# Perceptrons, SVMs, Neural Networks ECE 448/ CS 440

Ishan Deshpande

Nov 9

### Perceptrons, SVMs, Neural Networks

### Ishan Deshpande

Supervised Classification

#### erceptrons

Training Algorithm
Multi-class
classification

Support Vector Machines

Picking the best boundary Beyond linear boundaries - the Karnel Trick

## Outline

### Perceptrons. SVMs. Neural Networks

### Ishan Deshpande

## Supervised Classification

### Perceptrons

Linear Separability Training Algorithm Multi-class classification

### Support Vector Machines

Picking the best boundary Beyond linear boundaries - the Kernel Trick

### Neural Networks

Hidden layers

## Learn to tell apart

Perceptrons, SVMs, Neural Networks

Ishan Deshpande

### Supervised Classification

Perceptrons

Training Algorithm
Multi-class
classification

Support Vector Machines

Picking the best boundary Beyond linear boundaries - the

Neural Networks Hidden layers

▶ Given a set of tuples  $\{X_i, Y_i\}$ , learn a function f which tells us  $Y_i$  for a given  $X_i$ .

Training Algorithm
Multi-class
classification

Support Vector Machines

Picking the best boundary Beyond linear boundaries - the

Neural Networks

- ▶ Given a set of tuples  $\{X_i, Y_i\}$ , learn a function f which tells us  $Y_i$  for a given  $X_i$ .
- ► *X<sub>i</sub>* is the *feature vector*, *Y<sub>i</sub>* is the *label*, *f* is the *classifier*.

Training Algorithm
Multi-class
classification

Support Vector Machines

Picking the best boundary Beyond linear boundaries - the

Neural Networks

- ▶ Given a set of tuples  $\{X_i, Y_i\}$ , learn a function f which tells us  $Y_i$  for a given  $X_i$ .
- ► X<sub>i</sub> is the feature vector, Y<sub>i</sub> is the label, f is the classifier.
- ightharpoonup e.g. X = (vectorized) pixel intensity, Y = image type

Linear Separability Training Algorithm

Support Vector

Picking the best boundary Beyond linear

Neural Networks

Supervised Classification

## Perceptrons

Linear Separability

Training Algorithm

Multi-class classification

## Support Vector Machines

Picking the best boundary
Beyond linear boundaries - the Kernel Trick

## What's the classifier

### Perceptrons, SVMs, Neural Networks

### Ishan Deshpande

Supervised

Percentrons

### Linear Separability

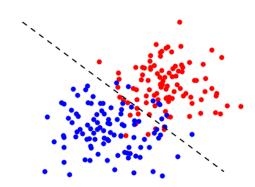
Multi-class classification

### Support Vector Machines

Picking the best boundary Beyond linear boundaries - the

Neural Networks

## What's the classifier



▶ Simply check which side are we on.

Perceptrons, SVMs, Neural Networks

Ishan Deshpande

Supervised

Danasakusas

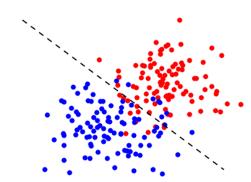
Linear Separability

Training Algorithm
Multi-class

Support Vector Machines

Picking the best boundary Beyond linear boundaries - the Kernel Trick

## What's the classifier



### Perceptrons, SVMs, Neural Networks

### Ishan Deshpande

Supervised

Percentrons

## Linear Separability

Multi-class classification

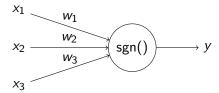
### Support Vector Machines

Picking the best boundary Beyond linear boundaries - the Kernel Trick

- ▶ Simply check which side are we on.
- ▶ Predict  $sgn(\mathbf{w}^T\mathbf{x})$ , where  $\mathbf{w}$  is the **normal** to the boundary.



