Primer parcial Inteligencia artificial

Preguntas parcial 2014

¿Qué es un histograma y que utilidad presenta la ecualización de histogramas?

El histograma de niveles de gris de una imagen digital corresponde a la distribución de probabilidad de intensidad de la imagen.

Un histograma es un gráfico que representa las frecuencias relativas de los niveles de gris de una imagen. El histograma es empleado entre otras cosas para el cálculo de coeficientes estadísticos tales como el valor medio, la desviación estándar y la entropía.

Si tenemos una imagen digital con niveles de grises en el rango [1,L] su histograma es una función discreta de valores,

$$h(g_i) = \frac{n_i}{n} \qquad i = 1,2,3,\ldots,L\ (Distribución\ de\ Frecuencia)$$

---

iendo

$$g_i = el\ i - esimo\ nivel\ de\ gris$$

$$n_i = número\ de\ pixeles\ en\ la\ imagen\ con\ ese\ nivel\ de\ gris$$

$$n = número\ total\ de\ pixeles\ de\ la\ imagen$$

La función h representa la probabilidad de que un pixel elegido al azar pertenezca a un determinado nivel de gris.

A partir del histograma podemos obtener también valores estadísticos como el valor medio de niveles de gris:

$$mg = \sum g\, h(g)$$

y la desviación estándar

$$\sigma^2 = \sum (g - mg)^2\, h(g)$$

Utilidad de ecualización de histograma:

La ecualización es un procedimiento de mejoramiento de imágenes que utiliza el histograma de niveles de grises de una imagen. Justifica su aplicación en imágenes de bajo contraste con áreas grandes de niveles medios de gris. Este proceso se basa en una transformación g→g'. Se reacomodan las líneas del histograma con nuevos intervalos Δg de manera de obtener niveles de densidad correspondientes entre todos los valores de g y g'.

Basándonos en

$$\frac{h(g')}{\Delta g'} = constante \qquad \text{y} \quad g' \text{ niveles de gris de la imagen transformada.}$$

Si tenemos

$g =$ niveles de gris
$h(g) =$ frecuencia de niveles de gris
$\Delta g = k \cdot h(g)$ $\quad k =$ número de niveles de gris sobre la sumatoria de $h(g)$

El nuevo histograma es $h(g')$ donde tratamos los valores originales de $h(g)$.

De esta manera, cuanto más alta sea $h(g)$ mayor será la separación entre las barras de $h(g')$.

Suma de distribución de $h(g)$: $\sum h(g) = 0{,}3 + 0{,}5 + 0{,}2 = 1$

Constante $k = \dfrac{niveles\ de\ gris}{\sum h(g)} = 10$

$$g_{nuevo} = g_{viejo} + k \cdot h(g)$$

¿Para que se usa la convolución de imágenes digitales? Explicar

La convolución de imágenes digitales es un método mediante el cual se puede hacer resaltar o disminuir la presencia de ciertos valores en una región deseada de una imagen.

Las aplicaciones más importantes de la convolución es el suavizado de imágenes que consiste en la pérdida de nitidez de la imagen y en la detección de bordes.

Consiste en recalcular el valor de un pixel en base a los pixeles de su vecindad.

El método consiste en definir una máscara cuya dimensión sea igual a la dimensión de la vecindad elegida para procesar la imagen.

Considerando a los vecinos de un pixel representados por la matriz $G(x,y)_{m\,x\,n}$ y la máscara a aplicar $F(x,y)_{m\,x\,n}$.

Calculamos el nuevo valor del pixel *P*, perteneciente a la matriz *G*, como la suma de los productos entre cada uno de los elementos de $G(x,y)_{m\,x\,n}$ por cada uno de los elementos de $F(x,y)_{m\,x\,n}$ y reemplazamos el valor de *P* por el nuevo valor calculado *P´*.

$$P' = \sum g(x,y) \cdot f(x,y)$$

3. Convolución de imágenes digitales definición, ejemplo, Mascaras, Aplicaciones.

La convolución de imágenes digitales es un método mediante el cual se puede hacer resaltar o disminuir la presencia de ciertos valores en una región deseada. Consiste en definir una máscara de dimensión igual a la dimensión de la vecindad elegida para procesar la imagen. Luego para calcular el valor de un pixel P' de la imagen transformada se debe realizar la suma de los productos entre los componentes, de la imagen y de la máscara, que ocupan la misma posición en la imagen, y reemplazar el pixel de referencia P de la vecindad por el nuevo valor.

