CHAPTER 9

NONSTATIONARY STREAM DATA LEARNING WITH IMBALANCED CLASS DISTRIBUTION

Sheng Chen* and Haibo He†

* Merrill Lynch, Bank of America; † University of Rhode Island.

Abstract

Please include an Abstract for your chapter. Here is just an example...

The ubiquitous imbalanced class distribution occurred in real world data sets has excited considerable interests in the study of imbalanced learning. However it is still a relatively uncharted area when it is nonstationary data stream with imbalanced class distribution that needs to be processed. Difficulties in this case are generally two-fold. First, dynamically structured learning framework is required to catch up with the evolution of unstable class concepts, i.e., concept drifts. Second, imbalanced class distribution over data stream demands a mechanism to intensify the underrepresented class concepts for improved overall performance. For instance, in order to design an

intelligent spam filtering system, one needs to not only make the system to be capable of self-tuning its learning parameters to keep pace with the rapid evolution of spam mail patterns, but also the system has to tackle the fundamental problem that in some situations normal emails are severely outnumbered by spam emails, but it is so much more expensive to misclassify a normal email into spam, e.g., confirmation of a business contract, than the other way around. This chapter introduces learning algorithms that were specifically proposed to tackle the problem of learning from nonstationary data set with imbalanced class distribution. System-level principle and framework of these methods are described in algorithmic level, the soundness of which are furthered validated through theoretical analysis as well as simulations on both synthetic and real-world benchmarks with varied level of imbalanced ratio and noise level.

9.1 INTRODUCTION

Learning from data stream has been featured in many practical applications such as network traffic monitoring and credit fraud identification [1]...

This is just an example – please add all your contents here...

9.2 SAMPLING APPROACHES

9.2.1 under-sampling

9.3 ACKNOWLEDGEMENT

For any acknowledgement (such as grants, etc.), please add them here...

REFERENCES

1. S. Grossberg, *Neural Networks and Natural Intelligence.* Cambridge, MA: MIT Press, 1988.