

Nombre: Carrillo Pacheco Francisco Javier

Matricula: 2143008102

Profesor: Rincón García Eric Alfredo

Materia: Técnicas heurísticas bio-inspiradas en la optimización

Tema: Tarea2 Sistema clasificador utilizando coloración de gráficas suaves

Fecha de entrega:29/01/2020

APRENDIZAJE MAQUINAL Y SISTEMAS CLASSIFICADORES

Una de las aplicaciones de la inteligencia artificial consiste en el aprendizaje automático mediante distintos algoritmos para el reconocimiento de patrones, la clasificación de diferentes tipos de objetos físicos o abstractos con la finalidad de analizar y obtener información útil. La ventaja de la automatización del aprendizaje es que el algoritmo se vuelve autoadaptable a diferentes tipos de problemas aprendiendo a lo largo de la ejecución.

El aprendizaje automático tiene una amplia gama de aplicaciones: En el área de la medicina se a utilizado para diagnosticar enfermedades, motores de búsqueda, fraudes y análisis del mercado de valores, en general clasificadores.

Los clasificadores son un tipo especial de reconocimiento de patrones, el objetivo de estos sistemas es enseñarles a las computadoras en ordenar una serie de objetos reales o inanimados, base a características y etiquetas previas en una base de conocimiento.

La coloración de las gráficas suaves es una generalización y representación de problemas de clasificación en el que se busca encontrar reducir la suma de las distancias entre vértices de nodos idénticos. El reconocimiento de patrones se refiere al problema de encontrar la estructura interna de los objetos etiquetados. Una de las formas para llevar a cabo la tarea es definir una distancia entre dos objetos, de esta manera si dos objetos están muy cerca respecto a una métrica dada, es muy probable que estén en la misma clase. Si los objetos están lejos uno del otro, van a estar en clases diferentes.

CLASIFICADOR AGRUPADOR (NO SUPERVISADO)

Un clasificador es la abstracción humana de reconocer objetos, clases o categorías base a patrones y características propias de dichos objetos. Para alcanzar una buena solución, previamente en el sistema se introduce un conocimiento del problema reduciendo el costo computacional de la búsqueda de la solución.

El objetivo es encontrar la clase más adecuada para el conjunto sin etiquetar. Utilizando el grado de similitud en sus atributos podemos determinar si los objetos van a una misma clase o en caso de encontrarse diferencias, colocarlos en distintas clases.

TIPOS DE CLASIFICADORES

Los clasificadores se dividen por la forma en que van aprendiendo:

Bayes ingenuo: Aplica el teorema de Bayes, esta asume independencia probabilística entre cada atributo de los objetos

Parzen: Utiliza el concepto de líneas de características, este permite generalizar el espacio de similitudes o diferencias, posteriormente estimar densidades de probabilidad mediante el método Parzen

Red Neuronal: Tratando de seguir el comportamiento de una neurona biológica. Utilizando el modelo de percepción multicapa para la clasificación tenemos un conjunto de señales de entrada multiplicadas por sus pesos correspondientes y luego sumadas a cada una de las nuevas entradas en el proceso de sinapsis o nivel de activación de la neurona

COLORACIÓN DE GRAFICAS SUAVES

Los algoritmos de agrupamiento (clustering) son un tipo de clasificadores no supervisados. Consiste en dividir un conjunto de datos dentro de k grupos de tal modo que los datos de un mismo grupo sean lo mas parecido posible entre ellos, los grupos pueden estar determinados por el usuario o puede decidirlo el propio algoritmo.

La coloración de graficas suaves es un problema de etiquetado de grafos. Consiste en colocar una etiqueta a un vértice del grafo de tal modo que ningún vértice adyacente contenga el mismo color.

En ella se busca reducir la dureza de la coloración, es decir, una coloración que minimice la suma de las penalizaciones entre las aristas adyacentes con vértices del mismo color.

CLASIDICADORES HILL CLIMBIN Y RECOSIDO SIMULADO CON IRIS

Utilizando la base de datos Iris se crearon dos clasificadores, la primera con el método de Hill climbing y el segundo con el método recosido simulado.

Con ayuda de nuestros generadores de números aleatorios le dimos el etiquetado a nuestros valores de nuestra base de datos y base a esto se efectuaron 30 corridas efectivas, cada iteración será corrida con una semilla diferente para nuestros generadores de números aleatorios, para determinar datos estadísticos para determinar y comparar la calidad de los algoritmos construidos en clase.

La semilla para nuestro generador de numero aleatorio, cambiaba en un orden de 300 en cada iteración. Para ver claramente estas comparaciones se realizará diagramas de cajas para ambos algoritmos

|  |  |  |
| --- | --- | --- |
| Hill climbing | Recocido Simulado |  |
| 5745.6 | 5745.6 |  |
| 5745.6 | 5745.6 |  |
| 5745.6 | 5745.6 |  |
| 5745.6 | 5745.6 |  |
| 5745.6 | 5745.6 |  |
| 5745.6 | 5745.6 |  |
| 5745.6 | 5745.6 |  |
| 5745.6 | 5745.6 |  |
| 5745.6 | 5745.6 |  |
| 5745.6 | 5745.6 |  |
| 5745.6 | 5745.6 |  |
| 5745.6 | 5745.6 |  |
| 5745.6 | 5745.6 |  |
| 5745.6 | 5745.6 |  |
| 5745.6 | 5745.6 |  |
| 5745.6 | 5745.6 |  |
| 5745.6 | 5745.6 |  |
| 5745.6 | 5745.6 |  |
| 5745.6 | 5745.6 |  |
| 5745.6 | 5745.6 |  |
| 5745.6 | 5745.6 |  |
| 5745.6 | 5745.6 |  |
| 5745.6 | 5745.6 |  |
| 5745.6 | 5745.6 |  |
| 5745.6 | 5745.6 |  |
| 5745.6 | 5745.6 |  |
| 5745.6 | 5745.6 |  |
| 5745.6 | 5745.6 |  |
| 5745.6 | 5745.6 |  |
| 5745.6 | 5745.6 |  |
| 5745.6 | 5745.6 | Promedio |
| 1.60195E-11 | 6.38436E-12 | Desviación estándar |
| 5745.6 | 5745.6 | Mejor solución |
| - | - | Peor solución |
|  |  |  |

Esto quiere decir que sin importar como se va variando la semilla de nuestro generador de ambos algoritmos encuentran la mejor solución para nuestra base de datos IRIS. En mi caso el algoritmo Hill climbing se tarda mas que el algoritmo de recosido simulado