

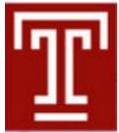
Utilizing Temporal Patterns for Estimating Uncertainty in Interpretable Early Decision Making

Mohamed Ghalwash,
Temple University
Philadelphia, PA. USA

Vladan Radosavljevic,
Yahoo Labs
Sunnyvale, CA. USA

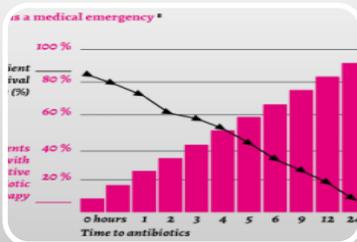
and Zoran Obradovic
Temple University
Philadelphia, PA. USA





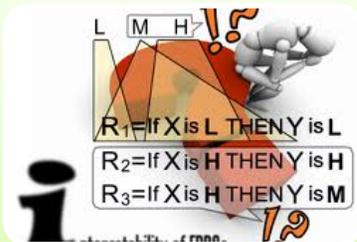
Motivation and Objectives

Although predictive accuracy is clearly an important objective in time series classification, additional criteria that are often considered include:



Earliness

- Timely diagnosis is essential for allowing physicians to design appropriate therapeutic strategies at early stages of diseases
- Early therapies are usually the most effective and the least costly



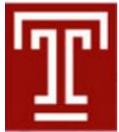
Interpretability

- It is important to attain decisions that are not only accurate and obtained early, but can also be easily interpreted
- Physicians prefer interpretable methods rather than more accurate black-box methods.



Uncertainty Estimation

- Uncertainty associated with a given prediction is important in clinical diagnosis
- Data mining methods assist clinical experts in making decisions and optimizing therapy



Motivation and Objectives

Existing methods are limited to addressing at most two of these aspects

1. Interpretability (I)
2. Earliness (E)
3. Uncertainty estimation (U)

	Borgwardt et al 2006	Lin et al 2008	Ye and Keogh 2009	Xing et al 2010	Xing et al 2011	Wiens et al 2012	Parrish et al 2014
I	X	X	✓	X	✓	X	X
E	X	X	X	✓	✓	✓	✓
U	X	X	X	X	X	✓	✓

Our objective is to develop a simple but effective method that satisfies all three aspects.

This is achieved by extending a recently proposed interpretable early classification method, called *early distinctive shapelet classification* (EDSC) [Xing et al., 2011], to **estimate the temporal uncertainty associated with the prediction.**



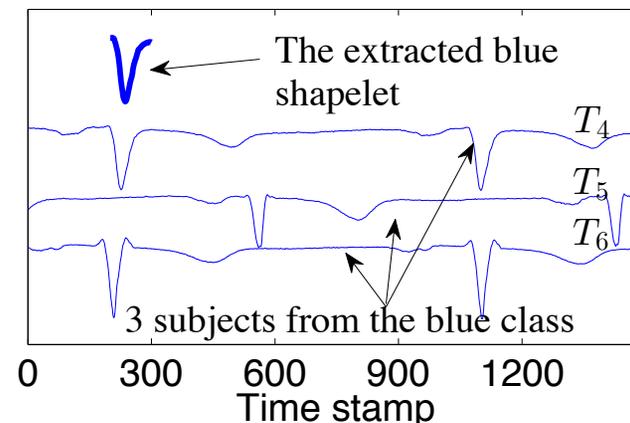
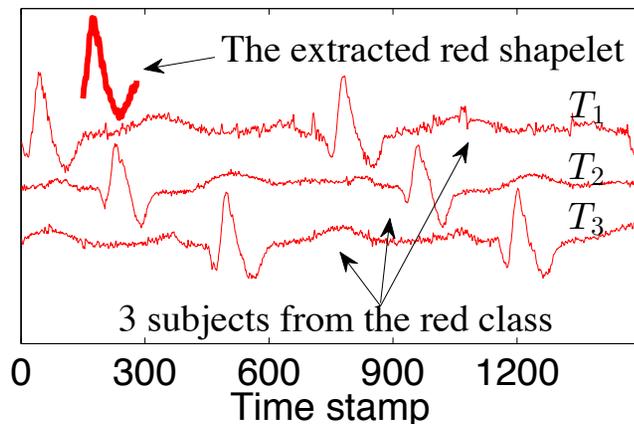
EDSC Framework [Xing et al, 2011]

Training:

Extract patterns (shapelets) that provides **example-specific** classification of unknown time series as **early** as possible.

