

# FKDC - Fermat Kernel Density Classifier

## Análisis

A fines de evaluar las bondades de la técnica propuesta, evaluaremos diferentes *estimadores* a través de una *métrica de performance* en distintas *tareas* de clasificación. La métrica más obvia, por caso, sería la *exactitud* (“accuracy”) de un estimador: el % de observaciones de evaluación clasificadas en la categoría correcta. Otras podrán tomar su lugar.

Para comparar equitativamente diferentes *algoritmos* de clasificación, elegiremos un *estimador* mediante la búsqueda de un set de *hiperparámetros óptimos* por *validación cruzada* en un *espacio de búsqueda* de hiperparámetros acorde a las necesidades del algoritmo.

A priori, quisiéramos comparar FKDC con su primo hermano, el GKDC, clasificador for KD gaussiano, caso particular del FKDC, cuando  $\alpha == 1$

A fin de obtener resultados de utilidad en el contexto más amplio de las tareas de clasificación, incluiremos entre los algoritmos a evaluar algunos de uso bien extendido en el ámbito industrial y académico, y otros cercanos a FKDC en el conjunto de supuestos y técnicas utilizadas:

- regresión lineal
- SVC
- GBTs
- NNs simples
- Naive Bayes

Para tener una idea “sistémica” de la performance de los algoritmos, evaluaremos su performance con diferentes *datasets*. Muchos factores en la definición de un dataset pueden afectar la exactitud de la clasificación; nos interesará explorar en particular 3 que a su vez figuran en el cálculo de la densidad en variedades:

- $n$ , el tamaño de la muestra,
- $d$ , la dimensión de las observaciones y
- $k$ , la cantidad de categorías.