## What's a perceptron



### Perceptrons, SVMs, Neural Networks

### Ishan Deshpande

Supervised Classification

#### Perceptrons

# Linear Separability Training Algorithm Multi-class

Support Vector Machines

> Picking the best boundary Beyond linear boundaries - the Kernel Trick

Neural Networks

Training Algorithm

## Perceptrons

Training Algorithm

Multi-class classification

Beyond linear boundaries - the Kernel Trick

### Perceptrons, SVMs, Neural Networks

### Ishan Deshpande

Supervised Classification

#### Percentrons

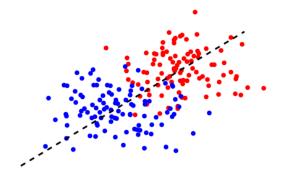
Linear Separabil

Training Algorithm
Multi-class
classification

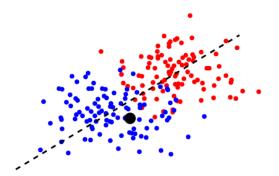
Support Vector Machines

Picking the best boundary Beyond linear boundaries - the

Neural Networks
Hidden layers



▶ Start with a random guess.



▶ Cycle through the training set. Check prediction y' vs actual label y.

Perceptrons, SVMs, Neural Networks

Ishan Deshpande

Supervised Classification

Darcontrone

1:---- C----

Training Algorithm
Multi-class

Support Vector Machines

Picking the best boundary Beyond linear boundaries - the Kernel Trick

### Perceptrons, SVMs, Neural Networks

### Ishan Deshpande

Supervised Classification

Darcontrone

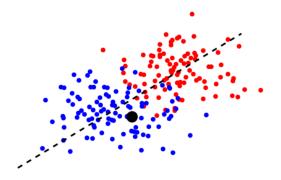
Linear Separability
Training Algorithm

Multi-class classification

Support Vector Machines

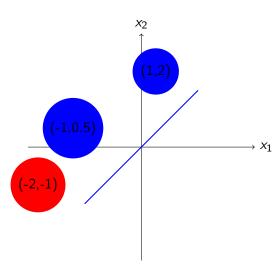
Picking the best boundary Beyond linear boundaries - the

Neural Networks



▶ Update line with the rule:

$$\mathbf{w} = \mathbf{w} + \alpha(y - y')\mathbf{x} \tag{1}$$



Perceptrons, SVMs, Neural Networks

Ishan Deshpande

Supervised

D.....

Perceptrons

Training Algorithm Multi-class

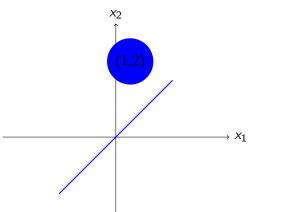
Support Vector Machines

Picking the best boundary Beyond linear boundaries - the Kernel Trick

Consider (1,2) with y=1. For this y'=-1. For  $\alpha=1$  the new **w** is

$$\mathbf{w} = (1, -1) + 1 \times (1 - (-1)) \times (1, 2) \tag{2}$$

$$\mathbf{w} = (3,3) \tag{3}$$



Perceptrons, SVMs, Neural Networks

Ishan Deshpande

Supervised Classificatio

erceptrons

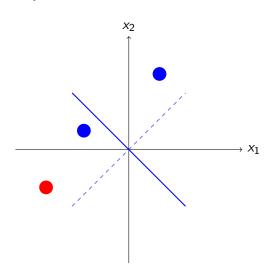
Training Algorithm Multi-class

Support Vector Machines

Picking the best boundary Beyond linear boundaries - the Kernel Trick



### The boundary is now:



Perceptrons, SVMs, Neural Networks

### Ishan Deshpande

Supervised

#### Parcantranc

Linear Separability
Training Algorithm
Multi-class

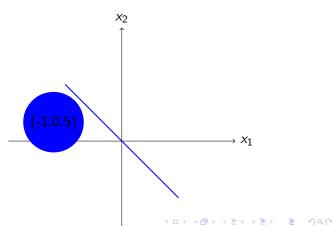
## Support Vector Machines

Picking the best boundary Beyond linear boundaries - the Kernel Trick

Consider (-1, 0.5) with y = 1. For this y' = -1. We update again:

$$\mathbf{w} = (3,3) + 1 \times (1 - (-1)) \times (-1,0.5) \tag{4}$$

$$\mathbf{w} = (1,4) \tag{5}$$



Perceptrons, SVMs, Neural Networks

Ishan Deshpande

Supervised Classificatio

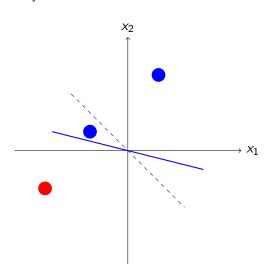
rceptrons

Training Algorithm
Multi-class

Support Vector Machines

Picking the best boundary Beyond linear boundaries - the Kernel Trick

### The boundary is now:



Perceptrons, SVMs, Neural Networks

Ishan Deshpande

Supervised

.....