Porción de imagen / Máscara / Resultado

$$\begin{bmatrix} a & b & c \\ d & P & e \\ f & g & h \end{bmatrix} \quad \begin{bmatrix} 1 & 0 & -1 \\ 1 & 0 & -1 \\ 1 & 0 & -1 \end{bmatrix} \quad \begin{bmatrix} a & b & c \\ d & P' & e \\ f & g & h \end{bmatrix}$$

con $P' = 1a + 0b - 1c + 1d + 0P - 1e + 1f + 0g - 1h$

Las aplicaciones más importantes de la convolución es el suavizado de imágenes q consiste en la perdida de nitidez de la imagen y en la detección de bordes.

Para esto se usan distintos tipos de mascaras.

Prewit / Sobel

$$\begin{bmatrix} -1 & 0 & 1 \\ -1 & 0 & 1 \\ -1 & 0 & 1 \end{bmatrix} \begin{bmatrix} -1 & 0 & 1 \\ -2 & 0 & 2 \\ -1 & 0 & 1 \end{bmatrix}$$ Bordes Verticales

Prewit / Sobel

$$\begin{bmatrix} 1 & 1 & 1 \\ 0 & 0 & 0 \\ -1 & -1 & -1 \end{bmatrix} \begin{bmatrix} 1 & 2 & 1 \\ 0 & 0 & 0 \\ -1 & -2 & -1 \end{bmatrix}$$ Bordes Horizontales

-Transformada de Hough. Explique el significado de la recta en el plano secundario.

11. Transformada de Hough. Como se calcula y para qué sirve?

La transformada de Hough es un método que resulta útil para la detección de ciertas morfologías simples de los bordes tales como bordes rectilíneos o círculos. Modo de operación es principalmente estadístico y consiste en que para cada punto que se desea averiguar si es parte de una línea se aplica una operación dentro de cierto rango, averiguando cuales son las posibles líneas que pasan por él. Esto se continúa para todos los puntos en la imagen, al final se determina qué líneas fueron las que más puntos posibles tuvieron y esas son las líneas en la imagen.

Por ejemplo: dado el punto $P_i = (x_i,\ y_i)$ sabemos que $y_i = ax_i + b$ es la familia de ecuaciones de todas las rectas que pasan por él. Aplicando la Transformada de Hough obtenemos $b = y_i - ax_i$, que es la ecuación de la recta que representa a todas las rectas que pasan por P_i en el dominio (a, b). Es decir, cada punto de esta recta en (a, b) representa una recta en (x, y). Si hay varios puntos en (x, y) graficando las rectas $b = y_i - ax_i$ para cada uno de ellos podríamos observar el punto en el que el mayor número de rectas se cortan, igualar las ecuaciones de dos de esas rectas nos permitiría entonces hallar a y b de la recta en (x, y) que más puntos une.

¿Qué es el reconocimiento de patrones y que tipos de funciones de decisión... Para clasificar las entradas? Explicar

El reconocimiento de patrones puede ser concebido como la actividad de discriminar, clasificar o categorizar la información de entrada, no entre patrones individuales sino entre poblaciones, por medio de búsqueda de características o atributos invariantes entre los miembros de una población.

Definimos una **clase** como una categoría determinada por algunos atributos comunes.

Un **patrón** como la descripción de cualquier miembro de una categoría que representa a una clase de patrones.

