



Mestrado em Engenharia Eletrotécnica

Previsão da Produção Média de Energia para o Dia D+2

Trabalho de Projeto apresentado para a obtenção do
grau de Mestre em Engenharia Eletrotécnica
Especialização em Automação e Comunicações em Sistemas de Energia

Autor

André Filipe dos Santos Martins

Orientadores

Doutora Cristina Isabel Ferreira Figueiras Faustino Agreira

Professora do Departamento de Engenharia Eletrotécnica
Instituto Superior de Engenharia de Coimbra

Engenheiro Rui José Oliveira Nóbrega Pestana

REN – Rede Eléctrica Nacional, S.A.

Coimbra, maio, 2018

In Memoriam

Ao meu pai José, avó Amália, irmã Marília e José Cabral

AGRADECIMENTOS

A concretização deste projeto não seria possível sem o apoio e motivação de um conjunto de pessoas, ao qual não poderia deixar de prestar algumas palavras de agradecimento.

Agradeço ao Instituto Superior de Engenharia de Coimbra e ao seu Departamento de Engenharia Eletrotécnica por me terem proporcionado as condições necessárias para a realização deste projeto.

O meu agradecimento à REN – Rede Elétrica Nacional, S.A. por me ter acolhido durante os dois meses e por me ter facultado os dados necessários que me permitiram desenvolver a minha investigação e levá-la a bom porto.

O meu sincero agradecimento à minha orientadora, Professora Doutora Cristina Isabel Ferreira Figueira Faustino Agreira e ao meu orientador, Engenheiro Rui José Oliveira Nóbrega Pestana que me propuseram e orientaram neste trabalho. Obrigado por esta possibilidade e pela confiança que depositaram em mim para a realização deste projeto. Todo o vosso apoio, partilha de conhecimentos, disponibilidade e confiança tornaram-se fulcrais para a sua realização.

Agradeço ainda à minha mãe por todo o apoio dado até ao dia de hoje. Sem ela sei que nada disto teria sido possível. Todo o sacrifício que tem feito por mim, toda a força e amor que me tem dado. Devo-lhe muito pela pessoa que sou hoje.

Às minhas irmãs, Ana Martins e Carla Cabral, por todo o apoio, força e incentivo que me têm dado ao longo do meu percurso académico e da minha vida. Um obrigado não chega.

À minha namorada, Joana Oliveira, agradeço o apoio incondicional que me tem dado e todo o carinho, paciência e motivação que sempre me deu.

Agradeço de uma forma geral a toda a família e amigos que contribuíram para a realização deste projeto e conclusão do curso. Todos vocês tiveram um contributo importantíssimo no meu percurso académico, do qual eu não me esquecerei.

RESUMO

A energia elétrica desempenha uma função vital na sociedade. Prever com antecedência o que ocorre na rede elétrica é essencial de modo a garantir o acesso contínuo e seguro desta energia. Há um conjunto de procedimentos essenciais para atingir estes requisitos, dos quais se destaca o DACF – *Day Ahead Congestion Forecast*. Este é responsável pela previsão de congestionamentos na rede elétrica, essencial para garantir a segurança do sistema de transmissão Europeu. Este é realizado diariamente pelos TSOs, de acordo com os procedimentos da ENTSO-E. Efetuar o mesmo procedimento para o dia depois de amanhã, dia D+2, permite que o Operador do Sistema dos TSOs tenham um maior período de tempo para realizar as alterações necessárias na rede elétrica, de modo a garantir a sua segurança e fiabilidade.

A falta de dados previstos para este horizonte temporal condiciona a determinação do D2CF – *Day 2 Ahead Congestion Forecast*. Neste projeto é desenvolvida e aplicada uma metodologia de previsão da produção média das centrais a carvão, gás natural, albufeiras, fio-de-água, bombagem e da importação e exportação de energia, para o dia depois de amanhã (dia D+2). Esta é uma primeira abordagem para a determinação dos dados em falta, referentes às produções, para este horizonte temporal.

Este projeto está assente na Teoria dos Conjuntos Aproximativos, através da aplicação de cenários distintos, com base em diferentes códigos de equivalência e algoritmos de indução de regras. Os resultados obtidos permitiram determinar qual a previsão dos atributos em estudo, para o dia D+2, através de técnicas de *Data Mining*. A aplicação da Teoria dos Conjuntos Aproximativos permitiu extrair um conjunto de regras fundamentais de modo a obter o conhecimento necessário para classificar novos objetos.

Palavras-chave: Descoberta de Conhecimento em Bases de Dados, Mineração de Dados, Teoria dos Conjuntos Aproximativos, Previsão de produção de energia

ABSTRACT

Electricity plays a vital role in society. Predicting in advance what will happen in the electrical grid is essential in order to guarantee the supply and safe access to electricity. There are a set of procedures to meet these requirements, of which the DACF – Day Ahead Congestion Forecast stands out. It is responsible for the prediction of congestion in the grid, essential to ensure the safety of the European transmission system. This is done on a daily basis by the TSOs, according to the ENTSO-E procedures. Performing the same procedure for the day after tomorrow, D+2, allows the System Operator of the TSOs to have a longer period of time to make the necessary changes to the electrical grid, in order to guarantee its safety and reliability.

The lack of forecasted data for this time horizon conditions the determination of the D2CF – Day 2 Ahead Congestion Forecast. In this project a methodology is developed to forecast the average production of coal, natural gas, reservoirs, run-of-river, pumped-storage and energy import and export for the day after tomorrow (D+2). This is the first approach for the determination of the missing production data for this time span.

This project is based on the Rough Set Theory, focused on the application of different scenarios, based on different equivalence codes and rule induction algorithms. The results obtained allowed to determine the prediction of the attributes studied, for the day D+2, using Data Mining techniques. The application of the Rough Set Theory allowed to extract a set of fundamental rules in order to obtain the necessary knowledge to classify new objects.

Key-words: Knowledge Discovery in Databases, Data Mining, Rough Set Theory, Energy Production Forecasting

ÍNDICE

Capítulo 1 – Introdução	1
1.1 – Enquadramento	1
1.2 – Objetivos da Tese.....	3
1.3 – Estrutura da Tese.....	3
Capítulo 2 – A Descoberta de Conhecimento em Bases de Dados	5
2.1 – Introdução	5
2.2 – <i>Data Mining</i> e o processo de <i>Knowledge Discovery in Databases</i>	5
2.3 – Etapas do KDD	6
2.3.1 – Seleção de Dados	7
2.3.2 – Pré-processamento dos dados	7
2.3.3 – Transformação de dados	8
2.3.4 – <i>Data Mining</i>	8
2.3.5 – Interpretação/ Avaliação do Conhecimento	8
2.3 – Origem dos dados.....	9
2.5 – Tarefas da mineração de dados	10
2.5.1 – Classificação	10
2.5.2 – Regressão (ou Estimativa)	11
2.5.3 – Segmentação ou <i>Clustering</i>	11
2.5.4 – Associação ou Dependência	11
2.5.5 – Sumarização.....	12
2.6 – Modelos e Técnicas de <i>Data Mining</i>	12
2.6.1 – A aprendizagem automática.....	12
2.7 – Conclusão.....	13
Capítulo 3 – Teoria dos Conjuntos Aproximativos	15
3.1 – Introdução	15
3.2 – Conceitos Fundamentais da Teoria dos Conjuntos Aproximativos.....	16
3.3 – Sistema de Informação ou Tabela de Informação	16
3.4 – Relação de indiscernibilidade.....	17
3.5 – Aproximação de um conjunto	18
3.6 – Redução de atributos	20
3.7 – Tabelas de decisão	21
3.8 – Aplicação da Teoria dos Conjuntos Aproximativos a um Centro de Controlo	22
3.9 – Sistema de Aprendizagem LERS	30

3.10 - Conclusão	30
Capítulo 4 – O MIBEL e a Produção de Energia	31
4.1 – Introdução	31
4.2 – O Mercado Ibérico de Eletricidade – MIBEL	31
4.2.1 – Objetivos do MIBEL	31
4.2.2 – Fases de Implementação	32
4.2.3 – Operadores do MIBEL.....	33
4.2.3.1 – OMIE	33
4.2.3.2 – OMIP.....	34
4.2.4 – Índices de Eletricidade do MIBEL	34
4.3 – Produção de energia elétrica	36
4.4 – Conclusão.....	38
Capítulo 5 – Aplicação da Teoria dos Conjuntos Aproximativos	39
5.1 – Introdução	39
5.2 – Dados e tratamento de dados.....	39
5.3 – Discretização.....	40
5.4 – O <i>software</i> ROSE2.....	44
5.4.1 – Algoritmo de indução de regras	45
5.4.2 – Parâmetros e interpretação das regras	45
5.5 – Aplicação da Teoria dos Conjuntos Aproximativos.....	46
5.5.1 – Definição dos atributos	47
5.5.2 – Casos de estudo em análise.....	48
5.5.3 – Avaliação dos casos de estudo	48
5.5.4 – Carvão - Caso de estudo 1	49
5.5.4.1 – Algoritmo de Indução Laplace	50
5.5.4.2 – Algoritmo de Indução Entropia.....	52
5.5.5 – Carvão - Caso de estudo 2	53
5.5.5.1 – Algoritmo de Indução Laplace	54
5.5.5.2 – Algoritmo de Indução Entropia.....	55
5.5.6 – Carvão – Conclusões	57
5.5.7 – Gás Natural - Caso de estudo 1.....	57
5.5.7.1 – Algoritmo de Indução Laplace	58
5.5.7.2 – Algoritmo de Indução Entropia.....	59
5.5.8 – Gás Natural - Caso de estudo 2.....	60

5.5.8.1 – Algoritmo de Indução Laplace	60
5.5.8.2 – Algoritmo de Indução Entropia	62
5.5.9 – Gás Natural – Conclusões	63
5.5.10 – Albufeira - Caso de estudo 1	63
5.5.10.1 – Algoritmo de Indução Laplace	64
5.5.10.2 – Algoritmo de Indução Entropia	65
5.5.11 – Albufeira - Caso de estudo 2	66
5.5.11.1 – Algoritmo de Indução Laplace	66
5.5.11.2 – Algoritmo de Indução Entropia	68
5.5.12 – Albufeira – Conclusões	69
5.5.13 – Fio-de-água - Caso de estudo 1	70
5.5.13.1 – Algoritmo de Indução Laplace	70
5.5.13.2 – Algoritmo de Indução Entropia	71
5.5.14 – Fio-de-água - Caso de estudo 2	72
5.5.14.1 – Algoritmo de Indução Laplace	73
5.5.14.2 – Algoritmo de Indução Entropia	74
5.5.15 – Fio-de-água – Conclusões	76
5.5.16 – Bombagem - Caso de estudo 1	76
5.5.16.1 – Algoritmo de Indução Laplace	77
5.5.16.2 – Algoritmo de Indução Entropia	78
5.5.17 – Bombagem - Caso de estudo 2	79
5.5.17.1 – Algoritmo de Indução Laplace	80
5.5.17.2 – Algoritmo de Indução Entropia	81
5.5.18 – Bombagem – Conclusões	83
5.5.19 – Importação - Caso de estudo 1	83
5.5.19.1 – Algoritmo de Indução Laplace	84
5.5.19.2 – Algoritmo de Indução Entropia	85
5.5.20 – Importação - Caso de estudo 2	86
5.5.20.1 – Algoritmo de Indução Laplace	87
5.5.20.2 – Algoritmo de Indução Entropia	88
5.5.21 – Importação – Conclusões	89
5.5.22 – Exportação - Caso de estudo 1	90
5.5.22.1 – Algoritmo de Indução Laplace	90

5.5.22.2 – Algoritmo de Indução Entropia.....	91
5.5.23 – Exportação - Caso de estudo 2.....	92
5.5.23.1 – Algoritmo de Indução Laplace	93
5.5.23.2 – Algoritmo de Indução Entropia.....	94
5.5.24 – Exportação – Conclusões.....	95
5.5.25 – Core e Reduto.....	95
5.7 – Conclusão.....	96
Capítulo 6 – Conclusões Gerais e Sugestões para Trabalhos Futuros.....	99
6.1 – Conclusões Gerais.....	99
6.2 – Sugestões Para Trabalhos Futuros.....	100
Referências Bibliográficas	101
Anexo I – Tabelas dos Códigos de Equivalência.....	105
Anexo II – Artigo ICEUBI 2017	121

Índice de Figuras

Figura 2.1 – <i>Knowledge Discovery in Databases</i> , adaptado de (Fayyad <i>et al</i> , 1996)	6
Figura 3.1 – Conjunto das <i>aproximações a R</i> e as <i>regiões de R</i> (Lambert-Torres, 2002).	19
Figura 4.1 – Evolução da produção de energia elétrica, entre 2011 e 2015 (REN, 2016)	38
Figura 5.1 – Casos de estudo e respetivos cenários em análise	48

ÍNDICE DE TABELAS

Tabela 3.1 – Conjunto de informação reduzido de um sistema de controlo (Lambert-Torres <i>et al</i> , 1996a).....	23
Tabela 3.2 – Sistema de informação equivalente (Lambert-Torres <i>et al</i> , 1996a).....	24
Tabela 3.3 – Conjunto resultante dos exemplos (Lambert-Torres <i>et al</i> , 1996a).....	26
Tabela 3.4 – Indispensabilidade do atributo B (Lambert-Torres <i>et al</i> , 1996a).....	27
Tabela 3.5 – Conjunto de objetos com os atributos indispensáveis (I) (Lambert-Torres <i>et al</i> , 1996a).....	27
Tabela 3.6 – Conjunto de objetos com os atributos indispensáveis (II) (Lambert-Torres <i>et al</i> , 1996a).....	27
Tabela 3.7 – Core do conjunto de objetos (Lambert-Torres <i>et al</i> , 1996a).....	28
Tabela 3.8 – Reduto do conjunto de objetos (Lambert-Torres <i>et al</i> , 1996a).....	28
Tabela 5.1 – Código de equivalência do Índice PTEL.....	41
Tabela 5.2 – Código de equivalência do método 1 – Mini-hídrica, Eólica e Consumo.....	41
Tabela 5.3 – Código de equivalência do método 1 – Carvão, Gás Natural, Importação e Exportação.....	42
Tabela 5.4 – Código de equivalência do método 1 – Albufeira, Fio-de-água e Bombagem.....	42
Tabela 5.5 – Código de equivalência do método 2 – Mini-hídrica, Eólica e Consumo.....	43
Tabela 5.6 – Código de equivalência do método 2 – Carvão, Gás Natural, Importação e Exportação.....	43
Tabela 5.7 – Código de equivalência do método 2 – Albufeira, Fio-de-água e Bombagem.....	44
Tabela 5.8 – Base de dados relativa ao Carvão, aplicando códigos de equivalência do método 1.....	49
Tabela 5.9 – Base de dados dos atributos do dia (D+2) referente ao ano 2015.....	52
Tabela 5.10 – Base de dados relativa ao Carvão, aplicando códigos de equivalência do método 2.....	54
Tabela 5.11 – Base de dados relativa ao Gás Natural, aplicando códigos de equivalência do método 1.....	58
Tabela 5.12 – Base de dados relativa ao Gás Natural, aplicando códigos de equivalência do método 2.....	60
Tabela 5.13 – Base de dados relativa às Albufeiras, aplicando códigos de equivalência do método 1.....	64
Tabela 5.14 – Base de dados relativa às Albufeiras, aplicando códigos de equivalência do método 2.....	66
Tabela 5.15 – Base de dados relativa às Fio-de-água, aplicando os códigos de equivalência do método 1.....	70
Tabela 5.16 – Base de dados relativa às Fio-de-água, aplicando códigos de equivalência do método 2.....	73
Tabela 5.17 – Base de dados relativa à Bombagem, aplicando códigos de equivalência do método 1.....	77
Tabela 5.18 – Base de dados relativa à Bombagem, aplicando códigos de equivalência do método 2.....	80
Tabela 5.19 – Base de dados relativa à Importação, aplicando códigos de equivalência do método 1.....	84
Tabela 5.20 – Base de dados relativa à Importação, aplicando códigos de equivalência do método 2.....	86
Tabela 5.21 – Base de dados relativa à Exportação, aplicando códigos de equivalência do método 1.....	90
Tabela 5.22 – Base de dados relativa à Exportação, aplicando códigos de equivalência do método 2.....	92

ABREVIATURAS

AA	Aprendizagem Automática
DACF	<i>Day Ahead Congestion Forecast</i>
D2CF	<i>Day 2 Ahead Congestion Forecast</i>
DW	<i>Data Warehouse</i>
ENTSO-E	<i>European Network of Transmission System Operators for Electricity</i>
HOES	Hora Oficial Espanhola
KDD	<i>Knowledge Discovery in Database</i>
LEERS	<i>Learning from Examples based on Rough Sets</i>
MIBEL	Mercado Ibérico de Eletricidade
OMI	Operador de Mercado Ibérico
OMIE	Operador de Mercado Ibérico de Energia, polo Espanhol
OMIP	Operador de Mercado Ibérico de Energia, polo Português
PRO	Produção em Regime Ordinário
REN	Rede Elétrica Nacional
SEN	Sistema Elétrico Nacional
TCA	Teoria dos Conjuntos Aproximativos
TSO	<i>Transmission System Operator</i>
VTP	Verificação Técnica do Programa de Exploração

Capítulo 1 – Introdução

1.1 – Enquadramento

A energia é a necessidade básica para o desenvolvimento económico e social de um país. As sociedades modernas encontram-se cada vez mais dependentes do consumo de energia elétrica, havendo estudos que correlacionam o índice de qualidade de vida e o uso de eletricidade (Pasten e Santamarina, 2012). A energia elétrica é indiscutivelmente um bem essencial na indústria, habitações, hospitais, escolas, iluminação pública, etc. Desta forma, manter a continuidade e qualidade do fornecimento de energia elétrica representa um peso de grande importância na economia global e na sua competitividade.

Em Portugal, atualmente, a eletricidade é produzida com recurso a um *mix* de diferentes tecnologias e fontes primárias de energia. O aumento do número de centrais de produção de energia renovável tem sido uma aposta inevitável ao longo dos últimos anos, contribuindo para a redução da dependência energética do exterior, cuja energia é maioritariamente proveniente de combustíveis fósseis, de impacto ambiental negativo e de preços voláteis. O setor da energia elétrica encontra-se num processo complexo e dinâmico de transformação. O processo de liberalização do setor elétrico, a diversificação do *mix* energético e o aumento da produção de energia de origem renovável trouxeram novos desafios para o Operador de Sistema. Este tem como função garantir a continuidade e a segurança do fornecimento de eletricidade, respeitando os parâmetros de qualidade exigidos.

Vários procedimentos necessitam de ser cumpridos, de modo a assegurar o fornecimento de eletricidade. A cada dia, o Operador de Sistema informa o MIBEL (Mercado Ibérico de Eletricidade) da capacidade de interligação para o dia seguinte, entre Portugal e Espanha. Além disso, é utilizada uma ferramenta denominada VTP (Verificação Técnica do Programa de Exploração), que utiliza como *input* os *snapshots* da rede elétrica. Estes *snapshots* são tirados a cada hora do dia, perfazendo um total de 24 *savecases* por dia. Para além deste *input*, é também utilizado o perfil de geração da PRO – Produção em Regime Ordinário (com base nas informações disponibilizadas pelo MIBEL), as indisponibilidades da rede, a previsão de carga e a previsão da produção eólica, solar e pequena hídrica. A ferramenta VTP permite criar 24 *savecases*, com a previsão do que irá ocorrer no dia de amanhã (D+1). Esta informação é fulcral

para o procedimento DACF – *Day Ahead Congestion Forecast*, que é realizado diariamente pelos TSOs – *Transmission System Operators*, de acordo com a Política 4 do *ENTSO-E RG CE Operational Handbook* (ENTSO-E, 2004).

A liberalização do Mercado Europeu de Eletricidade trouxe várias mudanças a nível da rede de transmissão de eletricidade da Europa Continental, coordenada atualmente pela ENTSO-E. O sistema evoluiu, deixando de ser um monopólio onde a produção, comercialização, transmissão e distribuição estavam a cargo de apenas uma empresa. Sistema este, relativamente fácil de gerir e de evitar congestionamentos. A produção era a maior parte do tempo bastante estável de um dia (ou semana) para o outro, facilitando a previsão dos fluxos de carga. A liberalização veio abrir as fronteiras à comercialização de energia, permitindo a compra de uma determinada quantidade de eletricidade num país e a sua venda noutra. O fluxo de energia aumentou e tornou-se mais dinâmico, devido ao aumento da troca de eletricidade a curto-prazo (mercado intradiário) e da produção de energia eólica. Por forma a manter a segurança, garantir o fornecimento de eletricidade e rentabilizar a utilização da rede elétrica, foi necessário implementar um sistema de partilha de informação entre os TSOs (Panciatici e Vrsecky, 2005).

O procedimento DACF foi posto em funcionamento no ano 2002, resultante de uma cooperação entre vários países Europeus. Todo o procedimento requer troca de dados e um trabalho conjunto entre os TSOs envolvidos. O principal objetivo deste método é que cada país disponibilize a todos os outros uma previsão que compreenda um conjunto de dados de fluxo de carga da sua rede (Panciatici e Vrsecky, 2005).

