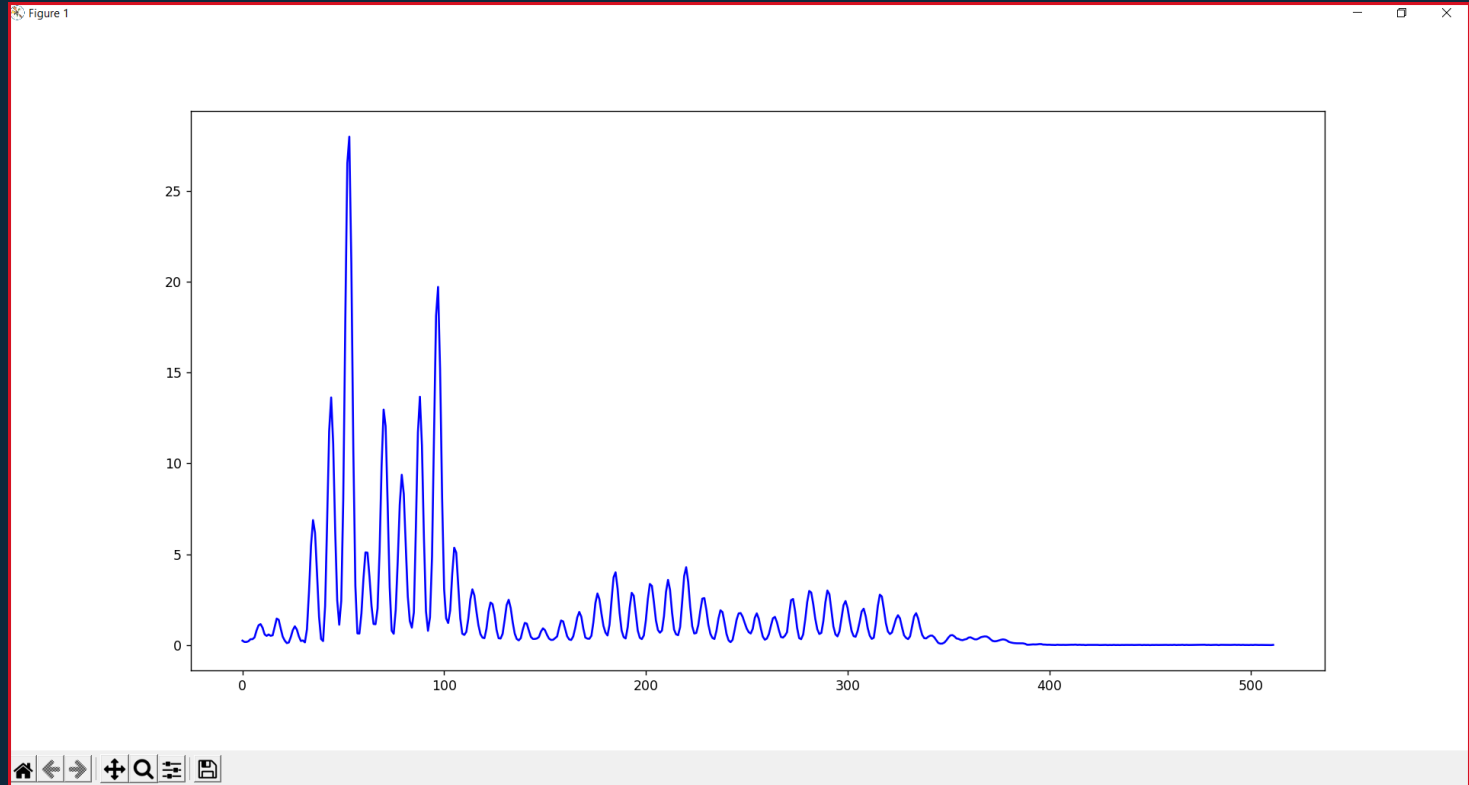


# Nhận dạng nguyên âm bằng vector đặc trưng FFT

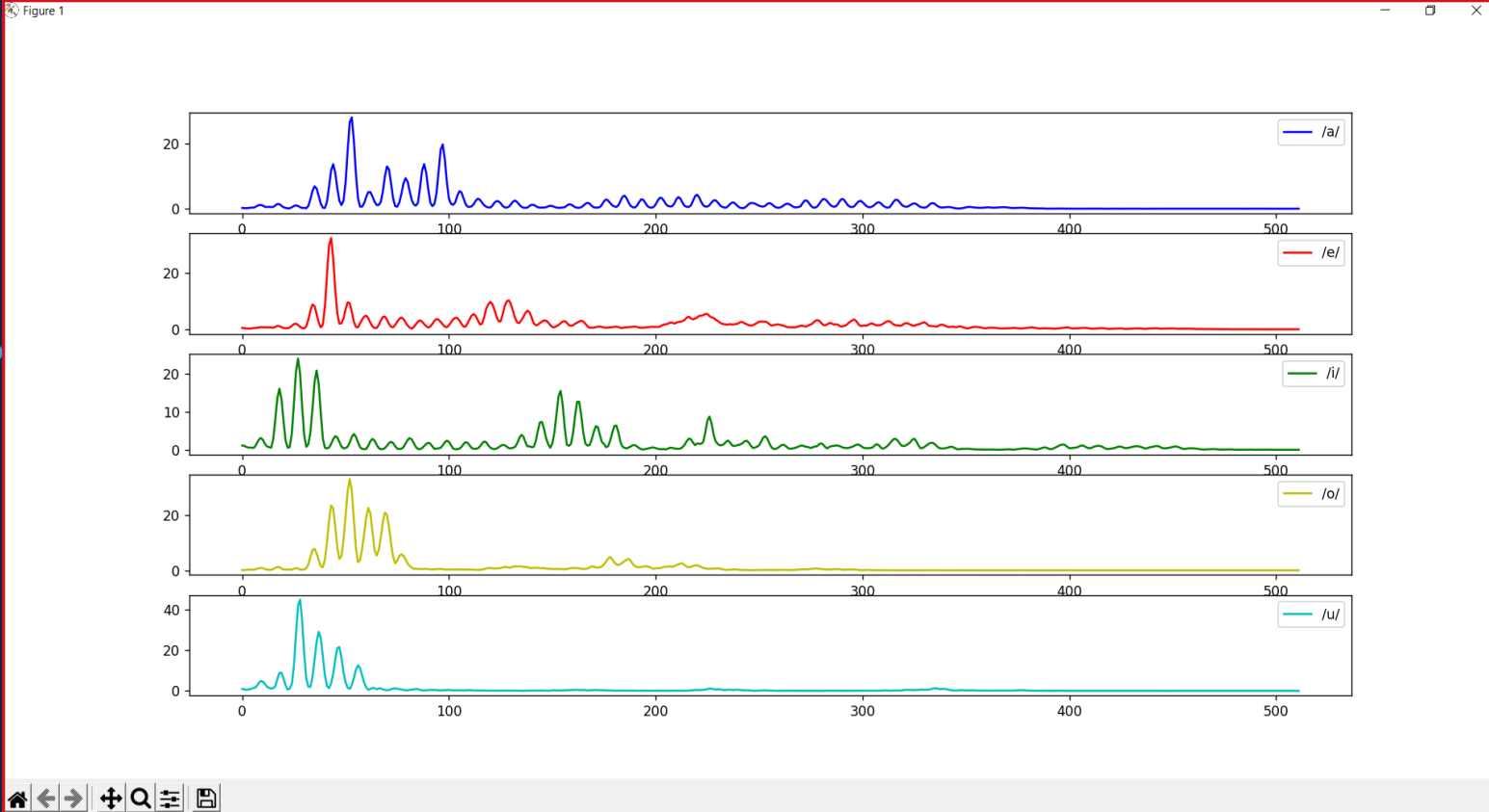




# Vector đặc trưng FFT của 1 nguyên âm



## An abstract graphic featuring a dark blue background with several hexagons of varying shades of blue and teal. The hexagons contain white icons: a lightbulb, a thumbs-up, a network of nodes, a smartphone, a magnifying glass, a gear, and a speech bubble. A large, stylized white double quote mark is centered in the middle of the composition.





