



Universidad  
Industrial de  
Santander

# Predicción de compra por cliente a partir de cotizaciones de una empresa

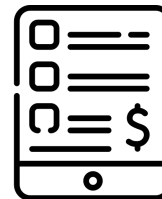
---

Miguel Ángel Plata Rodríguez  
Adriana Marcela Cujia Reyes  
Sebastián Suarez

# Dataset

100 Cotizaciones asignadas a mano, con las siguientes características:

- **N\_cot:** Número de cotización que solía repetirse de vez en cuando. (*int*)
- **Cliente:** Nombre del cliente o la empresa. (*String*)
- **Fecha** (*formato date*)
- **N\_Producto:** Número de productos adquiridos. (*int*)
- **Capacitación:** Si cobraba capacitaciones de uso o no. (*bool*)
- **V\_Total:** El valor total de cobro y nuestro valor a predecir. (*float*)
- **Descuento** (*bool*)
- **document:** Nombre del documento PDF. (*String*)



	N_Cot	Cliente	Fecha	N_Producto	Capacitación	V_Total	Descuento	document
2	751	Eduardo Ra	3/5/2016	2	FALSE	273.760	FALSE	Cotización 751 -Eduardo Ramirez- (PLA) 2.85mm
3	752	DAR Soluci	3/5/2016	2	FALSE	6.641.000	FALSE	Cotización 752 - DAR Soluciones (PLA-ABS)
4	753	Paola Buitri	4/5/2016	3	TRUE	3.089.520	FALSE	Cotización 753 - Paola Buitrago - Micro + insumos
5	754	BPM Globa	4/5/2016	6	TRUE	24.831.784	TRUE	Cotización 754 -BPM Global Solutions- Makerbot Replicator 2 y 2X
6	755	Diego Salaz	5/5/2016	2	TRUE	21.451.000	TRUE	Cotización 755 -Diego Salazar- Scanners 3D (Fuel 3D + David2)
7	756	UNIVERSID	5/5/2016	3	FALSE	888.000	FALSE	Cotización 756 - UNIVERSIDAD DE LOS ANDES- ABS EasyFil Formfutura, ABS IC3D
8	756	Daniela gor	5/5/2016	2	TRUE	6.042.080	TRUE	Cotización 756 -Daniela Gomez- Software FUEL 3D
9	757	ASSA ABL'O	4/4/2017	6	FALSE	1.281.327	FALSE	Cotización 757 -YALE- Repuestos MakerGear M2
10	758	UNIVERSID	5/5/2016	2	FALSE	640.400	FALSE	Cotización 758 - Uniandes (Buildtak)
11	759	Colegio Col	10/5/2016	11	TRUE	61.210.890	FALSE	Cotización 759 - Colegio Colombo Británico - Portfolio impresoras
12	760	Iproyect Ca	10/5/2016	2	TRUE	5.739.976	TRUE	Cotización 760 -Iproyect - FUEL 3D + Software
13	761	Jhon Fredy	10/5/2016	1	TRUE	1.899.000	FALSE	Cotización 761 - Jhon Fredy Correa Wi3 V2
14	762	Nubia Godc	10/5/2016	2	TRUE	1.944.650	TRUE	Cotización 762 - Nubia Godoy Wi3 V2
15	764	Leonel José	5/5/2016	3	FALSE	159.500	FALSE	Cotización 764 - Leonel Paredes - PLA + Grasa + Lubricante - copia-1
16	765	Pontificia U	16/5/2016	1	TRUE	2.896.000	FALSE	Cotización 765 - Javeriana Bogotá - MICRO 3D + 3Doodler
17	766	RTR S.A.S.	17/5/2016	6	TRUE	30.873.560	FALSE	Cotización 766 - RTR - U2+, M2, Dreamer
18	767	Felipe Fran	17/5/2016	2	TRUE	1.944.650	TRUE	Cotización 767 - Felipe Franco Wi3 V2
19	768	Jose Fernar	17/5/2016	8	FALSE	6.296.240	FALSE	Cotización 768 - Tecnoparque CALI u2+ Upgrade kit + PLA
20	769	Tecnologia	18/5/2016	2	TRUE	1.944.650	TRUE	Cotización 769 - Tecnologia Ambiental NOW Wi3 V2
21	770	Tecnologia	18/5/2016	2	TRUE	5.958.400	TRUE	Cotización 770 - Tecnologia Ambiental NOW- Liquid Crystal 10" + resina
22	771	RTR S.A.S.	18/5/2016	3	TRUE	10.786.794	TRUE	Cotización 771 - RTR D55
23	772	Freddy Parr	23/5/2016	10	TRUE	17.439.873	TRUE	Cotización 772 - Freddy Parra - U2+
24	773	Juan Carlos	23/5/2016	1	TRUE	5.196.580	TRUE	Cotización 773 - Juan Carlos Fonseca - DAVID Starter Kit
25	774	Sandra Per	20/5/2016	4	TRUE	19.299.120	TRUE	Cotización 774 - Sandra Perez - D5S Mini + Dreamer + D6
26	775	Universidac	20/5/2016	6	TRUE	30.874.160	TRUE	Cotización 775 - U. Manuela Beltrán - U2+, M2, DReamer
27	778	TEK SOLUCI	20/5/2016	5	FALSE	13.357.000	TRUE	Cotización 778 - TEK Soluciones - 3Doodler
28	779	Universidac	20/5/2016	1	FALSE	63.800	FALSE	Cotización 779 - USTA - Bloque extrusor M2