Este procedimento permite que os TSOs tenham apenas 24 horas de antecedência para a resolução de congestionamentos, uma vez que é feito hoje para amanhã. O aumento deste espectro temporal, de modo a que o Operador de Sistema seja capaz de planear e saber antecipadamente possíveis problemas no futuro próximo, é uma necessidade. Através do procedimento D2CF – *Day 2 Ahead Congestion Forecast* seria possível satisfazer essa necessidade, prevendo hoje para depois de amanhã (D+2).

Apesar de parte da informação necessária para efetuar o procedimento D2CF já estar disponível, ainda há uma lacuna de *inputs* cruciais para o seu funcionamento. Esta lacuna provém da informação disponibilizada pelo MIBEL, uma vez que este não fornece o perfil de geração da PRO para o dia depois de amanhã ou os dados relativos à troca de energia entre Portugal e Espanha. Consequentemente, a falta desta informação não permite que o Operador de Sistema

seja capaz de prever os 24 *savecases* necessários para esse dia. Surge deste modo a necessidade de criar uma metodologia capaz de prever estes *inputs*.

1.2 – Objetivos da Tese

A presente tese tem como finalidade a apresentação de uma nova metodologia destinada à previsão do perfil de geração da PRO, das centrais hidroelétricas a fio-de-água, das centrais hidroelétricas com bombagem, e da energia elétrica exportada e importada, para o dia depois de amanhã. Pretendeu-se assim criar uma ferramenta capaz de fornecer ao Operador de Sistema a informação necessária relativamente à produção média de energia das centrais cuja produção para o dia depois de amanhã é ainda desconhecida e a quantidade de energia elétrica que é transacionada entre Portugal e Espanha.

O objetivo inicial do trabalho passa pela extração de conhecimento, utilizando técnicas de *Data Mining*, recorrendo assim a uma base de dados de informação disponibilizada pela REN (Rede Elétrica Nacional, S.A.), relativa a anos anteriores. Neste sentido, espera-se ser possível identificar padrões de correlação entre os dados em estudo. Ainda com o intuito de rentabilizar a extração de conhecimento, surge a necessidade de definir códigos de equivalência para a utilização da Teoria dos Conjuntos Aproximativos. Outro dos objetivos consiste na aplicação do pacote computacional ROSE2 para a obtenção de regras de decisão. O terceiro e último objetivo passa pela utilização deste pacote computacional na classificação/previsão da produção e troca de energia entre Portugal e Espanha para o dia depois de amanhã.

1.3 – Estrutura da Tese

A presente tese é constituída por seis capítulos, diretamente relacionados com os objetivos anteriormente anunciados. Apresenta no início um resumo e um *abstract* que sintetiza todo o trabalho concretizado. Após o índice, apresenta-se o índice de figuras, índice de tabelas e as abreviaturas utilizadas.

No presente capítulo, Capítulo 1, pretende-se fazer uma introdução ao trabalho, apresentando genericamente o trabalho realizado. Para além desta descrição da estrutura do documento, é

também feito o enquadramento e são delineados os objetivos pretendidos com a realização deste trabalho.

No Capítulo 2, é feita uma abordagem ao processo de Descoberta de Conhecimento em Bases de Dados. Destacou-se ainda a etapa de *Data Mining*, cuja importância na literatura tem tido um maior destaque, permitindo induzir facilmente uma descoberta de padrões para extrair conhecimento.

No Capítulo 3 são apresentados e explicados os conceitos relativos à Teoria dos Conjuntos Aproximativos, exemplificando através da sua aplicação a um Centro de Controlo.

No Capítulo 4 são enunciados alguns dos objetivos e fases da implementação do Mercado Ibérico de Eletricidade. São ainda destacados os diferentes operadores deste mercado, dando ênfase ao operador OMIP, responsável pela gestão do mercado de contratação a prazo e aos índices de eletricidade PTEL Base e PTEL Peak. É também feita uma análise do diagrama de cargas dos últimos anos, focando de forma mais detalhada os anos 2014 e 2015.

No Capítulo 5 são apresentados os resultados dos estudos efetuados para determinar as previsões para do dia D+2. Inicialmente são definidos todos os processos inerentes à aplicação da metodologia e as ferramentas para análise e interpretação dos resultados. Descrevem-se também os diferentes cenários aplicados e os códigos de equivalência. Cada cenário é aplicado utilizando o programa ROSE2, gerando conjuntos de regras para a previsão da produção das diferentes centrais e da exportação e importação para o dia depois de amanhã.

Por fim, no Capítulo 6 são apresentadas as conclusões gerais do presente trabalho, bem como as propostas para trabalhos futuros que visem a melhoria do trabalho desenvolvido.

Capítulo 2 – A Descoberta de Conhecimento em Bases de Dados

2.1 – Introdução

Atualmente, a facilidade de armazenar dados em diferentes formatos e base de dados tem contribuído para o aumento significativo de informação no mundo. Estes dados representam informação potencialmente útil que, no entanto, é raramente explícita ou aproveitada. É necessária uma aplicação de processos de análise inteligentes, visando a manipulação automática.

Para descobrir conhecimento é necessário extrair de grandes bases de dados, informações relevantes e desconhecidas, que podem ser utilizadas na tomada de decisões. Isto só é possível através do processo de Descoberta de Conhecimento em Bases de Dados.

As técnicas tradicionais de tratamento de dados não têm capacidade de manipular grandes quantidades de dados. A abordagem proporcionada pelas técnicas de *Data Mining* vem colmatar a lacuna existente neste tipo de desafios. Estas técnicas apresentam capacidade de inter-relacionar informações passadas com informações do presente, capazes de predeterminar acontecimentos futuros.

Data Mining ou mineração de dados é o processo de extração de informação válida, de um grande conjunto de dados, de forma a encontrar padrões, correlações e anomalias, com o objetivo de suportar a tomada de decisões fundamentais e proporcionar vantagens estratégicas. É um método que proporciona mais do que uma simples consulta de uma base de dados, uma vez que permite aos utilizadores explorar e inferir informação útil a partir de dados, descobrindo relacionamentos escondidos numa base de dados (Agreira, 2010).

2.2 – *Data Mining* e o processo de *Knowledge Discovery in Databases*

As técnicas de *Data Mining* são uma parte integral do processo de Descoberta de Conhecimento em Base de Dados, *Knowledge Discovery in Databases* (KDD). O termo KDD é usado para representar o processo geral de conversão de um conjunto de dados de baixo nível em conhecimento de alto nível, enquanto mineração de dados é o processo de aplicação de algoritmos específicos para extração de padrões ou modelos de dados observados. Segundo

Fayyad (Fayyad *et al*, 1996), o KDD pode ser definido como um processo de extração de conhecimentos válidos, novos, potencialmente úteis e compreensíveis, visando melhorar o entendimento de um problema.

No fundo, a mineração de dados é apenas uma etapa específica no conjunto de processos que constituem o KDD, como se pode observar na figura 2.1. Este processo engloba três grandes fases, nomeadamente, o Pré-processamento (Seleção, Pré-processamento, Transformação), *Data Mining* e o Pós-processamento (Interpretação e Consolidação) (Fayyad *et al*, 1996).

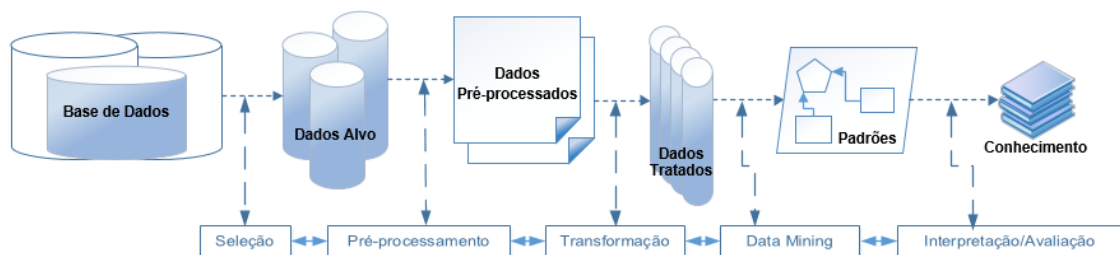


Figura 2.1 – *Knowledge Discovery in Databases*, adaptado de (Fayyad *et al*, 1996)

Com o acumular de informação armazenada, as organizações entenderam que não bastava apenas armazená-la, era necessário interpretá-la, analisá-la e relacioná-la a fim de extrair conhecimento de modo a sustentar a decisão. O KDD é um processo composto por etapas que devem ser desenvolvidas para atingir o objetivo final, a extração de conhecimento.

2.3 – Etapas do KDD

O processo KDD possui numerosos passos, que envolvem um número elevado de decisões a serem tomadas, sendo assim iterativo. É também um processo iterativo, ou seja, ao longo do processo KDD um passo será repetido as vezes que forem necessárias para que se chegue a um resultado satisfatório (Fayyad *et al*, 1996).

É necessário serem definidos os objetivos deste processo de descoberta de conhecimento. Isto inclui, analisar quais os dados disponíveis, obter dados adicionais (se necessário) e integrar todos estes dados para a descoberta de conhecimento numa base de dados. Este é um passo

fulcral, em que todos os atributos relevantes necessitam de estar incluídos, aumentando assim a possibilidade de sucesso do estudo.

Como referido, o KDD é composto pelas seguintes etapas: seleção dos dados, pré-processamento dos dados, transformação dos dados, mineração de dados (*Data Mining*) e interpretação/avaliação do conhecimento. Todas as etapas deste processo encontram-se interligadas, deste modo, os resultados obtidos numa etapa são utilizados para melhorar os resultados das próximas etapas.

2.3.1 – Seleção de Dados

A primeira etapa deste processo tem como objetivo a criação de um conjunto de dados-alvo. Nesta fase, seleciona-se um conjunto de dados ou foca-se num subconjunto de atributos (variáveis), no qual a descoberta será realizada (Fayyad *et al*, 1996).

Na seleção de dados escolhem-se apenas os atributos relevantes do conjunto de atributos da base de dados. Esta seleção de um subconjunto de atributos relevantes permite otimizar o tempo de processamento do algoritmo de *Data Mining*. Assim, este apenas trabalha com um subconjunto de atributos, diminuindo o seu espaço de pesquisa.

2.3.2 – Pré-processamento dos dados

As bases de dados são suscetíveis a ruídos, valores em falta e inconsistentes. De forma a assegurar a qualidade dos dados selecionados, é necessário fazer uma limpeza dos dados. Este processo envolve a verificação da consistência das informações, a correção de possíveis erros e o preenchimento ou a eliminação de valores nulos e redundantes. Nesta fase são identificados e removidos os dados, duplicados e/ou corrompidos. A execução desta fase corrige assim a base de dados, eliminando processos desnecessários que seriam executados pela mineração de dados, afetando o seu processamento (Sassi, 2006).

Os métodos de limpeza utilizados são dependentes do domínio da aplicação. Um exemplo seria a definição de um intervalo de possíveis valores para um determinado atributo, $\{0 \dots 10\}$. Caso surgisse um valor diferente dos definidos no intervalo, esse dado seria retirado.

2.3.3 – Transformação de dados

A fase de transformação de dados tem como objetivo principal converter o conjunto bruto de dados numa forma padrão de uso. Os dados de um atributo podem ser padronizados (normalizados), de forma a serem inseridos numa faixa de valores (escala), como por exemplo: -1,0 a 1,0 ou 0,0 a 1,0.

A transformação de dados para o atributo idade poderia ser feito da seguinte forma: {0...18}, Intervalo 1; {19...25}, Intervalo 2; {25...30} Intervalo 3, e assim sucessivamente. Os valores contínuos das idades são assim discretizados em intervalos.

Este processo de transformação ou codificação de dados permite melhorar a compreensão do conhecimento descoberto; reduzir o tempo de processamento para o algoritmo de *Data Mining* e facilitar o algoritmo a tomar decisões. Apesar destas vantagens, esta etapa ao compactar de certa forma os dados, leva à perda de informação relevante extraída.

2.3.4 – Data Mining

Apesar de todas as etapas do KDD serem fundamentais, o *Data Mining* destaca-se neste processo, apresentando a maior importância na literatura. *Data Mining* refere-se à exploração e análise, por meio automático ou semiautomático, de grandes quantidades de dados, a fim de descobrir padrões e regras relevantes (Berry e Linoff, 1997). As técnicas de mineração de dados têm como principal objetivo encontrar padrões e relações entre dados (regras), de forma a auxiliar na previsão, decisão e estimativa de futuros acontecimentos.

O *Data Mining* usa técnicas baseadas em descobertas por meio de procura de padrões nos dados, o que é feito recorrendo a uma série de algoritmos inteligentes para encontrar relações fundamentais entre os dados. Os seus principais objetivos são descobrir relacionamentos entre dados e fornecer auxílios para que possam ser feitas previsões de tendências futuras baseadas no passado. Deste modo, permite a tomada de decisão baseada em fatos e não em suposições.

2.3.5 – Interpretação/ Avaliação do Conhecimento

Esta etapa é também conhecida como pós-processamento. Após a mineração de dados, é necessária a interpretação do conhecimento descoberto, ou algum processamento desse conhecimento. No fundo, nesta fase pretende-se melhorar a compreensão do conhecimento

descoberto pelo algoritmo aplicado. Pode ser necessário retornar a uma das etapas anteriores, remover padrões redundantes ou irrelevantes, sendo que os resultados relevantes devem ser traduzidos numa linguagem compreensível para os utilizadores (Fayyad *et al*, 1996).

As técnicas de visualização são uma forma utilizada para obter a compreensão e interpretação dos resultados. Estas estimulam a percepção e a inteligência humana, aumentando a capacidade de entendimento e a associação de novos padrões (Sassi, 2006).

2.3 – Origem dos dados

Para utilizar as técnicas de mineração de dados, é necessário recorrer a enormes bases de dados. É possível rentabilizar a aplicação destas técnicas, utilizando bases de dados projetadas para esse fim, tais como *Data Warehouse* (DW) ou *Data Mart*. Os dados provêm de bases de dados operacionais, convertidos num formato adequado para serem armazenados no DW. Através deste processo de conversão, geralmente obtém-se uma base de dados mais consistente, organizada e clara. Estas técnicas de mineração de dados podem também ser aplicadas sobre um *data set*, que pode ser definido como uma “base de dados” (num sentido fraco do termo) contendo apenas um conjunto específico de dados, para um determinado tipo de investigação a ser realizada (Agreira, 2010).

Segundo W. H. Inmon, considerado um pioneiro neste tema, “um DW é um conjunto de dados baseado em assuntos, integrado, variante em relação ao tempo, não volátil, que tem por objetivo dar apoio aos processos de tomada de decisão” (Inmon, 2002).

Um DW é no fundo uma base de dados que fornece e centraliza dados de diferentes fontes. Deste modo, é possível ter uma base de dados mais robusta, consistente, oferecendo organização e processamento de bases de dados. A sua utilização permite auxiliar na tomada de decisões mais rápidas e eficazes, proporcionando às organizações uma melhoria dos processos atuais e futuros. Uma *Data Mart* é um subconjunto de dados de um DW, ou seja, é um DW departamental, contendo informação destinada a um departamento em particular (Inmon, 2002).

Os dados contidos no *Data Mart* são mais compactos, sumarizados e relacionais ou multidimensionais não voláteis. Esta base de dados utiliza uma estrutura de dados moldada aos

requisitos específicos de cada departamento, facilitando a toma de decisões ao nível departamental (Inmon, 2002).

2.5 – Tarefas da mineração de dados

A aplicação das técnicas de mineração de dados tem como objetivo atingir um determinado resultado. Este resultado só será alcançado caso sejam aplicadas as técnicas de mineração às tarefas adequadas, de modo a atingir os objetivos pretendidos. As tarefas subdividem-se em dois grupos: previsão, responsável pela procura constante de padrões que permitam prever situações/comportamentos e a descrição que assenta numa procura de padrões que apresentem o conhecimento de forma compreensível.

Para o objetivo de previsão, as técnicas mais comuns são: segmentação (ou *clustering*), associação e sumarização. Por outro lado, para o objetivo de descrição, assenta na classificação e regressão (ou estimativa). Estas técnicas são descritas de uma forma mais detalhada nos seguintes pontos.

2.5.1 – Classificação

Classificação é uma das tarefas mais comuns na mineração de dados. A tarefa de classificação consiste em determinar a que classe um novo objeto pertence, de acordo com um modelo de classificação e dado um conjunto de classes predefinidas (Goebel e Gruenwald, 1999).

Inicialmente constroem-se modelos de classificação, com base num conjunto de dados de treino, para descrever as classes, que posteriormente são aplicados a dados não classificados. A classe de mamíferos, por exemplo, possui atributos que descrevem o mamífero; se um animal satisfizer as propriedades de classificação do mamífero, então este pode ser classificado como mamífero (Santos e Azevedo, 2005).

A aplicação da tarefa de classificação pode ser aplicada em diversas áreas, tais como deteção de fraude; diagnósticos médicos, identificando a forma de tratamento na qual um paciente está mais propício a reagir positivamente, baseando-se em classes de pacientes que reagem bem a determinado tipo de tratamento médico; ou classificar pedidos de crédito como de baixo, médio e alto risco (Goebel e Gruenwald, 1999).

2.5.2 – Regressão (ou Estimativa)

A regressão ou estimativa é a análise da dependência de alguns atributos de um objeto, em função dos seus restantes atributos, conduzindo à produção automática de um modelo capaz de prever os valores destes atributos para novos objetos (Goebel e Gruenwald, 1999). A classificação trata resultados de carácter discreto: sim ou não; sarampo, rubéola ou varicela. A regressão lida com resultados contínuos, ou seja, através de um conjunto de dados de *input*, a regressão determina qual o valor da variável contínua desconhecida, tal como a altura ou o saldo do cartão de crédito (Berry e Linoff, 1997). De uma forma mais sucinta, a regressão mapeia um *item* de dados numa variável de previsão, de valor real (Fayyad *et al*, 1996).

A regressão ou estimativa é aplicada em vários casos, tais como: previsão da quantidade de biomassa presente numa floresta, com base nas medições de micro-ondas feitas remotamente; estimar a probabilidade de um paciente sobreviver, dado um conjunto de diagnósticos médicos; prever a demanda de um produto novo, por um consumidor, em função das despesas de publicidade (Fayyad *et al*, 1996).

2.5.3 – Segmentação ou *Clustering*

Clustering é a tarefa de segmentação de uma população heterogénea em subgrupos ou *clusters* mais homogéneos. Segmentação distingue-se da classificação por esta não depender de classes predefinidas, os registos são agrupados com base na sua semelhança.

Este tipo de tarefa por vezes precede outra tarefa de *Data Mining*, por exemplo: identificar subgrupos de clientes que têm hábitos de compra semelhantes e posteriormente perguntar qual as promoções que são mais adequadas para cada subgrupo (Berry e Linoff, 1997).

2.5.4 – Associação ou Dependência

A tarefa de associação ou dependência permite identificar a relação entre dados aparentemente não relacionáveis, i.e., encontrar um modelo que descreva dependências significativas entre variáveis (itens tipicamente associados) (Santos e Azevedo, 2005).

Um exemplo clássico da aplicação desta tarefa é na determinação de qual o produto que um cliente irá adicionar ao seu carrinho de compras, tendo por base o último produto adicionado,

esta tarefa é denominada de *Market Basket Analysis*. Esta tarefa de associação permite que seja possível efetuar mudanças na disposição dos produtos nas prateleiras das lojas ou num catálogo, criar pacotes de produtos apelativos e agrupar produtos e serviços num só pacote (Berry e Linoff, 1997).

2.5.5 – Sumarização

Segundo Fayyad (Fayyad *et al*, 1996), a tarefa de sumarização envolve métodos para encontrar uma descrição compacta para um subconjunto de dados. Os métodos de sumarização mais sofisticados derivam de regras de resumo, técnicas de visualização multivariadas e das descobertas de relações funcionais entre variáveis.

As técnicas de sumarização são frequentemente aplicadas à análise exploratória de dados e à geração automática de relatórios.

2.6 – Modelos e Técnicas de *Data Mining*

Data Mining não é um processo com uma só metodologia ‘universal’ que resolva de forma eficiente todos os problemas. Segundo (Fayyad *et al*, 1996), a escolha do conjunto mais adequado é no fundo uma arte, uma vez que existem vários modelos e técnicas para as mesmas tarefas de mineração de dados, com vantagens e desvantagens inerentes.

Um modelo é definido como a representação matemática de um sistema (ou processo) com o objetivo de o estudar, sendo a descrição formal e simplificada de um sistema, ou seja, a estruturação da representação do conhecimento com vista a atingir um determinado resultado (Santos e Azevedo, 2005).

Como técnica compreende-se o conjunto de processos baseados em conhecimento científico, de cálculo ou experimentação, utilizados para a obtenção de um resultado.

2.6.1 – A aprendizagem automática

A Aprendizagem Automática (AA) visa criar algoritmos e técnicas que permitam melhorar o desempenho das máquinas em determinadas tarefas, nomeadamente através do conceito “experiência”. O tipo de AA utilizado em *Data Mining* corresponde à aprendizagem indutiva,

i.e., a criação de um padrão (modelo) aplicável a novos casos. Definem-se as principais características do sistema, que devem representar o mais fielmente possível a realidade e recolhem-se os dados necessários para a sua construção e posterior validação. A este conjunto de dados são aplicadas as técnicas de amostragem em que há uma divisão do conjunto de dados em dois subconjuntos, um de treino para ajustar ou induzir o modelo e outro de teste para efeitos de validação. Ao modelo estão associados algoritmos para identificação de padrões e relacionamentos. Após a construção do modelo, aplica-se o conjunto de dados de validação de forma a validar o modelo criado (Santos e Azevedo, 2005).

