



EDUCACIÓN
SECRETARÍA DE EDUCACIÓN PÚBLICA



TECNOLÓGICO
NACIONAL DE MÉXICO®

TECNOLÓGICO NACIONAL DE MÉXICO

INSTITUTO TECNOLÓGICO DE TIJUANA

SUBDIRECCIÓN ACADÉMICA

DEPARTAMENTO DE SISTEMAS Y COMPUTACIÓN

NOMBRE DE LOS ALUMNOS:

RAYMUNDO HIRALES LAZARENO (N. CONTROL: 17212339)

GALAVIZ LONA OSCAR EDUARDO (N.CONTROL: 17212993)

Carrera: Ingeniería Informática

Semestre: 9no

MATERIA: Minería de datos

PROFESOR: JOSE CHRISTIAN ROMERO HERNANDEZ

TRABAJOS: Practica 1

FECHA: 26/11/21

Visualización de datos de modelo de machine learning regresión lineal:

Usando una fuente de datos de los salarios y la experiencia de los trabajadores con los cuales visualizamos los modelos básicos de regresión lineal.

Código:

```
# importar el repositorio
dataset <- read.csv('Salary_Data.csv')

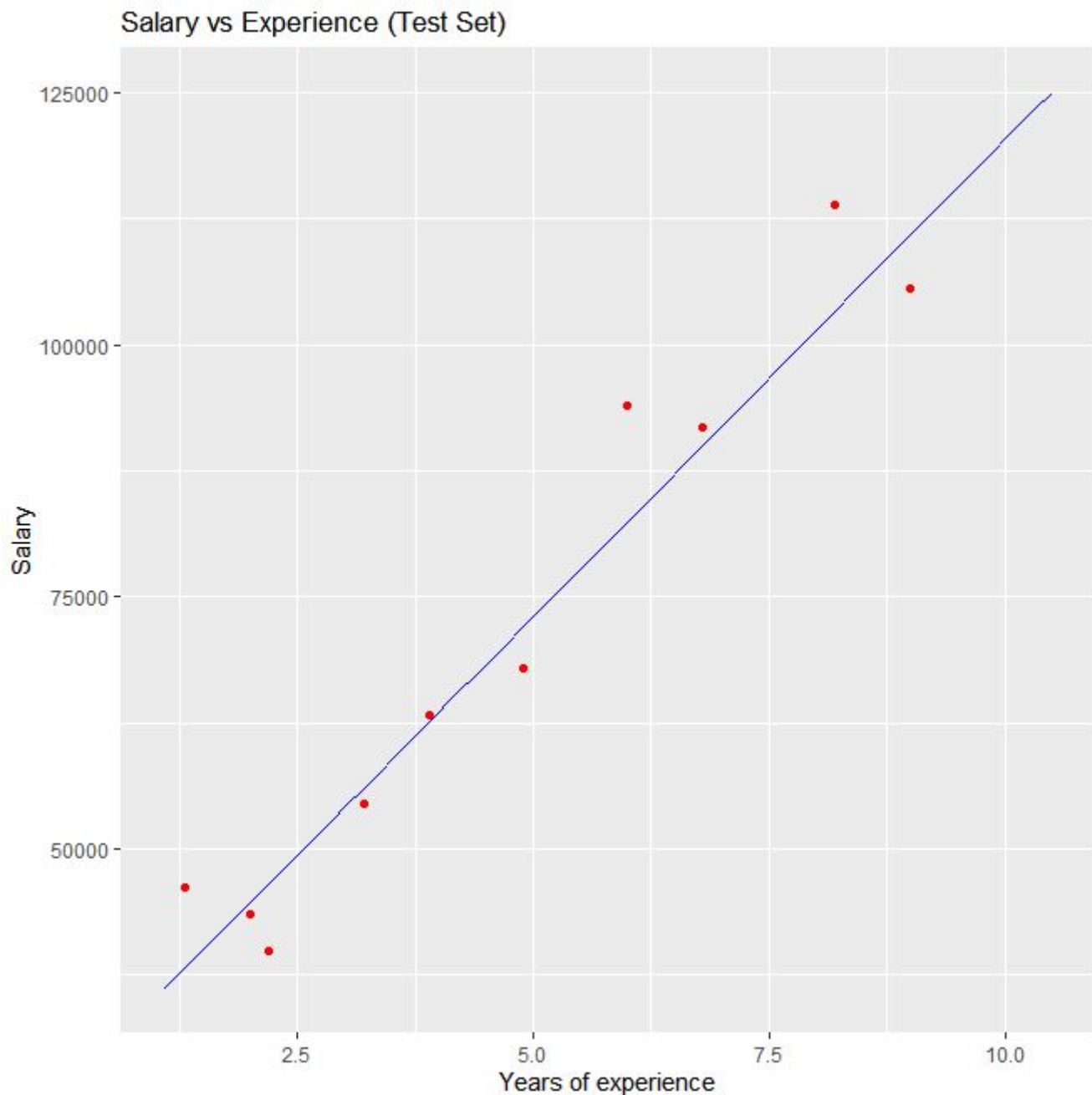
# separar el datasets dentro del training set y el test set
# instalar.packages('caTools')
library(caTools)
set.seed(123)
split <- sample.split(dataset$Salary, SplitRatio = 2/3)
training_set <- subset(dataset, split == TRUE)
test_set <- subset(dataset, split == FALSE)

# acomodando la regresion linear simple al Training set
regressor = lm(formula = Salary ~ YearsExperience,
               data = dataset)
summary(regressor)

# Prediccion de reusltados del Test set
y_pred = predict(regressor, newdata = test_set)

# Visualizando los resultados del Training set
library(ggplot2)
ggplot() +
  geom_point(aes(x=training_set$YearsExperience, y=training_set$Salary),
            color = 'red') +
  geom_line(aes(x = training_set$YearsExperience, y = predict(regressor, newdata =
training_set)),
            color = 'blue') +
  ggtitle('Salary vs Experience (Training Set)') +
  xlab('Years of experience') +
  ylab('Salary')
```