[2, 10, 50, 100, 500]

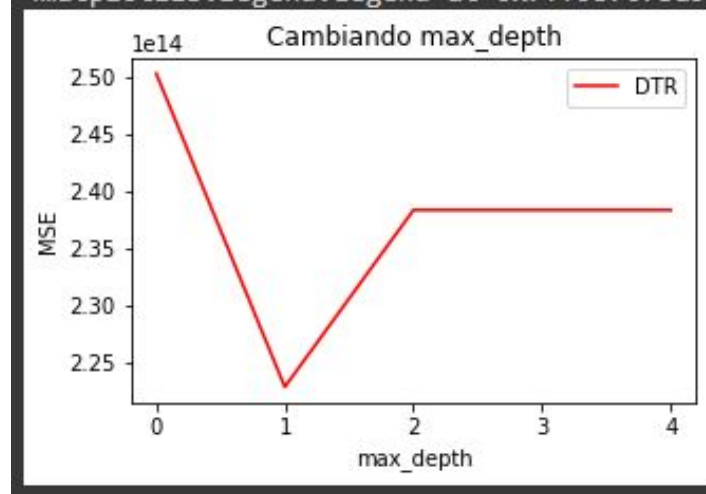
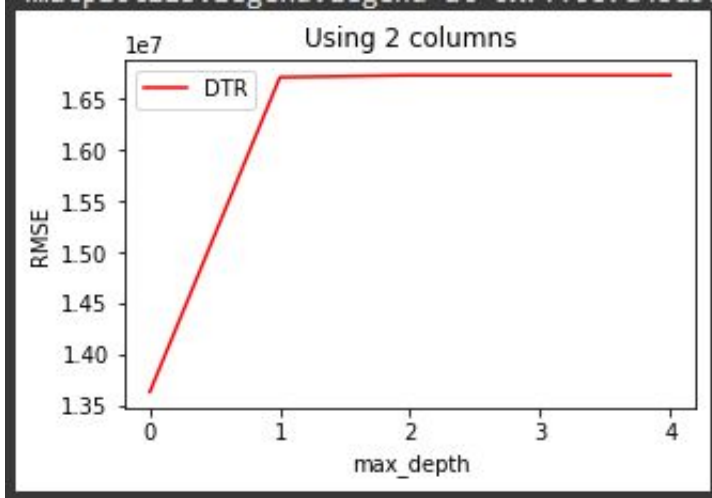
# Decision Tree Regressor