2

Số sánh hiệu quả của 2 loại vector  
đặc trưng FFT và MFCC

# Ma trận nhầm lẫn

FFT					
	a	e	i	o	u
a	17	3	0	1	0
e	0	15	5	1	0
o	0	1	18	1	1
i	4	2	1	13	1
u	0	0	1	1	19
Độ chính xác 78.0952380952381 %					

MFCC(Average)					
	a	e	i	o	u
a	19	1	0	1	0
e	0	18	3	0	0
o	0	0	20	0	1
i	4	2	0	15	0
u	0	0	1	0	20
Độ chính xác 87.61904761904762 %					



# Nhận xét

- MFCC, FFT đều có khả năng nhận dạng nguyên âm
- MFCC cho kết quả nhận dạng nguyên âm tốt hơn FFT 9.6%





3

Khảo sát hiệu quả nhận dạng khi  
số cụm KMean thay đổi và nMFCC  
thay đổi

# Thống kê ảnh hưởng của các tham số

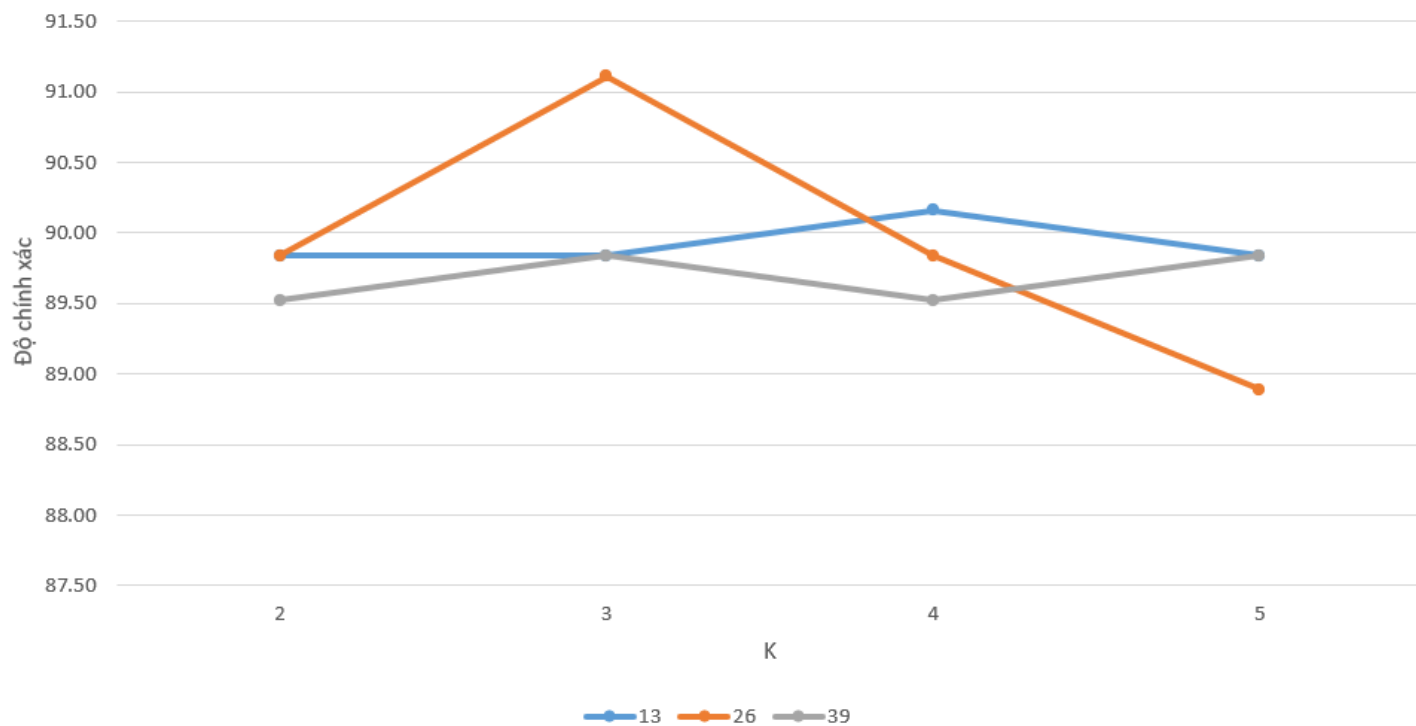
nMFCC \ K	13			
	Lần 1	Lần 2	Lần 3	Trung bình
2	89.52	89.52	90.48	89.84
3	91.43	89.52	89.52	90.16
4	89.52	90.48	89.52	89.84
5	85.71	87.62	90.48	87.94

