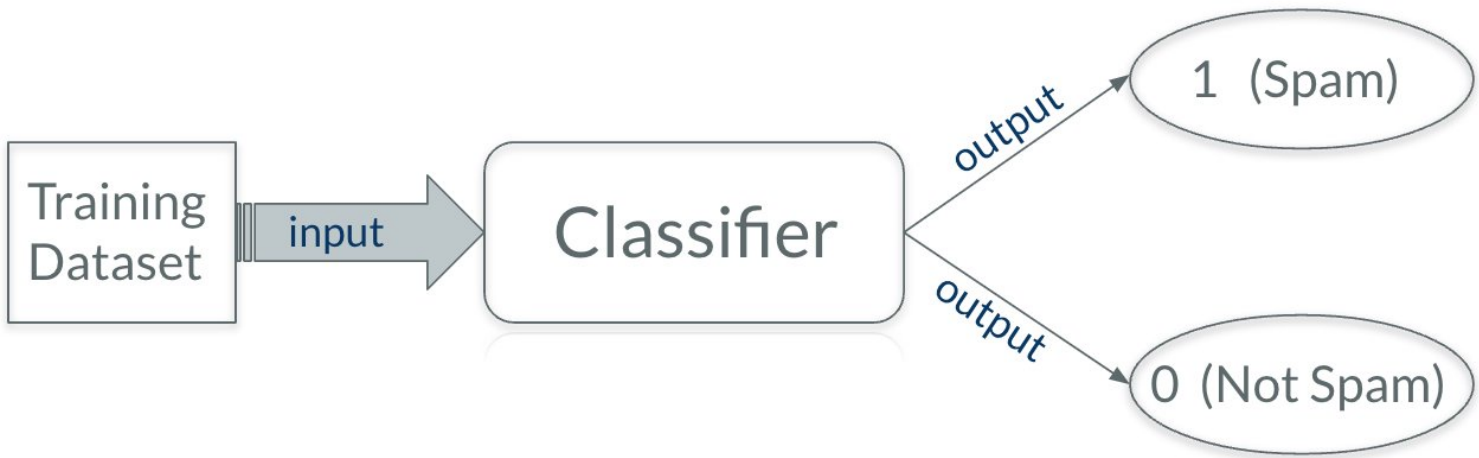
Problem: Binary Classification of High Dimensional Data

- **Classification**: Grouping objects into preset categories, based on their characteristics.
  Objects ⇒ **samples**, Their characteristics ⇒ **features**.

  Applications: Spam Filtering, Medical Diagnosis (Tumor detection based on MRIs), etc.



- **Goal**: To train a "good" classifier
- **Challenge**: High Dimension (Curse of Dimensionality)

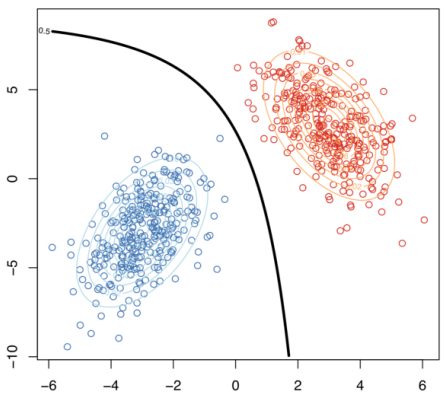
Arman Jahangiri, Department of Mathematical Sciences

UNIVERSITY OF CALGARY

Proposed Approach

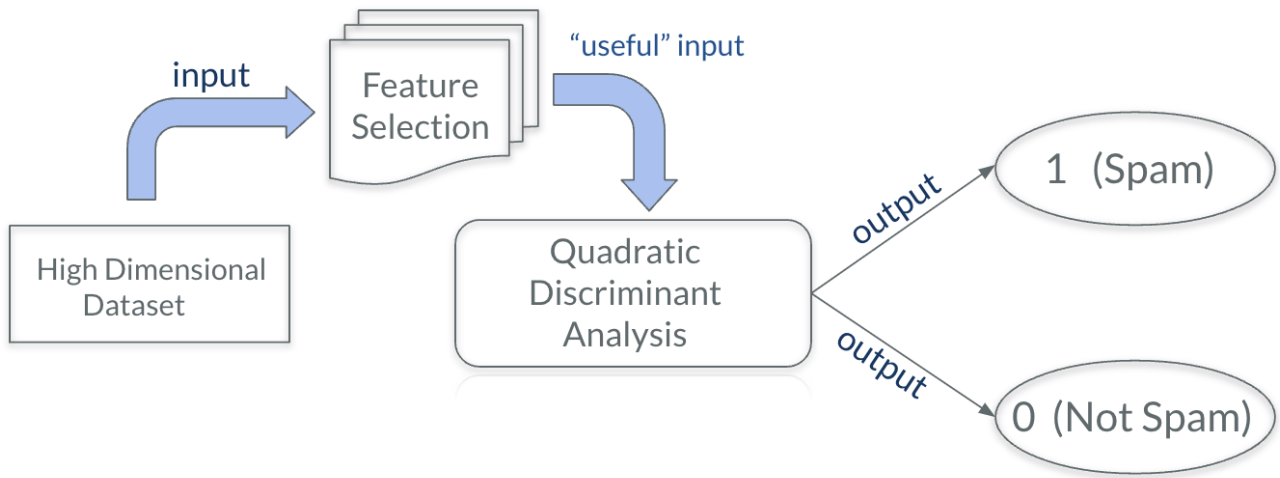
Quadratic Discriminant Analysis (QDA):
Assume (spam/non-spam) emails are normally distributed.

Feature Selection: Only select the "useful" features



QDA in $\mathbb{R}^2$

- **Goals**:
  1. To find QDAFS successful regions
  2. To compare QDAFS with other possible classifiers.



QDAFS

Arman Jahangiri, Department of Mathematical Sciences

UNIVERSITY OF CALGARY