Linear Separability
Training Algorithm

Multi-class classification

Support Vector Machines

Picking the best boundary Beyond linear boundaries - the Kernel Trick

If the data is indeed linearly separable, it will eventually converge!

### Perceptrons, SVMs, Neural Networks

### Ishan Deshpande

Supervised Classificatio

#### Percentrons

inear Separat

Training Algorithm
Multi-class

Support Vector Machines

Picking the best boundary Beyond linear boundaries - the Kernel Trick

- If the data is indeed linearly separable, it will eventually converge!
- ▶ If the data is NOT linearly separable, training can diverge for a fixed  $\alpha$ .

### Perceptrons, SVMs, Neural Networks

### Ishan Deshpande

Supervised Classificatio

erceptrons

Training Algorithm Multi-class

Support Vector Machines

Picking the best boundary Beyond linear boundaries - the Kernel Trick

#### Perceptrons, SVMs, Neural Networks

### Ishan Deshpande

Supervised Classification

erceptrons

Training Algorithm
Multi-class

Support Vector Machines

Picking the best boundary Beyond linear boundaries - the

leural Networks

- ▶ If the data is indeed linearly separable, it will eventually converge!
- ▶ If the data is NOT linearly separable, training can diverge for a fixed  $\alpha$ .
- Solution: Use  $\alpha = \frac{1}{t}$  finds the best separator if data is actually separable, otherwise finds the MMSE solution .

Support Vector Machines

Picking the best boundary Beyond linear boundaries - the

- ▶ If the data is indeed linearly separable, it will eventually converge!
- ▶ If the data is NOT linearly separable, training can diverge for a fixed  $\alpha$ .
- Solution: Use  $\alpha = \frac{1}{t}$  finds the best separator if data is actually separable, otherwise finds the MMSE solution .
- ▶ Include a bias term, i.e. offset of line.

$$y = sgn(\mathbf{w}^T \mathbf{x} + b) \tag{6}$$

Training Algorithm
Multi-class

Support Vector Machines

Picking the best boundary Beyond linear boundaries - the

Neural Networks
Hidden lavers

- ▶ If the data is indeed linearly separable, it will eventually converge!
- ▶ If the data is NOT linearly separable, training can diverge for a fixed  $\alpha$ .
- Solution: Use  $\alpha = \frac{1}{t}$  finds the best separator if data is actually separable, otherwise finds the MMSE solution .
- ▶ Include a bias term, i.e. offset of line.

$$y = sgn(\mathbf{w}^T \mathbf{x} + b) \tag{6}$$

▶ Use the same training algorithm with  $\tilde{\mathbf{x}}$  and  $\tilde{\mathbf{w}}$  as

$$\tilde{\mathbf{x}} = \{\mathbf{x}, 1\}, \quad \tilde{\mathbf{w}} = \{\mathbf{w}, b\} \tag{7}$$

## Differentiable Variant

▶ Instead of sgn(.), use a differentiable non-linear function, such as the sigmoid  $\sigma(x) = \frac{1}{1+e^{-x}}$ .

### Perceptrons, SVMs, Neural Networks

### Ishan Deshpande

Supervised Classificatio

Б.

Linear Separability
Training Algorithm

Classification
Support Vector

Picking the best boundary Beyond linear

Neural Networks

▶ Instead of sgn(.), use a differentiable non-linear function, such as the sigmoid  $\sigma(x) = \frac{1}{1+e^{-x}}$ .

Minimize

$$E(\mathbf{w}) = \sum_{i} (y_i - f(\mathbf{x}_i))^2$$
 (8)

Training Algorithm

Support Vector Machines

Picking the best boundary Beyond linear boundaries - the

Neural Networks

▶ Instead of sgn(.), use a differentiable non-linear function, such as the sigmoid  $\sigma(x) = \frac{1}{1+e^{-x}}$ .