En esta imagen podemos denotar la linea de progresion lineal del training set



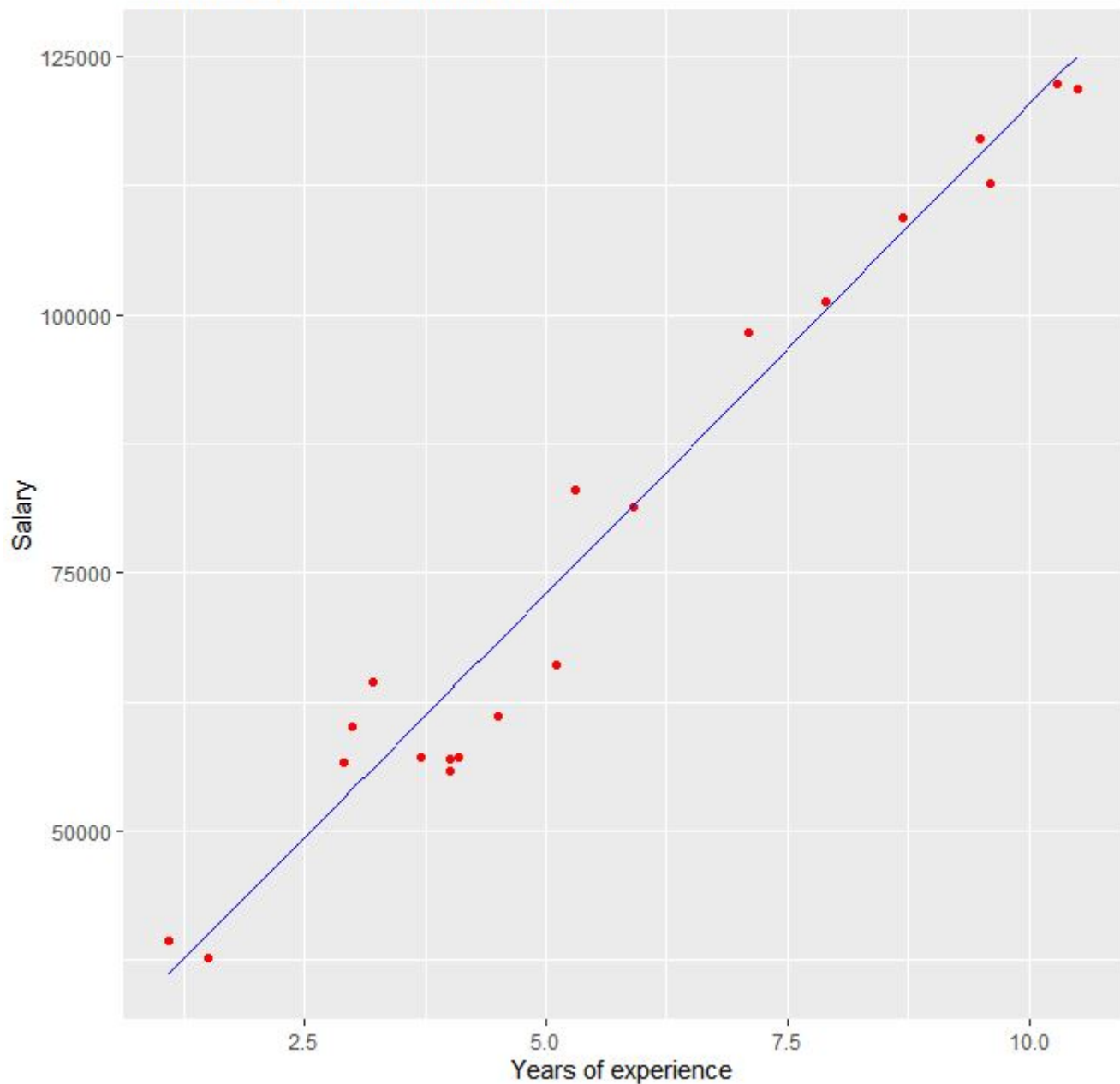
```
# Visualizando los resultados del Test set
ggplot() +
  geom_point(aes(x=test_set$YearsExperience, y=test_set$Salary),
             color = 'red') +
  geom_line(aes(x = training_set$YearsExperience, y = predict(regressor, newdata =
training_set)),
            color = 'blue') +
  ggtitle('Salary vs Experience (Test Set)') +
  xlab('Years of experience') +
  ylab('Salary')
```

R Graphics: Device 2 (ACTIVE)

— □ ×

Archivo Histórico Redimensionar

Salary vs Experience (Training Set)



Analisis y conclusion

Se puede denotar la regresion lineal en el ejercicio a partir de las imagenes presentadas lo que demuestra el uso de este tema para el analisis de mineria de datos.