

Inteligência Artificial aplicada ao mapeamento das redes de drenagem

Luis Fernando Bueno ^{1,2}
Janielson da Silva Lima³
Tatiane Emílio Chechia ^{2,3}
Tony Vinicius Moreira Sampaio ²

¹Centro Gestor e Operacional do Sistema de Proteção da Amazônia - CENSIPAM
Setor Policial Sul, Área 05 – Quadra 03 – Bloco K – CEP 70610-200 - Brasília – DF
luis.bueno@sipam.gov.br

² Universidade Federal do Paraná - UFPR
Programa de Pós-Graduação em Geografia
Av. Cel. Francisco H dos Santos, 100 - Centro Politécnico - sala 108 do Edifício João José
Bigarella - CEP 81531-980 Caixa Postal 19001 - Curitiba - PR
tonysampaio@ufpr.br

³Fundação Universidade Federal de Rondônia – UNIR
Núcleo de Ciências Exatas e da Terra – NCET - Departamento de Geografia
Rodovia BR 364, Km 9,5 – Porto Velho – Rondônia
janyelson@gmail.com
tati.chechia@gmail.com

Abstract. This article discusses the use of Artificial Intelligence techniques applied to the question of mapping the drainage networks. This study aimed to assess the relevance of using the computational tools of Artificial Intelligence in the study of issues related to improving the accuracy of the mapping of drainage networks. In this context, exploratory spatial data analysis was performed in Data Mining environment where thematic attributes derived from data on hydrography, geology, geomorphology, hydrology, vegetation and soil were considered. It was found that the different data mining activities described relationships between the variables used, and pointed out the importance of the attributes drainage system, hydrogeological units and soil in the study of the drainage area of research. The findings also point to the opportunity of building an Artificial Neural Network to represent the drainage study that considers several variables involved in the derivation of drainage networks process area.

Palavras-chave: drainage network mapping, Artificial Intelligence, Data Mining , exploratory analysis

1. Introdução

O mapeamento automatizado da rede de drenagem vem sendo aperfeiçoado ao longo dos anos, visando a obtenção simplificada e com menor custo de aquisição de dados. O desenvolvimento de técnicas e ferramentas matemáticas computacionais possibilitaram que as redes de drenagem, bem como os respectivos padrões morfométricos, sejam extraídos diretamente de Modelos Digitais de Terenos – MDT.

Em detrimento dos avanços tecnológicos e conceituais no campo dos algoritmos para extração de redes de drenagem, Sampaio (2008) discutiu aspectos inerentes à subjetividade do processo de mapeamento. Para Sampaio (2008) o mapeamento da rede de drenagem, manual ou automático, continua sujeito à processos subjetivos. É relevante o estudo da acurácia dos mapeamentos da rede de drenagem extraídos por meio de processos automáticos, com vistas a ultrapassar os obstáculos inerentes aos atuais métodos disponíveis. Pode-se, portanto, pensar em novos arranjos tecnológicos que possam contornar as limitações e contribuir para o aumento da acurácia.

O avanço tecnológico da Computação quando aplicado à Geografia pode contribuir significativamente para melhorar o poder de expressividade dos modelos, potencializar o uso de metodologias consagradas e aumentar a capacidade de análise. Como argumentou Ross (2006) a Geografia necessita do uso rotineiro das Tecnologias da Informação para desempenhar seu objeto principal de análise, que é conhecer e analisar as relações da humanidade entre si, e desta com o meio natural.

Ao longo do tempo as bases de dados estão se tornando cada vez mais volumosas. O crescimento de tais bases de dados sugere que, pouco a pouco, ficará impraticável manipular manualmente os dados em diversos domínios, incluindo o de Geografia. Faz-se necessário, portanto, considerar o uso de ferramentas automatizadas para manipular grandes volumes de dados, extrair informações oportunas e gerar conhecimento atualizado. Fayyad, Piatetsky-Shapiro e Smyth (1996) defendiam em seu trabalho que a necessidade de intensificar as capacidades de análise humana para lidar com o grande número de bytes é de natureza econômica e financeira.

