Maximum Margin Separations in Finite Closure Systems

Florian Seiffarth
Tamás Horváth
Stefan Wrobel

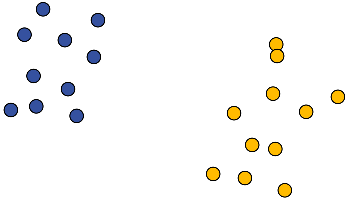
UNIVERSITÄT BONN

ML AI lab

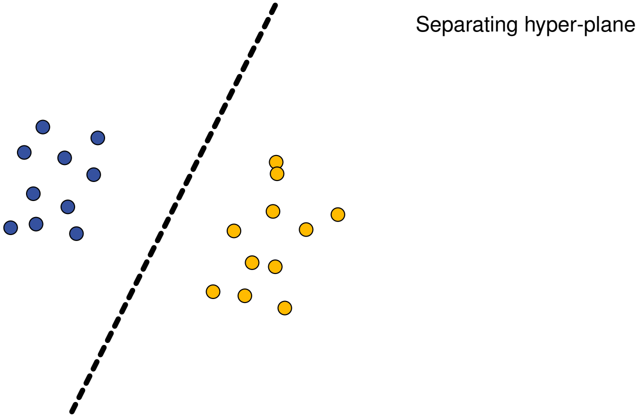
Fraunhofer IAIS

Hyperplane Separation in $\mathbb{R}^d$

Hyperplane Separation in $\mathbb{R}^d$

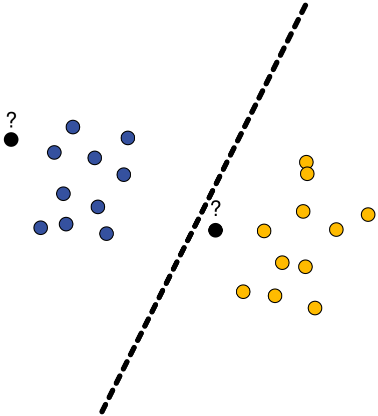


Hyperplane Separation in $\mathbb{R}^d$



- classical machine learning problem (Rosenblatt, 1958)

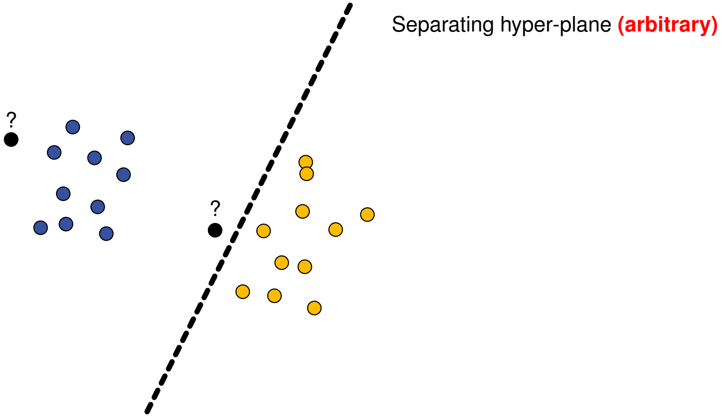
Hyperplane Separation in $\mathbb{R}^d$



Separating hyper-plane **(arbitrary)**

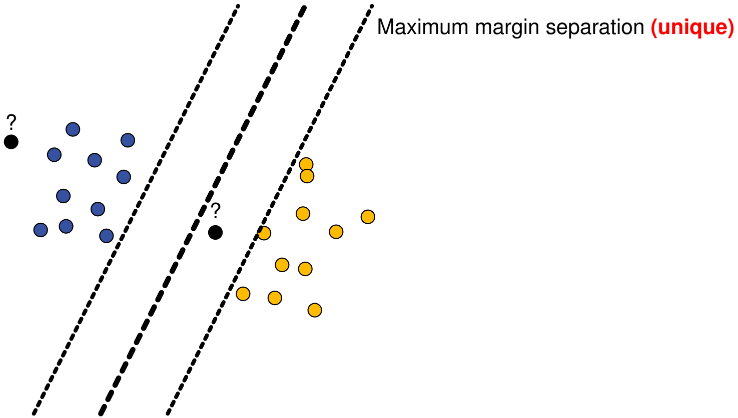
- classical machine learning problem (Rosenblatt, 1958), not unique, over-fitting

Hyperplane Separation in $\mathbb{R}^d$



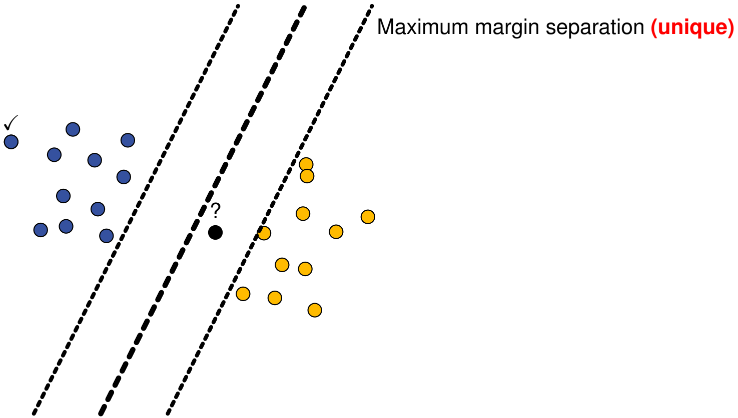
- classical machine learning problem (Rosenblatt, 1958), not unique, over-fitting

Hyperplane Separation in $\mathbb{R}^d$



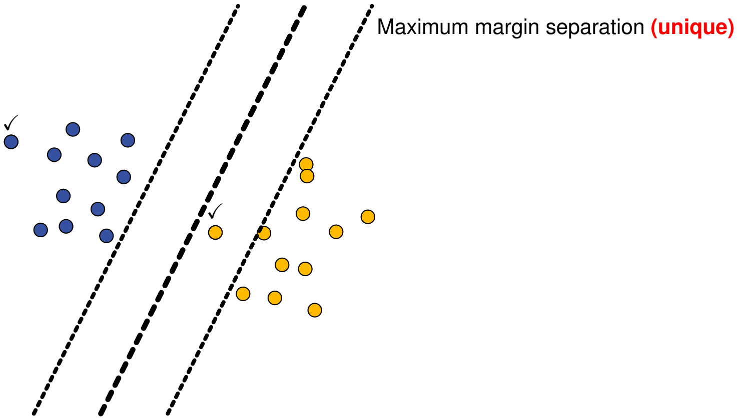
- classical machine learning problem (Rosenblatt, 1958), not unique, over-fitting
- support vector machines (Vapnik, 1992), unique

Hyperplane Separation in $\mathbb{R}^d$



- classical machine learning problem (Rosenblatt, 1958), not unique, over-fitting
- support vector machines (Vapnik, 1992), unique

Hyperplane Separation in $\mathbb{R}^d$



- classical machine learning problem (Rosenblatt, 1958), not unique, over-fitting
- support vector machines (Vapnik, 1992), unique

This Work: Maximum margins in finite closure systems

This Work: Maximum margins in finite closure systems

Adaptions:

This Work: Maximum margins in finite closure systems

Adaptions:

$\mathbb{R}^d \quad \rightarrow$

This Work: Maximum margins in finite closure systems

Adaptions:

$\mathbb{R}^d \quad \rightarrow \quad E$ (*finite* ground set)

This Work: Maximum margins in finite closure systems

Adaptions:

$\mathbb{R}^d$	$\rightarrow$	E (*finite* ground set)
convexity	$\rightarrow$	

This Work: Maximum margins in finite closure systems

Adaptions:

$\mathbb{R}^d$	$\rightarrow$	E (*finite* ground set)
convexity	$\rightarrow$	closure operator

This Work: Maximum margins in finite closure systems

Adaptions:

$\mathbb{R}^d$	$\rightarrow$	E (*finite* ground set)
convexity	$\rightarrow$	closure operator
convex hull	$\rightarrow$	

This Work: Maximum margins in finite closure systems

Adaptions:

$\mathbb{R}^d$	$\rightarrow$	E (*finite* ground set)
convexity	$\rightarrow$	closure operator
convex hull	$\rightarrow$	closed set

This Work: Maximum margins in finite closure systems

Adaptions:

$\mathbb{R}^d$	$\rightarrow$	E (*finite* ground set)
convexity	$\rightarrow$	closure operator
convex hull	$\rightarrow$	closed set
all convex hulls	$\rightarrow$	

Definition:

closure operator c over E:

- $A \subseteq c(A)$
- $A \subseteq B \Rightarrow c(A) \subseteq c(B)$
- $c(c(A)) = c(A)$

Definition:

closure system $(E, \mathcal{C})$:

- $\emptyset, E \in \mathcal{C}$
- $A, B \in \mathcal{C} \Rightarrow A \cap B \in \mathcal{C}$

This Work: Maximum margins in finite closure systems

Adaptions:

$\mathbb{R}^d$	$\rightarrow$	E (*finite* ground set)
convexity	$\rightarrow$	closure operator
convex hull	$\rightarrow$	closed set
all convex hulls	$\rightarrow$	closure system

