DETERMINACIÓN MEDIANTE TÉCNICAS DE APRENDIZAJE AUTOMÁTICO DE LOS ATRIBUTOS RELEVANTES PARA LA CLASIFICACION OFICIAL SEUROP DE CANALES BOVINAS LIGERAS Y PESADAS

Jorge Díez[1,2], Iván Fernández[2], Antonio Bahamonde[1], Jaime Alonso[1] Luis José Royo[2], Isabel Álvarez[2] y Félix Goyache[2*]

[1] Centro de Inteligencia Artificial de la Universidad de Oviedo, Campus de Viesques, 33271 Gijón (Asturias); [2]SERIDA-CENSYRA-Somió, C/ Camino de los Claveles 604, 33203 Gijón (Asturias); *e-mail: fgoyache@serida.org

INTRODUCCIÓN

Nuestro grupo aplica técnicas de Aprendizaje Automático a la clasificación de la conformación bovina tanto en animales vivos (GOYACHE et al., 2001a) como en canal (GOYACHE et al., 2001b). El Aprendizaje Automático es uno de los aspectos básicos de la Inteligencia Artificial y puede utilizarse para determinar qué atributos de la canal tienen mayor importancia para la obtención de una calificación concreta sin importar que las relaciones existentes entre los diferentes atributos entre sí y con la nota de clasificación pueda no ser lineal. El sistema de clasificación oficial de canales bovinas SEUROP se diseñó para ser utilizado en canales provenientes de animales a partir de 300 kg de peso vivo. Este hecho hace que se comparen entre sí animales de muy distinto grado de madurez fisiológica lo que puede inducir a un aumento en el error de las clasificaciones. La conformación de las canales de animales en los primeros momentos de la pubertad es difícilmente comparable con las de animales con suficiente madurez fisiológica. El objetivo de la presente comunicación es comprobar la validez de la metodología SEUROP en canales bovinas ligeras y pesadas en el marco del proyecto MCyT TIC2001-3579.

MATERIAL Y MÉTODOS

Se han utilizado datos procedentes de 29 canales ligeras y 68 canales pesadas. En la clasificación de la conformación SEUROP (±) participaron tres clasificadores de la Empresa Asturiana de Servicios Agrarios (EASA) responsable del control de la D.O. "Ternera Asturiana". Cada clase de conformación se puntuó numéricamente de 1 (P) a 6 (S). Los signos + y – se consideraron como +0.25 o –0.25 puntos de conformación respectivamente, sobre la puntuación de la clase correspondiente. Cada clasificación individual ha sido considerada como un ejemplo independiente por lo que el conjunto de entrenamiento del sistema de AA incluyó 65 ejemplos de canales ligeras y 162 de canales pesadas. Siguiendo la práctica habitual en mercados de referencia como los de Binefar o Lleida se consideraron canales ligeras aquellas con un peso canal menor de 200 kg. Los clasificadores de EASA obtuvieron tres fotografías digitales de cada canal en vistas dorsal, lateral interna y lateral externa. Esas fotografías se procesaron mediante la aplicación informática descrita en DÍEZ et al. (2001) marcándose en cada canal 21 puntos y 5 perfiles para obtener los siguientes atributos de cada canal: anchura del muslo, longitud del muslo, ratio entre la anchura y longitud del muslo, longitud de la pierna, longitud de la espalda, anchura de la espalda, altura lateral de la espalda, área de la espalda, volumen de la espalda, perfil de la pierna, perfil de la espalda, perfil de la tapa, longitud del lomo, anchura anterior del lomo, anchura posterior del lomo, ratio entre la anchura anterior

y posterior del lomo, anchura lateral del lomo en la cadera, anchura lateral del lomo en la cruz, ratio entre las anchuras laterales del lomo, volumen del lomo, perfil de la cadera, perfil del muslo, perfil del morrillo, longitud de la canal, profundidad del pecho, profundidad del flanco. Estos atributos se añadieron al peso de la canal y a la nota de conformación obtenida por cada clasificador, por lo que los conjuntos de entrenamiento incluían finalmente 27 atributos y una clase (nota de conformación). La búsqueda de atributos relevantes para obtener los conjuntos de entrenamiento utilizables como entrada de los algoritmos de aprendizaje se realizó mediante la combinación de herramientas las herramientas BETS y FA que se describen en GOYACHE et al. (2001b). BETS permite establecer una ordenación de los atributos en función de su relevancia (calidad de las reglas clasificatorias) permitiendo distinguir entre atributos relevantes e irrelevantes. FA estudia el conjunto de atributos seleccionados por BETS eliminando de forma iterativa el atributo menos relevante y comprobando la bondad del conjunto resultante mediante una validación cruzada aplicando el método del vecino más próximo. El conjunto de atributos adoptado finalmente como solución será aquel que menor error genere. Si se admite un cierto incremento en el error obtenido con el conjunto de partida cuando se reduce el número de atributos se realiza un *filtrado agresivo*. Los conjuntos de entrenamiento de canales ligeras y pesadas se analizaron con tres sistemas de AA independientes: Cubist, SAFE y BETS. Todos estos sistemas generan reglas de regresión que permiten obtener un valor numérico de las categorías y están descritos someramente en GOYACHE et al. (2001a,b). Para medir la calidad del aprendizaje en la clasificación de las canales de cada sistema de AA se calculó el error absoluto medio cometido mediante validaciones cruzadas. Cada conjunto de entrenamiento se dividió en 10 particiones. Cada una de esas particiones se utilizó sucesivamente como test del aprendizaje realizado con las 9 restantes. Cada experimento se repitió 5 veces considerándose la media de los errores obtenidos como una correcta estimación del rendimiento de los diferentes sistemas de AA empleados. Además, se analizaron los conjuntos de entrenamiento seleccionados mediante regresión lineal (SAS, 1999).