Complementarmente, considera-se os aspectos abordados por Cunico e Oka-Fiori (2009) que sugeriram em seu trabalho a adoção do conceito de totalidade, em oposição à perspectiva unitária, para análise e avaliação sobre os recursos naturais, no qual os elementos envolvidos não se apresentam de maneira dissociada, e sim interação de maneira dinâmica e em diferentes escalas. Acredita-se que o ferramental computacional, sobretudo do ramo da Inteligência Artificial, pode ajudar significativamente este tipo de análise.

Este artigo apresenta resultados parciais de uma pesquisa que propõe a reunião de diversas tecnologias computacionais num ambiente efetivamente integrado, que permita configurar cenários variados para o estudo da acurácia das redes de drenagem extraídas de forma automática. Especificamente, objetivou-se aqui verificar a relevância do uso do ferramental computacional da Inteligência Artificial no estudo das questões inerentes à melhoria da acurácia do mapeamento das redes de drenagem. Desta forma, técnicas de Inteligência Artificial foram utilizadas para explorar um banco de dados espaciais construído com dados de uma bacia hidrográfica localizada na Amazônia Brasileira.

2. Metodologia de Trabalho

A área de estudo corresponde à Bacia Hidrográfica do Rio Mutum-Paraná, localizada no noroeste do Estado de Rondônia, estendendo-se pelos municípios de Porto Velho e Nova Mamoré e delimitada pelas coordenadas 9°34'40" e 10°01'49" de latitude Sul e 65°15'34" e 64°57'51" de longitude Oeste, com área de aproximadamente 3.560 km², conforme figura 1 apresentada a seguir.

Quanto à geologia da área Rizzoto et al (2005) descreveram as seguintes unidades litoestratigráficas: Complexo Jamari; Formação Mutum-Paraná; Granito Serra da Muralha; Suíte Intrusiva Serra da Providência; Suíte Metamórfica Nova Mamoré; Suíte Laje; Suíte Intrusiva São Lourenço-Caripunas; Suíte Intrusiva Rondônia; Formação Palmeiral; Coberturas Cenozóicas (Formação rio Madeira, Formação Jaci-Paraná, Cobertura Detrito-Laterítica, Sedimentos Aluvionares Argilosos e Arenosos, Sedimentos Aluvionares Indiscriminados).

Adamy e Dantas (2005) descreveram a geomorfologia da Bacia do Mutum-Paraná, dividindo-a em duas partes: Bacia do Alto Rio Mutum-Paraná e Bacia do Baixo Rio Mutum-Paraná. Em relação a Bacia do Alto Rio Mutum-Paraná, afirmaram Adamy e Dantas (2005) que caracteriza-se por um relevo colinoso medianamente dissecado apresentando áreas com uma dissecação variável entre alta a baixa. Ao contrário das bacias dos rios Jaci-Paraná e Candeias, que também são afluentes do rio Madeira entre Porto Velho e Jirau, a bacia do rio Mutum-Paraná, de menor abrangência, não drena as vertentes escarpadas da serra dos Pacaás Novos, mas apenas alguns de seus contrafortes mais rebaixados (ADAMY E DANTAS,

2005). Adamy e Dantas (2005) apresentaram duas unidades geomorfológicas maiores neste subambiente, representadas pelas Superfícies de Aplanamento e Planícies Aluviais de Rios Secundários.

Segundo os supramencionados autores, nesta parte da Bacia do Rio Mutum, os terrenos são embasados por rochas do Complexo Jamari, onde predominam solos Podzólicos Vermelho-Amarelos álicos, que de acordo com os autores foram posteriormente reinterpretados como Argissolos Vermelho-Amarelos Alumínicos e Latossolos Amarelo Alumínicos, sendo caracterizados por solos espessos, argilosos, bem estruturados e com expressiva variação textural entre os horizontes A e Bt.

