

UNIVERSIDADE CANDIDO MENDES – UCAM
PROGRAMA DE PÓS-GRADUAÇÃO EM PESQUISA OPERACIONAL E
INTELIGÊNCIA COMPUTACIONAL
CURSO DE MESTRADO EM PESQUISA OPERACIONAL E INTELIGÊNCIA
COMPUTACIONAL

Luiz Alberto Louzada Hosken

AVALIAÇÃO DO POTENCIAL TÉCNICO DAS ENERGIAS EÓLICA E
SOLAR NO MUNICÍPIO DE CAMPOS DOS GOYTACAZES/RJ
ATRAVÉS DE MINERAÇÃO DE DADOS METEOROLÓGICOS.

CAMPOS DOS GOYTACAZES, RJ.
Março de 2014

UNIVERSIDADE CANDIDO MENDES – UCAM
PROGRAMA DE PÓS-GRADUAÇÃO EM PESQUISA OPERACIONAL E
INTELIGÊNCIA COMPUTACIONAL
CURSO DE MESTRADO EM PESQUISA OPERACIONAL E INTELIGÊNCIA
COMPUTACIONAL

Luiz Alberto Louzada Hosken

AVALIAÇÃO DO POTENCIAL TÉCNICO DAS ENERGIAS EÓLICA E
SOLAR NO MUNICÍPIO DE CAMPOS DOS GOYTACAZES/RJ
ATRAVÉS DE MINERAÇÃO DE DADOS METEOROLÓGICOS.

Dissertação apresentada ao Programa de
Pós-Graduação em Pesquisa Operacional
e Inteligência Computacional, da
Universidade Candido Mendes –
Campos/RJ, para obtenção do grau de
MESTRE EM PESQUISA OPERACIONAL
E INTELIGÊNCIA COMPUTACIONAL.

Orientadora: Prof^a. Geórgia Regina Rodrigues Gomes, D.Sc.

CAMPOS DOS GOYTACAZES, RJ.
Março de 2014

LUIZ ALBERTO LOUZADA HOSKEN

AVALIAÇÃO DO POTENCIAL TÉCNICO DAS ENERGIAS EÓLICA E SOLAR NO
MUNICÍPIO DE CAMPOS DOS GOYTACAZES/RJ ATRAVÉS DE MINERAÇÃO DE
DADOS METEOROLÓGICOS.

Dissertação apresentada ao
Programa de Pós-Graduação em
Pesquisa Operacional e
Inteligência Computacional, da
Universidade Candido Mendes –
Campos/RJ, para obtenção do
grau de MESTRE EM PESQUISA
OPERACIONAL E INTELIGÊNCIA
COMPUTACIONAL.

Aprovada em ____/____/ 2014.

BANCA EXAMINADORA

Prof^a Geórgia Regina Rodrigues Gomes, D.Sc.
Universidade Candido Mendes

Prof. Milton Erthal Júnior, D.Sc.
Universidade Candido Mendes

Prof. Helder Gomes Costa, D.Sc.
Universidade Federal Fluminense

CAMPOS DOS GOYTACAZES, RJ.
2014

À minha querida mãe Acedália,
que sempre me motivou mesmo
não podendo mais contar com a
sua doçura e seu carinho
presente, mas eternamente no
coração.

À minha esposa Marília, que
confiou, com paciência, na vitória
deste trabalho desde sempre.

Aos meus filhos Hugo, Jessica e
Luísa, que entenderam as
ausências provocadas pelos
estudos

AGRADECIMENTOS

À minha orientadora, professora Geórgia, pela amizade, paciência e dedicação na realização deste trabalho.

Aos amigos do mestrado, especialmente Luiz Fernando, que compreenderam e motivaram sempre nos estudos de forma carinhosa.

À minha família, pelos momentos de força e coragem na caminhada que percorremos juntos.

RESUMO

AVALIAÇÃO DO POTENCIAL TÉCNICO DAS ENERGIAS EÓLICA E SOLAR NO MUNICÍPIO DE CAMPOS DOS GOYTACAZES / RJ ATRAVÉS DE MINERAÇÃO DE DADOS METEOROLÓGICOS

Dentre as várias fontes primárias de produção de energia elétrica, as fontes renováveis vem, na atualidade, se destacando por serem limpas e baixo impacto ao meio ambiente. Neste contexto, o Brasil se destaca no potencial eólico e solar, apresentando índices favoráveis a estes modais de produção energética. O país apresenta 12% das residências brasileiras sem acesso à energia elétrica convencional fornecida por concessionária, obtendo a mesma através da queima de combustíveis fósseis em geradores de pequeno porte. Este trabalho tem como objetivo verificar a viabilidade de implementação de energia renovável através das técnicas de mineração de dados aplicadas a um banco de dados de uma estação meteorológica instalada na região norte do estado do Rio de Janeiro, no município de Campos dos Goytacazes. A ferramenta utilizada é o WEKA (Waikato Environment for Knowledge Analysis) da Universidade Waikato da Nova Zelândia. O banco de dados meteorológicos utilizado foi constituído pelas medições de uma estação automática durante sete anos. Os resultados advindos das técnicas de mineração de dados apontaram para a viabilidade de geração de energia para consumo de pequeno porte de residência isolada não atendida por rede elétrica convencional, com arranjo simples de equipamentos eólico e solar. Mas, há de ressaltar a utilização de acumuladores de energia para suprir a demanda no período noturno, visto que a produção híbrida acontece apenas no período diurno. Também destacou que incentivos ao uso destes modais geradores de energia acarretariam na produção maior destes equipamentos e, consequentemente, redução dos custos finais de fabricação.

PALAVRAS-CHAVE: Energia Renovável, Dados Meteorológicos, Sistema Isolado.

ABSTRACT

EVALUATION OF TECHNICAL POTENTIAL POWER WIND AND SOLAR FIELDS IN THE MUNICIPALITY CAMPOS DOS GOYTACAZES / RJ THROUGH METEOROLOGICAL DATA MINING

Among the various sources of primary production of electricity, renewable sources has been highlighted in the news for being clean and low impact to the environment. In this context, Brazil stands out in wind power and solar, presenting rates favorable to this modal of energy production. It is noteworthy, too, that the country has 12% of households without access to conventional electricity supplied by dealer, getting the same through the burning of fossil fuels in generating small. This study aims to determine the feasibility of implementing renewable energy through data mining techniques applied to a significantly large database of a meteorological station installed in the northern region of the state of Rio de Janeiro , in the municipality of Campos dos Goytacazes. The open source tool for data mining WEKA (Waikato Environment for Knowledge Analysis) University of Waikato in New Zealand was used to extract the knowledge of the values of wind speed and insolation useful for producing electricity. The meteorological databank used was comprised of an automatic measurement station for seven years. The results arising from data mining techniques indicated the feasibility of power generation to consumption of small isolated residence not served by the conventional electric network, with simple arrangement of wind and solar equipments. But, there is the use of the accumulator to meet demand during the night, since the hybrid production happens only during daytime. Also noted that incentives to use these generator's energy modals would cause the increased production of these equipments and thus reducing the final cost of manufacture.

KEYWORDS: Renewable Energy, Weather Data, Isolated System.

LISTA DE FIGURAS

FIGURA 1: Residência isolada abastecida por energia eólica.....	15
FIGURA 2: Residência rural atendida pelo Programa “Luz Para Todos” no interior da Bahia.....	19
FIGURA 3: Sistema híbrido de geração de energia: eólico e solar.....	21
FIGURA 4: Velocidade média anual (m/s) dos ventos no estado do Rio de Janeiro-Brasil.....	23
FIGURA 5: Radiação solar global diária – média anual típica (MJm ² .dia) na região Sudeste do Brasil.....	24
FIGURA 6: Etapas do processo KDD.....	27
FIGURA 7: Etapas do <i>Data Mining</i>	28
FIGURA 8: Arquitetura típica de Data Mining.....	30
FIGURA 9: Tela inicial do WEKA Explorer.....	35
FIGURA 10: Tela do WEKA Explorer após carregamento do arquivo ARFF.....	37
FIGURA 11: Estação Automática Meteorológica A607.....	39
FIGURA 12: Insolação diária em Campos dos Goytacazes-RJ.....	40
FIGURA 13: Aero gerador modelo H3.8-2kW, Hummer e Módulo FV modelo M 140P, Martifer.....	42
FIGURA 14: Gráfico VEL_MAN(velocidade dos ventos da manhã).....	43
FIGURA 15: Gráfico VEL_TAR(velocidade dos ventos da tarde).....	44
FIGURA 16: Gráfico VEL_NOI(velocidade dos ventos da noite).....	44
FIGURA 17: Gráfico RAD (radiação solar).....	45

LISTA DE QUADROS

QUADRO 1: Evolução dos Bancos de dados.....	25
QUADRO 2: Tarefas de Mineração de Dados e Algoritmos.....	33
QUADRO 3: Consumo Energético Estimado para Residência Isolada.....	38
QUADRO 4: Potência Elétrica Gerada pelas Médias da velocidade do vento e radiação por período.....	48
QUADRO 5: Resultados da 1ª Tarefa de Agrupamento utilizando Algoritmo <i>SimpleKMeans</i>	49
QUADRO 6: Resultados da 2ª Tarefa de Agrupamento utilizando Algoritmo <i>SimpleKMeans</i>	50
QUADRO 7: Resultados da 3ª Tarefa de Agrupamento utilizando Algoritmo <i>SimpleKMeans</i>	51

LISTA DE SIGLAS

Ah	Ampère-hora
ANEEL	Agência Nacional de Energia Elétrica
ARFF	Attribute-Relation File Format
ASCII	<i>American Standard Code for Information Interchange</i>
BD	Banco de dados
BDMEP	Banco de Dados Meteorológicos para Ensino e Pesquisa
COELBA	Companhia Elétrica da Bahia
CRESESB	Centro de Referência em Energia Solar e Eólica Sérgio de Slavo Brito
CSV	<i>Comma-separated values</i>
EPE	Empresa de Pesquisas Energéticas
FCMC	Fórum Capixaba de Mudanças Climáticas
FV	Fotovoltaico
GMT	<i>Greenwich Mean Time</i>
IBGE	Instituto Brasileiro de Geografia e Estatística
IEA	<i>International Energy Agency</i>
IFPE	Instituto Federal de Pernambuco
INMET	Instituto Nacional de Meteorologia
INPE	Instituto Nacional de Pesquisas Espaciais
KDD	<i>Knowledge Discovery in Databases</i>
kJ/m^2	Quilo Joule por Metro Quadrado
km	Quilômetro
kWh	Quilowatt-hora
m/s	Metro por Segundo
m^2	Metro Quadrado
MD	Mineração de Dados
MJ/m^2	Mega Joule por Metro Quadrado
PNE 2030	Plano Nacional de Energia 2030
PROCEL	Programa Nacional de Conservação de Energia Elétrica
PROINFA	Programa de Incentivo às Fontes Alternativas de Energia Elétrica
SIMPEP	Simpósio de Engenharia de Produção
SGBD	Sistema de Gerenciamento de Banco de Dados

UFPE	Universidade Federal de Pernambuco
UTC	<i>Universal Time Coordinated</i>
W	Watt
WEKA	<i>Waikato Environment for Knowledge Analysis</i>
Wh	Watt-hora
Wp	Watt pico

SUMÁRIO

INTRODUÇÃO	13
1.1 O PROBLEMA	13
1.2 MOTIVAÇÃO	14
1.3 OBJETIVOS	15
1.3.1 Objetivo Geral	15
1.3.2 Objetivos Específicos	16
1.4 ORGANIZAÇÃO DA DISSERTAÇÃO	16
2 FUNDAMENTAÇÃO TEÓRICA	18
2.1 ENERGIAS RENOVÁVEIS SOLAR, EÓLICA E HÍBRIDA	18
2.1.1 Energia Solar	18
2.1.2 Energia Eólica	20
2.1.3 Energia Híbrida Solar/eólica Autônoma	20
2.1.4 Acumuladores de Energia (Baterias)	22
2.2 ATLAS CLIMATOLÓGICO DO ESTADO DO RIO DE JANEIRO	22
2.3 MINERAÇÃO DE DADOS	24
2.3.1 Descoberta do Conhecimento em Bancos de dados (KDD)	25
2.3.2 Atributos	29
2.3.3 Overfitting e Underfitting	29
2.3.4 Arquitetura de um <i>Data Mining</i>	30
2.4 TAREFAS DA MINERAÇÃO DE DADOS	31
2.5 FERRAMENTAS DE MINERAÇÃO DE DADOS	34
2.5.1 Waitako Environment for Knowledge Analysis (WEKA)	34
2.5.2 Arquivos ARFF	36
3 METODOLOGIA	38
3.1 MÉTODOS E MATERIAIS	38
3.1.1 Equipamentos Geradores de Eletricidade	41
3.2 LIMPEZA E ADAPTAÇÃO DOS DADOS	42
3.2.1 Especificação dos Atributos	43
3.2.1.1 Velocidade dos Ventos no Período da Manhã (VEL_MAN)	43
3.2.1.2 Velocidade dos Ventos no Período da Tarde (VEL_TAR)	44
3.2.1.3 Velocidade dos Ventos no Período da Noite (VEL_NOI)	44
3.2.1.4 Radiação Solar (RAD)	45

3.3	ETAPAS DE UTILIZAÇÃO DA FERRAMENTA WEKA	46
3.4	ALGORITMO DE CLUSTERIZAÇÃO	47
4	RESULTADOS E DISCUSSÕES	48
	CONSIDERAÇÕES FINAIS	52
5.1	TRABALHOS FUTUROS	54
	REFERÊNCIAS BIBLIOGRÁFICAS	56

INTRODUÇÃO

1.1 O PROBLEMA

Atualmente, há a nível mundial sérios problemas no âmbito ambiental. Parte destes problemas está ligado a geração de energia, uma vez que a maior parcela da energia consumida no planeta é proveniente de hidrocarbonetos. Mistura resultantes de processos físico-químicos sofridos por matéria orgânica, durante milhões de anos (PEDROTI, 2007).

Os hidrocarbonetos naturais são compostos químicos constituídos apenas por átomos de carbono (C) e de hidrogênio (H), aos quais se podem juntar átomos de oxigênio (O), nitrogênio (N) e enxofre (S) dando origem a diferentes compostos de outros grupos funcionais. Os hidrocarbonetos naturais formam-se a grandes pressões no interior da terra, abaixo de 150 km de profundidade, e são trazidos para zonas de menor pressão através de processos geológicos, onde podem formar acumulações comerciais (petróleo, gás natural, carvão etc.) (PERUZZO et al. 1999).