Minimize

$$E(\mathbf{w}) = \sum_{i} (y_i - f(\mathbf{x}_i))^2$$
 (8)

▶ Update via gradient descent

$$\mathbf{w} = \mathbf{w} - \alpha \frac{d}{d\mathbf{w}} E(\mathbf{w}) \tag{9}$$

Training Algorithm Multi-class

Support Vector Machines

Picking the best boundary Beyond linear boundaries - the

Neural Networks Hidden layers

▶ Instead of sgn(.), use a differentiable non-linear function, such as the sigmoid  $\sigma(x) = \frac{1}{1+a-x}$ .

Minimize

$$E(\mathbf{w}) = \sum_{i} (y_i - f(\mathbf{x}_i))^2$$
 (8)

▶ Update via gradient descent

$$\mathbf{w} = \mathbf{w} - \alpha \frac{d}{d\mathbf{w}} E(\mathbf{w}) \tag{9}$$

► For the sigmoid, this is:

$$\mathbf{w} = \mathbf{w} - \alpha(y - f(\mathbf{x}))f(\mathbf{x})(1 - f(\mathbf{x}))\mathbf{x}$$
 (10)

Multi-class classification

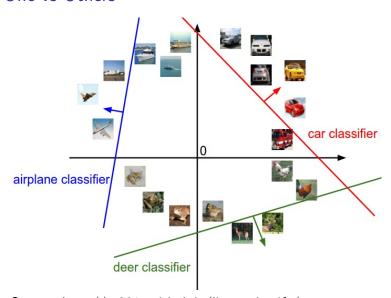
Ishan Deshpande

### Perceptrons

Multi-class classification

Beyond linear boundaries - the Kernel Trick

### One vs Others



Source: http://cs231n.github.io/linear-classify/

Perceptrons, SVMs, Neural Networks

Ishan Deshpande

Supervised Classification

Percentrons

Linear Separability
Training Algorithm
Multi-class
classification

upport Vector

Picking the best boundary Beyond linear

Beyond linear boundaries - the Kernel Trick

### One vs Others

One classifier for one class.

### Perceptrons, SVMs, Neural Networks

### Ishan Deshpande

Supervised Classification

#### Perceptrons

Linear Separability
Training Algorithm
Multi-class
classification

Support Vector Machines

Picking the best boundary Beyond linear boundaries - the Kernel Trick

Neural Networks

Picking the best boundary Beyond linear boundaries - the

- One classifier for one class.
- Predict

$$c = \operatorname{argmax}_{c'} \ \mathbf{w}_{c'}^T \mathbf{x} \tag{11}$$

classification

### Predict

$$c = argmax_{c'} \quad \mathbf{w}_{c'}^T \mathbf{x} \tag{11}$$

▶ If c is misclassified as c', update using

$$\mathbf{w}_c = \mathbf{w}_c + \alpha \mathbf{x} \tag{12}$$

$$\mathbf{w}_{c'} = \mathbf{w}_{c'} - \alpha \mathbf{x} \tag{13}$$

## Outline

Perceptrons, SVMs, Neural Networks

### Ishan Deshpande

Supervised Classification

#### Perceptrons

Linear Separability Training Algorithm Multi-class classification

Support Vector Machines

#### Picking the best boundary Beyond linear

Beyond linear boundaries - the

Neural Networks Hidden layers

Supervised Classification

### Perceptrons

Linear Separability
Training Algorithm
Multi-class classification

## Support Vector Machines

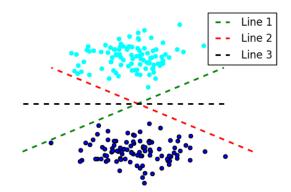
Picking the best boundary

Beyond linear boundaries - the Kernel Trick

Neural Networks

Hidden layers

## Which boundary is better?



### Perceptrons, SVMs, Neural Networks

### Ishan Deshpande

Supervised Classification

#### Perceptrons

Training Algorithm Multi-class classification

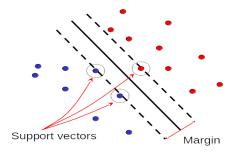
Support Vector Machines

### Picking the best boundary Beyond linear

Beyond linear boundaries - the Kernel Trick

## Which boundary is better?

Intuitively, pick the one that is equally distant from both classes.