- 1. Extraction:** Extract all shapelets of different lengths, where for each shapelet a distance threshold is learned such that the shapelet discriminates between classes,
- 2. Ranking:** Rank the shapelets using a utility function that incorporates earliness and accuracy of the shapelet,
- 3. Pruning:** Prunes the shapelets by selecting the top shapelets that cover the entire dataset.

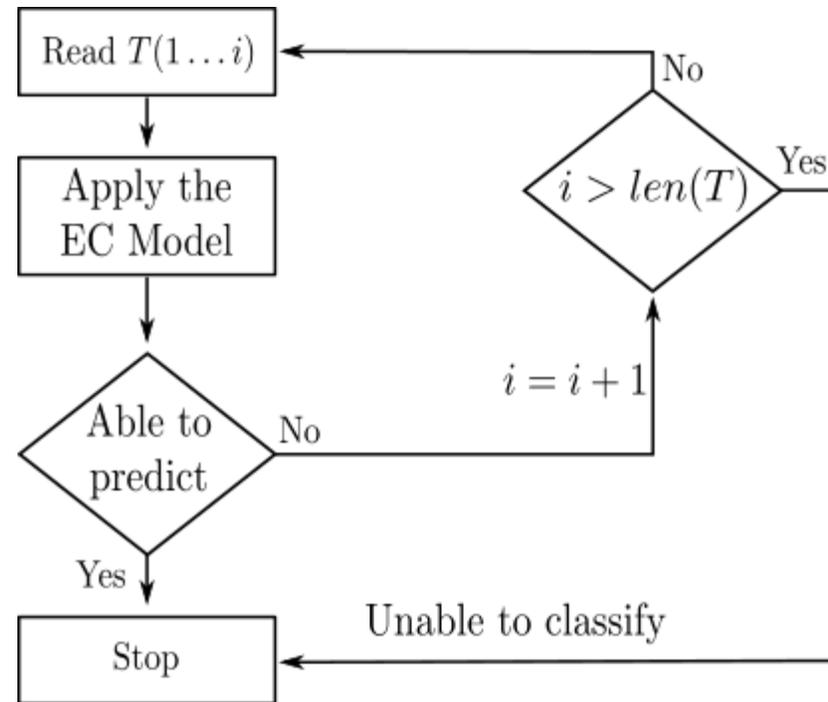
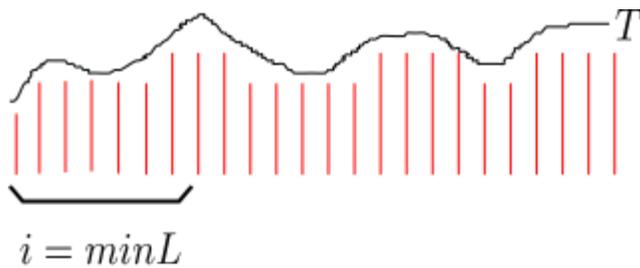
Shapelets are subsequences of time series, and thus highly interpretable, that have been used as features representing the characteristics of the time series.

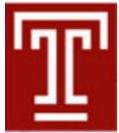




EDSC Framework [Xing et al, 2011]

Classification: classifies unknown time series based on the closest matching shapelet.





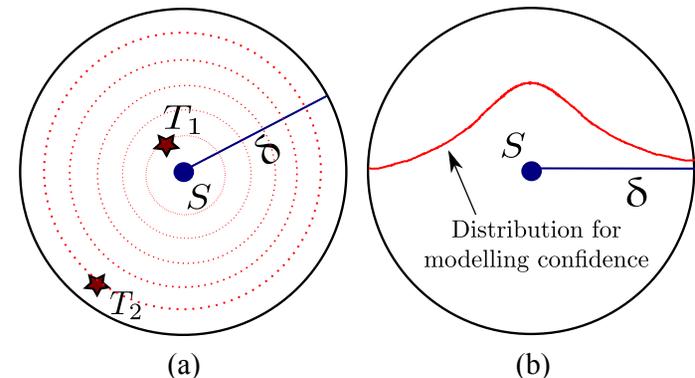
Motivating Example

Assume that we have a shapelet $S = (s, l, \delta, c)$ where

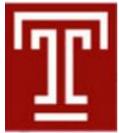
- s is subsequence of the time series
- l is the length of the shapelet
- δ is the distance threshold (computed by EDSC)
- c is the class label of the shapelet

The shapelet S is represented as a point and the radius of the circle around S represents the shapelet's distance threshold δ .