Definition:

closure operator c over E:

- $A \subseteq c(A)$
- $A \subseteq B \Rightarrow c(A) \subseteq c(B)$
- $c(c(A)) = c(A)$

Definition:

closure system $(E, \mathcal{C})$:

- $\emptyset, E \in \mathcal{C}$
- $A, B \in \mathcal{C} \Rightarrow A \cap B \in \mathcal{C}$

This Work: Maximum margins in finite closure systems

Adaptions:

$\mathbb{R}^d$	$\rightarrow$	E (*finite* ground set)
convexity	$\rightarrow$	closure operator
convex hull	$\rightarrow$	closed set
all convex hulls	$\rightarrow$	closure system
separation	$\rightarrow$	

Definition:

closure operator c over E:

- $A \subseteq c(A)$
- $A \subseteq B \Rightarrow c(A) \subseteq c(B)$
- $c(c(A)) = c(A)$

Definition:

closure system $(E, \mathcal{C})$:

- $\emptyset, E \in \mathcal{C}$
- $A, B \in \mathcal{C} \Rightarrow A \cap B \in \mathcal{C}$

This Work: Maximum margins in finite closure systems

Adaptions:

$\mathbb{R}^d$	$\rightarrow$	E (*finite* ground set)
convexity	$\rightarrow$	closure operator
convex hull	$\rightarrow$	closed set
all convex hulls	$\rightarrow$	closure system
separation	$\rightarrow$	disjoint closed sets

Definition:

closure operator c over E:

- $A \subseteq c(A)$
- $A \subseteq B \Rightarrow c(A) \subseteq c(B)$
- $c(c(A)) = c(A)$

Definition:

closure system $(E, \mathcal{C})$:

- $\emptyset, E \in \mathcal{C}$
- $A, B \in \mathcal{C} \Rightarrow A \cap B \in \mathcal{C}$

This Work: Maximum margins in finite closure systems

Adaptions:

$\mathbb{R}^d$	$\rightarrow$	E (*finite* ground set)
convexity	$\rightarrow$	closure operator
convex hull	$\rightarrow$	closed set
all convex hulls	$\rightarrow$	closure system
separation	$\rightarrow$	disjoint closed sets
metric	$\rightarrow$	

Definition:

closure operator c over E:

- $A \subseteq c(A)$
- $A \subseteq B \Rightarrow c(A) \subseteq c(B)$
- $c(c(A)) = c(A)$

Definition:

closure system $(E, \mathcal{C})$:

- $\emptyset, E \in \mathcal{C}$
- $A, B \in \mathcal{C} \Rightarrow A \cap B \in \mathcal{C}$

This Work: Maximum margins in finite closure systems

Adaptions:

$\mathbb{R}^d$	$\rightarrow$	E (*finite* ground set)
convexity	$\rightarrow$	closure operator
convex hull	$\rightarrow$	closed set
all convex hulls	$\rightarrow$	closure system
separation	$\rightarrow$	disjoint closed sets
metric	$\rightarrow$	monotone linkage function

Definition:

closure operator c over E:

- $A \subseteq c(A)$
- $A \subseteq B \Rightarrow c(A) \subseteq c(B)$
- $c(c(A)) = c(A)$

Definition:

closure system $(E, \mathcal{C})$:

- $\emptyset, E \in \mathcal{C}$
- $A, B \in \mathcal{C} \Rightarrow A \cap B \in \mathcal{C}$

This Work: Maximum margins in finite closure systems

Adaptions:

$\mathbb{R}^d$	$\rightarrow$	E (*finite* ground set)
convexity	$\rightarrow$	closure operator
convex hull	$\rightarrow$	closed set
all convex hulls	$\rightarrow$	closure system
separation	$\rightarrow$	disjoint closed sets
metric	$\rightarrow$	monotone linkage function
separation margin	$\rightarrow$	

Definition:

closure operator c over E:

- $A \subseteq c(A)$
- $A \subseteq B \Rightarrow c(A) \subseteq c(B)$
- $c(c(A)) = c(A)$

Definition:

closure system $(E, \mathcal{C})$:

- $\emptyset, E \in \mathcal{C}$
- $A, B \in \mathcal{C} \Rightarrow A \cap B \in \mathcal{C}$

This Work: Maximum margins in finite closure systems

Adaptions:

$\mathbb{R}^d$	$\rightarrow$	E (*finite* ground set)
convexity	$\rightarrow$	closure operator
convex hull	$\rightarrow$	closed set
all convex hulls	$\rightarrow$	closure system
separation	$\rightarrow$	disjoint closed sets
metric	$\rightarrow$	monotone linkage function
separation margin	$\rightarrow$	?

- **How to define margins in finite closure systems?**

Definition:

closure operator c over E:

- $A \subseteq c(A)$
- $A \subseteq B \Rightarrow c(A) \subseteq c(B)$
- $c(c(A)) = c(A)$

Definition:

closure system $(E, \mathcal{C})$:

- $\emptyset, E \in \mathcal{C}$
- $A, B \in \mathcal{C} \Rightarrow A \cap B \in \mathcal{C}$

This Work: Maximum margins in finite closure systems

Adaptions:

$\mathbb{R}^d$	$\rightarrow$	E (*finite* ground set)
convexity	$\rightarrow$	closure operator
convex hull	$\rightarrow$	closed set
all convex hulls	$\rightarrow$	closure system
separation	$\rightarrow$	disjoint closed sets
metric	$\rightarrow$	monotone linkage function
separation margin	$\rightarrow$	?

- **How to define margins in finite closure systems?**
- **Which properties should be preserved?**

Definition:

closure operator c over E:

- $A \subseteq c(A)$
- $A \subseteq B \Rightarrow c(A) \subseteq c(B)$
- $c(c(A)) = c(A)$

Definition:

closure system $(E, \mathcal{C})$:

- $\emptyset, E \in \mathcal{C}$
- $A, B \in \mathcal{C} \Rightarrow A \cap B \in \mathcal{C}$

This Work: Maximum margins in finite closure systems

Adaptions:

$\mathbb{R}^d$	$\rightarrow$	E (*finite* ground set)
convexity	$\rightarrow$	closure operator
convex hull	$\rightarrow$	closed set
all convex hulls	$\rightarrow$	closure system
separation	$\rightarrow$	disjoint closed sets
metric	$\rightarrow$	monotone linkage function
separation margin	$\rightarrow$	?

- **How to define margins in finite closure systems?**
- **Which properties should be preserved?**
- **Goal: Efficient algorithm!**

Definition:

closure operator c over E:

- $A \subseteq c(A)$
- $A \subseteq B \Rightarrow c(A) \subseteq c(B)$
- $c(c(A)) = c(A)$

Definition:

closure system $(E, \mathcal{C})$:

- $\emptyset, E \in \mathcal{C}$
- $A, B \in \mathcal{C} \Rightarrow A \cap B \in \mathcal{C}$

Maximum Margins in Finite Closure Systems

How to define margins in finite closure systems?

Monotone Linkage Functions (Mullat (1976))

Function $l : 2^E \to E$ such that

$$X \subseteq Y \Rightarrow l(X, e) \geq l(Y, e)$$

for all $X, Y \subseteq E$.

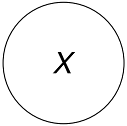
How to define margins in finite closure systems?

Monotone Linkage Functions (Mullat (1976))

Function $l : 2^E \to E$ such that

$$X \subseteq Y \Rightarrow l(X, e) \geq l(Y, e)$$

for all $X, Y \subseteq E$.



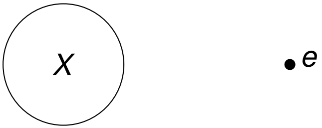
How to define margins in finite closure systems?

Monotone Linkage Functions (Mullat (1976))

Function $l : 2^E \to E$ such that

$$X \subseteq Y \Rightarrow l(X, e) \geq l(Y, e)$$

for all $X, Y \subseteq E$.



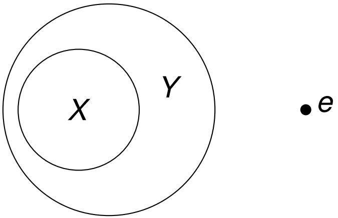
How to define margins in finite closure systems?

Monotone Linkage Functions (Mullat (1976))

Function $l : 2^E \to E$ such that

$$X \subseteq Y \Rightarrow l(X, e) \geq l(Y, e)$$

for all $X, Y \subseteq E$.



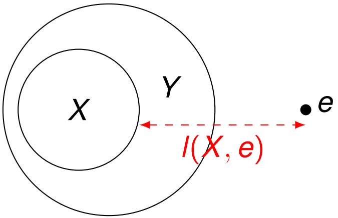
How to define margins in finite closure systems?

Monotone Linkage Functions (Mullat (1976))

Function $l : 2^E \to E$ such that

$$X \subseteq Y \Rightarrow l(X, e) \geq l(Y, e)$$

for all $X, Y \subseteq E$.



