

## Laboratorio 5

### Regresión Lineal Regularizada

**Instrucciones:** Usted usará Python a través de un Jupyter Notebook para llevar a cabo este laboratorio. Recuerde utilizar comentarios para describir lo que está haciendo en cada parte del proceso y así dejar constancia del entendimiento de su trabajo. Estará usando el juego de datos proporcionado dentro del portal. Al finalizar recuerde subir al portal un link a su repositorio en el que se pueda correr su notebook.

El juego de datos proporcionado es parte de la plataforma Kaggle, dentro del cual se muestran diferentes parámetros que suelen ser considerados por las universidades para admitir a los postulantes en los programas de postgrado (maestrías), además se muestra la probabilidad de admisión de estas personas.

1. GRE Scores ( out of 340 )
2. TOEFL Scores ( out of 120 )
3. University Rating ( out of 5 )
4. Statement of Purpose and Letter of Recommendation Strength ( out of 5 )
5. Undergraduate GPA ( out of 10 )
6. Research Experience ( either 0 or 1 )
7. Chance of Admit ( ranging from 0 to 1 )

Para este ejercicio se le pide que proporcione un modelo de regresión lineal multidimensional **regularizada** que prediga la admisión del alumno usando todas las características que considere necesarias. A continuación se mencionan las generalidades de los pasos sugeridos a realizar.

1. Leer el archivo CSV proporcionado (Admission\_Predict.csv, <https://www.kaggle.com/mohansacharya/graduate-admissions>) y almacenarlo en un **np.array** para ser trabajado en el notebook.
2. Ajustar un modelo lineal (polinómico) en base al juego de datos cargado que relacione cualquier subconjunto propio de las variables de los indicadores seleccionados con la probabilidad de admisión.
3. Utilice la implementación regularizada de la regresión lineal que mejor se ajuste a los datos.
4. Asegúrese de usar el *lambda* que mejor apoye al modelo que ha implementado para describir y predecir la información sobre la nube de datos.
5. Haga un análisis sobre sus hallazgos, donde mencione claramente las razones por las que considera que su modelo es bueno justificando adecuadamente así como el *lambda* usado y las evidencias correspondientes.