

UDA Experiments

Michael Morris

November 22, 2019

1 Introduction

Semi supervised learning is a method of training machine learning models which leverages unlabelled data. There are many ways of doing this, typically by using unlabelled samples to create a decision boundary.

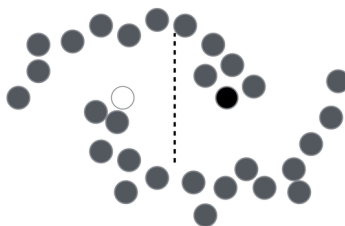


Figure 1: Decision boundary with unsupervised examples

It's intuitive which class each unlabelled sample would be in, and can be mathematically found based on some assumptions:

- Points which are close together are likely to be in the same class
- The data tends to form distinct clusters
- The data can be expressed in much lower dimensions than its inputs.