2.7 – Conclusão

O presente capítulo procurou descrever todo o processo de KDD através da análise das suas tarefas, etapas e origens dos dados. Destacou-se a etapa de *Data Mining*, que embora as restantes etapas sejam de igual importância, esta tem tido um maior destaque na literatura. A aplicação de apenas esta etapa induz facilmente a uma descoberta de padrões inválidos e inconclusivos. A descrição deste processo contribuiu para fundamentar as várias etapas do processo utilizado para o desenvolvimento deste trabalho.

O termo Descoberta de Conhecimento em Base de Dados veio realçar que o conhecimento é o produto final de uma descoberta baseada em dados. KDD evoluiu e continua a evoluir, através da sua aplicação em distintas áreas de estudo, permitindo extrair conhecimento de alto nível de uma extensa base de dados de baixo nível.

Capítulo 3 – Teoria dos Conjuntos Aproximativos

3.1 – Introdução

Devido à imensa variedade de modelos e técnicas de *Data Mining* existentes, torna-se difícil escolher a mais apropriada para resolver um determinado problema em questão. As técnicas mais utilizadas são: Árvores de Decisão, Redes Neuroniais, Indução de Regras, Conjuntos Difusos, Redes de Bayes, Sistemas de Classificação e Conjuntos Aproximativos.

Optou-se por utilizar a Teoria dos Conjuntos Aproximativos (*Rough Sets Theory*), proposta por Zdzislaw Pawlak, em 1982. Tendo surgido a partir de estudos em Sistemas de Informação, esta é uma nova ferramenta matemática, em constante desenvolvimento, dedicada à classificação e análise de informação imprecisa, incerta ou incompleta (Pawlak, 1982) (Pawlak, 1995) (Cios *et al*, 1998).

A Teoria dos Conjuntos Aproximativos (TCA) apresenta como principal vantagem o facto de não necessitar de informação preliminar ou adicional sobre a informação a tratar, tais como são necessários a distribuição de probabilidade na Estatística, atribuição básica de probabilidade na Teoria de Dempster-Shafer e na Teoria de Bayes, e o grau de pertinência ou valor de possibilidade na Teoria de Conjuntos Difusos (Rissino e Lambert-Torres, 2009).

A sua utilização tem sido preciosa em vários problemas tais como: na representação de conhecimento incerto ou impreciso; na análise do conhecimento; avaliação da qualidade e coerência das informações; e em áreas como a medicina, análise de atributos de base de dados e no controlo de processos (Rissino e Lambert-Torres, 2009).

A TCA utiliza informação associada a objetos ou acontecimentos. Um conjunto de objetos caracterizados com a mesma informação são denominados de indiscerníveis (semelhantes), dada a informação conhecida sobre eles. É esta a filosofia que sustenta a TCA, descobrir todos os objetos que possuem o mesmo tipo de informação (atributos), ou seja, que são indiscerníveis. A partir deste conceito foram então geradas as bases matemáticas desta teoria.

Esta teoria tem sido aplicada em vários domínios, tais como no apoio à decisão, meio ambiente, medicina, engenharia e outros. Na engenharia eletrotécnica destaca-se o seu papel nas mais variadas áreas ligadas aos Sistemas Eléctricos de Energia. A TCA tem sido usada para vários estudos que vão desde a análise de segurança até a estudos de risco financeiro nos novos

mercados de energia, passando por aplicações como a classificação de distúrbios na qualidade da energia, diagnóstico de falhas nos transformadores, entre outros (Pisicâ e Postolache, 2010).

3.2 – Conceitos Fundamentais da Teoria dos Conjuntos Aproximativos

A premissa central da TCA parte da filosofia de que o conhecimento consiste na habilidade de classificar objetos. Ao fazer isso é possível formar classes de objetos indiscerníveis, usados para construir conhecimento sobre um mundo real abstrato. Qualquer conjunto de todos os objetos indiscerníveis é denominado de conjunto elementar, e forma um átomo de conhecimento sobre o universo. Qualquer união de alguns conjuntos elementares é referenciado como um conjunto exato, caso contrário, o conjunto é *rough* (impreciso, vago) (Pawlak, 2002).

A análise de dados aplicando a TCA tem origem numa tabela de dados, denominado de sistema de informação. Esta tabela de informação contém dados sobre objetos (acontecimentos, estados, contingências, ...), correspondentes às linhas da tabela e caracterizados por um conjunto finito de atributos (características, condições, capacidades, potências, ...), correspondentes às colunas da tabela. A cada par (objeto, atributo) corresponde um valor designado por descritor.

Analisando a tabela de informação é possível descobrir relações de dependência (padrões). Com este intuito, distingue-se um conjunto de atributos de condição C e um conjunto de atributos de decisão D , de modo a analisar de que forma os valores dos atributos de C se relacionam com os valores dos atributos D . Uma tabela de informação onde os atributos de condição e os atributos de decisão são diferenciados, é denominada tabela de decisão. Desta tabela de decisão é possível deduzir padrões no seguinte formato, “*se ..., então ...*”, criando assim regras de decisão. Mais precisamente, estas regras definem que, a um determinado valor de um atributo de condição, corresponde um valor de um atributo de decisão específico (Pawlak, 2002).

3.3 – Sistema de Informação ou Tabela de Informação

Um sistema de informação ou tabela de informação pode ser visto como uma tabela, constituída por objetos (exemplos, pacientes, estados, ...) e atributos (características, variáveis, qualificadores, ...). Este é o ponto de partida para a representação de dados e pode ser representado por $S = (U,A)$, onde:

$U \rightarrow$ é um conjunto finito de objetos $\{\text{Obj}_1, \text{Obj}_2, \text{Obj}_3, \dots\}$, designados Universo e representados numa tabela como sendo as linhas;

$A \rightarrow$ é um conjunto finito de atributos $\{A_1, A_2, A_3, \dots\}$, representados numa tabela como sendo as colunas;

A cada atributo $a \in A$ está associado um conjunto (*set*) de possíveis valores que qualquer objeto pode ter, denominado de domínio de a e representado por V_a . Considere-se:

$$V = \cup_{a \in A} V_a, V_a \text{ é o domínio do atributo } a.$$

Todo o atributo $a \in A$, sendo a uma função total tal que $a: U \rightarrow V_a$ onde:

$$V_a \rightarrow \text{é o conjunto de valores de } a, \text{ chamado o domínio de } a.$$

A seguinte função de informação:

$$f: U \times A \rightarrow V \text{ é uma função total tal que } f(x, a) \in V_a \text{ para todo o } a \in V_a, x \in U.$$

Esta função associa cada par (objeto x , atributo a) ao valor correspondente do seu domínio V_a .

Qualquer par (a, v) , $a \in A$, $v \in V_a$ é denominado descritor em S .

A estrutura $S = \langle U, A, V, f \rangle$ é definida como um sistema de informação. No fundo, um sistema de informação não é mais do que, uma tabela de dados, finita, em que as colunas representam os atributos, e as linhas representam os objetos. A entrada na coluna a e na linha x tem o valor de $f(x, a)$. O Sistema de Informação também é conhecido por Sistema de Representação do Conhecimento.

3.4 – Relação de indiscernibilidade

A TCA mede a similaridade entre dois ou mais objetos através da relação de indiscernibilidade. Esta é no fundo uma relação de equivalência que particiona o universo, U , de acordo com classes. Estas classes são rótulos atribuídos a elementos que possuem determinadas características em comum, expressas por meio de um atributo de decisão. Isto é, dado um sistema de informação $S = \langle U, A, V, f \rangle$ e $P \subseteq A$, define-se que dois objetos $x, y \in U$ são indiscerníveis para o conjunto de atributos P em S se, e somente se, $f(x, a) = f(y, a)$, para todo o $a \in P$. Desta forma, x e y são indiscerníveis em P se possuírem os mesmos valores para todos os atributos em P . Assim, todo

o $P \subseteq A$ gera uma relação binária em U , à qual é denominada de relação de indiscernibilidade, denotada como $IND(P)$ e definida pela seguinte equação:

$$IND(P) = \{(x,y) \in U \times U : \forall a \in P, a(x) = a(y)\}$$

$IND(P)$ é uma relação de equivalência para qualquer P . O conjunto de todas as classes de equivalência da relação $IND(P)$ é representado por $U/IND(P)$, denominado quociente de U pela relação $IND(P)$. Este pode também ser abreviado por uma notação mais simples, U/P , ou ainda por P^* . Este conjunto de classes é também denominada por conhecimento P-básico de U em S . Tal simplificação justifica-se pelo facto de que a cada conjunto elementar pode-se associar exatamente uma única descrição e vice-versa (Pawlak, 1991).

As classes de equivalência da relação $IND(P)$ são denominadas conjuntos P-elementares (conceitos ou categorias do conhecimento P-básico) em S .

$Desp(X)$ denota a descrição do conjunto P-elementar $X \in U/IND(P)$ em termos dos pares (atributo, valor), isto é:

$$Desp(X) = \{(a,v) : f(x,a) = v, \forall x \in X, \forall a \in P\}$$

3.5 – Aproximação de um conjunto

A ideia de relação de equivalência induz a partição do Universo. Estas partições podem ser usadas para construir novos subconjuntos do universo. Por vezes, os elementos são indiscerníveis e têm valores de decisão contraditórios, não sendo possível induzir uma descrição precisa. Surge aqui a noção de conjunto aproximativo ou *rough set*.

Supondo um dado sistema de informação $S = (U,R)$, a cada subconjunto $X \subseteq U$ a relação de equivalência $R \in IND(S)$ associa três subconjuntos, denominados de Aproximação Inferior ou Região Positiva, Aproximação Superior ou Região Negativa e a Região Fronteira (Pawlak, 1991).

- Aproximação P-inferior $\underline{RX} = \cup\{Y \in U/R : Y \subseteq X\}$
- Aproximação P-superior $\overline{RX} = \cup\{Y \in U/R : Y \cap X \neq \emptyset\}$
- Região P-fronteira $BN_R(X) = \underline{RX} - \overline{RX}$

O que significa que os elementos que pertencem à Aproximação Inferior \underline{RX} são classificados, com certeza, como elementos de X . Os elementos da Aproximação Superior \overline{RX} podem ser classificados como possíveis elementos de X (Pawlak, 1991).

Há cinco regiões (ou conjuntos) de interesse: \overline{RX} e \underline{RX} , $POS_R(X)$, $BN_R(X)$ e $NEG_R(X)$. Estes conjuntos são definidas da seguinte forma (Lambert-Torres, 2002):

\underline{RX} é o conjunto de todos os elementos de U , que podem certamente ser classificados como elementos de X empregando o conjunto de atributos P .

\overline{RX} é o conjunto de elementos de U , que podem possivelmente ser classificados como elementos de X empregando o conjunto de atributos P .

$BN_R(X)$ é o conjunto de elementos que não podem ser certamente classificados como elementos de X , empregando o conjunto de atributos P .

São também definidos segundo os seguintes termos (Pawlak, 1991):

- Região P-positiva $POS_R(X) = \underline{RX} \Rightarrow$ certamente membro de X ;
- Região P-negativa $NEG_R(X) = U - \overline{RX} \Rightarrow$ certamente não é membro de X ;
- Região P-fronteira $BN_R(X) = \overline{RX} - \underline{RX} \Rightarrow$ possivelmente membro de X .

A figura 3.1 representa o conjunto das aproximações a R e as regiões de R .

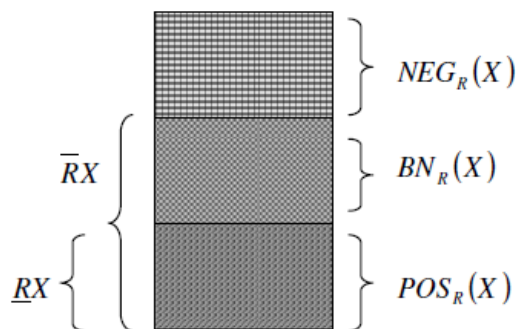


Figura 3.1 – Conjunto das aproximações a R e as regiões de R (Lambert-Torres, 2002).

Um conjunto X é definido como *rough* (impreciso) se a sua Região de Fronteira for diferente do conjunto vazio ($BN_R(X) \neq 0$). No entanto, se o conjunto for vazio ($BN_R(X) = 0$), o conjunto X é definido como *crisp* (exato).

A inexatidão de um conjunto deve-se à existência de uma região fronteira. Quanto maior for essa região, menor a precisão do conjunto. A precisão da aproximação do conjunto $X \subseteq U$ por R em K ou, simplesmente precisão de X é dada por (Pawlak, 1991):

$$\alpha_R(X) = \frac{\text{card}(\underline{RX})}{\text{card}(RX)} \quad (3.1)$$

A precisão de X poderá tomar valores entre $0 \leq \alpha_R \leq 1$, para cada R e $X \subseteq U$. Se $\alpha_R = 1$, X é considerado *crisp* em relação ao conjunto de atributos R . Se $\alpha_R < 1$, X é *rough*, em relação ao conjunto de atributos R .

3.6 – Redução de atributos

Nos sistemas de representação de conhecimento, um dos problemas fundamentais é o de encontrar subconjuntos do conjunto original de atributos com o mesmo poder discriminatório. A TCA é utilizada para a análise de dependência entre atributos, visando principalmente identificar e eliminar os redundantes ou supérfluos. A eliminação dos redundantes permite que se encontre subconjuntos mínimos de atributos (denominados de redutos) que possuam o mesmo valor discriminatório do conjunto original.

Na TCA, a obtenção de redutos está intimamente ligada à análise de dependência entre atributos. O conjunto de atributos $R \subseteq Q$ depende do conjunto de atributos $P \subseteq Q$ em K (notação $P \rightarrow R$) se $IND(P) \subseteq IND(R)$. A descoberta de relações de dependência entre atributos é de extrema importância para análise de dados no contexto dos conjuntos aproximativos (Agreira, 2010).

A redução de atributos é, normalmente, conseguida de modo que o conjunto reduzido de atributos possua a mesma qualidade de classificação do conjunto original de atributos. O

subconjunto mínimo $R \subseteq P \subseteq Q$, tal que, $\gamma_P(X) = \gamma_R(X)$ é denominado reduto. O reduto é também chamado de conjunto mínimo (ou subconjunto) em K .

É de salientar que um sistema de informação pode ter mais do que um reduto, conseguindo-se representar a base de dados com todos os seus atributos originais, independentemente do reduto escolhido. $RED_R(P)$ representa a família de todos os redutos de P . A intersecção de todos os redutos de P é designada de *core* de P , i.e., $CORE_X(P) = \cap RED_X(P)$, onde $RED(P)$ representa a família de todos os redutos de P (Agreira, 2010).

O *core* é o conjunto dos atributos mais importantes, os quais não podem ser eliminados do sistema K sem diminuir a qualidade da aproximação da classificação X .

3.7 – Tabelas de decisão

Um sistema de informação pode ser analisado como uma tabela de decisão, contendo dois tipos de atributos, designados de atributos de condição e atributos de decisão. Tem-se então, que $Q = C \cup D$ e $C \cap D = \emptyset$, onde C representa o subconjunto de Q formado por atributos que são condições ou conjuntos de condições (atributos condicionais) e D o subconjunto de Q formado por atributos que são decisões ou conjuntos de decisões (atributos decisão). As tabelas de decisão apresentam como principal vantagem a possibilidade de representação de decisões num dado sistema para o qual ainda não existem regras explícitas para a tomada de decisão (Agreira, 2010).

A tabela de decisão $DT = \langle U, C \cup D, V, f \rangle$ pode ser determinística (consistente) ou não-determinística (inconsistente). Os atributos D dependem totalmente de C ($C \rightarrow D$), se todos os valores nominais de D forem univocamente determinados por valores nominais dos atributos de C , classificando-se como determinística, caso contrário, é não-determinística. No caso de uma tabela não-determinística, as decisões não são unicamente determinadas pelas condições. Ao contrário, é definido um subconjunto de decisões que podem ser tomadas de acordo com determinadas circunstâncias. As decisões podem ainda depender parcialmente das condições. Neste caso, não existe a necessidade da presença de todos os atributos condicionais para gerar os valores nominais de D , abrindo espaço para uma redução de atributos, como referido no ponto 3.7 (Agreira, 2010).

Através da tabela de decisão é possível derivar um conjunto de regras de decisão. Seja $U/IND(C)$ uma família de todos os conjuntos C – elementares chamados classes condicionais em DT , com notação X_i ($i=1, \dots, k$). Seja também, $U/IND(D)$ uma família de todos os conjuntos D – elementares chamados classes de decisão em DT , e com notação Y_j ($j=1, \dots, n$).

$Des_C(X_i) \Rightarrow Des_D(Y_j)$ é chamada regra de decisão (C,D) .

As regras são afirmações lógicas do tipo (se...então...) relacionando descrições das classes condicionais e de decisão. O conjunto de regras de decisão para cada classe de decisão Y_j ($j=1, \dots, n$) tem a notação $\{r_{ij}\}$. Assim:

$$\{r_{ij}\} = \{Des_C(X_i) \Rightarrow Des_D(Y_j) : X_i \cap Y_j \neq \emptyset, i=1, \dots, k\}$$

A regra r_{ij} é determinística (consistente ou certa) em DT se $X_i \subseteq Y_j$.

A regra r_{ij} é não-determinística (inconsistente) em DT , no caso contrário.

O conjunto das regras de decisão para todas as classes de decisão é denominado algoritmo de decisão.

3.8 – Aplicação da Teoria dos Conjuntos Aproximativos a um Centro de Controlo

A TCA tem vindo a destacar-se nas mais diversas áreas dos Sistemas Eléctricos de Energia. A sua aplicação implica a passagem por várias etapas. Como exemplo da utilização desta teoria, recorreu-se à sua aplicação num Centro de Controlo de Sistemas Eléctricos de Energia, tendo como objetivo a redução do volume de informação que as bases de dados possuem nos centros de controlo. Através da aplicação da TCA, o número de atributos inicialmente considerado é reduzido e o Sistema de Informação é classificado segundo um conjunto de regras em três estados de segurança (Lambert-Torres *et al*, 1996a), (Lambert-Torres *et al*, 1996b), (Lambert-Torres, 2002).

Para que o exemplo seja de fácil compreensão, Germano Lambert-Torres *et al*, adotou um sistema com poucos atributos e poucos casos (Lambert-Torres *et al*, 1996a) (Lambert-Torres, 2002). Contudo, para um sistema de informação real, os procedimentos são os mesmos.

O sistema de informação é composto por um conjunto de medidas e informação analógica como pode ser visualizada na tabela 3.1. Esta contém como objetos o número de casos que são 25. Os

atributos são representados pelo conjunto $\{A, \dots, H\}$, e a decisão é representada por S (Lambert-Torres *et al*, 1996a), (Lambert-Torres *et al*, 1996b).

Os atributos da tabela 3.1 representam o seguinte:

- A, B e $C \rightarrow$ são a percentagem do trânsito de potências real nas linhas de transmissão de acordo com a capacidade nominal, em [%];
- $D, E,$ e $F \rightarrow$ são as tensões nos barramentos em [p.u.];
- G e $H \rightarrow$ são o estado do circuito (0 – aberto, 1 – fechado).

A classificação do estado, representado por S , é usualmente feita de acordo com um perito (usualmente um operador sénior / engenheiro), e os três possíveis *outputs* (S, U_1 e U_2) podem ser seleccionados como estado operacional do Sistema Eléctrico de Energia, onde:

- $S \rightarrow$ Seguro
- $U_1 \rightarrow$ Inseguro de nível 1
- $U_2 \rightarrow$ Inseguro de nível 2

Tabela 3.1 – Conjunto de informação reduzido de um sistema de controlo (Lambert-Torres *et al*, 1996a)

Nº	Atributos								S
	A	B	C	D	E	F	G	H	
1	65	45	29	0,91	1,06	0,95	1	1	S
2	35	67	63	1,08	0,82	1,06	0	1	U2
3	47	57	32	0,98	1,07	0,99	1	1	S
4	61	32	38	0,89	1,07	0,94	1	1	S
5	39	75	57	1,06	0,99	1,07	0	1	U1
6	29	65	60	1,07	0,80	1,06	1	1	U1
7	27	45	32	1,07	1,00	1,06	0	1	U2
8	23	76	47	0,98	0,95	0,95	0	1	S
9	61	35	37	0,99	1,06	1,03	1	1	S
10	34	83	65	1,07	1,03	1,06	0	1	U1
11	32	45	33	1,07	1,03	1,07	0	1	U2
12	35	65	23	1,07	0,97	1,06	0	1	U2
13	57	78	54	0,84	0,82	0,84	1	1	U1
14	32	48	49	1,07	0,83	1,06	0	1	U2
15	65	36	32	0,93	1,06	1,03	1	1	S
16	49	79	25	0,87	1,07	1,00	1	1	S
17	21	65	49	1,08	0,82	1,06	1	1	U1
18	32	65	56	1,06	1,01	1,06	0	1	U2
19	34	45	65	1,07	0,84	1,08	1	1	U1
20	65	60	38	1,00	1,07	1,01	1	1	S
21	22	49	52	1,08	0,83	1,07	1	1	U1
22	82	67	69	0,82	1,06	0,83	1	1	U2
23	32	45	67	1,07	0,84	1,07	1	1	U1
24	75	47	49	0,97	1,06	0,98	1	1	S
25	32	72	49	0,98	0,99	1,01	0	1	S

Primeiro Passo:

O primeiro passo do algoritmo consiste em redefinir os valores de cada atributo de acordo com um conjunto de equivalências. Neste exemplo, foram usados valores típicos de operação nos Sistemas Elétricos de Energia (Lambert-Torres *et al*, 1996a):

- Valores reais da potência:
 - menos de 40% da capacidade nominal = Baixo (B)
 - entre 40% e 70% = Médio (M)
 - mais de 70% da capacidade nominal = Alta (A)

- Valores reais da tensão nos barramentos:
 - menos de 0.85 p.u. = Baixo (B)
 - entre 0.85 e 1.05 p.u. = Normal (N)
 - mais de 1.05 p.u. = Alto (A)

O estado dos disjuntores é mantido. Usando então estas redefinições para este exemplo, é possível reescrever a tabela 3.1, o que pode ser visualizado na tabela 3.2.