Segundo a IEA (2010), em um relatório estatístico sobre energia, as matrizes energéticas da maioria dos países são fortemente galgadas em combustíveis fósseis. Os combustíveis fósseis são substâncias de origem mineral, formados pelos compostos de carbono. São originados pela decomposição de materiais orgânicos, porém este processo leva milhões de anos. Logo são considerados recursos naturais não renováveis. Os combustíveis fósseis mais conhecidos são: gasolina, óleo diesel, gás natural e carvão mineral. A queima destes compostos é usada para gerar energia e movimentar motores de máquinas, veículos e até mesmo na geração de energia elétrica (no caso das usinas termoeletricas).

As energias renováveis, a exemplo da eólica e fotovoltaica, são processos de produção limpa de energia, pois possuem capacidade de reposição e servem aos

objetivos energéticos com um impacto ambiental menor que o provocado pelos combustíveis fósseis (MENDES et al., 2011).

As pesquisas de fontes renováveis de energia solar e eólica, na região norte do estado do Rio de Janeiro, através de uma estação meteorológica, poderão acrescentar subsídios na decisão pelas energias alternativas. Estas estações produzem um número elevado de dados necessários para uma melhor amostragem do clima de uma região. Devido a quantidade de dados gerados, a utilização de ferramenta de mineração de dados é recomendada. As ferramentas computacionais de mineração de dados têm a capacidade de extrair conhecimentos a partir de um banco de dados. Estas, utilizadas em dados meteorológicos, podem auxiliar no estudo de viabilidade técnica e econômica, e também ajudar na escolha e localização dos sistemas de renováveis de geração de energia.

1.2 MOTIVAÇÃO

A motivação deste trabalho advém da importância da produção de energia elétrica associada às fontes alternativas renováveis no contexto ambiental no mundo. Energia elétrica na vida moderna é de extrema necessidade, sendo que no Brasil ainda encontra-se 12% dos lares brasileiros sem este bem para uso cotidiano (IBGE, 2011). Este percentual refere-se às zonas rurais especialmente, onde a energia elétrica convencional não está disponível. Segundo Mendes et al. (2012), alternativas tecnológicas podem ser implementadas para minimizar este problema. Energias limpas e renováveis podem atingir lugares afastados a custos acessíveis.

As fontes renováveis, principalmente **eólica e solar**, são abundantes no Brasil, mas estudos da incidência destes parâmetros devem validar a produção energética em escala domiciliar e assim proporcionar a diminuição dos índices de carência de energia elétrica no interior do país.

Tendo como foco ambiental o uso cada vez menor de combustíveis fósseis na produção de energia elétrica, e o Brasil apresentando clima favorável, tanto para ventos como para insolação, a proposta deste trabalho está voltada para a quantificação do uso de energia renovável na produção de energia elétrica.

A Figura 1 retrata a motivação deste trabalho, tendo as regiões não atendidas pela energia elétrica convencional. O foco na pesquisas de energias alternativas

para levar a estes domicílios o desenvolvimento que a eletricidade pode proporcionar.



Figura 1: Residência isolada abastecida por energia eólica

Disponível em: <[http:// www.energiapura/aerogeradores_airbreeze.html](http://www.energiapura/aerogeradores_airbreeze.html)>, acessado em 11/07/2013.

1.3 OBJETIVOS

1.3.1 Objetivo Geral

Este trabalho tem como objetivo, através da mineração de dados, a extração do conhecimento contido numa determinada quantidade de registros de uma estação meteorológica situada no município de Campos dos Goytacazes, região norte do estado do Rio de Janeiro. As informações obtidas proporcionarão verificar a viabilidade da implantação de um sistema renovável na produção de energia elétrica no âmbito doméstico, tendo como referência o consumo de uma residência não atendida por rede elétrica de concessionária. Abordará os modais eólico, solar e sistema híbrido (eólico/solar).

1.3.2 Objetivos Específicos

O estudo atingirá o objetivo geral, pontuando-se aspectos específicos:

- Obter os registros solarimétricos e eólicos de uma estação meteorológica;
- Identificar os parâmetros necessários à produção de energia elétrica por fontes renováveis;
- Realizar o pré-processamento dos dados e colocá-lo no formato na utilização em mineração de dados;
- Utilizar ferramenta computacional de mineração de dados para extrair conhecimento a respeito dos valores de insolação e de ventos que podem contribuir para produção de energia renovável em escala doméstica;
- Propor a discussão da utilização do método para outras localidades com estações meteorológicas no interior do Brasil;

1.4 ORGANIZAÇÃO DA DISSERTAÇÃO

Além do presente capítulo, esta dissertação, seguindo recomendações da Universidade Candido Mendes (UCAM-Campos), está organizada da seguinte forma:

O capítulo 2 apresenta um referencial teórico sobre a produção de energia renovável, especialmente as energias eólica e fotovoltaica. Também é descrita a fundamentação do processo de descoberta de conhecimento em banco de dados, detalhando suas etapas e tarefas que serão aplicadas neste trabalho. Neste capítulo, também são mostrados os mapas climatológicos de ventos e insolação da região objeto deste estudo, no caso, região norte do estado do Rio de Janeiro. Inclui-se ainda neste capítulo a apresentação da ferramenta computacional WEKA e sua aparência gráfica.

O capítulo 3 descreve a metodologia utilizada neste trabalho, demonstrando métodos e materiais. Esta metodologia apresenta o arquivo do banco de dados meteorológicos da estação automática A607, as etapas de limpeza dos dados e de criação dos arquivos necessários para utilização no WEKA. Também são descritos a seleção dos atributos e o algoritmo de clusterização e associação selecionados.

Para o capítulo 4, os resultados obtidos utilizando-se as técnicas de mineração de dados através do WEKA, são apresentados de forma analítica para obtenção do conhecimento sobre os valores minerados de ventos e insolação.

Nas considerações finais são desenvolvidas as interpretações dos resultados para a validação do trabalho, no que tange à produção doméstica de energia renovável através dos valores meteorológicos de ventos e insolação.

Também será neste capítulo apresentada uma proposta para trabalhos futuros de mineração de dados neste segmento de produção de energia, principalmente em áreas não atendidas por concessionários de energia elétrica convencional.

Por fim serão apresentadas as referências bibliográficas utilizadas nesta trabalho.

2 FUNDAMENTAÇÃO TEÓRICA

2.1 ENERGIAS RENOVÁVEIS SOLAR, EÓLICA E HÍBRIDA

2.1.1 Energia Solar

O sol é fonte de energia renovável e o aproveitamento desta energia, tanto como fonte de calor quanto de luz, é uma das alternativas energéticas mais promissoras para o enfrentamento dos desafios do novo milênio (ROBAINA, 2011).

A energia solar é abundante, não polui e não prejudica o ecossistema. A energia solar é uma solução bastante razoável para áreas distantes e ainda não eletrificadas, especialmente num país como o Brasil onde se encontram bons índices de insolação em grande parte do território nacional (DE LOYS, 2012).

A energia solar é importante na preservação do meio ambiente, pois tem muitas vantagens sobre as outras formas de obtenção de energia, como: baixo nível poluente, não influir no efeito estufa, não necessitar de máquinas secundárias para a produção de energia elétrica, mas tem como desvantagem a exigência de altos investimentos para o seu aproveitamento. Para cada 1m² de coletor solar instalado evita-se a inundação de 56 m² de terras férteis, na construção de novas usinas hidrelétricas. Uma parte de energia solar que o Brasil recebe durante o ano poderia dar suprimento de energia equivalente a: 54% do petróleo nacional; 2 vezes a energia obtida com o carvão mineral; 4 vezes a energia gerada no mesmo período por uma usina hidrelétrica (GOULART, 2013).

A Energia Solar Fotovoltaica é a energia da conversão direta da luz em eletricidade (Efeito Fotovoltaico). O efeito fotovoltaico é o aparecimento de uma diferença de potencial nos extremos de uma estrutura de material semicondutor, produzida pela absorção da luz. A célula fotovoltaica é a unidade fundamental do processo de conversão (DE LOYS, 2012).

Atualmente o custo das células solares é um grande desafio para a indústria e o principal obstáculo para a difusão dos sistemas fotovoltaicos em larga escala. A tecnologia fotovoltaica está se tornando cada vez mais competitiva, tanto porque seus custos estão decrescendo, quanto porque a avaliação dos custos das outras formas de geração está se tornando mais real, levando em conta fatores que eram anteriormente ignorados, como a questão dos impactos ambientais.

O atendimento de comunidades isoladas tem impulsionado a busca e o desenvolvimento de fontes renováveis de energia. No Brasil, por exemplo, 12% da população não tem acesso a energia elétrica. Coincidentemente, esta parcela da população vive em regiões onde o atendimento por meio da expansão do sistema elétrico convencional é economicamente inviável. Trata-se de núcleos populacionais esparsos e pouco densos, típicos das regiões Centro-Oeste, Nordeste e Norte (IBGE, 2010). Na Figura 2 é apresentada uma residência rural do interior da Bahia abastecida eletricamente por painel fotovoltaico:



Figura 2: Residência rural atendida pelo “Programa Luz Para Todos” no interior da Bahia
Fonte: Coelba, 2009

No Brasil a geração de energia elétrica por conversão fotovoltaica teve um impulso notável, através de projetos privados e governamentais, atraindo interesse de fabricantes pelo mercado brasileiro. A quantidade de radiação incidente no Brasil é outro fator muito significativo para o aproveitamento da energia solar.

O Brasil possui um grande potencial energético solar fotovoltaico e, na maior parte do território nacional, o INMET possui estações meteorológicas para medição da radiação solar. O aproveitamento racional desta energia no sentido de produzir instalações bem dimensionadas e economicamente viáveis só será possível a partir de informações solarimétricas consistentes da região a ser estudada (INMET, 2013).

2.1.2 Energia Eólica

A energia eólica é a energia obtida pelo movimento do ar (vento). É uma abundante fonte de energia renovável, limpa e na maioria das vezes disponível. Os ventos são gerados pela diferença de temperatura da terra e das águas, das planícies e das montanhas, das regiões equatoriais e dos polos do planeta Terra (ROBAINA, 2011).

A quantidade de energia disponível no vento varia de acordo com as estações do ano e as horas do dia. A topografia e a rugosidade do solo também têm grande influência na distribuição de frequência de ocorrência dos ventos e de sua velocidade em um local. Além disso, a quantidade de energia eólica extraível numa região depende das características de desempenho, altura de operação e espaçamento horizontal dos sistemas de conversão de energia eólica instalados (AMARAL, 2011).

A avaliação do potencial de vento em uma região é o primeiro e fundamental passo para o aproveitamento do recurso eólico como fonte de energia. Portanto para esta avaliação de uma região é necessária a coleta de dados de vento com precisão e qualidade, capaz de fornecer um mapeamento eólico da região. A energia eólica é considerada uma energia limpa, disponível em diversos lugares e em diferentes intensidades. Uma alternativa sustentável para as energias não-renováveis (FCMC, 2013).

2.1.3 Energia Híbrida Solar/eólica Autônoma

A utilização somente de um sistema isolado de energia: fotovoltaico ou eólico, desprende um dimensionamento para as ocasiões em que a captação energética não se faça disponível (valores de irradiação solar ou da velocidade do vento insuficientes para a geração da energia demandada pelo consumidor), de forma a obter-se confiabilidade aceitável do sistema, porém, isso elevaria os custos da instalação inviabilizando muitas vezes seu uso (PIANEZZOLA, 2006).

Um sistema fotovoltaico-eólico autônomo utiliza ambas fontes geradoras de energia: solar e eólica. Por ser autônomo, não possui alimentação com a rede elétrica da concessionária local. Estas fontes geradoras podem estar em operação conjuntamente ou separadas entre si, atribuindo um complemento mútuo quanto à obtenção de geração de energia elétrica. Este sistema torna-se complexo devido à necessidade de um sistema de gerenciamento e controle que facilite a operação e a interligação destas fontes de energia. Na Figura 3 é apresentado um sistema híbrido de geração de energia composto de turbina eólica e painel fotovoltaico



Figura 3: Sistema Híbrido de geração de energia: eólico e solar
Fonte: Disponível em: <<http://www.google.com.br>>, acessado em 2013.

Acumuladores de energia, explicados no tópico a seguir, são fundamentais para armazenamento da energia gerada, a fim de abastecer o consumidor na ausência de radiação solar ou vento, necessárias para a geração de energia.

2.1.4 Acumuladores de Energia (Baterias)

Devido à instabilidade na oferta da radiação solar e vento, é importante a utilização de equipamentos de armazenamento de energia, a fim de suprir, de forma constante, o consumidor final nos instantes em que a demanda é maior que a geração ou quando da oportunidade de manutenção dos elementos geradores, onde seria nula a geração de energia. Nos instantes em que a geração é superior a demanda, o potencial excedente seria transferido eletricamente para o equipamento de armazenamento de energia.

A utilização de um banco de baterias de acumulação química é o método mais difundido para armazenar de energia, devido ao seu custo relativamente baixo, manutenção, confiabilidade e facilidade de operação.

A bateria de acumulação química possui a característica de converter a energia química acumulada em energia elétrica. A unidade básica de uma bateria é denominada de célula, que é constituída por um recipiente onde se encontram dois eletrodos imersos em uma solução eletrolítica, conhecidos como materiais ativos, encarregados da transformação da energia química em elétrica ou vice-versa, segundo estejam em estado de carga ou descarga. A diferença de potencial nominal de uma célula é de 2 Volts, de tal forma que, para a obtenção de tensões maiores associa-se células em série, formando o que denominamos bateria, comumente constituída por seis células em série, a qual nos fornece uma tensão de 12 Volts (CHAVES, 2002).

2.2 ATLAS CLIMATOLÓGICO DO ESTADO DO RIO DE JANEIRO

De acordo com estudos realizados pela Empresa de Pesquisas Energéticas (EPE) no Plano Nacional de Energia 2030 (PNE 2030), Os dados relativos ao comportamento dos ventos, que auxiliam na determinação do potencial eólico da região, são relativos à intensidade da velocidade e à direção do vento. Para obter estes dados, é necessário também analisar os fatores que influenciam o regime dos

quando a radiação é varia de 8 a 18 MJ/m². Na Figura 5, verifica-se a distribuição de radiação solar ao longo da região Sudeste.

