

Inteligencia artificial

Resumen para el segundo parcial

Ultima modificación el martes 3 de julio de 2007 a las 03:51:05
Copyright © 2007, Kronoman – In loving memory of my father - <http://kronoman.kicks-ass.org/apuntes/>

Ontología

Definición Filosófica

El término Ontología proviene de la filosofía, y proviene de dos vocablos de la lengua griega:

Ontos: Ser, Ente, Existencia

Logos: Palabra

En filosofía, Ontología es la rama que se ocupa de la naturaleza y organización de la realidad.

Es el estudio del ente. Es llamada la teoría del ser, del estudio de las cosas: Qué es, Cómo es, Cómo es posible.

Establece las categorías fundamentales de las cosas.

Representación del Conocimiento

Define el vocabulario de un área, mediante un conjunto de términos básicos y relaciones entre ellos, y las reglas que combinan términos y relaciones

Es un conjunto de términos estructurados jerárquicamente para la descripción de un dominio, que puede utilizarse como esqueleto básico de una base de conocimiento.

Provee los significados para definir explícitamente la conceptualización que está detrás del conocimiento representado en la base de conocimiento.

Es una base de datos con información sobre:

- Las categorías o conceptos que existen en el mundo o un dominio acotado.
- Las propiedades que estas categorías poseen
- Las relaciones que pueden existir entre ellas

Proveen una comprensión compartida y consensuada del conocimiento de un dominio, que puede ser comunicado entre personas y/o sistemas heterogéneos. Fueron desarrolladas para el intercambio y uso del conocimiento y pueden representar relaciones complejas entre los objetos.

Es una especificación formal y explícita de una conceptualización compartida.

Formal: debe ser representada con un lenguaje de representación formalizado y legible (por máquinas y personas).

Explícita: especificación consciente de los diferentes conceptos que conforman la ontología.

Conceptualización: modelo abstracto de uno o varios fenómenos del mundo, en el que se identifican conceptos relevantes.

Compartida: el conocimiento debe ser aceptado y consensuado entre, por lo menos, las personas que la utilizarán.

¿Por qué surgen las ontologías?

Necesidad de reutilizar conocimiento.

Existen diferentes contextos, puntos de vista y suposiciones, en un dominio determinado.

Falta de entendimiento o lenguaje común entre aplicaciones o personas que utilizan el mismo conocimiento con diferentes fines (inter-operatividad).

¿Para qué se usan las ontologías?

Reutilizar de conocimiento.

Explicitar conocimiento implícito.

Representar el conocimiento de un dominio y razonar sobre él.

Definir un vocabulario común.

Definir un entendimiento común entre agentes de software y personas.

Establecer un formato de intercambio de conocimiento y un protocolo de comunicación.

Clasificaciones de las Ontologías

Clasificaciones posibles

Existen diversas clasificaciones según el autor y las características tomadas en cuenta.

Clasificación de:

- Ontologías de un dominio.
- Ontologías genéricas.
- Ontologías representacionales.
- Ontologías de tareas.
- Ontologías de un dominio

Se representa el conocimiento de un dominio o subdominio: medicina, aplicaciones militares, cardiología, empresa, comidas, vinos, etc.

Ontologías genéricas

(Upper-Level)

Se representan conceptos generales y fundacionales del conocimiento.

Estructuras parte/todo.

Cuantificación.

Procesos.

Tipos de objetos.

Describen las categorías de más alto nivel de las entidades existentes en el mundo:

Conceptos de tiempo, espacio, materia, evento, etc.

Ontologías representacionales

(Meta-Ontologies)

Se especifica lo que tiene que ver con los formalismos de representación de conocimiento:

Clases, relaciones, funciones, axiomas, cardinalidad, dominio, rango, etc.

Ontologías de tareas

Creadas para una actividad específica dentro de un dominio:

Diagnóstico de enfermedades en el dominio de la medicina.

Pueden incluir métodos de resolución de problemas.

Criterios de diseño de ontologías

Claridad

debe comunicar de manera efectiva y precisa el significado de los términos.

Debe contener definiciones objetivas y comentadas en lenguaje natural.

Coherencia

Se deben poder hacer inferencias consistentes con las definiciones.

Compleitud:

debe ser expresada en términos necesarios y suficientes.

