Feature optimization using Linear Discriminant Analysis for phoneme classification

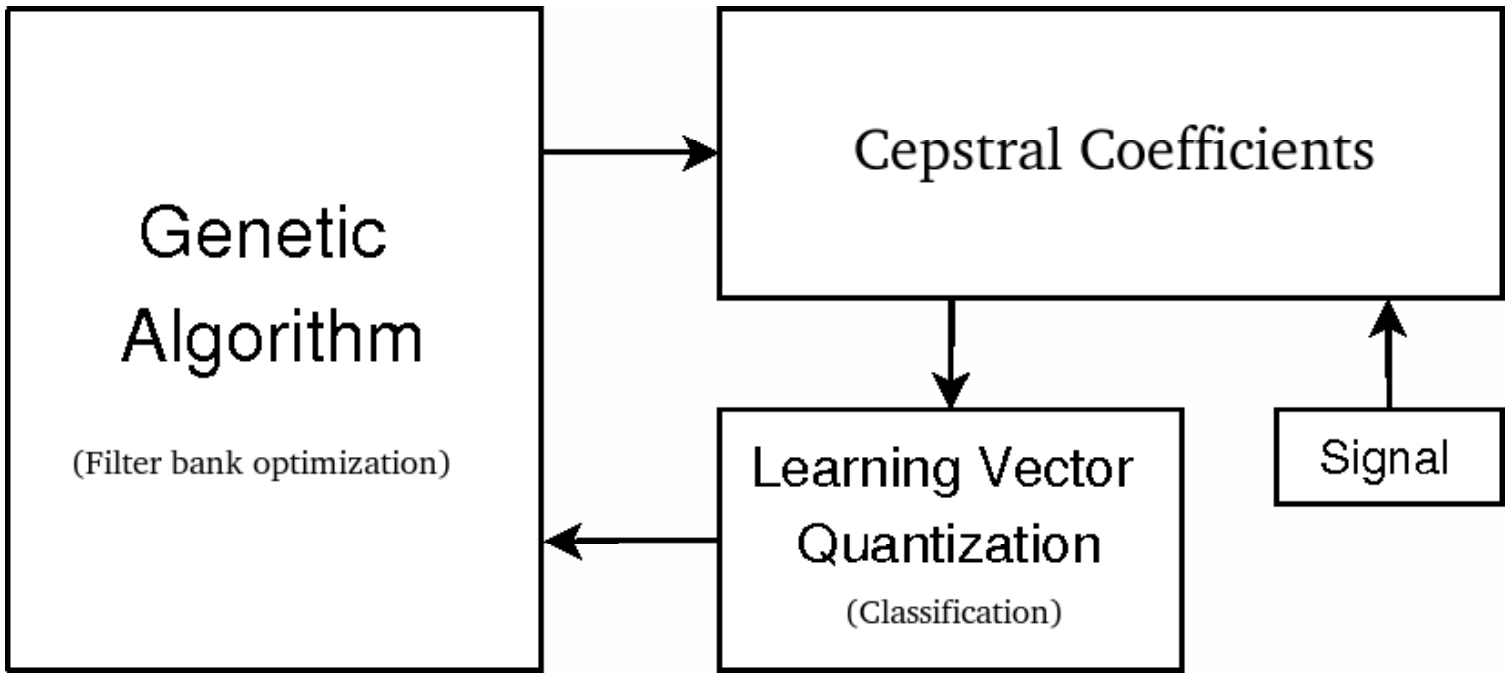
Leandro D. Vignolo

sinc(*i*)

