

Trabajo de minería de datos

Objetivo: Encontrar en menor número de variables que permitan lograr la mayor tasa de reconocimientos

Los datos: El experimento E-GEOD-3307 que consiste de 121 muestras de tejido muscular de un total de 11 condiciones además de la condición normal. Éstas son:

- Acute quadriplegic myopathy
- Juvenile dermatomyositis
- Amyotrophic lateral sclerosis
- Spastic paraplegia
- **Fascioscapulohumeral muscular dystrophy**
- Emery Dreifuss muscular dystrophy
- Becker muscular dystrophy
- Duchenne muscular dystrophy
- Calpain 3
- Dysferlin
- FKRP

Dos clases: FMD (14 muestras) y Normal (18 muestras) por 22,283 variables

Trabajo de minería de datos

Los métodos:

- Reducción de variables
- Algoritmos de clasificación: Naive Bayes, kNN, LDC, QDC, Neural Network
- 10 xvalidation
- Listado de variables finales
- Visualización del modelo usando PCA
- Generar un PDF con formato de artículo con el experimento
- USEN TODO LO QUE MATLAB LES OFRECE
- EN EQUIPO
- Una semana