Foundations of Machine Learning AI2000 and AI5000

FoML-18
Classification with the Basis Functions

Dr. Konda Reddy Mopuri
Department of AI, IIT Hyderabad
July-Nov 2025

భారతీయ సాంకేతిక విజ్ఞాన సంస్థ హైదరాబాద్
भारतीय प्रौद्योगिकी संस्थान हैदराबाद
Indian Institute of Technology Hyderabad

DiL
Data-driven Intelligence & Learning Lab

So far in FoML

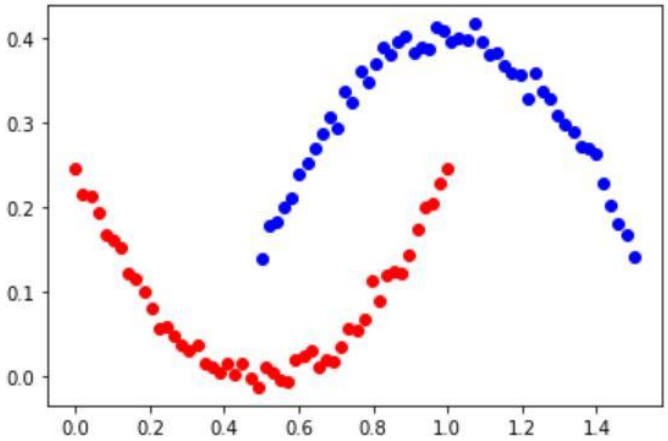
- Intro to ML and Probability refresher
- MLE, MAP, and fully Bayesian treatment
- Supervised learning
    a. Linear Regression with basis functions (regularization, model selection)
    b. Bias-Variance Decomposition (Bayesian Regression)
    c. Decision Theory - three broad classification strategies
        - Probabilistic Generative Models - Continuous & discrete data
        - (Linear) Discriminant Functions - least squares solution, Perceptron

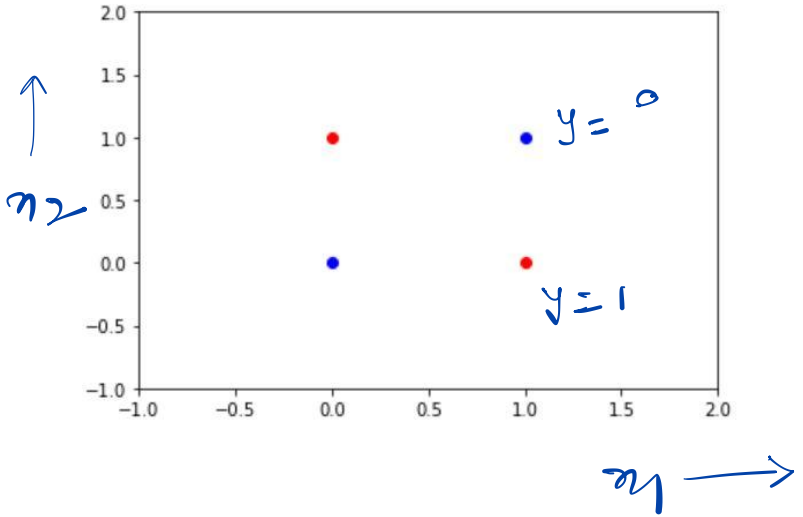
Classification with Basis functions

భారతీయ సాంకేతిక విజ్ఞాన సంస్థ హైదరాబాద్
भारतीय प्रौद्योगिकी संस्थान हैदराबाद
Indian Institute of Technology Hyderabad

DiL
Data-driven Intelligence
& Learning Lab

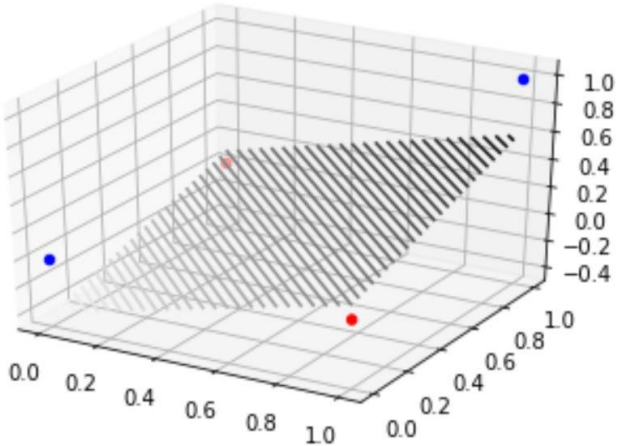
When the data is not linearly separable





Sometimes preprocessing (feature design) helps

- Consider the XOR function with $\phi(\mathbf{x}) = (x_1, x_2, x_1 x_2)^T$
- Perceptron in the new space will classify the data