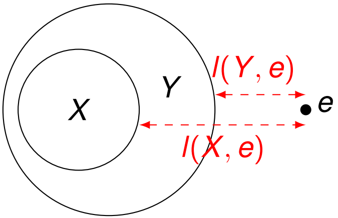
How to define margins in finite closure systems?

Monotone Linkage Functions (Mullat (1976))

Function $l : 2^E \rightarrow E$ such that

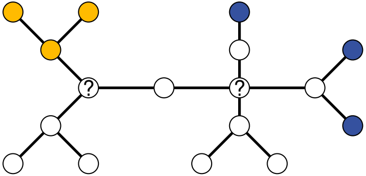
$$X \subseteq Y \Rightarrow l(X, e) \geq l(Y, e)$$

for all $X, Y \subseteq E$.

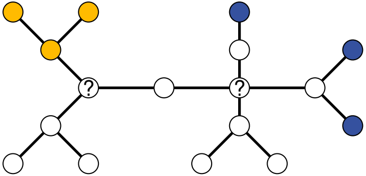


Some Practical Motivations:

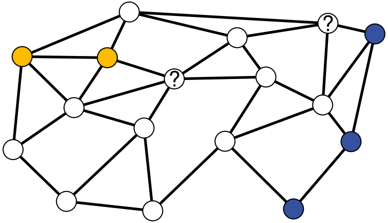
Some Practical Motivations:



Trees

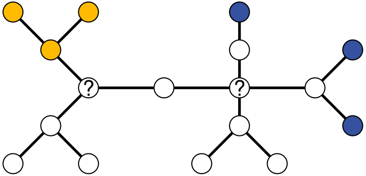


Trees

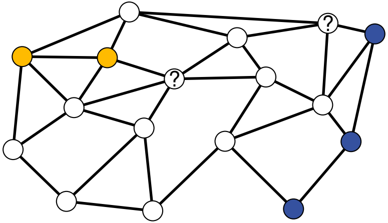


Graphs (e.g. molecule, social graph, . . .)

Some Practical Motivations:



Trees



Graphs (e.g. molecule, social graph, . . .)

- requires **semantically** meaningful definition of closure system/operator and linkage function

Example: Shortest path closure in graphs

Example: Shortest path closure in graphs

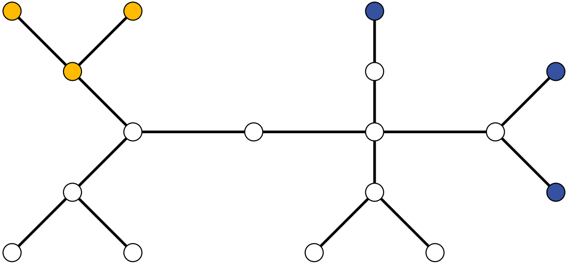
- A node set X is **closed** if it contains all shortest paths between nodes in X.

Example: Shortest path closure in graphs

- A node set X is **closed** if it contains all shortest paths between nodes in X.
- The **closure** $c(Y)$ is the smallest closed set containing Y.

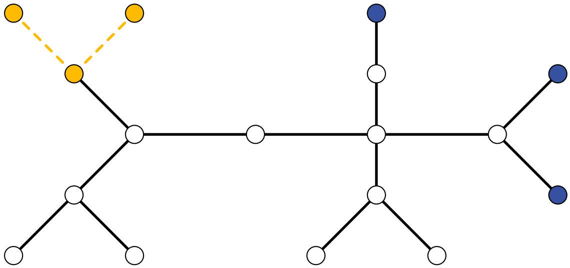
Example: Shortest path closure in graphs

- A node set X is **closed** if it contains all shortest paths between nodes in X.
- The **closure** $c(Y)$ is the smallest closed set containing Y.



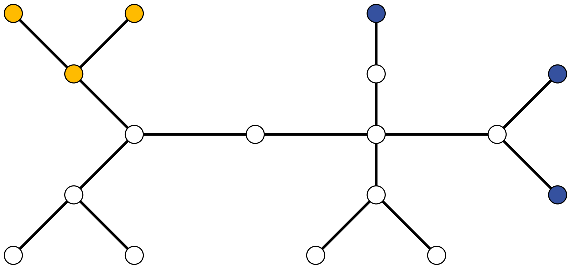
Example: Shortest path closure in graphs

- A node set X is **closed** if it contains all shortest paths between nodes in X.
- The **closure** $c(Y)$ is the smallest closed set containing Y.



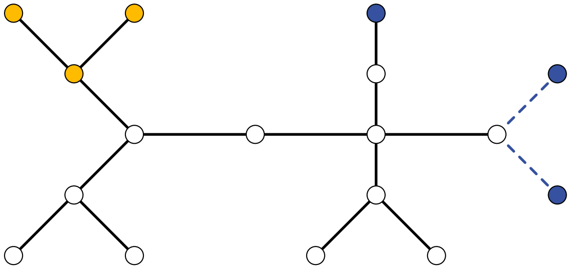
Example: Shortest path closure in graphs

- A node set X is **closed** if it contains all shortest paths between nodes in X.
- The **closure** $c(Y)$ is the smallest closed set containing Y.



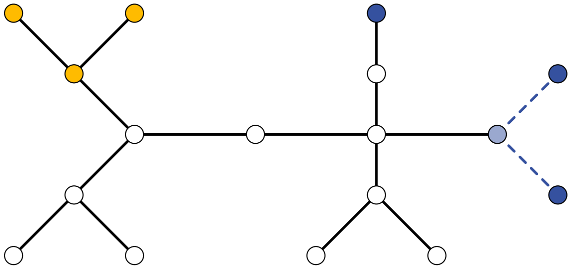
Example: Shortest path closure in graphs

- A node set X is **closed** if it contains all shortest paths between nodes in X.
- The **closure** $c(Y)$ is the smallest closed set containing Y.



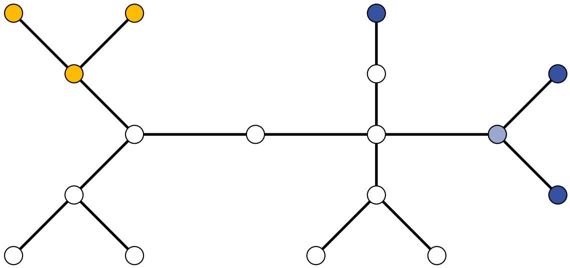
Example: Shortest path closure in graphs

- A node set X is **closed** if it contains all shortest paths between nodes in X.
- The **closure** $c(Y)$ is the smallest closed set containing Y.



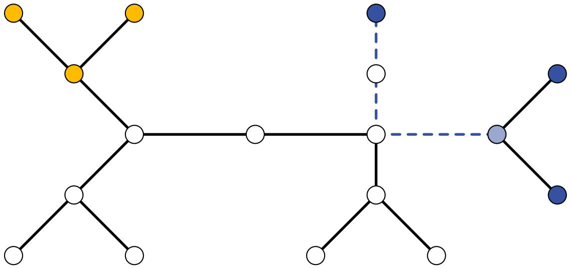
Example: Shortest path closure in graphs

- A node set X is **closed** if it contains all shortest paths between nodes in X.
- The **closure** $c(Y)$ is the smallest closed set containing Y.



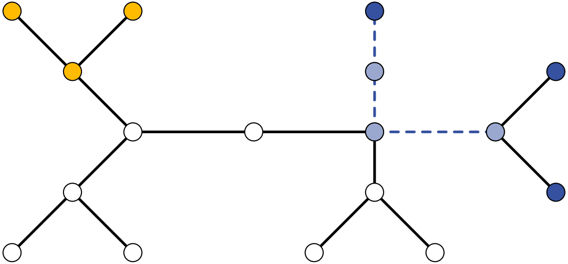
Example: Shortest path closure in graphs

- A node set X is **closed** if it contains all shortest paths between nodes in X.
- The **closure** $c(Y)$ is the smallest closed set containing Y.



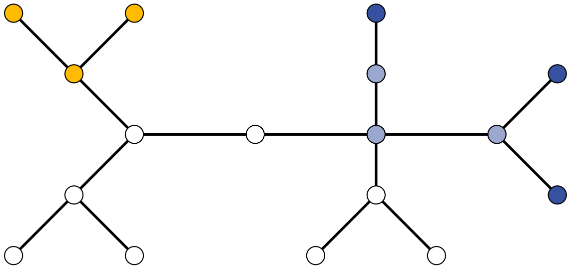
Example: Shortest path closure in graphs

- A node set X is **closed** if it contains all shortest paths between nodes in X.
- The **closure** $c(Y)$ is the smallest closed set containing Y.



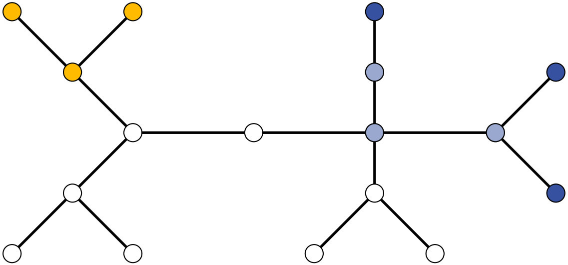
Example: Shortest path closure in graphs

- A node set X is **closed** if it contains all shortest paths between nodes in X.
- The **closure** $c(Y)$ is the smallest closed set containing Y.



