

Campos gravitatorios en el espacio muestral

José Luis Triviño Rodríguez

Dept. Lenguajes y Ciencias de la
Computación
Universidad de Málaga

Septiembre, 2004

● Introducción

- Los ingredientes
- La receta
- El resultado
- Conclusiones

Enfocando el problema

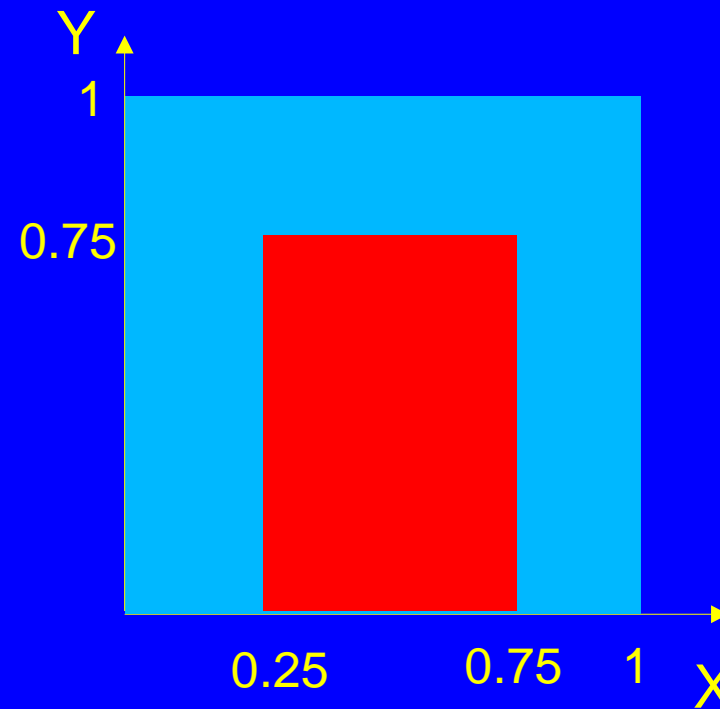
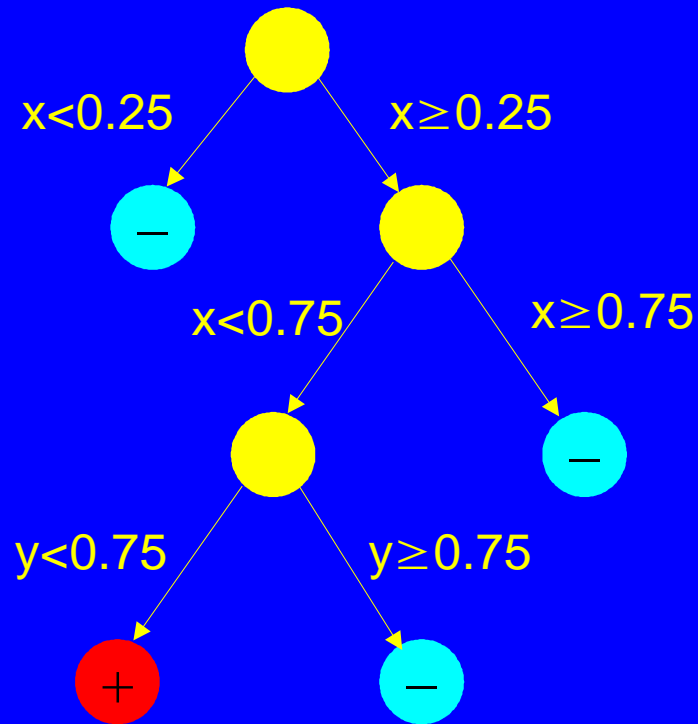
- Aprendizaje supervisado
- Clasificación dicotómica (2 clases)
- Atributos de tipo continuo
- Valores normalizados entre 0 y 1



Introducción

- Los ingredientes
- La receta
- El resultado
- Conclusiones

Clasificar = Dividir el espacio muestral



• Introducción

- Los ingredientes
- La receta
- El resultado
- Conclusiones

Aprendizaje = Cálculo de fronteras

- El proceso de aprendizaje consiste en el cálculo de las fronteras entre las particiones del espacio muestral
- Las particiones constituyen una abstracción del conjunto de muestras de aprendizaje



● Introducción

- Los ingredientes
- La receta
- El resultado
- Conclusiones

El vecino más cercano

- Constituye el mecanismo de aprendizaje más sencillo
- Las particiones se calculan según la distancia a cada elemento del conjunto de aprendizaje
- Las fronteras corresponden a puntos equidistantes a muestras de signo contrario



• Introducción

- Los ingredientes
- La receta
- El resultado
- Conclusiones

El vecino más cercano

Ventajas

- El proceso de aprendizaje es muy sencillo: no hacer nada
- La forma de las fronteras no esta limitada por el modelo de aprendizaje