T_1 and T_2 are two time series that are less than δ -distant apart from the shapelet.



The distance between T_1 and S is less than the distance between T_2 and S which reflects the fact that the shapelet is more certain about the classification of T_1 than the classification of T_2 .



Uncertainty Estimation for Shapelet-Based Methods

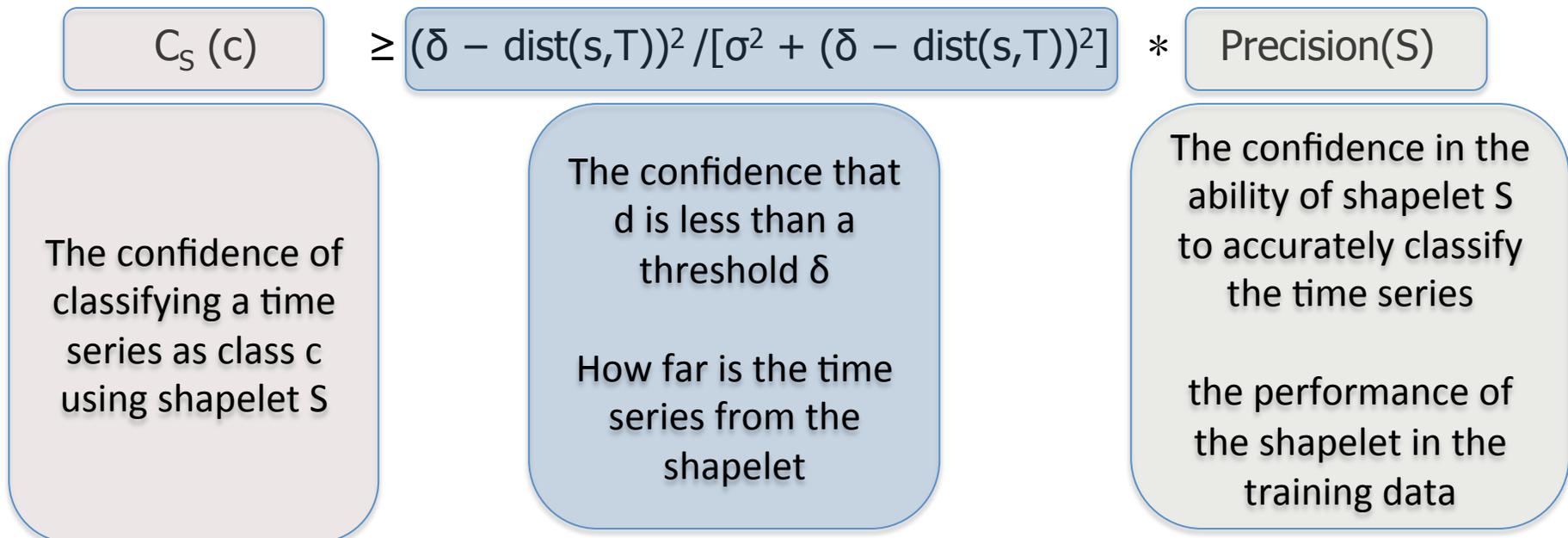
We model the confidence $C(c)$ of classifying a time series as class c . The uncertainty $U(c)$ of classifying a time series as class c can be computed as

$$U(c) = 1 - C(c)$$

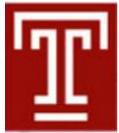
We define the distance between T and S as a random variable d

$$d = \text{dist}(s, T) + \varepsilon$$

where ε is some random variable with mean equal to 0 and standard deviation equal to σ



Since both terms in this product take value between 0 and 1, the highest value of $C_S(c)$ is 1.



Aggregated Class Confidence

Assume that $S^c = \{S_1, S_2, \dots, S_N\}$ is the set of all shapelets from class c that **match** the current time series.

The confidence for classifying the time series as class c is increased by having **multiple matched** shapelets as computed by

$$\begin{aligned} C(c) &= C_{S^c}(c) \\ &= C_{S_1 \cup S_2 \cup \dots \cup S_N}(c) \\ &= \sum_{k=1}^N (-1)^{k+1} \sum_{\substack{I \subset \{1, 2, \dots, N\} \\ |I|=k}} C_{S_I}(c) \end{aligned}$$

The class confidence $C(c)$ satisfies all properties of the confidence measure, i.e. **takes values on the range $[0, 1]$ and has value higher than any individual shapelet.**



Modified EDSC with Uncertainty Estimates (MEDSC-U)

Pruning:

The MEDSC-U method sorts the shapelets descending based on their utility scores and starts with the highest ranked shapelet S .

