**Introducción**

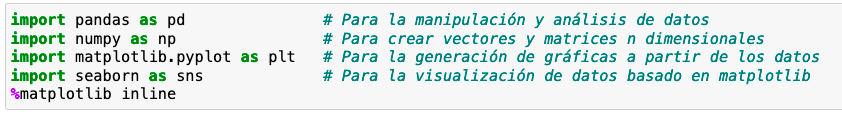
En esta práctica se generarán pronósticos utilizando regresión lineal múltiple para registros geofísicos sobre la saturación de aceite remanente. Utilizando la regresión lineal múltiple se generará un modelo que pueda predecir RC4 a partir de la profundidad, RC1, RC2 y RC3.

**Objetivos**

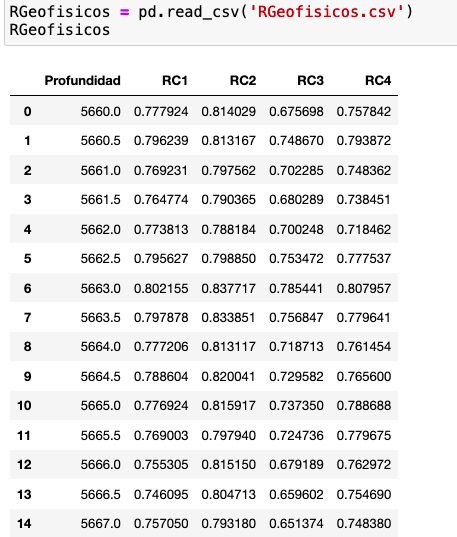
Se tienen mediciones de registros geofísicos convencionales: RC1 (Registro Neutrón), RC2 (Registro Sónico), RC3 (Registro Densidad-Neutrón) y RC4 (Registro Densidad -corregido por arcilla-). Se desea obtener el pronóstico de la saturación de aceite remanente (ROS, Residual Oil Saturation).

**Desarrollo**

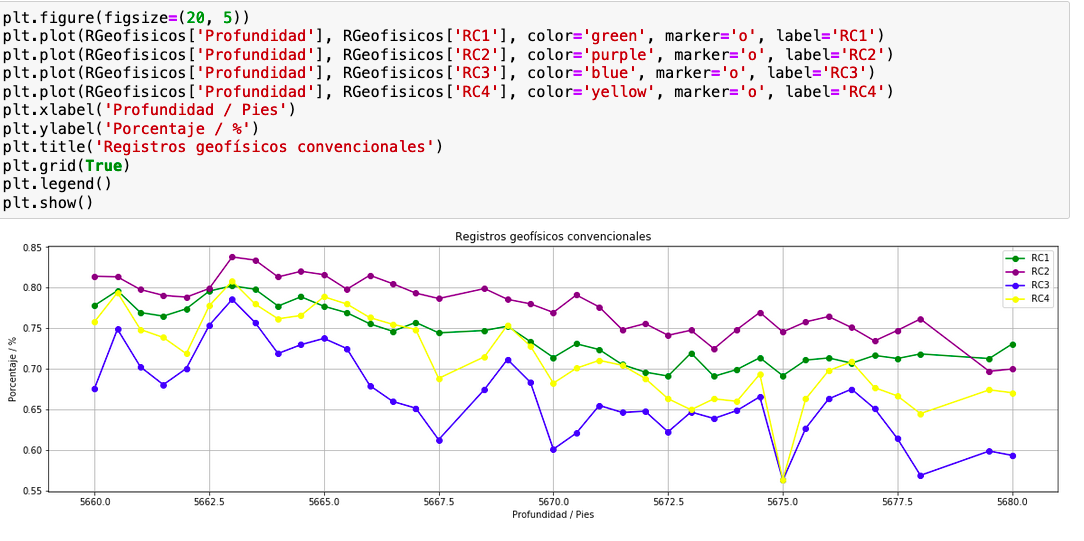
Se importaron todas las bibliotecas relevantes para la realización de la práctica.



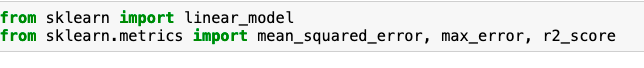
Se muestran los datos donde se puede observar cada profundidad y los diferentes registros que se tienen de un pozo.