A Tutorial on Support Vector Machines for Pattern Recognition

### Perceptrons, SVMs, Neural Networks

Ishan Deshpande

Supervised Classification

Percentrons

Linear Separability
Training Algorithm
Multi-class
classification

Support Vector

Machines

Picking the best boundary Beyond linear boundaries - the

### Perceptrons, SVMs, Neural Networks

Ishan Deshpande

Supervised Classification

Classification

Linear Separabil

Multi-class classification

Support Vector Machines

Picking the best boundary Beyond linear boundaries - the

Neural Networks

Perpendicular distance of support vectors from the boundary is:

$$\frac{\left|\mathbf{w}^{\mathsf{T}}\mathbf{x} + b\right|}{\|\mathbf{w}\|}\tag{14}$$

Support Vector Machines

Picking the best boundary Beyond linear

Kernel Trick Neural Networks

Neural Networks Hidden layers

Perpendicular distance of support vectors from the boundary is:

$$\frac{\left|\mathbf{w}^{T}\mathbf{x} + b\right|}{\|\mathbf{w}\|} \tag{14}$$

▶ Suppose we require, for all support vectors, that :

$$\left|\mathbf{w}^{\mathsf{T}}\mathbf{x} + b\right| = 1 \tag{15}$$

Training Algorithm Multi-class classification

Support Vector Machines

Picking the best boundary Beyond linear

Kernel Trick Neural Networks

Veural Networks Hidden layers

Perpendicular distance of support vectors from the boundary is:

$$\frac{\left|\mathbf{w}^{T}\mathbf{x} + b\right|}{\|\mathbf{w}\|} \tag{14}$$

▶ Suppose we require, for all support vectors, that :

$$\left|\mathbf{w}^{\mathsf{T}}\mathbf{x} + b\right| = 1 \tag{15}$$

► The margin is then  $\frac{2}{\|\mathbf{w}\|}$ 

► Formulated as:

$$min_{\mathbf{w},b} \quad \frac{1}{2}\mathbf{w}^T\mathbf{w}$$
 (16)

subject to:

$$y_i(\mathbf{w}^T\mathbf{x}_i + b) \ge 1 \tag{17}$$

Perceptrons, SVMs, Neural Networks

Ishan Deshpande

Supervised

erceptrons

Linear Separability Training Algorithm Multi-class classification

Support Vector Machines

Picking the best boundary Beyond linear boundaries - the

Neural Networks

► Formulated as:

$$min_{\mathbf{w},b} \quad \frac{1}{2}\mathbf{w}^T\mathbf{w}$$
 (16)

subject to:

$$y_i(\mathbf{w}^T\mathbf{x}_i + b) \ge 1 \tag{17}$$

► Eq. 16 ensures maximum margin, while Eq. 17 ensures correct classification.

### Perceptrons, SVMs, Neural Networks

Ishan Deshpande

Supervised Classificatio

erceptrons

Linear Separability Training Algorithm Multi-class classification

Support Vector Machines

Picking the best boundary Beyond linear

Perceptrons.

Formulated as:

$$min_{\mathbf{w},b} \quad \frac{1}{2}\mathbf{w}^{T}\mathbf{w} \tag{16}$$

subject to:

$$y_i(\mathbf{w}^T\mathbf{x}_i + b) \ge 1 \tag{17}$$

- ▶ Eq. 16 ensures maximum margin, while Eq. 17 ensures correct classification.
- The classifier is of the form:

$$\mathbf{w} = \sum_{i} \alpha_{i} y_{i} \mathbf{x}_{i} \tag{18}$$

and

$$y = \sum_{i} \alpha_{i} y_{i} \mathbf{x}_{i}^{\mathsf{T}} \mathbf{x} + b \tag{19}$$

where  $\alpha_i$  are learned weights.

► This can be relaxed if the data is not actually separable

$$min_{\mathbf{w},b} \frac{1}{2} \mathbf{w}^T \mathbf{w} + C \sum_i max(0, 1 - y_i(\mathbf{w}^T \mathbf{x}_i + b))$$
 (20)

Perceptrons, SVMs, Neural Networks

Ishan Deshpande

Supervised Classification

Parcantranc

Linear Separability Training Algorithm Multi-class classification

Support Vector
Machines

Picking the best boundary Beyond linear boundaries - the

► This can be relaxed if the data is not actually separable

$$min_{\mathbf{w},b} \frac{1}{2} \mathbf{w}^T \mathbf{w} + C \sum_{i} max(0, 1 - y_i(\mathbf{w}^T \mathbf{x}_i + b))$$
 (20)

► *C* allows you to give weight to one over the other.