● Introducción

- Los ingredientes
- La receta
- El resultado
- Conclusiones

El vecino más cercano Inconvenientes

- No hay un aprendizaje real
- La hipótesis generada por el modelo no es más sencilla que la muestra de aprendizaje (Principio de la navaja de Occam)
- No hay extracción del conocimiento



Fundamentos

- **Introducción**

- Los ingredientes
- La receta
- El resultado
- Conclusiones

- Agrupar muestras supone una abstracción del espacio muestral (Aprendizaje)
- Aplicando el mecanismo del vecino más cercano a los grupos de muestras se definen fronteras



● Introducción

- Los ingredientes
- La receta
- El resultado
- Conclusiones

Cuestiones

- ¿Como agrupar las muestras?
- ¿Existen modelos de hipótesis que no limiten la complejidad de las fronteras?
- ¿Que función distancia utilizar?



- Introducción
- **Los ingredientes**
- La receta
- El resultado
- Conclusiones

Los ingredientes

- Representar agrupación de muestras mediante su centro de gravedad (Modelo de Hipótesis)
- Ecuación de la fuerza gravitatoria como distancia
- Desintegración de masa (Olvido)



- Introducción
- **Los ingredientes**
- La receta
- El resultado
- Conclusiones

Centro de masas

- Constituye una abstracción de un conjunto de masas (Aprendizaje)
- Fácil de calcular
- Calculable de forma incremental (Aprendizaje incremental – En línea)



- Introducción
- **Los ingredientes**
- La receta
- El resultado
- Conclusiones

Fuerza gravitatoria

- Constituye un concepto natural de distancia
- Incluye el concepto de peso
- Utiliza valores continuos
- Estrechamente ligado al concepto de campo gravitatorio (Región del espacio)



- Introducción
- **Los ingredientes**
- La receta
- El resultado
- Conclusiones

Desintegración de masa

- Permite modelar el olvido durante el aprendizaje
- Aprendizaje en línea
- Modelado de sistemas dinámicos



Algoritmo de aprendizaje básico

- Introducción
- Los ingredientes
- La receta
- El resultado
- Conclusiones

- Una muestra es clasificada por el centro de masas que genera una fuerza mayor sobre la muestra
- Las muestras se añaden de una en una al modelo



- Introducción
- Los ingredientes
- La receta
- El resultado
- Conclusiones

Nueva muestra consistente con el modelo

- Si la muestra añadida es clasificada correctamente por un centro de masas existente en el modelo, entonces la muestra se añade a ese centro de masas



- Introducción
- Los ingredientes
- La receta
- El resultado
- Conclusiones

Nueva muestra inconsistente con el modelo

- Si la muestra añadida no es clasificada correctamente por un centro de masas existente en el modelo, entonces se crea un nuevo centro de masas a partir de la nueva muestra



- Introducción
- Los ingredientes
- **La receta**
- El resultado
- Conclusiones

División de centros de masa

- La naturaleza incremental del modelo de aprendizaje puede dar lugar a que dos centros de masa se consideren uno sólo
- Es necesario un mecanismo para dividir estos centros de masa