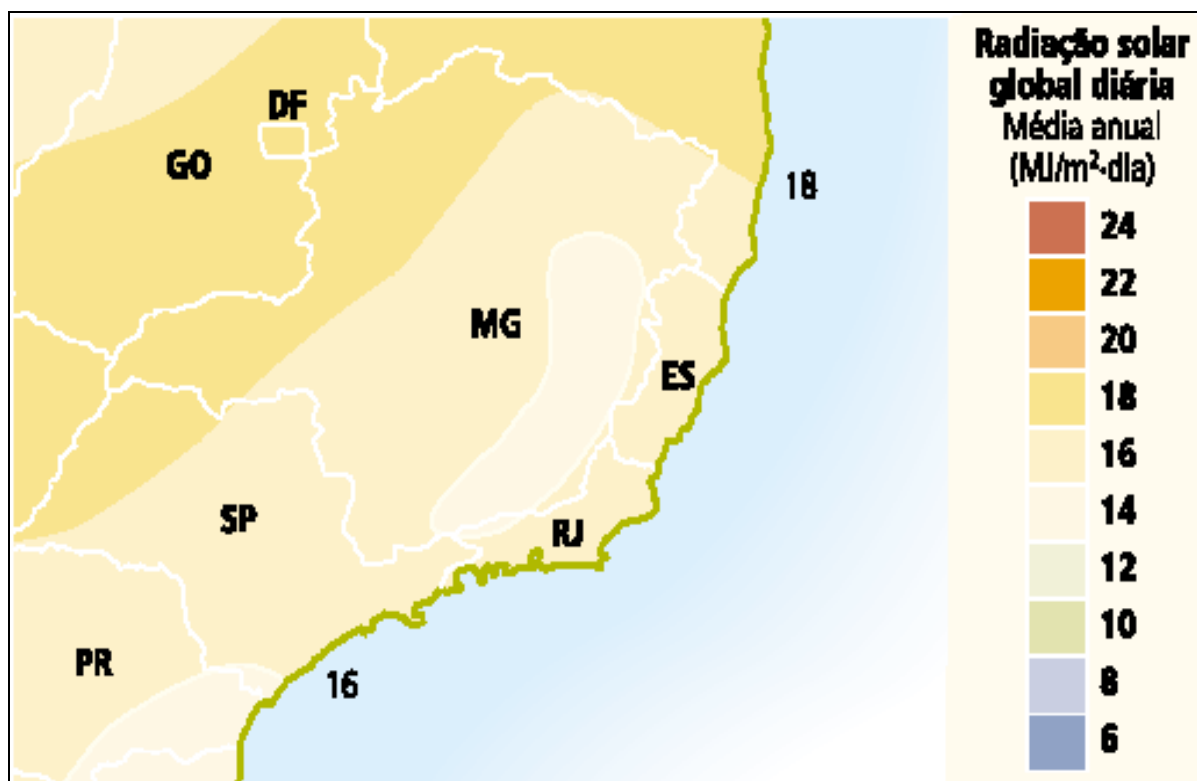


Figura 5 – Radiação solar global diária - média anual típica (MJ/m².dia) na região Sudeste do Brasil.
Fonte: Adaptado do Atlas da energia Elétrica no Brasil - 2^a edição – 2005 (ANEEL).

Nesta imagem, percebe-se que a radiação média anual do Estado do Rio de Janeiro é de 16 MJ/ m².dia, ou seja, radiação solar é favorável à geração de energia solar.

2.3 MINERAÇÃO DE DADOS

A Mineração de Dados (*Data Mining*) consiste de um conjunto de técnicas reunidas da estrutura de Banco de Dados, Estatística e da Inteligência Artificial com o objetivo de descobrir conhecimento novo, útil, relevante e não-trivial que, porventura, esteja escondido em uma grande massa de dados (GOLDSCHMIDT e PASSOS, 2005). São muitos os exemplos de sua utilização em áreas como o Marketing, Economia, Engenharia e até Medicina (CARVALHO, 2001). Um conjunto de dados tratados sob técnicas de Mineração de Dados resulta em informação (ou conjunto de informações) útil. Descobrir padrões e tendências escondidos em

grandes massas de dados não é trivial. Em Mineração de Dados esse processo envolve o uso de diversas tarefas (FAYYAD et al., 1996).

Segundo Batista (2003), a área da inteligência artificial que utiliza técnicas e ferramentas para a extração de conhecimento a partir de um banco de dados é conhecida como *Knowledge Discovery in Databases* (KDD).

Pode-se inferir a evolução dos Bancos de dados apresentado no Quadro 1.

Quadro1: Evolução dos Bancos de Dados

Evolução	Perguntas	Tecnologia disponível	Características
Coleção de dados 1960	"Qual o rendimento total nos últimos 5 anos?"	Computadores, Fitas, discos	Retrospectiva, Dados estáticos como resposta
Acessos aos dados 1980	"Qual o rendimento no Brasil no último verão?"	RDBMS, SQL, ODBC	Retrospectiva, dados dinâmicos a nível de registros como resposta
Data Warehousing & suporte a decisão 1990	"Qual o rendimento no Brasil no último verão, do sul ao nordeste?"	Processamento analítico on-line, banco de dados multidirecionais, data warehousing	Retrospectiva, dados dinâmicos em múltiplos níveis como resposta
Data Mining dias atuais	"Por que alguns produtos são mais vendidos na região sul no verão?"	Algoritmos avançados, computadores multiprocessadores, BD grandes e poderosos	Prospectivo, informações (perspectivas) como resposta

Fonte: Goulart Junior, Fernando; Fidalgo, Robson; UFPE (1998), adaptado

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

2.3.1 Descoberta do Conhecimento de Bancos de Dados (KDD)

Atualmente, com o crescimento expressivo das informações, o mundo se enquadra na “era da informação”. O conhecimento detalhado das informações impulsiona cada vez mais a humanidade em eficiência nos seus afazeres. A sociedade científica, empresas e instituições necessitam de dados apurados extraídos de uma grande quantidade de registros para desenvolverem suas atividades (TAKECIAN et al., 2012).

O avanço tecnológico na área de informática com o desenvolvimento dos sistemas operacionais, a redução física dos componentes de armazenamento e processadores cada vez mais potentes gerou e gera constantemente grandes bancos de dados. Isso proporcionou uma ramificação dentro da informática no final dos anos 80 chamada Mineração de Dados, cujo objetivo principal era descobrir informações interessantes nas bases de dados volumosas. A definição de mineração de dados para Han e Kamber é “extração ou mineração do conhecimento em um grande amontoado de dados”. (HAN & KAMBER, 2006, p.05).

Ainda de acordo com Quoriam et al. (2001)

a capacidade de armazenamento em banco de dados, assim como sua utilização, vem crescendo na mesma proporção dos avanços em novas tecnologias de informação e comunicação. A atividade de extrair informações relevantes, por conseguinte, está se tornando bastante complexa. Este processo de ‘garimpagem’ é chamado de *Knowledge Discovery in Databases* – KDD (p.20).

No começo dos anos 70, a metodologia KDD aplicava recursos de estatística, reconhecimento de padrões, máquinas de aprendizado e métodos de visualização. Esta técnica obtinha formas de regras para dar suporte de análise de dados e descobrir princípios que estavam embutidos em dados (FAYYAD et al.,1996).

A descoberta do conhecimento em banco de dados é um processo de etapas na mineração de dados que permite ao usuário obter informações relevantes dentro de uma gama enorme de dados. Mesmo assim, segundo Dias (2001), nem sempre a mineração de dados pode se encaixar na solução de um problema focado, já uma suposição antevista pode não advir com esta mineração ou de forma simples os dados não favorecem à solução proposta pelo observador.

Segundo Liebstein (2001), o processo de KDD destina-se encontrar e interpretar padrões nos dados através da repetição de algoritmos e da análise de resultados. A Figura 6 representa as etapas do processo de KDD.

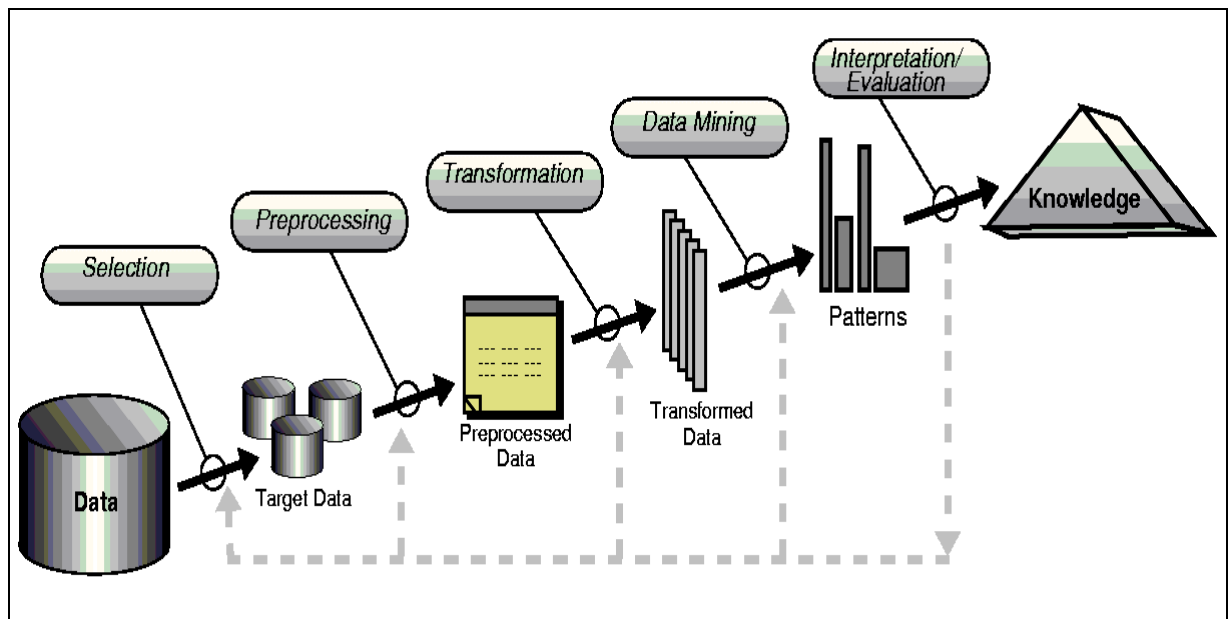


Figura 6: Etapas do processo KDD.
 Fonte: Arakaki, E.M. Notas de Aula do IFPE, adaptado.

Segundo Han e Kamber (2006, p.07) podemos descrever as fases do KDD, conforme Figura 6:

- Seleção da base de dados: fase onde são selecionados ou segmentados os dados de acordo com a informação a ser obtida;
- Preparação (pré-processamento): limpeza dos dados para remoção de informações julgadas desnecessárias;
- Transformação: os dados são transformados em formatos a serem utilizados pelas ferramentas computacionais;
- Mineração de Dados: é a fase de aplicação dos softwares para extração comportamental da informação;
- Padrão de Avaliação: utilização de definição de fatos, medidas padronizadas e relacionamentos entre dados;
- Aproveitamento do conhecimento: interpretação da informação nova relevante que darão suporte às tomadas de decisões.

Segundo Quoniam et al. (2001), também as etapas do *Data mining* podem ser ilustradas pela Figura 7:

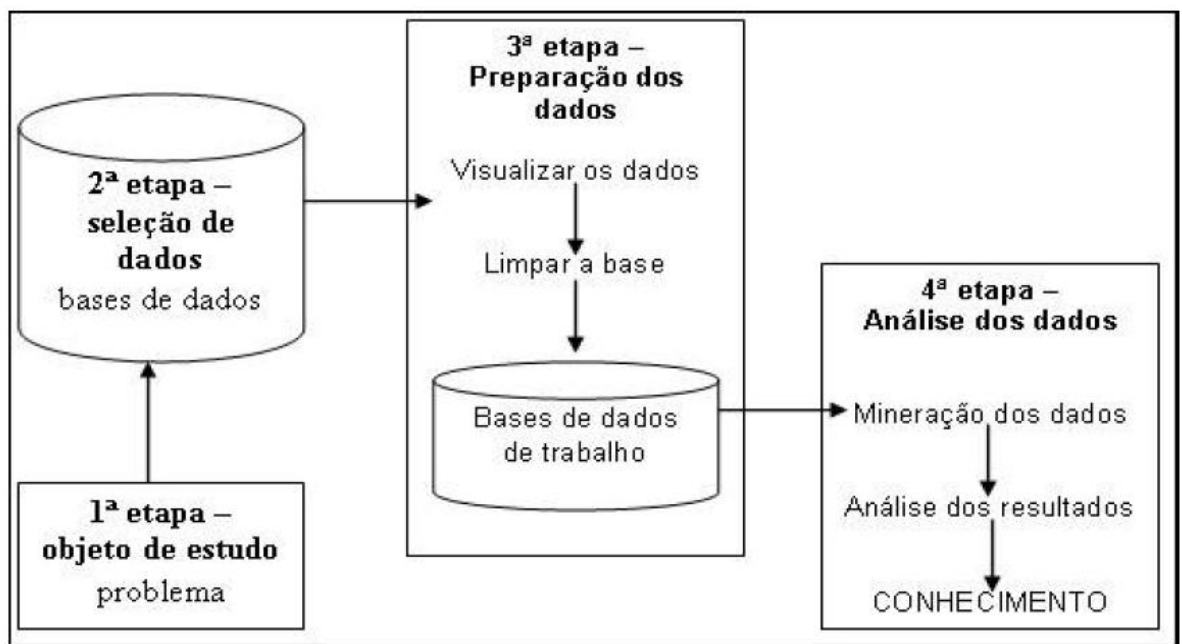


Figura 7: Etapas do Data Mining (QUORIAM et al. 2001, adaptado)

De acordo com a Figura 7, as etapas do Data Mining são as seguintes:

1ª Etapa – Objeto de Estudo - é a definição do problema de forma objetiva a ser explorado pelas técnicas de mineração de dados;

2ª Etapa – Seleção dos dados - envolve a escolha, dentro de um banco de dados, os itens de interesse a serem manipulados que foram apontados pela primeira etapa, principalmente sendo este banco composto por uma grande quantidade de registros;

3ª Etapa – Preparação dos dados - é a etapa onde é requerida uma preparação dos dados a serem minerados, incluindo nesta etapa a limpeza e adequação dos mesmos, para a utilização dos softwares de *Data Mining*;

4ª Etapa – Análise dos dados - nesta etapa são avaliados os resultados obtidos que devem considerar:

- a) a informação nova seja apresentada de forma clara para ser explorada;
- b) a relevância desta nova informação como fator modificador e/ou qualificador;
- c) a análise do conjunto de dados obtidos na mineração por especialistas para o melhor aproveitamento e tomada de decisões.

2.3.2 Atributos

A definição de atributo, segundo Tan et al. (2009): “um atributo é uma propriedade ou característica de um objeto que pode variar, seja de um objeto para outro ou de tempo para outro.

Por exemplo, a cor dos olhos varia de pessoa para pessoa, enquanto que a temperatura de um objeto varia com o tempo”.

Pode-se definir dois grupos de atributos: categorizados (qualitativos) e numéricos (quantitativos). Os categorizados dividem-se em nominais e ordinais, enquanto os numéricos se dividem em intervalar e proporcional. Atributos qualitativos não possuem a maioria das propriedades dos números. Já os atributos quantitativos são representados por números e possuem a maioria das propriedades dos mesmos.

