

Minería de Datos

Reglas de asociación



ZOO

Alberto Castillo Lamas

alcasla90@gmail.com

Este artículo contiene un análisis del dataset Zoo. Constituido por 17 características más una de clase, con un total de 101 instancias representando cada una las características individuales de un animal.

Voy a trabajar siempre con una generación de reglas de máxima longitud 6.

1 Inicio y análisis general

El dataset no viene en forma de transacciones, por tanto antes de transformarlo voy a proceder a preprocesarlo para trabajar más cómodo.

Voy a eliminar la característica *name*, la cual no nos proporciona información útil, transformo las características lógicas asignándoles cadenas de caracteres. Finalmente voy a generar una nueva característica a partir de *legs* para distinguir los animales con y sin patas, así me será útil para el estudio de esta información en conjunto (introducir el número de piernas nos complicaría la comprensión de las reglas puesto que puede tomar hasta 6 valores distintos además de dejarnos un soporte muy dividido).

Una vez preprocesado lo transformo en transacciones y veo la frecuencia de cada ítem como muestra la figura 1. Donde la línea horizontal situada en el 0.75 nos indica los ítems que se encuentran presentes al menos en 3 de cada 4 muestras, muy presentes, estos son animales sin plumaje (no feather), no aéreos (no airborne), con columna vertebral (backbone), respiran (breathe), no venenosos (no venomous), sin alas (no fin), con piernas (legs), y no domésticos (no domestic). También se repite bastante con cola (tail), pero por el momento queda en el aire su estudio.

Estos nos indican la existencia de subconjuntos opuestos (variables lógicas) muy particulares. Estos subgrupos serán los empezaré estudiando.

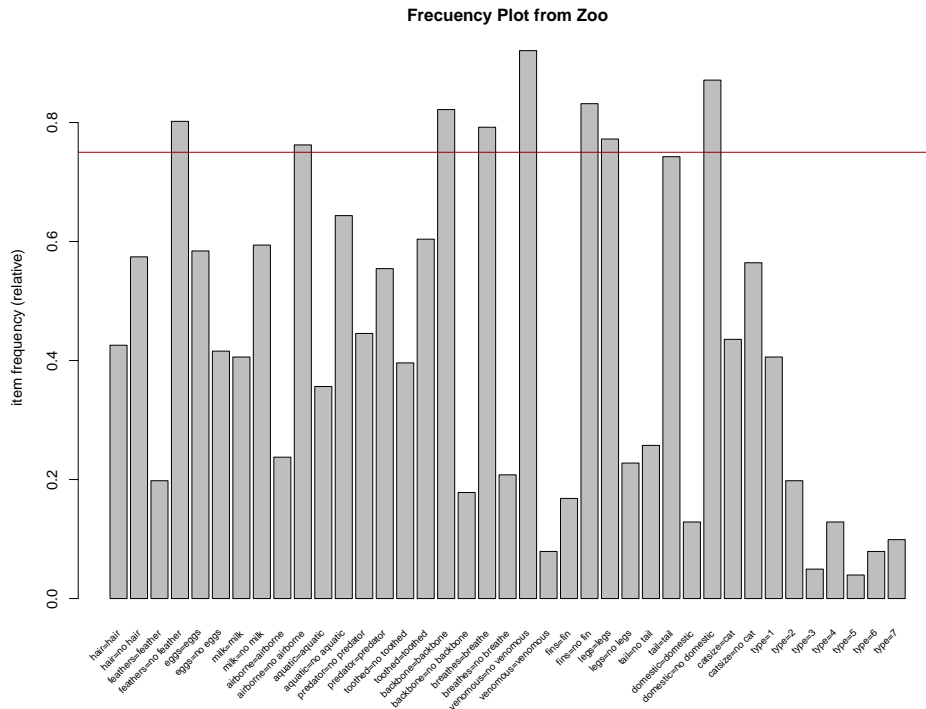


Figure 1: Plot frecuencia de todos los ítems del dataset Zoo. Delimitador horizontal situado en el 3º cuartil (0.75)

Legs (piernas) Como se ha visto por los ítems más frecuentes nos encontramos ante una mayoría de animales con unas características posiblemente comunes que nos indica mamíferos

vertebrados terrestres, con piernas y cola, además salvajes.

Una característica, como poseer piernas fue procesada y su transformación nos proporciona un grupo grande de animales con piernas, suma del resto de grupos con distintos números de piernas, dejando en su contra uno de animales sin piernas entorno al 20% de la población. Obtengo las reglas donde *no tener piernas* aparezca en el consecuente y elimino las reglas redundantes, lo que me deja un conjunto de 120 reglas que ordeno por *índice de certeza*, obteniendo reglas con dependencia completa completa (índice de certeza, CF=1) que voy a analizar.

