A. Alemán

Introduction

Aprendizaje

y paradigmas

Aprendizaje

Aprendizaie N

Supervisado

Clustorin

Anlicacione

Charterian have

distancias

Clustering er

The Markov Cluster

Detección de

Preguntas

Clustering en Biología de Sistemas Máster de Bioinformática (UV)

Autor: Alejandro Alemán Ramos aaleman@cipf.es

Instituto de Genómica Computacional (CIPF)

12 de Diciembre de 2013

Supervisado

Definicio

Aplicacion

Clustering basado e

distancias

The Markov Cluste

Detección de

Pregunt

Índice

Introduction

Aprendizaje Tipos de Aprendizaje y paradigmas

2 Aprendizaje

Aprendizaje Supervisado Aprendizaje No Supervisado

3 Clustering

Definición Aplicaciones Clustering basado en distancias

4 Clustering en Grafos The Markov Cluster Algorithm

- 6 Detección de comunidades
- **6** Preguntas

A. Alemán

Introductio Aprendizaie

Tipos de Aprendia

Aprendizaj

Aprendizaj

Aprendizaje N

Supervisado

Definici

Aplicacione

Clustering basado

Clustering en

The Markov Cluste

Detección de comunidades

Preguntas

- Introduction
 Aprendizaje
 Tipos de Aprendizaje y paradigmas
- 2 Aprendizaje Aprendizaje Supervisado Aprendizaje No Supervisado
 - 3 Clustering Definición Aplicaciones Clustering basado en distancias
- 4 Clustering en Grafos
 The Markov Cluster Algorithm
- 6 Detección de comunidades
- 6 Preguntas

A. Alemán

Introductio

Aprendizaje

Tipos de Aprendia

Aprendiz

Aprendizaje

Anrendizaie I

Aprendizaje l Supervisado

Clusterin

Dofinici

Anlicacione

Clustering basado

Clustering er

The Markov Cluste

Detección de

Preguntas

Aprendizaje

Definiciones de aprendizaje:

- Cualquier cambio en un sistema que le permita realizar la misma tarea de manera más eficiente la próxima vez (H. Simon)
- Modificar la representación del mundo que se está percibiendo (R. Michalski)
- Realizar cambios útiles en nuestras mentes (M. Minsky)

A. Alemán

Introductio

Aprendizaje

lipos de Aprendia

Aprendiza

Aprendizaje Supervisado

Aprendizaje N

Supervisado

0.000

Definicio

Aplicaciones

Clustering basado e distancias

Clustering en

The Markov Cluste

Detección de

Pregunta:

- Cualquier cambio en un sistema que le permita realizar la misma tarea de manera más eficiente la próxima vez (H. Simon)
- Modificar la representación del mundo que se está percibiendo (R. Michalski)
- Realizar cambios útiles en nuestras mentes (M. Minsky)

A Alemán

Aprendizaie

- Cualquier cambio en un sistema que le permita realizar la misma tarea de manera más eficiente la próxima vez (H. Simon)
- Modificar la representación del mundo que se está percibiendo (R. Michalski)
- Realizar cambios útiles en nuestras mentes (M. Minsky)

A Alemán

Aprendizaie

- Cualquier cambio en un sistema que le permita realizar la misma tarea de manera más eficiente la próxima vez (H. Simon)
- Modificar la representación del mundo que se está percibiendo (R. Michalski)
- Realizar cambios útiles en nuestras mentes (M. Minsky)

A Alemán

Aprendizaie

Aprendizaje

- Cualquier cambio en un sistema que le permita realizar la misma tarea de manera más eficiente la próxima vez (H. Simon)
- Modificar la representación del mundo que se está percibiendo (R. Michalski)
- Realizar cambios útiles en nuestras mentes (M. Minsky)

Aprendizaje Automático: construir aplicaciones que mejoran automáticamente con la experiencia.

A. Alemán

Introduction

Tipos de Aprendizaje y paradigmas

Aprendizaj

, tprenaizaj

Aprendizaje N

Supervisado

Clusteri

Definició

Aplicaciones

Clustering basa

Clustering basado distancias

Clustering er Grafos

The Markov Cluster

Detección de

Preguntas

Tipos de Aprendizaje y paradigmas

Tipos de aprendizaje

Supervisado Se conocen los distintos valores de clasificación

- Paradigmas
 - Árboles de decisión
 - Aprendizaje de reglas
 - Redes neuronales
 - Aprendizaje bavesiano
 - Clasificación (Clústering)
 -

A. Alemán

Introduction

Tipos de Aprendizaje y paradigmas

Aprendizaj

A ----- di---i-

Anrendizaie N

Supervisado

D=6=1=14

Anlicaciones

Clustering basade

distancias

The Markov Cluste

Algorithm

Detección de

Pregunta

Tipos de Aprendizaje y paradigmas

Tipos de aprendizaje

Supervisado Se conocen los distintos valores de clasificación

- Paradigmas
 - Árboles de decisión
 - Aprendizaje de reglas
 - Redes neuronales
 - Aprendizaje bayesiano
 - Clasificación (Clústering)
 - ...

A. Alemán

Introduction

Tipos de Aprendizaje y paradigmas

Aprendizaj

Supervisado

Aprendizaje N Supervisado

Ciuste

Definicio

Aplicaciones

Clustering basado el distancias

Grafos
The Markov Clust

Algorithm

Detección de

Pregunta

Tipos de Aprendizaje y paradigmas

Tipos de aprendizaje

Supervisado Se conocen los distintos valores de clasificación

- Paradigmas
 - Árboles de decisión
 - Aprendizaje de reglas
 - Redes neuronales
 - Aprendizaie bavesiano
 - Clasificación (Clústering)
 - . . .

A. Alemán

Introduction

Tipos de Aprendizaje y paradigmas

Aprendizaj

Aprendizaje Supervisado

Aprendizaje N

Supervisado

Cidote

Definicio

Aplicaciones

Clustering basado (distancias

Grafos

Algorithm

Detección de comunidades

Preguntas

Tipos de Aprendizaje y paradigmas

Tipos de aprendizaje

Supervisado Se conocen los distintos valores de clasificación

- Paradigmas
 - Árboles de decisión
 - Aprendizaje de reglas
 - Redes neuronales
 - Aprendizaje bayesiano
 - Clasificación (Clústering)
 - . .

