

# Choosing an Accurate Number of Mel Frequency Cepstral Coefficients for Audio Classification Purpose



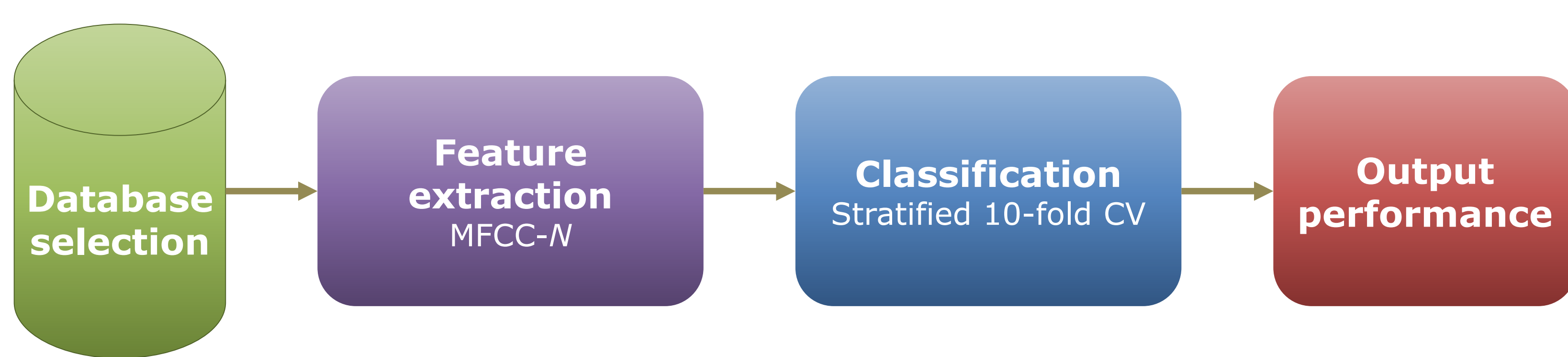
Lăcrimioara GRAMA, Corneliu RUSU

Signal Processing Group  
Faculty of Electronics, Telecommunications and Information Technology  
Technical University of Cluj-Napoca

## Abstract

In this paper, we study several audio classification schemes applied on different number of features for multiclass classification with imbalanced datasets. As features, we proposed the liftering Mel frequency cepstral coefficients, while for classification we use probabilistic methods, instance-based learning algorithms, support vector machines, neural networks,  $L^\infty$ -norm based classifier, fuzzy lattice reasoning classifier, and trees. The final goal is to find the appropriate number of liftering Mel frequency cepstral coefficients to provide the desired accuracy for audio classification purpose.

## Audio classification system



## Database

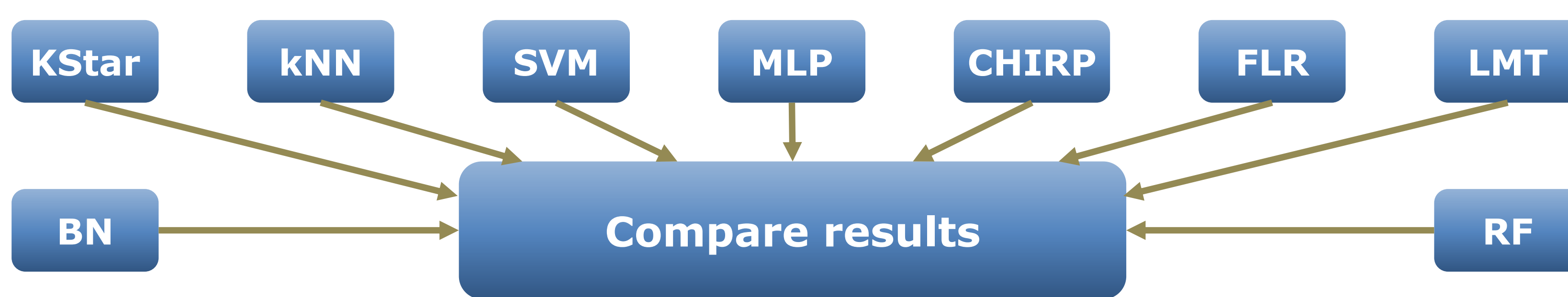
Bird	654 (40.95%)	16 kHz, 16-bit None of the audio signals are studio recordings ⇒ they are subject to some additive noise from surroundings
Chainsaw	356 (22.29%)	
Gunshot	120 (7.51%)	
Human voice	207 (12.96%)	
Tractor	260 (16.28%)	