Tabela 3.2 – Sistema de informação equivalente (Lambert-Torres *et al*, 1996a)

Nº	Atributos								S
	A	B	C	D	E	F	G	H	
1	M	M	B	N	A	N	1	1	S
2	B	M	M	A	B	A	0	1	U2
3	M	M	B	N	A	N	1	1	S
4	M	B	B	N	A	N	1	1	S
5	B	A	M	A	N	A	0	1	U1
6	B	M	M	A	B	A	1	1	U1
7	B	M	B	A	N	A	0	1	U2
8	B	A	M	N	N	N	0	1	S
9	M	B	B	N	A	N	1	1	S
10	B	A	M	A	N	A	0	1	U1
11	B	M	B	A	N	A	0	1	U2
12	B	M	B	A	N	A	0	1	U2
13	M	A	M	B	B	B	1	1	U1
14	B	M	M	A	B	A	0	1	U2
15	M	B	B	N	A	N	1	1	S
16	M	A	B	N	A	N	1	1	S
17	B	M	M	A	B	A	1	1	U1
18	B	M	M	A	N	A	0	1	U2
19	B	M	M	A	B	A	1	1	U1
20	M	M	B	N	A	N	1	1	S
21	B	M	M	A	B	A	1	1	U1
22	A	M	M	B	A	B	1	1	U2
23	B	M	M	A	B	A	1	1	U1
24	A	M	M	N	A	N	1	1	S
25	B	A	M	N	N	N	0	1	S

Segundo e Terceiro Passo:

Considerando o conjunto de atributos A, B e C, notamos que os objetos 1, 3 e 20 têm os mesmos valores (M, M e B) para estes três atributos. Por esta razão, 1, 3 e 20 são indiscerníveis em relação a este conjunto de atributos. Os objetos 2 e 6 são, igualmente, indiscerníveis entre si.

A relação de indiscernibilidade é, então, uma relação de equivalência.

Os elementos que são indiscerníveis formam conjuntos denominados de conjuntos elementares. Assim sendo, na tabela 3.2, o conjunto de atributos A, B, e C definem sete conjuntos elementares:

$$\begin{aligned} & \{1, 3, 20\} \\ & \{2, 6, 14, 17, 18, 19, 21, 23\} \\ & \{4, 9, 15\} \\ & \{5, 8, 10, 25\} \\ & \{7, 11, 12\} \\ & \{22, 24\} \\ & \{13\} \\ & \{16\} \end{aligned}$$

A qualquer união finita de conjuntos elementares chama-se conjunto definido. Por exemplo, o conjunto $\{2, 6, 13, 14, 17, 18, 19, 21, 23\}$ é definido pelos atributos A, B e C. Qualquer elemento deste conjunto é caracterizado pelo atributo A = Baixo, B = Médio e C = Médio ou pelo atributo A = Médio, B = Alto, C = Médio (Agreira, 2010).

O seguinte passo do algoritmo tem como objetivo verificar se algum atributo pode ser eliminado por repetição. Na tabela 3.2 pode-se verificar que os atributos D e F são idênticos para todos os exemplos. Desta forma, um dos dois pode ser eliminado. Existem outros exemplos idênticos, como por exemplo os objetos $\{6, 17, 19, 21, 23\}$ e $\{1, 3, 20\}$. Os exemplos semelhantes são então suprimidos e surge na tabela 3.3 o conjunto resultante dos objetos (Lambert-Torres *et al*, 1996a).

Tabela 3.3 – Conjunto resultante dos exemplos (Lambert-Torres *et al*, 1996a)

Nº	Atributos							S
	A	B	C	D	E	G	H	
1	B	M	M	A	B	0	1	U2
2	B	M	B	A	N	0	1	U2
3	B	M	M	A	N	0	1	U2
4	A	M	M	B	A	1	1	U2
5	B	M	M	A	B	1	1	U1
6	M	A	M	B	B	1	1	U1
7	B	A	M	A	N	0	1	U1
8	M	B	B	N	A	1	1	S
9	M	M	B	N	A	1	1	S
10	M	A	B	N	A	1	1	S
11	A	M	M	N	A	1	1	S
12	B	A	M	N	N	0	1	S

Quarto Passo:

A relação de indiscernibilidade é utilizada para definir atributos redundantes, ou seja, atributos dispensáveis. Designa-se de conjunto mínimo (ou independente) o conjunto de atributos em que nenhum é redundante.

Este passo tem como objetivo verificar se a tabela de decisão contém apenas os atributos indispensáveis. Esta tarefa pode ser realizada eliminando passo a passo cada atributo e verificando se a tabela continua a dar a classificação correta.

Por exemplo, se o atributo A for eliminado, a tabela continua a dar a classificação correta. Pode-se então dizer que A é um atributo dispensável para esta tabela de decisão.

Contudo, quando o atributo B é eliminado (tabela 3.4), pode-se verificar que os objetos 3 e 7 têm o mesmo conjunto de atributos, mas apresentam classificações diferentes. Neste caso, pode-se dizer que o atributo B é indispensável para esta tabela de decisão.

Findo este passo para todos os atributos, pode-se concluir que os atributos B, D, e G são indispensáveis e os atributos A, C, E, e H são dispensáveis para esta tabela de decisão.

Tabela 3.4 – Indispensabilidade do atributo B (Lambert-Torres *et al.*, 1996a)

Nº	Atributos						S
	A	C	D	E	G	H	
1	B	M	A	B	0	1	U2
2	B	B	A	N	0	1	U2
3	B	M	A	N	0	1	U2
4	A	M	B	A	1	1	U2
5	B	M	A	B	1	1	U1
6	M	M	B	B	1	1	U1
7	B	M	A	N	0	1	U1
8	M	B	N	A	1	1	S
9	M	B	N	A	1	1	S
10	M	B	N	A	1	1	S
11	A	M	N	A	1	1	S
12	B	M	N	N	0	1	S

Na tabela 3.5 apresenta-se o conjunto resultante dos objetos, com os atributos indispensáveis.

Tabela 3.5 – Conjunto de objetos com os atributos indispensáveis (I) (Lambert-Torres *et al.*, 1996a)

Nº	Atributos			S
	B	D	G	
1	M	A	0	U2
2	M	A	0	U2
3	M	A	0	U2
4	M	B	1	U2
5	M	A	1	U1
6	A	B	1	U1
7	A	A	0	U1
8	B	N	1	S
9	M	N	1	S
10	A	N	1	S
11	M	N	1	S
12	A	N	0	S

Observando a tabela 3.5 é possível analisar que a mesma pode ser novamente reduzida, uma vez que os objetos {1, 2, 3} são idênticos. Na tabela 3.6 pode-se visualizar o novo conjunto de objetos.

Tabela 3.6 – Conjunto de objetos com os atributos indispensáveis (II) (Lambert-Torres *et al.*, 1996a)

Nº	Atributos			S
	B	D	G	
1	M	A	0	U2
2	M	B	1	U2
3	M	A	1	U1
4	A	B	1	U1
5	A	A	0	U1
6	B	N	1	S
7	M	N	1	S
8	A	N	1	S
9	M	N	1	S
10	A	N	0	S

Quinto e Sexto Passo:

Utilizando a tabela 3.6, é possível calcular o *core* do conjunto de objetos do exemplo. Este cálculo pode ser feito eliminando cada atributo passo a passo, e verificando se a tabela de decisão continua a dar as respostas corretas (i.e. se continua a ser consistente).

Um exemplo deste cálculo pode ser visto no exemplo 1 (tabela 3.5). Ao eliminar-se o atributo B, é possível verificar que a tabela de decisão se torna inconsistente porque o objeto 1 e 5 têm os mesmos atributos (o atributo D toma o valor de A, e G o valor de 0) e diferentes decisões (U_2 e U_1 , respetivamente).

No entanto, quando o atributo D é eliminado, a tabela de decisão permanece consistente. Na tabela 3.7 observa-se o *core* da tabela de decisão, enquanto na tabela 3.8 tem-se o Reduto de cada objeto.

Tabela 3.7 – *Core* do conjunto de objetos (Lambert-Torres *et al*, 1996a)

Nº	Atributos			S
	B	D	G	
1	M	–	0	U2
2	M	B	–	U2
3	–	A	1	U1
4	A	B	–	U1
5	A	A	–	U1
6	–	–	–	S
7	–	N	–	S
8	–	N	–	S
9	–	N	–	S
10	–	N	–	S

Tabela 3.8 – Reduto do conjunto de objetos (Lambert-Torres *et al*, 1996a)

Nº	Atributos			S
	B	D	G	
1	M	×	0	U2
2	M	B	×	U2
3	×	A	1	U1
4	A	B	×	U1
5	A	A	×	U1
6'	B	×	×	S
6''	×	×	×	S
7	×	N	×	S
8	×	N	×	S
9	×	N	×	S
10	×	N	×	S

Sétimo e Oitavo Passo:

Segundo a tabela 3.8, é possível expressar o conhecimento existente na tabela 3.1 através do seguinte conjunto de regras:

- Se (B é M e G é 0) ou (B é M e D é B) então, $S = U_2$.
- Se (D é A e G é 1) ou (B é A e D é B) ou (B é A e D é A) então, $S = U_1$.
- Se (D é B) ou (D é N) então, $S = S$.

Utilizando a lógica aritmética é possível escrever o conjunto de regras de outra forma. Incorporando também uma escala de valores, no qual o final das regras contém o conhecimento da tabela 3.1 e pode ser expresso da seguinte maneira:

- Se (B é M) e (D é B ou G é 0) então, $S = U_2$.
- Se (D é A e G é 1) ou (B é A e (D é B ou A)) então, $S = U_1$.
- Se (D é B) ou (D é N) então, $S = S$.

Usando uma formulação completa das regras tem-se:

- Se (o trânsito de potências na linha de transmissão B está entre 40% e 70%) e ((a tensão no barramento D está abaixo de 0.85) ou (o disjuntor G está aberto)) então, a classificação do estado do sistema é Inseguro de Nível 2.
- Se (a tensão no barramento D está acima de 1.05) e (o disjuntor está fechado) então, a classificação do estado do sistema é Inseguro de Nível 1.
- Se (o trânsito de potências na linha de transmissão B está acima de 70%) e (a tensão no barramento D está abaixo de 0.85 ou acima de 1.05) então, a classificação do estado do sistema é Inseguro de Nível 1.
- Se (o trânsito de potências na linha B está abaixo de 40%) então, a classificação do estado do sistema é Seguro.
- Se (a tensão no barramento D está entre 0.85 e 1.05) então, a classificação do estado do sistema é Seguro.

Os passos anteriormente descritos serão aplicados com o intuito de tratar um amplo conjunto de informação fornecido pela Rede Elétrica Nacional, deste modo é possível gerar um conhecimento inicial, fulcral para o processo de previsão.

3.9 – Sistema de Aprendizagem LERS

O sistema de aprendizagem LERS (*Learning from Examples based on Rough Sets*), desenvolvido por Grzyrnala-Busse, tem como finalidade induzir um conjunto de regras com base em exemplos e classificar novos exemplos usando o conjunto de regras anteriormente induzidas. O sistema calcula assim as aproximações inferiores e superiores de cada conceito. Seguidamente, cria regras exatas e regras possíveis com base nessas aproximações. Estas regras são posteriormente utilizadas para classificar novos dados, permitindo assim classificar eventos futuros.

3.10 - Conclusão

Através da revisão bibliográfica realizada ao longo do presente capítulo, facilmente se compreendem os conceitos básicos associados à Teoria dos Conjuntos Aproximativos, proposta por Z. Pawlak em 1982. Através da sua abordagem para a descoberta de conhecimento é possível analisar uma base de dados, identificar e remover dados redundantes e classificar dados incompletos, vagos e incertos.

A presente teoria tem vindo a ser aplicada nas mais diversas áreas de estudo. Na área dos Sistemas Eléctricos de Energia, a sua aplicação tem vindo a dar provas da sua utilidade, como podemos observar na sua aplicação a um Centro de Controlo de Sistemas Eléctricos de Energia ou mesmo noutros trabalhos já desenvolvidos na área da Análise de Segurança dos Sistemas Eléctricos de Energia (Agreira, 2010).

Capítulo 4 – O MIBEL e a Produção de Energia

4.1 – Introdução

O sistema elétrico tem vindo a sofrer profundas reestruturações ao longo das últimas décadas em todo o mundo. No sentido de estabelecer um mercado de energia elétrica mais competitivo, estas mudanças incidiram na liberalização do mercado, verificando-se em muitos países o fim da organização em monopólio. As alterações também se têm verificado na produção de energia, na qual tem ocorrido um aumento da diversificação das fontes primárias de energia, contribuindo para aumentar estruturalmente a segurança de abastecimento de cada país. O sistema elétrico está assim em constante alteração de modo a fazer face às necessidades do conforto e satisfação do consumidor e da evolução económica do país.

4.2 – O Mercado Ibérico de Eletricidade – MIBEL

O mercado de energia tem estado em constante desenvolvimento ao longo dos últimos anos. Desde 1996 que a União Europeia iniciou gradualmente a liberalização do mercado de energia, tal como já tinha feito noutros setores. O seu objetivo primordial tem sido a criação de um único mercado integrado Europeu, o Mercado Interno de Energia, de forma a reduzir os custos globais da rede e assegurar a segurança no fornecimento. No âmbito da criação deste mercado, têm sido constituídos mercados regionais no espaço da União Europeia. Estes visam ser uma etapa intermédia para a integração num mercado único.

Neste contexto surge o Mercado Ibérico de Eletricidade – MIBEL, este constitui uma iniciativa conjunta de cooperação entre os Governos de Portugal e de Espanha, com o objetivo de promoverem a integração dos sistemas elétricos dos dois países. Esta cooperação culminou num mercado de energia elétrica a nível ibérico, constituindo assim mais um mercado regional e mais um passo significativo para a construção do Mercado Interno de Energia.

4.2.1 – Objetivos do MIBEL

A concretização do MIBEL veio permitir a qualquer consumidor do espaço ibérico adquirir energia elétrica, num regime de livre concorrência, a qualquer produtor ou comercializador que atue em Portugal ou Espanha.

Os principais objetivos do MIBEL baseiam-se nas seguintes metas (OMIP, 2016a):

- Beneficiar os consumidores de eletricidade dos dois países, através do processo de integração dos respetivos sistemas elétricos;
- Estruturar o funcionamento do mercado com base nos princípios de transparência, livre concorrência, objetividade, liquidez, autofinanciamento e auto-organização;
- Favorecer o desenvolvimento do mercado de eletricidade de ambos os países, com a existência de uma metodologia única e integrada, para toda a península ibérica, de definição dos preços de referência;
- Permitir a todos os participantes o livre acesso ao mercado, em condições de igualdade de direitos e obrigações, transparência e objetividade;
- Favorecer a eficiência económica das empresas do setor elétrico, promovendo a livre concorrência entre as mesmas.

4.2.2 – Fases de Implementação

O processo de criação do MIBEL iniciou em 1998 através de conversações e estudos para progressivamente eliminarem obstáculos e favorecerem a sua criação.

Nessa sequência, em novembro de 2001 foi iniciado formalmente a criação do MIBEL através da celebração do “Protocolo de colaboração entre as Administrações espanhola e portuguesa para a criação do Mercado Ibérico de Eletricidade”. Esta foi a primeira fase, onde foram estabelecidas as bases necessárias para o início da cooperação entre as diversas entidades.

Em outubro de 2002, em Valência na XVIII Cimeira Luso-Espanhola, ficou decidido o modelo de organização do MIBEL, assente na existência do Operador de Mercado Ibérico (OMI) e foram estabelecidas as principais metas de concretização do MIBEL. A gestão dos mercados organizados do MIBEL ficaram assim a cargo do OMI, sendo que esta assentaria numa estrutura bipolar interligada, na qual:

- O OMIE (polo espanhol) ficou com a competência da gestão do mercado diário e intradiário;
- O OMIP (polo português) ficou encarregue da gestão dos mercados a prazo.

Seguidamente, em janeiro de 2004 foi firmado o “Acordo entre a República Portuguesa e o Reino de Espanha para a Constituição de um Mercado Ibérico de Energia Elétrica”. Diversas

circunstâncias de carácter político e legal obrigaram à introdução de novos ajustes e à substituição do acordo firmado.

O arranque do MIBEL foi sofrendo alguns atrasos devido às dificuldades e morosidade da resolução de certos obstáculos. No entanto, em novembro de 2016, aquando da realização da XXII Cimeira Luso-Espanhola em Badajoz, foram tomadas um conjunto de decisões, por parte dos dois governos, que permitiram definir as condições efetivas para o arranque do MIBEL.

No dia 1 de julho de 2007 deu-se o arranque do MIBEL, resultante de um longo trabalho de harmonização de condições entre os dois sistemas elétricos nacionais de Portugal e Espanha. Criado numa perspetiva de originar benefícios para todos os consumidores de ambos os países (OMIP, 2016a).

4.2.3 – Operadores do MIBEL

4.2.3.1 – OMIE

O OMIE – Operador de Mercado Ibérico de Energia (Polo Espanhol), S. A. é a entidade gestora do mercado *spot* de contratação à vista, com uma componente diária (Mercado Diário) e uma componente de ajustes intradiários (Mercado Intradiário). Neste mercado estabelece-se a compra e venda de energia elétrica a ser entregue no dia seguinte ao da negociação. Ao nível do funcionamento dos mercados, o OMIE é responsável por (MIBEL, 2016):

- Receber as ofertas de aquisição e de venda de energia elétrica para cada período de programação, das diferentes entidades que participam no Mercado Diário e Intradiário;
- Realizar o encontro das ofertas de venda e de aquisição de energia elétrica partindo da oferta mais barata até igualar a procura para cada período de programação;
- Determinar o preço de energia elétrica resultante do encontro de ofertas no Mercado Diário e Intradiário, para cada período de programação, bem como, comunicá-lo a todos os agentes envolvidos;
- Comunicar ao Operador de Sistema as ofertas de venda e de aquisição de energia elétrica que tiverem sido aceites, para cada período de programação;
- Comunicar ao Operador de Sistema, com a devida antecedência, todas as modificações relacionadas com os agentes e as ofertas envolvidas.

O Mercado Diário forma o preço para cada uma das 24 horas de cada dia, correspondendo a 24 horas de transação, tendo como referência a hora espanhola. Nos dias em que houver mudança de hora, esse número de períodos é modificado para 23 ou 25 conforme o ajuste horário necessário.

4.2.3.2 – OMIP

O OMIP – Operador de Mercado Ibérico de Energia (Polo Português), SGMR, S.A. assegura a gestão do mercado de contratação a prazo do MIBEL. O OMIP tem como principais objetivos contribuir para o desenvolvimento do MIBEL, promover preços de referência ibéricos, disponibilizar instrumentos eficientes de gestão de risco e disponibilizar um modelo de mercado adequado a todos os agentes. É também da sua responsabilidade a prestação de informação relevante aos participantes e ao público em geral, relativamente ao funcionamento do mercado a prazo e ao estabelecimento de referenciais de preço oficiais, designadamente através da publicação do Boletim de Mercado (MIBEL, 2016).

Através do Boletim de Mercado é possível ter acesso aos preços de referência de negociação dos Contratos de Futuros, o produto mais transacionado pelo OMIP. Os produtos transacionados neste mercado a prazo são contratos de compra e venda de energia, que têm um horizonte temporal futuro, nomeadamente, dias, semanas, meses, trimestres ou mesmo anos. Os preços de referência que cobrem esse espaço temporal são apresentados no Boletim de Mercado. São também apresentados neste boletim os preços de referência relativos aos três dias seguintes, o qual permite saber no dia de hoje os valores previstos para o preço da energia do dia depois de amanhã. O valor destes preços da energia elétrica é calculado com base em índices de eletricidade. Esta informação é de elevada importância para a metodologia desenvolvida neste projeto.