A. Alemán

Introduction

Tipos de Aprendizaje y paradigmas

Aprendizaj

Supervisado

Aprendizaio N

Aprendizaje I Supervisado

Clusteria

Definici

Aplicacione

Clustering basado er

Clustering en

The Markov Cluster Algorithm

Detección de

Preguntas

Tipos de Aprendizaje y paradigmas

Tipos de aprendizaje

Supervisado Se conocen los distintos valores de clasificación

- Paradigmas
 - Árboles de decisión
 - Aprendizaje de reglas
 - Redes neuronales
 - Aprendizaje bayesiano
 - Clasificación (Clústering)

A. Alemán

Introduction

Tipos de Aprendizaje y paradigmas

Aprendizaj

Aprendizaie N

Aprendizaje I Supervisado

Clusteri

Definici

Aplicacione

Clustering basado e

Grafos Grafos

The Markov Cluste Algorithm

Detección de comunidades

Preguntas

Tipos de Aprendizaje y paradigmas

Tipos de aprendizaje

Supervisado Se conocen los distintos valores de clasificación

- Paradigmas
 - Árboles de decisión
 - Aprendizaje de reglas
 - Redes neuronales
 - Aprendizaje bayesiano
 - Clasificación (Clústering)

A. Alemán

Introduction

Tipos de Aprendizaje y paradigmas

Aprendizaj

Aprendizaie N

Aprendizaje l' Supervisado

Clusterin

Definicio

Aplicaciones

Clustering basado e

Grafos
The Markov Cluste

Algorithm

Detección de comunidades

Pregunta

Tipos de Aprendizaje y paradigmas

Tipos de aprendizaje

Supervisado Se conocen los distintos valores de clasificación

- Paradigmas
 - Árboles de decisión
 - Aprendizaje de reglas
 - Redes neuronales
 - Aprendizaje bayesiano
 - Clasificación (Clústering)
 - . . .

A. Alemán

Introduction

Tipos de Aprendizaje y paradigmas

Aprendizaj

Supervisado

Aprendizaje N Supervisado

CI . .

Definicio

Aplicaciones

Clustering basado er

Clustering en Grafos

The Markov Cluster Algorithm

Detección de comunidades

Preguntas

Tipos de Aprendizaje y paradigmas

Tipos de aprendizaje

Supervisado Se conocen los distintos valores de clasificación

- Paradigmas
 - Árboles de decisión
 - Aprendizaje de reglas
 - Redes neuronales
 - Aprendizaje bayesiano
 - Clasificación (Clústering)
 - . . .

A Alemán

Tipos de Aprendizaie y paradigmas

Tipos de Aprendizaje y paradigmas

Tipos de aprendizaje

Supervisado Se conocen los distintos valores de clasificación

- Paradigmas
 - Árboles de decisión
 - Aprendizaje de reglas
 - Redes neuronales
 - Aprendizaje bavesiano
 - Clasificación (Clústering)

A. Alemán

Introduction

Tipos de Aprendizaje y paradigmas

Aprendizaj

Aprendizaie N

Aprendizaje N Supervisado

Cluste

Definició

Aplicaciones Clustering basado e

distancias

The Markov Cluste

Algorithm

Detección de comunidades

Preguntas

Tipos de Aprendizaje y paradigmas

Tipos de aprendizaje

Supervisado Se conocen los distintos valores de clasificación

- Paradigmas
 - Árboles de decisión
 - Aprendizaje de reglas
 - Redes neuronales
 - Aprendizaje bayesiano
 - Clasificación (Clústering)
 - . .

A Alemán

Aprendizaje

Introduction Tipos de Aprendizaje y paradigmas

2 Aprendizaje

Aprendizaje Supervisado Aprendizaje No Supervisado

- 4 Clustering en Grafos
- 6 Detección de comunidades.

A. Alemán

Introduction

Aprendizaje Tipos de Aprendiz

Aprendizaje Aprendizaje

Supervisado

Aprendizaje N

Supervisado

Ciusto

Definicio

Aplicaciones
Clustering basado er

distancias

The Markov Cluster

Detección de

Preguntas

- Obtener un conjunto de instancias positivas y negativas del concepto que queremos aprender.
- Analizar las características comunes de las instancias positivas que las diferencian de las instancias negativas.
- Buscar hipótesis expresables que discriminen entre el conjunto de instancias conocidas.
- Seleccionar la mejor hipótesis.

A. Alemán

Introduction

Aprendizaje

y paradigmas

Aprendizaje

Supervisado Aprendizaje I

Supervisado

Clusteri

Definici

Aplicaciones

Clustering basado e distancias

The Markov Cluste

Algorithm

comunidades

Pregunta

- Obtener un conjunto de instancias positivas y negativas del concepto que queremos aprender.
- Analizar las características comunes de las instancias positivas que las diferencian de las instancias negativas.
- Buscar hipótesis expresables que discriminen entre el conjunto de instancias conocidas.
- Seleccionar la mejor hipótesis.

A. Alemán

Introduction

Aprendizaje

Aprendizaj

Aprendizaje Supervisado

Aprendizaje I

Supervisado

D C . .

Aplicaciones

Clustering basado e distancias

The Markov Cluste

Detección de

Preguntas

- Obtener un conjunto de instancias positivas y negativas del concepto que queremos aprender.
- Analizar las características comunes de las instancias positivas que las diferencian de las instancias negativas.
- Buscar hipótesis expresables que discriminen entre el conjunto de instancias conocidas.
- Seleccionar la mejor hipótesis.

A. Alemán

Introduction

Aprendizaje
Tipos de Aprendiz

Aprendizaje Aprendizaje

Supervisado

Aprendizaio N

Aprendizaje I Supervisado

CI . .

Definici

Aplicaciones

Clustering basado e distancias

The Markov Cluste

Detección de

Preguntas

- Obtener un conjunto de instancias positivas y negativas del concepto que queremos aprender.
- Analizar las características comunes de las instancias positivas que las diferencian de las instancias negativas.
- Buscar hipótesis expresables que discriminen entre el conjunto de instancias conocidas.
- Seleccionar la mejor hipótesis.

A. Alemán

Introduction

Aprendizaje
Tipos de Aprendiza

Aprendizaje

Aprendizaie

Supervisado Aprendizaje N

Supervisado

CI . .