Outra distinção entre atributos é pelo número de valores que eles podem receber:

- Discretos: um atributo discreto possui um conjunto de valores finito ou contavelmente infinito. Podem ser categorizados ou numéricos;
- Contínuos: cujos valores são números do tipo real. Exemplos incluem temperatura, altura ou peso;
- Assimétricos: apenas se a presença – um valor de atributo diferente de zero – é considerado importante. (TAN et al., 2009).

2.3.3 Overfitting e Underfitting

Como etapa importante, mostrada na Figura 7, a limpeza dos dados é necessária para verificar a potencialidade, em termos de energias renováveis, da cidade de Campos dos Goytacazes – RJ no que tange tais energias. Segundo GOLDSCHMIDT et al. (2005), é importante perceber que a qualidade dos dados possui grande influência na qualidade dos modelos de conhecimento a serem abstraídos a partir destes dados. Quanto pior for a qualidade dos dados informados ao processo de KDD, pior será a qualidade dos modelos de conhecimento gerados (GIGO – *Garbage in, Garbage out*).

A percepção sobre como os dados devem ser pré processados a fim de melhorar a qualidade dos mesmos e, conseqüentemente, dos resultados da mineração constitui-se em uma questão de grande relevância no processo do KDD.

Tem-se o fenômeno do *Underfitting* quando devido a uma amostra muito pouco representativa, elementos de grande participação/importância são desconsiderados ou tem menor peso que o ideal fazendo assim que o classificador cubra uma extensão menor que a adequada. O fenômeno do *Overfitting* ocorre devido à consideração excessiva de um ruído (dados em branco ou inconsistente) na amostra ou de simplesmente uma amostra anômala, acarretando com que o classificador decida considerar uma extensão maior de amostras que a ideal (GOLDSCHMIDT e PASSOS, 2005).

2.3.4 Arquitetura Típica de um *Data Mining*

A arquitetura típica de um *Data Mining*, ainda segundo Han e Kamber (2006), é apresentada na Figura 8:

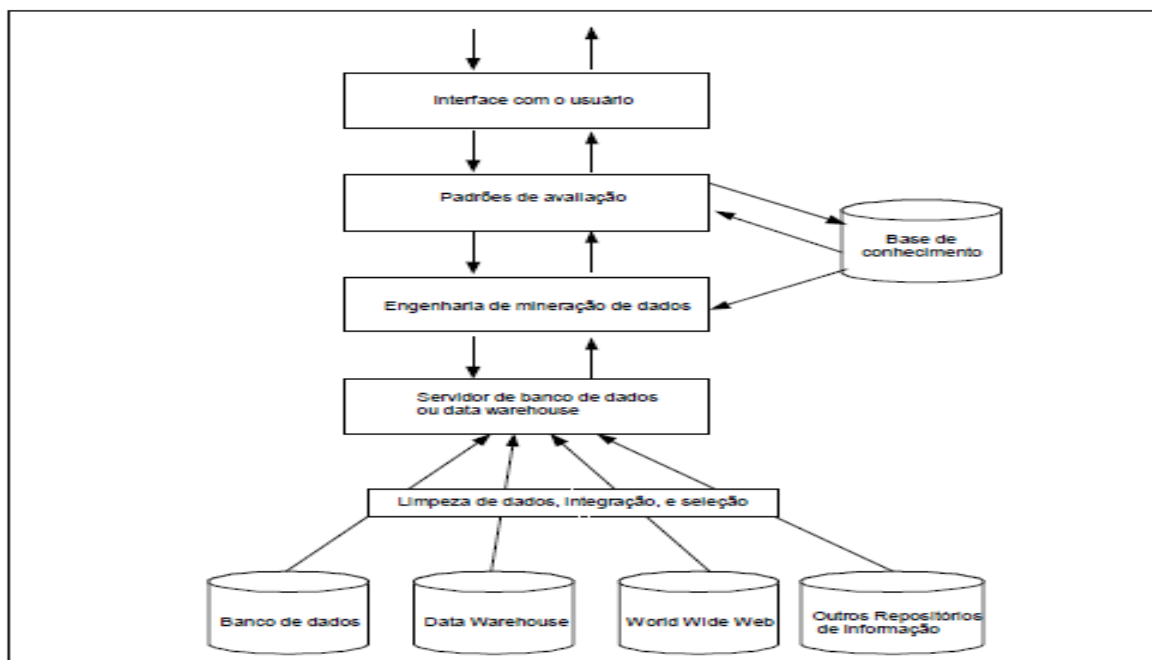


Figura 8: Arquitetura Típica de Data Mining.

Fonte: HAN & KAMBER, 2006, p.08, adaptado

A descrição, segundo Han e Kamber (2006) desta Arquitetura é feita do seguinte modo:

→ *Banco de dados, Data Warehouse, World Wide Web e outros repositórios de informação*: é um conjunto de banco de dados, Data Warehouse, planilhas e outros repositórios de informação. A limpeza e preparação dos dados deverão ser feitos neste arquivos;

→ *Servidor de banco de dados ou Data Warehouse*: é responsável por gerenciar os dados, com base na solicitação do usuário;

→ *Engenharia de Mineração de Dados*: é o conjunto de módulos funcionais para tarefas tais como: caracterização, associação, análise de correlação, classificação, predição, análise de cluster, análise de *outliers* e análise de evolução;

→ *Base de conhecimento*: é o domínio do conhecimento que é usado para guiar as pesquisas ou avaliar os interesses dos padrões resultantes;

→ *Padrões de Avaliação*: este módulo emprega medidas de interesse e interação com os módulos de mineração de dados;

→ *Interface com usuário*: faz a comunicação entre usuários e a execução do processo de mineração de dados, onde é permitido ao usuário interagir com este processo, e especificar uma consulta de mineração de dados ou tarefa.

Empregam-se as técnicas de mineração de dados realizando uma tarefa nesta arquitetura típica de *data mining*, podendo encontrar informações de relevância para determinada tarefa, na necessidade desejada.

2.4 TAREFAS DA MINERAÇÃO DE DADOS

Tarefa, no contexto da Mineração de dados, é um tipo de problema de descoberta de conhecimento a ser solucionado. Destacam-se as tarefas de classificação, agrupamento, estimativas, sumarização e associação, que podem ser resolvidas de forma individual ou combinadas. A escolha da tarefa mais adequada

depende da natureza da aplicação que se pretende desenvolver (GOLDSCHMIDT e PASSOS, 2005).

Nesta seção, serão apresentados as principais tarefas utilizadas em mineração de dados. Entre os principais tem-se: Classificação e Regressão (tarefas preditivas), Agrupamento (clustering), Regras de Associação e Sumarização (tarefas descritivas).

Classificação – consiste em relacionar determinado item dentro de um banco de dados em classes, classificando-o por vários atributos pré-definidos. Por exemplo, tendo um banco de dados de multas de trânsito, em que cada multa contém data, hora, placa do veículo, marca do veículo, código da infração, tipo da infração, entre outras, classificar cada multa por tipo de veículo: de passeio, de carga ou motocicleta;

Regressão – assemelha-se à classificação, porém são analisados dados numéricos, objetivando principalmente obter uma preditiva de dados históricos num banco de dados. Busca uma função que relaciona os dados em valores reais.

Clustering (Regras de agrupamento) – os clusters são definidos através de similaridade ou frequência. Atinge uma associação de um item a uma classe ou várias classes de categorias. Difere da classificação, pois nessa as classes são pré-definidas. Ao analisar os clusters verifica-se a ocorrência de diferentes grupos num banco de dados, onde existindo tal similaridade, identifica-se quais são.

Segundo Groth (1998), a clusterização na mineração de dados normalmente é um exemplo de aprendizado de máquina não supervisionado ou indireto, com a finalidade de formar grupos similares dentro do banco de dados ou reconhecer suas exceções.

Um cluster define-se como um conjunto de dados agrupados pela similaridade e pela relação entre tais dados, sendo a clusterização definida como a tarefa desta segmentação de uma população heterogênea em uma determinada quantidade de subgrupos (clusters) mais homogêneos possíveis, de acordo com algum parâmetro (DILLY, 1995).

Regras de Associação – tem como objetivo obter, dentro de um banco de dados, itens que ocorram de forma conjunta, formando regras para que o analista possa utilizar conhecimento novo no projeto estudado. Regras de associação não são muito diferentes de regras de classificação, exceto que elas podem prever qualquer atributo, e não apenas a classe. Isso lhes dá a liberdade de prever

combinações de atributos também. Além disso, regras de associação não se destinam ser usadas como um conjunto, como as regras de classificação são (WITTEN et al., 2005).

Sumarização – identificação e indicação de características comuns num conjunto de dados (GOLDSCHMIDT & PASSOS, 2005).

O Quadro 2 apresenta os métodos acima descritos relacionados com os principais algoritmos (BOENTE, GOLDSCHMIDT & ESTRELA, 2006):

Quadro 2:Tarefas de Mineração de Dados e Algoritmos

Métodos	Algoritmos
Classificação	Redes Neurais, Algoritmos Genéticos
Regressão	Estatística, Redes Neurais
Clustering	K-Means, K-Modes, K-Medoids
Regras de Associação	Apriori, J-48, Partition
Sumarização	Lógica Indutiva, Algoritmos Genéticos

Fonte: Boente et al. (2006), adaptado.

Segundo Fayyad et al.(1996), as duas metas primárias da mineração de dados, na prática, consistem na previsão e descrição:

- Previsão envolve usar algumas variáveis ou campos no banco de dados para prever valores desconhecidos ou futuros de outras variáveis de interesse;
- Descrição se concentra em encontrar padrões interpretáveis pelo homem descrevendo os dados.

Embora a fronteira entre previsão e descrição não seja grande (alguns modelos preditivos podem ser descritivos, na medida em que são compreensíveis, e vice-versa), em geral, a distinção é necessária para a compreensão do objetivo a ser descoberto.

A importância relativa de previsão e descrição nas aplicações de *data mining* pode variar consideravelmente. As tarefas preditivas e descritivas podem ser obtidas utilizando uma variedade de métodos de mineração de dados.

2.5 FERRAMENTAS DE MINERAÇÃO DE DADOS

Atualmente existem diversas ferramentas capazes de lidar com os diversos algoritmos de mineração de dados. Muitas delas são genéricas da inteligência artificial ou da comunidade de Estatística. Tais ferramentas operam separadamente da fonte de dados, requerendo uma quantidade significativa de tempo gasto com exportação e importação de dados, pré e pós-processamento e transformações de dados. Entretanto, a conexão entre a ferramenta de descoberta de conhecimentos e a base de dados analisados, utilizando o suporte do SGBD (Sistema de Gerenciamento de Banco de Dados) existente, é desejável. Para Goebel e Gruenwald (1999 apud ESCOVAR, 2004), as características a serem consideradas na escolha de uma ferramenta de descoberta de conhecimento devem ser:

- A habilidade de acesso a uma variedade de fontes de dados, de uma forma on-line e off-line;
- A capacidade de incluir modelos de dados orientados a objetos ou modelos não padronizados (tal como multimídia, espacial ou temporal);
- A capacidade de processamentos com relação ao número máximo de tabelas/tuplas/atributos;
- A capacidade de processamento com relação ao tamanho do banco de dados;
- A variedade do tipo de atributos que a ferramenta pode manipular; e o tipo de linguagem da consulta.

2.5.1 Waitako Environment for Knowledge Analysis (WEKA)

A ferramenta WEKA (Waikato Environment for Knowledge Analysis), tem sido utilizada na realização da etapa de Mineração de Dados, por ser de domínio público e prover um conjunto de algoritmos que implementam diversas técnicas para resolver problemas reais de Mineração de Dados. Esta ferramenta foi implementada na linguagem Java e desenvolvida no meio acadêmico da Universidade de Waikato, na Nova Zelândia, em 1999. Suas principais características são herdadas do fato de ser uma ferramenta desenvolvida em Java, uma linguagem multi-plataforma orientada a objetos. A portabilidade da linguagem Java permite ao WEKA ser

executado em diversas plataformas diferentes, e sua orientação a objetos produz vantagens como modularidade, polimorfismo, encapsulamento, reutilização de código entre outras. O WEKA é composto por dois pacotes que podem ser embutidos em outros programas escritos em Java, permitindo que um desenvolvedor possa criar seu próprio ambiente de Mineração de Dados. O primeiro pacote possui interfaces para a manipulação interativa de algoritmos de Mineração de Dados e o segundo possui classes Java responsáveis pelo “encapsulamento” desses algoritmos. A Figura 9 apresenta uma das principais interfaces da ferramenta WEKA:

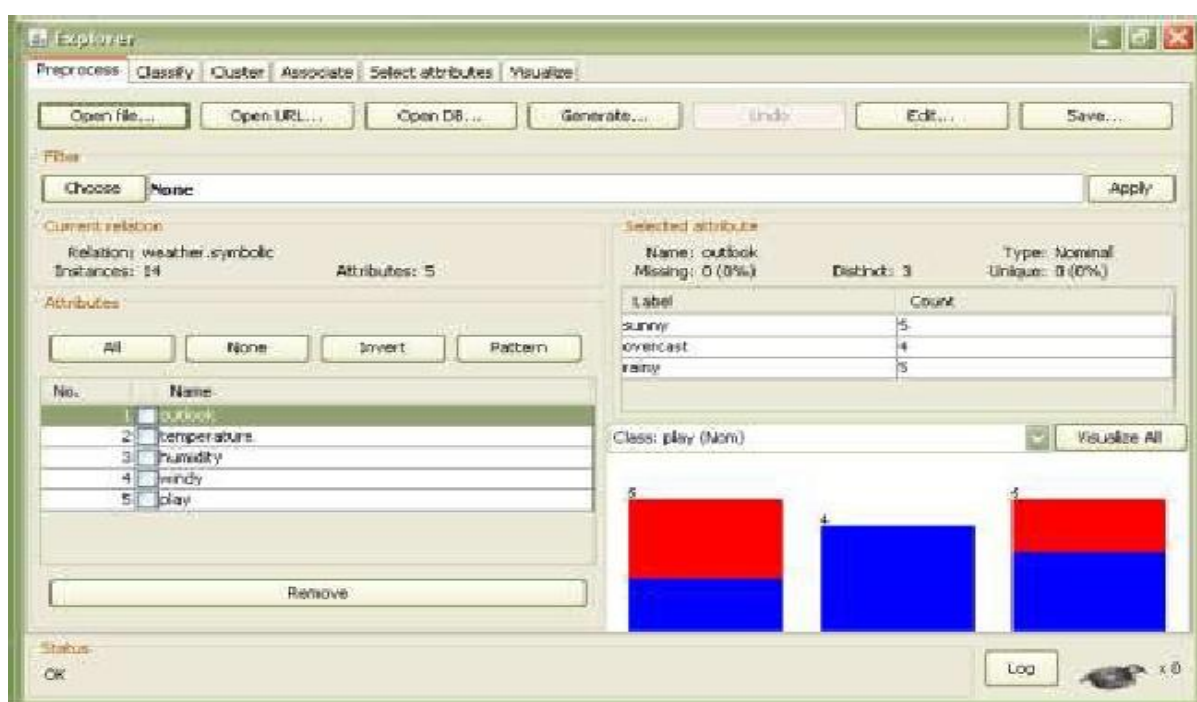


Figura 9: Tela inicial do WEKA Explorer.
 Fonte: Universidade de Waikato – Nova Zelândia

A interface do WEKA apresenta painéis interativo com o usuário, descrito da seguinte forma:

Preprocess - tem facilidade para a importação de dados de um arquivo, extensão ARFF, e pré-processamento de dados usando um algoritmo de filtragem;

Classify - permite ao usuário aplicar algoritmos de classificação e regressão (indiscriminadamente chamados classificadores em Weka) para o conjunto de dados resultante, para estimar a precisão do resultado do modelo preditivo;

Cluster - permite acesso às técnicas de agrupamento em Weka, por exemplo, o algoritmo *SimpleKMeans*;

Associate - fornece ao usuário acesso às regras de associação que tentam identificar as interrelações importantes entre os atributos dos dados;

Select attributes – permite acessar algoritmos para identificar os atributos mais preditivos em um conjunto de dados;

Visualize - mostra gráficos de dispersão da matriz, onde gráficos de dispersão individuais podem ser selecionados e ampliada.