4.2.4 – Índices de Eletricidade do MIBEL

Tendo em vista a obtenção de referências fiáveis e consistentes do Mercado Ibérico de Eletricidade, de modo a permitir a negociação de derivados, o OMIP calcula e divulga índices de eletricidade, que representam o comportamento dos preços deste ativo ao longo do tempo. Estes índices podem referir-se à eletricidade espanhola (índices SPEL) e eletricidade portuguesa (índices PTEL). O cálculo separado dos índices para as duas zonas ocorre devido

ao mecanismo de *market splitting*, implementado na gestão de congestionamentos entre Portugal e Espanha. Este mecanismo faz com que o OMEL passe a publicar diferentes preços horários zonais relativos a Espanha e Portugal o que leva o OMIP a calcular e publicar os índices SPEL e PTEL (OMIP, 2016b) (OMIP, 2016c).

Por forma a cobrir as formas típicas de atuação no mercado, o OMIP divulga dois tipos de índices, o *Base Load* e o *Peak Load*. O primeiro refere-se ao fornecimento de eletricidade durante as 24 horas do dia, enquanto o segundo diz respeito apenas ao fornecimento durante o período das 8 horas às 20 horas (horário espanhol).

Para cada dia de calendário, o índice PTEL Base corresponde à média aritmética dos preços marginais horários da zona portuguesa formados no mercado diário gerido pelo OMEL, para as 24 (23 ou 25) horas do dia, definidas de acordo com a Hora Oficial Espanhola (HOES), arredondada a duas casas decimais, de acordo com a seguinte fórmula de cálculo (OMIP, 2016b) (OMIP, 2016c):

$$\text{ÍndicePTELBase} = \frac{\sum_{i=1}^n \text{SMP}(p)_i^B}{n} \quad (4.1)$$

Em que:

- n = número de horas num dia de calendário (24, exceto no último domingo de março, em que são 23 e no de outubro, que são 25);
- i = corresponde a cada hora do dia de calendário para o qual se está a calcular o índice PTEL Base, sendo ambos (horas e dias) definidos de acordo com a HOES;
- $\text{SMP}(p)_i^B$ = preço marginal horário da zona portuguesa (Precio Marginal Horario del Mercado Diário – Sistema Português) para a hora i , definido em € por MWh com duas casas decimais e de acordo com a HOES.

Para cada dia da semana, segunda-feira a sexta-feira (incluindo feriados, de segunda a sexta-feira), o índice PTEL Peak corresponde à média aritmética dos preços marginais horários do sistema português, para o período entre as 8 horas e as 20 horas (HOES), arredondada a duas casas decimais. A respetiva fórmula de cálculo é a seguinte (OMIP, 2016b) (OMIP, 2016c):

$$\text{ÍndicePTELPeak} = \frac{\sum_{j=1}^n \text{SMP}(p)_j^P}{n} \quad (4.2)$$

Em que:

- n = número de horas *peak* num dia útil (período entre as 8 horas e as 20 horas (HOES), o que equivale a 12 horas);
- j = corresponde a cada hora *peak* do dia de calendário, definida de acordo com a HOES, para o qual se está a calcular o índice PTEL Peak;
- $\text{SMP}(p)_j^P$ = preço marginal horário da zona portuguesa (Precio Marginal Horario del Mercado Diário – Sistema Português) para a hora j , definido em € por MWh com duas casas decimais e de acordo com a HOES.

O procedimento é análogo para a zona espanhola, havendo apenas alteração na variável referente ao preço marginal horário, caso esta difira da zona portuguesa.

4.3 – Produção de energia elétrica

O SEN (Sistema Eléctrico Nacional) tem como objetivo fundamental a disponibilização de energia eléctrica em termos adequados às necessidades dos consumidores. Os principais elos de cadeia de valor do SEN são a produção, transporte, distribuição e comercialização de electricidade.

A produção de energia eléctrica em Portugal é feita com recurso a diversas tecnologias e diferentes fontes primárias de energia (carvão, gás, água, vento, entre outros). A produção é exercida num contexto de livre concorrência, submetida a controlo prévio através da emissão de licença ou através de comunicação prévia, dividindo-se nos dois regimes seguintes:

- Produção em Regime Ordinário – energia eléctrica produzida através de fontes tradicionais não renováveis e de grandes centros eletroprodutores hídricos. Esta pode ser dividida em duas categorias, as térmicas e as grandes hídricas. Fazem parte das térmicas as centrais a carvão e a gás natural. Por outro lado, as grandes hídricas são formadas pelas centrais a fio-de-água e as centrais hidroeléctricas de albufeira.

- Produção em Regime Especial – energia elétrica produzida através de centrais de cogeração e de centrais cujos recursos são de fonte renovável, excluindo os grandes centros eletroprodutores hídricos.

A produção de energia elétrica tem por objetivo a satisfação do consumo em qualquer momento, o que implica um bom planeamento do sistema de energia elétrica. É por isso necessário estudar quais os centros produtores que idealmente devem ser os mais vantajosos em termos económicos e ambientais. Esta filosofia tem originado uma mudança ao longo dos últimos anos no parque eletroprodutor em Portugal.

Analisando os dados no qual incide o presente projeto (2014 e 2015), é possível constatar que no final de 2015, o valor total da potência instalada era de 18 533 MW, dos quais 12 014 MW de origem renovável e 6 519 MW não renovável. Um aumento de 699 MW de potência instalada face a 2014. Este aumento verificou-se apenas na produção de origem renovável que sofreu um aumento de 792 MW. A produção de origem não renovável teve um decréscimo de 93 MW (REN, 2016).

O consumo de energia elétrica em Portugal entre 2011 e 2015, analisado na figura 4.1, desagregado nas diferentes naturezas de produção, tem oscilado nos 50 TWh ao longo destes cinco anos. Em 2014, as condições foram particularmente favoráveis para a produção hídrica, com uma quota de 31% e a eólica com 24%. As centrais térmicas a carvão abasteceram 23% do consumo, enquanto as centrais a gás natural abasteceram 13%. O saldo de trocas com a rede elétrica espanhola foi importador, 2% do consumo, o valor mais reduzido desde 2001. No ano 2015, a produção hídrica baixou para 17% devido ao tempo muito seco. A eólica manteve-se nos 23%, não apresentando uma variação significativa face a 2014. As centrais térmicas a carvão abasteceram 28%, representando um aumento de 5%. Enquanto as centrais a gás natural abasteceram 20%, um aumento de 7%. O saldo de trocas com a rede elétrica espanhola foi importador, 4,7% do consumo (REN, 2016).

Com esta análise é possível destacar a importância das centrais térmicas e das importações com Espanha. Apesar do contributo da produção hídrica e da restante produção de origem renovável, esta está fortemente dependente do estado do tempo.

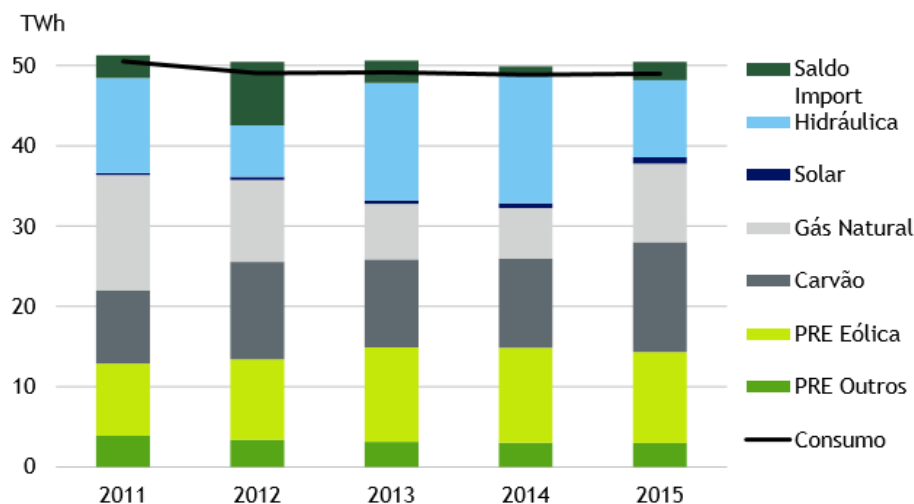


Figura 4.1 – Evolução da produção de energia elétrica, entre 2011 e 2015 (REN, 2016)

A produção de energia elétrica ao longo do ano está condicionada devido a vários fatores, desde o preço dos combustíveis, às condições climatéricas e até mesmo devido a questões políticas. Por esses e outros motivos, a previsão da produção de energia elétrica é um desafio. Desafio este que tem vindo a ser colmatado através de técnicas de previsão que atualmente já permitem a previsão da produção de energia para o dia de amanhã e a previsão para o dia depois de amanhã para as centrais eólicas, fotovoltaicas e mini-hídricas. Existe, no entanto, ainda uma lacuna relativamente à previsão da produção de energia elétrica, para o dia depois de amanhã. Esta abrange a PRO, nomeadamente a grande hídrica, a fio-de-água e a bombagem, tal como a produção térmica, a carvão e a gás natural. Abrange ainda a troca de energia entre Portugal e Espanha (importação e exportação).

4.4 – Conclusão

O MIBEL representou uma grande e importante mudança no sector elétrico. Através da informação presente neste capítulo foi possível enunciar alguns dos objetivos e fases marcantes da implementação deste mercado de energia.

Salienta-se ainda a importância e responsabilidade dos operadores de mercado e dos índices de eletricidade, dando ênfase aos índices PTEL que desempenham um papel crucial para o desenvolvimento da presente metodologia. Tendo em conta que o presente trabalho se foca nos anos 2014 e 2015, é ainda feita uma análise do diagrama de carga deste período.

Capítulo 5 – Aplicação da Teoria dos Conjuntos Aproximativos

5.1 – Introdução

A falta de informação não permite ter os dados necessários para efetuar a análise do *Day Ahead Congestion Forecast* para o dia depois de amanhã. Consequentemente, é necessário criar uma metodologia capaz de prever a informação para a aplicação deste procedimento.

A Teoria dos Conjuntos Aproximativos é uma ferramenta poderosa para extração de informação em grandes bases de dados, desenvolvida a partir da década de oitenta. A sua aplicação pode ser feita através do *software* ROSE2, permitindo analisar grandes bases de dados e aplicar técnicas de descoberta de regras capazes de classificar novos objetos.

Neste capítulo são apresentados vários cenários onde são aplicados diferentes métodos de discretização e algoritmos de indução de regras. Através destes cenários e com base nesta teoria, são desenvolvidos vários conjuntos de regras que ao serem aplicadas permitem classificar a produção média dos vários tipos de centrais em estudo e as trocas de produção entre Portugal e Espanha, para o dia depois de amanhã.

Este capítulo pretende ainda descrever as etapas necessárias para a aplicação desta metodologia, desde o tratamento de dados, o processo de discretização e a aplicação do *software*. Para isso, é feita uma descrição pormenorizada dos vários cenários aplicados às centrais cuja produção para o dia depois de amanhã é desconhecida, são elas as centrais a Carvão, a Gás Natural, de Albufeira, a Fio-de-água, Bombagem e ainda à Exportação e Importação.

5.2 – Dados e tratamento de dados

A metodologia aplicada neste trabalho, para a previsão da produção de energia elétrica e troca de energia entre Portugal e Espanha para o dia depois de amanhã, baseia-se numa base de dados referente aos anos 2014 e 2015, cedida pela REN. Os dados presentes nesta base de dados referem-se à produção de energia (Carvão, Gás Natural, Albufeiras, Fio-de-água, Mini-hídrica, Eólica, Bombagem), à energia transacionada (Importação e Exportação), ao consumo de energia (todos fornecidos na unidade de MW) e ao preço de energia em Portugal (em €/MWh), ao longo dos 365 dias de cada ano.

Para o tratamento dos dados foi inicialmente feita uma conversão dos dados relativos à produção, consumo e troca de energia elétrica. Este passo foi necessário uma vez que a base de dados inicial apresentava os valores com um intervalo de tempo de 15 minutos, perfazendo um total de 96 amostras por dia. Foi feita a média destas amostras, culminando assim num único valor, representando a potência média para cada dia. O mesmo tratamento de dados foi feito relativamente ao consumo e à troca de energia elétrica, mantendo assim a coerência dos dados.

5.3 – Discretização

A extração de conhecimento em base de dados com grandes volumes de informação, utilizando a TCA, requer a transformação dos valores contínuos dos atributos em intervalos discretos. As variáveis presentes nestas bases de dados apresentam valores contínuos que variam de forma aleatória. A utilização destes valores na sua forma primitiva, para extrair regras de mineração, gera um número elevado de regras. Esta quantidade de regras aumenta o esforço computacional e condiciona a sua posterior aplicação, dado serem regras muito minuciosas para o conjunto de aprendizagem e pouco abrangentes para posteriormente classificarem os objetos de teste. Assim sendo, todos os valores contínuos dos atributos necessitam de ser reduzidos a um número finito de intervalos e posteriormente aplicado o processo de mineração de regras. Este processo, denominado discretização, é responsável pela transformação de valores contínuos em intervalos discretos. É uma técnica de redução de dados, que reduz a gama de valores de um atributo com valores contínuos num número mínimo de intervalos discretos (Ali *et al*, 2015).

Apesar de existirem diferentes tipos de métodos de discretização aplicados à TCA, estes não têm a sensibilidade desejada para o tipo de informação que está a ser processada. Para discretizar os valores contínuos em questão (relativos à produção, consumo e troca de energia), mantendo a fiabilidade da informação, foi feito um estudo de comparação entre dois métodos desenvolvidos, com distintos códigos de equivalência para os atributos. O índice PTEL, cuja gama de valores tinha uma variação mais pequena (variando aproximadamente entre 0 e 80 €/MWh), optou-se pela utilização de apenas um código de equivalência fixo, conforme se pode observar na tabela 5.1.

Tabela 5.1 – Código de equivalência do Índice PTEL

Índice PTEL	Código
$0 \leq \text{PTEL} < 4.95$	0
$4.95 \leq \text{PTEL} < 14.95$	10
$14.95 \leq \text{PTEL} < 24.95$	20
$24.95 \leq \text{PTEL} < 34.95$	30
$34.95 \leq \text{PTEL} < 44.95$	40
$44.95 \leq \text{PTEL} < 54.95$	50
$54.95 \leq \text{PTEL} < 64.95$	60
$64.95 \leq \text{PTEL} < 74.95$	70
$74.95 \leq \text{PTEL} < 84.95$	80
$84.95 \leq \text{PTEL} < 94.95$	90

Para os restantes atributos, cuja gama de valores tem uma variação bastante elevada dependendo do tipo de central e da sua potência instalada, foi necessário aplicar os métodos desenvolvidos que utilizam códigos de equivalência que compreendem toda essa gama. Os códigos de equivalência para o primeiro método estão repartidos pelas tabelas 5.2, 5.3 e 5.4 onde estão representados de modo parcial.

Na tabela 5.2 estão definidos os códigos de equivalência para aplicar nos valores das potências das centrais mini-hídricas, eólicas e no consumo. Estes códigos encontram-se agrupados na mesma tabela dado que são os únicos atributos cuja previsão para o dia depois de amanhã (D+2) é conhecida. Esta tabela pode ser visualizada na totalidade no Anexo I como AI.1.

Tabela 5.2 – Código de equivalência do método 1 – Mini-hídrica, Eólica e Consumo

Mini-hídrica		Eólica		Consumo	
$14.95 \leq P < 24.95$	20	$134.95 \leq P < 144.95$	140	$4044.95 \leq P < 4054.95$	4050
$24.95 \leq P < 34.95$	30	$144.95 \leq P < 154.95$	150	$4054.95 \leq P < 4064.95$	4060
$34.95 \leq P < 44.95$	40	$154.95 \leq P < 164.95$	160	$4064.95 \leq P < 4074.95$	4070
$44.95 \leq P < 54.95$	50	$164.95 \leq P < 174.95$	170	$4074.95 \leq P < 4084.95$	4080
$54.95 \leq P < 64.95$	60	$174.95 \leq P < 184.95$	180	$4084.95 \leq P < 4094.95$	4090
...
...
$314.95 \leq P < 324.95$	320	$3634.95 \leq P < 3644.95$	3640	$6724.95 \leq P < 6734.95$	6730
$324.95 \leq P < 334.95$	330	$3644.95 \leq P < 3654.95$	3650	$6734.95 \leq P < 6744.95$	6740
$334.95 \leq P < 344.95$	340	$3654.95 \leq P < 3664.95$	3660	$6744.95 \leq P < 6754.95$	6750
$344.95 \leq P < 354.95$	350	$3664.95 \leq P < 3674.95$	3670	$6754.95 \leq P < 6764.95$	6760
$354.95 \leq P < 364.95$	360	$3674.95 \leq P < 3684.95$	3680	$6764.95 \leq P < 6774.95$	6770

Seguidamente, na tabela 5.3 estão definidos os códigos de equivalência para aplicar nos dados referentes às potências das centrais térmicas (carvão e gás natural) e nos das trocas de energia entre Portugal e Espanha (importação e exportação). Esta tabela pode ser visualizada na totalidade no Anexo I como AI.2.

Tabela 5.3 – Código de equivalência do método 1 – Carvão, Gás Natural, Importação e Exportação

Carvão		Gás Natural		Importação		Exportação	
$174.95 \leq P < 184.95$	180	$0 \leq P < 4.95$	0	$0 \leq P < 4.95$	0	$0 \leq P < 4.95$	0
$184.95 \leq P < 194.95$	190	$4.95 \leq P < 14.95$	10	$4.95 \leq P < 14.95$	10	$4.95 \leq P < 14.95$	10
$194.95 \leq P < 204.95$	200	$14.95 \leq P < 24.95$	20	$14.95 \leq P < 24.95$	20	$14.95 \leq P < 24.95$	20
$204.95 \leq P < 214.95$	210	$24.95 \leq P < 34.95$	30	$24.95 \leq P < 34.95$	30	$24.95 \leq P < 34.95$	30
$214.95 \leq P < 224.95$	220	$34.95 \leq P < 44.95$	40	$34.95 \leq P < 44.95$	40	$34.95 \leq P < 44.95$	40
...
...
$1714.95 \leq P < 1724.95$	1720	$1044.95 \leq P < 1054.95$	1050	$1474.95 \leq P < 1484.95$	1480	$1804.95 \leq P < 1814.95$	1810
$1724.95 \leq P < 1734.95$	1730	$1054.95 \leq P < 1064.95$	1060	$1484.95 \leq P < 1494.95$	1490	$1814.95 \leq P < 1824.95$	1820
$1734.95 \leq P < 1744.95$	1740	$1064.95 \leq P < 1074.95$	1070	$1494.95 \leq P < 1504.95$	1500	$1824.95 \leq P < 1834.95$	1830
$1744.95 \leq P < 1754.95$	1750	$1074.95 \leq P < 1084.95$	1080	$1504.95 \leq P < 1514.95$	1510	$1834.95 \leq P < 1844.95$	1840
$1754.95 \leq P < 1764.95$	1760	$1084.95 \leq P < 1094.95$	1090	$1514.95 \leq P < 1524.95$	1520	$1844.95 \leq P < 1854.95$	1850

Na tabela 5.4 estão definidos os códigos de equivalência para aplicar nos valores das potências das restantes centrais hídricas (albufeira, fio-de-água e bombagem). Esta tabela pode ser visualizada na totalidade no Anexo I como AI.3.

Tabela 5.4 – Código de equivalência do método 1 – Albufeira, Fio-de-água e Bombagem

Albufeira		Fio-de-água		Bombagem	
$114.95 \leq P < 124.95$	120	$184.95 \leq P < 194.95$	190	$0 \leq P < 4.95$	0
$124.95 \leq P < 134.95$	130	$194.95 \leq P < 204.95$	200	$4.95 \leq P < 14.95$	10
$134.95 \leq P < 144.95$	140	$204.95 \leq P < 214.95$	210	$14.95 \leq P < 24.95$	20
$144.95 \leq P < 154.95$	150	$214.95 \leq P < 224.95$	220	$24.95 \leq P < 34.95$	30
$154.95 \leq P < 164.95$	160	$224.95 \leq P < 234.95$	230	$34.95 \leq P < 44.95$	40
...
...
$2114.95 \leq P < 2124.95$	2120	$2214.95 \leq P < 2224.95$	2220	$374.95 \leq P < 384.95$	380
$2124.95 \leq P < 2134.95$	2130	$2224.95 \leq P < 2234.95$	2230	$384.95 \leq P < 394.95$	390
$2134.95 \leq P < 2144.95$	2140	$2234.95 \leq P < 2244.95$	2240	$394.95 \leq P < 404.95$	400
$2144.95 \leq P < 2154.95$	2150	$2244.95 \leq P < 2254.95$	2250	$404.95 \leq P < 414.95$	410
$2154.95 \leq P < 2164.95$	2160	$2254.95 \leq P < 2264.95$	2260	$414.95 \leq P < 424.95$	420

Os códigos de equivalência para o segundo método são mais abrangentes e estão repartidos pelas tabelas 5.5, 5.6 e 5.7 onde estão representados de modo parcial.

Na tabela 5.5 estão definidos os códigos de equivalência para aplicar aos dados das potências das centrais mini-hídricas, eólicas e aos do consumo. Estes códigos encontram-se agrupados na mesma tabela dado que são os únicos atributos cuja previsão para o dia depois de amanhã (D+2) é conhecida. Esta tabela pode ser visualizada na totalidade no Anexo I como AI.4.

