

## INTRODUCCIÓN A SISTEMAS INTELIGENTES

---

*Profesor Adjunto:* Ochoa María Alejandra

*Ayudante:* Vazquez, Diego Anibal

### **Trabajo Práctico: “Redes Neuronales basadas en Clustering”**

- 1) Aplicar el algoritmo de red neuronal SOM, mediante WEKA, sobre una base de datos del repositorio University of California at Irvine Machine Learning Repository (<https://archive.ics.uci.edu/ml>). Para utilizar el algoritmo será necesario descargar e instalar en la aplicación el plugin: [https://sourceforge.net/projects/wekann/files/SelfOrganizingMap/SelfOrganizingMap1.0.3.zip/download?use\\_mirror=ufpr&download=](https://sourceforge.net/projects/wekann/files/SelfOrganizingMap/SelfOrganizingMap1.0.3.zip/download?use_mirror=ufpr&download=)
- 2) Se pretende que el alumno realice la preparación de los datos (Fase III – Preparación de los datos – Metodología CRISP – DM) que sea necesaria para el correcto procesamiento mediante la técnica de minería de datos.
- 3) Realice un informe donde indique:
  - Descripción de los datos: cantidad de observaciones, nombre de los atributos, rango de valores, tipo de variable (categórico, continuo o clase), indicar el significado de los valores que asume el atributo. Describir puntos clave a tener en cuenta en el análisis posterior observando los gráficos de distribución de variables y distribución de la clase.
  - Especificar parámetros de aprendizaje: Número de variables, cantidad de neuronas del mapa, cantidad de ciclos de entrenamiento, learning rate.
  - Informe final con los resultados obtenidos
    - Mencionar la cantidad de neuronas obtenidas y la cantidad de observaciones en cada una.
    - Gráfico con los valores promedio de cada atributo variable, y el análisis correspondiente.
    - Analizar el clúster con la menor cantidad de observaciones, determinar diferencias con los demás.
- 4) La base de datos será la del repositorio correspondiente a pacientes indios con problemas de hígados: <https://www.kaggle.com/uciml/indian-liver-patient-records>  
Las columnas presentes en el set de datos son:
  - Age of the patient: edad de los pacientes

- Gender of the patient: sexo del paciente ( 'Male' o 'Female')
- Total Bilirubin : La bilirrubina es una sustancia amarillenta que se forma durante el proceso normal de descomposición de los glóbulos rojos por el cuerpo. La bilirrubina se encuentra en la bilis, un líquido segregado por el hígado que ayuda a digerir los alimentos. Cuando el hígado funciona normalmente, elimina la mayoría de la bilirrubina del cuerpo. Si el hígado está dañado, la bilirrubina puede salirse del hígado y pasar a la sangre. Incluye la bilirrubina directa e indirecta.
- Direct Bilirubin: bilirrubina directa o conjugada: es la forma combinada de la bilirrubina y azúcares que es creada por el hígado y liberada en la bilis. De esta forma, está presente en bajas cantidades en la sangre y se altera cuando existe un problema a nivel del hígado;
- Alkaline Phosphatase: fosfatasa alcalina

- Alanine Aminotransferase
- Aspartate Aminotransferase
- Total Proteins
- Albumin
- Albumin and Globulin Ratio
- Dataset: 1 significa que es paciente y 2 que no.