El conjunto de reglas obtenido tiene en su top por fiabilidad reglas que determinan animales sin patas mediante unno o dos ítems en el antecedente, se amplían conforme pierden fiabilidad. Las reglas más interesantes son:

- Si es de tipo 4 no tienen patas
- Si no respira y tiene aleta, entonces no tiene patas
- Si no tiene pelo y tiene aleta, entonces no tiene patas
- Si pone huevos y tiene aleta, entonces no tiene patas
- Si no da leche y tiene aleta, entonces no tiene patas
- Si es dentado y no respira, entonces no tiene patas
- Si no respira y tiene cola, entonces no tiene patas
- Si tiene columna vertebral y no respira, entonces no tiene patas
- Si tiene aleta, entonces no tiene patas

Entre las reglas más destacadas he marcado con **rojo** las reglas que no me aportan información o más bien que la información que aportan es trivial. Todas estas reglas se relacionan con animales sin patas, y en estos casos claramente señalan a animales acuáticos como peces, puesto que no respiran. Las reglas solo nombran características de este tipo de animales, como que *no respiran*, tienen *cola*, *aleta*, *columna vertebral*, o ponen *huevos*.

Esto quiere decir que desde una regla podemos ir a la otra, es del tipo 4 y no tiene patas obtiene el mismo soporte que si insertando que tenga aleta, así nos indica que todo los animales del *tipo 4 no tienen patas y tienen aleta*. Buscar qué determina ser de *tipo 4*, y claramente muestra con un dependencia total según si índice de certeza que este tipo de animales tienen aleta, no respiran, no tienen piernas, son acuáticos, no tienen pelo, ponen huevos, tienen dientes, columna vertebral y cola, pues nos aclara indiscutiblemente que son peces. Las reglas que determinan estos datos son reglas de un antecedente y un consecuente con ser de tipo 4 en el antecedente, al tener todas una confianza de 1 las características comentadas se cumplen en todos los casos de *tipo 4* pero llama la atención dos reglas más (tabla 1) que muestran unas anomalías dentro del grupo de peces donde por la bajada del soporte indica que no todos son salvajes y no venenosos.

lhs	rhs	support	confidence	lift	certainty
		0.1287	1.0		
Tipo 4 – >	no domestico	0.1188	0.9230	1.0594	0.4023
Tipo 4 – >	no venenoso	0.1188	0.9230	1.0024	0.0288

Table 1: Reglas con medidas a observar.

Por la poca ocurrencia de las anomalías detectadas por las dos reglas de la tabla 1 hay que reducir mucho los requisitos mínimos para obtener las reglas y generar casi todas, así si puedo obtener reglas para animales de *tipo 4 venenosos*, hay uno entre el grupo de peces, caracterizado también por su tamaño y no ser doméstico, reglas con confianza uno, lo que me indica que solo este es venenoso y al ser no doméstico no representa la anomalía detectada por la otra regla, por tanta ese es otro individuo perteneciente a este grupo de peces pero de carácter doméstico.

En resumen, se comenzó estudiando animales sin piernas, estudio que desembocó en un subgrupo mayoritario, los de *tipo 4* que son peces, en su mayoría, *no domésticos* y entre los que se encuentran dos individuos distintos, uno venenoso, y otro doméstico. Seguramente este zoo cuente con un estanque de peces con distintas especies seguramente raras y quizá algo agresivas o no típicas de pequeños peces de pecera, con un caso de pez venenoso y uno doméstico que podría ser una carpa con pocas esperanzas de sobrevivir si se encuentra compartiendo agua con sus compañeros salvajes.

Feather (plumaje) Otra característica que poseen más de 4 de cada 5 animales es que no tiene plumaje. Por lo pronto el grupo de "peces" (*tipo 4*) completo no lo tiene y seguramente por el soporte que muestran las reglas asociadas a tener plumaje, cerca del 0.2, estas pertenecen a un grupo de aves. Voy a comprobar a quien pertenecen estas reglas y como los caracteriza. Para ello obtengo las reglas con un soporte mínimo de 0.15 y confianza de 0.65.