nMFCC \ K	26			
	Lần 1	Lần 2	Lần 3	Trung bình
2	89.52	90.48	89.52	89.84
3	91.43	90.48	91.43	91.11
4	87.62	91.43	90.48	89.84
5	91.43	87.62	87.62	88.89

nMFCC \ K	39			
	Lần 1	Lần 2	Lần 3	Trung bình
2	89.52	89.52	89.52	89.52
3	90.48	89.52	89.52	89.84
4	87.62	89.52	91.43	89.52
5	90.48	88.57	90.48	89.84



Biểu đồ sự thay đổi độ chính xác khi thay đổi nMFCC và kMeans








# Nhận xét

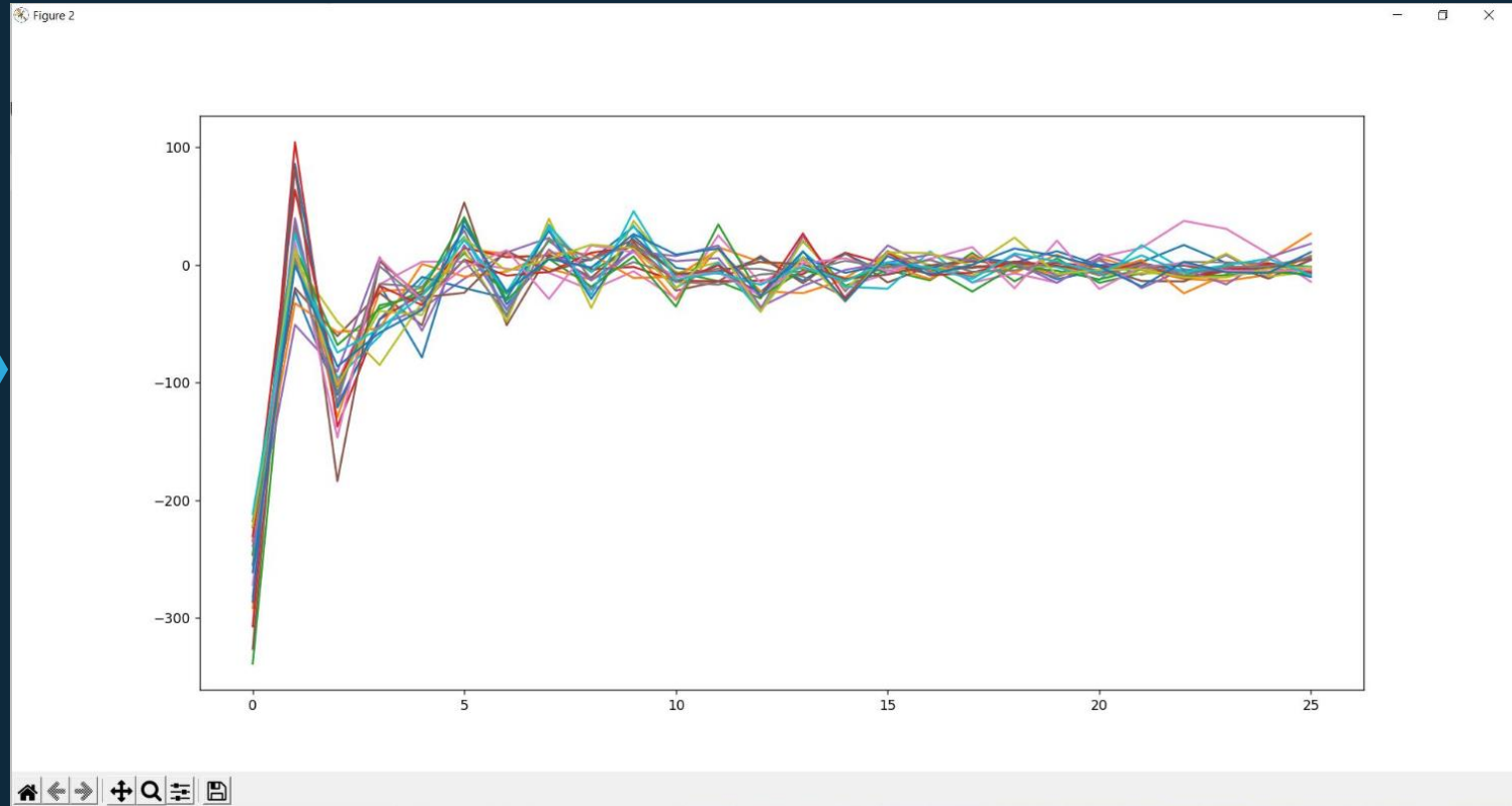
- ◇ Việc thay đổi nMFCC và số cụm Kmean có thể thay đổi hiệu suất bài toán từ 88.89 - 91.11%
- ◇ Với tham số nMFCC=26 k=3 cho hiệu suất nhận dạng tốt nhất là 91.11%
- ◇ Phương pháp Kmeans cải thiện 3.51% so với sử dụng trung bình các vector đặc trưng



