

## **Aplicação da Mineração de Dados com a Ferramenta WEKA na Extração de Conhecimentos de Medidas Morfométricas de Cavalos Pantaneiros<sup>8</sup>**

**Eduardo Lourenço dos Santos<sup>9</sup>**

**Bruno Ruiz de Barros<sup>10</sup>**

**Glória Maria Gutierrez Vargas Freitas<sup>11</sup>**

**João Batista Garcia<sup>12</sup>**

**Sandra Aparecida Santos<sup>13</sup>**

O núcleo de criação de cavalos Pantaneiros da EMBRAPA Pantanal, criado em 1988, possui uma grande quantidade de dados relacionados com caracterização morfométrica da raça. Com o intuito de extrair conhecimentos e padrões de 15 medidas lineares: altura da cernelha (AC), altura do dorso (AD), altura da garupa (AG), altura dos costados (ACost), comprimento da cabeça (CC), comprimento do pescoço (CP), comprimento dorso-lombar (CD), comprimento da garupa (CG), comprimento da espádua (CE), comprimento do corpo (Ccorp), largura da cabeça (LCD), largura do peito (LP), largura da anca (LA), perímetro torácico (PT) e perímetro da canela (PC), comumente tomadas no registro da raça, objetivou-se desenvolver métodos de mineração de dados adaptados às necessidades de descoberta de conhecimentos da base de dados, processo KDD. A mineração dos dados consiste na eficiência da ferramenta de domínio público WEKA em qualificar os dados segundo a interferência de seu analista e da fonte de abastecimento dos dados tidos pelo núcleo de criação de cavalos Pantaneiros. Para a mineração, dados de 15 medidas lineares de cavalo Pantaneiro e seguiu em 4 etapas definidas: entrada de dados, análise, tratamento e geração de gráficos estatísticos. Os dados de entrada foram organizados em arquivo formato (ARFF), interpretável pela ferramenta WEKA. Em seguida os dados das medidas do cavalo pantaneiro foram minerados, analisados estatisticamente e visualizados na forma de gráficos. A aplicação da mineração dos dados sobre as medidas lineares dos cavalos Pantaneiros proporcionou a descoberta de padrões e conhecimentos da base de dados dos cavalos registrados. Análise dos gráficos permitiu ao analista verificar as relações de variáveis de maior e menor ocorrências, indicando os registros de maior ocorrências estar dentro do padrão da raça, e os de menor ocorrências, fora do padrão. Os processos de mineração dados sobre as 15 medidas lineares é de extrema importância como auxílio no manejo e à tomada de decisão na seleção e melhoramento do cavalo Pantaneiro.

<sup>8</sup> Embrapa Pantanal - Conservação e manejo adaptativo de pastagens nativas – Cavalo Pantaneiro

<sup>9</sup> Acadêmico do Instituto Federal do Mato Grosso do Sul, Corumbá, MS (eduardolourenco\_1@hotmail.com)

<sup>10</sup> Acadêmico do Instituto Federal do Mato Grosso do Sul, Corumbá, MS (brunnoruiz@hotmail.com)

<sup>11</sup> Acadêmica da Universidade Federal de Mato Grosso do Sul e bolsista Pibic/CNPq na Embrapa Pantanal, Corumbá, MS (gloriamfreitas@msm.com)

<sup>12</sup> Analista da Embrapa Pantanal, Caixa Postal 109, 79320-900, Corumbá, MS (joao.garcia@embrapa.br)

<sup>13</sup> Pesquisadora da Embrapa Pantanal, Caixa Postal 109, 79320-900, Corumbá, MS (sandra.santos@embrapa.br)