The method removes all time series from the dataset that are covered by the shapelet and stores the shapelet S and **all other shapelets** that have the same utility score as S .

Classification:

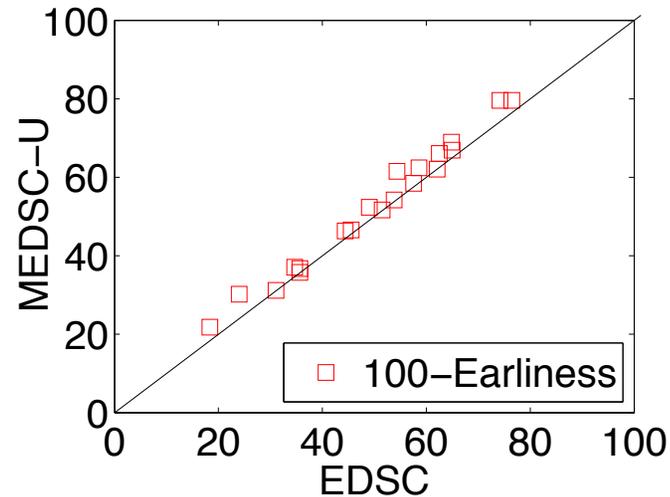
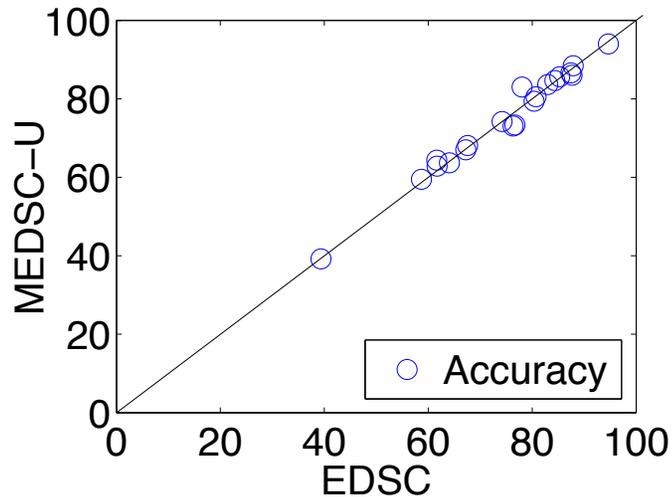
We compute the distance between the current stream of the time series and ***all*** discriminative shapelets extracted by MEDSC-U (we do not start with the highest one until a match is found as EDSC).

Then we compute the uncertainty for each class based on ***all matched*** shapelets from that class.



Accuracy and Earliness Characterization of MEDCS-U

- MEDSC-U includes more equal-performance shapelets than EDSC in order to have reliable uncertainty estimates
- The effect of including more shapelets on the accuracy and earliness of the MEDSC-U method without considering uncertainty estimates vs. EDSC is considered on 20 time series datasets from the UCR time series archive [Keogh et al, 2011].

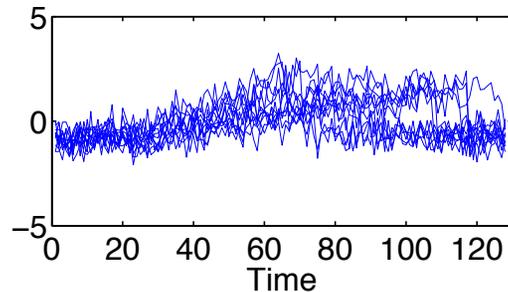


- Including equal-performance shapelets
 - does not negatively effect the accuracy of the EDSC method
 - slightly improved the earliness of the classification decision as this allows for variability among time series' patterns.