2.5.2 Arquivos ARFF

ARFF é um formato desenvolvido na *University of Waikato* para ser utilizado no projeto *Weka Machine Learning Project*. São arquivos de texto na codificação ASCII que descreve as relações e seus atributos (ATTRIBUTE-RELATION, 2008). Na primeira parte do arquivo, conhecida como *header*, são declaradas as relações e os atributos, onde a declaração das relações segue o formato: @relation <nome-da-relação>, e a declaração dos atributos segue o formato: @attribute <nome-do-atributo> <tipo>.

Os tipos podem ser:

- **numeric**: usado para números reais e inteiros;
- **<nominal-specification>**: especifica uma lista pré-definida de valores possíveis separados por vírgula. Por exemplo: {amarelo, vermelho, azul};
- **string**: tipo que contém valores textuais;
- **date**: usado para datas.

Inicialmente, carrega-se o arquivo *labsql.ARFF* a partir do modo Explorer, que é a principal interface do WEKA para a mineração de dados, sendo mostrada a tela como observado na Figura 10:

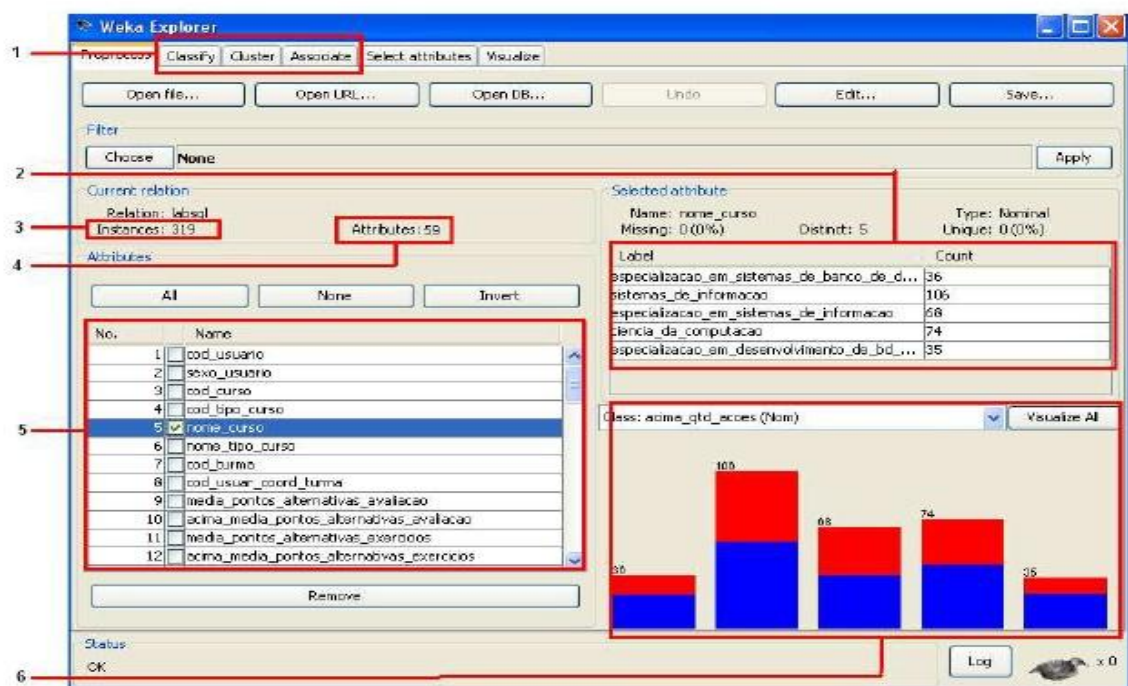


Figura 10: Tela do WEKA Explorer após carregamento do arquivo ARFF

Fonte: Universidade de Waikato – Nova Zelândia

A partir da Figura 10, pode-se observar:

1. As tarefas de MD disponíveis: classificação, Clusterização e associação.
2. O número de instâncias de cada atributo presente na base de dados.
3. A quantidade de registros (instâncias) trabalhados.
4. A quantidade de atributos presentes na base de dados.
5. A listagem de todos os atributos presentes na base de dados.
6. Um histograma com informações sobre a distribuição dos exemplos para o atributo selecionado. Após carregar o arquivo, selecionou-se a aba Classify (classificação), tela onde é realizada a escolha do algoritmo de classificação desejado.

3 METODOLOGIA

3.1 MATERIAIS E MÉTODO

Neste trabalho, utilizou-se de dados do Banco de Dados Meteorológicos para Ensino e Pesquisa (BDMEP), do Instituto Nacional de Meteorologia (INMET), coletados por 07 (sete) anos (período entre 25/09/2006 a 31/08/2013) pela estação meteorológica automática, código **A607** (latitude: 21.4147° S; longitude: 41.3441° W e altitude: 25 metros), que entrou em operação em 22/09/2006. A referida estação está situada em uma área próxima à Rodovia BR 356 (Campos-Itaperuna) na cidade de Campos dos Goytacazes - RJ, região Norte Fluminense.

Como referência de uma residência isolada, não abastecida por rede elétrica convencional, o Quadro 3 estima o consumo energético desta moradia, levando em conta hipoteticamente os equipamentos mínimos de uso doméstico. As potências elétricas dos equipamentos envolvidos foram obtidas pela tabela de eficiência energética do Programa Nacional de Conservação de Energia Elétrica (PROCEL).

Quadro 3: Consumo Energético Estimado para Residência Isolada

Cargas(*)	Potência da carga (W)	Diurno Horas por dia(*)	Consumo diurno (Wh)	Noturno Horas por dia(*)	Consumo noturno (Wh)
(*) estimativa					
Lâmpadas (5)	66	0	0	5,0	330,0
TV (colorida) 14"	100	0	0	5,0	500,0
Receptor Parabólico	30	0	0	5,0	150,0
Rádio Relógio	5	12,0	60	12,0	60,0
Geladeira 1 porta	35	12,0	420	12,0	420,0
		Total Diurno	480,0	Total Noturno	1460,0

Fonte: Procel - Procel Info – Adaptado pelo autor

A Figura 11 destaca a estação meteorológica deste trabalho. Interessante observar que a mesma é alimentada eletricamente por um painel fotovoltaico.



Figura 11: Estação Automática Meteorológica A607
Fonte: INMET- SONABRA

O BDMEP correspondeu aos registros de 2532 dias ao longo do tempo de apuração dos dados, lançados em uma planilha Excel da Microsoft Inc. Cada dia obtinha-se 24 registros, correspondentes a cada hora. Cada registro continha 19 itens, a saber: **Data**, **Hora UTC**, **Temperatura** (°C) [Instantânea, Máxima e Mínima], **Umidade** (%) [Instantânea, Máxima e Mínima], **Ponto de Orvalho** (°C) [Instantânea, Máxima e Mínima], **Pressão** (hPa) [Instantânea, Máxima e Mínima], **Vento** (m/s) [Velocidade, Direção e Rajada], **Radiação** (kJm²), **Chuva** (mm). Portanto, o banco de dados meteorológicos em questão foi formado por 1.154.592 dados apurados. Neste caso, para aplicação da ferramenta computacional WEKA, esta quantidade de dados a serem minerados é significativa.

Para produção de energia elétrica através de fontes renováveis, focadas neste trabalho, em primeiro lugar propõe-se identificar quais parâmetros serão observados em uma quantidade expressiva de dados. Neste caso, as fontes renováveis por serem a solar e a eólica, tiveram a 'radiação solar' e a 'velocidade

dos ventos' como dados principais observados para análise na produção energética. Mas, como dado secundário, a 'hora' do registro também foi analisada em função da energia solar, que depende exclusivamente da radiação do sol para geração elétrica através das placas fotovoltaicas. Conforme apresentado na Figura 12, relacionam-se hora e radiação solar, os registros de radiação estão durante a incidência da luz do sol como era esperado. A crista desta curva está por volta do horário das 16:00 horas (Hora UTC), o que significa em torno das 12:00 horas em Campos dos Goytacazes – RJ. Isto está de acordo com a literatura sobre incidências de radiação solar no hemisfério sul (TIBA et al., 2000). Na seção 3.2.1, Especificação dos Atributos, será feita uma abordagem de como os atributos foram agrupados em quatro grupos para que demonstrassem os valores médios de cada grupo. A Figura 12 apresenta a curva característica de insolação diária na localidade de Campos dos Goytacazes-RJ, obtida através do WEKA após a fase do pré-processamento do arquivo de dados no *explorer*:

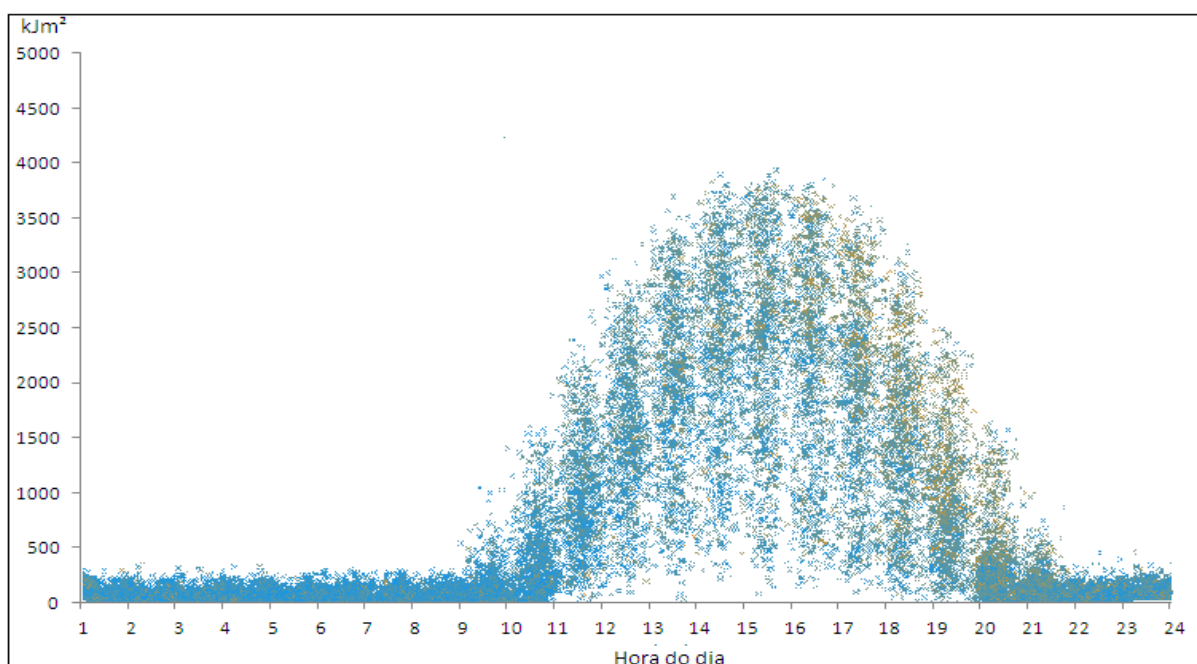


Figura 12: Insolação diária em Campos dos Goytacazes-RJ

Fonte: WEKA - Adaptado pelo autor

Já os registros de 'Data' tiveram a função de determinar em qual período do ano foram coletados.

O item registrado no banco de dados meteorológicos (Hora UTC) é **Tempo Universal Coordenado** (em inglês: *Coordinated Universal Time*), ou **UTC** (acrônimo de *Universal Time Coordinated*), também conhecido como tempo civil, é o fuso

horário de referência a partir do qual se calculam todas as outras zonas horárias do mundo. Corresponde à hora de inverno de Portugal Continental e Arquipélago da Madeira e à hora de verão do Arquipélago dos Açores. É o sucessor do *Tempo Médio de Greenwich (Greenwich Mean Time)*, cuja sigla é *GMT*. A nova denominação foi cunhada para eliminar a inclusão de uma localização específica num padrão internacional, assim como para basear a medida do tempo nos padrões atômicos, mais do que nos celestes. Como observação, deve-se considerar que para o fuso horário brasileiro sem horário de verão, onde a estação está localizada, há uma diferença de 04(quatro) horas.

Em INMET (2013), é disponibilizado um ambiente amigável e interativo ao usuário, onde se tem acesso às estações meteorológicas automáticas e convencionais posicionadas em todo território nacional.

3.1.1 Equipamentos Geradores de Eletricidade

Os parâmetros de velocidade dos ventos e radiação solar foram relacionados com os valores mínimos para produção de energia elétrica dos seguintes equipamentos geradores de eletricidade: **turbina eólica e placa fotovoltaica**, respectivamente. Para a velocidades dos ventos, utilizou-se as características técnicas de uma Turbina Eólica de pequeno porte, modelo H3.8-2kW, com taxa de potência de 250 W, a partir 3 m/s de velocidade do vento(valor mínimo de produção), marca Hummer. A produção de energia fotovoltaica foi estabelecida com relação às características técnicas de um Módulo Solar FV, modelo M Prime M140P, com taxa de potência de 140 Wp para incidência constante maior que 1000 kJ/m² de radiação solar difusa a 20° de declinação, marca Martifer, apresentados na Figura 13. Neste trabalho utilizou-se a configuração mínima de 01 aerogerador e 01 painel fotovoltaico.

Para os acumuladores de energia foram relacionados com os dados técnicos de um conjunto de Baterias Estacionárias Ácido-Chumbo, modelo DF300, capacidade para 30Ah, marca Freedom. A configuração do conjunto de acumuladores dependeria do arranjo necessário para uma carga estipulada de consumo.



