Primer Laboratorio de Introduction to Machine Learning

Universidad de Ingenieria y Tecnologia (UTEC) – Lima, Peru

August 2023 – Seccion Profesor Arturo Deza

Tema de Laboratorio : Linear Regression

5 points	What is your dataset ? *(Cual es tu dataset?)*
2 points	What is your regression problem? *(Cual es el problema de regresion que estas resolviendo?) [Ejemplo : Prediccion de Valor de Propiedades]* — Nota: No puede ser un dataset de prediccion de valor de propiedades inmobiliarias.
0.5 points	How many data points are there in the dataset? Nota : Minimo 100 data points; Maximo 10'000 data points.
1 points	What is the β term assuming a zero-th point intersection (no bias)
2 points	What is the bias term *not* assuming a zero-th point intersection (bias included)
0.5 points	What are the number of independent variables (minimum 1, maximum 2) [Nota: 2 puntos extra si hacen regresion de variable multiple con 2 variables], and what are they?
0.5 points	What is the dependent variable?
0.5 points	What are the unit of measurements for each variable(s)?
3 points	Please plot the raw data and a superimposing line on the data that passes through the origin. What is the mean square error (MSE)?
2 points	Please plot the raw data and a superimposing line on the data that does not pass through the origin. What is the mean square error (MSE)?
3 points	Repeat the regression line randomly 100 times. At each iteration randomly remove with replacement one data point and perform a regression. What are the average Beta values found for both cases where the fitting is done by fitting it through the origin (bias removed), and not fitting it through the origin (bias included). Hint : Look-up statistical bootstrapping.
Obligatory	Please list in your 2 page report :

(a) The contributions of each author.

(b) The list of all the python packages used.

(c) The list of all toolboxes used (and links to datasets and dataset license)

(d) The list of any AI tools (*e.g.* ChatGPT, Perplexity, You) used in your homework and how.

(e) The list of all academic references used in your homework.

(f) Attach a copy of all your code (this may extend 2 pages).