Definició

Aplicaciones

Clustering basac distancias

Clustering 6

Grafos

The Markov Cluste Algorithm

Detección de comunidades

Preguntas

Aprendizaje Supervisado ARCHES - P.Winston 1975

Ejemplos

A. Alemán

Introduction

Aprendizaje
Tipos de Aprendiza

Aprendizaj

Aprendizaie

Supervisado

Aprendizaje P

Supervisado

Cluste

Delillicio

Clustoring hasa

Clustering basa distancias

Grafos

The Markov Cluste Algorithm

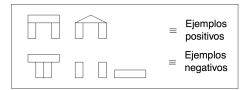
Detección de

Proguntas

Aprendizaje Supervisado

ARCHES - P.Winston 1975

Ejemplos



A. Alemán

Introduction

Tipos de Aprend

Aprendizai

Aprendizaie

Supervisado Aprendizaje N

Supervisado

D C . .

A - I: - - - : - - -

Clustering basad

Clustering

Grafos

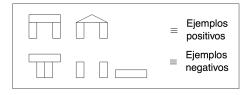
Algorithm

Detección de comunidades

Preguntas

Aprendizaje Supervisado

Ejemplos



A. Alemán

Introduction

Aprendizaje
Tipos de Aprendiza

Aprendizai

Aprendizaie

Supervisado Aprendizaje N

Supervisado

Clustering

Definicio

Aplicacione

Clustering basado e

Grafos
The Markov Clust

The Markov Cluster Algorithm

Detección de comunidades

Pregunta

Aprendizaje Supervisado

Ejemplos





A. Alemán

Introduction

Aprendizaje
Tipos de Aprendiza

Aprendizaio

Aprendizaje

Aprendizaje Supervisado

Aprendizaje N

Supervisado

D (. . . /

A -1:---:--

CI . . .

Clustering basad distancias

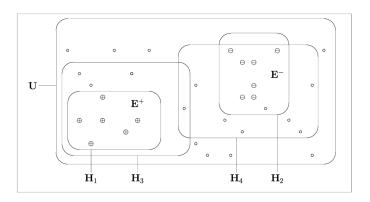
Grafos Grafos

The Markov Cluste Algorithm

Detección de comunidades

Pregunta:

Gráficamente



Introduction

Aprendizaje Tipos de Aprendiza

y paradigmas

Aprendiz

Aprendizaje Supervisado

Aprendizaje N

C1 . .

Ciusterii

Aplicaciones

Clustering basa distancias

Clustering e

The Markov Cluste

Detección de

Proguntas

•
$$E^+ = \{2, 4, 6, 8\}$$

•
$$E^- = \{11, 17\}$$

•
$$H_1 = E^+$$

•
$$H_2 = \mathbb{N} - E^-$$

•
$$H_3 = \{n \in \mathbb{N} | n \text{ es par } \}$$

•
$$H_4 = \{ n \in \mathbb{N} | n \le 10 \}$$

Aprendizaie Supervisado

•
$$E^+ = \{2, 4, 6, 8\}$$

•
$$E^- = \{11, 17\}$$

•
$$H_1 = E^+$$

•
$$H_2 = \mathbb{N} - E^-$$

•
$$H_3 = \{n \in \mathbb{N} | n \text{ es par } \}$$

•
$$H_4 = \{ n \in \mathbb{N} | n < 10 \}$$

Aprendizaie Supervisado

•
$$E^+ = \{2, 4, 6, 8\}$$

•
$$E^- = \{11, 17\}$$

•
$$H_1 = E^+$$

•
$$H_2 = \mathbb{N} - E^-$$

•
$$H_3 = \{n \in \mathbb{N} | n \text{ es par } \}$$

•
$$H_4 = \{ n \in \mathbb{N} | n \le 10 \}$$

Introduction

Aprendizaje

Tipos de Aprendiza

Aprendiza

Aprendizaje Supervisado

Aprendizaje N

Supervisado

Clustering

Definici

Aplicaciones

Clustering basa

Clustering e

Gratos The Markov Clust.

Algorithm

Detección de comunidades

Proguntas

•
$$E^+ = \{2, 4, 6, 8\}$$

•
$$E^- = \{11, 17\}$$

•
$$H_1 = E^+$$

•
$$H_2 = \mathbb{N} - E^-$$

•
$$H_3 = \{n \in \mathbb{N} | n \text{ es par } \}$$

•
$$H_4 = \{ n \in \mathbb{N} | n \le 10 \}$$

Introduction

Aprendizaje

Tipos de Aprendiza

Aprendiza

Aprendizaje Supervisado

Anrendizaie

Aprendizaje N Supervisado

Clustering

Definici

Anlicaciones

Clustering basa

Clustering (

Grafos

The Markov Cluste Algorithm

Detección de

Preguntas

•
$$E^+ = \{2, 4, 6, 8\}$$

•
$$E^- = \{11, 17\}$$

•
$$H_1 = E^+$$

•
$$H_2 = \mathbb{N} - E^-$$

•
$$H_3 = \{n \in \mathbb{N} | n \text{ es par } \}$$

•
$$H_4 = \{ n \in \mathbb{N} | n \le 10 \}$$

Aprendizaie Supervisado

•
$$E^+ = \{2, 4, 6, 8\}$$

•
$$E^- = \{11, 17\}$$

•
$$H_1 = E^+$$

•
$$H_2 = \mathbb{N} - E^-$$

•
$$H_3 = \{n \in \mathbb{N} | n \text{ es par } \}$$

•
$$H_4 = \{ n \in \mathbb{N} | n \le 10 \}$$

A Alemán

Aprendizaje No Supervisado

distancias

Aprendizaje No supervisado

- El aprendizaje no supervisado no hace uso de los valores de clasificación

A. Alemán

Introduction

Aprendizaje
Tipos de Aprendiza

Aprendizaj

Supervisado

Aprendizaje No Supervisado

Supervisado

.

Detinici

Clustoring basado

Clustering basado e distancias

Gratos The Markov Clust.

Algorithm

Detección de comunidades

Pregunta:

Aprendizaje No supervisado

- El aprendizaje no supervisado no hace uso de los valores de clasificación
- En muchos casos no existen los valores de clasificación
- El Clustering es un caso de aprendizaje no supervisado

A. Alemán

Introduction

Aprendizaje
Tipos de Aprendiza

Aprendizaj

Supervisado

Aprendizaje No Supervisado

Supervisado

Defining

Definici

Aplicaciones Clustering basado

distancias

The Markov Cluste

Detección de

Preguntas

Aprendizaje No supervisado

- El aprendizaje no supervisado no hace uso de los valores de clasificación
- En muchos casos no existen los valores de clasificación
- El Clustering es un caso de aprendizaje no supervisado

A Alemán

Clustering

Introduction Tipos de Aprendizaje y paradigmas

2 Aprendizaje

3 Clustering Definición Aplicaciones Clustering basado en distancias

- 4 Clustering en Grafos
- Detección de comunidades

A Alemán

Definición

distancias

Clustering

... in cluster analysis a group of objects is split up into a number of more or less homogeneous subgroups on the basis of an often subjectively chosen meausure of similarity (i.e., chosen subjectively based on its ability to create "interesting" clusters), such taht the similarity between objects within a subgroup is larger than the similarity between objects belonging to different subgroups. (Backer & Jain, 1981)

A Alemán

Definición

Clustering

Se trata de dividir un conjunto de datos de entrada en subconjuntos (clusters), de tal manera que los elementos de cada subconjunto compartan cierto patrón o características a priori desconocidas.