### Perceptrons, SVMs, Neural Networks

Ishan Deshpande

Supervised Classification

Perceptrons

Training Algorithm Multi-class classification

Support Vector
Machines

Picking the best boundary Beyond linear boundaries - the

## Which boundary is better?

▶ This can be relaxed if the data is not actually separable

$$min_{\mathbf{w},b} \frac{1}{2} \mathbf{w}^T \mathbf{w} + C \sum_{i} max(0, 1 - y_i(\mathbf{w}^T \mathbf{x}_i + b))$$
 (20)

- ► *C* allows you to give weight to one over the other.
- Here we judge classification accuracy with the 'hinge' loss

$$max(0, 1 - y_i(\mathbf{w}^T \mathbf{x}_i + b))$$
 (21)



### Perceptrons, SVMs, Neural Networks

Ishan Deshpande

Supervised Classificatio

Percentrons

Training Algorithm
Multi-class

Support Vector
Machines

Picking the best boundary Beyond linear

Neural Networks

### Outline

Perceptrons, SVMs, Neural Networks

### Ishan Deshpande

Supervised Classification

### Perceptrons

Linear Separability
Training Algorithm
Multi-class

Support Vector Machines

boundary

Beyond linear boundaries - the Kernel Trick

Neural Networks Hidden lavers

### Supervised Classification

### Perceptrons

Linear Separability
Training Algorithm
Multi-class classification

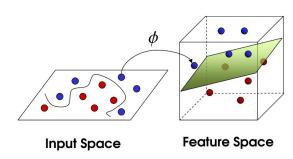
## Support Vector Machines

Picking the best boundary

Beyond linear boundaries - the Kernel Trick

## What if the data is not linearly separable?

 $\blacktriangleright$  Try mapping it to a space where it is! Use a  $\phi$  such that



Perceptrons, SVMs, Neural Networks

Ishan Deshpande

Supervised Classificatio

Perceptrons

Training Algorithm
Multi-class
classification

Support Vector Machines

boundary

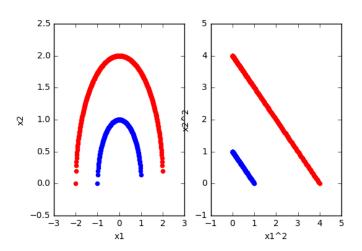
Beyond linear

Beyond linear boundaries - the Kernel Trick

Perceptrons,

Beyond linear boundaries - the Kernel Trick

▶ Consider points in concentric circles. Map  $x_i$  to  $x_i^2$ .



► Eq. 19 will be rewritten as:

$$y = \sum_{i} \alpha_{i} y_{i} \phi(\mathbf{x}_{i})^{T} \phi(\mathbf{x}) + b$$
 (22)

Perceptrons, SVMs, Neural Networks

Ishan Deshpande

Supervised Classification

Percentrons

Linear Separability
Training Algorithm
Multi-class

Support Vector Machines

Beyond linear boundaries - the

Kernel Trick Neural Networks ► Eq. 19 will be rewritten as:

$$y = \sum_{i} \alpha_{i} y_{i} \phi(\mathbf{x}_{i})^{T} \phi(\mathbf{x}) + b$$
 (22)

▶ Instead of explicitly defining  $\phi$ , we can also define a  $K(\mathbf{x}, \mathbf{x'})$  such that

$$y = \sum_{i} \alpha_{i} y_{i} K(\mathbf{x}_{i}, \mathbf{x}) + b$$
 (23)

Note: K must satisfy Mercer's conditions. Examples include polynomial kernels  $(1 + \mathbf{x}^T \mathbf{x'})^d$ , Gaussian kernels  $exp(\frac{-1}{2}(\mathbf{x}-\mathbf{x'})^T(\mathbf{x}-\mathbf{x'}))$ 

Perceptrons, SVMs, Neural Networks

Ishan Deshpande

Supervised Classificatio

Perceptrons

Training Algorithm

Multi-class

apport Vector lachines

icking the best oundary

Beyond linear boundaries - the Kernel Trick

## Outline

Perceptrons, SVMs, Neural Networks

### Ishan Deshpande

Supervised Classification

### Perceptrons

Linear Separability
Training Algorithm
Multi-class

### Support Vector Machines

Picking the best boundary Beyond linear boundaries - the

### Neural Networks

Hidden layers

Supervised Classification

### Perceptrons

Linear Separability
Training Algorithm
Multi-class classification

## Support Vector Machines

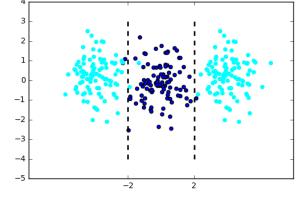
Picking the best boundary
Beyond linear boundaries - the Kernel Trick