Como se observa en la figura 3 tenemos 13 reglas sin repetidas de las cuales 11 determinan sin temor de duda cualidades de los animales con plumaje como no tener dientes, pelo, aleta, o tener cola, piernas, columna, poner huevos, no dar leche, respirar, no venenosos y aéreo. Además, y más esclarecedor, tener plumaje queda determinado por ser de tipo 2, por tanto esto nos deja bastante claro que es un grupo de aves, por el soporte y la confianza de la primera regla se que hay exactamente 20 individuos del grupo este de aves. Voy a visualizar estas reglas en forma de grafo para verlo con mayor claridad, figura 2.

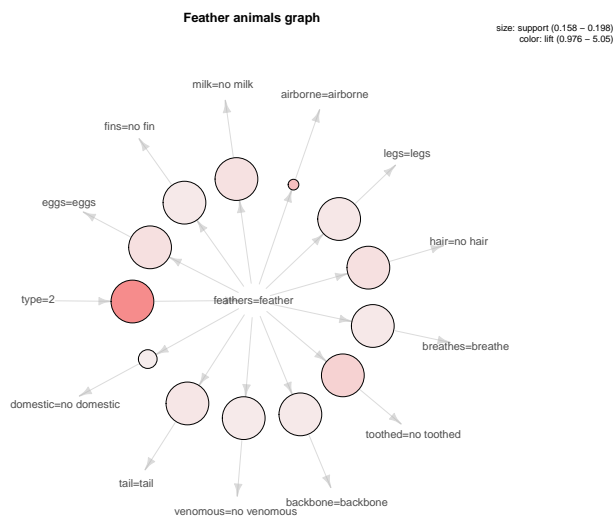


Figure 2: Reglas que contienen el item *feather* graficado en forma de grafo

También, he obtenido la primera regla pero con antecedente y consecuente invertidos, muestra las mismas medidas indicando así que todo animal de tipo 2 tendrá plumaje y todo animal con plumaje será de tipo 2, esto nos permite estudiar con seguridad las reglas [12] y [13] sabiendo que siendo todas aves, por el soporte, hay algunos individuos no voladores y/o domésticos, paso a estudiar estos casos más infrecuentes.

Para poder estudiarlos por su baja repetición me voy obligado a reducir los requisitos a la hora de obtener las reglas, obteniendo así varios millones más. Obtengo reglas para animales de *tipo 2* y de estas filtro en las que aparezca *no voladores*, puedo encontrar aparte de las características ya mencionadas por ser ave información interesante como que la mayoría este subconjunto son según la característica *catsize* de tamaño considerable, como un felino o un gato, y depredador. Aparte, a través de las anomalías detectadas por la regla [13] mencionada anteriormente, tras obtener aves domésticas se obtiene un pequeño grupo de 3 caracterizadas principalmente por no se predadoras y de pequeño tamaño.

	lhs	rhs	support	confidence	lift	certainty
[1]	{type=2}	=> {feathers=feather}	0.1980198	1.00	5.0500000	1.0000000
[2]	{feathers=feather}	=> {toothed=no toothed}	0.1980198	1.00	2.5250000	1.0000000
[3]	{feathers=feather}	=> {hair=no hair}	0.1980198	1.00	1.7413793	1.0000000
[4]	{feathers=feather}	=> {eggs=eggs}	0.1980198	1.00	1.7118644	1.0000000
[5]	{feathers=feather}	=> {milk=no milk}	0.1980198	1.00	1.6833333	1.0000000
[6]	{feathers=feather}	=> {tail=tail}	0.1980198	1.00	1.3466667	1.0000000
[7]	{feathers=feather}	=> {legs=legs}	0.1980198	1.00	1.2948718	1.0000000
[8]	{feathers=feather}	=> {breathes=breathe}	0.1980198	1.00	1.2625000	1.0000000
[9]	{feathers=feather}	=> {backbone=backbone}	0.1980198	1.00	1.2168675	1.0000000
[10]	{feathers=feather}	=> {fins=no fin}	0.1980198	1.00	1.2023810	1.0000000
[11]	{feathers=feather}	=> {venomous=no venomous}	0.1980198	1.00	1.0860215	1.0000000
[12]	{feathers=feather}	=> {airborne=airborne}	0.1584158	0.80	3.3666667	0.7376623
[13]	{feathers=feather}	=> {domestic=no domestic}	0.1683168	0.85	0.9755682	-0.1653846

Figure 3: Reglas que contienen el item *feather* ordenadas por orden de certeza descendiente

Ante toda esta información puedo resumirla, tras comenzar con una búsqueda de animales con plumaje, se que todos estos pertenecen al *tipo 2* de animales según su clasificación y que este grupo es de aves. Podemos dividirlo en dos subgrupos donde tenemos los voladores por norma general salvajes como podrían ser los típicos en un zoo como alcones, buitres o más tipos parecidos. Aparte tenemos un pequeño grupo de no voladores, son poco comunes y no hay muchas especies de aves no voladoras, al ser un zoo no especializado en este tipo de animales lo más usual es que haya algunos pingüinos, o avestruces.