A. Alemán

Introductio

Aprendizaje Tipos de Aprendiza

Aprendizaje

Aprendizaj

Aprendizaie

Supervisado

Ciustei

Anlicaciones

Clustering basado e distancias

Clustering er Grafos

The Markov Cluster Algorithm

Detección de comunidades

Preguntas

- **Ingeniería.** Reconocimiento biométrico, análisis de la señal de radares, eliminación de ruido...
- Computación. Web mining, segmentación de imágenes...
- Biología. Indentificación de la función de genes y proteínas, taxonomías...
- Ciencias Sociales. Patrones en los modelos de comportamiento, patologías criminales...
- Economía. Tipos de clientes, reconocimiento de patrones...

A Alemán

Anlicaciones

- Ingeniería. Reconocimiento biométrico, análisis de la señal de radares, eliminación de ruido...
- Computación. Web mining, segmentación de imágenes. . .

A Alemán

Anlicaciones

- Ingeniería. Reconocimiento biométrico, análisis de la señal de radares, eliminación de ruido...
- Computación. Web mining, segmentación de imágenes. . .
- **Biología.** Indentificación de la función de genes y proteínas, taxonomías...

A Alemán

Anlicaciones

- Ingeniería. Reconocimiento biométrico, análisis de la señal de radares, eliminación de ruido...
- Computación. Web mining, segmentación de imágenes. . .
- **Biología.** Indentificación de la función de genes y proteínas, taxonomías...
- Ciencias Sociales. Patrones en los modelos de comportamiento, patologías criminales...

A. Alemán

Introduction

Aprendizaje
Tipos de Aprendiza

Aprendizaj

A in it

Aprendizaje I

Supervisado

Clusterin

Definició

Aplicaciones

Clustering basado e distancias

Grafos
The Markov Cluste

Algorithm

Detección de

comunidades

Pregunta

- **Ingeniería.** Reconocimiento biométrico, análisis de la señal de radares, eliminación de ruido...
- Computación. Web mining, segmentación de imágenes. . .
- Biología. Indentificación de la función de genes y proteínas, taxonomías...
- Ciencias Sociales. Patrones en los modelos de comportamiento, patologías criminales...
- **Economía.** Tipos de clientes, reconocimiento de patrones. . .

A Alemán

Introduction

Aprendizaje

Tipos de Aprendiz y paradigmas

Aprendizaj

Supervisado

Aprendizaje N

Supervisado

Clusterin

Definición

Clustering basado en

distancias

Grafos
The Markov Cluster

Algorithm

Detección de comunidades

Preguntas

- Idea: dado el número k de grupos o clusters, buscar k puntos o centros representantes de cada cluster, de manera que cada dato se considera en el cluster correspondiente al centro que tiene a menor "distancia".
- El concepto de distancia será específico en cada problema:
 - Expresará la medida de similitud
 - La distancia más usada es la euclídea

A. Alemán

Introduction

Aprendizaje

y parauiginas

Aprendizaj

Aprendizaie N

Supervisado

Clustering

Definición

Aplicaciones

Clustering basado en distancias

Clustering er

The Markov Cluster Algorithm

Detección de

Pregunta

- Idea: dado el número k de grupos o clusters, buscar k puntos o centros representantes de cada cluster, de manera que cada dato se considera en el cluster correspondiente al centro que tiene a menor "distancia".
- El concepto de distancia será específico en cada problema:
 - Expresará la medida de similitud
 - La distancia más usada es la euclídea

A Alemán

Introduction

Aprendizaje

Aprendizai

, ipromanzaj

Anrendizaie N

Supervisado

Dofinición

A =1:===:===

Clustering basado en

distancias

Clustering er Grafos

The Markov Cluster Algorithm

Detección de

Preguntas

- Idea: dado el número k de grupos o clusters, buscar k puntos o centros representantes de cada cluster, de manera que cada dato se considera en el cluster correspondiente al centro que tiene a menor "distancia".
- El concepto de distancia será específico en cada problema:
 - Expresará la medida de similitud
 - La distancia más usada es la euclídea

A. Alemán

Introduction

Aprendizaje

y paradigmas

Aprendizaj

Aprendizaje 1

Supervisado

Definition

Aplicaciones

Clustering basado en

Clustering e

The Markov Cluster

Detección de

Preguntas

- Idea: dado el número k de grupos o clusters, buscar k puntos o centros representantes de cada cluster, de manera que cada dato se considera en el cluster correspondiente al centro que tiene a menor "distancia".
- El concepto de distancia será específico en cada problema:
 - Expresará la medida de similitud
 - La distancia más usada es la euclídea

A Alemán

Clustering basado en distancias

Un algoritmo clásico: k-medias

- Entrada: un número k de clusters, un conjunto de datos $\{x_i\}_{i=1}^N$ y una función de distancia
- Salida: un conjunto de k centros m_1, \ldots, m_k

A Alemán

Clustering basado en distancias

Un algoritmo clásico: k-medias

- Entrada: un número k de clusters, un conjunto de datos $\{x_i\}_{i=1}^N$ y una función de distancia
- Salida: un conjunto de k centros m_1, \ldots, m_k

A Alemán

Clustering basado en

distancias

Un algoritmo clásico: k-medias

- Entrada: un número k de clusters, un conjunto de datos $\{x_i\}_{i=1}^N$ y una función de distancia
- Salida: un conjunto de k centros m_1, \ldots, m_k

- 1 Inicializar m_i (i=1,..., k) (aleatoriamente o con algún critero heurístico)

A Alemán

Clustering basado en distancias

Un algoritmo clásico: k-medias

- Entrada: un número k de clusters, un conjunto de datos $\{x_i\}_{i=1}^N$ y una función de distancia
- Salida: un conjunto de k centros m_1, \ldots, m_k

- 1 Inicializar m_i (i=1,..., k) (aleatoriamente o con algún critero heurístico)
- **2** REPETIR (hasta que los m_i no cambien):



A Alemán

Clustering basado en

distancias

Un algoritmo clásico: k-medias

- Entrada: un número k de clusters, un conjunto de datos $\{x_i\}_{i=1}^N$ y una función de distancia
- Salida: un conjunto de k centros m_1, \ldots, m_k

- 1 Inicializar m_i (i=1,..., k) (aleatoriamente o con algún critero heurístico)
- 2 REPETIR (hasta que los m_i no cambien):
 - PARA j=1,...,N HACER: Calcular el cluster correspondiente a x_i , escogiendo, de entre todos los m_i el m_h tal que distancia (x_i, m_h) sea mínima.