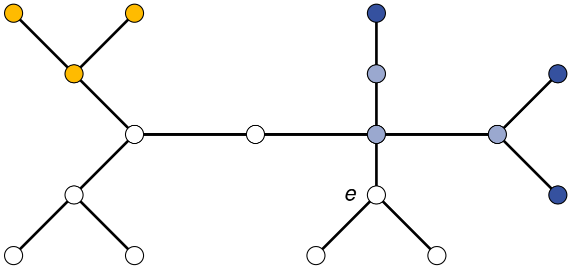
Example: Shortest path closure in graphs

- A node set X is **closed** if it contains all shortest paths between nodes in X.
- The **closure** $c(Y)$ is the smallest closed set containing Y.
- The **linkage** $l(X, y)$ is the minimum length of a shortest path from node y to node set X



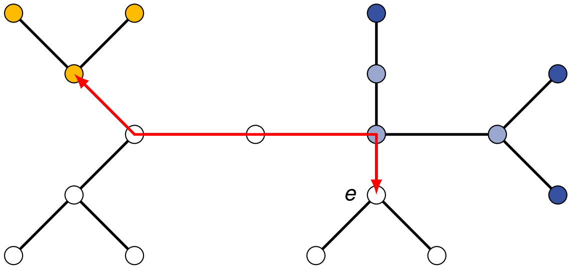
Example: Shortest path closure in graphs

- A node set X is **closed** if it contains all shortest paths between nodes in X.
- The **closure** $c(Y)$ is the smallest closed set containing Y.
- The **linkage** $l(X, y)$ is the minimum length of a shortest path from node y to node set X , e.g. $l($◯$, e) =$



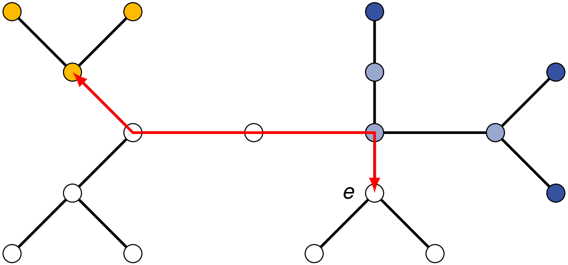
Example: Shortest path closure in graphs

- A node set X is **closed** if it contains all shortest paths between nodes in X.
- The **closure** $c(Y)$ is the smallest closed set containing Y.
- The **linkage** $l(X, y)$ is the minimum length of a shortest path from node y to node set X , e.g. $l($◯$, e) =$



Example: Shortest path closure in graphs

- A node set X is **closed** if it contains all shortest paths between nodes in X.
- The **closure** $c(Y)$ is the smallest closed set containing Y.
- The **linkage** $l(X, y)$ is the minimum length of a shortest path from node y to node set X , e.g. $l(\bigcirc, e) = 4$.



Maximum Margin considerations I

Extend linkage functions to sets:

Linkage from X to Y:

$$l(X, Y) := \min_{y \in Y} l(X, y)$$

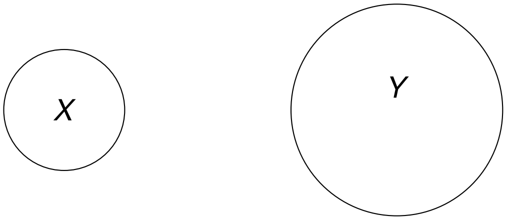
Maximum Margin considerations I

Extend linkage functions to sets:

Linkage from X to Y:

$$l(X, Y) := \min_{y \in Y} l(X, y)$$

Example:



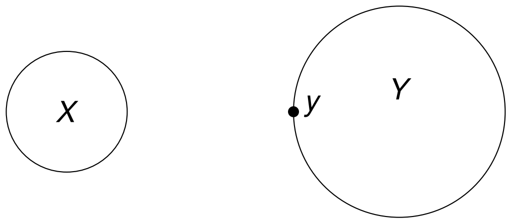
Maximum Margin considerations I

Extend linkage functions to sets:

Linkage from X to Y:

$$l(X, Y) := \min_{y \in Y} l(X, y)$$

Example:



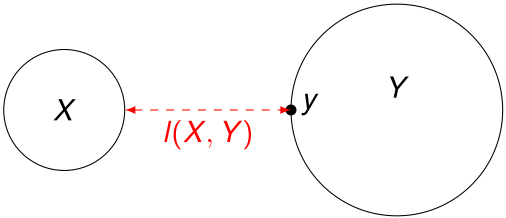
Maximum Margin considerations I

Extend linkage functions to sets:

Linkage from X to Y:

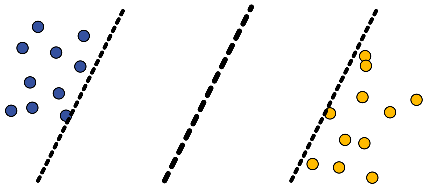
$$l(X, Y) := \min_{y \in Y} l(X, y)$$

Example:



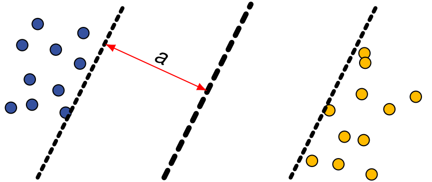
Maximum Margin considerations II

SVM:



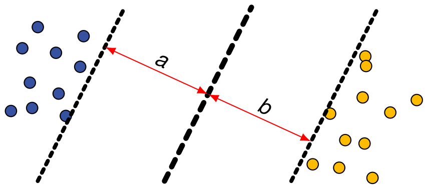
Maximum Margin considerations II

SVM:



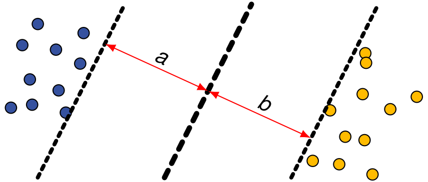
Maximum Margin considerations II

SVM:



Maximum Margin considerations II

SVM:



Maximum-Margin Property:

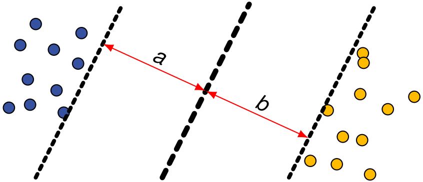
$$a + b \text{ is } \textit{maximal} \text{ and } a = b$$

$$\Leftrightarrow$$

$$\min(a, b) \text{ is } \textit{maximal}$$

Maximum Margin considerations II

SVM:



Finite Closure Systems:

Maximum-Margin Property:

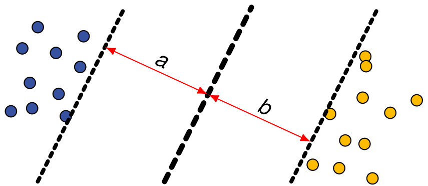
$$a + b \text{ is } \textit{maximal} \text{ and } a = b$$

$$\Leftrightarrow$$

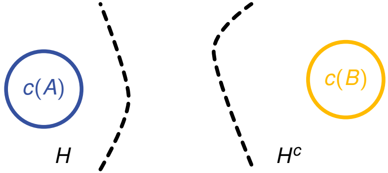
$$\min(a, b) \text{ is } \textit{maximal}$$

Maximum Margin considerations II

SVM:



Finite Closure Systems:



Maximum-Margin Property:

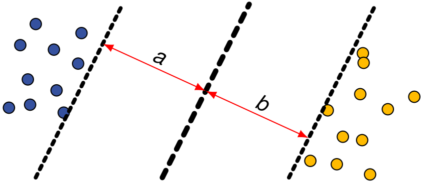
$$a + b \text{ is } \textit{maximal} \text{ and } a = b$$

$$\Leftrightarrow$$

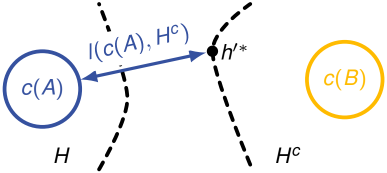
$$\min(a, b) \text{ is } \textit{maximal}$$

Maximum Margin considerations II

SVM:



Finite Closure Systems:



Maximum-Margin Property:

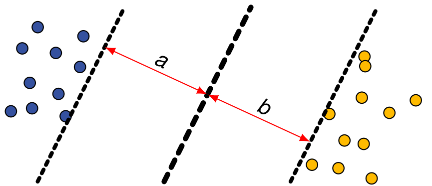
$$a + b \text{ is } \textit{maximal} \text{ and } a = b$$

$$\Leftrightarrow$$

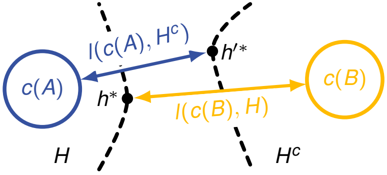
$$\min(a, b) \text{ is } \textit{maximal}$$

Maximum Margin considerations II

SVM:



Finite Closure Systems:



Maximum-Margin Property:

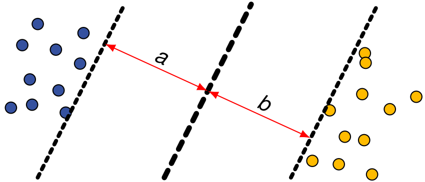
$$a + b \text{ is } \textit{maximal} \text{ and } a = b$$

$$\Leftrightarrow$$

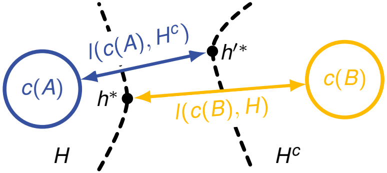
$$\min(a, b) \text{ is } \textit{maximal}$$

Maximum Margin considerations II

SVM:



Finite Closure Systems:



Maximum-Margin Property:

$$a + b \text{ is } \textit{maximal} \text{ and } a = b$$

$$\Leftrightarrow$$

$$\min(a, b) \text{ is } \textit{maximal}$$

Maximum-Margin Property:

$$\min(\longleftrightarrow, \longleftrightarrow) \text{ is } \textit{maximal}$$

Margins in finite closure systems

Margin

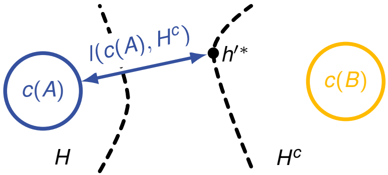
Given $A, B \subseteq E$, the *margin* for H and H^c is defined by

$$\mu_{H,H^c}(A, B) := \min\{l(c(A), H^c), l(c(B), H)\} \ .$$

Margin

Given $A, B \subseteq E$, the *margin* for H and H^c is defined by

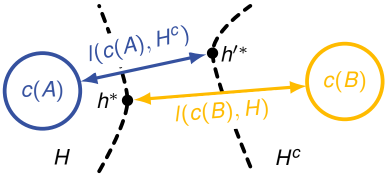
$$\mu_{H,H^c}(A,B) := \min\{l(c(A),H^c), l(c(B),H)\} \ .$$



Margin

Given $A, B \subseteq E$, the *margin* for H and H^c is defined by

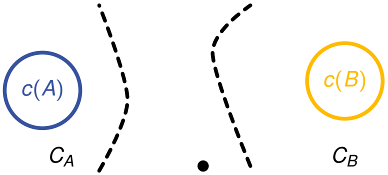
$$\mu_{H,H^c}(A,B) := \min\{l(c(A),H^c), l(c(B),H)\} \ .$$



Margin

Given $A, B \subseteq E$, the *margin* for C_A and C_B is defined by

$$\mu_{C_A,C_B}(A, B) := \min\{l(c(A), C_A^c), l(c(B), C_B^c)\} \ .$$



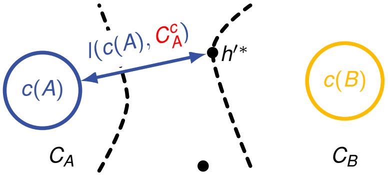
- the margin for disjoint closed sets $C_A \supseteq A$, $C_B \subseteq B$ can be defined similarly

Margins in finite closure systems

Margin

Given $A, B \subseteq E$, the *margin* for C_A and C_B is defined by

$$\mu_{C_A,C_B}(A,B) := \min\{l(c(A), C_A^c), l(c(B), C_B^c)\} \ .$$



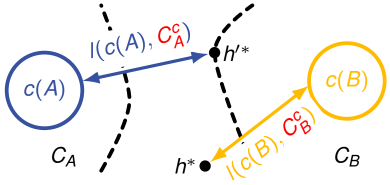
- the margin for disjoint closed sets $C_A \supseteq A$, $C_B \subseteq B$ can be defined similarly

Margins in finite closure systems

Margin

Given $A, B \subseteq E$, the *margin* for C_A and C_B is defined by

$$\mu_{C_A,C_B}(A,B) := \min\{l(c(A), C_A^c), l(c(B), C_B^c)\} \ .$$



- the margin for disjoint closed sets $C_A \supseteq A$, $C_B \subseteq B$ can be defined similarly

Problem Definition

Maximum Margin Half-Space Separation Problem:

Given a monotone linkage closure system $(E, \mathcal{C}, l)$ and two sets $A, B \subseteq E$, **find** a half-space H which **maximizes** the margin, i.e.

$$H = \underset{H_1, H_1^c \in \mathcal{C}_c}{\operatorname{argmax}} \; \mu_{H_1, H_1^c}(A, B)$$

if A, B are half-space separable; o/w return "NO".

Problem Definition

Maximum Margin Half-Space Separation Problem:

Given a monotone linkage closure system $(E, \mathcal{C}, l)$ and two sets $A, B \subseteq E$, **find** a half-space H which **maximizes** the margin, i.e.

$$H = \underset{H_1, H_1^c \in \mathcal{C}_c}{\operatorname{argmax}} \mu_{H_1, H_1^c}(A, B)$$

if A, B are half-space separable; o/w return "NO".

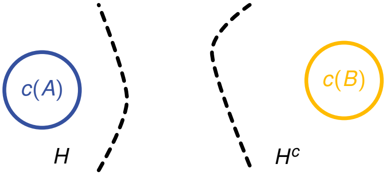
Problem Definition

Maximum Margin Half-Space Separation Problem:

Given a monotone linkage closure system $(E, \mathcal{C}, l)$ and two sets $A, B \subseteq E$, **find** a half-space H which **maximizes** the margin, i.e.

$$H = \underset{H_1, H_1^c \in \mathcal{C}_c}{\operatorname{argmax}} \mu_{H_1, H_1^c}(A, B)$$

if A, B are half-space separable; o/w return "NO".



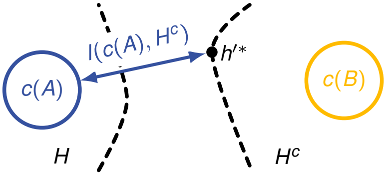
Problem Definition

Maximum Margin Half-Space Separation Problem:

Given a monotone linkage closure system $(E, \mathcal{C}, l)$ and two sets $A, B \subseteq E$, **find** a half-space H which **maximizes** the margin, i.e.

$$H = \underset{H_1, H_1^c \in \mathcal{C}_c}{\operatorname{argmax}} \; \mu_{H_1, H_1^c}(A, B)$$

if A, B are half-space separable; o/w return "NO".



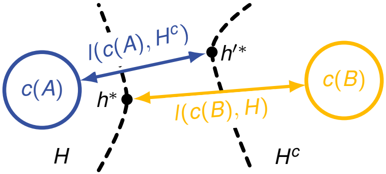
Problem Definition

Maximum Margin Half-Space Separation Problem:

Given a monotone linkage closure system $(E, \mathcal{C}, l)$ and two sets $A, B \subseteq E$, **find** a half-space H which **maximizes** the margin, i.e.

$$H = \underset{H_1, H_1^c \in \mathcal{C}_c}{\operatorname{argmax}} \; \mu_{H_1, H_1^c}(A, B)$$

if A, B are half-space separable; o/w return "NO".



Maximum Margin Half-Spaces

Problem: In general it is **hard** to find even separating half-spaces!

Maximum Margin Half-Spaces

Problem: In general it is **hard** to find even separating half-spaces!

Solutions:

Maximum Margin Half-Spaces

Problem: In general it is **hard** to find even separating half-spaces!

Solutions:

1. restrict to Kakutani closure systems (e.g. trees, distributive lattices)

Kakutani closure system:

All disjoint closed sets are half-space separable.

Maximum Margin Half-Spaces

Problem: In general it is **hard** to find even separating half-spaces!

Solutions:

1. restrict to Kakutani closure systems (e.g. trees, distributive lattices)
2. relax the problem to maximal disjoint closed sets

Kakutani closure system:

All disjoint closed sets are half-space separable.

Maximum Margin Disjoint Closed Sets

Maximum-Margin Closed Set Separation Problem:

Given non-empty subsets A, B of E, *find* disjoint closed sets $C_A, C_B \in \mathcal{C}_c$ with $A \subseteq C_A, B \subseteq C_B$ that **maximize** the margin if $c(A) \cap c(B) = \emptyset$; o/w return "No".

Maximum Margin Disjoint Closed Sets

Maximum-Margin Closed Set Separation Problem:

Given non-empty subsets A, B of E, *find* disjoint closed sets $C_A, C_B \in \mathcal{C}_c$ with $A \subseteq C_A, B \subseteq C_B$ that **maximize** the margin if $c(A) \cap c(B) = \emptyset$; o/w return "No".

Result: Very efficient algorithm.

Maximum Margin Disjoint Closed Sets

Maximum-Margin Closed Set Separation Problem:

Given non-empty subsets A, B of E, *find* disjoint closed sets $C_A, C_B \in \mathcal{C}_c$ with $A \subseteq C_A, B \subseteq C_B$ that **maximize** the margin if $c(A) \cap c(B) = \emptyset$; o/w return "No".

Result: Very efficient algorithm.