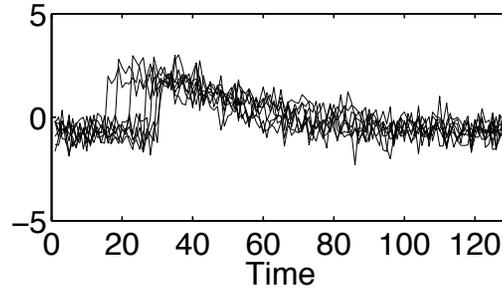


Case Study 1: Classification Based on a Confident Shapelet

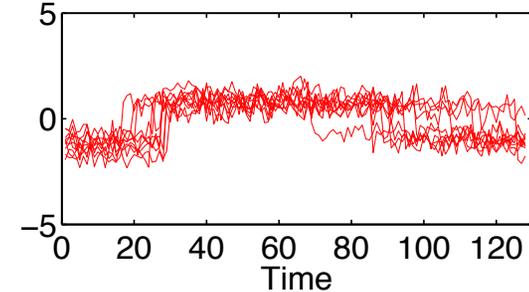
CBF dataset has 3 classes where black and red classes are similar to each other especially at the beginning of the time series



(a) Bell



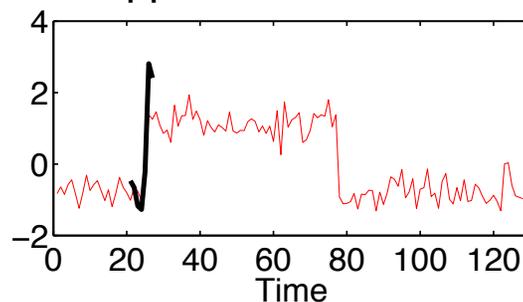
(b) Funnel



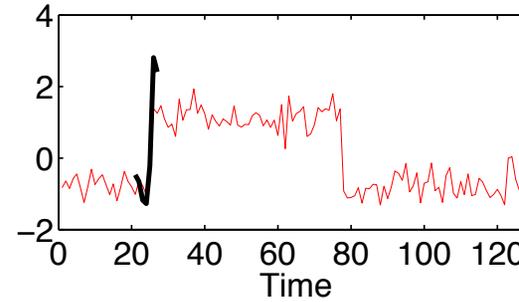
(c) Cylinder

The EDSC method

- has extracted one shapelet from each class and achieved 88% accuracy.
- incorrectly classified the cylinder (red) example as funnel (black) example.
- The same error happened if we used the MEDSC-U with uncertainty 1.



(d) EDSC



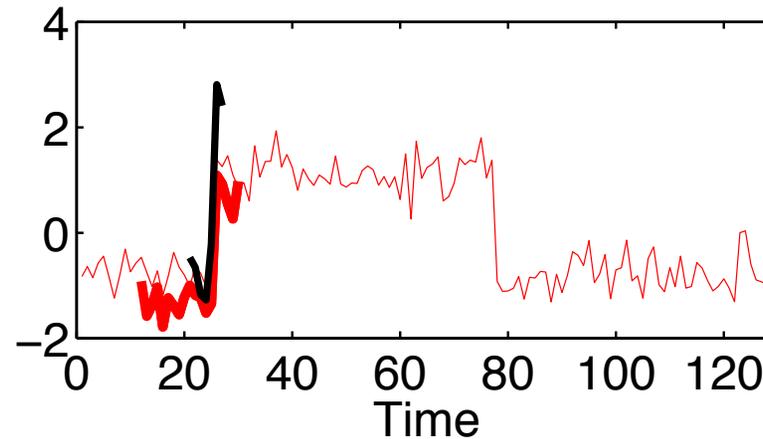
(e) MEDSC-U with uncertainty 1

To overcome this problem we measure the uncertainty for the classification decision, *especially to distinguish red from black classes.*



Case Study 1: Classification Based on a Confident Shapelet

The uncertainty associated with the classification using the black shapelet at time point 27 is 0.49.



0.49 is not satisfactory

wait and not provide classification at this point

A cylinder (red) shapelet matches the time series with uncertainty 0.13 at time 30

because we know that these two classes are similar to each other

So MEDSC-U is more confident to *'correctly'* classify the time series as a cylinder (red) class than to classify it as a funnel (black) class.