Máxima extensibilidad monótona:

Las especializaciones y generalizaciones deben ser incluidas en la ontología, de manera que no se requiera una revisión de las definiciones pre-existentes.

Sesgo de codificación mínimo:

Se debe especificar el conocimiento sin depender de una codificación particular a nivel de símbolo.

Principio de distinción ontológico:

Las clases deben ser disjuntas.

Modularidad:

Manteniendo mínimo acoplamiento entre módulos.

Mínimo compromiso ontológico:

Debe hacerse la menor cantidad de presunciones sobre el mundo modelado.

Áreas de aplicación de las ontologías

Existen tres grandes áreas:

- Comunicación entre personas y organizaciones.
- Interoperabilidad entre sistemas.
- Especificación de sistemas.

Áreas de aplicación

Las ontologías se usan en todos aquellos casos en que los sistemas informáticos donde se requiere una representación del dominio sobre el que actúan:

- Inteligencia artificial
- Paradigma Orientado a Objetos en la Ingeniería de software.

Algunos ejemplos de áreas de aplicación

Ingeniería del conocimiento

Gestión del conocimiento

Procesamiento del lenguaje natural

Sistemas de información cooperativos

Integración inteligente de información

Recuperación de información

Comercio electrónico

Ejemplos de ontologías

WordNet

Es un diccionario conceptual disponible gratuito en Internet.

Desarrollado por el Cognitive Science Laboratory en Princeton (1999).

La versión europea (EuroWordNet) es multilingüe:

Holandés, italiano, castellano, alemán, francés, checo y estonio.

En desarrollo: sueco, noruego, danés, griego, portugués, vascuence, catalán, rumano, lituano, ruso, búlgaro, esloveno

WordNet: estructura

Su estructura puede usarse para procesamiento del lenguaje natural y lingüística computacional.

Los sustantivos, verbos, adjetivos y adverbios, se agrupan en sinónimos cognitivos (synsets), cada uno de los cuales expresa un concepto diferente.

Estos synsets se vinculan con relaciones conceptuales semánticas

La red generada de conceptos vinculados semánticamente se puede navegar en el browser.

Lenguajes de marcado de ontologías

Hoy en día hay disponibles diversos tipos de lenguajes para la creación de ontologías, que se dividen en los siguientes grupos:

Lenguajes tradicionales de especificación de ontologías

Lenguajes Estándares y recomendados por el WC3

Lenguajes de especificación de ontologías basados en la Web.

Herramientas para la creación de ontologías

Existen diversos tipos de herramientas relacionadas con el diseño de ontologías:

- Herramientas de desarrollo de ontologías
- Herramientas de fusión e Integración de ontologías
- Herramientas de evaluación de ontologías
- Herramientas basadas en la anotación
- Herramientas de almacenamiento y de preguntas
- Herramientas de aprendizaje
- Herramientas para la creación de ontologías
- Herramientas de desarrollo de ontologías:
- Herramientas para la construcción de nuevas ontologías o bien para la reutilización de las existentes.

Componentes de una Ontología

Los componentes comúnmente identificados en una ontología son:

Clases

Son independientes del lenguaje: Ejemplo: gato y cat se refieren al mismo concepto

También llamadas “Conceptos”

Están organizadas en taxonomías.

Son ideas o conceptos básicos del dominio que se intenta formalizar.

Están relacionadas con otras clases.

Pueden ser objetos, métodos, planes, estrategias, procesos de razonamiento, etc.

Propiedades

Describen características o atributos de los conceptos de la ontología.

También llamadas “roles” o “slots”

Relaciones

Funciones

Son un tipo concreto de relación donde se identifica un elemento mediante el cálculo de una función que considera varios elementos de la ontología.

Pueden aparecer funciones como categorizar-clase, asignar fecha, etc.

Instancias

Se utilizan para representar objetos determinados de un concepto.

Es la representación ejemplar de los conceptos definidos en la ontología

Axiomas

Son sentencias lógicas cuya veracidad se asume por cierta sin ningún tipo de prueba.

Permiten junto con la herencia de conceptos, inferir conocimiento que no esté indicado explícitamente en la taxonomía de conceptos.

Permiten describir aquellas propiedades de la clase que no pueden ser definidas mediante slots y valores.

Tips a tener en cuenta en el diseño de ontologías

No existe una única forma de modelar un dominio, ni a una única aproximación lógica.

La mejor solución casi siempre depende de la aplicación, de la extensión requerida para la misma y del contexto.

