

XVI SEMINÁRIO  
DE INTEGRAÇÃO

25 E 26  
OUTUBRO  
de 2017

DO GLOBAL AO LOCAL: O PODER DAS ESCALAS SOBRE O TERRITÓRIO



## Estudo das Condições Climáticas em uma Planta Geradora de Energia Eólica Utilizando Mineração de Dados

Elias Rocha Gonçalves Júnior

Virgínia Siqueira Gonçalves

Cláudio Luiz Melo de Souza

Geórgia Regina Rodrigues Gomes

Grupo de Trabalho: ST3. Transformações econômicas recentes: Crise, Indústria do Petróleo e Portos

### Resumo

Devido às condições climáticas atuais e às preocupações quanto à geração de energia por fontes renováveis em âmbito mundial, a geração eólica torna-se uma alternativa interessante do ponto de vista ambiental. Com isso, este artigo tem como objetivo apresentar análises de um conjunto de dados climatológicos de uma estação meteorológica automática, situada no município de Natal/RN, com o intuito de verificar parâmetros básicos, tais como velocidade média dos ventos e temperatura ambiente. Para tal, foi utilizado um processo de descoberta de conhecimento (KDD) com uso de técnicas de mineração de dados e a ferramenta *open source* de *data mining* WEKA 3.8.1 para determinar o período mais adequado do dia para utilização da planta eólica em questão, garantindo melhor eficiência na geração de energia, aumentando a vida útil e reduzindo as intervenções de manutenção. Segundo as análises realizadas, pode-se afirmar que o período da tarde é mais adequado à utilização dos equipamentos de geração eólica de energia, além de demonstrar que temperaturas altas garantem que a velocidade média dos ventos seja adequada, ou seja, adequada e garantindo eficiência na geração de energia elétrica.

**Palavras-chave:** KDD. Energias renováveis. Mineração de dados. WEKA.

Mestrando em Engenharia de Produção – Ucam-Campos – [ucam-campos.br](http://ucam-campos.br)

Mestranda em Engenharia de Produção – Ucam-Campos – [ucam-campos.br](http://ucam-campos.br)

Professor do Mestrado em Engenharia de Produção – Ucam-Campos – [ucam-campos.br](http://ucam-campos.br)

Professora do Mestrado em Engenharia de Produção – Ucam-Campos – [ucam-campos.br](http://ucam-campos.br)

## **1. Introdução**

A humanidade enfrenta um grande desafio que é suprir a demanda de energia evitando agressões ao meio ambiente. A energia eólica é parte da solução deste problema pelo fato de se tratar de uma fonte renovável de energia, abundante e limpa, além de ter origem na própria atmosfera terrestre, ou seja, a dinâmica desta fonte de energia deriva da força dos ventos originada a partir do aquecimento da superfície terrestre pelo Sol.

No entanto, a origem da energia eólica gera dificuldades na integração deste tipo de geração à rede elétrica, pois, por depender das condições atmosféricas, a quantidade de energia que será gerada é de difícil previsão e as condições do vento não podem ser controladas (DALMAZ et al., 2008).

Um grande impulso à geração eólica de eletricidade ocorreu após a primeira crise do petróleo, com vários países investindo em pesquisas sobre novas formas de geração de energia, o que permitiu que esta fonte de energia se destacasse. A fim de aumentar a participação da energia eólica na matriz energética, vários países estabeleceram metas específicas de ampliação desta fonte, sendo que, até 2020, pretende-se, na Europa e nos Estados Unidos, atingir 10% e 6%, respectivamente, de eletricidade de origem eólica (PECORA, 2006).

O Brasil está iniciando a exploração da energia eólica com um total de 15 parques eólicos, localizados em sete estados, totalizando uma potência de 240 MW, o que representa 0,24% da potência instalada no país. Para o desenvolvimento das fontes alternativas, foi regulamentado, em 2004, o PROINFA (Programa de Incentivo

às Fontes Alternativas de Energia Elétrica), que, por meio de financiamentos e garantias de compra e preço da energia, visa aumentar a participação das fontes alternativas na matriz energética brasileira (DALMAZ et al., 2008).