Tabela 5.5 – Código de equivalência do método 2 – Mini-hídrica, Eólica e Consumo

Mini-hídrica		Eólica		Consumo	
$0 \leq P < 49.95$	0	$49.95 \leq P < 149.95$	100	$4049.95 \leq P < 4149.95$	4100
$49.95 \leq P < 149.95$	100	$149.95 \leq P < 249.95$	200	$4149.95 \leq P < 4249.95$	4200
$149.95 \leq P < 249.95$	200	$249.95 \leq P < 349.95$	300	$4249.95 \leq P < 4349.95$	4300
$249.95 \leq P < 349.95$	300	$349.95 \leq P < 449.95$	400	$4349.95 \leq P < 4449.95$	4400
$349.95 \leq P < 449.95$	400	$449.95 \leq P < 549.95$	500	$4449.95 \leq P < 4549.95$	4500
-----	-----
-----	-----
-----	-----	$3249.95 \leq P < 3349.95$	3300	$6349.95 \leq P < 6449.95$	6400
-----	-----	$3349.95 \leq P < 3449.95$	3400	$6449.95 \leq P < 6549.95$	6500
-----	-----	$3449.95 \leq P < 3549.95$	3500	$6549.95 \leq P < 6649.95$	6600
-----	-----	$3549.95 \leq P < 3649.95$	3600	$6649.95 \leq P < 6749.95$	6700
-----	-----	$3649.95 \leq P < 3749.95$	3700	$6749.95 \leq P < 6849.95$	6800

Na seguinte tabela 5.6 estão definidos os códigos de equivalência para as potências das centrais térmicas (carvão e gás natural) e das trocas de energia entre Portugal e Espanha (importação e exportação). Esta tabela pode ser visualizada na totalidade no Anexo I como AI.5.

Tabela 5.6 – Código de equivalência do método 2 – Carvão, Gás Natural, Importação e Exportação

Carvão		Gás Natural		Importação		Exportação	
$149.95 \leq P < 249.95$	200	$0 \leq P < 49.95$	0	$0 \leq P < 49.95$	0	$0 \leq P < 49.95$	0
$249.95 \leq P < 349.95$	300	$49.95 \leq P < 149.95$	100	$49.95 \leq P < 149.95$	100	$49.95 \leq P < 149.95$	100
$349.95 \leq P < 449.95$	400	$149.95 \leq P < 249.95$	200	$149.95 \leq P < 249.95$	200	$149.95 \leq P < 249.95$	200
$449.95 \leq P < 549.95$	500	$249.95 \leq P < 349.95$	300	$249.95 \leq P < 349.95$	300	$249.95 \leq P < 349.95$	300
$549.95 \leq P < 649.95$	600	$349.95 \leq P < 449.95$	400	$349.95 \leq P < 449.95$	400	$349.95 \leq P < 449.95$	400
...	...	$449.95 \leq P < 549.95$	500
...	...	$549.95 \leq P < 649.95$	600
$1349.95 \leq P < 1449.95$	1400	$649.95 \leq P < 749.95$	700	$1049.95 \leq P < 1149.95$	1100	$1449.95 \leq P < 1549.95$	1500
$1449.95 \leq P < 1549.95$	1500	$749.95 \leq P < 849.95$	800	$1149.95 \leq P < 1249.95$	1200	$1549.95 \leq P < 1649.95$	1600
$1549.95 \leq P < 1649.95$	1600	$849.95 \leq P < 949.95$	900	$1249.95 \leq P < 1349.95$	1300	$1649.95 \leq P < 1749.95$	1700
$1649.95 \leq P < 1749.95$	1700	$949.95 \leq P < 1049.95$	1000	$1349.95 \leq P < 1449.95$	1400	$1749.95 \leq P < 1849.95$	1800
$1749.95 \leq P < 1849.95$	1800	$1049.95 \leq P < 1149.95$	1100	$1449.95 \leq P < 1549.95$	1500	$1849.95 \leq P < 1949.95$	1900

Na tabela 5.7 estão definidos os códigos de equivalência para aplicar nos valores das potências das restantes centrais hídricas (albufeira, fio-de-água e bombagem). Esta tabela pode ser visualizada na totalidade no Anexo I como AI.6.

Tabela 5.7 – Código de equivalência do método 2 – Albufeira, Fio-de-água e Bombagem

Albufeira		Fio-de-água		Bombagem	
$49.95 \leq P < 149.95$	100	$149.95 \leq P < 249.95$	200	$0 \leq P < 49.95$	0
$149.95 \leq P < 249.95$	200	$249.95 \leq P < 349.95$	300	$49.95 \leq P < 149.95$	100
$249.95 \leq P < 349.95$	300	$349.95 \leq P < 449.95$	400	$149.95 \leq P < 249.95$	200
$349.95 \leq P < 449.95$	400	$449.95 \leq P < 549.95$	500	$249.95 \leq P < 349.95$	300
$449.95 \leq P < 549.95$	500	$549.95 \leq P < 649.95$	600	$349.95 \leq P < 449.95$	400
...	----	----
...	----	----
$1749.95 \leq P < 1849.95$	1800	$1849.95 \leq P < 1949.95$	1900	----	----
$1849.95 \leq P < 1949.95$	1900	$1949.95 \leq P < 2049.95$	2000	----	----
$1949.95 \leq P < 2049.95$	2000	$2049.95 \leq P < 2149.95$	2100	----	----
$2049.95 \leq P < 2149.95$	2100	$2149.95 \leq P < 2249.95$	2200	----	----
$2149.95 \leq P < 2249.95$	2200	$2249.95 \leq P < 2349.95$	2300	----	----

5.4 – O software ROSE2

O programa ROSE2 (*Rough Sets Data Explorer*) visa implementar a Teoria dos Conjuntos Aproximativos e técnicas de descoberta de regras. Desenvolvido pelo *Laboratory of Intelligent Decision Support Systems of the Institute of Computing Science* na Posnânia, este *software* baseia os seus cálculos nos fundamentos introduzidos pelo Z. Pawlak (Predki *et al*, 1998) (Predki e Wilk, 1999).

A introdução dos dados é feita através de uma tabela de informação onde estão definidos os atributos, objetos e o atributo de decisão, conforme definido no subcapítulo 3.3. Esta entrada de dados tem uma extensão “*.isf” e é com base nesta informação que é feito todo o processamento de dados, tal como a pesquisa do *core* e os redutos dos atributos, permitindo a redução dos dados. São também induzidos conjuntos de regras de decisão, resultantes da análise entre os atributos de condição e os de decisão. Regras estas que permitem estabelecer de uma forma clara e sucinta os dados da tabela de informação.

5.4.1 – Algoritmo de indução de regras

A indução de regras de decisão constitui um passo fulcral e minucioso na metodologia aplicada. O *software* ROSE2 permite a aplicação de um algoritmo capaz de realizar esta tarefa, o ModLEM. Este algoritmo é análogo ao LEM2 (Learning From Examples Module, version 2) e permite gerar regras com uma sintaxe mais abrangente. Em vez de regras simples, na forma *atributo = valor*, o ModLEM permite gerar regras na forma (*atributo ≤ valor*), (*atributo ≥ valor*) ou (*limite inferior ≤ atributo ≤ limite superior*). Permite ainda, caso o atributo seja escolhido duas vezes enquanto é construída uma única regra, definir a seguinte condição (*atributo = [valor1, valor2]*) que resulta da interseção de duas condições (*atributo < valor2*) e (*atributo ≥ valor1*) de tal modo que *valor1 < valor2* (Predki *et al*, 1998) (Predki e Wilk, 1999) (Grzymala-Busse e Stefanowski, 2001).

O algoritmo ModLEM disponibiliza duas medidas de avaliação, denominadas classe Entropia e precisão Laplaciana, utilizadas para avaliar as condições geradas. As condições com a melhor avaliação são adicionadas ao conjunto de regras geradas. As duas medidas disponibilizadas permitem que o utilizador opte pela mais adequada. A Entropia tem tendência a induzir regras “puras” abrangendo um reduzido número de exemplos. Por sua vez, a Laplaciana permite induzir regras com uma abrangência superior (Grzymala-Busse e Stefanowski, 2001).

5.4.2 – Parâmetros e interpretação das regras

A interpretação correta das regras é essencial para que se possa retirar o maior conhecimento. As regras são geradas pelo *software* ROSE2, num ficheiro com a extensão “*.rlf”. Cada regra é iniciada com ‘*rule No.*’, nomenclatura que identifica cada regra. Seguidamente, é apresentada a parte da condição da regra, usando a seguinte sintaxe:

- (*atributo<espaço>relação<espaço>valor*) & (*atributo<espaço>relação<espaço>valor*)

Em que, atributo representa o atributo de condição, valor é o seu valor e relação é um dos seguintes elementos: ‘=’, ‘>’, ‘<’, ‘>=’ ou ‘<=’. A quantidade de atributos de condição utilizados varia de regra para regra.

Terminada a parte condicional, são definidas as classes de decisão, representadas pela seguinte sintaxe:

- => (decisão=valor) ou (decisão=valor); ...

Uma regra pode ter uma ou mais decisões, caso se trate de uma regra aproximada, em que a mesma satisfaz a classificação de uma ou mais classes.

Cada regra contém ainda quatro parâmetros com informação adicional. Esta informação é apresentada depois da sintaxe das classes de decisão, na seguinte ordem:

- [*absolute cover, absolute support, relative support, level of discrimination*]

Em que:

- *Absolute cover* – representa o número de objetos de aprendizagem correspondentes à parte condicional da regra;
- *Absolute support* – representa o número de objetos de aprendizagem que correspondem tanto à parte condicional como à parte da decisão da regra. Caso o parâmetro *absolute support* seja inferior ao *absolute cover*, significa que os dados de aprendizagem são inconsistentes;
- *Relative cover* – é definido pelo rácio entre o *absolute support* e o número total de objetos da base de dados de aprendizagem, que apresentam a mesma classe de decisão;
- *Level of discrimination* – representa o grau de confiança e é definida pelo rácio entre os parâmetros *absolute support* e *absolute cover*. Regras induzidas através de dados inconsistentes apresentam um *level of discrimination* inferior a 1 ou 100%.

5.5 – Aplicação da Teoria dos Conjuntos Aproximativos

A Teoria dos Conjuntos Aproximativos foi aplicada com o objetivo de retirar o maior conhecimento possível de uma base de dados e aplicar esse conhecimento na previsão de dados futuros. Para apurar quais os códigos de equivalência mais fiáveis para a discretização dos dados da base de dados de aprendizagem (referente ao ano 2014), ambos os códigos foram aplicados e comparados. Foram ainda testados os diferentes algoritmos de indução de regras, permitindo assim analisar o desempenho de cada um dos algoritmos (Entropia ou Laplace) e analisar a eficácia das suas regras na previsão dos valores dos atributos para o dia D+2.

5.5.1 – Definição dos atributos

A TCA, descrita no Capítulo 3, foi aplicada a um conjunto de atributos específicos da base de dados de aprendizagem. Os atributos relevantes para esta fase de aprendizagem eram relativos à produção proveniente das centrais mini-hídricas, eólicas e ao consumo e índice PTEL. Os restantes atributos serão considerados atributos de decisão. O atributo de decisão varia consoante a previsão que se pretende fazer.

Desta forma, caso o elemento em estudo seja o Carvão, é este o atributo de decisão. Se o elemento em estudo for o Gás Natural, Albufeira, Fio-de-água, Bombagem, Importação ou a Exportação, este é definido como atributo de decisão. Os restantes atributos (Mini-hídrica, Eólica, Consumo e Índice PTEL) são os únicos atributos cuja previsão para o dia depois de amanhã (D+2) é conhecida, servindo estes como atributos condicionais para formar as condições das regras.

Os atributos de condição, considerados para a aplicação da teoria dos conjuntos aproximativos foram os seguintes:

- PTEL – Valor do índice PTEL;
- HID – Valor da produção das centrais mini-hídricas;
- EOL – Valor da produção das centrais eólicas;
- CONS – Valor do consumo de eletricidade.

Os atributos de decisão, considerados os de maior importância uma vez que representam o atributo a prever para o dia depois de amanhã, são representados pelos seguintes acrónimos:

- CARV – Valor da produção das centrais a carvão;
- GN – Valor da produção das centrais a gás natural;
- ALB – Valor da produção das centrais de albufeira;
- FA – Valor da produção das centrais a fio-de-água;
- BOMB – Valor da produção das centrais com bombagem;
- EXP – Valor da energia elétrica exportada;
- IMP – Valor da energia elétrica importada.

5.5.2 – Casos de estudo em análise

Para analisar a metodologia mais fiável a ser aplicada a cada um dos atributos a prever, foram definidos quatro cenários diferentes, utilizando a mesma base de dados inicial, com o respetivo atributo de decisão e os mesmos atributos de condição, mas com códigos de equivalência e algoritmos de indução de regras distintos. Esta variação dos códigos de equivalência e dos algoritmos de indução de regras permite testar e validar a melhor opção para a classificação de cada um dos atributos a prever.

Estes cenários podem ser agrupados em dois casos de estudo mais abrangentes, definidos pelos códigos de equivalência, ao qual são aplicados os dois algoritmos, conforme representado no esquema da figura 5.1.

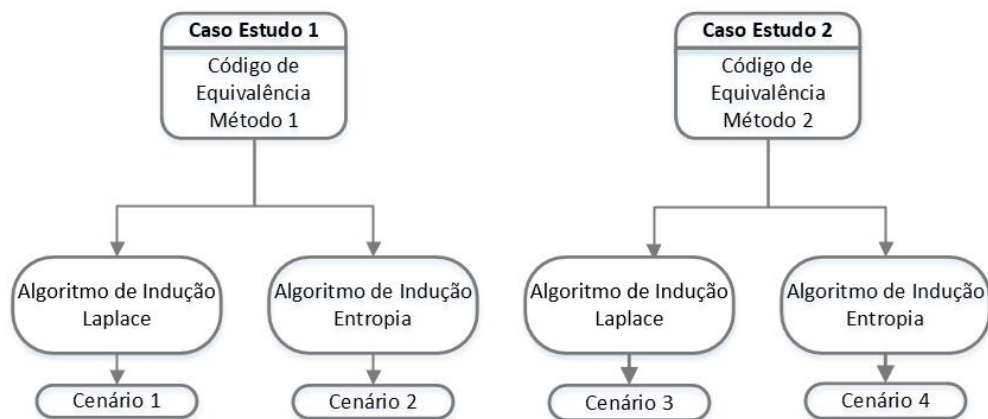


Figura 5.1 – Casos de estudo e respetivos cenários em análise

5.5.3 – Avaliação dos casos de estudo

Para se poder avaliar o desempenho de um modelo de previsão de modo a determinar qual a metodologia mais eficiente é necessário realizar estudos do erro. De modo a avaliar qual o cenário e metodologia mais favorável, é aplicado o critério *Root Mean Square Error*, normalizado com o valor médio medido. Este é dado pela seguinte expressão:

$$RMSE = \frac{\sqrt{\frac{1}{N} \times \sum_1^N (Y_t - \hat{Y}_t)^2}}{\frac{1}{N} \times \sum_1^N Y_t} \quad (5.1)$$

Nesta equação 5.1, Y_t representa o valor real (medido) da grandeza em análise e \hat{Y}_t representa o valor previsto da grandeza em análise, ambos para o instante de tempo t . Ainda nesta equação, N representa o número de amostras consideradas para a análise.

5.5.4 – Carvão - Caso de estudo 1

A primeira análise teve como elemento de estudo o carvão. O primeiro passo deste caso de estudo foi redefinir os valores de cada um dos atributos de condição, definidos no subcapítulo 5.5.1. O atributo de decisão em análise é o CARV, referente aos valores da produção das centrais a Carvão. Esta etapa é realizada de acordo com os códigos de equivalência do primeiro método, presentes nas tabelas 5.2 e 5.3 do subcapítulo 5.3. Na tabela 5.2 encontram-se os códigos de equivalência dos atributos de condição. Na tabela 5.3 estão presentes os códigos de equivalência para o atributo de decisão.

Como referido, o atributo PTEL, representando o índice PTEL, utiliza o mesmo código de equivalência (tabela 5.1) para os dois casos de estudo.

Aplicando estas redefinições aos atributos em estudo, é construída a tabela 5.8 com os dados necessários para o *input* no *software* ROSE2. Estes são apresentados de forma parcial na seguinte tabela 5.8.

Tabela 5.8 – Base de dados relativa ao Carvão, aplicando códigos de equivalência do método 1

Objetos	Atributos				
	PTEL	HID	EOL	CONS	CARV
1	10	330	2920	4860	260
2	30	310	2920	5930	820
3	40	300	1970	6000	1050
4	0	310	2930	5540	180
5	0	320	2700	5290	180
...
...
361	40	150	1720	5470	910
362	20	140	2990	5290	310
363	40	150	3060	6010	1130
364	40	140	2520	6160	1390
365	50	130	570	5700	1700

5.5.4.1 – Algoritmo de Indução Laplace

Através dos dados da tabela 5.8, aplica-se o algoritmo de indução Laplace utilizando o *software* ROSE2. Construiu-se assim um conjunto de regras em que, o conjunto final das regras exatas contém todo o conhecimento da tabela inicial. Este processo é definido como o cenário 1.

As 260 regras geradas têm capacidade para classificar 111 classes diferentes e cada uma dessas classes representa uma produção distinta. Como exemplo foram retiradas cinco regras cujo seu suporte é feito com mais de cinco objetos, apresentando-se seguidamente na sua forma primária.

- Regra 1: (PTEL < 5) & (CONS < 6425) => (CARV = 180); [10, 10, 62.50%, 100.00%];
- Regra 5: (PTEL < 5) & (EOL < 2925) => (CARV = 180); [6, 6, 37.50%, 100.00%];
- Regra 247: (PTEL >= 55) & (HID in [35, 65]) & (EOL < 715) & CONS >= 5605) => (CARV = 1750); [8, 8, 19.51%, 100.00%];
- Regra 248: (HID < 65) & (EOL < 315) & (CONS >= 4970) => (CARV = 1750); [9, 9, 21.95%, 100.00%];
- Regra 249: (PTEL >= 55) & (EOL < 635) & (CONS in [5695, 5835]) => (CARV = 1750); [11, 11, 26.83%, 100.00%].

Analisando os quatro parâmetros com informação adicional foi possível concluir que os parâmetros *absolute cover* e *absolute support* se mantêm iguais, resultando num *level of discrimination* de 100%. Este parâmetro permitiu concluir que os dados de aprendizagem são consistentes.

Estas regras podem, no entanto, ser reescritas da seguinte forma:

- Regra 1: Se (PTEL for inferior a 5 e CONS for inferior a 6425) então CARV = 180, com um suporte absoluto de 10;
- Regra 5: Se (PTEL for inferior a 5 e EOL for inferior a 2925) então CARV = 180, com um suporte absoluto de 6;
- Regra 247: Se (PTEL for maior ou igual a 55 e HID estiver entre 35 e 65 e EOL for inferior a 715 e CONS for maior ou igual a 5605) então CARV = 1750, com um suporte absoluto de 8;
- Regra 248: Se (HID for inferior a 65 e EOL for inferior a 315 e CONS for maior ou igual a 4970) então CARV = 1750, com um suporte absoluto de 9;

- Regra 249: Se (PTEL for maior ou igual a 55 e EOL for inferior a 635 e CONS estiver entre 5695 e 5835) então $CARV = 1750$, com um suporte absoluto de 11.

Estas regras podem também ser escritas de um modo mais compreensível e extenso. Tendo em conta o estudo efetuado, temos que:

Regra 1: Se o índice PTEL tiver um valor inferior a 5 e o valor do consumo de eletricidade for inferior a 6425, então o valor da produção média das centrais a carvão é de 180.

Regra 5: Se o índice PTEL tiver um valor inferior a 5 e o valor da produção das centrais eólicas for inferior a 2925, então o valor da produção média das centrais a carvão é de 180.

Regra 247: Se o índice PTEL for maior ou igual a 55, o valor da produção das centrais mini-hídricas estiver entre o valor 35 e 65, o valor da produção das centrais eólicas for inferior a 715 e o valor do consumo de eletricidade for maior ou igual a 5605, então o valor da produção média das centrais a carvão é de 1750.

Regra 248: Se o valor da produção das centrais mini-hídricas for inferior a 65, o valor da produção das centrais eólicas for inferior a 315 e o valor do consumo de eletricidade for maior ou igual a 4970, então o valor da produção média das centrais a carvão é de 1750.

Regra 249: Se o índice PTEL tiver um valor maior ou igual a 55, o valor da produção das centrais eólicas for inferior a 635 e o valor do consumo de eletricidade estiver entre 5695 e 5835, então o valor da produção média das centrais a carvão é de 1750.

Seguidamente foi feita a validação das regras. Recorrendo ao *software* ROSE2 e utilizando o seu método de validação *Single Pass*, onde é introduzido a base de dados de aprendizagem redefinida (tabela 5.8) e a base de dados de teste onde se encontram os dados de 2015, representada de modo parcial na tabela 5.9. Estes dados de 2015 incluem os atributos cuja previsão para o dia depois de amanhã (D+2) é conhecida. Estes são os atributos HID, EOL, CONS e PTEL, em que HID representa o valor da produção das centrais mini-hídricas, EOL representa o valor da produção das centrais eólicas, CONS o valor do consumo de eletricidade e o PTEL representa o valor do índice PTEL. É também introduzido o ficheiro com as regras de indução geradas.