Recommending an Uncertainty Threshold

Problem

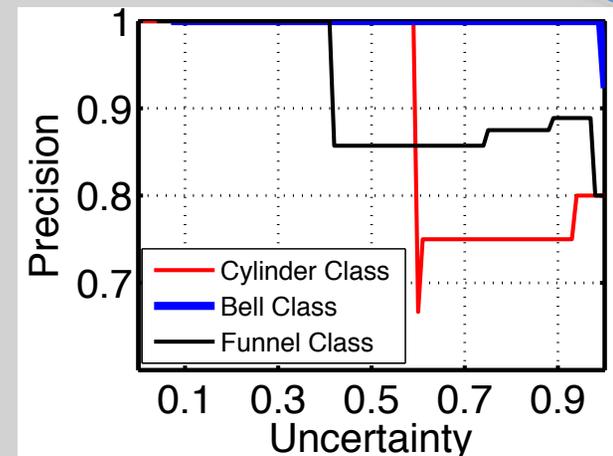
The domain expert might have no notion about the recommended uncertainty threshold for each class.

Solution

- ❑ We compute the precision of each class at different values of uncertainty thresholds.
- ❑ Based on a desired value of precision, the user can choose the corresponding uncertainty threshold.

Example

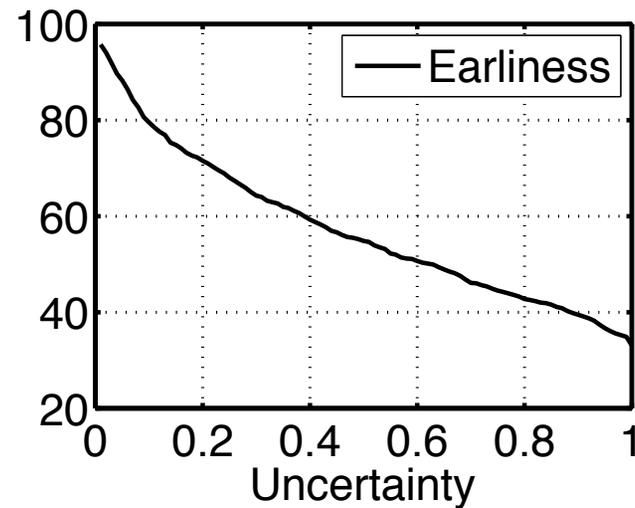
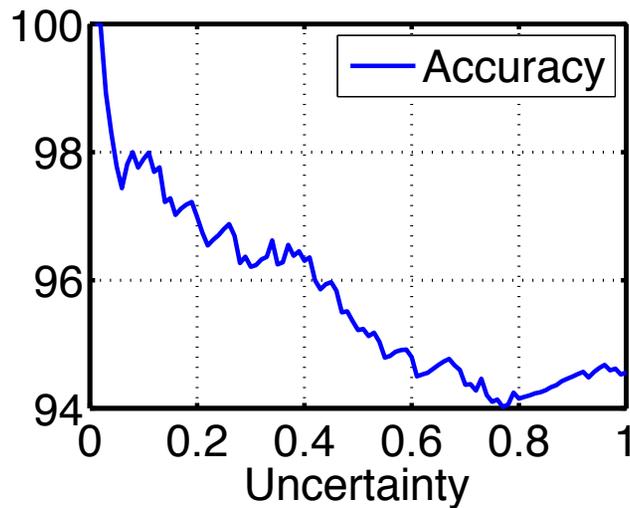
1. The blue class has 100% precision at each level of uncertainty while the precision of the other two classes drops earlier
 - ✓ The blue class is dissimilar to the other classes
 - ✓ The blue class can be recognized by our method even with high value of uncertainty such as 0.9.
2. The precision for the red (Cylinder) and black (Funnel) classes dropped at approximately 0.6 and 0.4 which would be a good estimate for the uncertainty thresholds, respectively.



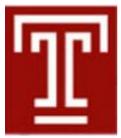


Uncertainty versus Accuracy and Earliness

The accuracy and the earliness of MEDSC-U for different uncertainty thresholds on CBF dataset.



Lower value of uncertainty threshold gives more accurate results but delays the decision