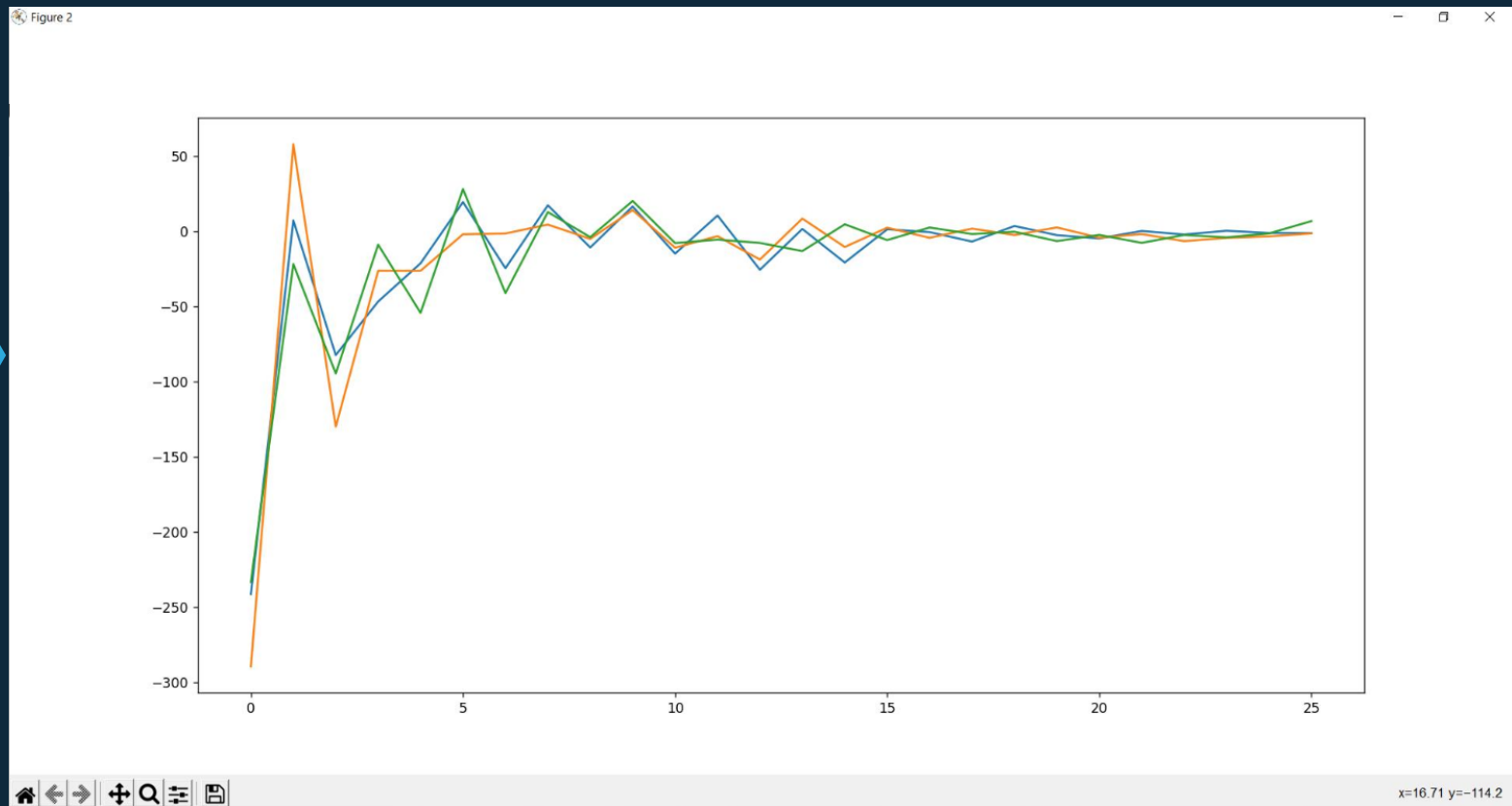
A decorative graphic on the left side of the slide consists of a large cyan hexagon in the center, surrounded by several smaller hexagons of varying shades of blue and cyan. These smaller hexagons contain white