Segundo o Atlas do Potencial Eólico Brasileiro (2001), o Brasil tem um potencial de geração eólico estimado em 143,47 GW, considerando apenas os locais com velocidade média anual dos ventos acima de 7 m/s. Algumas regiões que se destacam no Brasil são: o litoral do Nordeste, principalmente do Ceará e do Rio Grande do Norte, o litoral do Rio Grande do Sul, as Serras Gaúcha e Catarinense, alguns locais do litoral Catarinense e a região dos campos entre Paraná e Santa Catarina.

Diante deste panorama, este trabalho tem por objetivo avaliar, utilizando técnicas de Mineração de Dados, a variação da velocidade dos ventos e da temperatura ambiente, identificando a viabilidade de utilização da planta geradora de Natal/RN e o período do dia mais adequado para utilização dos equipamentos, com base em aspectos técnicos e econômicos.

## **2. Referencial teórico**

### **2.1 Viabilidade de geração elétrica**

A avaliação precisa do potencial de vento em uma região é o primeiro e fundamental passo para o aproveitamento do recurso eólico como fonte de energia.

Para que a energia eólica seja considerada tecnicamente aproveitável, o Plano Nacional de Energia-2030 (PNE 2030) aponta que a velocidade média dos ventos que podem indicar viabilidade técnica do aproveitamento eólico para equipamentos de pequeno porte para sistemas isolados é de 4m/s a 10m de altura, sendo este o caso da planta em questão.

Segundo Dutra (2004), estas informações de velocidade do vento são importantes, porém, para determinação do potencial eólico para geração elétrica, é necessária a análise desses dados durante vários anos

A fim de distinguir quais regiões possuem mais atratividade para instalação de novos parques eólicos, Reinhard (2009) afirma que diversos aspectos

determinantes, sendo estes econômicos, técnicos, sociais e ambientais, devem ser levados em consideração nos modelos de apoio à decisão. Desse modo, a decisão não pode ser baseada unicamente em um único eixo-dimensional de avaliação, como custo ou benefício, sendo utilizados múltiplos critérios, que, em alguns casos, são conflitantes de natureza distintas e inerentes à decisão.

## **2.2 Mineração de Dados**

As técnicas relacionadas à extração de conhecimento e padrões tendo como base um banco de dados é uma área da inteligência computacional denominada como *Knowledge Discovery in Databases* (KDD).

De acordo com Murasse e Tsunoda (2006), a mineração de dados, ou *data mining*, pode ser compreendida como:

[...] etapa do KDD responsável pela seleção dos métodos a serem utilizados para localizar padrões nos dados, seguida da efetiva busca por padrões de interesse numa forma particular de representação, juntamente com a busca pelo melhor ajuste dos parâmetros do algoritmo para a tarefa em questão.

Fayyad (1996) expõe a Mineração de Dados como um processo para identificação, nos dados utilizados como base, padrões válidos, úteis e compreensíveis.

O processo de KDD é composto pelos seguintes passos: seleção, pré-processamento e limpeza, transformação, mineração de dados (*data mining*) e interpretação/avaliação. Como se pode notar, o processo compreende todo o ciclo que o dado percorre até tornar-se conhecimento ou informação, vide Figura 1.