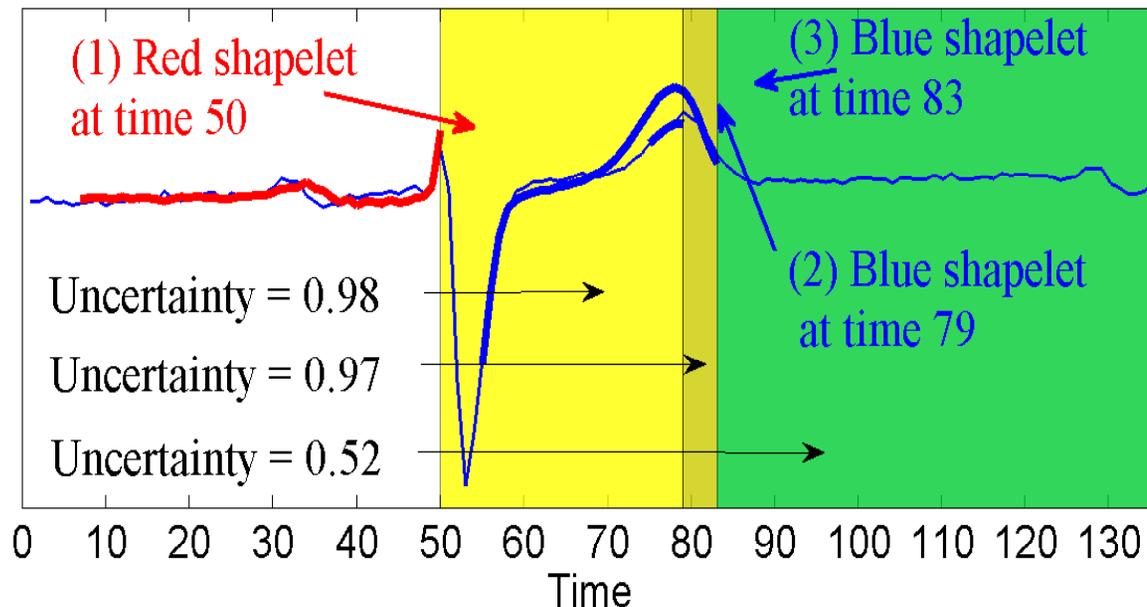


Case Study 2:

Classification based on Multiple Matched Shapelets

ECGFiveDays Case Study:

- (1) At time point 50 the time series from the blue class is matched by a red shapelet with uncertainty 0.98. The prediction is postponed due to high uncertainty.
- (2) at the time point 79 a blue shapelet matches the time series with uncertainty 0.97. The prediction is undecided due to similar uncertainties for red and blue classes.
- (3) At time point 83, a new blue shapelet matches the time series. The uncertainties from the two blue shapelets are aggregated resulting in uncertainty 0.52 for the blue class which is the correct class.

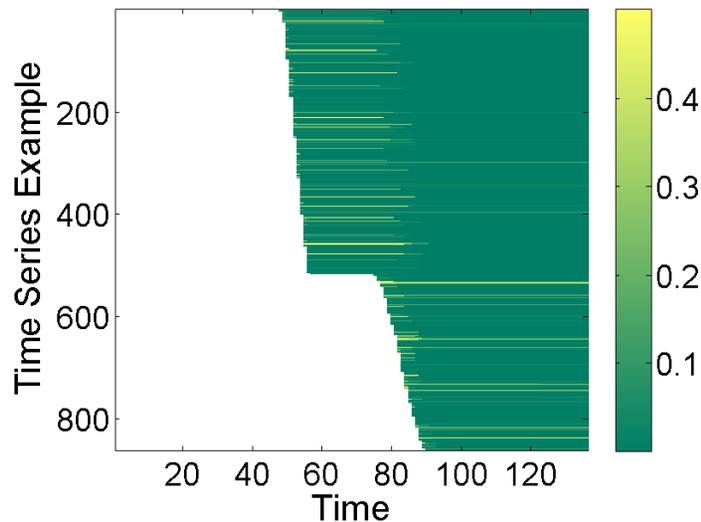
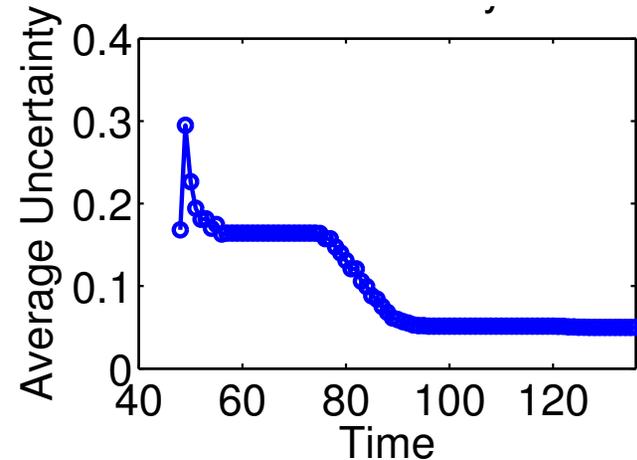




Case Study 2:

