**Instituto Tecnológico de Monterrey Campus Querétaro**

****

**Programming Languages Project**

**August – December 2021**

**Linear Regression Prediction**

**Teacher: Benjamín Valdés Aguirre PhD**

**Ricardo Antonio Vázquez Rodríguez**

**A01209245**

# 

# Context of the problem

When there is uncertainty of the future, it is very common for humans to take decisions that will have an impact for them. The best way to make decisions is based on data and previous history of the past. The goal is to predict the grades of a group of students considering different models and features for prediction.

# Solution

**Concepts:**

* Cost function:
* Root mean squared error:
* Minimize cost function:
* Hypothesis:

The first step is to analyze the data, to see the pattern it tends to. If we can see a linear relationship between two or more variables, then we can use a linear prediction to best fit those elements and make predictions of the future.

A good approach for designing a linear regression function in based on the Gradient Descent Method.

Most of the elements from the real world include many features or variables. That is why this needs to be taken into consideration while.

The implementation (steps) is the following:

1. Pick a random value for the intercept (value will be optimized)

2. Evaluate the fitting of the line with the data using "Sum of the squared residuals" (is a Loss Function)

3. Express this equation with respect of the intercept, using the given y value minus the equation of the predicted one with the line

4. Keep adding the square of each of these residuals

5. Get the derivative of this equation (using the chain rule) to get the lowest

6. Least squares uses slope = 0, Gradient descent uses steps until reaching the best value.

7. Evaluate different values for the intercept, until getting closer to 0 (baby steps)

8. Size of the step is related to the slope

9. STEP SIZE = slope \* LEARNING RATE (0.01)

10. NEW INTERCEPT = Old Intercept - step Size

GD estimate: 0.95

STOPS when Step Size is very close to 0 (when the slope is close to cero). In practice 0.001 or smaller OR MAXIMUM NUMER OF STEPS 1000 or greater

# Results

# Conclusions

Linear regressions are

# Setup instructions

## Code Setup and run

1. Clone GitHub repository on the terminal.
2. Run the python file.
3. Insert the values for the independent variables.
4. View the predicted result.

# References

[0] <https://www.youtube.com/watch?v=sDv4f4s2SB8>

[1] <https://www.coursera.org/lecture/machine-learning/gradient-descent-for-multiple-variables-Z9DKX>

[2] https://www.youtube.com/watch?v=sDv4f4s2SB8

[3] https://towardsdatascience.com/linear-regression-using-gradient-descent-97a6c8700931

[4] <https://www.geeksforgeeks.org/gradient-descent-in-linear-regression/>

<https://www.khanacademy.org/math/statistics-probability/describing-relationships-quantitative-data/more-on-regression/v/calculating-r-squared>