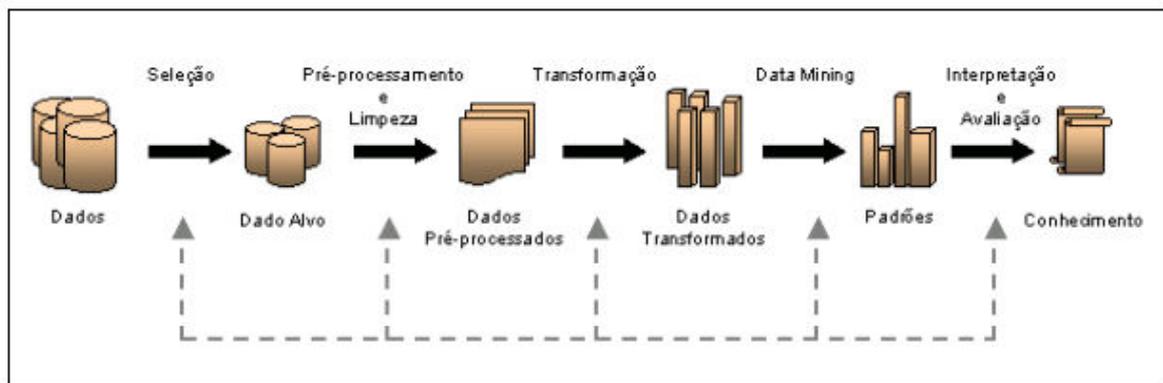


FIGURA 1: O ciclo do processo de KDD. Fonte: Adaptação de FAYYAD et al. (1996).

Para Salvador et al. (2009), a clusterização, uma das técnicas da mineração de dados, busca agrupar os dados de tal maneira que seja capaz de potencializar a similaridade dos objetos de um mesmo grupo e/ou diferença entre grupos distintos.

Segundo Fayyad et al. (1996), árvore de decisão pode ser definida modelo preditivo que pode ser visualizado na forma de uma árvore, daí seu nome. Cada ramo da árvore é uma questão de classificação e cada folha é uma partição do conjunto de dados com sua classificação.

A tarefa de associação tem como objetivo identificar associações entre registros de dados que, de alguma maneira, estão ou devem estar relacionados. Sua premissa básica é encontrar elementos que implicam na presença de outros em uma mesma transação (SCHUNEIDER, 2002). Alguns algoritmos que utilizam os conceitos desta tarefa são as regras de associação e os padrões sequenciais.

### 3. Metodologia

O presente artigo trata da aplicação de KDD sobre um conjunto de dados climatológicos que são essenciais para determinação do período do dia mais adequado a utilização da planta eólica de geração de energia. Para tal, os dados levantados foram velocidade do vento a uma altura de 10 metros e umidade do ar, ambos medidos a cada minuto do dia, ao longo de um determinado período. Deve-se salientar que este artigo não utiliza fatores econômicos e tecnológicos na avaliação do sistema renovável para geração de energia elétrica.

A ferramenta de *data mining* utilizada foi o WEKA versão 3.8.1 e foi desenvolvida pela Universidade de Waikato na Nova Zelândia. Ela pode ser definida como uma coleção de algoritmos do tipo *machine learning* para executar tarefas de *data mining*.

O WEKA vem sendo cada vez mais aplicado e algumas características interessantes ajudam a explicar seu sucesso (MARKOV e RUSSELL, 2006):

- Contém diversos algoritmos para *data mining*, *web mining* e *machine learning*;
- Tem código aberto e está disponível na Web gratuitamente;
- É relativamente fácil de usar, inclusive por pessoas que não sejam especialistas;
- Proporciona recursos flexíveis para experimentos;
- É mantido atualizado, pois novos algoritmos são adicionados logo que surgem na literatura.

O processo de mineração de dados compreende os seguintes (MURASSE e TSUNODA *apud* MARKOV e RUSSELL, 2006):

- Levantar a fontes de dados (bancos de dados, relatórios, etc.);
- Efetuar uma limpeza dos dados para “carregar” para o WEKA;
- “Carregar” para o WEKA o arquivo pós-limpeza dos dados;
- Buscar padrões relevantes.

A descoberta do conhecimento sobre as condições climatológicas foi realizada com o levantamento de 23.041 dados gerados nos primeiros quinze dias do mês de março (período entre 01/03/2016 a 15/03/2016), oriundos da estação meteorológica de superfície automática do Instituto Nacional de Meteorologia (INMET). A estação está situada em uma área costeira na cidade de Natal, capital do estado do Rio de Grande do Norte.