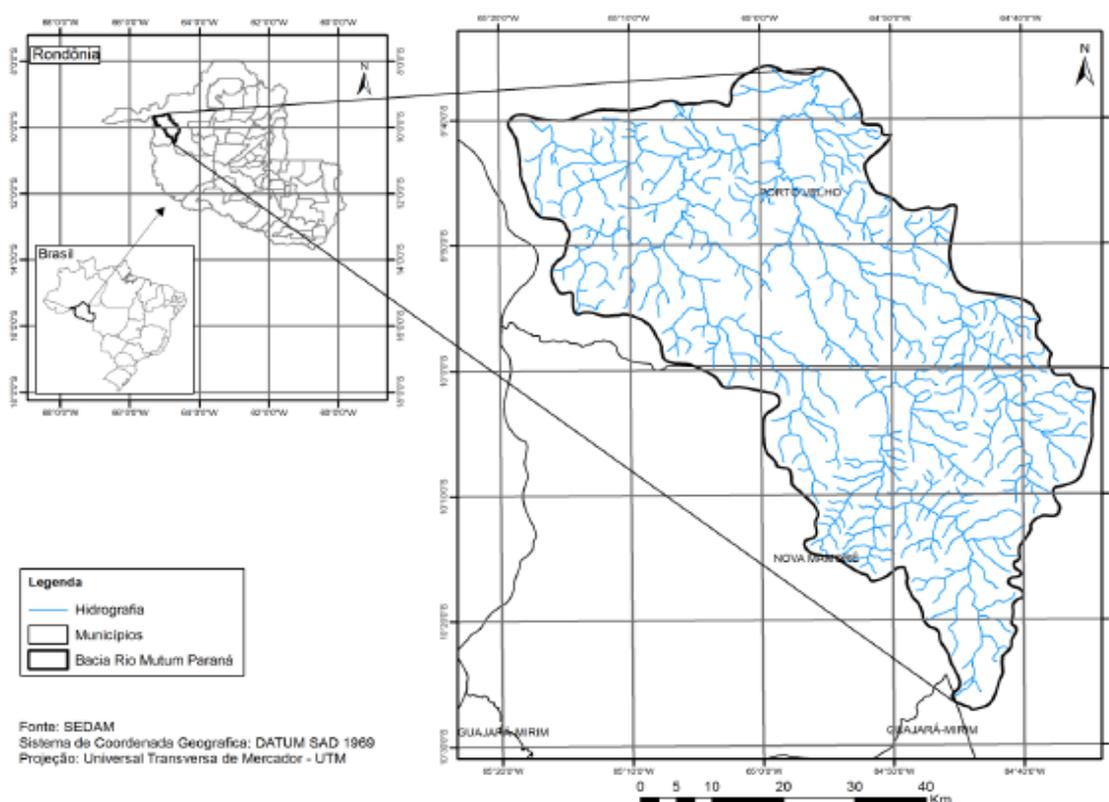


Figura 1. Localização da área de estudo.

O ambiente foi caracterizado como pouco alterado pela intervenção humana, recoberto por Floresta Tropical aberta, notabilizando-se por sua estabilidade morfodinâmica frente aos processos erosivo-deposicionais e a movimentos de massa (ADAMY E DANTAS, 2005). Para os autores citados, a fraca declividade das vertentes das colinas associada à descontinuidade hidráulica existente no contato dos horizontes A e B dos Argissolos podem desencadear algumas ocorrências erosivas, mas de pouco significado enquanto estes terrenos mantiverem-se florestados.

A Bacia do Baixo Rio Mutum-Paraná é constituída pela Bacia do rio Mutum-Paraná e pelas bacias de igarapés menores que drenam diretamente para a margem direita do rio Madeira tais como os igarapés Jirau e Cirilo, caracteriza-se por um relevo plano, muito pouco dissecado, inserido no Planalto Rebaixado da Amazônia Ocidental. Localmente, pode exibir faixas com um grau de dissecação mais acentuado (ADAMY E DANTAS, 2005).

Os supra mencionados autores descreverem que a área deste subambiente é caracterizada por uma ampla superfície de aplainamento, de relevo plano a muito suavemente ondulado, entre as cotas de 200 e 300 metros, e que apresentam graus de dissecação variando entre

baixo e alto. Os autores afirmaram ainda que foram identificados uma baixa ocorrência de relevos residuais, tais como inselbergs, hillocks e tors.

