

Product Profitability Predictions with R

Presentación: Luis Sánchez Peña
sanchezisp@gmail.com

Fecha: 29 de Agosto de 2017

Product Profitability Prediction with R

introduction

- +Blackwell Electronics'CTO Danielle Sherman quiere que preveamos la productividad de 17 candidatos a productos nuevos. De entre los candidatos elegiremos a 5 para ser vendidos en las tiendas.
- + Debemos realmente comparar el uso y resultados de WEKA frente a los de R. Los cálculos y conclusiones con WEKA se hicieron en el pasado (tarea anterior).
- + Debo aconsejar o no la incorporación de R y en su caso los pasos a dar.

Product Profitability Prediction with R

introduction

Deberé usar el software R para realizar este trabajo.
Usaré tres tipos de análisis para el cálculo:

- Linear Model
- Support Vector Machine (SVM)
- Random Forest

Debo reportar mi experiencia con el software R así como comparativa con los resultados de la tarea anterior.

Product Profitability Prediction with R

R experience

Pros

- ✓ Es un lenguaje de programación específico para estadística.
- ✓ Es Open Source (código libre y gratuito).

Documentación amplia.

No black-box (código disponible).

Ampliable/adaptable sin límites.

- ✓ Es muy popular.
- ✓ Es reproducible (= fiable).
- ✓ Nuestros resultados tendrían credibilidad ante terceros.
- ✓ Usada ampliamente por empresas de reconocido prestigio.
- ✓ Independiente de plataforma/S.O. (Linux, Mac, Windows).

Product Profitability Prediction with R

R experience

Contras

- × Tiene un curva de aprendizaje complicada al principio.
- × Algunos módulos son de pago.

Product Profitability Prediction with R

R experience

Nuestra experiencia

- + Instalación sencilla y rápida.
 - + Desaparece el efecto “caja negra” --> controlamos mejor nuestros análisis y datos.
 - + Permite (teóricamente) el trabajo en equipo (Ej.: Github).
 - Avisos de errores poco limpios.
 - TDD (debugging) poco claro.
 - Tutorial actual aceptable. No hay uno sencillo por culpa de R.
- Elegir en función de la experiencia y conocimientos previos del lector. Ofrecer opciones.

Product Profitability Prediction with R

R experience

Did you learn for business?

- Práctica densa. Debíamos aprender:

1. de errores;
2. R;
3. trabajo en equipo;
4. SVM, LM, RF y sus métricas;
5. limpiar datos...

.....en una sola práctica me resultó frustrante.

+ He aprendido github y me inicié en R

- De Data Analyst methods poco.

+ Limpiar, limpiar y limpiar los datos de entrada.

Product Profitability Prediction with R

R experience

Recommendations to get started

- Aprender antes del cambio algo del entorno y de programación básica (funcional y POO).
- Comenzar con datos/algoritmos conocidos y dominados.
- Comenzar con datos sin errores. Luego subir a datos con fallos.
- Comenzar con versiones probadas y estables (tanto R como algoritmos).
- El trabajo con los compañeros es clave (comunicación equipo).

Product Profitability Prediction with R

Analysis LM

¡Que hable el código!

Source: <https://github.com/meetsCode/Mod2T2Luis>
rama (branch): unificado
(<https://github.com/meetsCode/Mod2T2Luis.git>)

Product Profitability Prediction with R

Analysis LM

	ProductType	Product	ProductHeight	ProfitMargin	Volume	VolPredictions	beneficios	Profit SMOreg	Rank SMOref	profit similarity	Pos Similarity
11	Tablet	187	0,4	0,2 ?		11213,4560805264	446295	150125958,2	1	199	8
5	Laptop	176	0,88	0,23 ?		644,914162831454	296512	1839999,54	11	1999	2
1	PC	171	8,39	0,25 ?		942,136148696658	164638	67103825,25	3	699	13
2	PC	172	9,13	0,2 ?		631,205460598243	108567	35087484	6	860	7
4	Laptop	175	1,4	0,15 ?		270,1935970886	48594	5035979,85	8	1199	6
10	Tablet	186	0,37	0,1 ?		683,941978678657	43019	74473537,1	2	629	5
3	Laptop	173	0,68	0,1 ?		346,657615841799	41564	35490,4	13	1199	4
7	Netbook	180	0,95	0,09 ?		1310,87665057918	38815	36953398,44	5	329	12
16	GameConsole	199	13,2	0,09 ?		1713,86538698556	38560	41578314,3009	4	249,99	3
15	Smartphone	196	0,4	0,11 ?		647,494398508707	21367	6600033	7	300	1
12	Smartphone	193	0,4	0,11 ?		630,161012406386	13794	8668,44	14	199	16
13	Smartphone	194	0,37	0,12 ?		1924,42465074809	11315	2351994,12	10	49	9
8	Netbook	181	1,5	0,11 ?		79,5113602065567	3839	4442,68	15	439	10
14	Smartphone	195	0,4	0,15 ?		167,964129028318	3753	3754,8	16	149	14
17	Display	201	7,3	0,05 ?		402,570047965009	2817	111986	12	140	15
6	Netbook	178	1,2	0,08 ?		51,7332777646225	1655	2431907,2008	9	399,99	11
9	Netbook	183	0,97	0,09 ?		-131,9909171165	-3920	356,4	17	330	17

1	PC	171	0,25 ?	942,136148696658	164638	67103825,25	3	699	13	
2	PC	172	0,2 ?	631,205460598243	108567	35087484	6	860	7	
3	Laptop	173	0,1 ?	346,657615841799	41564	35490,4	13	1199	4	
4	Laptop	175	0,15 ?	270,1935970886	48594	5035979,85	8	1199	6	
5	Laptop	176	0,23 ?	644,914162831454	296512	1839999,54	11	1999	2	
6	Netbook	178	0,08 ?	51,7332777646225	1655	2431907,2008	9	399,99	11	
7	Netbook	180	0,09 ?	1310,87665057918	38815	36953398,44	5	329	12	
8	Netbook	181	0,11 ?	79,5113602065567	3839	4442,68	15	439	10	
9	Netbook	183	0,09 ?	-131,9909171165	-3920	356,4	17	330	17	
10	Tablet	186	0,1 ?	683,941978678657	43019	74473537,1	2	629	5	
11	Tablet	187	0,2 ?	11213,4560805264	446295	150125958,2	1	199	8	
12	Smartphone	193	0,11 ?	630,161012406386	13794	8668,44	14	199	16	
13	Smartphone	194	0,12 ?	1924,42465074809	11315	2351994,12	10	49	9	
14	Smartphone	195	0,15 ?	167,964129028318	3753	3754,8	16	149	14	
15	Smartphone	196	0,11 ?	647,494398508707	21367	6600033	7	300	1	
16	GameConsole	199	0,09 ?	1713,86538698556	38560	41578314,3009	4	249,99	3	
17	Display	201	0,05 ?	402,570047965009	2817	111986	12	140	15	
	ProductType	Product	ProfitMargin	Volume	VolPredictions	beneficios	Profit SMOreg	Rank SMOref	profit similarity	Pos Similarity