Un patrón es un conjunto de características. Una clase de patrones es un conjunto de patrones "similares". El objetivo del reconocimiento de patrones es asignar un patrón a la clase que pertenece. Los patrones se pueden almacenar en vectores patrón (para características cuantitativas):

$$\boldsymbol{x} = \begin{bmatrix} x_1 \\ x_2 \\ \dots \\ x_n \end{bmatrix}$$

donde $\boldsymbol{x}$ es el patrón y x_i son las características.

Los vectores patrón son definidos como vectores columna pero pueden ser expresados por ser traspuesta. X= (x1, x2, …, Xn)'

Funciones de decisión lineal:

La función principal de un sistema de reconocimiento de patrones es tomar decisiones acerca de la pertenencia a una clase dada de los patrones con que se lo confronta.

Considerar $d(\boldsymbol{x}) = w_1x_1 + w_2x_2 + w_3 = 0$ como la ecuación de una línea de separación, donde $\boldsymbol{w}$ es el vector de parámetros y $\boldsymbol{x}$ es el vector de las variables de la entrada.

Clasificador Polinomial

El clasificador lineal utilizado para el reconocimiento de patrones es inadecuado para resolver problemas de clasificación de mayor complejidad, en particular casos tales como clasificación no lineal y clasificación donde la información está contenida en vectores de dimensiones de más de dos dígitos.

Para poder comprender el clasificador no lineal a partir del concepto de función de decisión línea, se introduce el concepto de funciones de decisión generalizada, lo cual significa aplicar funciones a cada una de las componentes del vector, tomándose estas como las componentes de un nuevo vector. Si al menos una de estas funciones es no lineal, se obtiene una fusión de clasificación no lineal.

Dado un vector $\boldsymbol{x} = (x_1, x_2, \dots, x_n)'$ podemos transformarlo como

$$\boldsymbol{x}^* = (f_1(x_1), f_2(x_2), \dots, f_n(x_n))' = (x_1^*, x_2^*, \dots, x_n^*)'$$

Donde el superíndice (*) indica las nuevas componentes del vector luego de aplicar las funciones $f_i(x_i)$.

El caso más simple se da cuando estas funciones son lineales, es decir, cando las funciones de decisión se definen simplemente como $f_i(x_i) = x_i$. Bajo esta condición se obtiene la expresión lineal

$$d(x) = w_0 x + w_{n+1}$$

El otro caso se plantea cuando se tiene funciones cuadráticas. En el caso bidimensional transformamos los patrones según la expresión:

$$x^* = (x_1, x_2)^2 = (x_1^2, x_2^2, x_1 x_2, x_1, x_2, 1) = (x_1^*, x_2^*, x_3^*, x_4^*, x_5^*, 1)$$

Lo cual se puede explicar a partir de la expansión del vector $x = (x_1, x_2, 1)$ y el cálculo de las combinaciones posibles de sus elementos tomando de 2 en 2.

Con esta transformación la función polinomial se segundo orden toma la forma:

$$d(x^*) = w_1 x_1^2 + w_2 x_2^2 + w_3 x_1 x_2 + w_4 x_1 + w_5 x_2 + w_6 1$$

En el caso bidimensional puede generalizarse mediante una expresión que considere todas las combinaciones posibles de los componentes de **x** que forman términos de segundo orden o inferior, es decir:

$$d(x) = \sum_{j=1}^{n} w_{jj} x_j^2 + \sum_{j=1}^{n-1} \sum_{k=j+1}^{n} w_{jk} x_j x_k + \sum_{j=1}^{n} w_{jj} x_j + w_{n+1}$$

Donde para cada una de las funciones puede definirse f(x)= $x_p^s \, x_q^t$ con p y q=1,2,…n y s, t=0,1 .