Tabela 5.9 – Base de dados dos atributos do dia (D+2) referente ao ano 2015

Objetos	Atributos			
	PTEL	HID	EOL	CONS
1	47	126	453	4977
2	54	122	253	5725
3	54	123	133	5547
4	46	116	196	5420
5	60	120	223	6506
...
...
361	30	80	3509	4804
362	36	139	3603	5536
363	47	190	1238	5507
364	41	175	2857	5550
365	48	208	536	5070

Através deste processo é possível classificar todos os dias de 2015 com a sua produção média, permitindo assim que nenhum dos dias fique por classificar. Utilizando os valores previstos foi feito o estudo do erro entre os valores previstos e os valores reais, relativos à produção das centrais a carvão. Este cálculo permitiu determinar um erro de 28,78%

5.5.4.2 – Algoritmo de Indução Entropia

Com base nos dados da tabela 5.8, foi aplicado neste cenário 2 o algoritmo de indução Entropia, utilizando o *software* ROSE2. Este permitiu gerar um conjunto de regras em que, o conjunto final das regras exatas contém todo o conhecimento da tabela 5.8.

As 275 regras geradas classificam 111 classes, cada uma representando um valor de produção distinto. A título de exemplo são apresentadas cinco regras cujo suporte é feito com mais de cinco objetos.

- Regra 1: Se (PTEL for inferior a 5 e CONS for inferior a 6425) então CARV = 180, com um suporte absoluto de 7;
- Regra 259: Se (PTEL estiver entre 55 e 65 e HID for inferior a 65 e EOL estiver entre 445 e 715 e CONS for maior ou igual a 5605) então CARV = 1750, com um suporte absoluto de 7;

- Regra 260: Se (PTEL for maior ou igual a 55 e HID for inferior a 65 e EOL for inferior a 805 e CONS estiver entre 5695 e 5835) então CARV = 1750, com um suporte absoluto de 10;
- Regra 264: Se (HID for inferior a 55 e EOL estiver entre 155 e 345) então CARV = 1750, com um suporte absoluto de 7;
- Regra 268: Se (PTEL for maior ou igual a 65 e HID for maior ou igual a 35) então CARV = 1750, com um suporte absoluto de 6.

Como descrito anteriormente, as regras podem ser escritas de um modo mais extenso. Devido ao grande volume de dados e de regras, foram, no entanto, escolhidas apenas cinco regras e apresentadas de um modo sucinto.

Recorrendo ao *software* ROSE2 e utilizando o seu método de validação *Single Pass* foi possível validar as regras geradas aplicando-as aos dados de 2015, simulando os dados disponíveis para o dia D+2. Como *input* foi introduzido a base de dados de aprendizagem redefinida (tabela 5.8), a base de dados de teste onde se encontram os dados reais de 2015 (tabela 5.9) e as regras de indução geradas anteriormente.

Através deste processo é possível classificar todos os dias de 2015 com a sua produção média, permitindo assim que nenhum dos dias fique por prever. Através destes valores previstos foi feito um estudo do erro entre os valores previstos e os valores reais relativos à produção das centrais a carvão, resultando num erro de 26,41%.

5.5.5 – Carvão - Caso de estudo 2

Neste caso de estudo 2, analisando ainda as previsões para a produção das centrais a carvão, são aplicados os códigos de equivalência do segundo método aos atributos de condição (HID, EOL e CONS) e ao atributo de decisão (CARV). Redefinindo os valores de cada um dos atributos, aplicando os códigos de equivalência presentes nas tabelas 5.5 e 5.6 do subcapítulo 5.3 e mantendo o código de equivalência da tabela 5.1 para o atributo de condição PTEL. Construiu-se assim a tabela 5.10 com os dados necessários de *input* a serem introduzidos no *software* para análise nos dois seguintes cenários em estudo. Os valores redefinidos pela aplicação deste processo de discretização são apresentados de forma parcial nesta tabela 5.10.

Tabela 5.10 – Base de dados relativa ao Carvão, aplicando códigos de equivalência do método 2

Objetos	Atributos				
	PTEL	HID	EOL	CONS	CARV
1	10	300	2900	4900	300
2	30	300	2900	5900	800
3	40	300	2000	6000	1000
4	0	300	2900	5500	200
5	0	300	2700	5300	200
...
...
361	40	200	1700	5500	900
362	20	100	3000	5300	300
363	40	100	3100	6000	1100
364	40	100	2500	6200	1400
365	50	100	600	5700	1700

5.5.5.1 – Algoritmo de Indução Laplace

Neste cenário 3, com base nos dados da tabela 5.10 aplica-se o algoritmo de indução Laplace utilizando o *software* ROSE2. Construiu-se assim um conjunto de 109 regras exatas e 5 aproximadas, capazes de classificar 16 classes diferentes representando produções distintas. Como exemplo são expostas cinco regras cujo suporte é feito com mais de cinco objetos, apresentando-se na sua forma reescrita.

- Regra 1: Se (PTEL for inferior a 5) então CARV = 200, com um suporte absoluto de 13;
- Regra 2: Se (PTEL for inferior a 15 e HID for inferior a 350 e EOL for maior ou igual a 2850 e CONS for maior ou igual a 5150) então CARV = 200, com um suporte absoluto de 9;
- Regra 8: Se (PTEL estiver entre 5 e 25 e EOL estiver entre 2150 e 3050 e CONS estiver entre 4850 e 6350) então CARV = 300, com um suporte absoluto de 11;
- Regra 78: Se (PTEL for maior ou igual a 55 e EOL estiver entre 750 e 1150 e CONS estiver entre 5550 e 5950) então CARV = 1700, com um suporte absoluto de 7;
- Regra 99: Se (PTEL estiver entre 55 e 65 e HID for inferior a 250 e EOL inferior a 450 e CONS for maior ou igual a 5750) então CARV = 1800, com um suporte absoluto de 7.

As cinco regras aproximadas estão descritas seguidamente na sua forma reescrita, para que sejam de mais fácil compreensão.

- Regra 110: Se (PTEL for maior ou igual a 45 e EOL estiver entre 1350 e 1450 e CONS estiver entre 5450 e 5550) então $CARV = 1400$ ou $CARV = 1500$, com um suporte absoluto de 3;
- Regra 111: Se (PTEL for inferior a 45 e HID for inferior a 150 e EOL estiver entre 2150 e 2250) então $CARV = 1300$ ou $CARV = 1600$, com um suporte absoluto de 2;
- Regra 112: Se (HID for maior ou igual a 50 e EOL for inferior a 350 e CONS for inferior a 4650) então $CARV = 1500$ ou $CARV = 1800$, com um suporte absoluto de 2;
- Regra 113: Se (HID for inferior a 50 e EOL estiver entre 550 e 650 e CONS estiver entre 4950 e 5350) então $CARV = 1600$ ou $CARV = 1700$, com um suporte absoluto de 2;
- Regra 114: Se (EOL estiver entre 650 e 950 e CONS estiver entre 4950 e 5050) então $CARV = 1600$ ou $CARV = 1700$, com um suporte absoluto de 4.

Seguidamente foi feita a validação das regras, recorrendo à base de dados de aprendizagem redefinida, à base de dados de teste de 2015, com os atributos cuja previsão para o dia depois de amanhã (D+2) é conhecida, e às regras de indução geradas.

Este cenário 3 permitiu classificar todos os dias de 2015 com a sua produção média. Efetuando o estudo do erro, relativos à produção das centrais a carvão, é possível determinar que o erro foi de 21,15%.

5.5.5.2 – Algoritmo de Indução Entropia

Neste cenário 4, aplicou-se novamente aos dados da tabela 5.8 o algoritmo de indução Entropia. Este permitiu gerar um conjunto de 185 regras exatas e 8 regras aproximadas que representam todo o conhecimento da tabela inicial.

As 193 regras geradas permitem classificar 16 classes distintas, sendo que cada classe representa um valor de produção distinto. A título de exemplo foram seleccionadas cinco regras exatas, cujo suporte é feito com mais de seis objetos.

- Regra 2: Se (PTEL for inferior a 5) então $CARV = 200$, com um suporte absoluto de 13;
- Regra 6: Se (PTEL for inferior a 25 e EOL for maior ou igual a 3050) então $CARV = 200$, com um suporte absoluto de 10;
- Regra 10: Se (PTEL estiver entre 5 e 25 e EOL estiver entre 2150 e 3050 e CONS estiver entre 4850 e 6350) então $CARV = 300$, com um suporte absoluto de 11;
- Regra 143: Se (PTEL for maior ou igual a 45 e EOL estiver entre 750 e 1050 e CONS estiver entre 5550 e 5750) então $CARV = 1700$, com um suporte absoluto de 7;
- Regra 150: Se (PTEL for maior ou igual a 55 e EOL estiver entre 750 e 1050 e CONS estiver entre 5250 e 5950) então $CARV = 1700$, com um suporte absoluto de 7.

Das oito regras aproximadas foram selecionadas como exemplo as cinco seguintes, apresentando-se na sua forma reescrita.

- Regra 186: Se (PTEL for maior ou igual a 45 e EOL estiver entre 1350 e 1450 e CONS estiver entre 5450 e 5550) então $CARV = 1400$ ou $CARV = 1500$, com um suporte absoluto de 3;
- Regra 187: Se (PTEL for inferior a 45 e HID for inferior a 150 e EOL estiver entre 2150 e 2250) então $CARV = 1300$ ou $CARV = 1600$, com um suporte absoluto de 2;
- Regra 188: Se (HID for maior ou igual a 50 e EOL for inferior a 350 e CONS for inferior a 4650) então $CARV = 1500$ ou $CARV = 1800$, com um suporte absoluto de 2;
- Regra 190: Se (EOL estiver entre 650 e 750 e CONS estiver entre 5550 e 5650) então $CARV = 1700$ ou $CARV = 1800$, com um suporte absoluto de 2;
- Regra 191: Se (EOL estiver entre 550 e 950 e CONS estiver entre 4950 e 5050) então $CARV = 1600$ ou $CARV = 1700$, com um suporte absoluto de 4.

As regras geradas são agora aplicadas na previsão da produção das centrais a carvão para o ano 2015. Utilizando a base de dados de teste onde se encontram os dados reais de 2015 (tabela 5.9) e as regras induzidas utilizando o método de indução, foi possível classificar a produção das centrais a carvão para todos os dias D+2, do ano 2015.

Através destes valores previstos foi feito um estudo do erro, obtendo-se um erro de 18,70% para este cenário.

5.5.6 – Carvão – Conclusões

O presente estudo teve como finalidade analisar qual a metodologia mais adequada para prever a produção média de carvão para o dia depois de amanhã. Tendo em conta os resultados apresentados, conclui-se que o cenário 4 gera classificações com um menor erro. Este cenário utiliza o código de equivalência do método 2, com o algoritmo de indução Entropia. O número de regras geradas neste cenário é, no entanto, superior, comparativamente com o cenário 3. O erro neste cenário 3 é apenas 2,46% superior ao cenário 4. Se o esforço computacional necessário para efetuar a previsão for tido em conta, o cenário 4 poderá não compensar uma vez que tem mais 81 regras para testar, comparativamente com o cenário 3.

5.5.7 – Gás Natural - Caso de estudo 1

A seguinte análise tem como intuito o estudo da previsão da produção das centrais a Gás Natural, para o dia depois de amanhã. Por este motivo o atributo de decisão são os valores da produção das centrais a Gás Natural, representado por GN. Os atributos de condição mantêm-se inalterados, sendo eles o PTEL, HID, EOL e CONS, em que HID representa o valor da produção das centrais mini-hídricas, EOL representa o valor da produção das centrais eólicas, CONS representa o valor do consumo de eletricidade e o PTEL representa o valor do índice PTEL.

Inicialmente, para o caso estudo 1 foi aplicado o código de equivalência do primeiro método. Esta etapa é realizada de acordo com os códigos presentes nas tabelas 5.2 e 5.3 do subcapítulo 5.3, permitindo a redefinição dos valores de cada um dos atributos em análise. O atributo PTEL utiliza o mesmo código de equivalência (tabela 5.1) para os dois casos de estudo.

Aplicando estas redefinições aos atributos em estudo, é construída a tabela 5.11 com os dados necessários para o *input* no *software* ROSE2. Os valores redefinidos pela aplicação dos códigos de equivalência são apresentados nesta tabela de forma parcial.

Tabela 5.11 – Base de dados relativa ao Gás Natural, aplicando códigos de equivalência do método 1

Objetos	Atributos				
	PTEL	HID	EOL	CONS	GN
1	10	330	2920	4860	0
2	30	310	2920	5930	0
3	40	300	1970	6000	0
4	0	310	2930	5540	0
5	0	320	2700	5290	30
...
...
361	40	150	1720	5470	0
362	20	140	2990	5290	50
363	40	150	3060	6010	40
364	40	140	2520	6160	90
365	50	130	570	5700	60

5.5.7.1 – Algoritmo de Indução Laplace

Recorrendo aos dados da tabela 5.11, foi testado a aplicação do algoritmo de indução Laplace. Este permitiu gerar um conjunto de regras em que, o conjunto final das regras exatas contém todo o conhecimento da tabela inicial. Este processo foi definido como o cenário 1, relativo à produção das centrais a Gás Natural.

Este algoritmo de indução gerou 185 regras, aptas para classificar 69 classes diferentes, sendo que cada uma dessas classes representa uma produção distinta. A título de exemplo são apresentadas cinco regras, cujo suporte é feito com mais de treze objetos, apresentando-se reescritas de uma forma mais compreensível.

- Regra 1: Se (PTEL for inferior a 45 e HID for inferior a 285 e EOL estiver entre 1665 e 2235 e CONS for inferior a 5645) então GN = 0, com um suporte absoluto de 14;
- Regra 2: Se (HID estiver entre 185 e 315 e EOL for maior ou igual a 1315 e CONS for inferior a 5615) então GN = 0, com um suporte absoluto de 14;
- Regra 3: Se (PTEL for superior a 45 e HID estiver entre 55 e 255 e EOL estiver entre 920 e 1545) então GN = 0, com um suporte absoluto de 21;
- Regra 9: Se (PTEL estiver entre 5 e 55 e HID estiver entre 75 e 335 e EOL estiver entre 565 e 2925 e CONS estiver entre 4775 e 5410) então GN = 0, com um suporte absoluto de 31;

- **Regra 15:** Se (PTEL estiver entre 5 e 45 e HID estiver entre 205 e 325 e EOL for maior ou igual a 835 e CONS for inferior a 5615) então $GN = 0$, com um suporte absoluto de 18.

As regras apresentadas como exemplo são todas relativas à classe 0. Isto deve-se ao facto de haver 148 objetos pertencentes a esta classe, o que faz com que haja um número elevado de regras geradas para classificar esta classe.

A validação das regras foi feita recorrendo ao método de validação *Single Pass*, onde é introduzido a base de dados de aprendizagem redefinida (tabela 5.11), a base de dados de teste onde se encontram os atributos de condição referentes ao ano 2015 e as respetivas regras.

Através deste método de validação é possível classificar todos os dias de 2015 com a sua produção média, permitindo assim que nenhum dos dias fique por prever. Observando os valores previstos e os reais foi calculado um erro de 85,48% relativo à previsão da produção das centrais a Gás Natural.

5.5.7.2 – Algoritmo de Indução Entropia

Neste cenário 2 aplica-se o algoritmo de indução Entropia aos dados da tabela 5.11. Este gerou um conjunto de 212 regras divididas em 69 classes, cada uma representando uma produção diferente.

A título de exemplo foram escolhidas cinco regras com um suporte superior a cinco objetos.

- **Regra 2:** Se (PTEL for inferior a 45 e HID estiver entre 185 e 315 e EOL for maior ou igual a 645 e CONS for inferior a 5615) então $GN = 0$, com um suporte absoluto de 22;
- **Regra 6:** Se (PTEL for inferior a 45 e HID for maior ou igual a 55 e CONS for inferior a 5180) então $GN = 0$, com um suporte absoluto de 36;
- **Regra 11:** Se (PTEL for inferior a 55 e CONS estiver entre 4775 e 4910) então $GN = 0$, com um suporte absoluto de 17;
- **Regra 13:** Se (PTEL for maior ou igual a 25 e EOL for maior ou igual a 1435 e CONS for inferior a 5410) então $GN = 0$, com um suporte absoluto de 17;
- **Regra 29:** Se (CONS for inferior a 4555) então $GN = 0$, com um suporte absoluto de 15.

As regras geradas foram validadas utilizando os dados de 2015, que permitiram simular uma previsão para o dia depois de amanhã. Como *input* foi introduzida a base de dados de aprendizagem, a base de dados de teste e as regras de indução geradas. Este processo permitiu classificar todos os dias do ano 2015 com a sua produção média. Através destes valores previstos foi feito um estudo do erro, obtendo-se um erro de 84,49%.

5.5.8 – Gás Natural - Caso de estudo 2

Neste caso de estudo 2, analisando ainda as previsões para a produção das centrais a gás natural, aplicaram-se os códigos de equivalência do segundo método aos atributos de condição (HID, EOL e CONS) e ao atributo de decisão (GN). Redefinindo os valores de cada um dos atributos aplicando os códigos de equivalência presentes nas tabelas 5.5 e 5.6 do subcapítulo 5.3, construiu-se a tabela 5.12. Esta tabela apresentada de forma parcial, possui os dados necessários para análise dos dois seguintes cenários em estudo.

Tabela 5.12 – Base de dados relativa ao Gás Natural, aplicando códigos de equivalência do método 2

Objetos	Atributos				
	PTEL	HID	EOL	CONS	GN
1	10	300	2900	4900	0
2	30	300	2900	5900	0
3	40	300	2000	6000	0
4	0	300	2900	5500	0
5	0	300	2700	5300	0
...
...
361	40	200	1700	5500	0
362	20	100	3000	5300	100
363	40	100	3100	6000	0
364	40	100	2500	6200	100
365	50	100	600	5700	100

5.5.8.1 – Algoritmo de Indução Laplace

Com base nos dados da tabela 5.12, aplica-se o algoritmo de indução Laplace. Construiu-se assim um conjunto de 82 regras, divididas em 76 regras exatas e 6 regras aproximadas. Estas têm capacidade para classificar 12 classes diferentes, representando cada uma, uma produção

distinta. Como exemplo foram selecionadas cinco regras exatas cujo suporte é feito com mais de vinte e um objetos. Estas apresentam-se seguidamente na sua forma reescrita.

- Regra 1: Se (PTEL for inferior a 45 e CONS for inferior a 5050) então $GN = 0$, com um suporte absoluto de 40;
- Regra 2: Se (PTEL estiver entre 35 e 45 e EOL for inferior a 2650 e CONS for inferior a 5750) então $GN = 0$, com um suporte absoluto de 48;
- Regra 4: Se (PTEL for inferior a 55 e HID for inferior a 250 e EOL for maior ou igual a 650 e CONS for inferior a 5150) então $GN = 0$, com um suporte absoluto de 42;
- Regra 10: Se (PTEL for inferior a 45 e HID estiver entre 50 e 250 e EOL estiver entre 650 e 1550) então $GN = 0$, com um suporte absoluto de 30;
- Regra 29: Se (PTEL for inferior a 55 e EOL for maior ou igual a 450 e CONS estiver entre 4850 e 5150) então $GN = 0$, com um suporte absoluto de 22.

Foram ainda selecionadas como exemplo cinco das seis regras aproximadas, descritas na sua forma reescrita para que sejam de mais fácil compreensão.

- Regra 77: Se (PTEL for maior ou igual a 45 e EOL estiver entre 1350 e 1450 e CONS estiver entre 5450 e 5550) então $GN = 0$ ou $GN = 100$ ou $GN = 200$, com um suporte absoluto de 3;
- Regra 78: Se (HID for maior ou igual a 50 e EOL for inferior a 350 e CONS for inferior a 4650) então $GN = 0$ ou $GN = 100$, com um suporte absoluto de 2;
- Regra 79: Se (PTEL for inferior a 55 e EOL estiver entre 550 e 650 e CONS for maior ou igual a 5750) então $GN = 300$ ou $GN = 400$, com um suporte absoluto de 3;
- Regra 80: Se (HID for inferior a 50 e EOL estiver entre 1050 e 1150 e CONS for maior ou igual a 5550) então $GN = 100$ ou $GN = 600$, com um suporte absoluto de 2;
- Regra 82: Se (PTEL for maior ou igual a 55 e HID for maior ou igual a 50 e EOL for inferior a 350 e CONS estiver entre 5550 e 5650) então $GN = 300$ ou $GN = 500$, com um suporte absoluto de 2.

Seguidamente foi feita a validação das regras, introduzindo a base de dados de aprendizagem redefinida (tabela 5.12) e a base de dados de teste de 2015, representando os atributos cuja previsão para o dia depois de amanhã (D+2) é conhecida. Aplicando as regras de indução geradas é possível obter os resultados para este cenário 3.

O processo foi capaz de classificar todos os dias de 2015 com a sua produção média. Efetuando o estudo do erro entre os valores previstos e os valores reais, relativos à produção das centrais a Gás Natural, é possível determinar que o erro foi de 80,97%.

5.5.8.2 – Algoritmo de Indução Entropia

Neste cenário 4, de acordo com a base de dados da tabela 5.12, é aplicado o algoritmo de indução Entropia. Este permitiu construir um conjunto de regras em que o conjunto final das regras exatas e das regras aproximadas contém todo o conhecimento da tabela inicial. Esta análise gerou a informação essencial para o cenário 4.

