

## Classificação Andrológica de Touros avaliada por Machine Learning

Silvio André Isler<sup>1</sup>, Urbano Gomes Pinto de Abreu<sup>2</sup>, Ériklis Nogueira<sup>3</sup>, Vanessa Aparecida de Moraes Weber<sup>1</sup>,  
Felipe de Oliveira Pedro<sup>1</sup>, Fabiana de Andrade Melo Sterza<sup>1</sup>, Dauydisson Antonio Gonzalez Cordeiro<sup>4</sup>

<sup>1</sup>Universidade Estadual de Mato Grosso do Sul, Aquidauana, MS, Brasil; <sup>2</sup>Embrapa Pantanal, Corumbá, MS, Brasil; <sup>3</sup>Embrapa Gado de Corte, Campo Grande, MS, Brasil; <sup>4</sup>Universidade Federal de Mato Grosso do Sul, Campo Grande, MS, Brasil  
\*e-mail: islersilvio@gmail.com

O estudo foi realizado com o objetivo de verificar o comportamento da classificação dos Exames Andrológicos (EA) de touros utilizados em monta natural através do Aprendizado de Máquina (ML). Utilizou-se um banco de dados de EA de 2308 touros, com os seguintes atributos para cada exame: região, raça, genótipo, classe de idade, consistência testicular, circunferência escrotal, volume de sêmen, turbilhonamento espermático, motilidade espermática, vigor espermático, defeito de acrossoma, gota citoplasmática proximal, defeito de cabeça, defeito de peça média, defeitos maiores totais, gota citoplasmática distal, cabeça isolada normal, defeitos de cauda, defeitos menores totais, defeitos totais e espermatozoides normais totais. A classificação desses animais para cada teste foi feita como aptos (n = 1088), inaptos (n = 672) e questionáveis (n = 548). Esta base de dados foi submetida à ferramenta de ML Waikato Environment for Knowledge Analysis (WEKA) através da sua ferramenta Auto Weka, que sugeriu o algoritmo Random Forest, (uma coleção de árvores de decisão). O ML foi capaz de identificar resultados semelhantes ao EA para a maioria dos touros, principalmente para os animais classificados como aptos (94,1% - 1024/1088) e inaptos (95,2% - 522/548). Para os animais classificados como duvidosos, os resultados do ML coincidiram com os do EA em apenas 59,5% (400/672) dos touros. Assim, de um modo geral, 84% (1946/2308) dos touros têm os mesmos resultados utilizando o ML em relação aos EA, com um erro absoluto médio de 0,1651 e um erro quadrático médio de 0,2761. Concluímos que a ML pode ser utilizada para auxiliar os andrologistas na classificação dos resultados dos EA. Estudos futuros poderão aumentar a precisão e até determinar quais as características que melhor representam as classes na classificação final usando ML.

**Palavras-chave:** Exames Andrológicos, Monta Natural, Random Forest, WEKA.

**Agradecimentos:** À Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) e à Fundação para o Desenvolvimento do Ensino, Ciência e Tecnologia em Mato Grosso do Sul (FUNDECT/SECTEI, processo 16/2016).

## Andrological Classification of Bulls evaluated by Machine Learning

Silvio André Isler\*<sup>1</sup>, Urbano Gomes Pinto de Abreu<sup>2</sup>, Ériklis Nogueira<sup>3</sup>, Vanessa Aparecida de Moraes Weber<sup>1</sup>, Felipe de Oliveira Pedro<sup>1</sup>, Fabiana de Andrade Melo Sterza<sup>1</sup>, Daudyson Antonio Gonzalez Cordeiro<sup>4</sup>

<sup>1</sup>Universidade Estadual de Mato Grosso do Sul, Aquidauana, MS, Brasil; <sup>2</sup>Embrapa Pantanal, Corumbá, MS, Brasil; <sup>3</sup>Embrapa Gado de Corte, Campo Grande, MS, Brasil; <sup>4</sup>Universidade Federal de Mato Grosso do Sul, Campo Grande, MS, Brasil  
Rmail: \*islersilvio@gmail.com

The study was conducted with the aim of verifying the behavior of the classification of Andrological Exams (AE) of bulls used in natural mating through Machine Learning (ML). A database of AE of 2308 bulls was used, with the following attributes for each examination: region, breed, genotype, age class, testicular consistency, scrotal circumference, semen volume, sperm swirl, sperm motility, sperm vigor, acrosome defect, proximal cytoplasmic drop, head defect, midpiece defect, total major defects, distal cytoplasmic drop, normal isolated head, tail defects, total minor defects, total defects and total normal sperm. The classification of these animals for each test was made as fit (n = 1088), unfit (n = 672) and questionable (n = 548). This database was submitted to the ML tool Waikato Environment for Knowledge Analysis (WEKA) through its tool Auto Weka, which suggested the Random Forest algorithm, a collection of decision trees. ML was able to identify results similar to AE for most bulls, especially for animals classified as fit (94.1% - 1024/1088) and unfit (95.2% - 522/548). For animals classified as doubtful, the results of the ML matched those of the AE in only 59.5% (400/672) of the bulls. Thus, in general, 84% (1946/2308) of the bulls have the same results using ML or AE, with a mean absolute error of 0.1651 and a root mean square error of 0.2761. We concluded that ML can be used to assist andrologists in classifying the results of AE. Further studies could increase the accuracy and even determine which features best represent the classes in the final classification using ML.

**Keywords:** Andrological Exams, Natural Mating, Random Forest, WEKA.

**Acknowledgements:** To the Coordination for the Improvement of Higher Education Personnel (CAPES) and the Foundation for the Development of Education, Science and Technology in Mato Grosso do Sul (FUNDECT/SECTEI, process 16/2016).