Adapting to Dynamically Changing Noise During Learning of Heart Sounds:

An AIS-based Approach Using Systemic Computation

Yiqi Deng
University College London
Gower Street, London
United Kingdom
+44(0)7721871959
y.deng.11@ucl.ac.uk

Peter J Bentley
University College London
Gower Street, London
United Kingdom
+44(0)7734102352
p.bentley@cs.ucl.ac.uk

ABSTRACT

Real world machine learning, where data is sampled continuously, may in theory be classifiable into distinct and unchanging categories but in practice the classification becomes non-trivial because the nature of the background noise continuously changes. Applying distinct and unchanging categories for data ignores the fact that for some applications where the categories of data may remain constant, the background noise constantly changes, and thus the ability for a supervised learning method to work is limited. In this work, we propose a novel method based on an Artificial Immune System (AIS) and implemented on a systemic computer, which is designed to adapt itself over continuous arrival of data to cope with changing patterns of noise without requirement for feedback, as a result of its own experience.

Categories and Subject Descriptors

D.3.2 [**Programming Languages**]: Language Classifications – *Concurrent, distributed, and parallel languages.* F.1.2 [**Theory of Computation**]: Modes of Computation – Online computation, Parallelism and concurrency.

General Terms

Algorithms, Design, Experimentation, Languages, Theory.

Keywords

Artificial Immune Systems, Systemic Computer.

1. INTRODUCTION

In machine learning where there are distinct and unchanging categories for the data it would be normal to use a standard supervised learning approach. However, for problems such as the classification of heart sounds, although there may in theory be a limited number of specific categories of heart sound, in practice new sounds are gathered within such varied environments that the changing background noise can obscure or corrupt the features. For this application, a classifier trained on a static training set may be less effective than a method that can continuously learn and adapt. In this work we focus on heart sounds. According to the World Health Organization, cardiovascular diseases (CVDs) are the number one cause of death globally and primary care physicians are documented to have poor auscultatory skills [1]. Any method which can help to detect signs of heart disease could therefore have a significant impact on world health.

GECCO'14, July 12–16, 2014, Vancouver, BC, Canada.
ACM 978-1-4503-2881-4/14/07.
http://dx.doi.org/10.1145/2598394.2598461

2. Methodology

Le Martelot provided a simple implementation of "artificial tissue" for AIS using systemic computation [2,3], which serves as the foundation of our work. He created an artificial organism, a program with metabolism that "eats" data, expels waste, clusters cells based on the nature of its food and emits danger signals suitable for an artificial immune system. The whole ecosystem is visualized in Figure 1.

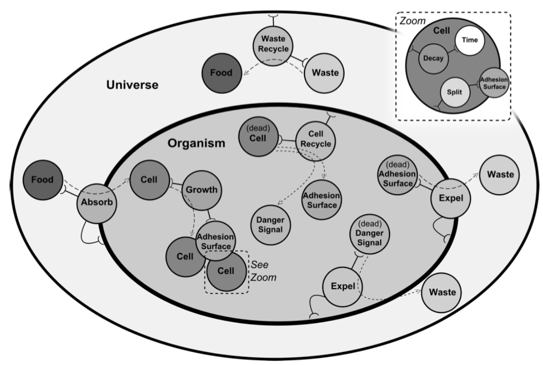

Figure 1 Systemic organization of Le Martelot's organism [3].

The immune system is a complex of cells, molecules and organs aiming at limiting damage to the host organism by pathogens, which elicit an immune response. We based on our model on [3] to implement the full function of AIS with antigens being "eaten" continuously into the organism and inspired the secretion of antibodies which serves as the classifier to recognize later-come antigens. More specifically, data from the same category is regarded as sharing a same/similar feature which in the scenario of AIS would correspond to antigen. Those data with same/similar antigen would then be recognized by memory antibodies trained previously through clonal selection.

As shown in Fig. 2, the whole immune system is organized in the scope of Universe which is an abstraction of the immune system. Inside it, B cell, antigens, antibody memories, clonal antibodies and antibody candidates (Wastes in the figure) are all represented by data systems, but with the bit 15 to bit 17 of their functions coded differently (see part of them in Fig 3). Among them, B cells are initialized with training data and antibody memories are assigned with random value at the beginning of the program. Meanwhile, antigens carrying test data 'float' outside B cells. The

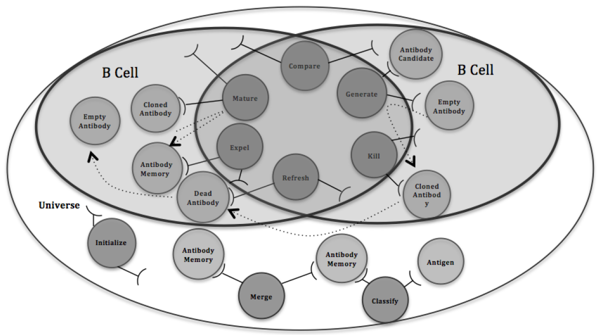

Figure 2 Systemic Organization of our model.

```
#label waste              %b0000000000000000?????????000000
#label antigen            %b000000000000000100????????000000
#label clonecandidate     %b00000000000000011?????????000000
#label bodymemory         %b000000000000000111????????000000
```