- Se entrenó con un con un 80% y se usó 20% en el test.
- Se imprime el `mean_squared_error`
- Se aplicó CrossOver
- Hubo tuning en el CrossOver y en el maxDepth.



accuracy 246884390980265.406 (+/- 603408088553196.87500)

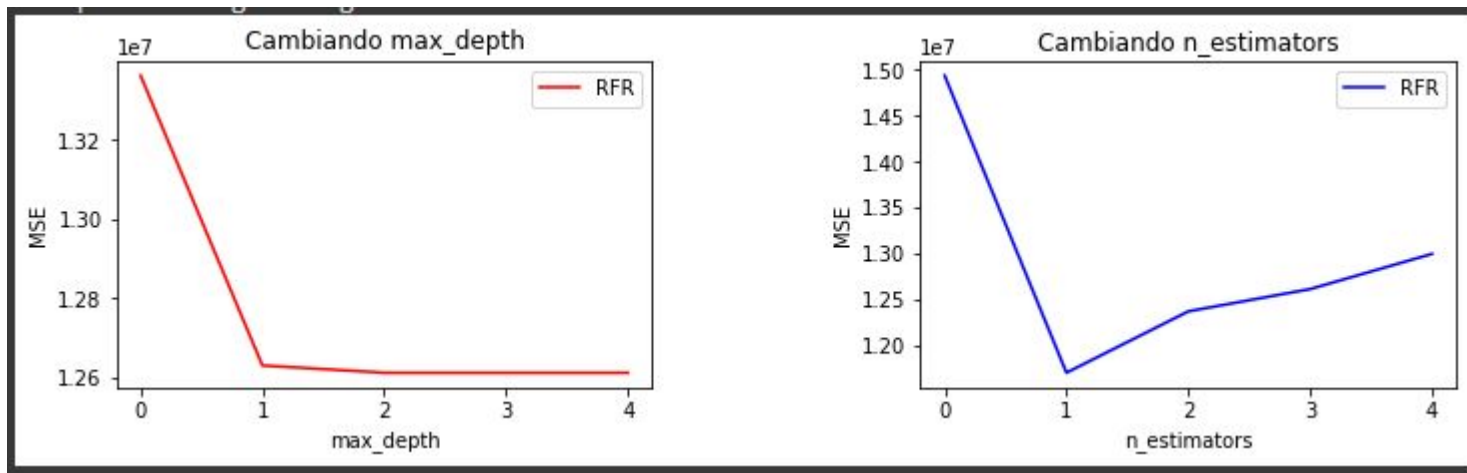
282967546354982.6



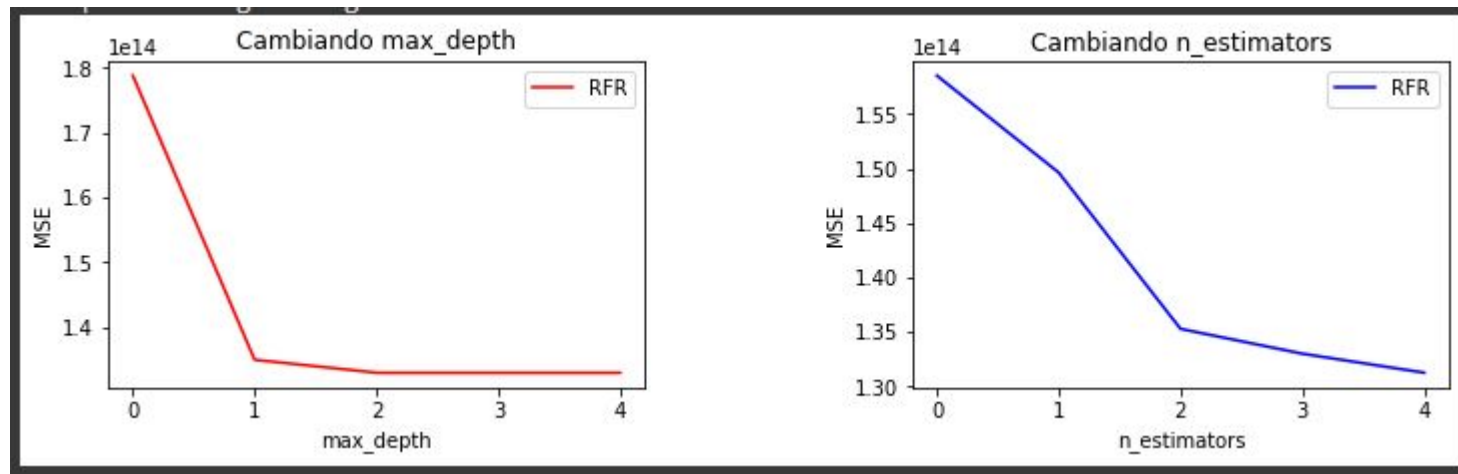
[2, 10, 50, 100, 500]