Na busca para identificar padrões e propriedades comuns nos dados climatológicos, foram utilizadas três variáveis, sendo período do dia, vide Quadro 1, velocidade média dos ventos, vide Quadro 2, e temperatura ambiente, vide Quadro 3. Foram utilizadas neste trabalho as de clusterização, associação e árvore de

decisão, para observar a relação entres as mesmas e a confiabilidade do resultado encontrado.

Quadro 1. Critérios de classificação para tratamento de dados do período do dia

Período do dia	Entre 00h00min e 05h59min	Entre 06h00min e 11h59min	Entre 12h00min e 17h59min	Entre 18h00min e 23h59min
Classificação	Madrugada	Manhã	Tarde	Noite

Fonte: Elaborado pelos autores.

Quadro 2. Critérios de classificação para tratamento de dados da velocidade média dos ventos

Velocidade média dos ventos	Abaixo de 4 m/s	Acima de 4 m/s
Classificação	Inadequada	Adequada

Fonte: Elaborado pelos autores.

Quadro 3. Critérios de classificação para tratamento de dados da temperatura ambiente

Temperatura ambiente	Valores entre 18,0°C e 23,9° C	Valores entre 24,0° C e 28,9° C	Valores acima de 29,0°C
Classificação	Amena	Média	Alta

Fonte: Elaborado pelos autores.

#### 4. Resultados e Discussões

O estudo teve como alvo dois atributos: velocidade média dos ventos e a radiação temperatura ambiente ao longo de um período de quinze dias, medindo-se as variáveis a cada minuto do dia.

Na criação de árvore de decisão, utilizou-se o algoritmo de J48, que demonstrou que o período da tarde é o mais adequado para utilização das turbinas geradoras, pois a velocidade média dos ventos é sempre classificada como “Adequada”, evitando usos desnecessários e aumentando os períodos entre as manutenções. Além disso, foi possível identificar que, no período da manhã, a velocidade dos ventos é adequada quando a temperatura ambiente for alta e

também que, no período da noite, esta só é considerada adequada com temperaturas altas e médias. Já, para o período da madrugada, apenas com temperaturas amenas, é que recomenda-se o uso da planta geradora. A referida árvore possui confiança de aproximadamente 72%, vide Figura 2.