El desarrollo de ontologías es necesariamente un proceso iterativo

Una ontología especifica una forma de ver el mundo, por lo cual cada ontología incorpora un punto de vista propio de su/sus autores.

Ontologías vs. Modelo O.O

Similitudes

En ambos dominios existe los conceptos de:

Clase o Concepto

Relación

Atributos o Propiedades

En ambos casos se trata de modelizar un realidad compleja, para que sea comprendida más fácilmente.

En ambos casos se piensa en la reutilización de componentes.

Diferencias:

El modelado O.O se centra principalmente en los métodos sobre las clases, basándose en las propiedades operativas de ella.

En el diseño de ontologías, el diseñador realiza estas mismas decisiones en base a las propiedades estructurales de la clase.

Las ontologías son:

Más generales

Tienen más reusabilidad.

Más fáciles de compartir.

Se acercan más a la semántica de los conceptos (la O.O se centra en las estructuras de integridad de los conceptos).

Tienen una mayor independencia de la aplicación o tarea (poseen un conocimiento genérico que puede ser reutilizado por diferentes tareas o aplicaciones).

CBR – Case Based Reasoning

Definiciones

Es una técnica que se basa en la idea de que las soluciones de problemas anteriores se pueden utilizar para resolver las nuevas situaciones.

Se utilizan algoritmos de recuperación y matching en bases de casos que consisten en pares de problema-solución pasados.

Las bases de casos y los atributos de un caso deben ser extensos para asegurar buenos resultados

Ventajas de CBR

La elicitación de conocimiento no es compleja, se restringe a recolectar los casos

La implementación consiste en definir los atributos que conforman un caso

Pueden manejar grandes volúmenes de casos utilizando tecnologías de Bases de Datos

Permite el aprendizaje: incorporación de nuevo conocimiento

Proceso de CBR

Recuperar los casos más similares

Reusar los casos recuperados para intentar resolver el problema

Revisar la solución propuesta para saber si tuvo éxito o no

Retener la solución como un nuevo caso

Este proceso por lo general requiere de la intervención humana, que no es una característica negativa de la técnica

Muchas veces se requiere una adaptación del caso recuperado para poder solucionar un problema

Definición y estructura de un Caso

¿Qué es un caso?

Es una pieza de conocimiento ubicada en un contexto, que representa una experiencia

Componentes de un caso

Problema

describe el estado de la situación en el momento que ocurre el problema

Describe el objetivo a ser alcanzado y las restricciones del objetivo

Solución

define la solución aplicada para resolver el problema

Incluye las justificaciones utilizadas para la selección o descarte de caso,

incluye el proceso de razonamiento.

La salida (Outcome)

Describe el estado de la situación luego de que se ha aplicado la solución propuesta

Incluye el feedback o evaluación de la solución dada

Puede incluir lo que se podría haber hecho para evitar el problema (reparación) y punteros a otras soluciones que se podrían haber utilizado.

Recuperación de casos

Consiste en recuperar los casos que más se parecen a la nueva situación

Es importante que el algoritmo de recuperación sea eficiente, sobre todo cuando la base de casos es muy grande

Se requiere:

- Una buena organización de la base de casos
- Identificación y selección de buenos índices
- Definición de la similitud entre dos casos
- Definición de sistemas de ranking de casos seleccionados

Indexación de casos

Son dimensiones particulares de un caso que facilitan la recuperación del mismo

Los índices son fundamentales para mejorar la recuperación de los casos

Los índices deben ser:

- Predictivos
- Dirigir hacia el propósito para el cual se utilizará el caso
- Ser suficientemente abstracto como para recuperar un caso relevante en gran variedad de situaciones futuras
- Ser suficientemente concreto como para ser reconocido con facilidad cuando ocurre la situación futura

Métodos de indexación de casos

Existen varios métodos, automáticos y manuales (los más efectivos son los manuales)

Un procedimiento común es:

- Tomar un conjunto representativo de casos
- Identificar los puntos que permiten aprender algún concepto
- Identificar las dimensiones que corresponden a esos puntos
- Describir los índices que permiten recuperar los casos en base a esas dimensiones

Almacenamiento de casos

La base de casos debe estar correctamente organizada para poder lograr una búsqueda y recuperación eficiente

Es importante lograr un equilibrio entre métodos que preserven la riqueza semántica de los casos y sus índices y que permitan el rápido acceso a los casos relevantes

Recuperación de casos

La recuperación consiste en utilizar los índices y la base de casos para recuperar los casos más similares a la situación o problema nuevo.