Extensión al espacio n-dimensional

El caso anterior puede ser fácilmente extendido a n-dimensiones definiendo una función de decisión lineal general de la forma:

$$d(\boldsymbol{x}) = w_1 x_1 + w_2 x_2 + \cdots + w_n x_n + w_{n+1}$$

Donde el vector $\boldsymbol{w_0} = (w_1, w_2, \dots, w_n)'$ es llamado *vector de pesos o vector de parámetros.*

Es además una convención muy difundida adicionar un 1 como última componente de todos los vectores patrones con lo que la igualdad anterior puede expresarse como:

$$d(\boldsymbol{x}) = \boldsymbol{w_0 x}$$

Donde $\boldsymbol{x} = (x_1, x_2, \dots, x_n, 1)'$ es llamado *vector patrón aumentado.*

¿Cuáles son las diferencias entre la red perceptron y la red adaline ?

7. Diferencia entre red Perceptrón y Adaline. (Acá mencionar la ley Delta y el cambio de F de escalón a lineal)

La principal diferencia entre Perceptrón y Adaline es la función de Salida.
En Perceptrón la salida es binaria, en Adaline es real
El Perceptrón utiliza la salida de la función umbral (binaria) para el aprendizaje. Sólo se tiene en cuenta si se ha equivocado o no.
En Adaline se utiliza directamente la salida de la red (real) teniendo en cuenta cuánto se ha equivocado.
Otra diferencia importante es la ley de aprendizaje q ambos utilizan.
Perceptrón usa la Ley de Aprendizaje q lleva su nombre.

$$\Delta w_{ij} = (y'_j - y_j)x_i$$

Adaline en cambio utiliza la Ley delta que es una combinación de la ley de Hebb y la de Perceptrón

$$\Delta w_{ij} = l_r(y'_j - y_j)x_i = l_r \delta x_i$$

En Adaline hay una razón de aprendizaje y para regular lo que va a afectar cada equivocación a la modificación de los pesos. Es siempre un valor entre 0 y 1 para ponderar el aprendizaje

Otras posibles preguntas

1. Que es un histograma. Explicar ecualización de histogramas.

Un histograma es un grafico que representa las frecuencias relativas de los niveles de gris de una imagen. Este es utilizado para el cálculo de coeficientes estadísticos tales como: el valor medio, la desviación estándar y la entropía. También es utilizado para aplicar un procedimiento de mejoramiento de imágenes denominado ecualización o igualación de histograma.
Este proceso se basa en una transformación g -> g' de manera de lograr una reacomodación de las líneas del histograma en nuevos intervalos Δg. Basándonos en: $\frac{\mathrm{h(g)}}{\Delta \mathrm{g}} = cte.$
Así, cuanto más alto sea h(g) mayor será la separación entre las barras.
Para el cálculo de el nuevo valor de g usaremos la siguiente fórmula: $g_{nuevo} = g_{previo} + k\ h(g)$

2. Explicar filtros de imagen (Filtro mediana, Filtro valor medio)

Estos operadores calculan un pixel para la imagen transformad, en función de la vecindad local en la imagen original.
El filtro de la mediana consiste en, por ejemplo tomar una ventana de 1x5, reemplazar el valor central por el valor de la mediana de los números q la componen. (La mediana se calcula ordenando los elementos y tomando el valor q queda al centro)
El filtro de la media es similar pero se reemplaza el valor del medio por el promedio del conjunto

3. Convolución de imágenes digitales definición, ejemplo, Mascaras, Aplicaciones.

La convolución de imágenes digitales es un método mediante el cual se puede hacer resaltar o disminuir la presencia de ciertos valores en una región deseada mediante el uso de mascaras.
La convolución discreta consiste en realizar la suma de los productos entre los componentes, de la imagen y de la máscara, que ocupan la misma posición en la imagen, y reemplazar el pixel de referencia de la vecindad por el nuevo valor.
Las aplicaciones más importantes de la convolución es el suavizado de imágenes q consiste en la perdida de nitidez de la imagen y en la detección de bordes.
Para esto se usan distintos tipos de mascaras.