A. Alemán

Introduction

Aprendizaje

Tipos de Aprendizajo

y paradigmas

Aprendizaj

^ ----di--i-

Supervisado

Aprendizaie 1

Aprendizaje I Supervisado

Clusterin

Definición Anlicacione

Clustering basado en distancias

Clustering en

The Markov Cluste

Detección de

Preguntas

Un algoritmo clásico: k-medias

- Entrada: un número k de clusters, un conjunto de datos $\{x_i\}_{i=1}^N$ y una función de distancia
- Salida: un conjunto de k centros m_1, \ldots, m_k

- 1 Inicializar m_i (i=1,..., k) (aleatoriamente o con algún critero heurístico)
- **2** REPETIR (hasta que los m_i no cambien):
 - 1 PARA j=1,...,N HACER: Calcular el cluster correspondiente a x_j, escogiendo, de entre todos los m_i el m_h tal que distancia(x_j, m_h) sea mínima.
 - **2** PARA i=1,...,k HACER: Asigar a m_i la media aritmética de los datos asignados al cluster i-ésimo.
- 3 Devolver m_1, \ldots, m_r



A. Alemán

Introduction

Aprendizaje
Tipos de Aprendizaje
v paradigmas

Aprendizaj

, premarzaj

Aprendizaje I

Aprendizaje I Supervisado

Clusterin

Definición

Aplicaciones Clustering basado en

distancias

Grafos
The Markov Cluste

Algorithm

comunidades

Preguntas

Un algoritmo clásico: k-medias

- Entrada: un número k de clusters, un conjunto de datos $\{x_i\}_{i=1}^N$ y una función de distancia
- Salida: un conjunto de k centros m_1, \ldots, m_k

- 1 Inicializar m_i (i=1,..., k) (aleatoriamente o con algún critero heurístico)
- **2** REPETIR (hasta que los m_i no cambien):
 - **1** PARA j=1,...,N HACER: Calcular el cluster correspondiente a x_j , escogiendo, de entre todos los m_i el m_h tal que $distancia(x_i, m_h)$ sea mínima.
 - **2** PARA i=1,...,k HACER: Asigar a m_i la media aritmética de los datos asignados al cluster i-ésimo.
- 3 Devolver m_1, \ldots, m_n



A. Alemán

Introduction

Aprendizaje

Aprendizai

Aprendizaje

Aprendizaje N

Clusterin

Delinicio

Clustering basado en

distancias

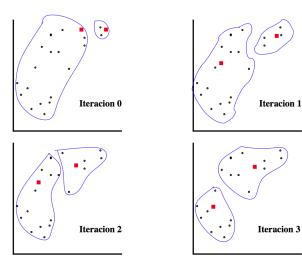
Clustering e Grafos

The Markov Cluste Algorithm

Detección de

Pregunta

Idea gráfica intuitiva en el algoritmo de k-medias



A. Alemán

Introduction

Tipos de Aprendizaj

Aprendizaj

, in .

Aprendizaje N

Supervisado

.

Anlicaciones

Clustering bas

Clustering en

Grafos The Markov Cluste

Algorithm

Detección de comunidades

Preguntas

- 2 Aprendizaje Aprendizaje Supervisado Aprendizaje No Supervisad
 - 3 Clustering
 Definición
 Aplicaciones
 Clustering basado en distancias
- 4 Clustering en Grafos The Markov Cluster Algorithm
- 5 Detección de comunidades
- 6 Preguntas

A. Alemán

Introduction

Aprendizaje
Tipos de Aprendiza

Aprendiza

Aprendizaj

Aprendizaje N

Supervisado

Ciable

Definició

Aplicaciones

Clustering basa distancias

Grafos Grafos

The Markov Cluster Algorithm

Detección de

Preguntas

Basado en la thesis de Stijn van Dongen

- MCL es un algoritmo de clustering de grafos
- Más información en http://www.micans.org/mc

A. Alemán

Introduction

Tipos de Aprendiza

Aprendiz

Aprendizaje

Aprendizaje N

Supervisado

Clusteri

Definici

Anlicacione

Clustering ba

Clustoring of

Grafos

The Markov Cluster Algorithm

Detección de comunidades

Preguntas

- Basado en la thesis de Stijn van Dongen
- MCL es un algoritmo de clustering de grafos
- Más información en http://www.micans.org/mcl

Introduction

Aprendizaje
Tipos de Aprendiza

Aprendizaj

Aprendizaje N

Supervisado

Ciusto

Clustering basado e

distancias

The Markov Cluster Algorithm

Detección de

Preguntas

- Basado en la thesis de Stijn van Dongen
- MCL es un algoritmo de clustering de grafos
- Más información en http://www.micans.org/mcl

A. Alemán

Introduction

Aprendizaje Tipos de Aprendiza

Aprendizaj

Aprendizaje N

Aprendizaje N Supervisado

Clusterin

Definición

Aplicaciones

Clustering basado e distancias

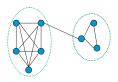
The Markov Cluster

Detección de

Preguntas

RandomWalks

Caminos aleatorios



- Dado un grafo, habrá muchas conexiones dentro del cluster y menos conexiones entre clusters.
- Si queremos empezar en un nodo y viajar (de manera aleatoria) a un nodo conectado, es más probable llegar a un nodo del cluster que a un nodo de otro cluster.
- MCL está basado en estos principios.

A. Alemán

Introduction

Aprendizaje Tipos de Aprendiz

Aprendizaj

Anrendizaie N

Aprendizaje N Supervisado

Clustorin

Definici

Aplicaciones Clustering basado

distancias

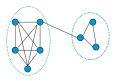
The Markov Cluster

Detección de

Preguntas

RandomWalks

Caminos aleatorios



- Dado un grafo, habrá muchas conexiones dentro del cluster y menos conexiones entre clusters.
- Si queremos empezar en un nodo y viajar (de manera aleatoria) a un nodo conectado, es más probable llegar a un nodo del cluster que a un nodo de otro cluster.
- MCL está basado en estos principios.