Figura 13: Aero gerador modelo H3.8-2kW, Hummer e Módulo FV modelo M 140P, Martifer
Fonte: Brasil Wind Service, Solar Brasil Tecnologia & Energia Fotovoltaica Ltda

A escolha dos equipamentos acima deveu-se ao fato dos mesmos atenderem às instalações de pequeno porte em residências isoladas, geralmente rurais, não atendidas por concessionária de energia elétrica, base deste trabalho. No Quadro 4 foi dimensionada a carga instalada e o consumo estimado deste tipo de habitação (CRESESB, 2003) que serviram de parâmetro do potencial energético necessário a ser produzido pelo sistema renovável.

3.2 LIMPEZA E ADAPTAÇÃO DOS DADOS

Neste trabalho a limpeza dos dados meteorológicos observou, dentro da estrutura dos registros da estação automática, aqueles cuja informação era de forma inconsistente, incompleta ou ruidosa (*Overfitting*). Como exemplo de informação ruidosa, os valores negativos para radiação no horário noturno que foram substituídos pelo valor 0 (zero). O valor negativo da radiação é medido no horário noturno pelo fato da onda de radiação ter seu sentido invertido, quando a radiação térmica do terreno, onde se encontra a estação, devolver a mesma para a atmosfera.

Na preparação dos dados para serem utilizados na ferramenta WEKA, aos registros de radiação solar que estavam no período noturno foi atribuído valor 0 (zero). Também registros inconsistentes foram eliminados por motivo de *overfitting* ou *underfitting*.

Os registros incompletos foram descartados, pois o quantitativo desses não influenciaria na observação dos resultados do estudo. Da mesma forma, foram

descartados os registros denominados *outliers*, ou seja, aqueles divergentes dos valores esperados para determinada variável.

3.2.1 Especificação dos Atributos

Após a etapa de limpeza dos dados, os registros a serem minerados para a produção de energia renovável proporcionaram o uso da técnica de mineração de dados sobre 73.428 registros.

Foram criados quatro grupos de dados de interesse na forma de atributos, sempre considerando a hora local da cidade de Campos dos Goytacazes-RJ:

3.2.1.1 Velocidade dos Ventos no período da manhã (VEL_MAN)

A quantidade de dados deste atributo, 17.724 registros, foram agrupados como a média aritmética dos dados dos ventos matinais anotados entre 6:00h e 12:00h, conforme a Figura 14:

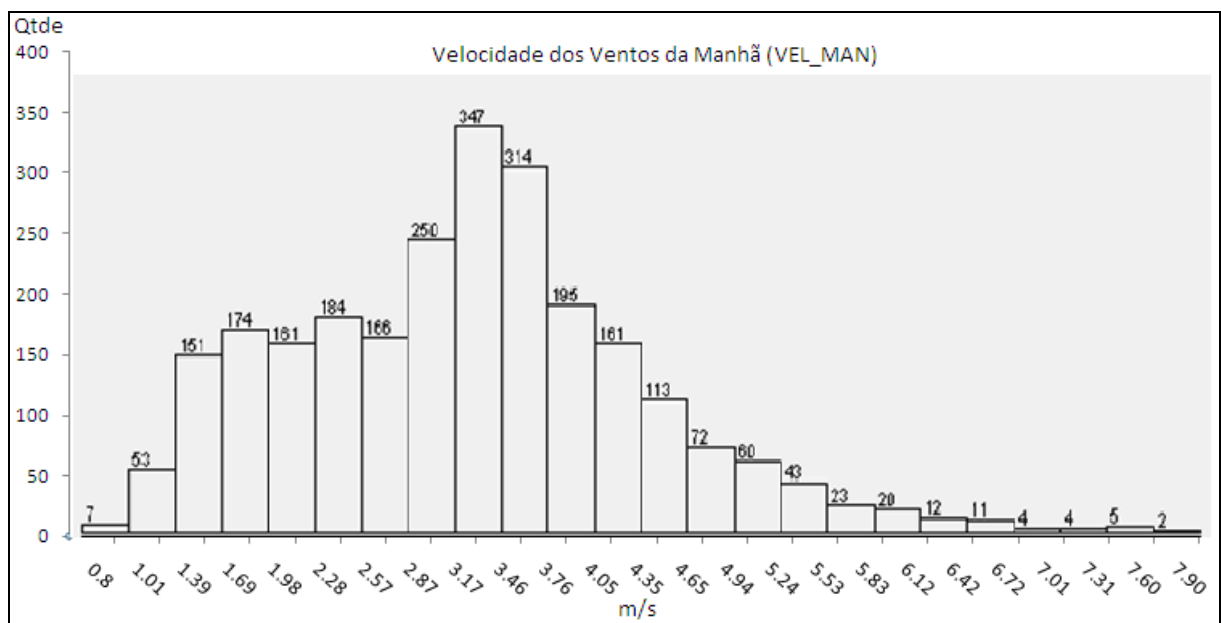


Figura 14: Gráfico VEL_MAN (velocidade dos ventos da manhã)

Fonte: WEKA: Pré-processamento - Adaptado pelo autor

3.2.1.2 Velocidade dos Ventos no período da tarde (VEL_TAR)

Da mesma forma, os dados dos ventos vespertinos coletados, 15.192 registros, foram considerados entre 13:00h e 18:00h, demonstrados na Figura 15:

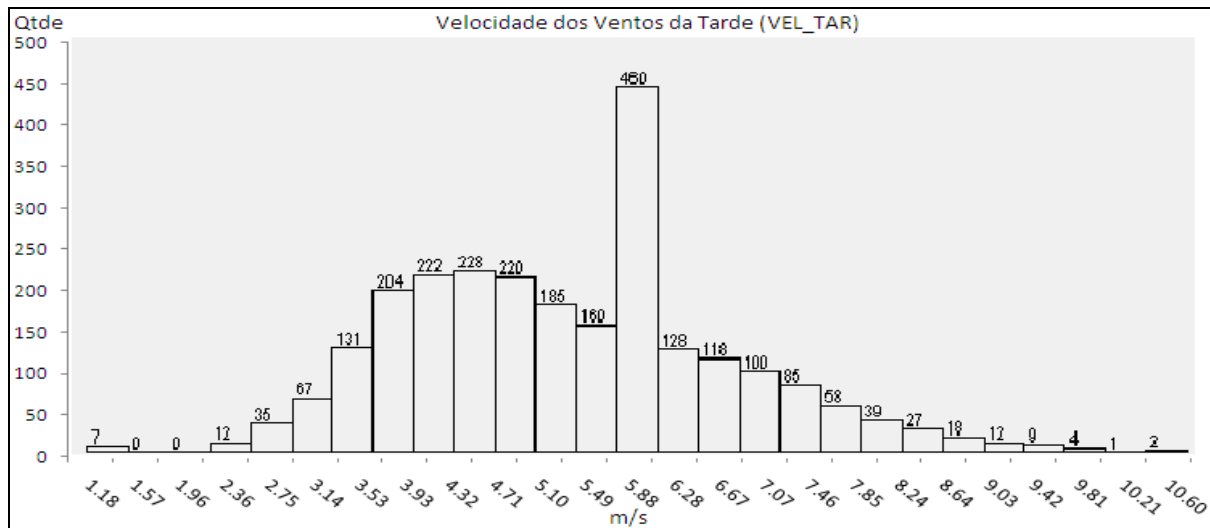


Figura 15: Gráfico VEL_TAR (velocidade dos ventos da tarde)

Fonte: WEKA: Pré-processamento - Adaptado pelo autor

3.2.1.3 Velocidade dos Ventos no período da noite (VEL_NOI)

Já para o período da noite, conforme a Figura 16, considerou-se o horário de 19:00h às 5:00h, onde a forma de média aritmética também foi utilizada para 27.852 dados registrados:

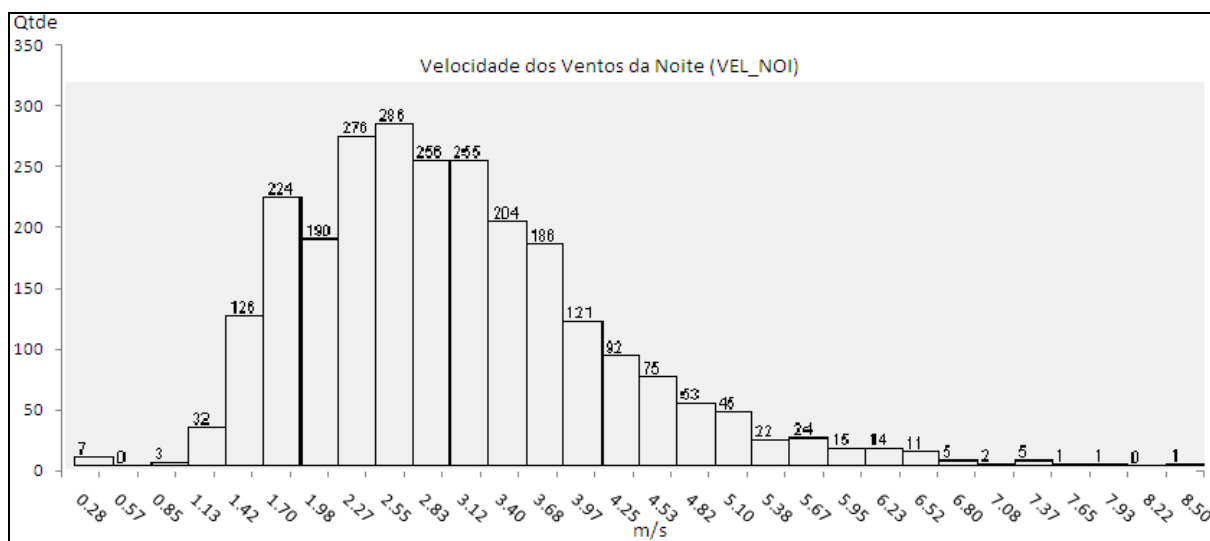


Figura 16: Gráfico VEL_NOI (velocidade dos ventos da noite)

Fonte: Fonte: WEKA: Pré-processamento - Adaptado pelo autor

3.2.1.4 Radiação Solar (RAD)

No atributo RAD foi considerada a média aritmética dos valores de radiação solar obtidos entre 10:00h e 14:00h dos 12.660 dados registrados. Desta forma, contemplou-se os valores próximos à crista da curva da Figura 12 que apresentam os melhores índices de radiação para aproveitamento fotovoltaico (TIBA et al., 2000), conforme a Figura 17:

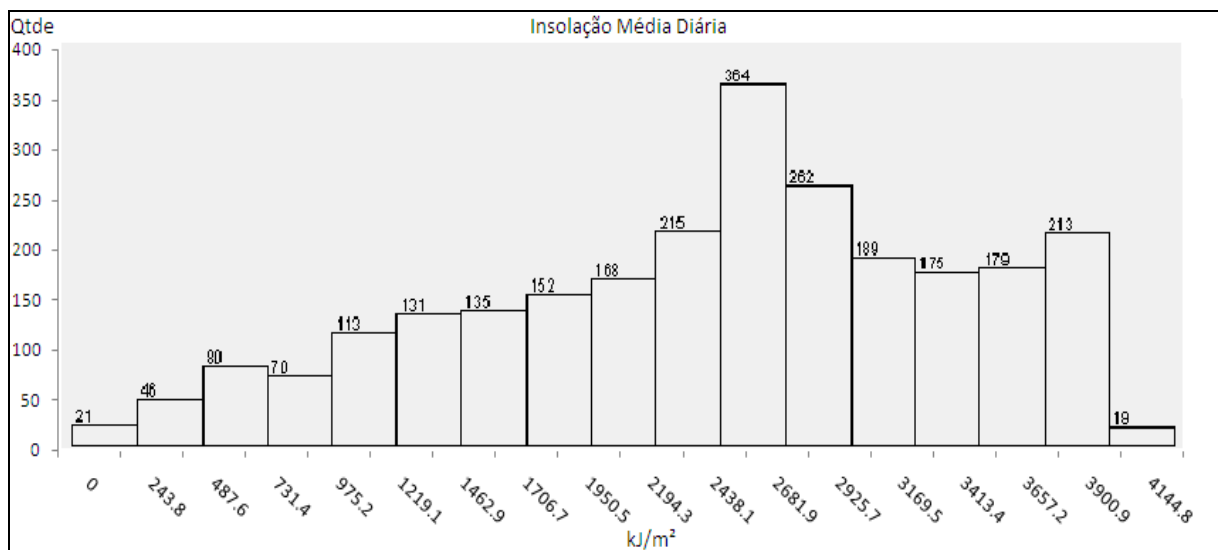


Figura 17: Gráfico RAD (radiação solar)

Fonte: WEKA: Pré-processamento - Adaptado pelo autor

Após a preparação dos dados e criação dos atributos, estes foram convertidos para um arquivo Excel CSV da Microsoft Inc., auxiliados por um arquivo Bloco de Notas e logo após transformados em arquivos ARFF (Attribute-Relation File Format) para utilização no WEKA. O arquivo criado foi denominado Renova.arff: (🐼 Renova). Nesta fase, o aplicativo foi executado sem ocorrência de erros, pois caso houvesse algum registro e/ou atributo não parametrizado de forma correta, o WEKA apresentaria mensagem e assim não geraria resultados para mineração dos dados de radiação solar e velocidade dos ventos. Parâmetros fundamentais para interpretação do objetivo geral deste trabalho.

3.3 ETAPAS DE UTILIZAÇÃO DA FERRAMENTA WEKA

Nesta Seção, são mostradas as etapas realizadas, neste estudo de caso, na ferramenta WEKA para a geração dos resultados, tendo como objetivo a verificar viabilidade da aplicação de energias renováveis eólica e solar através da mineração dos dados. Foram utilizadas as técnicas de Mineração de Dados para gerar um estudo inicial das condições climáticas do local. A metodologia proposta poderá contribuir para a tomada de decisão relacionada a este tipo de geração de energia conseqüentemente, impactará positivamente no meio ambiente.

Serão demonstrados: a aplicação do filtro de atributos, a escolha do algoritmo de clusterização e, principalmente, as informações geradas pela mineração dos dados para uma análise de resultados obtidos.

Já com o programa na área de trabalho e na fase de Pré-processamento, foi selecionado o filtro “*AllFilters*”. Um filtro genérico que passa através de todas as instâncias não modificadas para qualquer atributo. Ele é muito flexível e aponta para várias pesquisas e métodos de evolução a serem combinados.

A tarefa de mineração de dados utilizada neste trabalho foi *clustering* (regra de agrupamento). Esta tarefa foi sub-dividida em 3 sub-tarefas de agrupamento, ou seja, foi aplicada a técnica de agrupamento de 3 arranjos diferentes. O primeiro arranjo abordou o agrupamento das velocidades dos ventos nos períodos da manhã, tarde e noite, sendo ignorado o atributo radiação(RAD). A função da aplicação do algoritmo de clusterização foi encontrar, pela relação de similaridade das classes de atributos, a ocorrência da frequência dos registros de ventos que favorecessem a produção de energia no modal eólico exclusivamente.

