Depression Speaks

Automatic Discrimination Between Depressed and Non-Depressed Speakers Based on Nonverbal Speech Features

F.Scibelli, G.Roffo, M.Tayarani, L.Bartoli, G.De Mattia, A.Esposito and A.Vinciarelli

Universita' degli Studi della Campania (Italy)
Unita' Operative Salute Mentale (Italy)
University of Glasgow (UK)

Outline

- Introduction
- The Data
- The Approach
- Experiments and Results
- Conclusions

Outline

- Introduction
- The Data
- The Approach
- Experiments and Results
- Conclusions

Why Depression?

Why Depression?

- 300 million patients in 2015 (World Health Organisation)

Why Depression?

- 300 million patients in 2015 (World Health Organisation)
- Second cause of disability after ischaemia (European Commission)

Why Depression?

- 300 million patients in 2015 (World Health Organisation)
- Second cause of disability after ischaemia (European Commission)
- Most important suicide factor for elderly people (European Commission)

Why Nonverbal (Speech) Behaviour?

Why Nonverbal (Speech) Behaviour?

- Traditional methods (HRSD and BDI-II) not fully robust to biases of both clinicians and patients

Why Nonverbal (Speech) Behaviour?

- Traditional methods (HRSD and BDI-II) not fully robust to biases of both clinicians and patients
- Detection of nonverbal behavioural markers can limit the effect of biases while reducing the costs

Outline

- Introduction
- The Data
- The Approach
- Experiments and Results
- Conclusions

Depression

Control

Depression

20

Control

Depression

20 42

Control

Depression

20 42

Control

20

Depression

20 42

Control

20 37

Depression

20 42

DSM-based diagnosis by professional clinicians

Control

20 37

Depression

20 42

DSM-based diagnosis by professional clinicians

Control

20 37

Match based on gender, age and education

Depression

20 42

DSM-based diagnosis by professional clinicians

Control

20 37

Match based on gender, age and education

BDI-II Questionnaire

Depression

20 42

DSM-based diagnosis by professional clinicians

Control

20 37

Match based on gender, age and education

BDI-II Questionnaire

62 Diary recordings
57 Tale recordings

Depression

20 42

DSM-based diagnosis by professional clinicians

BDI-II Questionnaire

62 Diary recordings
57 Tale recordings

Control

20 37

Match based on gender, age and education

52 Diary recordings
54 Tale recordings

Depression

20 42

Control

20 37

Diary is an interview about the activities of the last week end (spontaneous speech)

62 Diary recordings
57 Tale recordings

52 Diary recordings
54 Tale recordings

Depression

20 42

Control

20 37

Diary is an interview about the activities of the last week end (spontaneous speech)

62 Diary recordings
57 Tale recordings

52 Diary recordings
54 Tale recordings

Depression

20 42

Control

20 37

Diary is an interview about the activities of the last week end (spontaneous speech)

Depression:

62 Diary recordings
57 Tale recordings

Control:

52 Diary recordings
54 Tale recordings

Depression

20 42

62 Diary recordings
57 Tale recordings

Control

20 37

52 Diary recordings
54 Tale recordings

Tale is reading a short story written by Aesopus (read speech)

Depression

20 42

Control

20 37

Tale is reading a short story written by Aesopus (read speech)