De acordo com Adamy e Dantas (2005) nestes terrenos embasados por rochas do Complexo Jamari, predominam Latossolos Vermelho-Amarelos álicos, que se caracterizam por solos muito espessos, argilosos, bem drenados e estruturados.

Novamente os autores destacam que o ambiente é recoberto pela Floresta Tropical Aberta, pouco alterado pela intervenção humana, excetuando-se ao longo da rodovia BR-364, onde as áreas apresentam estabilidade morfodinâmica frente aos processos erosivo-deposicionais e a movimentos de massa. A morfologia quase plana das áreas aplainadas e dos baixos platôs associada a solos e mantos de intemperismo espessos e bem drenados indica uma vulnerabilidade muito baixa com relação aos processos erosivo-deposicionais em terrenos florestados (ADAMY E DANTAS, 2005).

A hidrogeologia da área foi descrita por Melo Júnior (2005), que identificou diferentes unidades hidrogeológicas, a saber os Aquíferos Intergranulares Descontínuos Livres, os Aquíferos Intergranulares/Fraturados Contínuos Livres, os Aquíferos Locais Restritos as Zonas Fraturadas, os Aquíferos Fraturados Descontínuos Livres, e os Aquíferos.

Melo Júnior (2005) caracterizou ainda a vulnerabilidade natural das águas subterrâneas da região da Bacia do Mutum-Paraná, concluindo que os índices de vulnerabilidade variam de altos a extremamente altos nas porções centro-noroeste e centro-sudoeste da bacia.

Os dados espaciais necessários para a construção do banco de dados espaciais, foram obtidos junto à diversas instituições governamentais, tais como o Sistema de Proteção da Amazônia – SIPAM, o Serviço Geológico do Brasil – CPRM, o Instituto Nacional de Pesquisa Espaciais – INPE, a Prefeitura Municipal de Porto Velho, a Secretaria de Estado do Desenvolvimento Ambiental – SEDAM-RO, além de dados oriundos da empresa Santo Antonio Energia. Foi realizada uma seleção dos dados, de modo a definir apenas os dados que abrangiam a área geográfica relativa à bacia do estudo. Inicialmente, os dados vetoriais selecionados foram convertidos em ambiente SIG e reprojatados para o sistema SIRGAS 2000, UTM Zona 20 S.

O software escolhido para uso no projeto foi o QGIS, versão 2.4 Chugiak, que por meio das opções de processamento incorpora funcionalidades do SAGA versão 2.1, Grass versão 6.4.3 e Orfeo Toolbook. O QGIS também dispõe das funcionalidades providas pela biblioteca GDAL. Adicionalmente, funcionalidades do pacote TauDEM (Terrain Analysis Using Digital Elevation Models), que consiste num conjunto de ferramentas para a extração e análise de informações hidrológicas e topográficas a partir de modelos digitais de elevação, e que implementa diversos algoritmos para modelos hidrológicos, foram configuradas para funcionar no ambiente do QGIS.

Um banco de dados espaciais foi criado usando o Sistema Gerenciador de Banco de Dados PostgreSQL, habilitado com a extensão espacial PostGIS. PostgreSQL é um moderno sistema para banco de dados, que implementa conceitos avançados de administração de bases de dados, robusto e amplamente utilizado. PostGIS é uma extensão que habilita o PostgreSQL manipular dados espaciais em conformidade com os padrões estabelecidos pelo Open Geospatial Consortium, instituição que estabelece os padrões de interoperabilidade para dados geoespaciais. Desta forma, a base de dados reúne dados vetoriais que cobrem a área da Bacia do Mutum-Paraná.

Foi utilizado o software WEKA – Waikato Environment for Knowledge Analysis, versão 3.6 como ferramenta de mineração de dados. WEKA é um software desenvolvido e mantido pela Universidade de Waikato, que consiste numa coleção de algoritmos de aprendizado de máquina para tarefas de mineração de dados. Um requisito essencial para o uso do WEKA consiste em converter os dados para o formato de arquivo ARFF (Attribute Relation Format File). A partir daí, as funcionalidades providas pelo ambiente de operação do WEKA permite

explorar as tarefas básicas de mineração de dados para os propósitos desta pesquisa, a saber: processamento de dados; regressão; classificação; clustering; associação e; visualização e interpretação.

