ONLINE MASTERS IN **DATA SCIENCE**

DSC 257R - UNSUPERVISED LEARNING

INTRODUCTION

SANJOY DASGUPTA, PROFESSOR

UC San Diego

COMPUTER SCIENCE & ENGINEERING

HALICIOĞLU DATA SCIENCE INSTITUTE

Where Machine Learning Shines

Supervised learning with:

- highly specific tasks
- large labeled data sets

Where Machine Learning Shines

Our focus so far -

Supervised learning with:

- highly specific tasks
- large labeled data sets

Problems with this setting:

1. Large labeled data sets are usually hard to come by.
2. What about loosely related tasks?

Where Machine Learning Shines

Our focus so far -

Supervised learning with:

- highly specific tasks
- large labeled data sets

Problems with this setting:

1. Large labeled data sets are usually hard to come by.
2. What about loosely related tasks?

Humans seem to learn well in primarily unsupervised mode.

What Can Be Done Without Labels?

What Can Be Done Without Labels?

1 Remember

What Can Be Done Without Labels?

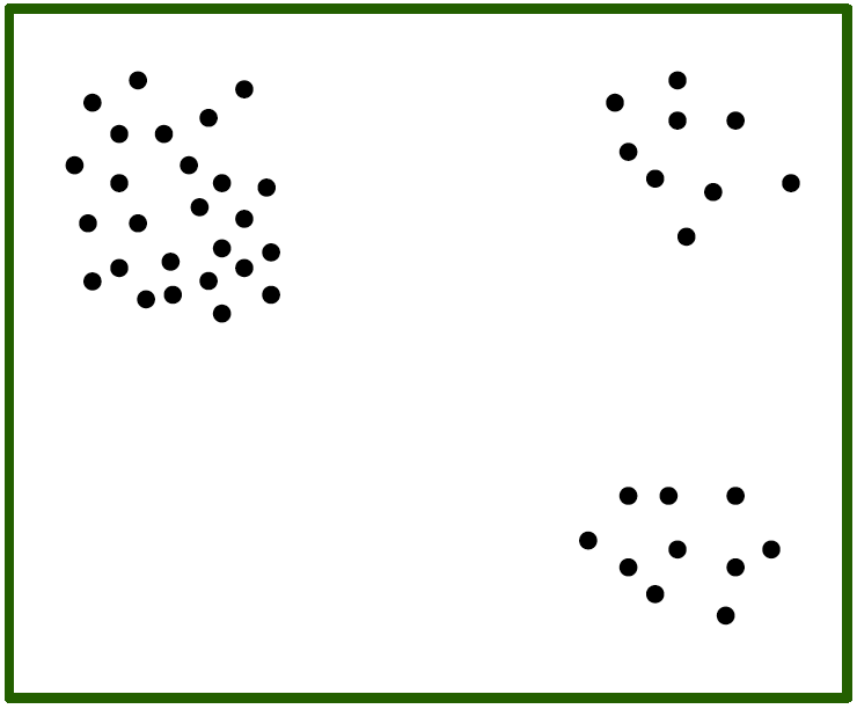
1 Remember

What Can Be Done Without Labels?

2 Summarize

What Can Be Done Without Labels?

2 Summarize

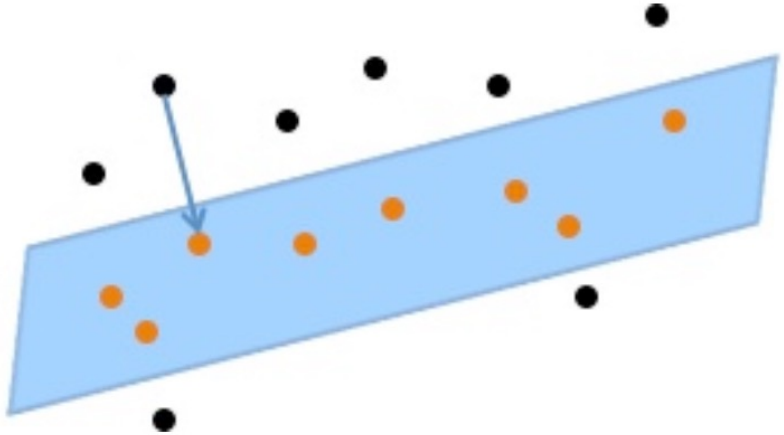


What Can Be Done Without Labels?

3 Compress

What Can Be Done Without Labels?