La recuperación debe incluir métodos heurísticos que permitan identificar casos que no son exactamente iguales pero que son parecidos, debido a que difícilmente un nuevo caso coincida 100% con uno existente

Comparación de casos

Consiste en encontrar y definir correspondencias de las dimensiones de los casos (nuevo contra todos los de la base)

Asignar una importancia a las dimensiones, no todas son relevantes

Definir un umbral que nos permita determinar cuando dos casos son similares

Establecer un ranking de casos de acuerdo a su similitud

Recuperación de casos similares

Existen varios métodos de recuperación de los casos más similares. Algunos son:

Vecinos más próximos (K Nearest neighbours)

Distancia euclidiana

Inducción

Inducción guiada

Recuperación de templates

Adaptación de casos

Una vez que se ha recuperado un caso similar al problema nuevo, se debe adaptar la solución del caso recuperado al nuevo problema

Se buscan las diferencias más importantes entre los casos y luego se aplican fórmulas o reglas que tengan en cuenta esas diferencias para sugerir soluciones.

Algunas Aplicaciones de CBR

- Clasificación: ej. Determinar el tipo de un organismo basándose en determinados atributos para decidir si aplicar tratamientos para cancer según casos pasados
- Tareas de diseño: ej. Determinar la disposición óptima de elementos en un horno
- Help Desk: se ingresa el problema de un usuario (ej. Mal funcionamiento de PC) y se extraen las mejores soluciones aplicadas en el pasado.

Redes neuronales

La red neuronal biológica

La red neuronal tiene unos 10 billones de neuronas, cada una de ellas conectada a otras 10.000 neuronas.

La neurona biológica

Cada neurona recibe señales electromecánicas a través de las dendritas

Si la suma de las señales electromecánicas es suficientemente poderosa como para activar la neurona, ésta transmite una señal electromecánica a través del axón.

Esta señal es pasada a todas las neuronas que están conectadas a través de las dendritas.

Estas neuronas receptoras pueden entonces activarse, pero solo si la suma de las señales de entrada supera cierto nivel.

Las neuronas tienen dos posibilidades: activarse o no activarse. No hay diferentes niveles de activación.

Partes de la neurona biológica

Cuerpo de la neurona

Dendritas son las ramas de extensión a través de las cuales la neurona recibe las señales de otras neuronas

El axón es lo que une dos neuronas. Algunas neuronas se comunican solo con las más cercanas y otras con miles de neuronas

Al final del axón se forman las llamadas sinapsis con las dendritas de otras neuronas. Las sinapsis son los puntos de contacto entre neuronas.

¿Cómo funciona una neurona biológica?

Una neurona envía una señal (onda de corriente) a otras a través del axón

La potencia de la señal depende del potencial de la neurona.

La neurona recoge todas las señales que llegan a ella a través de las sinapsis, que pueden ser excitatorias o inhibitorias.

La neurona suma las señales y si predominan las excitatorias envía una señal positiva

Nuestro cerebro

El cerebro está compuesto entonces de estas neuronas electromecánicas que se envían señales unas a otras.

Cada neurona realiza una tarea bien simple: sumar las entradas y generar una salida binaria enviándola a otras neuronas.

Así es como el cerebro realiza tareas muy complejas.

¿Cómo aprende la red biológica?

Cuando un axón de la célula A está cerca de producir la activación de B o toma parte en su disparo de forma persistente, tiene lugar algún proceso de crecimiento o algún cambio metabólico en una de las células o en las dos, de manera que la eficiencia de A, como una de las células que desencadena el disparo de B, se ve incrementada.

Redes neuronales artificiales

En el campo de la inteligencia artificial se refiere habitualmente de forma más sencilla como redes de neuronas o redes neuronales, las redes de neuronas artificiales (denominadas habitualmente como RNA o en inglés como: "ANN"[1]) son un paradigma de aprendizaje y procesamiento automático inspirado en la forma en que funciona el sistema nervioso de los animales. Se trata de un sistema de interconexión de neuronas en una red que colabora para producir un estímulo de salida.

Las redes neuronales artificiales no realizan las tareas complejas que es capaz de realizar nuestro cerebro.