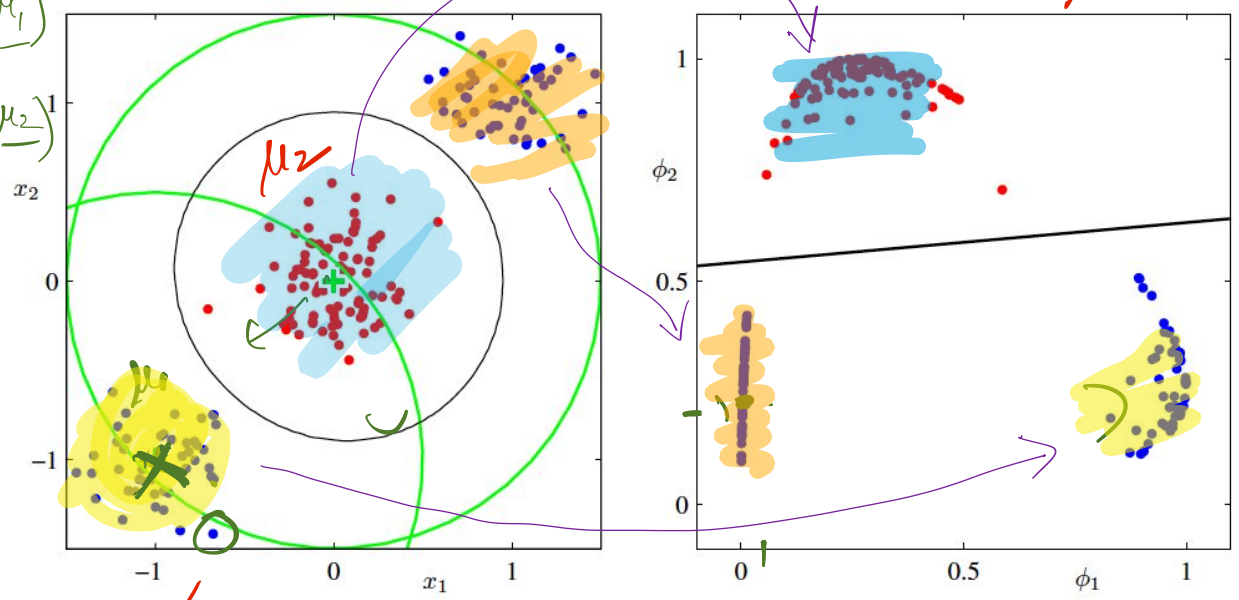
$$y(\mathbf{x}) = f(\mathbf{w}^T \phi(\mathbf{x}))$$

Use of Basis functions

$\phi = [\phi_1 \ \phi_2]^T$ space formed by Basis fns

$\phi_1(\underline{x}) = e^{-\frac{1}{2}(\underline{x}-\underline{\mu_1})^T \Sigma_1 (\underline{x}_1-\underline{\mu_1})}$

$\phi_2(\underline{x}) = e^{-\frac{1}{2}(\underline{x}-\underline{\mu_2})^T \Sigma_2 (\underline{x}_2-\underline{\mu_2})}$



Left: data in the original input space
Right: space of two Gaussian basis functions

i/p space

$x = [x_1 \ x_2]^T$

Basis functions - advantages

- Enable to learn nonlinear mapping functions from i/p to target
- Leads to closed form solution for LS problem and leads to tractable Bayesian treatment

Basis Functions - Limitations

- Basis functions are fixed (not learned)
- Hard to visualize high-dim data to design the basis functions
- In higher dimensions we need more basis functions (curse of dimensionality)

Next
Probabilistic Discriminative models
- Logistic Regression

భారతీయ సాంకేతిక విజ్ఞాన సంస్థ హైదరాబాద్
भारतीय प्रौद्योगिकी संस्थान हैदराबाद
Indian Institute of Technology Hyderabad

DiL
Data-driven Intelligence
& Learning Lab