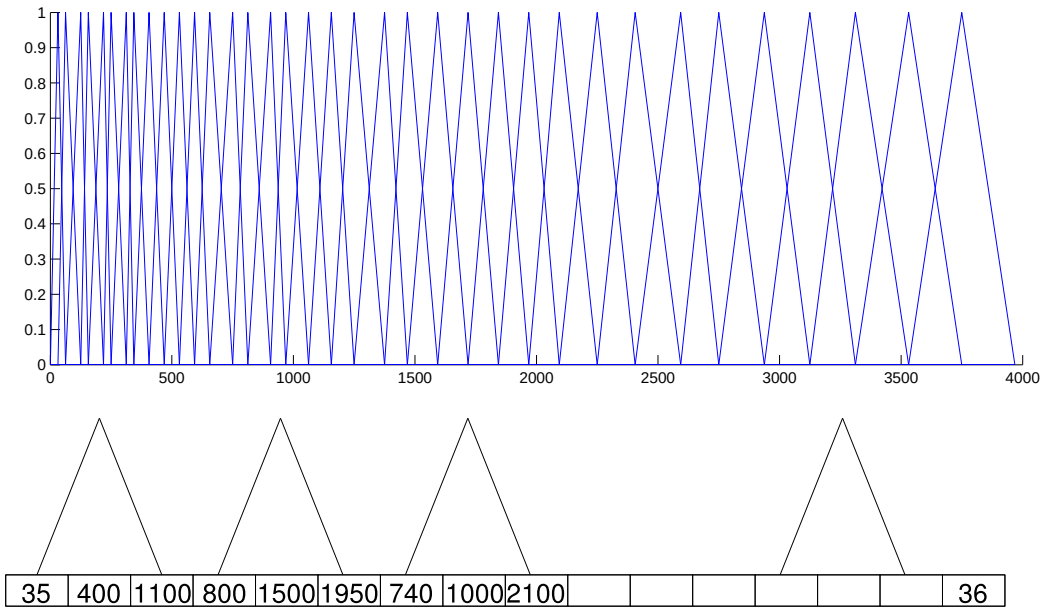
Grupo de Investigación en Señales e Inteligencia Computacional.
Departamento de Informática,
Facultad de Ingeniería y Ciencias Hídricas,
Universidad Nacional del Litoral

leandro.vignolo@gmail.com

The GA-FCC method



Mel Frequency Cepstral Coefficients



- Each chromosome represents a different filter bank.
- 3 parameters per filter are optimized.
- The total number of filters is also optimized.

Reducing dimensions with LDA

- Once we found an optimized filter bank we parameterize the data obtaining patterns in $\mathbb{R}^{16}$.
- Before classification we can use LDA to find an embedded space with:
  - Reduced dimensionality
  - Maximum class separation

Weighted LDA

- If one or more classes are far away from others, it is not necesary that their within-class scatter covariances be minimised in the transformed space.
- WLDA (Tang et al., 2004) uses different weights in the covariance estimation procedure.