icons: a lightbulb, a thumbs-up, a network of nodes, a smartphone, a magnifying glass, a gear, and a speech bubble.

Kmeans với  $K=3$  và  $n\text{MFCC}=26$

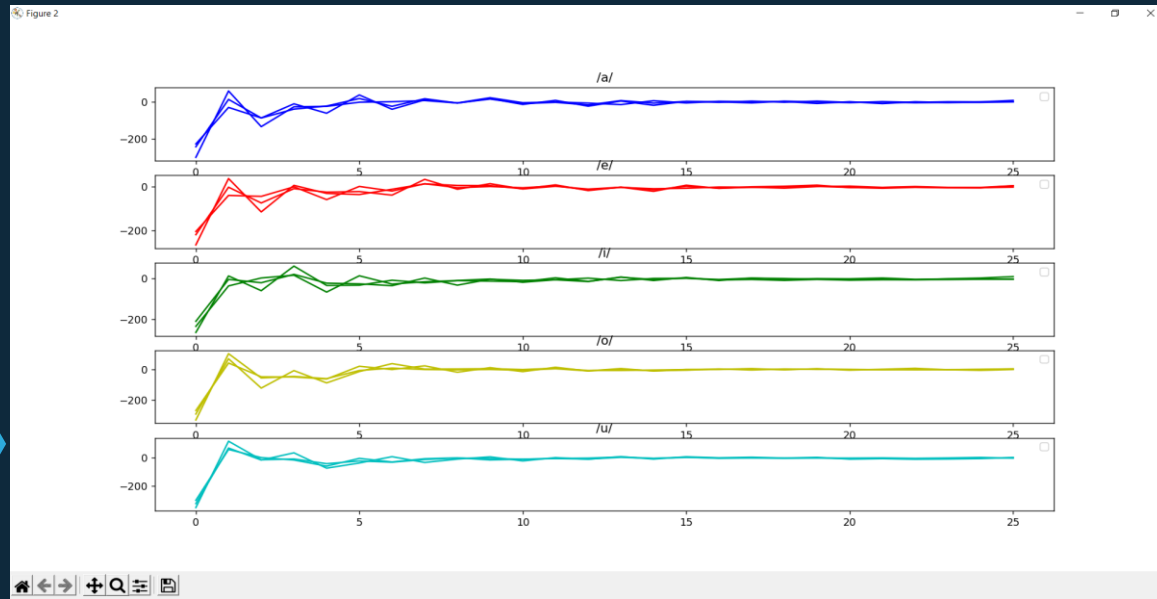
# Vector đặc trưng 1 nguyên âm của các người nói