- requires only the ground set the closure operator and the linkage function

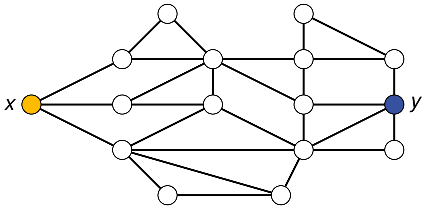
Maximum Margin Disjoint Closed Sets

Maximum-Margin Closed Set Separation Problem:

Given non-empty subsets A, B of E, *find* disjoint closed sets $C_A, C_B \in \mathcal{C}_c$ with $A \subseteq C_A, B \subseteq C_B$ that **maximize** the margin if $c(A) \cap c(B) = \emptyset$; o/w return "No".

Result: Very efficient algorithm.

- requires only the ground set the closure operator and the linkage function

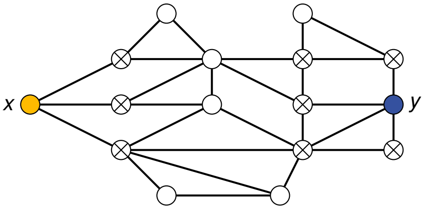


Maximum-Margin Closed Set Separation Problem:

Given non-empty subsets A, B of E, *find* disjoint closed sets $C_A, C_B \in \mathcal{C}_c$ with $A \subseteq C_A, B \subseteq C_B$ that **maximize** the margin if $c(A) \cap c(B) = \emptyset$; o/w return "No".

Result: Very efficient algorithm.

- requires only the ground set the closure operator and the linkage function



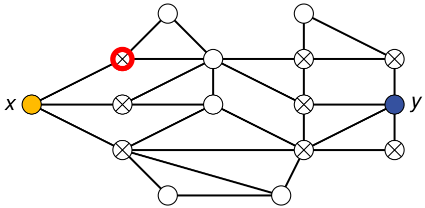
Maximum Margin Disjoint Closed Sets

Maximum-Margin Closed Set Separation Problem:

Given non-empty subsets A, B of E, *find* disjoint closed sets $C_A, C_B \in \mathcal{C}_c$ with $A \subseteq C_A, B \subseteq C_B$ that **maximize** the margin if $c(A) \cap c(B) = \emptyset$; o/w return "No".

Result: Very efficient algorithm.

- requires only the ground set the closure operator and the linkage function



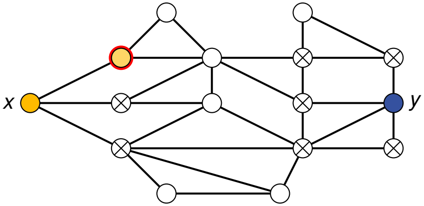
Maximum Margin Disjoint Closed Sets

Maximum-Margin Closed Set Separation Problem:

Given non-empty subsets A, B of E, *find* disjoint closed sets $C_A, C_B \in \mathcal{C}_c$ with $A \subseteq C_A, B \subseteq C_B$ that **maximize** the margin if $c(A) \cap c(B) = \emptyset$; o/w return "No".

Result: Very efficient algorithm.

- requires only the ground set the closure operator and the linkage function



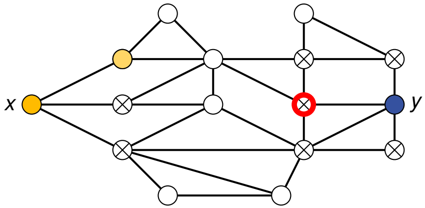
Maximum Margin Disjoint Closed Sets

Maximum-Margin Closed Set Separation Problem:

Given non-empty subsets A, B of E, *find* disjoint closed sets $C_A, C_B \in \mathcal{C}_c$ with $A \subseteq C_A, B \subseteq C_B$ that **maximize** the margin if $c(A) \cap c(B) = \emptyset$; o/w return "No".

Result: Very efficient algorithm.

- requires only the ground set the closure operator and the linkage function



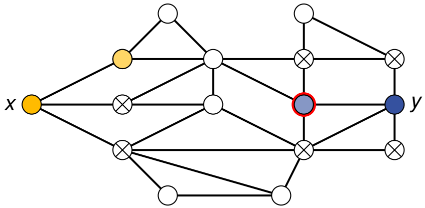
Maximum Margin Disjoint Closed Sets

Maximum-Margin Closed Set Separation Problem:

Given non-empty subsets A, B of E, *find* disjoint closed sets $C_A, C_B \in \mathcal{C}_c$ with $A \subseteq C_A, B \subseteq C_B$ that **maximize** the margin if $c(A) \cap c(B) = \emptyset$; o/w return “No”.

Result: Very efficient algorithm.

- requires only the ground set the closure operator and the linkage function



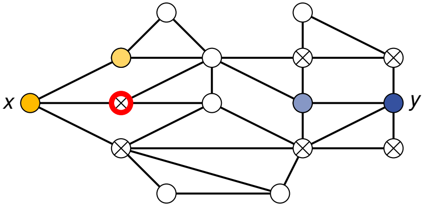
Maximum Margin Disjoint Closed Sets

Maximum-Margin Closed Set Separation Problem:

Given non-empty subsets A, B of E, *find* disjoint closed sets $C_A, C_B \in \mathcal{C}_c$ with $A \subseteq C_A, B \subseteq C_B$ that **maximize** the margin if $c(A) \cap c(B) = \emptyset$; o/w return "No".

Result: Very efficient algorithm.

- requires only the ground set the closure operator and the linkage function



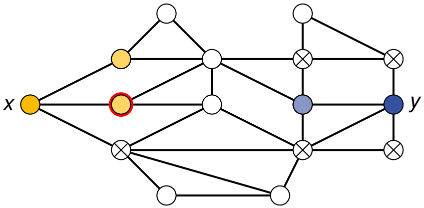
Maximum Margin Disjoint Closed Sets

Maximum-Margin Closed Set Separation Problem:

Given non-empty subsets A, B of E, *find* disjoint closed sets $C_A, C_B \in \mathcal{C}_c$ with $A \subseteq C_A, B \subseteq C_B$ that **maximize** the margin if $c(A) \cap c(B) = \emptyset$; o/w return "No".

Result: Very efficient algorithm.

- requires only the ground set the closure operator and the linkage function



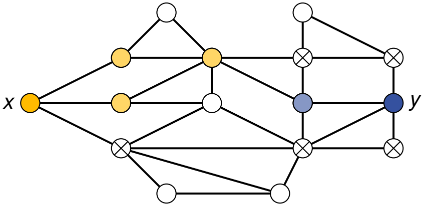
Maximum Margin Disjoint Closed Sets

Maximum-Margin Closed Set Separation Problem:

Given non-empty subsets A, B of E, *find* disjoint closed sets $C_A, C_B \in \mathcal{C}_c$ with $A \subseteq C_A, B \subseteq C_B$ that **maximize** the margin if $c(A) \cap c(B) = \emptyset$; o/w return "No".

Result: Very efficient algorithm.

- requires only the ground set the closure operator and the linkage function



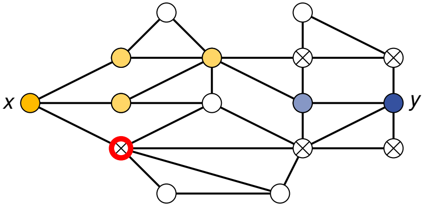
Maximum Margin Disjoint Closed Sets

Maximum-Margin Closed Set Separation Problem:

Given non-empty subsets A, B of E, *find* disjoint closed sets $C_A, C_B \in \mathcal{C}_c$ with $A \subseteq C_A, B \subseteq C_B$ that **maximize** the margin if $c(A) \cap c(B) = \emptyset$; o/w return “No”.

Result: Very efficient algorithm.

- requires only the ground set the closure operator and the linkage function



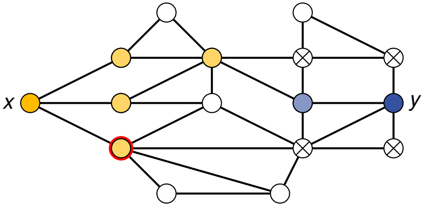
Maximum Margin Disjoint Closed Sets

Maximum-Margin Closed Set Separation Problem:

Given non-empty subsets A, B of E, *find* disjoint closed sets $C_A, C_B \in \mathcal{C}_c$ with $A \subseteq C_A, B \subseteq C_B$ that **maximize** the margin if $c(A) \cap c(B) = \emptyset$; o/w return "No".

Result: Very efficient algorithm.

- requires only the ground set the closure operator and the linkage function



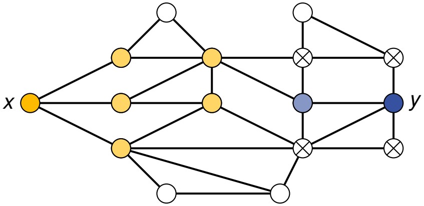
Maximum Margin Disjoint Closed Sets

Maximum-Margin Closed Set Separation Problem:

Given non-empty subsets A, B of E, *find* disjoint closed sets $C_A, C_B \in \mathcal{C}_c$ with $A \subseteq C_A, B \subseteq C_B$ that **maximize** the margin if $c(A) \cap c(B) = \emptyset$; o/w return “No”.