Classification based on Multiple Matched Shapelets

- The average uncertainty over all patients at each time point for the ECGFiveDays dataset.
- The method starts with high uncertainty and then becomes more certain about the classification as time evolves



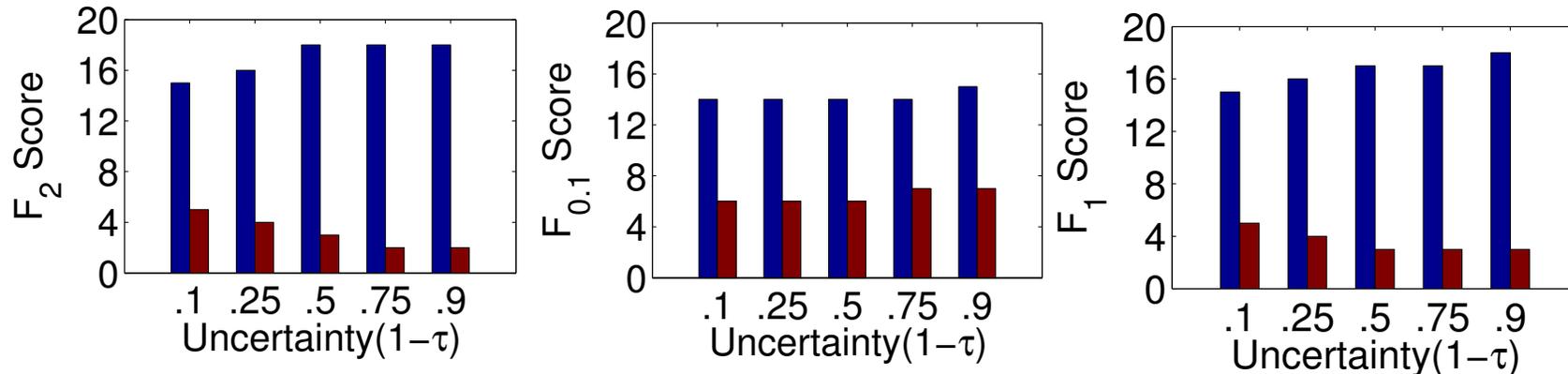
- The values of uncertainty over time for each patient from the ECGFiveDays dataset.
- The white bar indicates that there is no classification at that point and hence there is no uncertainty.
- The uncertainty for each patient reduces over time.
- The MEDSC-U method becomes more certain about the classification between time 82 - 86



Uncertainty Estimates Comparison to localQDA

[Parrish et al., JMLR 2014]

- The local quadratic discriminant analysis (localQDA) provides a reliability bound τ on the classifier's decision for every time point (the probability that the early classification will be the same as the classification at the end of the time series).
- We assume that $\text{Uncertainty} = 1 - \tau$.
- F_β score is the weighted average of the accuracy and 100-earliness where
 - $\beta = 2$ weights earliness more than accuracy,
 - $\beta = 0.1$ puts more emphasis on accuracy than earliness,
 - $\beta = 1$ is the balanced harmonic mean.
- Evaluated on 20 time series datasets. Left, middle and right figure show number of datasets where MEDSC-U (blue bar) has better F_2 , $F_{0.1}$ and F_1 score than the localQDA (red bar), and vice versa, at increasing uncertainty thresholds.



- MEDSC-U outperformed localQDA in most cases especially in F_1 and F_2 settings where the earliness is important.



Conclusion and Work in Progress

- ❑ We extended the interpretable early classification method (EDSC) and proposed the MEDSC-U method to measure the temporal uncertainty with the classification.

- ❑ MEDSC-U
 - ✓ Provides early, interpretable predictions for time series classification with uncertainty estimate.
 - ✓ It was more effective than the state-of-the-art method in our experiments on twenty datasets
 - ✓ It is very simple to implement.

- ❑ The temporal uncertainty estimates provided by MEDSC-U can be extended to the multivariate case where uncertainties from shapelets from different variables could be integrated.

- ❑ In progress: Application for early diagnosis of sepsis and accurate detection of change in the state of sepsis on large animals

- ❑ Acknowledgement: **Funded by DARPA DLT grant N66001-11-1-4183**



Thanks

More details:

Zoran Obradovic: zoran.obradovic@temple.edu

Collaboration:

Open research positions in Zoran's lab:

**Ph.D. students,
postdoctoral associates, and
visiting scholars**