## Kmeans với k=3



# Kmeans với k=3



➡ Tính khoảng cách từ các vector đặc trưng MFCC của tín hiệu kiểm thử đến các vector đặc trưng đã tìm được từ Kmeans

➡ Tìm khoảng cách ngắn nhất

➡ Kết luận

A decorative cluster of hexagonal icons in various shades of blue and cyan is located on the left side of the slide. The icons include a lightbulb, a thumbs-up, a network diagram, a smartphone, a magnifying glass, a gear, and a speech bubble.

4

## Báo cáo kết quả



# Kết quả thực thi với nMFCC=26 và k=3

	/a/	/e/	/i/	/o/	/u/
23MTL	✓	✓	✓	✓	✓
24FTL	✗	✗	✓	✓	✓
25MLM	✓	✗	✓	✓	✓
27MCM	✓	✓	✓	✓	✓
28MVN	✓	✓	✓	✗	✓
29MHN	✓	✓	✓	✓	✓
30FTN	✓	✓	✗	✓	✓
32MTP	✓	✓	✓	✓	✓
33MHP	✓	✓	✓	✓	✓
34MQP	✓	✓	✓	✗	✓
35MMQ	✗	✗	✓	✓	✓
36MAQ	✓	✗	✓	✓	✓
37MDS	✓	✓	✓	✓	✓
38MDS	✓	✓	✓	✓	✓
39MTS	✓	✓	✓	✓	✓
40MHS	✓	✓	✓	✓	✓
41MVS	✓	✓	✓	✓	✓
42FQT	✓	✓	✓	✓	✓
43MNT	✓	✓	✓	✓	✓
44MTT	✓	✓	✓	✓	✓
45MDV	✓	✓	✓	✓	✓

	a	e	i	o	u
a	19	1	0	1	0
e	1	17	3	0	0
o	0	0	20	0	1
i	2	0	0	19	0
u	0	0	0	0	21
Do chính xác 91.4 %					

## Nhận xét

- Nguyên âm /u/ có kết quả nhận dạng chính xác nhất (100%)
- Nguyên âm /e/ có kết quả nhận dạng kém chính xác nhất (80.95%)
- Hiệu suất nhận dạng là 91.4%



## ◇ NHẬN XÉT CHUNG

- Hoàn thành các yêu cầu cơ bản của bài toán
- Phân biệt nguyên âm dựa vào vector đặc trưng FFT và MFCC, so sánh hiệu quả
- Khảo sát được sự thay đổi hiệu suất bài toán khi thay đổi nMFCC và k (Kmean)
- Bài toán đạt hiệu suất tốt nhất khi sử dụng vector đặc trưng MFCC (nMFCC =26) và phương pháp Kmeans (k=3)





Thanks!