O reconhecimento da área de estudo relativa à Bacia do Mutum-Paraná foi feita por meio de visita *in-loco*. Assim, foi realizado trabalho de campo, no mês de setembro de 2013, com objetivo de coletar dados e verificar a qualidade da representação da hidrografia no material cartográfico e, da existência ou não do fluxo de água ou canal em campo, para verificar a acurácia da base cartográfica. O registro dos pontos dos locais visitados em campo foi feito com uso de equipamento GPS.

A representatividade amostral para a coleta de dados foi definida nos termos especificados por Andriotti (2003) (Equação 1) onde E representa o erro amostral tolerável. O conhecimento do número total de nascentes permitiu corrigir a fórmula em função do tamanho N (Equação 2), onde n0 é o tamanho da amostra, o que resultou em um n amostral mínimo de 140 bacias de drenagem (Equação 3).

$$n = \frac{1}{E^2} \quad (1)$$

$$n = \frac{(N * n0)}{(N + n0)} \quad (2)$$

$$n = \frac{(331 * 241)}{(331 + 241)} \quad (3)$$

3. Resultados e Discussão

Apesar do n amostral indicar a necessidade de 140 bacias de drenagem, os trabalhos de campo permitiram verificar 241 pontos amostrais, conforme apresentado na figura 2 a seguir, dos quais 41,91% foram devidamente mapeados na base cartográfica enquanto que os outros 58,09% não encontravam-se mapeados, conforme tabela 1 apresentada adiante. Observou-se ainda que todos os pontos não mapeados referiram-se à rios intermitentes.

Tabela 1. Pontos coletados em campo – mapeados x não mapeados

	Quantidade	Percentual	Regime Drenagem			
			Intermitente		Permanente	
			Quantidade	Percentual	Quantidade	Percentual
Mapeados	101	41,91	79	78,21	22	21,79
Não mapeados	140	58,09	140	100,00	-	-
Total	241	100,00	-	-	-	-

A análise exploratória executada em ambiente de mineração de dados permitiu realizar inferências sobre o conjunto de dados, considerando atributos temáticos diversos, a saber regime de drenagem (hidrografia); unidade litoestratigráfica, litotipo e grupos de rochas (geologia); padrão de relevo (geomorfologia); unidade hidrogeológica (hidrogeologia); tipo de vegetação (vegetação), tipo de solo (solo) e classificação do ponto (mapeado/não mapeado).