Airborne (volador) Característica que negada se encuentra muy presente en el conjunto de datos lo que nos deja un conjunto reducido de animales con la capacidad de volar. La primera impresión es que como acabo de ver en el subestudio anterior he descubierto como tenemos un conjunto, *tipo 2*, de aves por tanto es obvio que haya un grupo de animales voladores, pero la cuestión está en la diferencia de soportes entre el grupo de animales *tipo 2* y voladores con soportes aproximadamente de 0.2 y 0.25 respectivamente.

Esto indica la existencia en otro u otros grupos de animales con la capacidad de volar pero no clasificado como de *tipo 2*, quizá este grupo son aves pero con unas características comunes que no poseen los voladores de otro grupo, u otra posibilidad es que se encuentren en zonas contiguas dentro del Zoo que están divididos, como se apreciar en la figura 4 del típico mapa.



Figure 4: Mapa ejemplo de zonas en Zoos

Voy a estudiar estos casos de animales voladores pero que no pertenecen al grupo de aves, por lo pronto el único de aves encontrado. Para esto me encuentro como en situaciones anteriores, con un grupo de un soporte 0.25 al que quitaré las reglas asociadas a animales de *tipo 2* con un soporte cercano al 0.2, bajo soporte y confianza para generar estas reglas hasta 0.1 y 0.4 respectivamente.

Tras generar las reglas y filtrar las que aparezcan *airborne* y *tipo 2* elimino redundantes dejando así las reglas que muestra la figura 5.

Entre las reglas obtenidas la mayoría aportan información baladí acerca de las aves o carecen de confianza al no ser fiables, pero las reglas [4] y [6] sacan a la luz datos curiosos.

	lhs	rhs	support	confidence	lift	certainty
[1]	{airborne=airborne}	=> {legs=legs}	0.2376238	1.0000000	1.2948718	1.00000000
[2]	{airborne=airborne}	=> {breathes=breathe}	0.2376238	1.0000000	1.2625000	1.00000000
[3]	{airborne=airborne}	=> {fins=no fin}	0.2376238	1.0000000	1.2023810	1.00000000
[4]	{airborne=airborne}	=> {toothed=no toothed}	0.2178218	0.9166667	2.3145833	0.86202186
[5]	{airborne=airborne}	=> {eggs=eggs}	0.2178218	0.9166667	1.5692090	0.79960317
[6]	{airborne=airborne}	=> {milk=no milk}	0.2178218	0.9166667	1.5430556	0.79471545
[7]	{feathers=feather}	=> {airborne=airborne}	0.1584158	0.8000000	3.3666667	0.73766234
[8]	{airborne=airborne}	=> {catsize=no cat}	0.2079208	0.8750000	1.5504386	0.71306818
[9]	{airborne=airborne}	=> {predator=no predator}	0.1683168	0.7083333	1.5898148	0.47395833
[10]	{airborne=airborne}	=> {aquatic=no aquatic}	0.1881188	0.7916667	1.2301282	0.41550926
[11]	{airborne=airborne}	=> {hair=no hair}	0.1782178	0.7500000	1.3060345	0.41279070
[12]	{airborne=airborne}	=> {tail=tail}	0.1782178	0.7500000	1.0100000	0.02884615
[13]	{airborne=airborne}	=> {venomous=no venomous}	0.2178218	0.9166667	0.9955197	-0.05208333
[14]	{airborne=airborne}	=> {domestic=no domestic}	0.1980198	0.8333333	0.9564394	-0.29487179
[15]	{airborne=airborne}	=> {backbone=backbone}	0.1782178	0.7500000	0.9126506	-0.40277778

Figure 5: Reglas sin redundantes para animales voladores que no son de tipo 2

Primera, se puede observar por sus medidas como no todos los animales voladores no tienen dientes, esto se puede deber a dos hechos.

Las aves no tienen dientes, pero comenta el curioso artículo hay algunas especies de patos que pueden tener el pico serrado para ayudarse en su alimentación pero aún así no se consideran dientes aunque es un error que suele cometerse. Me inclino a pensar que en un Zoo hay especialistas que conocen esta información y por tanto no es un error cometido. Esto me empuja a la otra opción.