A. Alemán

Introduction

Aprendizaje Tipos de Aprendiz

Aprendizaj

Aprendizaie N

Supervisado

Clusterin

Definició

Aplicaciones
Clustering basado e

distancias

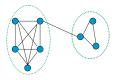
The Markov Cluster

Detección de

Preguntas

RandomWalks

Caminos aleatorios



- Dado un grafo, habrá muchas conexiones dentro del cluster y menos conexiones entre clusters.
- Si queremos empezar en un nodo y viajar (de manera aleatoria) a un nodo conectado, es más probable llegar a un nodo del cluster que a un nodo de otro cluster.
- MCL está basado en estos principios.

A. Alemán

Introductio

Aprendizaje
Tipos de Aprendiza

Aprendizaj

Supervisado

Aprendizaje I Supervisado

Supervisado

Defining

Anlicacion

Clustering ba

distancias

Grafos

The Markov Cluster Algorithm

Detección de

Preguntas

RandomWalks

- Mediante Random Walks podemos descubrir dónde tiende a concentrarse el flujo de los mismos, de esta forma sabremos dónde están los clusters.
- Los Random Walks se calculan mediante "Cadenas de Markov".

A. Alemán

Introductio

Aprendizaje
Tipos de Aprendiza

Aprendizaj

Supervisado Aprendizaie N

Aprendizaje N Supervisado

Ciusti

Anlicacione

Clustering basado distancias

Clustering en

The Markov Cluster Algorithm

Detección de

Preguntas

RandomWalks

- Mediante Random Walks podemos descubrir dónde tiende a concentrarse el flujo de los mismos, de esta forma sabremos dónde están los clusters.
- Los Random Walks se calculan mediante "Cadenas de Markov".

A Alemán

The Markov Cluster Algorithm

- 1 Entrada: un grafo no dirigido, potencia e y crecimientor

A. Alemán

Introductio

Aprendizaje
Tipos de Aprendiza

Aprendizaj

Anrendizaie

Aprendizaje N

Supervisado

CI . .

Definic

Aplicaciones

Clustering bas

Clustering e

The Markov Cluster

Detección de

Preguntas

- 1 Entrada: un grafo no dirigido, potencia e y crecimientor
- 2 Se crea una matriz de asociaciones
- 3 Se añaden bucles en cada nodo (consigo mismo) (optional)
- 4 Se normaliza la matriz
- **5** Se expande tomando la e_{th} potencia de la matriz
- 6 Se incrementa usando el crecimiento r
- 7 Se repiten pasos 5 y 6 hasta que el sistema converge
- 8 Se interpreta la matriz resultado para descubrir los clusters

A. Alemán

Introductio

Aprendizaje
Tipos de Aprendiza

Aprendizaj

.

Aprendizaje N

Supervisado

Ciuste

A =1:===:

Clustoring by

distancias

Clustering en Grafos

The Markov Cluster Algorithm

Detección de comunidades

Preguntas

- 1 Entrada: un grafo no dirigido, potencia e y crecimientor
- 2 Se crea una matriz de asociaciones
- 3 Se añaden bucles en cada nodo (consigo mismo) (optional)
- 4 Se normaliza la matriz
- Se expande tomando la e_{th} potencia de la matriz
- 6 Se incrementa usando el crecimiento r
- 7 Se repiten pasos 5 y 6 hasta que el sistema converge
- 8 Se interpreta la matriz resultado para descubrir los clusters

A. Alemán

Introduction

Aprendizaje Tipos de Aprendiza

Aprendizaj

Aprendizaj

Aprendizaje N

Supervisado

Clusterin

Definicio

Aplicaciones

Clustering ba

distancias

Grafos The Markov Cluster

The Markov Cluster Algorithm

Detección de comunidades

Preguntas

- 1 Entrada: un grafo no dirigido, potencia e y crecimientor
- Se crea una matriz de asociaciones
- 3 Se añaden bucles en cada nodo (consigo mismo) (optional)
- 4 Se normaliza la matriz
- **5** Se expande tomando la e_{th} potencia de la matriz
- 6 Se incrementa usando el crecimiento r
- 7 Se repiten pasos 5 y 6 hasta que el sistema converge
- 8 Se interpreta la matriz resultado para descubrir los clusters

A. Alemán

Introductio

Aprendizaje Tipos de Aprendiz

Aprendizaj

Aprendizaje

Aprendizaje N

Supervisado

Clusteri

Definici

Aplicaciones

Clustering basa

Clustering en

The Markov Cluster Algorithm

Detección de comunidades

Pregunta:

- 1 Entrada: un grafo no dirigido, potencia e y crecimientor
- 2 Se crea una matriz de asociaciones
- 3 Se añaden bucles en cada nodo (consigo mismo) (optional)
- 4 Se normaliza la matriz
- **5** Se expande tomando la e_{th} potencia de la matriz
- 6 Se incrementa usando el crecimiento r
- 7 Se repiten pasos 5 y 6 hasta que el sistema converge
- 8 Se interpreta la matriz resultado para descubrir los clusters

A. Alemán

Introduction

Aprendizaje Tipos de Aprendiz

Aprendiza

A III

Aprendizaje N

Supervisado

Clusteri

A = 1: ---:---

Clustering basado e

distancias

The Markov Cluster Algorithm

Detección de

Preguntas

- 1 Entrada: un grafo no dirigido, potencia e y crecimientor
- 2 Se crea una matriz de asociaciones
- 3 Se añaden bucles en cada nodo (consigo mismo) (optional)
- 4 Se normaliza la matriz
- **5** Se expande tomando la e_{th} potencia de la matriz
- 6 Se incrementa usando el crecimiento r
- 7 Se repiten pasos 5 y 6 hasta que el sistema converge
- 8 Se interpreta la matriz resultado para descubrir los clusters

A. Alemán

Introduction

Aprendizaje
Tipos de Aprendiza

Aprendizaj

Aprendizaje

Aprendizaje N

Supervisado

Definicio

Aplicaciones

Clustering basado e

distancias

The Markov Cluster Algorithm

Detección de

Preguntas

Algoritmo MCL

- 1 Entrada: un grafo no dirigido, potencia e y crecimientor
- 2 Se crea una matriz de asociaciones
- 3 Se añaden bucles en cada nodo (consigo mismo) (optional)
- 4 Se normaliza la matriz
- **5** Se expande tomando la e_{th} potencia de la matriz
- 6 Se incrementa usando el crecimiento r
- Se repiten pasos 5 y 6 hasta que el sistema converge
- Se interpreta la matriz resultado para descubrir los clusters