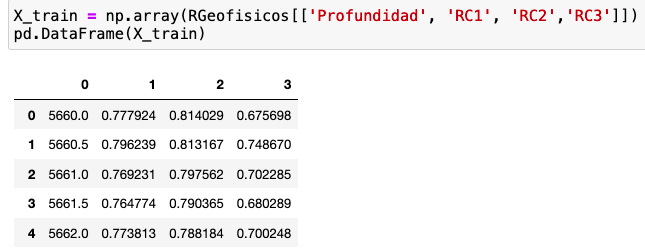
Utilizando matplotlib se generó una gráfica de la profundidad contra cada uno de los registros. Se puede observar una tendencia general de los diferentes registros, aunque no existe una tendencia muy clara.

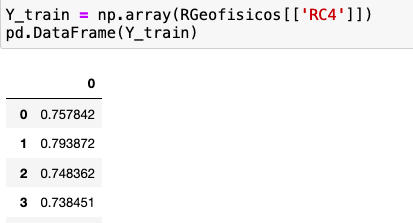


Para generar la regresión lineal múltiple se utilizó el módulo de linear\_model de la biblioteca de sklearn. Se importaron otras métricas que se utilizarán para analizar los datos.

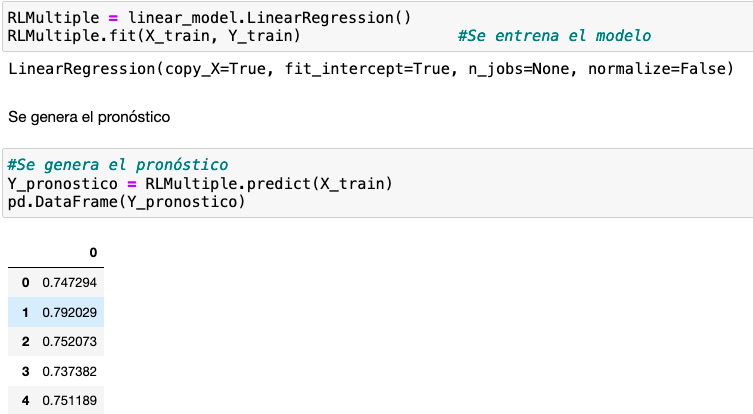


Se generan dos DataFrames con los datos de entrenamiento y se separan entre variables predictoras (x) y la variable a pronosticar. Para el caso de esta práctica se utilizará la profundidad, RC1, RC2 y RC3 como variables predictoras y RC4 como la variable a pronosticar.

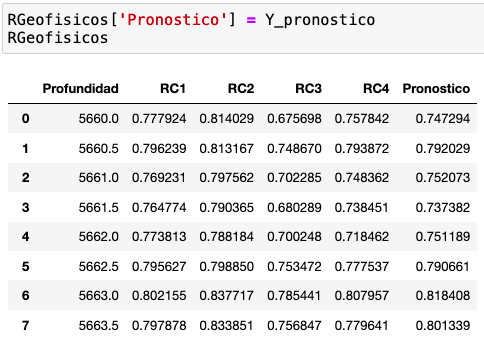




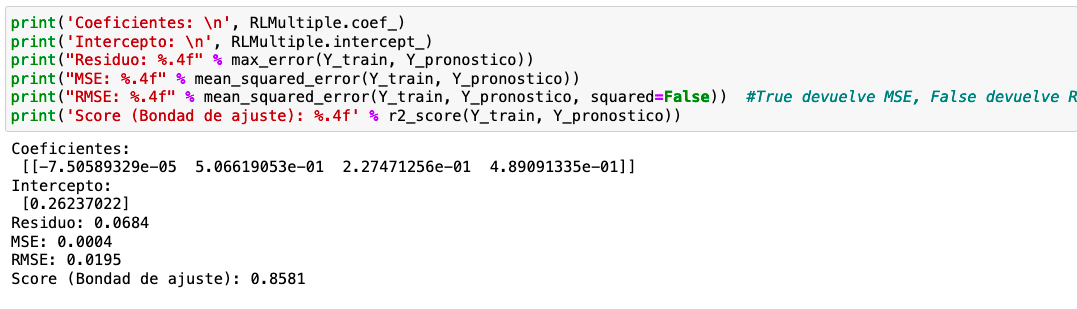
Utilizando la regresión lineal múltiple se entrena el modelo utilizando los datos y se genera el pronóstico a partir del modelo.



Se le adjunta a la tabla original la variable de Pronostico.