Hay un grupo de animales voladores pero no todos son aves, si, hay animales que vuelan o se considera que vuelan aunque no sean aves, como las ardillas, generalmente planean pero no vuelan por si solas. O también los murciélagos, son ejemplos muy comunes, sobre todo en el caso de las ardillas a ser animales comunes en un Zoo. Estos casos particulares recuerdo no corresponden al grupo de aves, por tanto pertenecen a otro grupo seguramente situado en otra zona (*type*) según la hipótesis planteada.

Segunda, la regla [6], muestra la existencia de animales voladores pero no del "grupo de aves" que producen leche, como es un Zoo y no Puleva la característica *milk* se referirá a animales que generan leche para alimentar a sus crías y no para producirla. Esta es una característica de los mamíferos, y restringiendo a voladores nos queda que el único es el murciélago. Esto indica que pertenecientes a otro grupo pueden estar dentro de uno con animales terrestres relacionados con un terreno que podría ser un espacio interno y cerrado, o un entorno mayor que pueda incluir este espacio, o incluso situarse con n grupo de animales sin relación pero situados contiguamente en espacios construidos artificialmente.

Volador mamífero Continuo con el estudio de las anomalías encontradas. Obtengo las reglas y elimino las redundantes quedando así un conjunto formado por 15 reglas que como puede verse en la figura 6 definen de forma unívoca estos animales.

	lhs	rhs	support	confidence	lift	certainty
[1]	{airborne=airborne,type=1}	=> {milk=milk}	0.01980198	1	2.463415	1
[2]	{milk=milk,airborne=airborne}	=> {eggs=no eggs}	0.01980198	1	2.404762	1
[3]	{milk=milk,airborne=airborne}	=> {hair=hair}	0.01980198	1	2.348837	1
[4]	{milk=milk,airborne=airborne}	=> {predator=no predator}	0.01980198	1	2.244444	1
[5]	{milk=milk,airborne=airborne}	=> {catsize=no cat}	0.01980198	1	1.771930	1
[6]	{milk=milk,airborne=airborne}	=> {toothed=toothed}	0.01980198	1	1.655738	1
[7]	{milk=milk,airborne=airborne}	=> {aquatic=no aquatic}	0.01980198	1	1.553846	1
[8]	{milk=milk,airborne=airborne}	=> {tail=tail}	0.01980198	1	1.346667	1
[9]	{milk=milk,airborne=airborne}	=> {legs=legs}	0.01980198	1	1.294872	1
[10]	{milk=milk,airborne=airborne}	=> {breathes=breathe}	0.01980198	1	1.262500	1
[11]	{milk=milk,airborne=airborne}	=> {feathers=no feather}	0.01980198	1	1.246914	1
[12]	{milk=milk,airborne=airborne}	=> {backbone=backbone}	0.01980198	1	1.216867	1
[13]	{milk=milk,airborne=airborne}	=> {fins=no fin}	0.01980198	1	1.202381	1
[14]	{milk=milk,airborne=airborne}	=> {domestic=no domestic}	0.01980198	1	1.147727	1
[15]	{milk=milk,airborne=airborne}	=> {venomous=no venomous}	0.01980198	1	1.086022	1

Figure 6: Reglas sin redundantes para animales voladores que no son de tipo 2 y producen leche

Pertenecientes al *tipo 1*, no ponen huevos, tienen pelo, cola, piernas, dientes y columna, no son predadores, ni acuáticos, no tienen plumaje, respiran y no son ni domésticos ni venenosos, además de producir leche y volar como se filtró. Queda bastante claro y todos los datos concuerdan. Ahora voy a ver que otros animales se integran en el *tipo 1*, estos solo representan dos individuos.

Type 1 Vamos a estudiar este grupo, primero veré que representan los murciélagos dentro de este. Veo las reglas de animales de *tipo 1* y voladores, muestra exactamente lo mismo que obtuve en la figura 6 por tanto dentro del subconjunto solo hay dos animales voladores que son los murciélagos así puedo pasar a su filtrado y quedarme con el resto de animales.

Aún filtrando la cantidad de reglas es inmanejable, vuelvo a generar las reglas subiendo los mínimos así pasarán a filtrarse solas las reglas que incluyan a los murciélagos. Me quedo con una muestra con la que pueda filtrar reglas redundantes generándose el conjunto de la figura 7. Éste muestra como los animales pertenecientes al *tipo 1* son mamíferos por tanto respiran, poseen columna vertebral, ninguno es venenoso y obviamente no tienen plumaje.