Result: Very efficient algorithm.

- requires only the ground set the closure operator and the linkage function



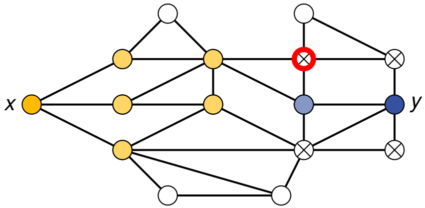
Maximum Margin Disjoint Closed Sets

Maximum-Margin Closed Set Separation Problem:

Given non-empty subsets A, B of E, *find* disjoint closed sets $C_A, C_B \in \mathcal{C}_c$ with $A \subseteq C_A$, $B \subseteq C_B$ that **maximize** the margin if $c(A) \cap c(B) = \emptyset$; o/w return "No".

Result: Very efficient algorithm.

- requires only the ground set the closure operator and the linkage function



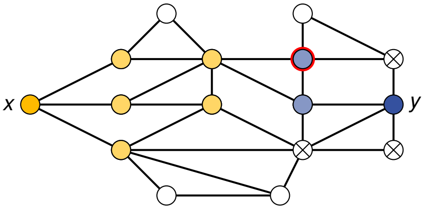
Maximum Margin Disjoint Closed Sets

Maximum-Margin Closed Set Separation Problem:

Given non-empty subsets A, B of E, *find* disjoint closed sets $C_A, C_B \in \mathcal{C}_c$ with $A \subseteq C_A, B \subseteq C_B$ that **maximize** the margin if $c(A) \cap c(B) = \emptyset$; o/w return "No".

Result: Very efficient algorithm.

- requires only the ground set the closure operator and the linkage function



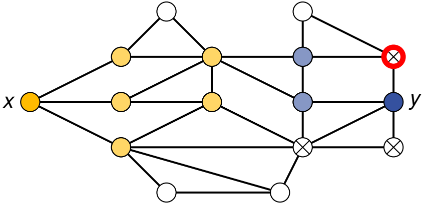
Maximum Margin Disjoint Closed Sets

Maximum-Margin Closed Set Separation Problem:

Given non-empty subsets A, B of E, *find* disjoint closed sets $C_A, C_B \in \mathcal{C}_c$ with $A \subseteq C_A, B \subseteq C_B$ that **maximize** the margin if $c(A) \cap c(B) = \emptyset$; o/w return “No”.

Result: Very efficient algorithm.

- requires only the ground set the closure operator and the linkage function



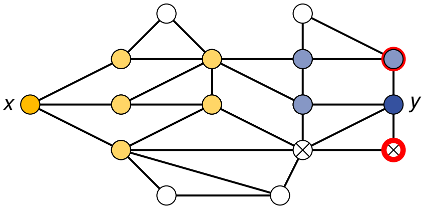
Maximum Margin Disjoint Closed Sets

Maximum-Margin Closed Set Separation Problem:

Given non-empty subsets A, B of E, *find* disjoint closed sets $C_A, C_B \in \mathcal{C}_c$ with $A \subseteq C_A, B \subseteq C_B$ that **maximize** the margin if $c(A) \cap c(B) = \emptyset$; o/w return “No”.

Result: Very efficient algorithm.

- requires only the ground set the closure operator and the linkage function



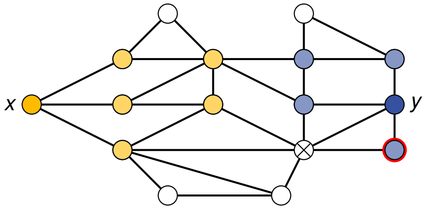
Maximum Margin Disjoint Closed Sets

Maximum-Margin Closed Set Separation Problem:

Given non-empty subsets A, B of E, *find* disjoint closed sets $C_A, C_B \in \mathcal{C}_c$ with $A \subseteq C_A, B \subseteq C_B$ that **maximize** the margin if $c(A) \cap c(B) = \emptyset$; o/w return "No".

Result: Very efficient algorithm.

- requires only the ground set the closure operator and the linkage function



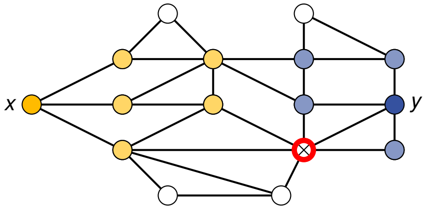
Maximum Margin Disjoint Closed Sets

Maximum-Margin Closed Set Separation Problem:

Given non-empty subsets A, B of E, *find* disjoint closed sets $C_A, C_B \in \mathcal{C}_c$ with $A \subseteq C_A, B \subseteq C_B$ that **maximize** the margin if $c(A) \cap c(B) = \emptyset$; o/w return "No".

Result: Very efficient algorithm.

- requires only the ground set the closure operator and the linkage function

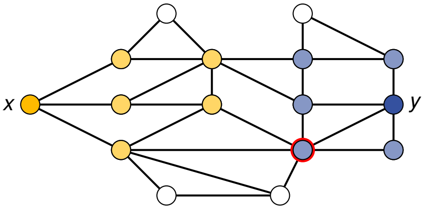


Maximum-Margin Closed Set Separation Problem:

Given non-empty subsets A, B of E, *find* disjoint closed sets $C_A, C_B \in \mathcal{C}_c$ with $A \subseteq C_A, B \subseteq C_B$ that **maximize** the margin if $c(A) \cap c(B) = \emptyset$; o/w return “No”.

Result: Very efficient algorithm.

- requires only the ground set the closure operator and the linkage function



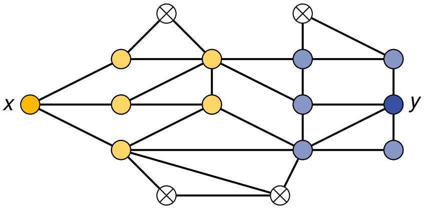
Maximum Margin Disjoint Closed Sets

Maximum-Margin Closed Set Separation Problem:

Given non-empty subsets A, B of E, *find* disjoint closed sets $C_A, C_B \in \mathcal{C}_c$ with $A \subseteq C_A, B \subseteq C_B$ that **maximize** the margin if $c(A) \cap c(B) = \emptyset$; o/w return "No".

Result: Very efficient algorithm.

- requires only the ground set the closure operator and the linkage function



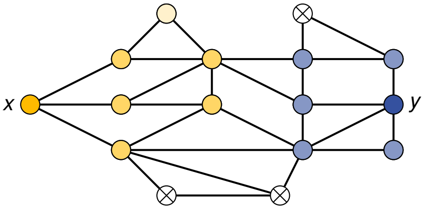
Maximum Margin Disjoint Closed Sets

Maximum-Margin Closed Set Separation Problem:

Given non-empty subsets A, B of E, *find* disjoint closed sets $C_A, C_B \in \mathcal{C}_c$ with $A \subseteq C_A, B \subseteq C_B$ that **maximize** the margin if $c(A) \cap c(B) = \emptyset$; o/w return "No".

Result: Very efficient algorithm.

- requires only the ground set the closure operator and the linkage function



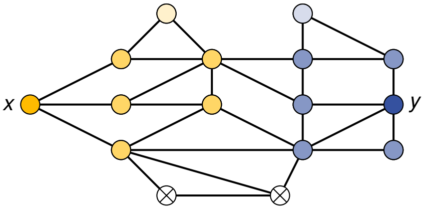
Maximum Margin Disjoint Closed Sets

Maximum-Margin Closed Set Separation Problem:

Given non-empty subsets A, B of E, *find* disjoint closed sets $C_A, C_B \in \mathcal{C}_c$ with $A \subseteq C_A, B \subseteq C_B$ that **maximize** the margin if $c(A) \cap c(B) = \emptyset$; o/w return “No”.

Result: Very efficient algorithm.

- requires only the ground set the closure operator and the linkage function



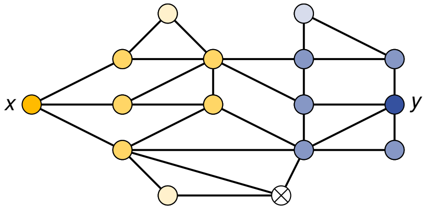
Maximum Margin Disjoint Closed Sets

Maximum-Margin Closed Set Separation Problem:

Given non-empty subsets A, B of E, *find* disjoint closed sets $C_A, C_B \in \mathcal{C}_c$ with $A \subseteq C_A, B \subseteq C_B$ that **maximize** the margin if $c(A) \cap c(B) = \emptyset$; o/w return "No".

Result: Very efficient algorithm.

- requires only the ground set the closure operator and the linkage function



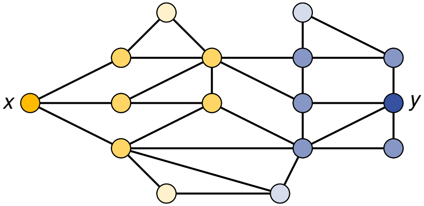
Maximum Margin Disjoint Closed Sets

Maximum-Margin Closed Set Separation Problem:

Given non-empty subsets A, B of E, *find* disjoint closed sets $C_A, C_B \in \mathcal{C}_c$ with $A \subseteq C_A, B \subseteq C_B$ that **maximize** the margin if $c(A) \cap c(B) = \emptyset$; o/w return "No".