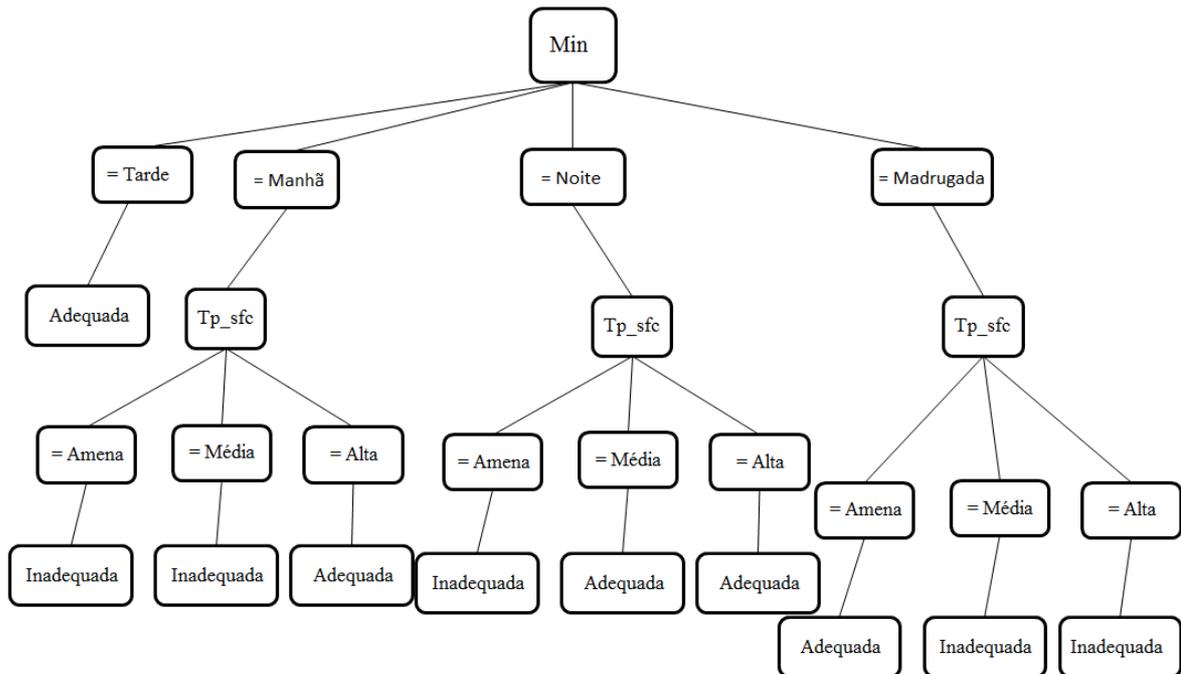


FIGURA 2. Tela do software WEKA versão 3.8.1: algoritmo J48

Já na etapa de clusterização, utilizou-se o algoritmo de SimpleKMeans, com a formação de cinco clusters, vide Figura 3. O cluster #1 identifica que se estiver no período da tarde e com temperatura alta, esta será adequada à utilização da planta geradora, com as melhores condições possíveis, vide Figura 3. Este apresenta 26% dos dados utilizados, o que, para este caso, é satisfatório, além de complementar a informação obtida na árvore de decisão, a qual indica que no período da tarde, a velocidade média dos ventos é sempre adequada.

O cluster #2 corrobora o fato apresentado na árvore de decisão, que indica o período da manhã, apresentando temperatura amena, como inadequado à geração de energia, bem como o cluster #4, que refere-se ao período da madrugada, com temperatura média, com velocidade média dos ventos inadequada.

No caso do cluster #3, há a indicação de que, quando a temperatura ambiente for média e no período da noite, indica-se a utilização dos equipamentos para gerar

energia, pois as condições são ideias para tal, assim como é explanado na árvore de decisão, apresentada anteriormente.

```

Final cluster centroids:
Attribute      Full Data      Cluster#
                (23040.0)      0          1          2          3          4
=====
min            Tarde  Madrugada  Tarde  Manhã  Noite  Madrugada
tp_sfc        Média   Média     Alta   Média  Média  Média
ws_10m        Adequada Adequada Adequada Inadequada Adequada Inadequada

Time taken to build model (full training data) : 0.09 seconds

=== Model and evaluation on training set ===

Clustered Instances

0      4904 ( 21%)
1      5897 ( 26%)
2      5237 ( 23%)
3      3797 ( 16%)
4      3205 ( 14%)

```

FIGURA 3. Tela do software WEKA versão 3.8.1: algoritmo SimpleKMeans

Para obtenção de regras de associação, foi utilizado o aplicativo Apriori, que gerou um total de dez regras, com confiança variando de 90% a 81%, ou seja, deveras relevante. Dentre as regras obtidas, destacam-se as seguintes:

1. min=Tarde tp\_sfc=Alta 4692 ==> ws\_10m=Adequada 4222 <conf:(0.9)> lift:(1.47) lev:(0.06) [1353] conv:(3.87)
3. min=Tarde 5760 ==> ws\_10m=Adequada 5113 <conf:(0.89)> lift:(1.45) lev:(0.07) [1591] conv:(3.45)
4. tp\_sfc=Alta 5764 ==> ws\_10m=Adequada 5071 <conf:(0.88)> lift:(1.44) lev:(0.07) [1547] conv:(3.23)
8. min=Tarde ws\_10m=Adequada 5113 ==> tp\_sfc=Alta 4222 <conf:(0.83)> lift:(3.3) lev:(0.13) [2942] conv:(4.3)

Há a indicação do período da tarde como aquele que apresenta condições de velocidade média dos ventos adequada, além de relacioná-lo a temperaturas altas,

com confiança elevada e significativa. Também foi indicado que, sempre que a temperatura for classificada como “Alta”, a velocidade dos ventos será adequada à geração de eletricidade.