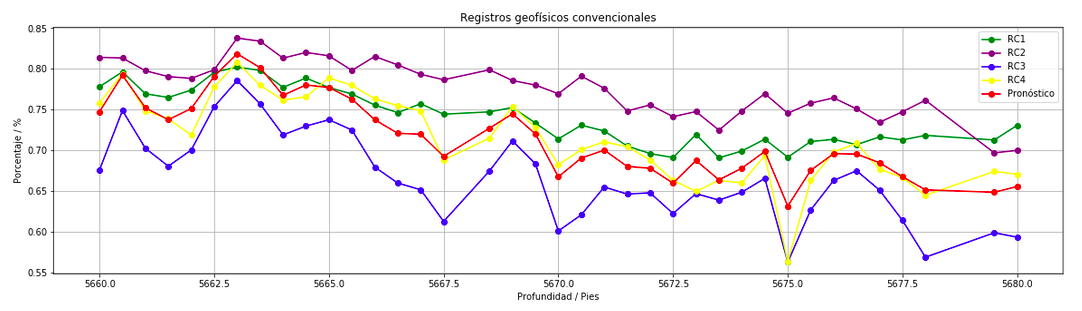
Utilizando algunos de los módulos que se importaron al principio se realiza el cálculo de los coeficientes, el intercepto, residuo, MSE, RMSE y la bondad de ajuste.



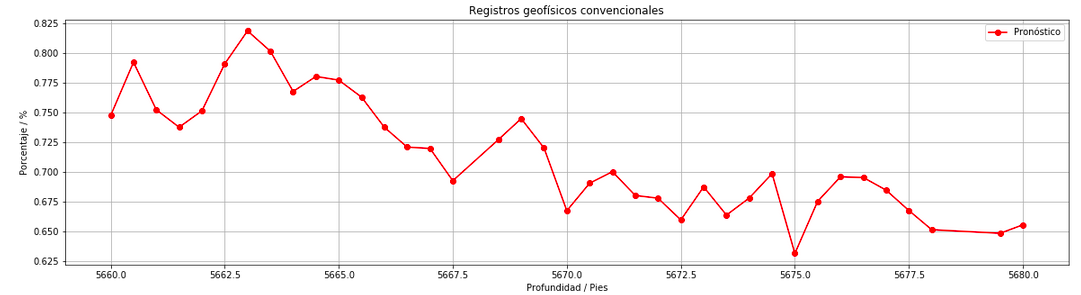
A partir de los datos anteriores se puede escribir el modelo como:

Con la bondad de ajuste se habla de cómo la predicción que se está generando tiene un 85% grado de efectividad. El MSE y RMSE indican cómo los pronósticos finales del modelo se alejan .0004 y .0195 unidades del valor real.

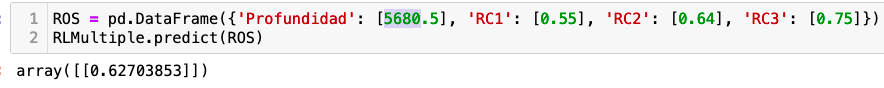
Se genera una nueva gráfica para visualizar los valores pronosticados y qué tan cercanos al valor real se encuentran. El pronóstico es la gráfica roja y RC4 es la amarilla por lo que se puede observar como sí se logró una gráfica cercana a la real excepto en pequeñas partes.



También se generó una gráfica donde solo se encuentran las predicciones para poder visualizar estos datos.



A partir del modelo generado también se pueden generar nuevas predicciones. Para ello se utiliza el método predict y se le debe alimentar al modelo con los cuatro valores que necesita para generar una predicción. Se muestra una predicción de RC4 cuando la profundidad es 5680.5, RC1 es 0.55, RC2 es 0.64 y RC3 es 0.75.



**Conclusiones**

En esta práctica se cumplieron los objetivos y se logró observar una aplicación de la regresión lineal múltiple. Además, se logró implementarla en una Jupyter Notebook utilizando diferentes módulos para obtener un modelo. Para dicho modelo se calculan datos como los coeficientes de la ecuación, su intercepto, su residuo, bondad de ajuste, MSE y RMSE. Cada uno de estos datos se interpretó en el caso específico de los registros de la saturación de aceite remanente en un pozo. Se pudo observar cómo el modelo generado tiene muy buena bondad de ajuste y esto permite tener buenas predicciones a partir del modelo. Con esta práctica se logró generar una regresión múltiple que ayuda a predecir el RC4 a partir de registros geofísicos.