Result: Very efficient algorithm.

- requires only the ground set the closure operator and the linkage function



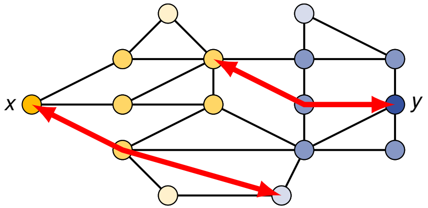
Maximum Margin Disjoint Closed Sets

Maximum-Margin Closed Set Separation Problem:

Given non-empty subsets A, B of E, *find* disjoint closed sets $C_A, C_B \in \mathcal{C}_c$ with $A \subseteq C_A, B \subseteq C_B$ that **maximize** the margin if $c(A) \cap c(B) = \emptyset$; o/w return "No".

Result: Very efficient algorithm.

- requires only the ground set the closure operator and the linkage function



Results

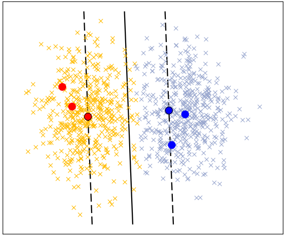
Theorem: The algorithm solves the Maximum Margin Closed Set Separation problem by calling the closure operator at most $2|E| - 2$ times and the linkage function at most $2|E|$ times.

Results

Theorem: The algorithm solves the Maximum Margin Closed Set Separation problem by calling the closure operator at most $2|E| - 2$ times and the linkage function at most $2|E|$ times.

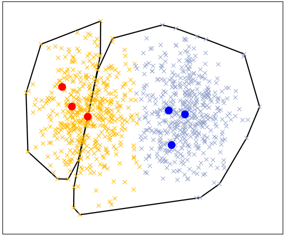
Corollary: The algorithm solves the Maximum Margin Half-Space Separation Problem if the closure system is Kakutani.

Experiments (Finite Point Sets)



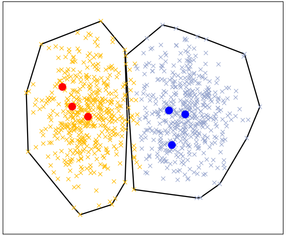
Support Vector Machine

(Vapnik, 1992)



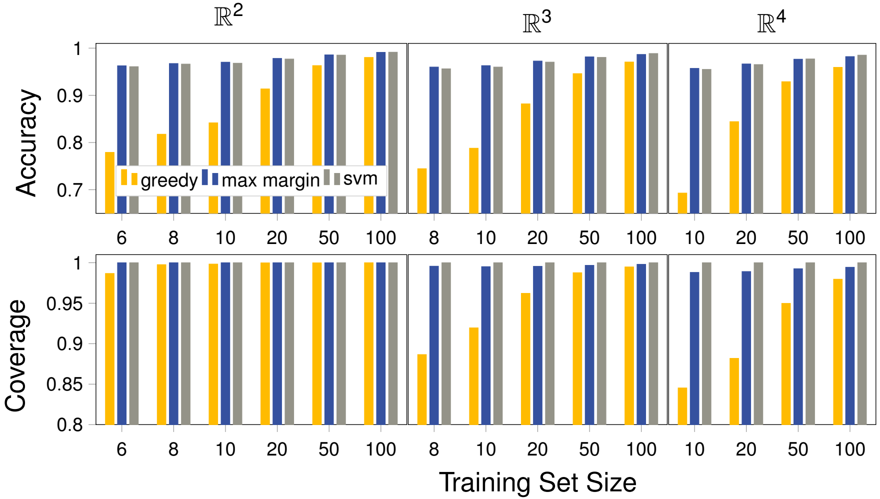
Greedy Algorithm

(ECML/PKDD 2019)



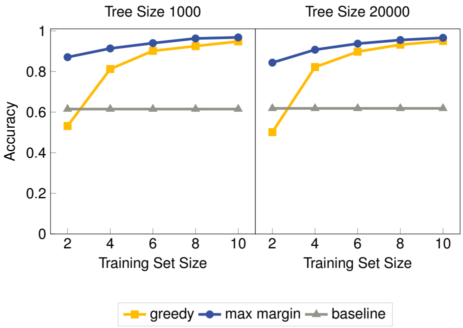
Maximum Margin Algorithm
(this work)

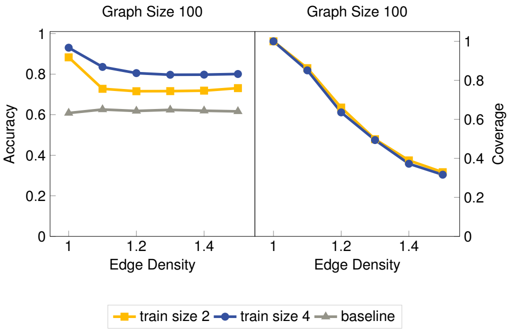
Experiments (Finite Point Sets)



svm: (Vapnik, 1992), greedy: (ECML/PDKK 2019), max margin: (this work)

Experiments (Trees and Graphs)





Maximum Half-Space separation on trees

Maximum closed set separation on small graphs

Summary

This work: Adaption of maximum margin separation from $\mathbb{R}^d$ to finite closure systems.

- Maximum margin concept **differs** from that in $\mathbb{R}^d$

Summary

This work: Adaption of maximum margin separation from $\mathbb{R}^d$ to finite closure systems.

- Maximum margin concept **differs** from that in $\mathbb{R}^d$
    - equivalence of maximum margin definition does **not** hold

Summary

This work: Adaption of maximum margin separation from $\mathbb{R}^d$ to finite closure systems.

- Maximum margin concept **differs** from that in $\mathbb{R}^d$
  - equivalence of maximum margin definition does **not** hold
  - maximum margin separation is **not unique**

Summary

This work: Adaption of maximum margin separation from $\mathbb{R}^d$ to finite closure systems.

- Maximum margin concept **differs** from that in $\mathbb{R}^d$
  - equivalence of maximum margin definition does **not** hold
  - maximum margin separation is **not unique**
- **no** specific structure is utilized (e.g. metric, vector space)

Summary

This work: Adaption of maximum margin separation from $\mathbb{R}^d$ to finite closure systems.

- Maximum margin concept **differs** from that in $\mathbb{R}^d$
  - equivalence of maximum margin definition does **not** hold
  - maximum margin separation is **not unique**
- **no** specific structure is utilized (e.g. metric, vector space)
- **monotone linkage** function provide great generality

Summary

This work: Adaption of maximum margin separation from $\mathbb{R}^d$ to finite closure systems.

- Maximum margin concept **differs** from that in $\mathbb{R}^d$
  - equivalence of maximum margin definition does **not** hold
  - maximum margin separation is **not unique**
- **no** specific structure is utilized (e.g. metric, vector space)
- **monotone linkage** function provide great generality
- Maximum Margin Half-Space Separation is in P for Kakutani closure systems (e.g. trees, distributive lattices)

Summary

This work: Adaption of maximum margin separation from $\mathbb{R}^d$ to finite closure systems.

- Maximum margin concept **differs** from that in $\mathbb{R}^d$
  - equivalence of maximum margin definition does **not** hold
  - maximum margin separation is **not unique**
- **no** specific structure is utilized (e.g. metric, vector space)
- **monotone linkage** function provide great generality
- Maximum Margin Half-Space Separation is in P for Kakutani closure systems (e.g. trees, distributive lattices)
- Maximum Margin Closed Set Separation is efficient solvable

Summary

This work: Adaption of maximum margin separation from $\mathbb{R}^d$ to finite closure systems.

- Maximum margin concept **differs** from that in $\mathbb{R}^d$
  - equivalence of maximum margin definition does **not** hold
  - maximum margin separation is **not unique**
- **no** specific structure is utilized (e.g. metric, vector space)
- **monotone linkage** function provide great generality
- Maximum Margin Half-Space Separation is in *P* for Kakutani closure systems (e.g. trees, distributive lattices)
- Maximum Margin Closed Set Separation is efficient solvable
  - outperforms simple greedy algorithm (ECML/PKDD, 2019)

On-going/Future Work

- further applications to the special case of (distributive) lattices

On-going/Future Work

- further applications to the special case of (distributive) lattices
- studying uniqueness conditions for maximum margin half-space separation

On-going/Future Work

- further applications to the special case of (distributive) lattices
- studying uniqueness conditions for maximum margin half-space separation
- reduce complexity of the algorithm for known closure systems (e.g. trees, lattices,. . .)

On-going/Future Work

- further applications to the special case of (distributive) lattices
- studying uniqueness conditions for maximum margin half-space separation
- reduce complexity of the algorithm for known closure systems (e.g. trees, lattices,. . .)
- allow approximate/fuzzy solutions

Further Details

Come to the Question and Answer Session or contact me:

Florian Seiffarth
University of Bonn
seiffarth@cs.uni-bonn.de