# Random Forest Regressor

- Se entrenó con un con un 80% y se usó 20% en el test.
- Se imprime el `mean_squared_error` 159082113784260.12
- Se aplicó CrossOver `accuracy 134357605282148.203 (+/- 238608053981478.81250)`
- Hubo tuning en el CrossOver, en el maxDepth y en el nstimators

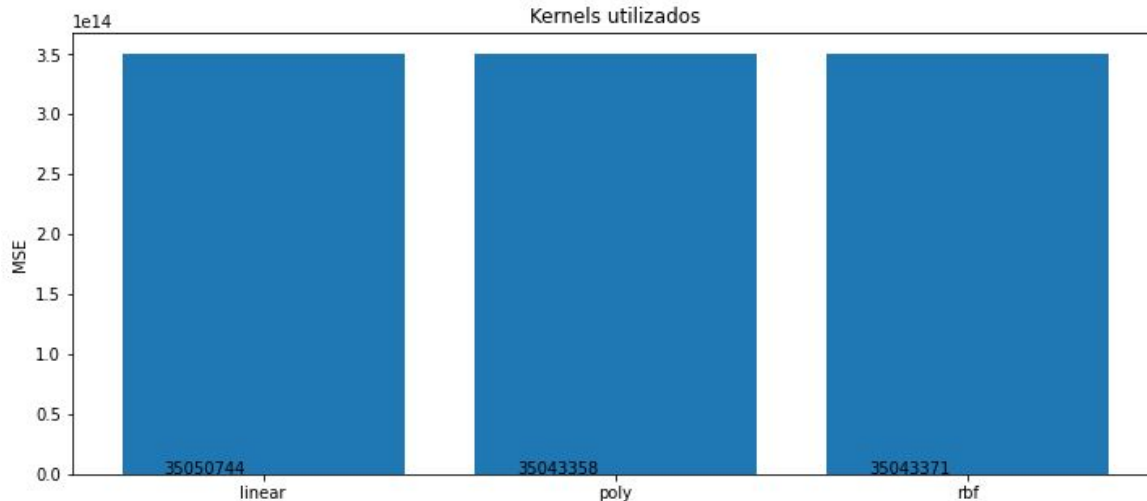


# Tunning Random Forest Regressor



# Support Vector Regression (SVR)

- Se entrenó con un con un 90% y se usó 10% para el test.
- Se usó un ciclo for para iterar en las principales Kernels: Lineal, Polinómico y RBF.
- Se obtuvieron los siguientes % de errores:



# Conclusiones

- Nuestro dataset no era apto para uso de Clasificación.
- Faltó un mayor tamaño de filas del dataset para aumentar las posibilidades de repetición de cliente.
- Si hubieran más clientes repetidos, el entrenamiento hubiera sido mejor y por lo tanto, daría mejores resultados.



Valores únicos repetidos en un porcentaje