Depression:
62 Diary recordings
57 Tale recordings

Control:
52 Diary recordings
54 Tale recordings

Depression

20 42

Control

20 37

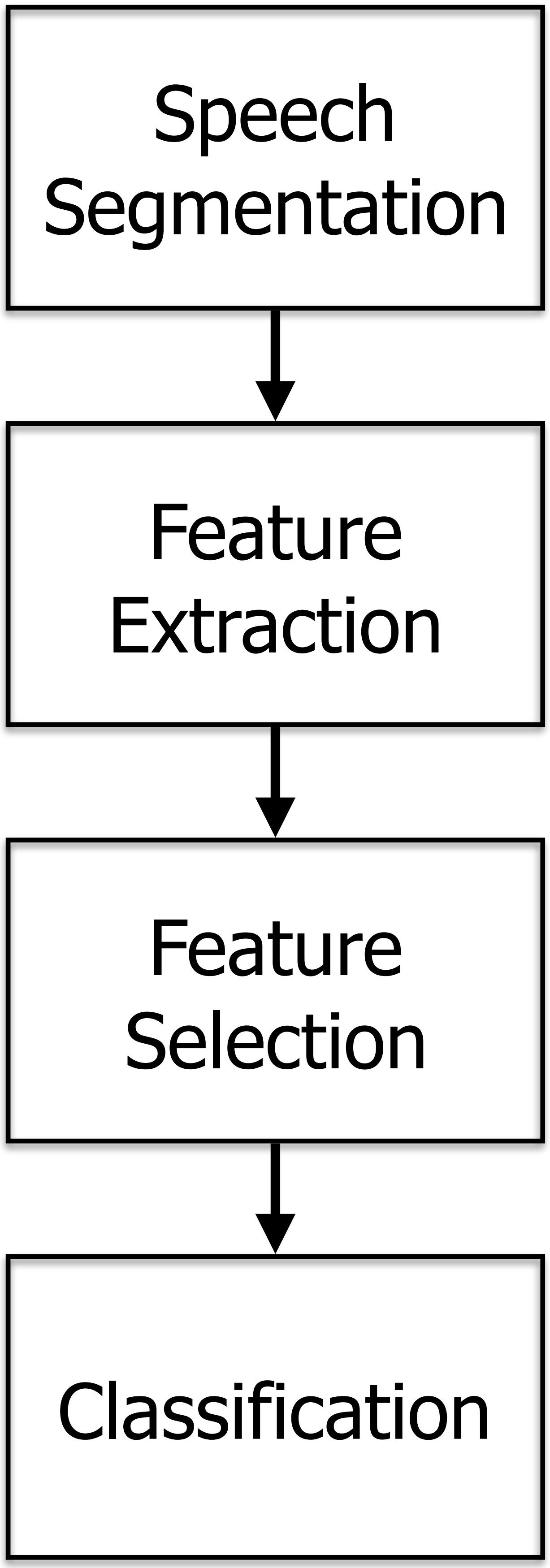
Tale is reading a short story written by Aesopus (read speech)