	lhs	rhs	support	confidence	lift	certainty
[1]	{type=1}	=> {milk=milk}	0.4059406	1.0000000	2.4634146	1.0000000
[2]	{type=1}	=> {breathes=breathe}	0.4059406	1.0000000	1.2625000	1.0000000
[3]	{type=1}	=> {feathers=no feather}	0.4059406	1.0000000	1.2469136	1.0000000
[4]	{type=1}	=> {backbone=backbone}	0.4059406	1.0000000	1.2168675	1.0000000
[5]	{type=1}	=> {venomous=no venomous}	0.4059406	1.0000000	1.0860215	1.0000000
[6]	{type=1}	=> {eggs=no eggs}	0.3960396	0.9756098	2.3461092	0.9582472
[7]	{type=1}	=> {toothed=toothed}	0.3960396	0.9756098	1.6153539	0.9384146
[8]	{type=1}	=> {hair=hair}	0.3861386	0.9512195	2.2342598	0.9150547
[9]	{type=1}	=> {airborne=no airborne}	0.3861386	0.9512195	1.2477035	0.7947154
[10]	{type=1}	=> {legs=legs}	0.3762376	0.9268293	1.2001251	0.6786850
[11]	{type=1}	=> {catsize=cat}	0.3168317	0.7804878	1.7915743	0.6110398
[12]	{type=1}	=> {aquatic=no aquatic}	0.3465347	0.8536585	1.3264540	0.5894309
[13]	{type=1}	=> {tail=tail}	0.3465347	0.8536585	1.1495935	0.4315197
[14]	{type=1}	=> {fins=no fin}	0.3663366	0.9024390	1.0850755	0.4203730
[15]	{type=1}	=> {domestic=no domestic}	0.3267327	0.8048780	0.9237805	-0.5159475

Figure 7: Muestra de reglas sin redundantes para animales de tipo 1 no voladores

Tras estas reglas comunes en todos, comienzan las interesantes. Me llama primero la atención la sexta regla que me da a entender la existencia de algún individuo mamífero ovíparo o también denominados monotremados, existen pocos tipos de estos.

	lhs	rhs	support	confidence	lift	certainty
[1]	{eggs=eggs,type=1}	=> {aquatic=aquatic}	0.00990099	1	2.805556	1
[2]	{toothed=no toothed,type=1}	=> {eggs=eggs}	0.00990099	1	1.711864	1
[3]	{eggs=eggs,type=1}	=> {toothed=no toothed}	0.00990099	1	2.525000	1
[4]	{eggs=eggs,type=1}	=> {milk=milk}	0.00990099	1	2.463415	1
[5]	{eggs=eggs,milk=milk}	=> {type=1}	0.00990099	1	2.463415	1
[6]	{eggs=eggs,type=1}	=> {hair=hair}	0.00990099	1	2.348837	1
[7]	{eggs=eggs,type=1}	=> {catsize=cat}	0.00990099	1	2.295455	1
[8]	{eggs=eggs,type=1}	=> {predator=predator}	0.00990099	1	1.803571	1
[9]	{eggs=eggs,type=1}	=> {tail=tail}	0.00990099	1	1.346667	1
[10]	{eggs=eggs,type=1}	=> {airborne=no airborne}	0.00990099	1	1.311688	1
[11]	{eggs=eggs,type=1}	=> {legs=legs}	0.00990099	1	1.294872	1
[12]	{eggs=eggs,type=1}	=> {breathes=breathe}	0.00990099	1	1.262500	1
[13]	{eggs=eggs,type=1}	=> {feathers=no feather}	0.00990099	1	1.246914	1
[14]	{eggs=eggs,type=1}	=> {backbone=backbone}	0.00990099	1	1.216867	1
[15]	{eggs=eggs,type=1}	=> {fins=no fin}	0.00990099	1	1.202381	1
[16]	{eggs=eggs,type=1}	=> {domestic=no domestic}	0.00990099	1	1.147727	1
[17]	{eggs=eggs,type=1}	=> {venomous=no venomous}	0.00990099	1	1.086022	1
[18]	{aquatic=aquatic,toothed=no toothed,type=1}	=> {eggs=eggs}	0.00990099	1	1.711864	1
[19]	{eggs=eggs,aquatic=aquatic,type=1}	=> {toothed=no toothed}	0.00990099	1	2.525000	1
[20]	{eggs=eggs,toothed=no toothed,type=1}	=> {aquatic=aquatic}	0.00990099	1	2.805556	1

Figure 8: Muestra de reglas con redundantes para animal ovíparo dentro del *type 1*

En concreto tengo un caso en los datos, coincidiendo sus datos, como tener piel cubierta de pelo, no tiene dientes o los tiene atrofiados, además se observa en las reglas que tiene cola y es acuático.