3 Compress

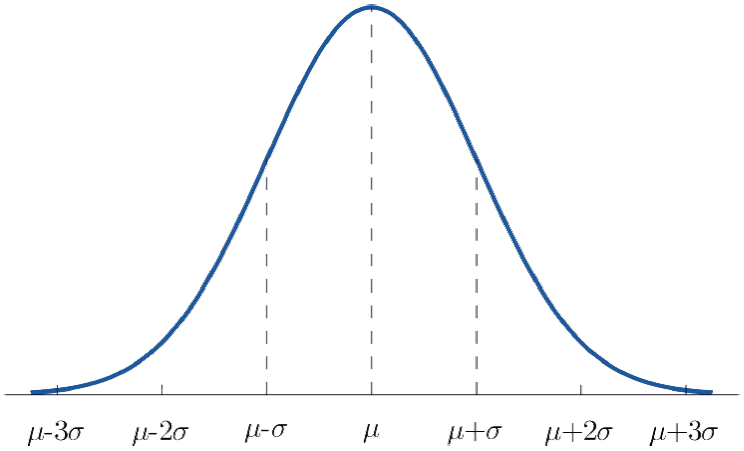


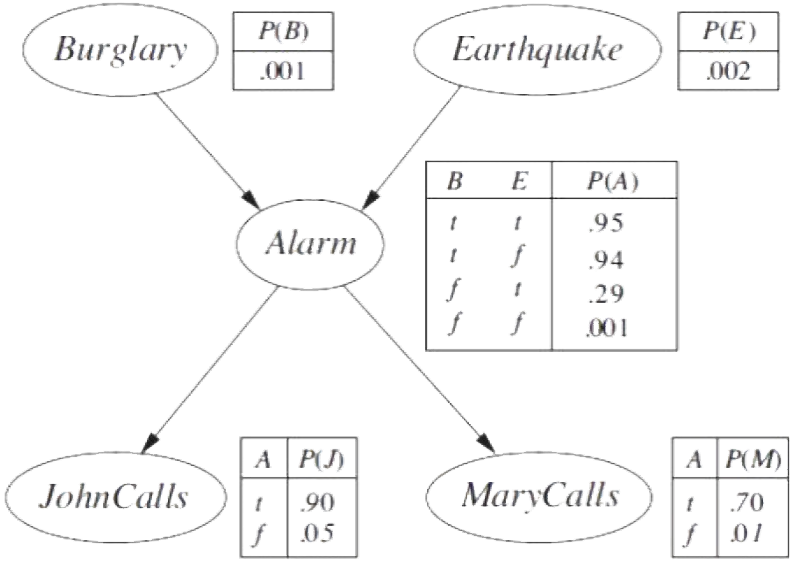
What Can Be Done Without Labels?

4 Generate

What Can Be Done Without Labels?

4 Generate



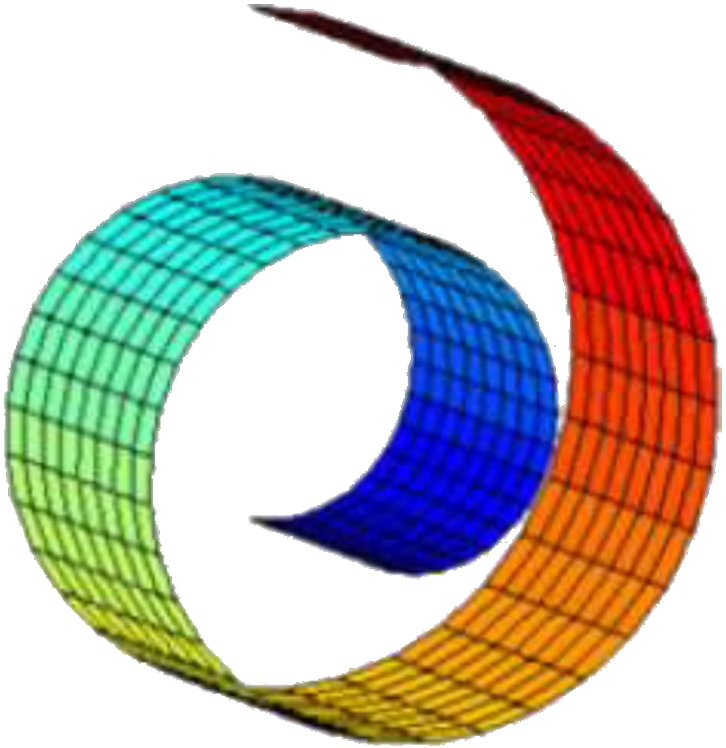


What Can Be Done Without Labels?

5 Embed

What Can Be Done Without Labels?

5 Embed



Some Approaches to Unsupervised Learning

1. Remember
2. Summarize
3. Compress
4. Generate
5. Embed