Depression:
62 Diary recordings
57 Tale recordings

Control:
52 Diary recordings
54 Tale recordings

Outline

- Introduction
- The Data
- The Approach
- Experiments and Results
- Conclusions


Speech Segmentation
Feature Extraction
Feature Selection
Classification

Speech Segmentation

↓

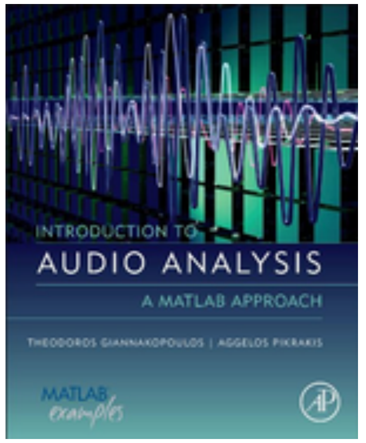
Feature Extraction

↓

Feature Selection

↓

Classification



Speech Segmentation

↓

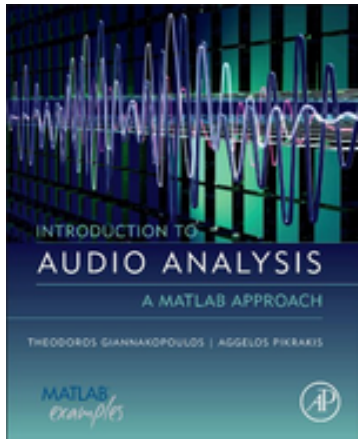
Feature Extraction

↓

Feature Selection

↓

Classification



Selection of patient or control speech

Speech Segmentation

↓

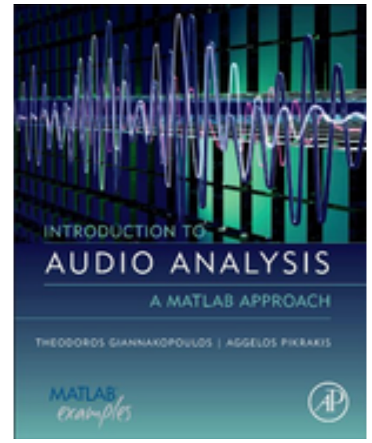
Feature Extraction

↓

Feature Selection

↓

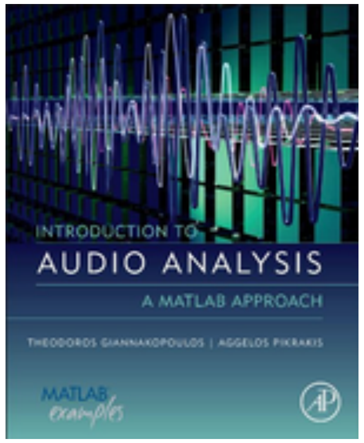
Classification



openSMILE:)
by audEERING™

Selection of patient or control speech

Speech Segmentation



Selection of patient or control speech

Feature Extraction

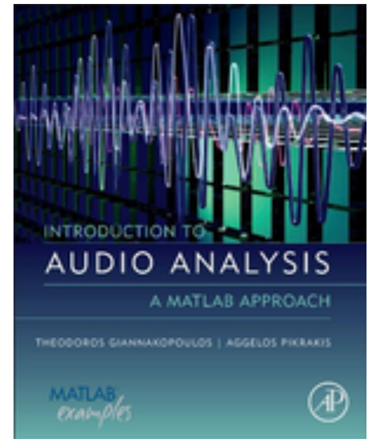
openSMILE:) by audEERING™

Extraction of 384 features (IS09)

Feature Selection

Classification

Speech Segmentation



Selection of patient or control speech

↓

Feature Extraction

openSMILE:) by audEERING™

Extraction of 384 features (IS09)

↓

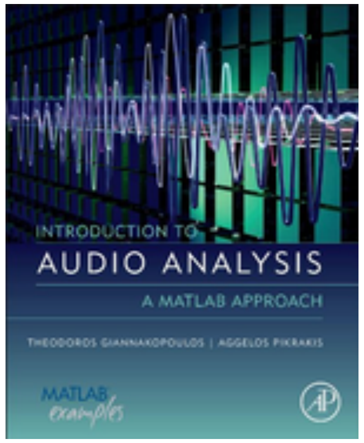
Feature Selection

Infinite Latent Feature Selection

↓

Classification

Speech Segmentation



Selection of patient or control speech

↓

Feature Extraction

openSMILE:) by audEERING™

Extraction of 384 features (IS09)