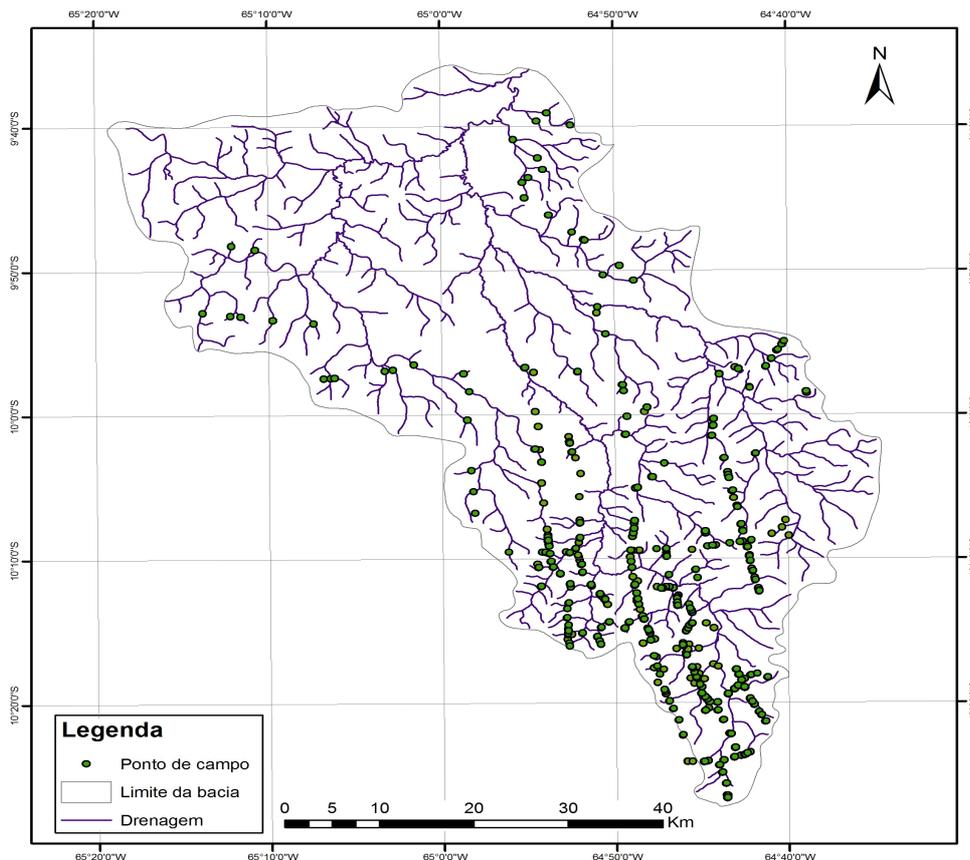


Figura 2. Pontos de coleta de dados em campo na área de estudo.

A análise exploratória executada em ambiente de mineração de dados permitiu realizar inferências sobre o conjunto de dados, considerando atributos temáticos diversos, a saber regime de drenagem (hidrografia); unidade litoestratigráfica, litotipo e grupos de rochas (geologia); padrão de relevo (geomorfologia); unidade hidrogeológica (hidrogeologia); tipo de vegetação (vegetação), tipo de solo (solo) e classificação do ponto (mapeado/não mapeado).

A aplicação do algoritmo J48 para classificação dos pontos coletados resultou na árvore apresentada na figura 3 a seguir.

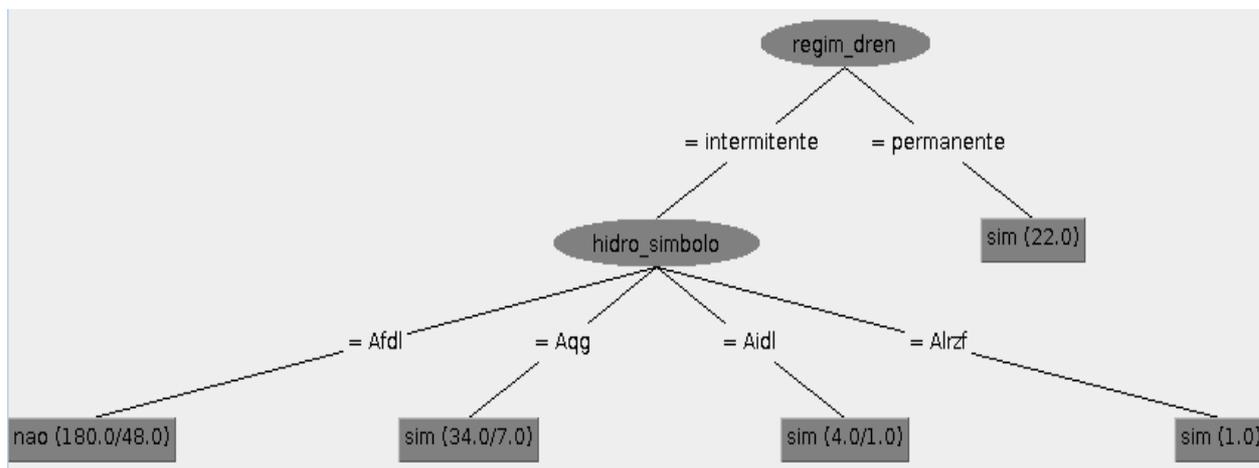


Figura 3. Árvore de decisão gerada com o algoritmo J48 do WEKA.