A. Alemán

Introduction

Aprendizaje Tipos de Aprendiz

Aprendizaj

Aprendizaje

Aprendizaje N

Supervisado

Clusterin

Definición

Aplicaciones

Clustering basado e distancias

The Markov Cluster

Algorithm

Detección de comunidades

Preguntas

Algoritmo MCL

- 1 Entrada: un grafo no dirigido, potencia e y crecimientor
- Se crea una matriz de asociaciones
- 3 Se añaden bucles en cada nodo (consigo mismo) (optional)
- 4 Se normaliza la matriz
- Se expande tomando la e_{th} potencia de la matriz
- 6 Se incrementa usando el crecimiento r
- 7 Se repiten pasos 5 y 6 hasta que el sistema converge
- 8 Se interpreta la matriz resultado para descubrir los clusters

A. Alemán

Introduction

Tipos de Aprendiza

Aprendizaje

. . . .

Aprendizaje N

Supervisado

D-6-1-14

Delinicio

Aplicacione

distancias

distancias

Grafos

The Markov Cluster Algorithm

Detección de

Pregunta:

- Los vértices se dividen en 2 tipos: Los vértices atractores (que atraen otros vértices) y los vértices que son atraídos por los atractores.
- Cada atractor atrae los vértices que tienen su mismo valor en su fila.
- En nuestro ejemplo tendríamos:
 - Nodo 1
 - Nodos 2-4
 - Nodo 3

A. Alemán

Introduction

Tipos de Aprendiza

Aprendizaj

Supervisado

Supervisado

Supervisado

Dofinicia

Anlicacio

Clustering h

distancias

Clustering en

The Markov Cluster Algorithm

Detección de

Preguntas

- Los vértices se dividen en 2 tipos: Los vértices atractores (que atraen otros vértices) y los vértices que son atraídos por los atractores.
- Cada atractor atrae los vértices que tienen su mismo valor en su fila.
- En nuestro ejemplo tendríamos:
 - Nodo 1
 - Nodos 2-4
 - Nodo 3

A. Alemán

Introduction

Tipos de Aprendiza

Aprendizaj

.

Aprendizaje N

Supervisado

.

Definició

. . . .

CI . . .

Clustering basado e

distancias

The Markov Cluster Algorithm

Detección de

Preguntas

- Los vértices se dividen en 2 tipos: Los vértices atractores (que atraen otros vértices) y los vértices que son atraídos por los atractores.
- Cada atractor atrae los vértices que tienen su mismo valor en su fila.
- En nuestro ejemplo tendríamos:
 - Nodo 1
 - Nodos 2-4
 - Nodo 3

A. Alemán

Introduction

Tipos de Aprendiza

Aprendizaj

Supervisado

Supervisado

Clustonin

Definicio

Aplicaciones

Clustering basado er

distancias

The Markov Cluster Algorithm

Detección de

Preguntas

- Los vértices se dividen en 2 tipos: Los vértices atractores (que atraen otros vértices) y los vértices que son atraídos por los atractores.
- Cada atractor atrae los vértices que tienen su mismo valor en su fila.
- En nuestro ejemplo tendríamos:
 - Nodo 1
 - Nodos 2-4
 - Nodo 3

A. Alemán

Introduction

Tipos de Aprendiza

Aprendizaj

Aprendizaie

Aprendizaje N

Supervisado

- -

Definicio

Aplicaciones

Clustering basado e

distancias

The Markov Cluster Algorithm

Detección de

Preguntas

- Los vértices se dividen en 2 tipos: Los vértices atractores (que atraen otros vértices) y los vértices que son atraídos por los atractores.
- Cada atractor atrae los vértices que tienen su mismo valor en su fila.
- En nuestro ejemplo tendríamos:
 - Nodo 1
 - Nodos 2-4
 - Nodo 3

A. Alemán

Introduction

Tipos de Aprendiza

Aprendizaj

Aprendizaje No

Supervisado

Definicio

Detinici

Aplicaciones

Clustering basado e distancias

Clustering en Grafos

The Markov Cluster Algorithm

Detección de

Pregunta

- Los vértices se dividen en 2 tipos: Los vértices atractores (que atraen otros vértices) y los vértices que son atraídos por los atractores.
- Cada atractor atrae los vértices que tienen su mismo valor en su fila.
- En nuestro ejemplo tendríamos:
 - Nodo 1
 - Nodos 2-4
 - Nodo 3

A. Alemán

Introduction

Aprendizaje
Tipos de Aprendiza

Aprendizaje

Aprendizaje

Aprendizaje N

Supervisado

Cluster

Definició

Aplicaciones

Clustering b

Cl

Grafos

The Markov Cluster Algorithm

Detección de comunidades

Pregunta:

Ejemplo con Cytoscape y Reactome

- Cytoscape
- Reactome FI
- Example Data

A. Alemán

Introduction

Tipos de Aprendiza

Aprendiz

Aprendizaje

Aprendizaje N

Supervisado

(3) Clusterin

Definici

Aplicaciones

Clustering ba

Clustering en

The Markov Cluste

Detección de comunidades

Preguntas

- Introduction
 Aprendizaje
 Tipos de Aprendizaje v paradigmas
- 2 Aprendizaje Aprendizaje Supervisado Aprendizaje No Supervisad
 - Definición
 Aplicaciones
 Clustering basado en distancias
- 4 Clustering en Grafos
 The Markov Cluster Algorithm
- 5 Detección de comunidades
- 6 Preguntas

A. Alemán

Introduction

Aprendizaje Tipos de Aprendiza

Aprendizaj

Aprendizaje

Aprendizaje N

Supervisado

Ciuste

Definicio

Aplicacione:

Clustering base

Clustering er

The Markov Cluster

Detección de comunidades

Preguntas

Detección de Comunidades

- ¿En internet?
- ¿En la gestión de procesos en informática?
- ¿En interacciones proteína-proteína?
- ; En redes metabólicas?

A. Alemán

Introduction

Aprendizaje Tipos de Aprendiz

Aprendizaj

Apropdizaio N

Supervisado

Definici

Anlicacione

.

distancias

Clustering er

The Markov Cluste

Detección de comunidades

Pregunta

Detección de Comunidades

- ¿En internet?
- ¿En la gestión de procesos en informática?
- ¿En interacciones proteína-proteína?
- ¿En redes metabólicas?

