**Maestría en Ingeniería con Énfasis en Ingeniería de Sistemas y Computación**

**Introduction to Pattern Recognition for Computer Vision**

**Universidad del Valle**

**Harold Armando Achicanoy Estrella**

**Resumen de Temas Vistos**

En la siguiente tabla se presenta un resumen de los métodos vistos hasta el momento, los cuales cubren los métodos generativos y discriminativos.

|  |  |  |  |
| --- | --- | --- | --- |
| **Método** | **Funcionamiento** | **Supuestos** | **Aplicación** |
| **Métodos generativos** | | | |
| Máxima verosimilitud y estimación bayesiana de parámetros | Métodos que permiten estimar probabilidades condicionadas y no condicionadas desde los puntos de vista frecuentista y bayesiano. Desde ambos punto de vista, se asume (y se tiene que verificar) la distribución de probabilidad asociada o esta tiene que ser estimada a partir de repeticiones (frecuentista) o dada por el conocimiento previo (bayesiano) | - Independencia entre las observaciones | Base probabilística para la construcción de modelos generativos y discriminativos, estimación de parámetros |
| Análisis de componentes principales | A partir de un conjunto de variables correlacionadas, el objetivo es construir unas nuevas variables independientes (no correlacionadas) a partir de combinaciones lineales que permitan describir las principales características de los datos de entrada en un espacio de menor dimensión | - Independencia entre las observaciones | Reducción de dimensionalidad |
| Análisis discriminante lineal de Fisher | Al igual que el análisis de componentes principales la idea principal de este método consiste en generar unas nuevas variables (combinaciones lineales de las originales) que permitan separar de forma máxima las clases para obtener clasificaciones con alta precisión | -Normalidad dentro de cada clase  -Homogeneidad entre matrices de varianzas-covarianzas | Clasificación |
| Algoritmo EM | Algoritmo usado para maximizar la función de verosimilitud bien sea en un proceso de estimación de parámetros o clasificación. Y consta de dos pasos, el primero denominado Esperanza, donde se busca estimar el valor esperado de los parámetros a partir de la información disponible en el modelo, posteriormente, se aplica el paso de Maximización donde a partir de la estimación previa se actualiza la función de verosimilitud buscando maximizarla hasta conseguir convergencia | - Independencia entre las observaciones  - Verificar que los datos siguen la distribución de probabilidad propuesta | Clasificación, estimación de parámetros, imputación de datos faltantes |
| Estimación no paramétricas de densidades | Sin necesidad de realizar supuestos sobre la distribución que siguen los datos, el objetivo de los presentes métodos, específicamente los métodos basados en Kernel buscan estimar una función de densidad suavizada teniendo en cuenta la concentración de datos en el rango de la variable de interés | - Independencia entre las observaciones | Estimación de densidades |
| **Métodos discriminativos** | | | |
| Clasificación por los vecinos más cercanos | Método no paramétrico que no asume ningún supuesto sobre los datos, en el cual se busca identificar cuáles son las observaciones más cercanas a una determinada observación (en términos de distancia) y a partir de estas se genera una clasificación, como el voto máximo en el caso de variables categóricas o el promedio para las variables numéricas |  | Clasificación |
| Métricas de distancia | Las medidas de distancia son de utilidad para diversas metodologías que van desde la clasificación hasta la segmentación o agrupamiento (más conocido como clustering). La idea principal de dichas métricas consiste en medir que tan alejados o cercanos son los observaciones de una tabla de datos con el fin de generar agrupamientos o segmentar los individuos a partir de su similaridad |  | Clustering |
| Clasificación Fuzzy | A partir de la construcción de funciones de pertenencia a determinadas clases de interés, el objetivo de estos métodos consiste en asignar un grado de creencia a la clasificación final de una observación en una clase determinada. Esto es de utilidad cuando los problemas que se están trabajando no presentan una distinción clara entre las clases |  | Clasificación, clustering |