↓

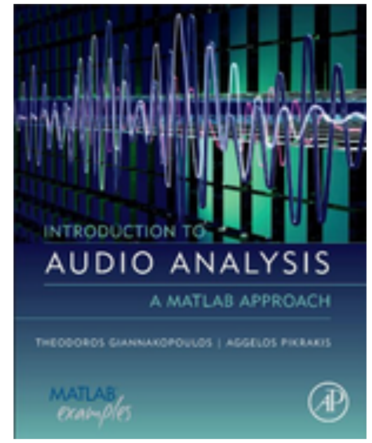
Feature Selection

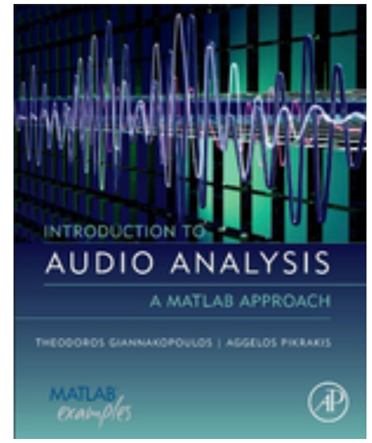
Infinite Latent Feature Selection

Identification of markers

↓

Classification

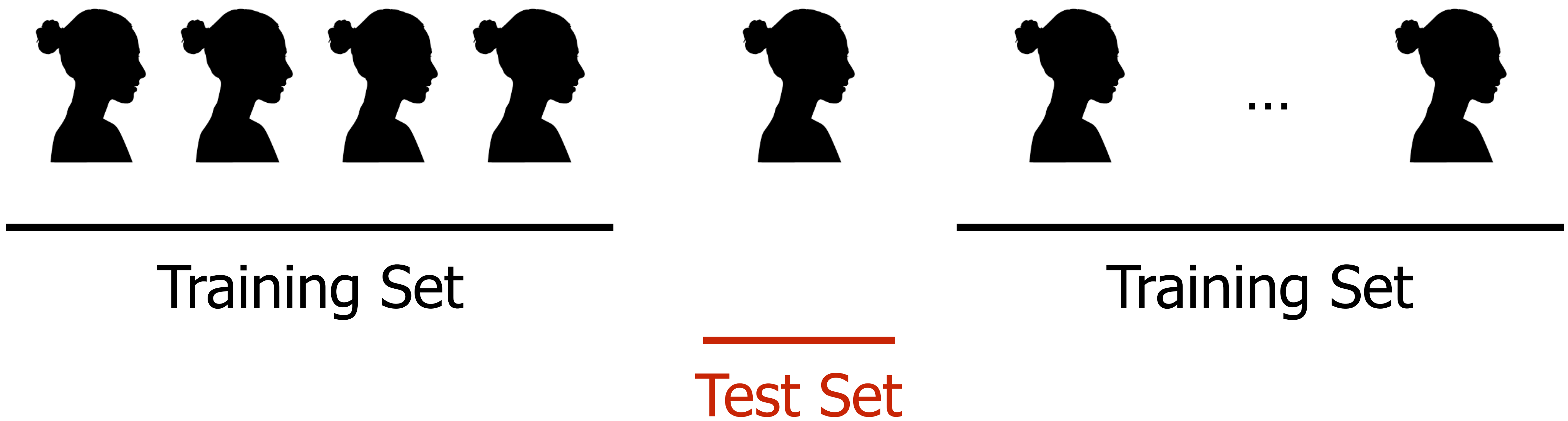
Speech Segmentation		Selection of patient or control speech
↓		
Feature Extraction	openSMILE:) by audEERING™	Extraction of 384 features (IS09)
↓		
Feature Selection	Infinite Latent Feature Selection	Identification of markers
↓		
Classification	Support Vector Machines	

Speech Segmentation		Selection of patient or control speech
Feature Extraction	openSMILE:) by audEERING™	Extraction of 384 features (IS09)
Feature Selection	Infinite Latent Feature Selection	Identification of markers
Classification	Support Vector Machines	Depressed vs Control

Outline

- Introduction
- The Data
- The Approach
- Experiments and Results
- Conclusions

Leave One Subject Out



Leave One Subject Out approach for both Linear Kernel SVM and Infinite Latent Feature Selection

Results

Task	Precision	Recall	Accuracy
Diary	66%	60%	68%
Tale	75%	74%	76%
Diary-FS	74%	65%	74%
Tale-FS	74%	80%	77%

Diary

	D	C
D	74.2%	25.8%
C	40.4%	59.6%

Diary-FS

	D	C
D	80.62%	19.4%
C	34.6%	65.4%

Tale

	D	C
D	77.2%	22.8%
C	25.9%	74.1%

Tale-FS

	D	C
D	73.7%	26.3%
C	20.4%	79.6%

Outline

- Introduction
- The Data
- The Approach
- Experiments and Results
- Conclusions

Conclusions

Conclusions

- While being preliminary, the approach discriminates well above chance

Conclusions

- While being preliminary, the approach discriminates well above chance
- Unlike previous works, the approach performs better on read speech (in line with the psychiatric literature)

Conclusions

- While being preliminary, the approach discriminates well above chance
- Unlike previous works, the approach performs better on read speech (in line with the psychiatric literature)
- Future work includes the adoption of Deep Networks and the use of the transcriptions

Thank You!

F.Scibelli, G.Roffo, M.Tayarani, L.Bartoli, G.De Mattia,
A.Esposito and A.Vinciarelli

Universita' degli Studi della Campania (Italy)
Unita' Operative Salute Mentale (Italy)
University of Glasgow (UK)

University of Glasgow

UNIVERSITÀ DEGLI STUDI DELLA CAMPANIA LUIGI VANVITELLI

EPSRC
Engineering and Physical Sciences Research Council