Actualmente se conocen cinco especies, el ornitorrinco y cuatro de hormigueros espinosos, todas ellas habitantes de la región Australiana en zonas próximas a las costas. El ornitorrinco es el único acuático, en realidad semiacuático.

Tras este estudio puedo afirmando la presencia de un ornitorrinco en el Zoo, como muestra la figura 9 y se menciona en su documentación es una animal protegido en todo su ámbito de distribución por tanto para poder verlo vivo tendría que ser en Australia, por tanto los datos de este Zoo provienen de Australia.

Grupo de reglas para mamíferos Para terminar voy a concluir con el trabajo y el estudio del grupo de animales *type 1*, sabemos que mamíferos, donde he encontrado un subgrupo de murciélagos y un ornitorrinco.

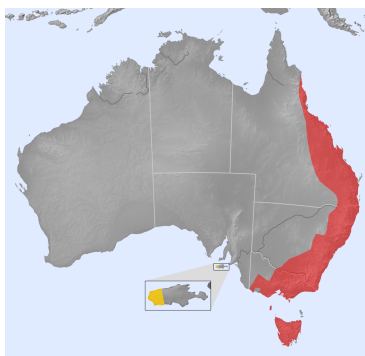


Figure 9: Distribución en la región Australiana de la presencia del Ornitorrinco

Entre las reglas selecciono las que se cumplan siempre (certainty=1) para buscar reglas cuyo antecedente sea igual al consecuente de estas primeras. Repetiré el proceso hasta encontrar grupos de reglas.

no tail->hair	->no venomous	->breathe ->backbone	->no venomous ->breathe ->milk	->no venomous ->breathe ->backbone ->no feather
no tail->no airborne no tail->toohted no tail->no eggs	->breathe ->backbone ->milk ->no feather	->milk ->no feather	->no feather	

Table 2: Despliegue de búsqueda para grupos de reglas.

Como se puede ver solo profundizando en 4 búsquedas ya se obtienen algunos grupos de reglas como se observa en la tabla 2.

- no venomous -> backbone
backbone -> no venomous
- no venomous -> backbone
backbone -> milk
milk -> no venomous
- backbone -> milk
milk -> backbone

Estos grupos de reglas no indican que dentro de los mamíferos los animales que poseen columna vertebral no son venenosos. Este grupo también está expandido con producir leche, pero como sabemos que estudiamos los mamíferos este grupo no nos aporta información útil. Finalmente tenemos otro grupo que me confirma la totalidad del grupo de mamíferos, todos producen leche, tienen columna vertebral, aunque tampoco importa información puesto que todos los mamíferos poseen esas dos cualidades.

Extensión Dejo aquí el estudio del este grupo y de tantos otros items que presentan valores potenciales para estudiar presentes en la figura 1.

En contrapartida voy a pasar a estudiar esta información que me parece más interesante y finalizar el estudio, ya se hace repetitivo el estudio aplicando las mismas técnicas y me queda poco papel. Más adelante se aplicarán el estudio por grupos de reglas e items negados.

2 Conclusión Australiana

Conociendo la información mencionada puedo aislar a una parte y reducir de forma drástica la incertidumbre acerca del Zoo. Como muestra la figura 10 (búsqueda GoogleMap) estamos entre uno de una lista de aproximadamente diez zoos, entre los que no todos tienen Ornitorrincos. El primero más famoso que cumple las restricciones es MelbourneZOO, muy probablemente este sea el zoo que he estudiado, suponiendo claro que la obtención de un dataset mundialmente conocido de un Zoo lo más normal es que se haga en todo caso sobre uno famoso antes que sobre uno desconocido.