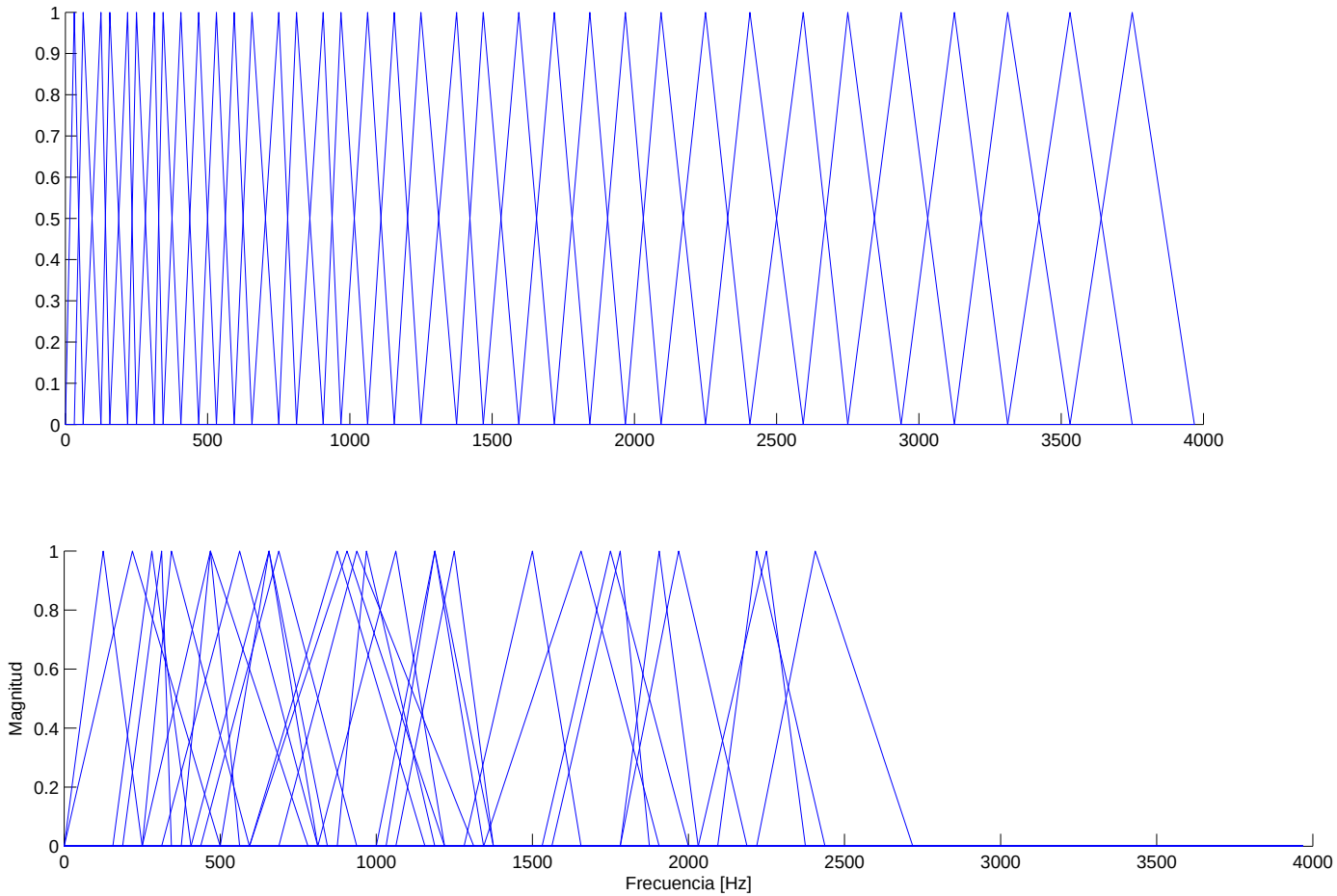
$$\mathbf{S}_W = \sum_{i=1}^{C} r_i p_i \mathbf{S}_{Wi}$$