Cabe neste momento destacar que apesar dos registros “Hora UTC” não ter qualquer influência na geração de energia elétrica quando se trata de energia eólica, é conveniente que esteja relacionado com a velocidade do vento para mostrar pelos resultados qual período do dia seria mais propício e/ou de maior produtividade neste modal de energia.

No segundo e terceiro arranjo, as tarefas utilizadas procuraram a relação da radiação registrada com os ventos da manhã e tarde. Neste caso, o atributo não associado foi a velocidade no período da noite (VEL_NOI), pois não faria sentido relacioná-lo com a radiação que, para efeitos de produção fotovoltaica, apenas são considerados horário diurnos. Porém, vale ressaltar que os ventos noturnos

contribuiriam para produção e armazenamento de energia no período de menor pico de consumo doméstico, geralmente na madrugada. Estas tarefas focaram a produção energética pelo sistema híbrido, pois os valores de radiação seriam auxiliados pelos valores de ventos diurnos na captação de diferença de potencial para os acumuladores. A utilização dos agrupamentos descritas na segunda e terceira tarefas intencionou mostrar os percentuais favoráveis na produção de energia que validassem a proposta de utilização de sistema híbridos autônomos.

3.4 ALGORITMO DE CLUSTERIZAÇÃO

Após a fase de pré-processamento, foi aplicado o algoritmo de clusterização *SimpleKMeans*, que se apresenta na sua forma computacional como: *SimpleKMeans - N 4 - A EuclideanDistance -R first-last" - I 500 - S 10*. Onde N é quantidade de clusters igual a 4 (N 4), utilizando a Distância Euclidianiana como padrão, número de interações igual a 500 (I 500) e alimentação randômica de 10 vezes (S 10). Distância Euclidianiana (ou distância métrica) é medida de proximidade bastante utilizada dada pela raiz quadrada do somatório dos quadrados da diferença dos valores para n variáveis.

De acordo com Kumar (2009), K-Means é um algoritmo de clusterização que particiona determinado conjunto de dados k em uma quantidade de clusters definido pelo operador. É a técnica mais simples de aprendizagem não-supervisionada que consiste em fixar k centróides de maneira aleatória para cada cluster.

Nas Considerações Finais deste trabalho, serão feitas as considerações em cada fase dos implementos, observando os resultados nas conclusões do trabalho tanto para energia eólica quanto para energia fotovoltaica. Também serão analisados os resultados para a geração de energia elétrica no sistema híbrido, ou seja, energia eólica associada à energia fotovoltaica ao mesmo tempo. Importante aspecto de geração de energia, pois aproveita ao máximo os recursos naturais dos ventos e do sol.

4 RESULTADOS e DISCUSSÕES

Após a aplicação do filtro na fase de pré-processamento, foi possível obter os valores médios da velocidade dos ventos e da radiação solar dos atributos considerados. Vale ressaltar, conforme alínea 'd' do item 3.2.1, que a média de radiação leva em consideração 5 horas diárias quando há incidência solar de melhores índices (TIBA et al., 2000).

Relacionando-os com os equipamentos geradores de referência, citados nos item 3.1.1, obteve-se o Quadro 4 que apresenta as potências elétricas produzidas por tais equipamentos:

Quadro 4: Potência Elétrica Gerada pelas Médias da velocidade do vento e radiação por período

Médias dos valores por período		Potência Média Gerada
Velocidade dos Ventos da Manhã:	3,24 m/s	270,0 watts
Velocidade dos Ventos da Tarde:	4,63 m/s	386,0 watts
Velocidade dos Ventos da Noite:	2,85 m/s	0,0 watts
Radiação do período considerado:	2407,8 kJ/m ²	140,0 watts

Fonte: WEKA: Etapa de Pré-processamento - Adaptado pelo autor

De acordo com o Quadro 4, onde apontam-se as médias obtidas nos quatro atributos, pode-se inferir que os ventos da manhã e da tarde são capazes de alimentar as cargas diurnas desta residência, mas que pela média os ventos noturnos devem ser usados os acumuladores no período noturno. Já para o modal fotovoltaico apenas, demonstram os resultados que durante o dia seria atendida a demanda e que para o período de não produção necessitaria também dos

acumuladores. Porém, outras finalidades para uso da energia produzida por sistema fotovoltaico poderiam ser atendidas, como por exemplo: irrigação de pequena plantação de subsistência, bombeamento de água de poços artesianos, dentre outras.

Na primeira tarefa de mineração de dados, onde se aplicou o algoritmo de agrupamento *SimpleKMeans*, considerando agrupamento apenas entre as velocidades dos ventos diurnos e noturnos, os resultados são apresentados no Quadro 5.

Quadro 5: Resultados da 1ª Tarefa de Agrupamento utilizando Algoritmo *SimpleKMeans*

AGRUPAMENTO VELOCIDADE DOS VENTOS	CLUSTERS				Registros
	0	1	2	3	Agrupados
PERCENTUAL	29	29	12	30	100
ATRIBUTO	VELOCIDADE DOS VENTOS				unidade
VEL_MAN	3,39	3,63	5,05	1,99	m/s
VEL_TAR	5,46	4,31	6,92	3,21	m/s
VEL_NOI	3,49	2,31	4,48	2,09	m/s
PERÍODO	POTÊNCIA MÉDIA GERADA				unidade
MANHÃ	0,283	0,303	0,421	0,000	kW
TARDE	0,455	0,359	0,577	0,268	kW
NOITE	0,291	0,000	0,373	0,000	kW

Fonte: WEKA: Tarefa de Agrupamento das Velocidades dos Ventos - Adaptado pelo autor.

Os resultados apresentados no Quadro 5, inferem que:

- 41% dos registros produzem energia durante o dia inteiro (Clusters 0 e 2);
- Todos os Clusters apontam que os ventos vespertinos são capazes de alimentar as cargas estimadas sem acumuladores;
- Os ventos noturnos em 59% das vezes não possibilitam alimentação autônoma, necessitando dos acumuladores (Clusters 1 e 3);
- 12% dos registros apontam produção elétrica considerável pela manhã e pela noite, capazes de suprir a demanda de carga do sistema (Cluster 2);
- 70% dos ventos da manhã produzem potência elétrica suficiente para as cargas dimensionadas (Clusters 0,1 e 2).

Os agrupamentos a seguir, conforme anteriormente dito, relacionam a radiação solar com os ventos da manhã e da tarde. Uma informação calculada foi introduzida,

a qual foi denominada “**GAP DE ENERGIA**”, que vem a ser a diferença entre a energia produzida e a consumida pelos eletrodomésticos desta residência em questão. Esta diferença vislumbra a quantidade de energia excedente ou deficitária que os acumuladores receberão ou terão de ceder ao sistema. O trabalho levou em consideração para os cálculos: 7 horas diárias de ventos matutinos, 6 horas por dia dos ventos vespertinos e 5 horas de insolação diária.

Para a segunda tarefa de *clustering*, onde a radiação relacionou-se com os ventos da manhã, o Quadro 6 foi construído a partir dos valores encontrados na aplicação do algoritmo:

Quadro 6: Resultados da 2ª Tarefa de Agrupamento utilizando Algoritmo *SimpleKMeans*

AGRUPAMENTO RAD x VEL_MAN	CLUSTERS				Registros
	0	1	2	3	Agrupados
PERCENTUAL	27	22	17	34	100
ATRIBUTO	VELOCIDADE DOS VENTOS				unidade
VEL_MAN	3,43	3,09	4,67	2,46	m/s
RAD	3457	1056	2563	2382	kJ/m ²
PERÍODO	POTÊNCIA MÉDIA GERADA				unidade
MANHÃ	0,426	0,397	0,529	0,140	kW
GAP DE ENERGIA	2,14	1,94	2,86	0,14	kWh

Fonte: WEKA: Tarefa de Agrupamento da Radiação com Velocidades dos Ventos na Manhã - Adaptado pelo autor.

São observados os dados expresso o Quadro 6:

- 17% na relação da radiação com os ventos da manhã produzem energia durante o período da manhã, tendo um excedente de energia de 2,86 kWh (Cluster 2);
- 34% dos registros têm a energia apenas produzida pelo sistema fotovoltaico (Cluster 3);
- Os Clusters 0, 1 e 2 apontam que os acumuladores serão carregados significativamente para o consumo noturno, com 66% dos registros meteorológicos;

Semelhantemente, o Quadro 7 é constituído pelos valores após a terceira tarefa de agrupamento: radiação com os ventos da tarde. Da mesma forma, a informação GAP DE ENERGIA foi introduzida para análise dos resultados:

Quadro 7: Resultados da 3ª Tarefa de Agrupamento utilizando Algoritmo *SimpleKMeans*

AGRUPAMENTO RAD x VEL_TAR	CLUSTERS				Registros
	0	1	2	3	Agrupados
PERCENTUAL	21	24	16	39	100
ATRIBUTO	VELOCIDADE DOS VENTOS				unidade
VEL_TAR	4,61	3,79	7,03	4,18	m/s
RAD	3357	1092	3253	2364	kJ/m²
PERÍODO	POTÊNCIA MÉDIA GERADA				unidade
TARDE	0,524	0,455	0,726	0,488	kW
GAP DE ENERGIA	2,34	1,93	3,55	2,13	kWh

Fonte: WEKA: Tarefa de Agrupamento da Radiação com Velocidades dos Ventos na Tarde - Adaptado pelo autor.

Inferem-se os resultados do Quadro 7 para as seguintes observações:

- Todos os registros apresentam produção de energia capazes de alimentar as cargas do sistema proposto;
- 16% dos registros têm produção considerável com excedente para os acumuladores de 3,55 kWh (Cluster 2);
- O arranjo do sistema híbrido se mostra eficaz, conseguindo ainda suprir a energia necessária no período da noite. pelos dados meteorológicos do período da tarde;

Para efeito de compreensão do cálculo do GAP DE ENERGIA, a seguir apresenta-se um exemplo do Quadro 7:

Cluster 0:

Energia gerada pelos ventos = 0,384 kW x 6 h = 2,30 kWh

Energia gerada por 2 h de radiação = 0,140 kW x 2 h = 0,28 kWh

Energia consumida estimada no período = 0,040 kW x 6 h = -0,24 kWh

GAP DE ENERGIA = 2,34 kWh

5 CONSIDERAÇÕES FINAIS

Nas análises dos resultados encontrados no Capítulo 4 são apontadas conclusões deste trabalho, onde serão confrontados o Objetivo Geral e os Objetivos Específicos para validação do uso da mineração de dados como apoio à decisão de viabilidade de implantação de energias renováveis em ambientes não atendidos pela energia elétrica convencional.

Na aplicação das técnicas de mineração de dados nos registros meteorológicos, foram encontrados resultados que proporcionaram a descoberta do conhecimento a respeito de tais dados. Tendo como objetivo central do trabalho a viabilidade de uso das energias renováveis - eólica e solar - numa determinada moradia não atendida pela energia elétrica de concessionária, os parâmetros a serem minerados foram: velocidade dos ventos e radiação solar. Deste modo, os dados meteorológicos de uma estação automática, situada em Campos dos Goytacazes-RJ, foram fornecidos pelo INMET.

Após a escolha dos parâmetros de interesse para produção de energia através dos ventos e do sol, foi procedida a limpeza dos dados, explicados no item 3.2. Ao final desta etapa o banco de dados meteorológicos proporcionou 73.428 dados para serem minerados. Para melhor dimensionamento dos parâmetros, foram especificados quatro atributos relacionando-os com o período do dia: manhã, tarde e noite, conforme item 3.2.1.

Ressalta-se que para produção e uso de energia renovável, foram usadas as características técnicas de um turbina eólica, um painel fotovoltaico e uma bateria estacionária, especificados no item 3.1.1 e uma demanda para uma residência isolada, dos equipamentos descritos no Quadro 3.

Com o intuito de validar, através da mineração de dados, o uso das fontes renováveis para produção doméstica de energia, o trabalho proporcionou as seguintes conclusões:

- Ser viável a obtenção da eletricidade para uso em pequeno porte, em áreas não atendidas por energia elétrica convencional, através dos ventos e do sol;
- Os equipamentos e a configuração do sistema são de fácil arranjo e instalação, não necessitando de espaçamentos especiais;
- A produção de energia por fontes renováveis satisfaz, na demanda deste trabalho, uma demanda mensal de até 50 kWh por mês (BRASIL – LUZ PARA TODOS, 2011);
- A produção confrontada com o consumo noturno, desta mesma demanda considerada, apenas seria atendido em 41% das vezes - pela força dos ventos da noite - o que não permite abrir mão de baterias estacionárias como acumuladores de energia;
- Tanto a fonte eólica como a fonte solar, isoladamente, poderiam alimentar equipamentos como: irrigadores, bombas de recalque, etc;
- O sistema híbrido proporcionaria os acumuladores serem recarregados no período do dia após consumo noturno, permitindo a reciclagem eficiente das cargas.

Este trabalho se baseou no consumo de residência isolada notoriamente de baixa renda (IBGE, 2010). Neste aspecto, há regiões no interior do Brasil que são afastadas, principalmente na região Norte e agreste do Nordeste do país, onde estações meteorológicas estão e poderão registrar uma quantidade significativa de dados para corroborar as pesquisas de obtenção de energia limpa e renovável.

Quanto mais pesquisas e validação das mesmas para produção de energia elétrica através do sol e dos ventos forem incentivadas, e consequentemente, a demanda de utilização for aumentada, os insumos técnicos e equipamentos terão redução expressiva nos custos. Programas do governo federal, como o PROINFA (BRASIL, 2003), poderão cada vez mais atingir um maior número de habitantes que hoje não tem acesso à energia elétrica. Além, é claro, que a produção particular

estará sendo incentivada também, mesmo naquelas residências atendidas por rede convencional de energia, chamado *grid tie*.

Portanto, o uso da ferramenta computacional de mineração de banco de dados mostrou-se eficaz na descoberta do conhecimento dos índices meteorológicos para produção de energia elétrica, através de fontes alternativas renováveis, como a fotovoltaica e eólica. A quantidade de registros coletados em sete anos permitiu à análise dos resultados percorrendo as quatro estações do ano na região em que estava instalada a referida estação meteorológica.

Os valores percentuais de radiação solar e velocidades dos ventos corroboraram com as expectativas motivadoras deste trabalho. Lembrando que para eficiência energética, quantitativo de produção e arranjos de cargas maiores, outros estudos são demandados. Ou seja, ainda tem bastante a descobrir com esta técnica computacional para demandas na matriz energética do Brasil, principalmente as renováveis. Os resultados parciais foram oriundos de um artigo publicado no SIMPEP em 2011, que focou naquela época os dados meteorológicos de uma estação meteorológica situada no distrito de Farol de São Tomé, no município de Campos dos Goytacazes/RJ.