Prewit Sobel

$$\begin{bmatrix} -1 & 0 & 1 \\ -1 & 0 & 1 \\ -1 & 0 & 1 \end{bmatrix}\begin{bmatrix} -1 & 0 & 1 \\ -2 & 0 & 2 \\ -1 & 0 & 1 \end{bmatrix}$$ Bordes Verticales

Prewit Sobel

$$\begin{bmatrix} 1 & 1 & 1 \\ 0 & 0 & 0 \\ -1 & -1 & -1 \end{bmatrix}\begin{bmatrix} 1 & 2 & 1 \\ 0 & 0 & 0 \\ -1 & -2 & -1 \end{bmatrix}$$ Bordes Horizontales

4. Clasificador polinomial: Fundamentos, graficar un clasificador polinomial con vector de entradas de 2do orden y explicar el funcionamiento

En casos en que deben resolverse problemas de clasificación tales como clasificación no lineal o clasificaciones donde la información está contenida en vectores de dimensiones de dos dígitos debe introducirse el concepto de funciones de decisión generalizada, lo cual significa aplicar funciones a cada una de las componentes del vector de patrones, tomando a estas como las componentes de un nuevo vector

Es decir q dado $\boldsymbol{x}= (x_1; x_2; x_3;\ldots;x_n)$ se transformara como $\boldsymbol{x}^* =(f_1(x_1);f_2(x_2);f_3(x_3);\ldots;f_n(x_n)) = (x_1^*; x_2^*; x_3^*;\ldots;x_n^*)$.

En el caso en que la función a aplicar es la función cuadrática transformamos los patrones según la expresión:

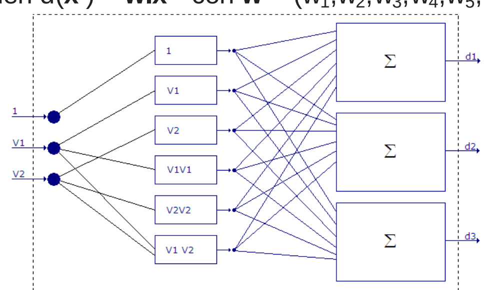
$$X^* = (x_1;x_2)^2 = (x_1^2;x_2^2;x_1x_2;x_1;x_2;1) = (x_1^*; x_2^*; x_3^*; x_4^*; x_5^*;1)$$

La cual surge del vector de patrones expandido $\boldsymbol{x}= (x_1; x_2; 1)$ y el cálculo de las combinaciones posibles de sus elementos tomados de 2 en 2.

Con esta transformación la función polinomial toma la forma:

$$D(x^*) = w_1.x_1^2 + w_2.x_2^2 + w_3.x_1x_2 + w_4x_1 + w_5x_2 + w_61$$

$$\text{O bien } d(\mathbf{x}^*) = \mathbf{w.x}^* \quad \text{con } \mathbf{w} = (w_1,w_2,w_3,w_4,w_5,w_6).$$



En el aumento de las características de los problemas reales se hace inviable el cálculo de la matriz de coeficiente, por lo que se aplica otro procedimiento. Se introduce un nuevo vector $z = (x_1,x_2,x_3,\ldots,x_n, y_1,y_2,y_3,\ldots,y_m)$ y una nueva matriz M_z.

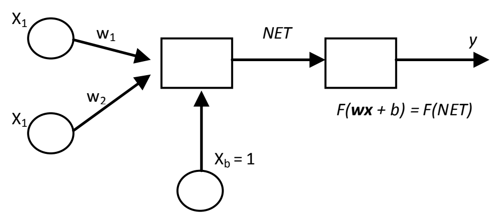
$$M_z = E\left\{zz^T\right\} = E\left\{\begin{pmatrix} x \\ y \end{pmatrix}\begin{pmatrix} x^T & y^T \end{pmatrix}\right\} = \begin{pmatrix} E\left\{xx^T\right\} & E\left\{xy^T\right\} \\ E\left\{yx^T\right\} & E\left\{yy^T\right\} \end{pmatrix} = \begin{pmatrix} I & w \\ 0 & e \end{pmatrix}$$

Para resolver el clasificador polinomial simplemente debemos construir la matriz M y aplicar transformaciones por filas.