RESULTADOS Y DISCUSIÓN

El acierto de los sistemas de predicción de la se evaluó como la desviación sobre la media de la clasificación de los expertos. Las clasificaciones individuales de los expertos se desviaban, a su vez, de la media en torno a 0,40 puntos. El error obtenido en la clasificación mediante regresión lineal es mayor que con los sistemas de AA especialmente cuando se reduce el número de atributos incluidos en el conjunto de entrenamiento. Cubist y BETS presentan un comportamiento similar para ambos tipos de canales, en cualquier caso comparable al error cometido individualmente por los expertos. El número de funciones utilizadas para obtener las clasificaciones señala que las relaciones entre atributos no es lineal, aunque este aspecto aparece más destacado en las canales pesadas. Los atributos utilizados para realizar la clasificación de canales pesadas es coherente con la metodología SEUROP al utilizar exclusivamente atributos de perfiles (tapa) y, especialmente, desarrollos musculares. Para la clasificación de canales ligeras los clasificadores usan atributos que caracterizan el desarrollo esquelético del animal (profundidad del flanco) y de desarrollo esquelético. Los clasificadores parecen experimentar cierto grado de incertidumbre al clasificar canales ligeras, lo que les lleva a tener en cuenta el grado de desarrollo del animal. Además de usar atributos diferentes los

clasificadores los aplican con criterios distintos. La utilización de las reglas de aprendizaje obtenidas con las canales ligeras sobre el conjunto de entrenamiento de canales pesadas y viceversa produjo un error inaceptable en el entorno del 1,0. Nuestros resultados sugieren que las canales ligeras son un producto diferente de las canales pesadas que debería tener un sistema de clasificación propio que evitara errores en la clasificación y permitiera la normalización de un mercado importante en la cuenca mediterránea. Por otra parte, nuestros resultados afectan a los sistemas de selección de sementales por la conformación de la canal de sus hijos.

Tabla 1: Errores medios absolutos y número de funciones seleccionadas para la predicción de la conformación en canales bovinas ligeras o pesadas obtenidos mediante regresión lineal y dos sistemas de AA mediante validación cruzada.

Numero de atributos	Regresión lineal	BETS		SAFE		Cubist	
	Error	Error	Prototipos	Error	Funciones	Error	Funciones
Canales ligeras							
27	0,58	0,38	9,58	0,37	1,34	0,40	3,54
11*	0,34	0,35	10,30	0,34	1,72	0,37	2,22
10**	0,33	0,32	11,04	0,35	1,96	0,36	1,92
3***	0,57	0,40	9,40	0,45	4,32	0,40	4,58
Canales pesadas							
27	0,45	0,44	24,06	0,43	3,06	0,42	5,58
8*	0,51	0,42	22,60	0,47	5,92	0,45	5,52
7**	0,53	0,44	23,78	0,46	6,54	0,44	6,00

*filtrado normal; ** primer filtrado agresivo, *** segundo filtrado agresivo

Tabla 2: Lista de atributos seleccionados para aprender a clasificar canales ligeras y pesadas tras un procedimiento de estimación de la relevancia

Canales	Número de atributos	Atributos utilizados
ligeras	3	Profundidad del flanco, anchura del muslo, relación largo/ancho del muslo
pesadas	7	Longitud del muslo, perfil de la espalda, anchura posterior del lomo, longitud del lomo, mantenimiento de la anchura del lomo, perfil de la tapa, desarrollo muscular de la espalda

REFERENCIAS BIBLIOGRÁFICAS

DÍEZ, J., ALONSO, J., LÓPEZ, S., BAHAMONDE, A., GOYACHE, F 2001. Variables de conformación que influyen en la clasificación oficial SEUROP de canales bovinas. ITEA, 22 Vol. Extra II, 550-552.

GOYACHE F., DEL COZ J.J., QUEVEDO J.R., LÓPEZ S., ALONSO J., RANILLA J., LUACES O., ALVAREZ I., BAHAMONDE A. 2001a. Using artificial intelligence to design and implement a morphological assessment system in beef cattle, Animal Science, 73, 49-60

GOYACHE, F., BAHAMONDE, A. ALONSO, J., LÓPEZ, S., ALONSO, J., DEL COZ J.J., QUEVEDO, J.R., RANILLA, J., LUACES, O., ALVAREZ, I., ROYO, L. DÍEZ J. 2001b, The usefulness of Artificial Intelligence techniques to assess subjective quality of products in the food industry. Trends in Food Science and Technology, 12: 370-381.