$$r_i = \sum_{j \neq 1} \frac{1}{L_{ij}}$$

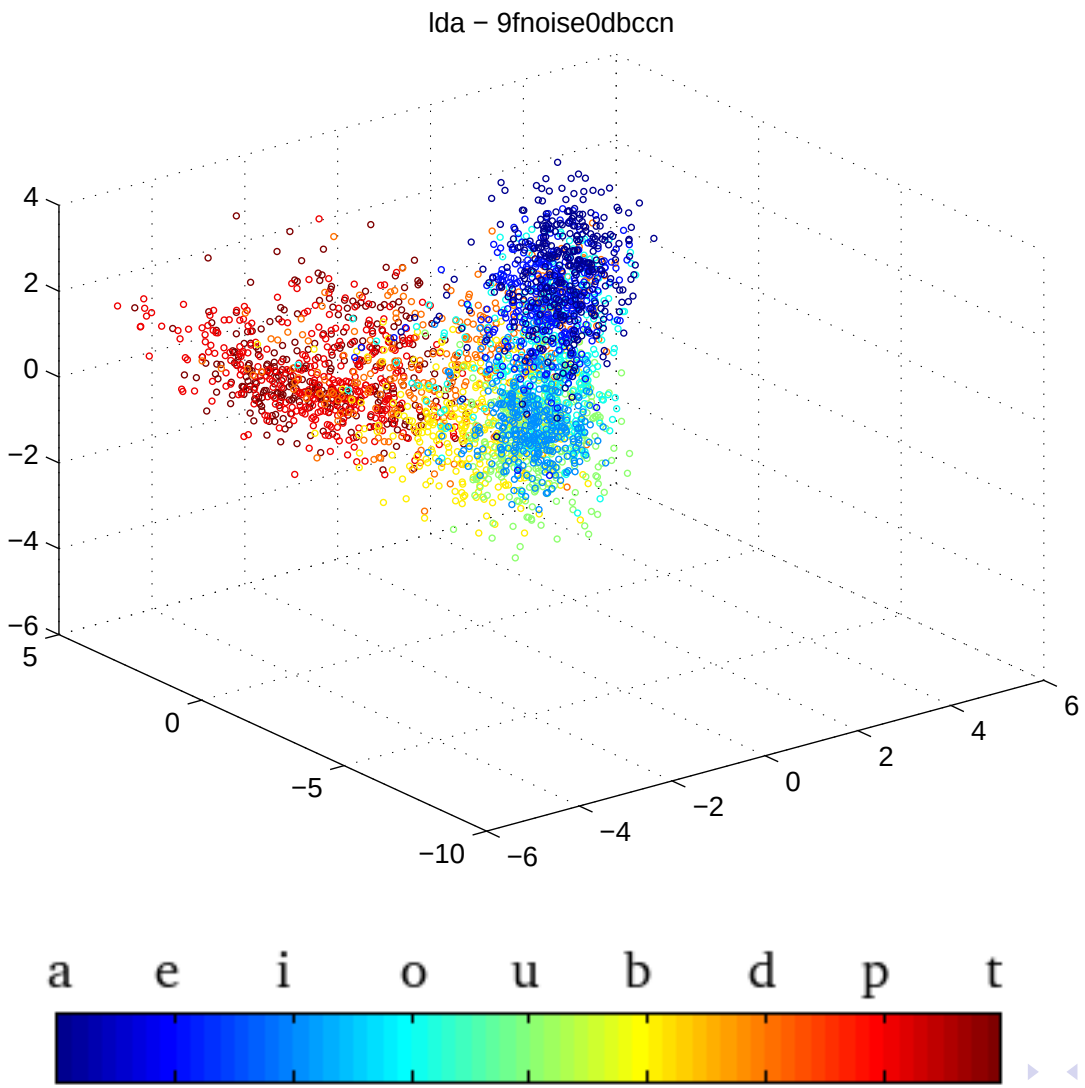
Experiments

- Albayzin database subset with 600 phonetically segmented voice files.
- 12 speakers (6 men, 6 women).
- Segments corresponding to different phonemes were extracted.
- 256 samples from the center of each phoneme.
- 9 Spanish phonemes: /a/, /e/, /i/, /o/, /u/, /b/, /d/, /p/ and /t/.