## Feature extraction

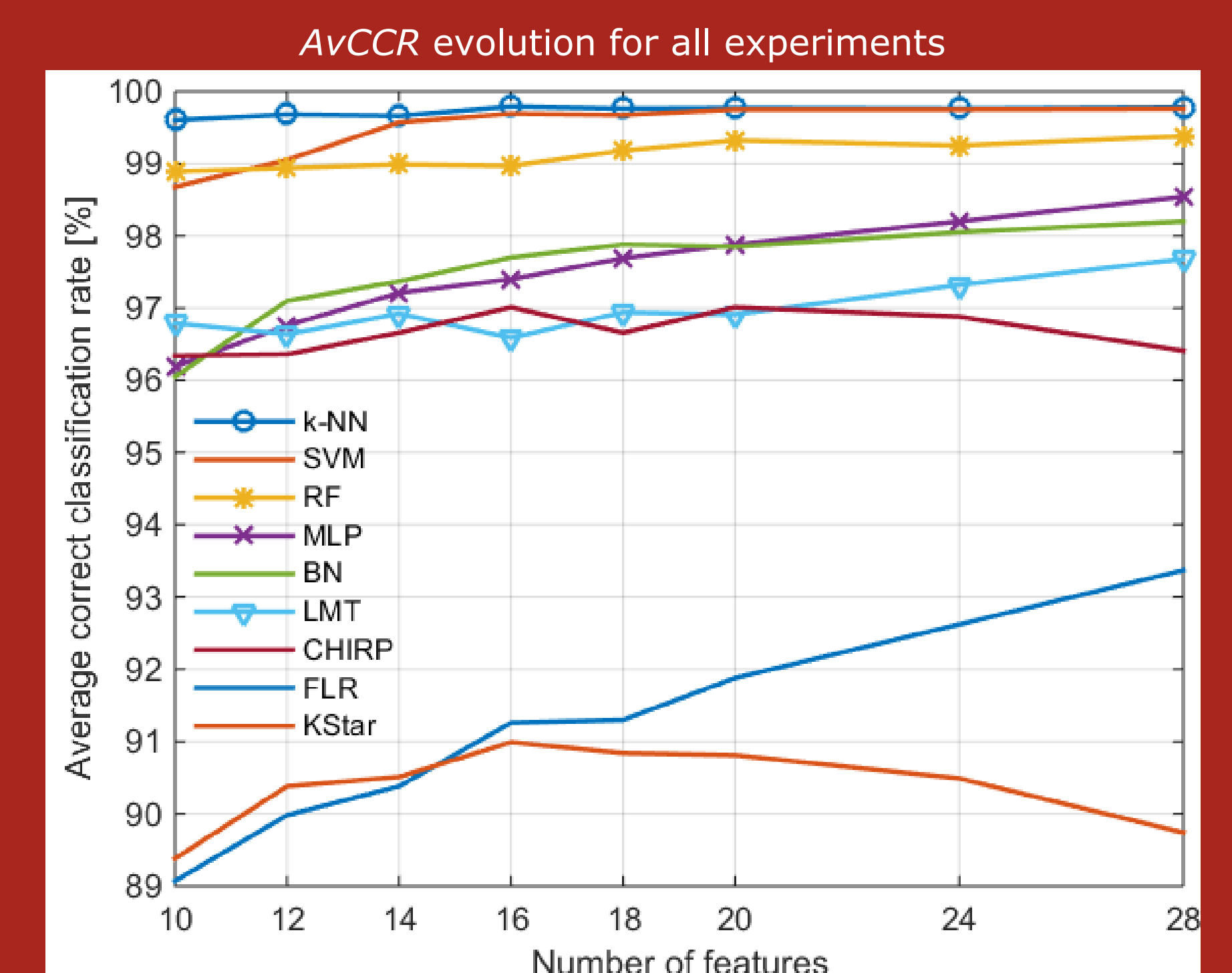
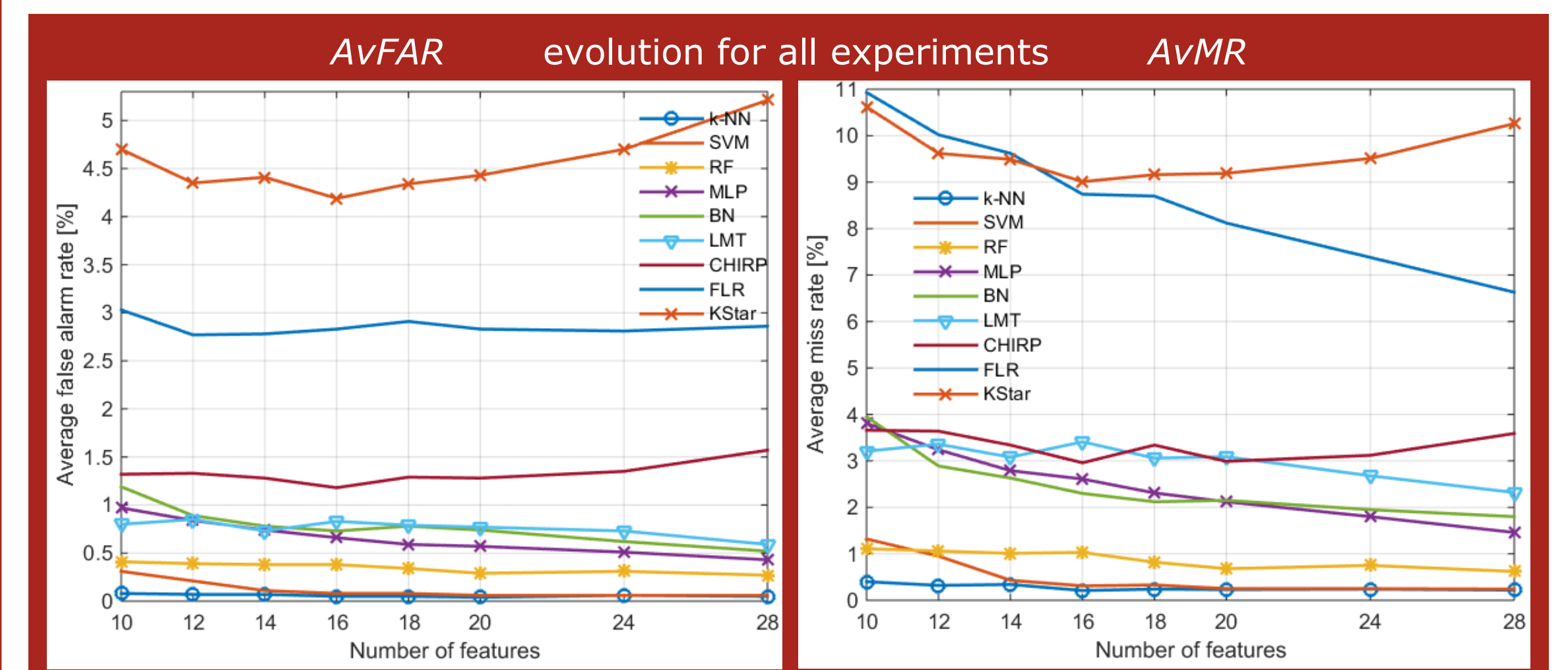
MFCC- $N$   
 $N = 10, 12, 14, 16, 18, 20, 24, 28$   
Frames: 25ms, 60% overlap  
Filterbank: 0÷8 kHz, 30 triang. filters

