Machine Learning in Advertising Technology

Vangie Shue

December 9, 2014

Abstract

This project will explore the application of Machine Learning in Technology. Specifically, we will seek to develop a reliable process for developing a classification algorithm on given data.

Introduction

Background on Advertising Technology use cases. Background on Cafemom and the specific problem: classify hispanic users

Classifying Rare Events

(1, pg141) Rare events are more statistically informative that zeros is seen in the variance matrix

Methods

Demdex has Traits/Segments, uuid and traits/segments collected Table generating en/es language to determine hispanic/non-hispanic, 1 for es, 0 for en

- segment_hispanic: 80324851 total uuids

- segment_hispanic: 1698878 hispanic (2 percent)

- segment_hispanic2: 9733751 total uuids

- segment_hispanic2: 1868091 hispanic (19 percent)

```
> data(example)
```

The above is a snippet of code used.

Results

The below is a sample graph of data.

Something like a plot centered.

Conclusion

We demonstrated the application of Machine Learning in Advertising Technology, in particular for rare events.

References

We used `RStudio Sweave` to build this LaTeX document

- http://gking.harvard.edu/files/gking/files/0s.pdf (rare events)