A. Alemán

Introduction

Aprendizaje
Tipos de Aprendiza

Aprendizaj

Aprendizaje N

Supervisado

Cluste

. . . .

Clustering basado

Clustoring on

The Markov Cluste

Detección de comunidades

Preguntas

Detección de Comunidades

- ¿En internet?
- ¿En la gestión de procesos en informática?
- ¿En interacciones proteína-proteína?
- ; En redes metabólicas?

A. Alemán

Introduction

Aprendizaje
Tipos de Aprendiz

Aprendizaj

Aprendizaje N

Supervisado

0.000

Anlicacion

Clustering basado

distancias

The Markov Cluster

Detección de comunidades

Preguntas

Detección de Comunidades

- ¿En internet?
- ¿En la gestión de procesos en informática?
- ¿En interacciones proteína-proteína?
- ¿En redes metabólicas?

A. Alemán

Introduction

Tipos de Aprendiz

Aprendizaj

Aprendizaje

Aprendizaje N

Supervisado

Clusterii

Definició

Aplicacione

Clustering basad

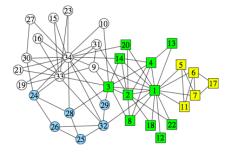
Clustering e Grafos

The Markov Cluster Algorithm

Detección de comunidades

Preguntas

Ejemplos Karate



A. Alemán

Introduction

Aprendizaje

Tinos de Aprendiz

Aprendizai

Aprendizaje Supervisado

Aprendizaje N

Anlicacion

. . . .

Clustering basado distancias

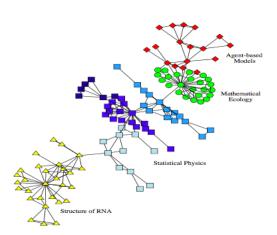
Grafos

The Markov Cluster Algorithm

Detección de comunidades

Preguntas

Ejemplos Instituto



A. Alemán

Introduction

Anrendi

Tipos de Aprendiza

Aprendizaio

Aprendizaje

Aprendizaje N

Supervisado

Clusterin

Definició

Anlicacion

Clustering basad

distancias

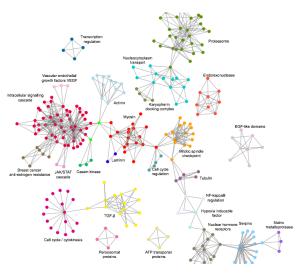
Grafos

The Markov Cluster Algorithm

Detección de comunidades

Preguntas

Ejemplos PPI



A Alemán

distancias

Detección de comunidades

Algoritmos para detección de comunidades

- Graph partitioning
- Partitional clustering

A. Alemán

Introduction

Aprendizaje
Tipos de Aprendiza

Aprendizaj

Supervisado

Aprendizaje N Supervisado

Supervisado

D-G-1-1

Aplicacion

Clustering ba

distancias

Clustering en

The Markov Cluster Algorithm

Detección de comunidades

Pregunta:

Algoritmos para detección de comunidades

- Graph partitioning
- Hierarchical clustering
- Partitional clustering
- Spectral clustering
- . . .

A. Alemán

Introduction

Aprendizaje
Tipos de Aprendiza

Aprendizaje

Supervisado

Supervisado

Supervisado

0.0...

Anlicacio

Clustering basad

distancias

The Markov Cluster

Detección de

comunidades

Preguntas

Algoritmos para detección de comunidades

- Graph partitioning
- Hierarchical clustering
- Partitional clustering
- Spectral clustering
- . . .

A. Alemán

Introduction

Aprendizaje
Tipos de Aprendiza

Aprendizaj

Supervisado

Supervisado

Supervisado

.

Delinici

Aplicacione

Clustering basado e distancias

Clustering en

The Markov Cluster Algorithm

Detección de comunidades

Pregunta

Algoritmos para detección de comunidades

- Graph partitioning
- Hierarchical clustering
- Partitional clustering
- Spectral clustering

. . . .

A. Alemán

Introduction

Aprendizaje
Tipos de Aprendiza

Aprendizaj

Supervisado

Supervisado

Definici

Anlicacione

Clustering basado e

distancias

The Markov Cluster

Detección de

Pregunta

Algoritmos para detección de comunidades

- Graph partitioning
- Hierarchical clustering
- Partitional clustering
- Spectral clustering
- . . .

A Alemán

Detección de comunidades

- Edge-Betweenness

- InfoMAP

A. Alemán

Introduction

Aprendizaje
Tipos de Aprendiza

Aprendizaj

, tpremarzaj

Aprendizaje I

Supervisado

D-6-1-1

A =1:===:=

.

distancias

Clustering er

The Markov Cluster Algorithm

Detección de comunidades

Pregunta

Ejemplos de algoritmos en igraph

- Edge-Betweenness
- Leading Eigenvector
- Fast-Greedy
- Multi-Leve
- Walktrap
- Label Propagation
- InfoMAP

A Alemán

Detección de comunidades

- Edge-Betweenness
- Leading Eigenvector
- Fast-Greedy

- InfoMAP

A. Alemán

Introduction

Aprendizaje
Tipos de Aprendiza

Aprendizaj

Aprendizaje

Aprendizaje I

Supervisado

Clusterii

Definici

Anlicacione

Clustering basado e

distancias

The Markov Cluste

Detección de

Pregunta

Ejemplos de algoritmos en igraph

- Edge-Betweenness
- Leading Eigenvector
- Fast-Greedy
- Multi-Level
- · Malktran
- Label Propagation
- InfoMAP

A Alemán

Detección de comunidades

- Edge-Betweenness
- Leading Eigenvector
- Fast-Greedy
- Multi-Level
- Walktrap
- InfoMAP

A Alemán

Detección de

comunidades

- Edge-Betweenness
- Leading Eigenvector
- Fast-Greedy
- Multi-Level
- Walktrap
- Label Propagation

A. Alemán

Introduction

Aprendizaje
Tipos de Aprendiza

Aprendizaj

Supervisado

Aprendizaio N

Supervisado

Charles

Definici

Aplicacion

Clustering basado e

distancias

The Markov Cluste

Detección de comunidades

Preguntas

- Edge-Betweenness
- Leading Eigenvector
- Fast-Greedy
- Multi-Level
- Walktrap
- Label Propagation
- InfoMAP

A. Alemán

Introduction

Tipos de Aprendiza

Aprendizaj

Supervisado

Aprendizaje N Supervisado

Definition

A = 1: == = : = =

Clustering basad

Clustering e

The Markov Clust

Detección de

Preguntas

Preguntas

¿Preguntas? Gracias