Sin embargo, imitando la arquitectura de la red neuronal biológica, se han podido resolver muchos problemas que son simples para una persona, pero muy complejos para una máquina. Por ejemplo: reconocer una imagen.

¿Qué es una red neuronal?

Es un conjunto de unidades de procesamiento interconectados (nodos), cuya funcionalidad está basada en la neurona biológica.

La capacidad de procesamiento de la red se almacena en las conexiones entre las unidades, que tienen un peso o fuerza, que se obtiene de un proceso de adaptación a, o aprendizaje desde, un conjunto de patrones.

Es como una especie de computadora con multiprocesadores con las siguientes características:

- Elementos de procesamiento simples
- Alto grado de interconexión
- Mensajes simples
- Interacción adaptable entre los elementos

Es un conjunto de unidades de procesamiento interconectados (nodos), cuya funcionalidad está basada en la neurona biológica.

La capacidad de procesamiento de la red se almacena en las conexiones entre las unidades, que tienen un peso o fuerza, que se obtiene de un proceso de adaptación a, o aprendizaje desde, un conjunto de patrones.

El perceptrón o single-layer network

Perceptrón: La neurona artificial

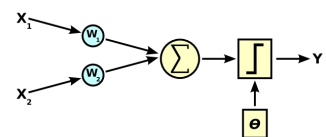
El perceptrón es un modelo matemático de una neurona biológica.

En la neurona biológica, las señales de entrada son señales electromagnéticas.

En la neurona artificial, estas entradas son valores numéricos.

El procesamiento de información de una neurona artificial puede resumirse de la siguiente manera:

- Las señales aparecen en las entradas (inputs) de las neuronas (sinapsis).
- El efecto que cada señal tendrá ante el input, puede calcularse multiplicando la señal por el peso (que indica la fuerza de la sinapsis).
- Los pesos de las señales se suman para producir una activación de la neurona
- Si la activación supera cierto valor, la neurona produce una respuesta.



Componentes de la neurona artificial

- Conjunto de Entradas X_i (inputs)
- Conjunto de Pesos W_i (weights)
- Un Umbral u (threshold)
- Una Función de activación f (activation function)
- Una salida y (output)

La función de activación

Cada tipo de función se utiliza para fines diferentes.

Unit Step: las salidas son valores discretos. Ejemplo: clasificación

Sigmoid: las salidas son una función continua.

PieceWise Linear: la salida tiene que ver con un intervalo. Ej: intervalos de tiempo

Gaussiana: la salida es una predicción.

¿Qué puede hacer un perceptrón?

Un perceptrón puede sumar las entradas y multiplicarlas por un peso.

Un ejemplo simple:

Tenemos dos entradas x e y

Tenemos dos pesos: A y B

La suma pesada será: $Ax + By$

El perceptrón da una salida diferente de cero, solo si se supera cierto valor umbral C

La salida podría ser:

1 si $Ax + By > C$

0 si $Ax + By \leq C$

Feed Forward Networks

Es una red con perceptrones organizados en capas. La primera es de entradas y la última de salidas.

Las del medio no tienen conexión con el exterior y se llaman capas ocultas (hidden layers).

Cada perceptrón de una capa se conecta con todos los de la capa siguiente.

La información es constantemente alimentada hacia adelante (feed forward), de una capa hacia la siguiente. (De ahí el nombre de las redes)

No hay conexión entre neuronas de la misma capa.

Un perceptrón puede clasificar puntos en dos regiones linealmente separables.

Considerar la siguiente red de una sola capa oculta:

Operador XOR con una red feed-forward

XOR se resuelve en una red usando los operadores NOT y OR

Tipos de redes

Acíclica, Supervisada o Feed-forward:

representa una función de las entradas actuales.

Su estado interno consiste en sus pesos.

Entrenamiento: supervisado.

Cíclica, No Supervisada o Recurrente:

permite que sus salidas realimenten sus propias entradas.

La red puede formar un sistema dinámico que puede alcanzar un estado estable, exhibir oscilaciones o tener un comportamiento caótico.

Entrenamiento: no supervisado

Tipos de Aprendizaje

Aprendizaje supervisado:

los datos para el entrenamiento son varios pares de patrones de entrenamiento de entrada y salida.

Como debe conocerse la salida, el entrenamiento necesita la supervisión de un experto.