5. Perceptrón: Dibujar el Perceptrón con 2 entradas, explicar proceso y fase de prueba que utiliza, dibujar el resultado en el plano cartesiano, si se modifica la función escalón, a qué tipo de red se convierte

Perceptrón es un esquema que resuelve el problema de clasificación lineal de patrones y que se corresponde con el modelo biológico de una neurona humana.

En este modelo en vez de expandir el vector **x** se adiciona una entrada a la neurona que siempre recibe valor 1



Al recibir el vector **x**, la red calcula la función de decisión lineal y le aplica 1 función umbral que la fuerza a obtener solo 2 valores de salida: 0 y 1, o 1 y -1 según como se defina la función F.

El vector **w** se calcula por un proceso iterativo de aproximación en vez de usar cálculos matriciales.

La ley de aprendizaje del Perceptrón considera que el ajuste de los coeficientes no interviene directamente en la salida **y**, sino la diferencia entre la salida obtenida al procesar la información y la salida deseada. O sea:

$$\Delta w_{ij} = x_i (y_i - y_j)$$

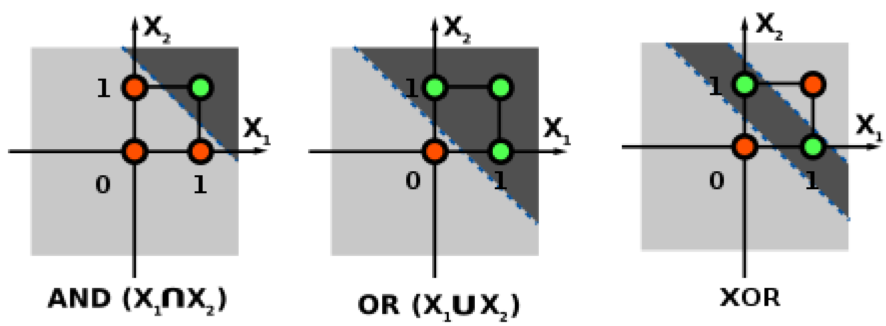
/ y_i = salida obtenida

y_j = salida deseada

de esta manera el nuevo valor de **W** se obtiene

$$\mathbf{W}_i = \mathbf{W}_{i-1} + x_i (y'_i - \mathbf{W'}_{i-1} x_i)$$

Si w1 y w2 son dos clase entre las q se quiere clasificar a **x**, si xЄw1 y w(k).x(j) ≤ 0 o xЄw2 y w(k).x(j) ≥ 0 se debe modificar W, sino no.



Al reemplazar la función escalón por una función rampa se obtiene el modelo ADALINE.

6. Perceptrón lineal (adaline) fundamentos, esquema, entrenamiento, se puede utilizar adaline en redes multicapa, como?

Se usa para mejorarla resolución de problemas no lineales usando redes multicapa.

Si reemplazamos la función de función de salida abrupta de un Perceptrón por una función rampa, podremos generalizar el Perceptrón para E/S continuas y además aplicar la técnica del gradiente o descenso iterativo.

Usamos una ley de aprendizaje denominada Ley Delta o Ley de Widrow/Hoff, que combina las de Hebb y la de Perceptrón:

$$\Delta w_{ij} = l_r (y_i - y_j).x_i = l_r.\delta\ x_i$$

En vez de Perceptrón los llamaremos ADALINAS

Ahora consideremos el error cuadrático medio

$$\text{LMSE} = \sqrt{\Sigma\left(y'_i - y_i\right)^2} = 1\,\frac{1}{2L}\Sigma(Y_k - O_k)^2$$

Siendo: Y_k = salida deseada

O_k = salida e la red

En cada paso modificaremos el valor de w hasta alcanzar el optimo haciendo mínimo el error.
Si trabajamos en Rn de orden n-1, la superficie del grafico del error sería muy irregular por lo que deberemos conformarnos con encontrar una solución lo suficientemente buena.

Existe una versión multicapa de la ADALINE denominada MADALINE (Múltiple ADALINE, múltiples Adalides) que consiste en una red neuronal con neuronas similares a las de la ADALINE pero que contiene capas de neuronas ocultas.