## Neural Networks

e.g. Consider a simple 1-d scenario

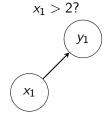
Perceptrons,

Hidden layers



Stacks of perceptrons can learn non-linear functions.

## Combine several perceptron units



### Perceptrons, SVMs, Neural Networks

### Ishan Deshpande

Supervised Classification

#### erceptrons

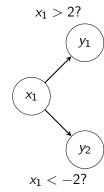
Training Algorithm
Multi-class
classification

## Support Vector Machines

Picking the bes boundary Beyond linear boundaries - the Kernel Trick

### Neural Networks

# Combine several perceptron units



### Perceptrons, SVMs, Neural Networks

### Ishan Deshpande

Supervised Classification

#### erceptrons

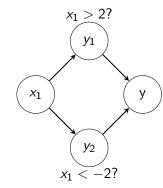
Training Algorithm
Multi-class
classification

## Support Vector Machines

Picking the be boundary Beyond linear boundaries - th Kernel Trick

### Neural Networks

## Combine several perceptron units



### Perceptrons, SVMs, Neural Networks

### Ishan Deshpande

Supervised Classificatio

#### Perceptrons

Training Algorithm
Multi-class
classification

### Support Vector Machines

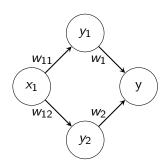
Picking the bes boundary Beyond linear boundaries - the Kernel Trick

### Neural Networks

 $w_{ij}$  represents weight to perceptron i in the hidden layer from input j, and  $w_i$  represents weight of perceptron i in the hidden layer to the output. Then:

$$y_i = sgn(\sum_j w_{ij} x_j) \tag{24}$$

$$y = sgn(\sum_{i} w_{i} \times y_{i}) \tag{25}$$



Perceptrons, SVMs, Neural Networks

Ishan Deshpande

upervised lassificatio

erceptrons

Linear Separability
Training Algorithm
Multi-class
classification

apport Vector

Picking the best boundary Beyond linear boundaries - the Kernel Trick



## How do we train this monster?

Use differentiable perceptrons.

### Perceptrons, SVMs, Neural Networks

### Ishan Deshpande

Supervised Classification

#### Percentrons

Linear Separability Training Algorithm Multi-class classification

### Support Vector Machines

Picking the bes boundary Beyond linear boundaries - th

### Neural Networks

### How do we train this monster?

- Use differentiable perceptrons.
- Minimize

$$E(\mathbf{w}) = \sum_{i} (y_i - f(\mathbf{x}_i))^2$$
 (26)

using gradient descent.

### Perceptrons, SVMs, Neural Networks

### Ishan Deshpande

Supervised Classification

#### erceptrons

Linear Separability
Training Algorithm
Multi-class
classification

### Support Vector Machines

Picking the best boundary Beyond linear boundaries - the

### Neural Networks

Hidden layers

- Use differentiable perceptrons.
- Minimize

$$E(\mathbf{w}) = \sum_{i} (y_i - f(\mathbf{x}_i))^2$$
 (26)

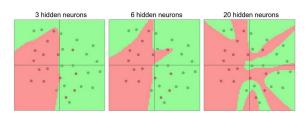
using gradient descent.

Use chain rule to recursively compute gradients from output layer to input - pass information backwards.

$$\frac{d}{dw_{11}}E(\mathbf{w}) = \left(\frac{d}{dy_1}E(\mathbf{w})\right)\frac{dy_1}{dw_{11}} \tag{27}$$

## How powerful is this hidden layer?

http://playground.tensorflow.org/



### Perceptrons, SVMs, Neural Networks

### Ishan Deshpande

Supervised Classification

#### ercentrons

Training Algorithm
Multi-class
classification

### Support Vector Machines

Picking the best boundary Beyond linear boundaries - the Kernel Trick

### Neural Networks