**Técnicas de aprendizaje estadístico**

**2019-1**

**Trabajo (25%). Máximo 4 integrantes. Fecha entrega: 27 Agosto 2019**

Seleccione una base de datos abierta que contenga mínimo 2000 muestras y 10 atributos, de los cuales al menos mínimo 3 sean categóricos y 3 numéricos.

1. **Descripción del conjunto de datos**
2. Realice una descripción para contextualizar acerca del conjunto de datos tomado (fuente, área, dominio, tiempo en que se tomó, método de recolección, etc.)
3. Realice un análisis descriptivo básico de cada una de las variables
4. Indique cual es la variable independiente y de interés para predecir
5. **Modelos de predicción**:

Realice un modelo para predicción de la variable independiente a través de un Random Forest, un modelo con Redes Neuronales y uno son Maquinas de soporte vectorial. Para cada uno, varíe los parámetros propios del método con el fin de encontrar el modelo que mejor se adapta a los datos. Indique el código realizado de cada uno de los modelos, los parámetros que varió y el análisis que hizo para llegar a la elección del mejor modelo.

Utilice las librerías scikit-learn, tensorflow y keras de pyhton para realizar los modelos.

1. **Predicción**:

Realice una aplicación sencilla que al correrla entrene el modelo seleccionado, y una vez finalice dicha etapa permita al usuario ingresar por consola, o interfaz gráfica, una nueva muestra y conocer su valor predicho. La aplicación debe solicitar al usuario los valores para cada variable e indicar los posibles valores que puede ingresar.