O algoritmo J48 é considerado um clone do algoritmo C4.5, e implementa no WEKA um classificador cujo aprendizado baseia-se em árvore de decisão. Desta forma, dado um conjunto de casos, uma árvore inicial é criada por meio do algoritmo J48 de divisão e conquista; se todos os casos no conjunto S pertencem à mesma classe, ou S é pequeno, a árvore é marcada com a classe mais frequente em S; caso contrário, deve-se escolher um teste baseado em um simples atributo, com dois ou mais resultados. Caso contrário, escolherá um teste com base em um único atributo com dois ou mais resultados. Aplica-se o teste para a raiz da árvore, formando um ramo para cada um dos resultados do teste, particionando S nos subconjuntos S1, S2, etc., de acordo com o resultado em cada caso, aplicando-se o mesmo procedimento recursivamente para cada subconjunto.

A árvore de decisão resultante da classificação com o J48 demonstrou que dentre os dez atributos avaliados apenas dois deles são de fato relevantes para determinar a classe dos pontos: regime de drenagem (regim_dren) e unidade hidrogeológica (hidro_simbolo). Conforme mencionado, o regime de drenagem foi intermitente para todos os pontos não mapeados. A árvore de decisão ressaltou que grande parte dos pontos não mapeados estavam inseridos na unidade hidrogeológica Aquíferos Fraturados Descontínuos livres.

Resultado semelhante foi obtido com o avaliador de atributos CfsSubsetEval. O algoritmo CfsSubsetEval implementado no WEKA avalia o valor de um subconjunto de atributos ao considerar a capacidade de previsão individual e o grau de redundância entre eles. Este tipo de algoritmo testa a correlação entre o atributo avaliado e o atributo classe. O método de busca usado na seleção de atributos foi o BestFirst que apontou os atributos regime de drenagem, unidade hidrogeológica e tipo de solo como os mais relevantes no conjunto de dados.

Para identificação de agrupamentos (cluster) foi executado o algoritmo EM. O algoritmo EM atribui uma probabilidade de distribuição para cada instância, indicando a probabilidade dela pertencer a cada um dos agrupamentos. Este algoritmo pode decidir quantos grupos para criar por validação cruzada, ou pela especificação a priori de um determinado número de grupos. De modo não supervisionado, o algoritmo resultou em sete agrupamentos distintos.

A formulação de regras de associação entre os dados também foi testada com o uso do algoritmo Apriori. De forma iterativa este algoritmo reduz o suporte mínimo até encontrar o número necessário de regras com confiança mínima estabelecida. Dentre as regras de associação sugeridas pelo Apriori, destacam-se aquelas que demonstraram os relacionamentos: a) entre os Aquíferos Fraturados Descontínuos Livres e as Superfícies de Aplanamento; b) entre o regime de drenagem intermitente, juntamente com os Aquíferos Fraturados Descontínuos Livres e as Superfícies de Aplanamento; c) entre a unidade litoestratigráfica Complexo Metamórfico Nova Mamoré e as Superfícies de Aplanamentos; e entre a unidade litoestratigráfica Complexo Metamórfico Nova Mamoré e os Aquíferos Fraturados Descontínuos Livres.

4. Conclusões

De fato a Inteligência Artificial, por meio de suas diversas técnicas, pode gerar conhecimentos novos e úteis para o estudo das redes de drenagem, os quais poderão embasar ações para melhorar a acurácia dos mapeamentos.

A mineração de dados demonstrou-se ser uma ferramenta útil para apoiar a análise exploratória, e permitiu descrever relação entre as variáveis do estudo. Os diversos algoritmos aplicados possibilitaram selecionar atributos, classificar e agrupar os pontos coletados e estabelecer regras de associações entre os dados. Os resultados obtidos com a mineração de dados sugerem o aprofundamento dos estudos para detalhar se as técnicas aplicadas efetivamente contribuirão na identificação de nascentes e mapeamento de cursos d'água.

