## PCA

主成分分析在我们的问题中有如下优缺点：

#### 优点

* 对数据进行降维，保留了主要影响因素，使得数据集更易可视化
* 使得结果容易理解
* 降低算法的计算开销

#### 缺点

* 丢失部分信息。被剔除的特征有可能对分类工作产生较大影响。

但是从两类点在二维特征空间中的分布图可以看出，两类点在特征空间中能较为明显的分隔开，因此可以认为PCA在我们的问题中是可取的，对分类任务造成的干扰可以忽略。

## 明显度

我们为了描述区域模糊程度而引入明显度函数，并且认为明显度低于一定阈值就算作模糊区域，它有如下优缺点：

#### 优点

* 能够正确反映出在某个区域内对坐标点进行分类的难度。如果区域内的两类训练集点数目接近，则分类难度高，模糊度高，反之亦然。
* 我们衡量明显度是在二维特征空间中进行，可以方便地进行可视化。

#### 缺点

* 计算某个坐标处的明显度依赖于选取的窗口宽度，并且没有直接的依据来指定窗口宽度。
* 得到的模糊区域是离散的方块区域，难以用统一的数学表达式描述

## 支持向量机

#### 优点

* 算法原理简单直观
* 仅仅使用一部分支持向量来做超平面的决策，无需依赖全部数据。
* 有大量的核函数可以使用

#### 缺点

* 训练集在特征空间上的分布呈现线性可分的特点。因此用基于线性核的支持向量机分类器有较好的效果；但是用基于非线性核的分类器则存在收敛慢，分隔线形状复杂等缺点。

## 改进

我们可以尝试其他的诸如线性判别分析、朴素贝叶斯分类器等线性分类法。