$$\mathbf{F}_{1597 \times N} = \begin{bmatrix} c'_{1,1} & c'_{1,2} & \dots & c'_{1,N} \\ c'_{2,1} & c'_{2,2} & \dots & c'_{2,N} \\ \vdots & \vdots & \ddots & \vdots \\ c'_{1597,1} & c'_{1597,2} & \dots & c'_{1597,N} \end{bmatrix}$$

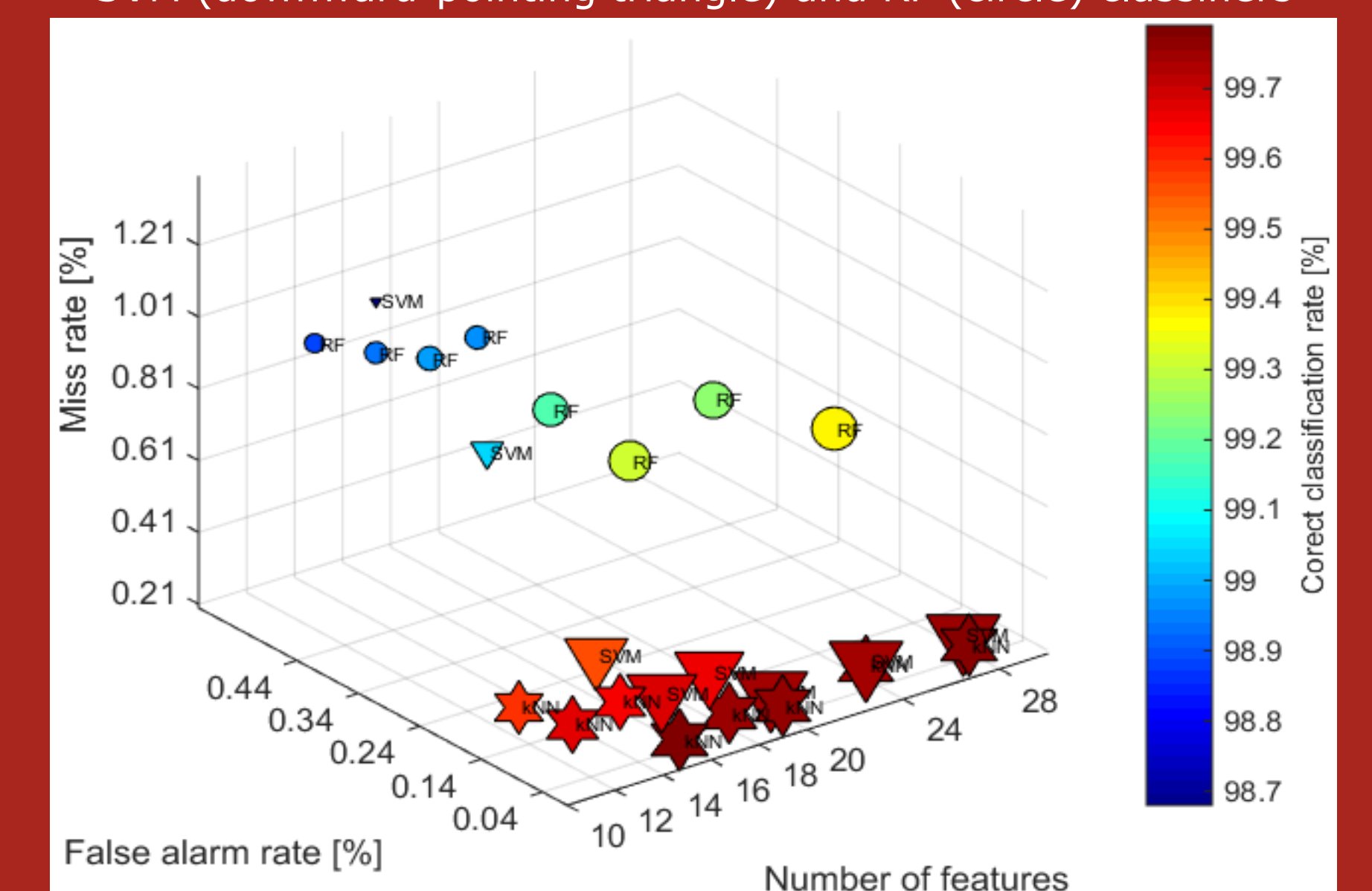
## Classification



## Output performance



(Number of features, AvFAR, AvMR, AvCCR) for  $k$ -NN (hexagram), SVM (downward-pointing triangle) and RF (circle) classifiers



## Conclusion

Regardless the number of features, the correct classification rate, for  $k$ -NN classifier is higher than 99.60, for SVM is higher than 98.68%, and for RF is higher than 98.89%.  
The best audio classification scheme is obtained with MFCC-16 and  $k$ -NN. In this case the correct classification rate is 99.79%, the false alarm rate is 0.05%, the miss rate is 0.21%, the precision is 99.80% and the  $F$ -measure is 99.79%.

## Acknowledgment

This work was supported by a grant of the Romanian National Authority for Scientific Research and Innovation, CNCS/CCCDI-UEFISCDI, project number PNIII-P2-2.1-BG-2016-0378, 54BG/2016, within PNCDI III.