De forma complementar, na sequência desta pesquisa pretende-se aplicar técnicas de mineração de dados em imagens de radar que cobrem a área da Bacia Hidrográfica do Rio Mutum-Paraná. Acredita-se que, assim como no caso da análise de dados espaciais vetoriais, esta técnica aplicada a dados matriciais de alta resolução poderá resultar em conhecimento útil ao estudo das redes de drenagem.

No decorrer da pesquisa vislumbrou-se a oportunidade de se utilizar a técnica de Rede Neural Artificial para representar a rede de drenagem da área de estudo. WEKA permite o trabalho conjunto com Redes Neurais Artificiais. No aprofundamento dos estudos as redes neurais serão usadas nesta pesquisa com a finalidade de identificar variáveis que influenciam a derivação das redes de drenagem. Opcionalmente, pretende-se investigar ainda o uso das redes neurais para a classificação das imagens usadas como insumos no trabalho.

Referências Bibliográficas

ADAMY, Amílcar; DANTAS, Marcelo Eduardo. Geomorfologia. In: **Projeto Rio Madeira. Levantamento de informações para subsidiar o estudo de viabilidade do aproveitamento hidrelétrico (AHE) do Rio Madeira. AHE Jirau: relatório final.** Coordenado por Gilmar José Rizzotto e José Guilherme Ferreira de Oliveira, organizado por Marcos Luiz E. S. Quadros, João Marcelo R. de Castro, Antônio Cordeiro, Amílcar Adamy, Homero Reis de Melo Junior e Marcelo Eduardo Dantas. - Porto Velho: CPRM - Serviço Geológico do Brasil, 2005.

ANDRIOTTI, J.L.S. **Fundamentos de Estatística e Geoestatística.** Rio Grande do Sul: Ed. UNISINOS, 2003.

CUNICO, Camila; OKA-FIORI, Chisato. Identificação e análise dos condicionantes físicos relevantes à vulnerabilidade ambiental: comparação entre a Serra do Mar e o Primeiro Planalto Paranaense. **Revista Geografar – Resumos do VII Seminário Interno de Pós-Graduação em Geografia**, p. 14 – 17, 2009.

FAYYAD, Usama; PIATETSKY-SHAPIRO, Gregory; SMYTH, Padhraic. From Data Mining to knowledge Discovery in Databases. In: *AI Magazine*. Fall, p. 37-54, 1996.

MELO JUNIOR, Homero Reis de. Hidrogeologia. In: **Projeto Rio Madeira. Levantamento de informações para subsidiar o estudo de viabilidade do aproveitamento hidrelétrico (AHE) do Rio Madeira. AHE Jirau: relatório final.** Coordenado por Gilmar José Rizzotto e José Guilherme Ferreira de Oliveira, organizado por Marcos Luiz E. S. Quadros, João Marcelo R. de Castro, Antônio Cordeiro, Amílcar Adamy, Homero Reis de Melo Junior e Marcelo Eduardo Dantas. - Porto Velho: CPRM - Serviço Geológico do Brasil, 2005.

RIZZOTTO, G.J. **Projeto Rio Madeira. Levantamento de informações para subsidiar o estudo de viabilidade do aproveitamento hidrelétrico (AHE) do Rio Madeira. AHE Jirau: relatório final.** Coordenado por Gilmar José Rizzotto e José Guilherme Ferreira de Oliveira, organizado por Marcos Luiz E. S. Quadros, João Marcelo R. de Castro, Antônio Cordeiro, Amílcar Adamy, Homero Reis de Melo Junior e Marcelo Eduardo Dantas. - Porto Velho: CPRM - Serviço Geológico do Brasil, 2005.

ROSS, Jurandy. **Ecogeografia do Brasil: Subsídios para planejamento ambiental.** Sao Paulo: Oficina de Textos, 2006. 207 pg.

SAMPAIO, Tony Vinicius Moreira. **Parâmetros morfométricos para melhoria da acurácia do mapeamento da rede de drenagem uma proposta baseada na análise da Bacia Hidrográfica do Rio Benevente – ES.** Belo Horizonte: UFMG, 2008. Tese (Doutorado em Geografia), Programa de Pós-Graduação em Geografia, Universidade Federal de Minas Gerais, 2008.