5.1 TRABALHOS FUTUROS

O potencial da mineração de dados para análise dos registros de estações meteorológicas pode ser expandido às regiões mais remotas do interior do Brasil. Regiões onde o atendimento da energia elétrica convencional não se faz presente.

A Região Norte e Nordeste, principalmente o agreste nordestino e a floresta amazônica, são áreas extremamente carente, do ponto de vista do atendimento de energia elétrica por concessionária. Também são nestas regiões onde os índices de velocidade dos ventos e/ou insolação estão com médias maiores, segundo os mapas de potencial eólico (CRESESB, 2001) e solarimétrico do Brasil (INPE, 2006).

O estudo nestas regiões, ora proposto neste trabalho, dimensionaria de forma confiável e precisa a viabilidade de implantação destas energias renováveis. De modo prático, os registros poderiam ser levantados através da rede mundial de computadores (internet) e serem minerados para obtenção do potencial energético para cada região específica.

Iniciativa como estas poderiam incrementar os projetos governamentais de distribuição de renda através do acesso ao desenvolvimento que a energia elétrica pode proporcionar. Postos de Saúde, Escolas e Sistemas de Comunicação poderiam ser alimentados eletricamente através das energias renováveis obtidas na natureza e “descobertas” por aplicativo computacional, como a mineração de dados.

O Programa “LUZ PARA TODOS” (BRASIL, 2011) do governo federal poderia ser expandido às regiões de difícil instalação ou impraticável instalação de energia elétrica convencional, fornecida por concessionária de energia elétrica.

A mineração de dados pode e deve ser utilizada para este tipo de desenvolvimento social, onde a falta de energia elétrica inviabiliza de forma significativa o progresso de uma região ou povoado. Os custos operacionais de implantação dos equipamentos seriam bem menores que todas as logísticas de políticas públicas, elaboradas pelos governos municipais, estaduais e federal.

REFERÊNCIAS BIBLIOGRÁFICAS

AGÊNCIA INTERNACIONAL DE ENERGIA. **Key World Energy Statistics 2010**. Disponível em: <http://www.iea.org/textbase/nppdf/free/2010/key_stats_2010.pdf>, acessado em 31/ago/2013.

AGÊNCIA NACIONAL DE ENERGIA ELÉTRICA (ANEEL). **Atlas da energia Elétrica no Brasil** - 3ª edição – 2008. Disponível em: <http://www.aneel.gov.br/arquivos/pdf/atlas_capa_sumario.pdf>, acessado em 14/set/2013.

AMARAL, B. M. **Modelos VARX para Geração de Cenários de Vento e Vazão Aplicados à Comercialização de Energia**. Dissertação (Mestrado em Engenharia Elétrica). Pontifícia Universidade Católica do Rio de Janeiro, Rio de Janeiro, 2011.

BATISTA, G.E.A.P.A. **Pré-Processamento de Dados em Aprendizado de Máquina Supervisionado**. 2003. Tese (Doutorado em Ciência de Computação e Matemática Computacional). Universidade de São Carlos, São Carlos, SP, 2003.

BOENTE, A.N.P.; GOLDSCHMIDT, R.; ESTRELA, V.V. **Uma Metodologia de Suporte ao Processo de Descoberta de Conhecimento em Bases de Dados**. In: SIMPÓSIO DE EXCELÊNCIA EM GESTÃO E TECNOLOGIA (5), 2008. Resende – RJ: V SEGeT, 2008. v. 1. p. 4-5.

BRASIL. **Lei nº 10.762**, de 11 de novembro de 2003. Institui o Programa de Incentivo às Fontes Alternativas de Energia Elétrica - PROINFA. Brasília, DF, 2003. Disponível em: <<http://www.mme.gov.br/programas/proinfa>>, acessado em 24/ago/2013.

BRASIL. **Decreto nº 6.135**, de 26 de junho de 2007. Dispõe sobre o Cadastro Único para Programas Sociais do Governo Federal e dá outras providências.. Brasília, DF, 2007. Disponível em: http://www.planalto.gov.br/ccivil_03/_ato2007-2010/2007/decreto/d6135.htm, acessado em 24/ago/2013.

BRASIL. **Decreto nº 7.520**, de 8 de julho de 2011. Institui o Programa Nacional de Universalização do Acesso e Uso da Energia Elétrica – “LUZ PARA TODOS” para o período de 2011 a 2014. Brasília, DF, 2011. Disponível em: <<http://luzparatodos.mme.gov.br/luzparatodos/downloads>>, acessado em 24/ago/2013.

BRASIL WIND SERVICE – **Aero Geradores, marca Hummer**. Disponível em: <<http://www.brasilwindservice.com>>, acessado em 14/set/2013.

CENTRO DE REFERÊNCIA PARA ENERGIA SOLAR E EÓLICA SÉRGIO DE SALVO BRITO (CRESESB). **Atlas do Potencial Eólico Brasileiro – 2001**. Disponível em: <http://www.cresesb.cepel.br/atlas_eolico/index.php>, acessado em 14/set/2013.

CENTRO DE REFERÊNCIA PARA ENERGIA SOLAR E EÓLICA SÉRGIO DE SALVO BRITO (CRESESB). **Coletânea de Artigos: Energias Solar e Eólica**. Rio de Janeiro: Imprinta, v.1, 2003.

COMPANHIA DE ELETRICIDADE DO ESTADO DA BAHIA (COELBA). **Programa Luz Para Todos**. Disponível em: <<http://servicos.coelba.com.br/residencial/luz-para-todos>>, acessado em 30/dez/2013.

CHAVES, C. R. **Apostila de Instrumentação Básica do Curso de Formação de Operadores de Refinaria**. Petrobrás. Curitiba: UnicenP, 2002. 98 p.

DE LOYS, I. **Utilização de Energia Solar Fotovoltaica em Áreas Urbanas: O Exemplo da Ecoducha Solar**. 2012. Dissertação (Mestrado em Engenharia Urbana). Universidade Federal do Rio de Janeiro. Rio de Janeiro, 2012.

DILLY, R. **Data mining: An Introduction**. Parallel Computer Centre. Queen's University of Belfast, 1995. Disponível em: <http://www.pcc.qub.ac.uk/tec/coursers/datamining/stu_notes/dm_book_2.html>, acessado em 31/mai/2013.

DUTRA, R.M. **Proposta de Políticas Específicas para Energia Eólica no Brasil após a Primeira Fase do Proinfa**. Tese (Doutorado em Ciências em Planejamento Energético). Universidade Federal do Rio de Janeiro, Rio de Janeiro, 2007

DUTRA, R.M.; TOLMASQUIM, M.T. **Estudo da Viabilidade Econômica para Projetos Eólicos com Base no Novo Contexto do Setor Elétrico**. In: CONGRESSO BRASILEIRO DE ENERGIA (9), 2002. Rio de Janeiro, 2002.

DIAS, M.M. **Um modelo de formalização do processo de desenvolvimento de sistemas de descoberta de conhecimento em banco de dados**. 2001. Tese (Doutorado em Engenharia de Produção). Universidade Federal de Santa Catarina. Florianópolis, 2001.

EMPRESA DE PESQUISA ENERGÉTICA (EPE). **Plano Nacional de Energia – 2030**. Disponível em: <<http://www.epe.gov.br>>, acessado em 14/set/2013.

ERTHAL JUNIOR, M.; HOSKEN, L.A.L.; MENDES, L.F.R. **Seleção de Alternativas Energéticas para Propriedades Rurais da Planície Litorânea do Norte do Estado do Rio de Janeiro**, 2012. ENCONTRO MINEIRO DE ENGENHARIA DE PRODUÇÃO (8). Itajubá, 2012.

FAYYAD, U. PIATETSKY-SHAPIO, G. and SMYTH, P. ***From Data Mining to Knowledge Discovery in Databases***. AI Magazine. p 37-54, 1996. Disponível em: <<http://www.aaai.org/Magazine/magazine.php>>, acessado em 14/set/2013.

FCMC. **Fórum Capixaba de Mudanças Climáticas e Uso Racional da Água**. Disponível em: <<http://www.fcmc.es.gov.br>>, acessado em 30/dez/2013.

GOEBEL, M.; GRUENWALD, L. A survey of data mining and knowledge discovery software tools **ACM SIGKDD**, San Diego, v. 1, n. 1, p. 20-33, 1999.

GOLDSCHIMIDT, R.; PASSOS, E. ***Data mining: Um guia prático***. Rio de Janeiro: Elsevier, 2005

GROTH, R. ***Data mining: a hand-on approach for business professionals***. New Jersey: Prentice Hall, 1998.

GOULART, S. **Sustentabilidade nas Edificações e Espaços Urbanos**. Apostila da Disciplina: Desempenho Térmico nas Edificações do Laboratório de Eficiência Energética em Edificações da Universidade Federal de Santa Catarina, 2013. Disponível em: <<http://www.labeee.ufsc.br>>, acessado em 30/dez/2013.

GOULART JUNIOR, F.S.; FIDALGO, R. ***Data mining: Apresentação de aula***. UFPE, 1998. Disponível em: <<http://www.di.ufpe.br/~compint/aulas-IAS/agentes/taci1-981/Datamining.ppt>>, acessado em 01/dez/2012.

HAN, J; KAMBER, M. ***Data Mining: Concepts and Techniques***. San Francisco – USA: Morgan Kaufmann, 2.ed, 2006.

HODGE, B.K. **Sistemas e Aplicações de Energias Alternativas**. São Paulo: LTC, 2011, 324p.

INSTITUTO BRASILEIRO DE GEOGRAFIA E ESTATÍSTICA (IBGE). **Resultados Preliminares do Universo do Censo Demográfico 2010**. Tabela 13 – Domicílios particulares permanentes, por existência de energia elétrica, segundo as Grandes Regiões e as Unidades da Federação – 2010. Disponível em: <<http://www.censo2010ibge.gov.br/sinopse>>, acessado em 12/out/2013.

INSTITUTO NACIONAL DE METEOROLOGIA (INMET) - SONABRA. **Monitoramento de Estações Automáticas**. Disponível em: <<http://www.inmet.gov.br/SONABRA/maps/automaticas.php>>, acessado em 14/set/2013.

INSTITUTO NACIONAL DE PESQUISA ESPACIAL (INPE). **Atlas Brasileiro de Energia Solar** – 2006. Disponível em: <http://www.sonda.ccst.inpe.br/publicacoes/atlas_solar.html>, acessado em 14/set/2013.

KUMAR, V. ***Data Mining and Knowledge Discovery Series***. Minneapolis – USA: Taylor; Francis Group, 2009.

LIEBSTEIN, L. H. **Data Mining – teoria e prática**. Porto Alegre, 2001. Disponível em: <http://www.inf.ufrgs.br/~clesio/cmp151/cmp15120021/artigo_lourdes.pdf>, acessado em 14/sete/2013.

MARKOV, Z.; RUSSELL, I. **An Introduction to the WEKA Data Mining System Proceedings**. In: ANNUAL SIGCSE CONFERENCE ON INNOVATION AND TECHNOLOGY IN COMPUTER SCIENCE EDUCATION(11). p. 367 - 368. Bologna, Itália, 2006.

MENDES, L.F.R.; HOSKEN, L.A.L.; GOMES, G.R.R. **Estudo de Viabilidade para Implantação de Energias Renováveis no Distrito de Farol de São Thomé em Campos dos Goytacazes-RJ – Utilizando Mineração de Dados**, 2011. In: SIMPÓSIO DE ENGENHARIA DE PRODUÇÃO(18). Bauru, SP, 2011.

MURASSE, C.M., TSUNODA, D.F. **Descoberta de conhecimento a partir de uma base de indicadores de desenvolvimento social utilizando WEKA**, 2010. Disponível em: <http://www.inf.pucminas.br/sbc2010/anais/pdf/wcge/st01_02.pdf>, acessado em 27/mai/2011.

PEDROTI, G.I. **Ensaio de Biodegradabilidade Aeróbia de Hidrocarbonetos Derivados do Petróleo em Solo**. Dissertação (Mestrado em Engenharia Ambiental). Universidade Federal do Espírito Santo, Vitória, 2007.

PIANEZZOLA, G. **Mapas de Complementaridade dos Potenciais Solar e Eólico no Estado do Rio Grande do Sul**. Dissertação (Mestrado em Engenharia). Universidade Federal do Rio Grande do Sul, Porto Alegre, 2006.

PROCEL – Programa Nacional de Conservação de Energia Elétrica. **Procel Info - Centro Brasileiro de Informação de Eficiência Energética. Potência de Eletrodomésticos**. Disponível em: <<http://www.procelinfo.com.br/main.asp>>, acessado em 14/set/2013.

QUONIAM L.; TARAPANOFF K.; ARAUJO JR R.H.; ALVARES L. **Inteligência obtida pela aplicação de data mining em base de teses francesas sobre o Brasil**. Ciência da Informação, v.30, 2001. Disponível em: < <http://www.scielo.br>>, acessado em 14/set/2013.

ROBAINA, F. **Apostila da Disciplina: Sistema de Potência**. Curso Técnico em Eletrotécnica do Cietec. Campos dos Goytacazes, 2011.

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. Disponível em: <http://www.scielo.br/scielo.php?script=sci_arttext&pid=S0102-77862009000400007&lng=pt&nrm=iso>, acessado em 14/set/2013.

SOLAR BRASIL – Tecnologia & Energia Fotovoltaica Ltda. **Painel Fotovoltaico M140P , marca Martifer**. Disponível em: <<http://www.solarbrasil.com.br>>, acessado em 14/set/2013.

TAKECIAN, P.L; FERREIRA J.E. **Methodological Guidelines and Adaptive Statistical Data Validation to Build Effective DataWarehouses**. SIMPÓSIO BRASILEIRO DE BANCOS DE DADOS – SBBD 2012. Workshop de Teses e Dissertações. Disponível em: <<http://www.lbd.dcc.ufmg.br/colecoes/wtdbd/2012/0010.pdf>>, acessado em 14/set/2013.

TAN, P; STEINBACH, M; KUMAR, V. **Introdução ao Data Mining Mineração de Dados**. Rio de Janeiro: Ciência Moderna, 2009.

TIBA, C.; FRAIDENRAICH, N. **Atlas Solarimétrico do Brasil: banco de dados terrestres**. Recife. Editora Universitária da UFPE, 2000.

UNIVERSIDADE DE WAIKATO. Disponível em: <<http://www.cs.waikato.ac.nz/ml/weka>>, acessado em 14/set/2013.

WEKA. **Data Mining Software in Java**. Disponível em: <<http://www.cs.waikato.ac.nz/ml/weka>>. 2008, acessado em: 30/dez/2013.

WITTEN, I. H.; FRANK, E. - **Data Mining – Practical Machine Learning Tools and Techniques** - San Francisco - USA: Morgan Kaufmann 2.ed, 2005.