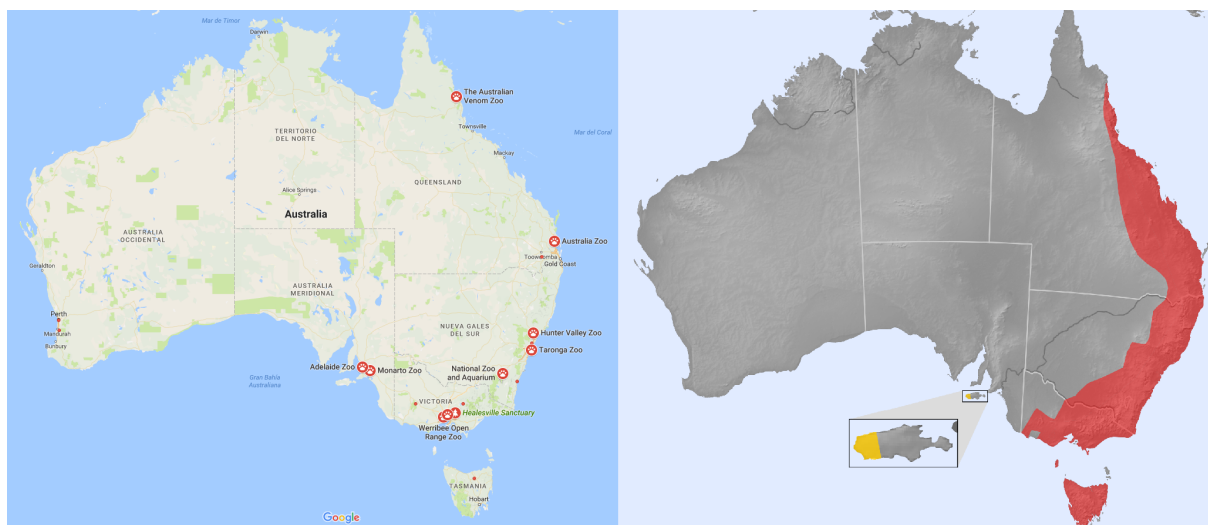


Figure 10: Distribución en la región Australiana de la presencia del Ornitorrinco (imagen derecha) y zoos de Australia (imagen izquierda)

3 Items negados

La utilidad de esta técnica reside en datasets con variables que tomen un conjunto de valores no muy grande. En mi caso las variables que tengo, en general, son de tipo lógico además de haber realizado esta operación pero de una forma distinta. En vez de seleccionar las tuplas para estudiar el contrario de algo que me pudiera interesar, como por ejemplo si estudio animales salvajes y quiero estudiar los no salvajes. Este proceso lo he ido realizando durante el estudio pero de otra forma que era a partir de todas las reglas filtrar mediante operadores lógicos, un ejemplo de ello es la figura 11. Durante el estudio de animales voladores de tipo 2 pasé a estudiar lo que fueran voladores no de tipo 2.

```
#airbone animals no bird
subset <- subset(rulesZoo, subset=lhs %in% "airborne=airborne" | rhs %in% "airborne=airborne")
subset <- subset(subset, subset=!(lhs %in% "type=2" | rhs %in% "type=2"))
```

Figure 11: Ejemplo de código para obtención de items negados a mi manera

Ahora que he pasado a aplicar la forma que me menciona el profesor Jesus Alcalá, generándose en igualdad de soporte y confianza mínimo una cantidad de reglas similar pero con la comodidad de trabajar directamente sobre reglas más simples contando con una restricción implícitamente.

Seleccionado las instancias de individuos sin cola y generando las reglas obtengo un conjunto de reglas como muestra la figura 12, donde en el antecedente de cada regla contaría con el item "tail=no tail".

```
set of 8998 rules
> inspect(head(rulesZooNoTail, n=100))
```

	lhs	rhs	support	confidence	lift	certainty
[1]	{predator=no predator}	=> {breathes=breathe}	0.4615385	1	1.368421	1
[2]	{predator=no predator}	=> {fins=no fin}	0.4615385	1	1.040000	1
[3]	{aquatic=no aquatic}	=> {fins=no fin}	0.6153846	1	1.040000	1
[4]	{hair=no hair}	=> {milk=no milk}	0.6153846	1	1.300000	1
[5]	{hair=no hair}	=> {eggs=eggs}	0.6153846	1	1.300000	1
[6]	{hair=no hair}	=> {domestic=no domestic}	0.6153846	1	1.130435	1
[7]	{hair=no hair}	=> {fins=no fin}	0.6153846	1	1.040000	1
[8]	{toothed=no toothed}	=> {backbone=no backbone}	0.6538462	1	1.529412	1
[9]	{backbone=no backbone}	=> {toothed=no toothed}	0.6538462	1	1.529412	1
[10]	{toothed=no toothed}	=> {milk=no milk}	0.6538462	1	1.300000	1
[11]	{toothed=no toothed}	=> {eggs=eggs}	0.6538462	1	1.300000	1
[12]	{toothed=no toothed}	=> {fins=no fin}	0.6538462	1	1.040000	1
[13]	{backbone=no backbone}	=> {milk=no milk}	0.6538462	1	1.300000	1
[14]	{backbone=no backbone}	=> {eggs=eggs}	0.6538462	1	1.300000	1
[15]	{backbone=no backbone}	=> {fins=no fin}	0.6538462	1	1.040000	1
[16]	{milk=no milk}	=> {eggs=eggs}	0.7692308	1	1.300000	1

Figure 12: Ejemplo de código para obtención de items negados a mi manera