Figure 3 Representation of different data system in SC

organism stimulates an adaptive immune response across different B cells simultaneously where a spectrum of clones of antibody candidates is produced with different affinity. Antibody memories, those with the highest affinity, are expelled to the outside environment to recognize antigens (classify data). Meanwhile, expelled antibody memories keep searching and merging with antibody memories who are within the threshold distances of them. To facilitate the process above, 9 contexts (shown as the darkest circles in Fig. 3) are implemented in side of which different functions are specified.

3. Experiments

To test the performance of the AIS-SC model to classify data despite changing background noise, we select 20 test cases with normal heartbeat and another 20 with murmur heartbeat from the PASCAL Challenge on Classifying Heart Sounds. The dataset comprises data collected using the digital stethoscope DigiScope [4]. For each of them, four clips of heartbeat are recorded in a different time and location. We reserve one sound per person for training and use the remaining 3 as test data. As a comparison, we also used the Multiple Kernel Anomaly Detection (MKAD) Algorithm [5]. We here set the number of interactions as 3000000 as it is the lowest number to guarantee the complete recognition of all the test data. Then we calculate the average result of 20 rounds of experiments on SC and MKAD. Table 1 and Fig 4 summarise the results.

Table 1 Classification by AIS-SC and MKAD.

	AIS-SC	MKAD
Precision of Normal	78.67%	81.29%
Precision of Murmur	65.72%	63.12%
Recall Rate of Normal	77.25%	82.37%
Recall Rate of Murmur	66.97%	63.01%

4. CONCLUSIONS

In this paper, we propose a novel method based on an AIS and implemented on a systemic computer, which is designed to adapt itself over continuous arrival of data to cope with changing patterns of noise without requirement for feedback, as a result of its own experience. Experiments on the heartbeat data classification shows the algorithm performs up to 3.60% better in the precision rate of murmur and 3.96% better in the recall rate of murmur than another standard anomaly detector. Results also show that with more data coming in, the model will better adapt itself to the heartbeat sound as well as the background noise.

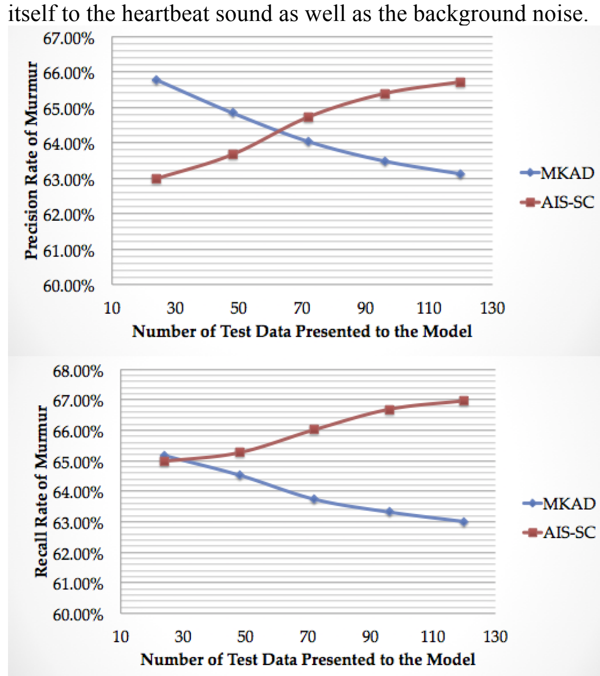

Figure 4 Precision Rate and Recall Rate of Murmur Against Number of Tested Data.

5. REFERENCES

[1] Choi, Samjin, and Zhongwei Jiang. Comparison of envelope extraction algorithms for cardiac sound signal segmentation. *Expert Systems with Applications 34.2* (2008): 1056-1069.

[2] Bentley, Peter J. Systemic computation: A model of interacting systems with natural characteristics. *The International Journal of Parallel, Emergent and Distributed Systems* 22.2 (2007): 103-121.

[3] Le Martelot, Erwan, Peter J. Bentley, and R. Beau Lotto. Eating data is good for your immune system: an artificial metabolism for data clustering using systemic computation. *Artificial Immune Systems*. Springer Berlin Heidelberg, 2008. 412-423.

[4] Bentley, P. and Nordehn, G. and Coimbra, M. and Mannor, S.(2011). The Pascal Classifying Heart Sounds Challenge. http://www.peterjbentley.com/heartchallenge/index.html

[5] Das, S., Matthews, B. L., Srivastava, A. N., & Oza, N. C. (2010, July). Multiple kernel learning for heterogeneous anomaly detection: algorithm and aviation safety case study. In Proceedings of the 16th ACM SIGKDD international conference on Knowledge discovery and data mining (pp. 47-56). ACM