Estas regras corroboram as informações obtidas a partir da utilização das outras duas técnicas de mineração de dados utilizados na realização deste trabalho.

## **5. Conclusão**

A aplicação do KDD utilizando o software WEKA 3.8.1 sobre a base de dados meteorológicos mostrou-se uma ferramenta útil para determinar o período mais adequado do dia para utilização da planta eólica em questão, garantindo melhor eficiência na geração de energia, aumentando a vida útil e reduzindo as intervenções de manutenção.

Segundo as análises realizadas, pode-se afirmar que o período da Tarde é mais adequado a utilização dos equipamentos de geração eólica de energia, além de demonstrar que temperaturas altas garantem que a velocidade média dos ventos seja adequada, ou seja, passível de geração de energia elétrica.

Como proposta para estudos futuros, sugere-se aumentar o número de dados meteorológicos para que os especialistas tenham um conhecimento mais real das condições climáticas da região nos outros meses do ano, para determinar se há sazonalidade dos ventos ou se há um padrão médio anual.

## Referências

CRESESB. **Atlas do Potencial Eólico brasileiro de 2001**. Disponível em: <<http://www.cresesb.cepel.br>>. Acesso em: 03 de julho de 2017.

DALMAZ, A.; PASSOS, J. C.; COLLE, S. Energia eólica para geração de eletricidade e a importância da previsão. **Revista ABCM–Engenharia**, Florianópolis 13.1, 2008.

DUTRA, R. M. Energia Eólica. In: **Alternativas Energéticas Sustentáveis no Brasil**. Rio de Janeiro: Relume Dumará, 2004.

EMPRESA DE PESQUISA ENERGÉTICA. **Anuário estatístico de energia elétrica 2012**. Rio de Janeiro: EPE, 2012. Disponível em: <[http://www.epe.gov.br/AnuarioEstatisticodeEnergiaEletrica/20120914\\_1.pdf](http://www.epe.gov.br/AnuarioEstatisticodeEnergiaEletrica/20120914_1.pdf)>. Acesso em: 01 de junho de 2017.

FAYYAD, U.; PIATETSKY-SHAPIRO, G.; SMYTH, P. From Data Mining to Knowledge Discovery in Databases. **AI magazine**, p. 37-54, 1996.

INSTITUTO NACIONAL DE METEOROLOGIA. **Monitoramento das Estações Automáticas**. Disponível em: <<http://www.inmet.gov.br/SONABRA/maps/automaticas.php>>. Acesso em: 01 de junho de 2017.

MARKOV, Z.; RUSSELL, I. An Introduction to the WEKA Data Mining System. ANNUAL SIGCSE CONFERENCE ON INNOVATION AND TECHNOLOGY IN COMPUTER SCIENCE EDUCATION, 11.,2006, BOLOGNA, ITÁLIA, 2006. **PROCEEDINGS...** p. 367-368.

PECORA, V. **Implantação de uma unidade demonstrativa de geração de energia elétrica a partir do biogás de tratamento do esgoto residencial da USP–Estudo de Caso**. São Paulo, 2006.

REINHARD, R.; ANTUNES, C. H.; Dias, L. C. Assessing the performance of biogas plants with multi-criteria and data envelopment analysis. **European Journal of Operational Research**, 197, p. 1084–1094, 2009.

SALVADOR, H. G.; CUNHA, A. M., CORRÊA, C. S. Vedalogic: um método de Verificação de Dados Climatológicos Apoiado em Modelos Minerados. **Revista Brasileira de Meteorologia**. São Paulo, v. 24, n. 4, dez. 2009.

SCHUNEIDER, L. F. **Mineração de Dados** - Conceitos. Universidade Federal do Rio Grande do Sul, UFRS, 2002.