- Introducción
- Los ingredientes
- La receta
- El resultado
- Conclusiones

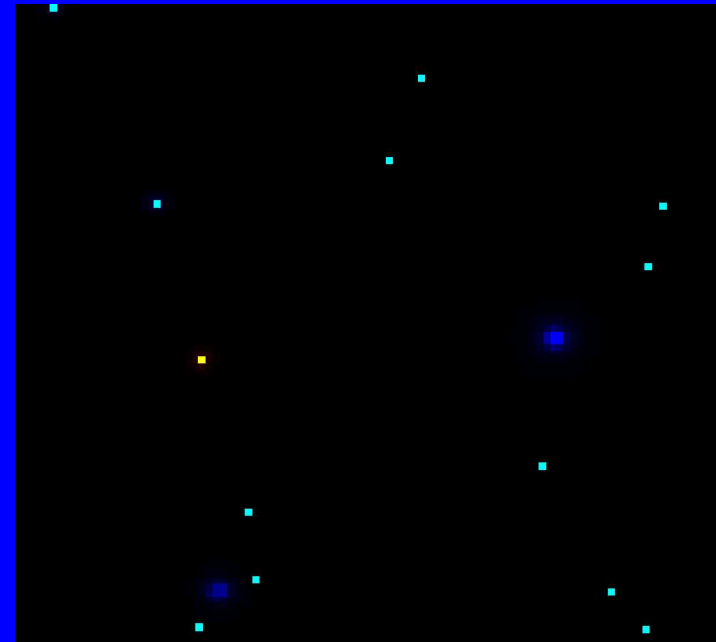
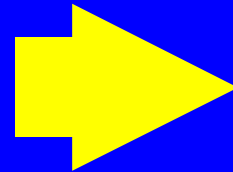
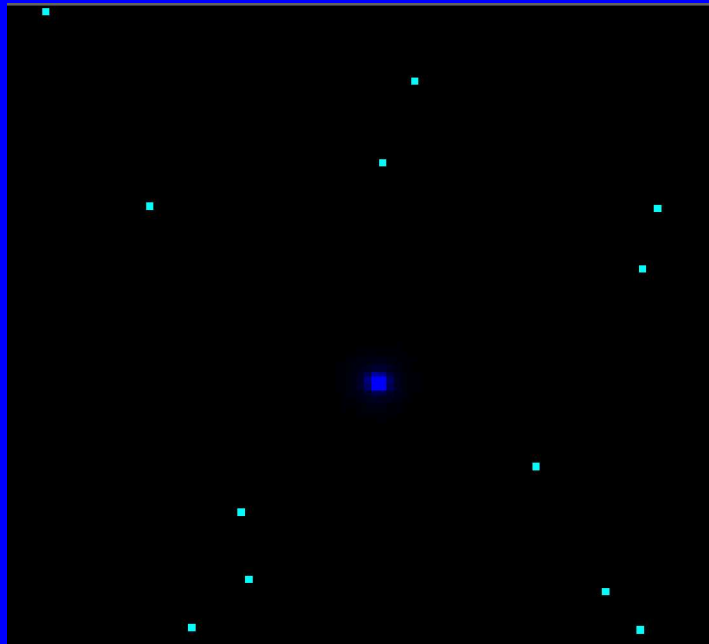
División de centros de masa

- Este mecanismo se activará cuando una muestra revele inconsistencias en el modelo
- Heurístico: Un centro de masa no puede estar más cerca de otro que la masa más lejana que lo forme



- Introducción
- Los ingredientes
- La receta
- El resultado
- Conclusiones

División de centros de masa



- Introducción
- Los ingredientes
- **La receta**
- El resultado
- Conclusiones

Fusión de centros de masa

- Es posible considerar un proceso inverso a la división de centros de masa
- Heurístico: Dos centros de masa pueden fusionarse si todas las muestras que lo forman son clasificadas correctamente por su unión



- Introducción
- Los ingredientes
- **La receta**
- El resultado
- Conclusiones

Fusión de centros de masa

- Es posible considerar un proceso inverso a la división de centros de masa
- Heurístico: Dos centros de masa pueden fusionarse si todas las muestras que lo forman son clasificadas correctamente por su unión



- Introducción
- Los ingredientes
- La receta
- El resultado
- Conclusiones

Desintegración de masa

- Cada vez que se añade una nueva muestra se disminuye la masa de todos los centros de masa (olvido)
- Permite eliminar progresivamente centros de masa que no reciben nuevas muestras



- Introducción
- Los ingredientes
- **La receta**
- El resultado
- Conclusiones

Memoria limitada

- Cada centro de masa almacena las muestras que lo forman
- El número de muestras que almacena un centro de masa es limitado
- Una vez alcanzado este máximo cada nueva muestra añadida elimina la más antigua



- Introducción
- Los ingredientes
- La receta
- **El resultado**
- Conclusiones

Modelo de hipótesis

- El resultado del proceso de aprendizaje es un conjunto de centros de masa
- Cada centro de masa tiene asociado un peso y una etiqueta
- Cada centro de masa representa un conjunto de muestras de aprendizaje



- Introducción
- Los ingredientes
- La receta
- **El resultado**
- Conclusiones

Modelo de hipótesis

- Cada centro de masa define una región en el espacio muestral
- Estas regiones corresponden a las zonas de mayor influencia gravitatoria de cada centro de masa



- Introducción
- Los ingredientes
- La receta
- **El resultado**
- Conclusiones

Modelo de hipótesis

- La forma de cada región viene determinada por la influencia mútua de los centros de masa entre sí
- La forma de estas regiones no está limitada por el modelo de hipótesis



Características del modelo

- Introducción
- Los ingredientes
- La receta
- El resultado
- Conclusiones

- Aprendizaje incremental
- Modelado de sistemas dinámicos (olvido)
- Memoria acotada
- Sencillo
- No limitado por el modelo de hipótesis



- Introducción
- Los ingredientes
- La receta
- El resultado
- Conclusiones

Desarrollo futuro

- Mejorar el proceso de fusión de centros de masa
- Estudiar la influencia de los parámetros en el resultado del aprendizaje
- Aplicar el modelo a problemas reales
- Extracción de reglas a partir de las regiones

