

Universidad Francisco Marroquín

Facultad de Ciencias Económicas

Ingeniería en Computer Science

Catedrático: Ing. Samuel Chávez

Contacto: samuelchavez@ufm.edu

Auxiliar: Paulo Mejía

Contacto: paulomejia@ufm.edu



Laboratorio 3

Regresión lineal

Instrucciones: Realice un programa en Python para resolver el siguiente problema.

Problema 1:

Anualmente en Estados Unidos, cerca de 1.5 millones de estudiantes de Educación Superior realizan un examen de aptitud escolar (SAT, por sus siglas en inglés). Aproximadamente el 80% de las universidades e instituciones de Educación Superior usan las calificaciones obtenidas por los estudiantes en este examen como criterio de admisión (College Board, marzo de 2006).

Un servicio de evaluaciones educativas ha recolectado la información del examen del SAT y la información del GPA de una muestra de 1000 estudiantes de una universidad (no se indica el nombre por temas de confidencialidad). El SAT consiste en múltiples secciones medibles, las cuales incluye: Matemática, lectura y escritura. Se prueba lectura y escritura juntas y Matemática se evalúa en una escala del 200 al 800. Una calificación perfecta es 1600 (800 en ambas secciones). En el caso de GPA es un término que se emplea para asignar un valor numérico a las puntuaciones acumuladas por un estudiante en el sistema estadounidense. Este valor puede ser anual o agruparse en períodos académicos y se calcula en una escala de 0 a 4 puntos (en algunos casos puede ser hasta 5).

Se tiene el siguiente juego de datos del archivo `satgpa.csv`, el cual tiene la siguiente información:

Variables:

- `sex` - sexo del estudiante (1=Mujer y 2=Hombre)
- `sat_v` - percentil SAT verbal
- `sat_m` - percentil SAT en Matemática
- `sat_sum` - total del percentil del SAT verbal y Matemática

- `hs_gpa` - promedio de calificaciones de la escuela secundaria
- `fy_gpa` - promedio de calificaciones del primer año de la universidad

Se tiene interés en determinar si el “sexo”, el “total del percentil de SAT verbal y Matemática” y el “promedio de calificaciones de la escuela secundaria” predicen en forma lineal el “promedio de calificaciones de primer año de la universidad”. Utilice el método del descenso del gradiente para encontrar los parámetros correspondientes (thetas). Estas características podrían utilizarse en un futuro para calificar a las personas para su ingreso a esta universidad.

Referencias:

- Diez, Cetinkaya-Rundel & Barr, OpenIntro Statistics, Four Edition. (OpenIntro, 2015).
- Anderson, Sweeney & Williams, Estadística para Negocios y Economía, 11a. ed. (Cengage Learning, 2011).
- <http://studentcaffe.com/preparate/estudiantes-de-la-secundaria/el-examen-sat?lang=es#:~:text=Se%20prueba%20lectura%20y%20escritura,en%20ambas%20de%20las%20secciones>).
- <https://www.academica.school/news/como-se-calcula-el-gpa/>