Aprendizaje no supervisado:

Los datos para el entrenamiento son pares de patrones de entrada.

La red aprende a adaptarse a partir de las experiencias obtenidas de los patrones de entrenamiento anteriores .

Ventajas de una red neuronal

- Permiten identificar con mayor precisión datos incompletos o “sucios”.
- Cuando no se puede resolver un problema, pueden acercarse a la solución razonablemente
- Son más robustas porque el conocimiento está distribuido a través de toda la red.
- Proveen un modelo de paralelismo porque cada neurona es una unidad independiente
- El cerebro alcanza la inteligencia a través de una arquitectura neuronal

Las redes neuronales artificiales (RNA) tienen muchas ventajas debido a que están basadas en la estructura del sistema nervioso, principalmente el cerebro.

- Aprendizaje: Las RNA tienen la habilidad de aprender mediante una etapa que se llama etapa de aprendizaje. Esta consiste en proporcionar a la RNA datos como entrada a su vez que se le indica cuál es la salida (respuesta) esperada.
- Auto organización: Una RNA crea su propia representación de la información en su interior, descargando al usuario de esto.
- Tolerancia a fallos: Debido a que una RNA almacena la información de forma redundante, ésta puede seguir respondiendo aceptablemente aún si se daña parcialmente.
- Flexibilidad: Una RNA puede manejar cambios no importantes en la información de entrada, como señales con ruido u otros cambios en la entrada (ej. si la información de entrada es la imagen de un objeto, la respuesta correspondiente no sufre cambios si la imagen cambia un poco su brillo o el objeto cambia ligeramente)
- Tiempo real: La estructura de una RNA es paralela, por lo cuál si esto es implementado con computadoras o en dispositivos electrónicos especiales, se pueden obtener respuestas en tiempo real.

Funcionamiento de una red neuronal

Modos de funcionamiento

La red neuronal tiene tres fases o modos de funcionamiento:

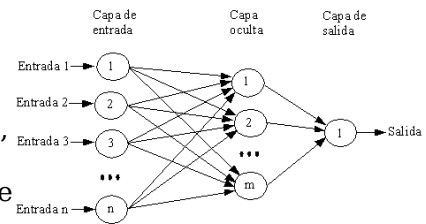
Fase de entrenamiento: la red es entrenada con los ejemplos, los pesos de las conexiones se van ajustando de acuerdo a ellos

Fase de testeo/simulación: la red se prueba con casos de prueba que no han sido utilizados para entrenar la red, de los cuales se tienen las salidas como con los casos de entrenamiento

Fase de producción: la red se utiliza para obtener un resultado no conocido.

Funcionamiento

Una de las misiones en una red neuronal consiste en simular las propiedades observadas en los sistemas neuronales biológicos a través de modelos matemáticos recreados mediante mecanismos artificiales (como un circuito integrado, un ordenador o un conjunto de válvulas). El objetivo es conseguir que las máquinas den respuestas similares a las que es capaz de dar el cerebro que se caracterizan por su generalización y su robustez.



Una red neuronal se compone de unidades llamadas neuronas. Cada neurona recibe una serie de entradas a través de interconexiones y emite una salida. Esta salida viene dada por tres funciones:

1. Una función de propagación (también conocida como función de excitación), que por lo general consiste en el sumatorio de cada entrada multiplicada por el peso de su interconexión (valor neto). Si el peso es positivo, la conexión se denomina excitatoria; si es negativo, se denomina inhibitoria.
2. Una función de activación, que modifica a la anterior. Puede no existir, siendo en este caso la salida la misma función de propagación.
3. Una función de transferencia, que se aplica al valor devuelto por la función de activación. Se utiliza para acotar la salida de la neurona y generalmente viene dada por la interpretación que queramos darle a dichas salidas. Algunas de las más utilizadas son la sigmoide (para obtener valores en el intervalo $[0,1]$) y la hiperbólica-tangente (para obtener valores en el intervalo $[-1,1]$).

Tipos de aplicaciones

En general, las redes neuronales se utilizan para diferentes tipos de problemas:

Clasificación

Diagnóstico médico, target recognition, reconocimiento de caracteres, detección de fraudes, reconocimiento de voz

Aproximación de funciones

Modelado de procesos, control de procesos, modelado de datos, diagnóstico de máquinas

Data Mining

Clustering, visualización de datos, extracción de datos