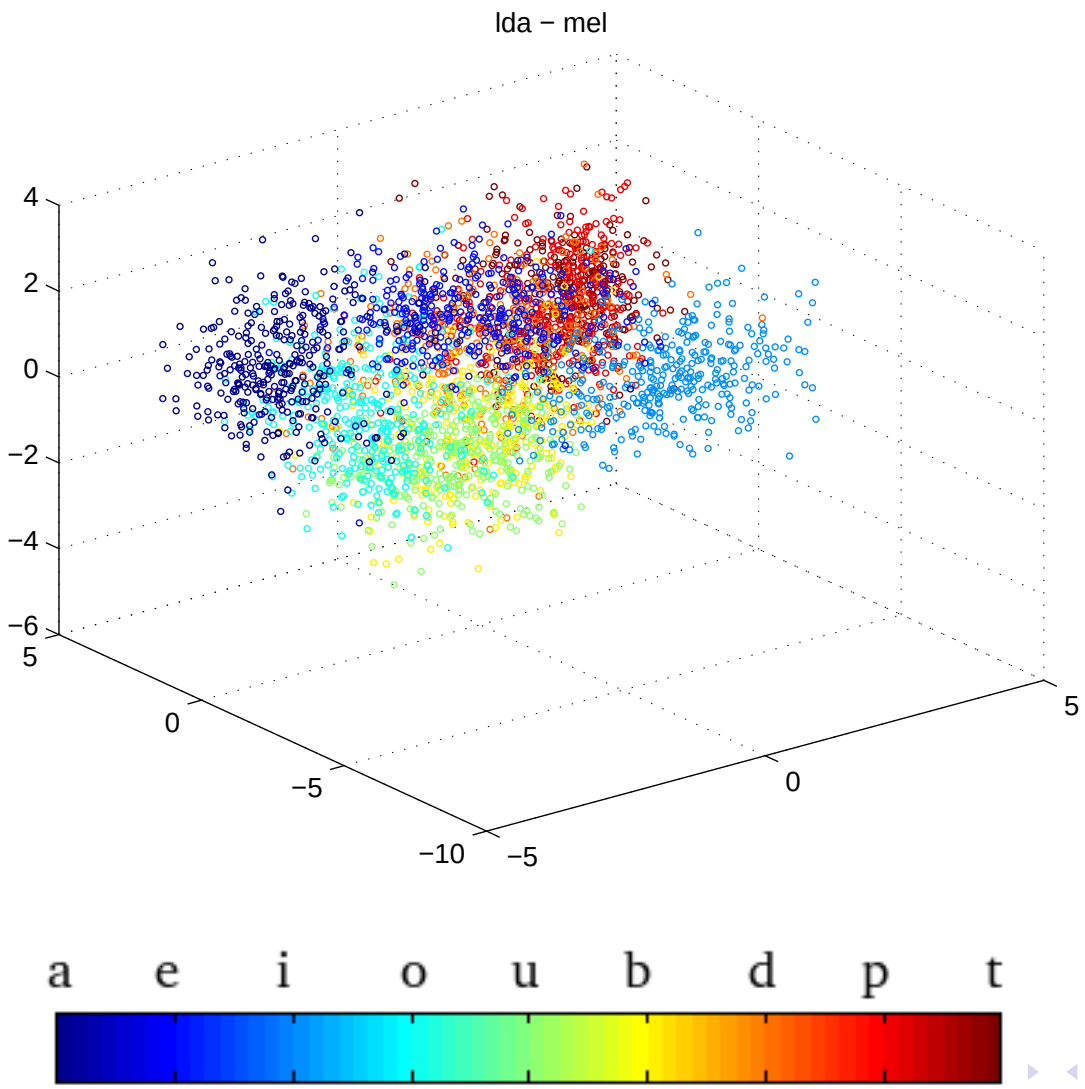
Optimized filter bank



Projected data: optimized filter bank



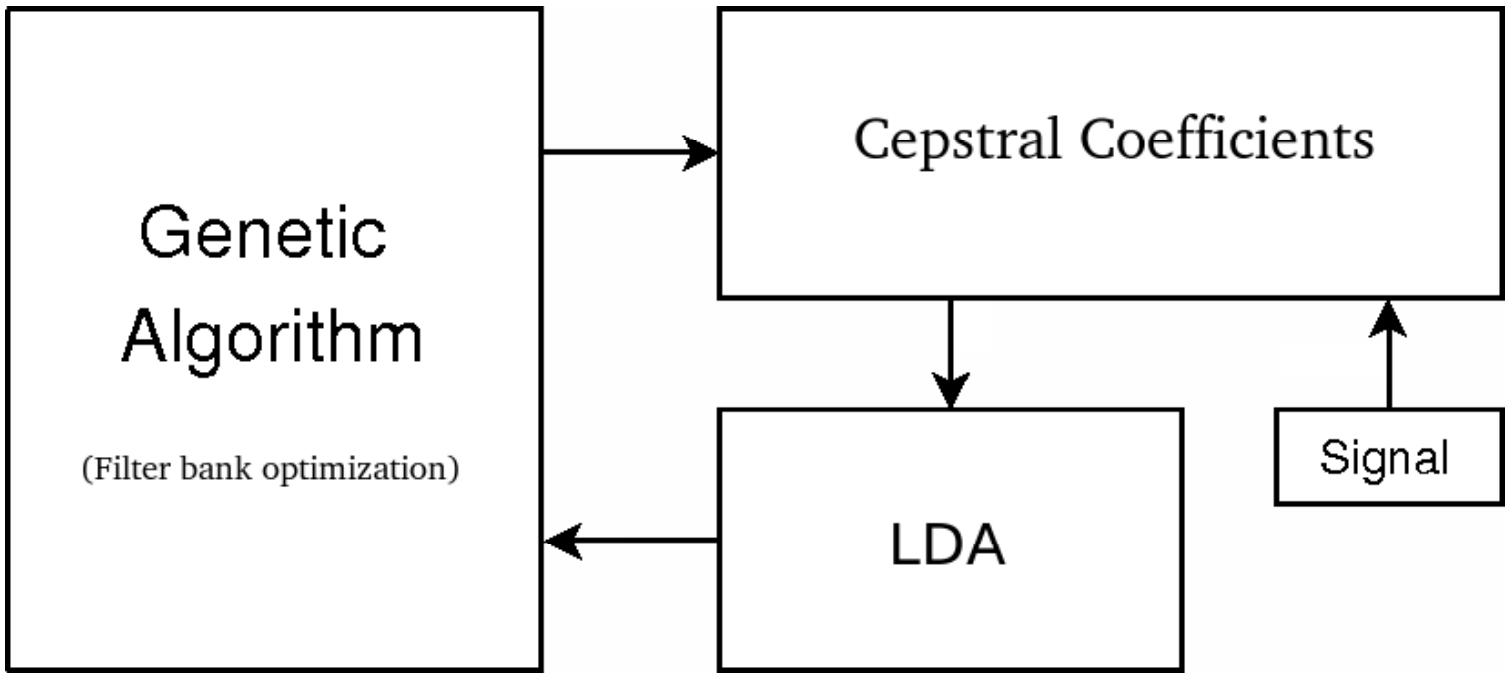
Projected data: mel scaled filter bank



Classification results

Phoneme	MFCC	GA-FCC	**LDA-MFCC**	**LDA-GA-FCC**
/a/	87.35	89.39	80.88	75.05
/e/	65.71	77.35	85.29	77.94
/i/	92.25	89.39	91.18	89.71
/o/	26.33	31.43	47.06	58.73
/u/	60.00	46.94	82.35	58.82
/b/	40.81	43.06	29.41	41.18
/d/	07.96	08.16	23.53	25.00
/p/	60.20	67.55	57.35	63.24
/t/	53.67	54.28	47.06	54.41
Average	53.94	55.39	**60.11**	**60.00**

LDA as fitness in GA-FCC



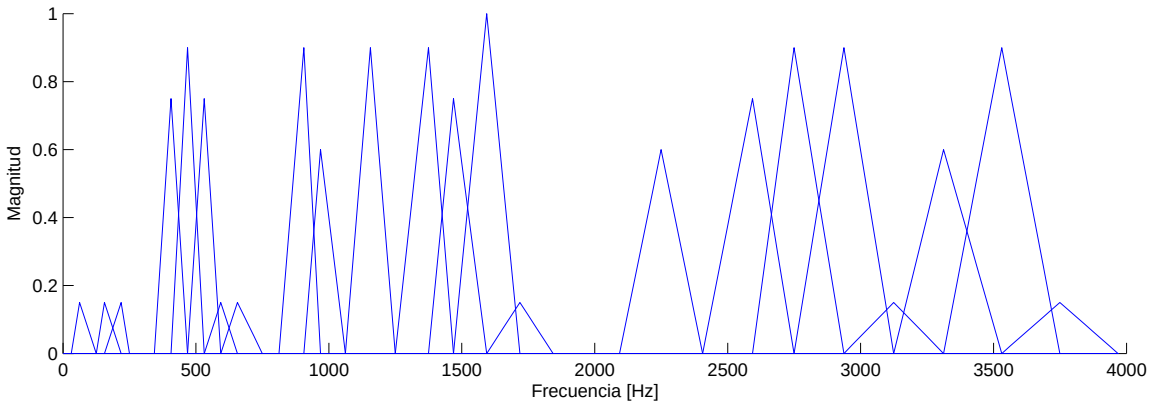
LDA as fitness in GA-FCC

As fitness function we use Fisher's criterion defined by

$$J(\mathbf{w}) = \mathrm{Tr}\left\{ \frac{(\mathbf{W}\mathbf{S}_B\mathbf{W}^T)}{(\mathbf{W}\mathbf{S}_W\mathbf{W}^T)} \right\},$$

where $\mathbf{S}_W$ and $\mathbf{S}_B$ are the *within-class* and *between-class* covariance matrices respectively.

LDA as fitness in GA-FCC



LDA as fitness in GA-FCC: classification results

Phoneme	MFCC	GA-FCC	Fitness LDA
/a/	87.35	89.39	85.51
/e/	65.71	77.35	65.92
/i/	92.25	89.39	84.49
/o/	26.33	31.43	41.84
/u/	60.00	46.94	44.90
/b/	40.81	43.06	34.08
/d/	07.96	08.16	09.39
/p/	60.20	67.55	62.04
/t/	53.67	54.28	53.88
Average	53.94	55.39	52.61