As 152 regras geradas são compostas por 143 regras exatas e 9 regras aproximadas. Estas têm capacidade para classificar 12 classes distintas, representando os valores de produção. A título de exemplo apresentam-se cinco regras cujo suporte é feito com mais de dez objetos.

- Regra 1: Se (PTEL estiver entre 35 e 45 e EOL for inferior a 2250 e CONS for inferior a 5750) então GN = 0, com um suporte absoluto de 48;
- Regra 3: Se (PTEL estiver entre 35 e 45 e EOL for maior ou igual a 1750 e CONS estiver entre 5550 e 6150) então GN = 0, com um suporte absoluto de 11;
- Regra 4: Se (PTEL estiver entre 5 e 35 e HID for maior ou igual a 150 e EOL for maior ou igual a 750 e CONS for inferior a 5450) então GN = 0, com um suporte absoluto de 21;
- Regra 10: Se (PTEL estiver entre 25 e 45 e EOL for inferior a 1350 e CONS estiver entre 5550 e 5850) então GN = 0, com um suporte absoluto de 14;
- Regra 27: Se (EOL for maior ou igual a 450 e CONS for inferior a 4650) então GN = 0, com um suporte absoluto de 19.

Das nove regras aproximadas foram selecionadas cinco regras capazes de classificar nove classes distintas. Estas são apresentadas na sua forma reescrita.

- Regra 144: Se (PTEL for maior ou igual a 45 e EOL estiver entre 1350 e 1450 e CONS estiver entre 5450 e 5550) então GN = 0 ou GN = 100 ou GN = 200, com um suporte absoluto de 3;
- Regra 146: Se (EOL estiver entre 350 e 450 e CONS estiver entre 5850 e 5950) então GN = 900 ou GN = 1100, com um suporte absoluto de 2;

- Regra 147: Se (PTEL for inferior a 55 e EOL estiver entre 550 e 650 e CONS for maior ou igual a 5750) então $GN = 300$ ou $GN = 400$, com um suporte absoluto de 3;
- Regra 151: Se (EOL estiver entre 550 e 650 e CONS estiver entre 5150 e 5250) então $GN = 300$ ou $GN = 600$, com um suporte absoluto de 2;
- Regra 152: Se (PTEL for maior ou igual a 55 e HID for maior ou igual a 50 e EOL for inferior a 350 e CONS estiver entre 5550 e 5650) então $GN = 300$ ou $GN = 500$, com um suporte absoluto de 2.

As regras geradas são agora aplicadas na previsão da produção das centrais a Gás Natural para o ano 2015, simulando assim as previsões para o dia depois de amanhã referente a esse ano. Aplicando o método de validação foi possível classificar todos os dias com a respetiva produção das centrais a Gás Natural. Efetuando o estudo do erro determinou-se um erro de 84,49%.

5.5.9 – Gás Natural – Conclusões

O presente estudo teve como finalidade analisar qual a metodologia mais adequada para prever a produção média do gás natural para o dia depois de amanhã. Tendo em conta os vários cenários e resultados, conclui-se que o cenário 4 apresenta um menor erro em comparação com os restantes cenários. Este cenário utiliza o código de equivalência do método 2, com o algoritmo de indução Entropia. Apesar de ter um menor erro este utiliza mais 70 regras, em comparação com o cenário 3, e o seu erro apenas melhora em 0,93%.

5.5.10 – Albufeira - Caso de estudo 1

A análise realizada neste caso de estudo pretende prever a produção das Albufeiras para o dia depois de amanhã. O primeiro passo deste caso de estudo foi redefinir os valores de cada um dos atributos de condição e do atributo de decisão. O atributo decisão é definido pelo acrónimo ALB e representa a produção das centrais de Albufeira. Esta etapa de redefinição dos valores da base de dados de aprendizagem é realizada de acordo com os códigos de equivalência do primeiro método, presentes nas tabelas 5.2 e 5.4 do subcapítulo 5.3. Na tabela 5.2 encontram-se os códigos de equivalência dos atributos de condição e na tabela 5.4 o código de equivalência do atributo de decisão. O atributo PTEL, representando o índice PTEL, utiliza o mesmo código de equivalência (tabela 5.1) para os dois casos de estudo.

Aplicando estas redefinições aos atributos em estudo é construída a tabela 5.13, apresentando parcialmente os dados aplicados.

Tabela 5.13 – Base de dados relativa às Albufeiras, aplicando códigos de equivalência do método 1

Objetos	Atributos				
	PTEL	HID	EOL	CONS	ALB
1	10	330	2920	4860	400
2	30	310	2920	5930	960
3	40	300	1970	6000	1080
4	0	310	2930	5540	940
5	0	320	2700	5290	900
6	10	320	2720	6280	1380
...
351	50	210	1530	6380	780
362	20	140	2990	5290	510
363	40	150	3060	6010	1070
364	40	140	2520	6160	1200
365	50	130	570	5700	940

5.5.10.1 – Algoritmo de Indução Laplace

O primeiro algoritmo a ser aplicado foi o de indução Laplace, definindo-se assim o cenário 1. De acordo com este algoritmo foram geradas um conjunto de 301 regras exatas, divididas em 138 classes diferentes. Cada uma dessas classes representa uma produção distinta.

Como exemplo são apresentadas cinco regras com um suporte absoluto de três objetos.

- Regra 56: Se (HID estiver entre 75 e 175 e EOL estiver entre 1510 e 1545) então ALB = 360, com um suporte absoluto de 3;
- Regra 79: Se (PTEL for inferior a 55 e HID for inferior a 95 e EOL estiver entre 595 e 1665 e CONS estiver entre 5675 e 5785) então ALB = 420, com um suporte absoluto de 3;
- Regra 90: Se (HID for inferior a 45 e EOL for inferior 935 e CONS estiver entre 5225 e 5565) então ALB = 450, com um suporte absoluto de 3;
- Regra 182: Se (HID estiver entre 105 e 125 e CONS estiver entre 5535 e 5555) então ALB = 790, com um suporte absoluto de 3;

- Regra 232: Se (PTEL for inferior a 35 e HID for maior ou igual a 295 e EOL estiver entre 685 e 2195 e CONS estiver entre 5155 e 5425) então ALB = 1010, com um suporte absoluto de 3.

Aplicando o método de validação *Single Pass* do *software* ROSE2 é feita a validação das regras utilizando como *input* a base de dados de aprendizagem redefinida (tabela 5.13), a base de dados de teste de 2015 (tabela 5.9) e as regras geradas.

Através deste cenário é possível classificar todos os dias de 2015, simulando assim a previsão da produção média das centrais de Albufeira para o dia depois de amanhã. Utilizando os valores previstos foi feito um estudo do erro entre os valores previstos e os reais, obtendo-se um erro de 104,66%.

5.5.10.2 – Algoritmo de Indução Entropia

Neste cenário 2 foi aplicado o algoritmo de indução Entropia à base de dados da tabela 5.13. Este permitiu gerar um conjunto de 313 regras em que o conjunto final das regras exatas contém o conhecimento presente na base de dados inicial.

A aplicação deste algoritmo à base de dados codificada pelo código de equivalência do método 1 permitiu classificar 138 conjuntos distintos.

A título de exemplo foram selecionadas cinco regras com um suporte absoluto de três objetos.

- Regra 56: Se (EOL estiver entre 1510 e 1545 e CONS for inferior a 4845) então ALB = 360, com um suporte absoluto de 3;
- Regra 92: Se (HID for inferior a 45 e EOL for inferior a 935 e CONS estiver entre 5225 e 5565) então ALB = 450, com um suporte absoluto de 3;
- Regra 189: Se (HID estiver entre 105 e 125 e CONS estiver entre 5535 e 5555) então ALB = 790, com um suporte absoluto de 3;
- Regra 243: Se (PTEL for maior ou igual a 15 e HID for maior ou igual a 325 e EOL for maior ou igual a 685) então ALB = 1010, com um suporte absoluto de 3;
- Regra 252: Se (CONS estiver entre 6025 e 6065) então ALB = 1060, com um suporte absoluto de 3.

As regras foram validadas aplicando-as aos dados de 2015 de modo a simular a previsão para o dia depois de amanhã. Através deste processo a produção média das albufeiras foi prevista para todos os dias de 2015, obtendo-se um erro de 106,58%.

5.5.11 – Albufeira - Caso de estudo 2

Neste caso de estudo 2, analisando ainda as previsões para a produção das centrais de albufeira, aplicaram-se os códigos de equivalência do segundo método aos atributos de condição (HID, EOL e CONS) e ao atributo de decisão (ALB). Redefinindo os valores de cada um dos atributos aplicando os códigos de equivalência presentes nas tabelas 5.5 e 5.7 do subcapítulo 5.3, construiu-se a tabela 5.14. Esta tabela apresenta os dados necessários de *input* a serem introduzidos no *software* para a análise a realizar nos dois seguintes cenários em estudo. Os valores redefinidos pela aplicação dos códigos de equivalência são apresentados de forma parcial nesta tabela 5.14.

Tabela 5.14 – Base de dados relativa às Albufeiras, aplicando códigos de equivalência do método 2

Objetos	Atributos				
	PTEL	HID	EOL	CONS	ALB
1	10	300	2900	4900	400
2	30	300	2900	5900	1000
3	40	300	2000	6000	1100
4	0	300	2900	5500	900
5	0	300	2700	5300	900
...
...
361	40	200	1700	5500	500
362	20	100	3000	5300	500
363	40	100	3100	6000	1100
364	40	100	2500	6200	1200
365	50	100	600	5700	900

5.5.11.1 – Algoritmo de Indução Laplace

Com base nos dados da tabela 5.14, aplica-se o algoritmo de indução Laplace. Construiu-se assim um conjunto de 154 regras, divididas em 147 regras exatas e 7 aproximadas. Estas regras

têm capacidade para classificar 21 classes diferentes, representando cada uma, uma produção distinta.

Como exemplo são apresentadas cinco regras exatas, cujo suporte é feito com quatro objetos. Estas apresentam-se seguidamente na sua forma reescrita.

- Regra 3: Se (HID for inferior a 50 e EOL for maior ou igual a 650 e CONS estiver entre 4550 e 4750) então $ALB = 200$, com um suporte absoluto de 4;
- Regra 9: Se (PTEL for inferior a 55 e HID for inferior a 50 e EOL for inferior a 550 e CONS estiver entre 4750 e 5750) então $ALB = 300$, com um suporte absoluto de 4;
- Regra 28: Se (PTEL for maior ou igual a 55 e EOL estiver entre 850 e 1150 e CONS estiver entre 5450 e 5650) então $ALB = 400$, com um suporte absoluto de 4;
- Regra 40: Se (PTEL for maior ou igual a 55 e HID for maior ou igual a 50 e EOL estiver entre 850 e 1450 e CONS estiver entre 5050 e 5750) então $ALB = 400$, com um suporte absoluto de 4;
- Regra 101: Se (PTEL estiver entre 25 e 35 e EOL for maior ou igual a 1450 e CONS estiver entre 5550 e 5950) então $ALB = 1000$, com um suporte absoluto de 4.

Foram ainda seleccionadas cinco das sete regras aproximadas, descritas na sua forma reescrita, para que sejam de mais fácil compreensão.

- Regra 148: Se (PTEL for maior ou igual a 45 e EOL estiver entre 1350 e 1450 e CONS estiver entre 5450 e 5550) então $ALB = 600$ ou $ALB = 800$, com um suporte absoluto de 3;
- Regra 150: Se (PTEL for inferior a 55 e HID for inferior a 50 e EOL estiver entre 550 e 650 e CONS for maior ou igual a 5150) então $ALB = 400$ ou $ALB = 500$, com um suporte absoluto de 5;
- Regra 151: Se (PTEL for maior ou igual a 45 e HID for maior ou igual a 50 e EOL for inferior a 350 e CONS estiver entre 5550 e 5650) então $ALB = 500$ ou $ALB = 700$, com um suporte absoluto de 2;
- Regra 153: Se (PTEL for maior ou igual a 55 e HID for inferior a 50 e EOL estiver entre 350 e 450 e CONS for maior ou igual a 5850) então $ALB = 300$ ou $ALB = 400$, com um suporte absoluto de 2;

- Regra 154: Se (HID for inferior a 50 e EOL for inferior a 850 e CONS estiver entre 4950 e 5050) então $ALB = 200$ ou $ALB = 300$, com um suporte absoluto de 2.

Seguidamente foi feita a validação das regras aplicando-as à base de dados de teste de 2015 (tabela 5.9), que representa os dados disponíveis para o dia depois de amanhã (D+2).

Aplicando as regras de indução geradas é possível obter os resultados para este cenário 3. Este cenário é capaz de classificar a produção média das Albufeiras para o dia depois de amanhã, referente a 2015. Efetuando o estudo do erro é possível determinar que este foi de 113,84%.

5.5.11.2 – Algoritmo de Indução Entropia

Neste cenário 4 o algoritmo de indução Entropia é aplicado à base de dados da tabela 5.14. A aplicação deste algoritmo permitiu construir um conjunto de regras, constituído por 221 regras exatas e 14 regras aproximadas.

Estas 235 regras permitem classificar 21 classes distintas, cada uma representando um valor de produção. A título de exemplo foram seleccionadas cinco regras, apresentadas seguidamente na sua forma simplificada.

- Regra 4: Se (HID for inferior a 50 e EOL for maior ou igual a 650 e CONS estiver entre 4550 e 4750) então $ALB = 200$, com um suporte absoluto de 4;
- Regra 14: Se (HID for inferior a 50 e EOL for maior ou igual a 1350 e CONS estiver entre 5350 e 5550) então $ALB = 300$, com um suporte absoluto de 4;
- Regra 22: Se (PTEL for maior ou igual a 35 e HID for maior ou igual a 50 e EOL estiver entre 1150 e 1450 e CONS for inferior a 5350) então $ALB = 300$, com um suporte absoluto de 4;
- Regra 60: Se (PTEL for maior ou igual a 45 e HID for inferior a 50 e EOL estiver entre 650 e 750) então $ALB = 500$, com um suporte absoluto de 4;
- Regra 156: Se (PTEL estiver entre 35 e 55 e HID for maior ou igual a 150 e EOL estiver entre 250 e 1050 e CONS estiver entre 5050 e 5550) então $ALB = 1000$, com um suporte absoluto de 4.

Como exemplo foram ainda selecionadas cinco regras aproximadas capazes de classificar sete classes distintas. Estas são apresentadas na sua forma reescrita.

- Regra 222: Se (PTEL for maior ou igual a 45 e EOL estiver entre 1350 e 1450 e CONS estiver entre 5450 e 5550) então ALB = 600 ou ALB = 800, com um suporte absoluto de 3;
- Regra 223: Se (PTEL for maior ou igual a 45 e EOL estiver entre 550 e 750 e CONS estiver entre 5150 e 5250) então ALB = 400 ou ALB = 500, com um suporte absoluto de 2;
- Regra 227: Se (HID for inferior a 150 e EOL estiver entre 1250 e 1350 e CONS estiver entre 5450 e 5550) então ALB = 500 ou ALB = 700, com um suporte absoluto de 2;
- Regra 233: Se (EOL estiver entre 350 e 450 e CONS estiver entre 5850 e 5950) então ALB = 300 ou ALB = 400, com um suporte absoluto de 2;
- Regra 234: Se (HID for inferior a 50 e EOL for inferior a 850 e CONS estiver entre 4950 e 5050) então ALB = 200 ou ALB = 300, com um suporte absoluto de 2;

As regras geradas são aplicadas na previsão da produção das centrais de Albufeira, para o ano 2015. Aplicando estas regras à base de dados de teste onde se encontram os dados de 2015, que representam os valores previstos para o dia depois de amanhã, é possível classificar a produção para todos esses dias. Comprova-se assim a capacidade destas regras classificarem a produção média para o dia depois de amanhã. Com base nos valores previstos e nos valores reais foi determinado um erro de 115,11% relativo a este cenário.

5.5.12 – Albufeira – Conclusões

O presente estudo teve como finalidade analisar qual a metodologia mais adequada para prever a produção média das centrais de Albufeira para o dia depois de amanhã. Das quatro metodologias testadas conclui-se que o cenário 1 apresenta melhores resultados. Esta conclusão é feita comparando o erro dos vários cenários. Este cenário utiliza o código de equivalência do método 1, com o algoritmo de indução Laplace. Apesar de apresentar o menor erro, este cenário apresenta um número elevado de regras, resultando num esforço computacional acrescido.

5.5.13 – Fio-de-água - Caso de estudo 1

Os valores da produção recorrendo a centrais a Fio-de-água são também uma incógnita para o dia depois de amanhã (D+2). Nesta secção é feita a discretização dos atributos em análise, de modo a serem aplicados os dois algoritmos em estudo. É definido como atributo de decisão os valores da produção das centrais a Fio-de-água, doravante designado pelo acrónimo FA. O primeiro passo foi redefinir os valores de cada um dos atributos de condição e de decisão. Esta etapa é realizada de acordo com os códigos de equivalência do primeiro método, presentes nas tabelas 5.2 e 5.4 do subcapítulo 5.3. Na tabela 5.2 encontram-se os códigos de equivalência dos atributos de condição e na tabela 5.3 os códigos referentes ao atributo de decisão. O atributo de condição representando o índice PTEL, utiliza o mesmo código de equivalência (tabela 5.1) para os dois métodos. Aplicando as redefinições aos atributos em estudo é construída a tabela 5.15.

Tabela 5.15 – Base de dados relativa às Fio-de-água, aplicando os códigos de equivalência do método 1

Objetos	Atributos				
	PTEL	HID	EOL	CONS	FA
1	10	330	2920	4860	1140
2	30	310	2920	5930	1340
3	40	300	1970	6000	1460
4	0	310	2930	5540	1730
5	0	320	2700	5290	1780
...
...
361	40	150	1720	5470	750
362	20	140	2990	5290	560
363	40	150	3060	6010	750
364	40	140	2520	6160	900
365	50	130	570	5700	1080

5.5.13.1 – Algoritmo de Indução Laplace

No primeiro estudo efetuado com os dados da tabela 5.15, foi aplicado o algoritmo de indução Laplace utilizando o *software* ROSE2. Construiu-se assim um conjunto de regras exatas contendo todo o conhecimento da tabela 5.15. Este primeiro estudo foi definido como o cenário 1.

As 301 regras geradas classificam 155 classes diferentes, sendo que cada uma dessas classes representa uma produção distinta. Como exemplo, foram retiradas cinco regras cujo parâmetro

suporte absoluto tem mais do que um objeto. Estas apresentam-se seguidamente reescritas de uma forma mais compreensível.

- Regra 11: Se (HID for inferior a 35 e EOL estiver entre 985 e 1010) então FA = 280, com um suporte absoluto de 2;
- Regra 23: Se (HID for inferior a 65 e EOL for maior ou igual a 585 e CONS estiver entre 4775 e 4830) então FA = 320, com um suporte absoluto de 3;
- Regra 42: Se (PTEL estiver entre 45 e 55 e HI estiver entre 35 e 45 e EOL estiver entre 645 e 1365 e CONS for maior ou igual a 5180) então FA = 360, com um suporte absoluto de 3;
- Regra 224: Se (PTEL estiver entre 15 e 25 e EOL estiver entre 2025 e 2415 e CONS for maior ou igual a 5555) então FA = 1550, com um suporte absoluto de 3;
- Regra 284: Se (PTEL for inferior a 35 e HID for inferior a 335 e EOL estiver entre 1465 e 2195 e CONS for maior ou igual a 5685) então FA = 2050, com um suporte absoluto de 3.

Seguidamente foi feita a validação das regras recorrendo ao método de validação *Single Pass* do programa. Para isso foi introduzida a base de dados de aprendizagem redefinida (tabela 5.15) e a base de dados de teste onde se encontram os dados de 2015 (tabela 5.9). Estes dados de 2015 incluem os atributos cuja previsão para o dia depois de amanhã (D+2) é conhecida. São eles os atributos HID, EOL, CONS e PTEL, em que HID representa o valor da produção das centrais mini-hídricas, EOL representa o valor da produção das centrais eólicas, CONS o valor do consumo de eletricidade e o PTEL representa o valor do índice PTEL.

Aplicando as regras de indução geradas anteriormente, é possível classificar todos os objetos referentes ao ano de 2015. Este cenário 1 permite então que nenhum dos dias fique com a sua produção proveniente das centrais a Fio-de-água por prever. Utilizando os valores previstos e os reais foi feito um estudo do erro, o qual permitiu concluir que o erro foi de 75,34%.

5.5.13.2 – Algoritmo de Indução Entropia

Na presente análise foi aplicado o algoritmo de indução Entropia à base de dados da tabela 5.15. Este permitiu gerar um conjunto de regras exatas que contêm todo o conhecimento da base de dados inicial.

O cenário 2 surge com a aplicação deste algoritmo à base de dados discretizada, através do código de equivalência do método 1. As 324 regras geradas permitem classificar 155 conjuntos distintos, cada um representando um valor de produção diferente. A título de exemplo foram retiradas cinco regras com um suporte absoluto de dois e três objetos. Estas foram reescritas para que sejam de mais fácil compreensão.

- Regra 24: Se (HID for inferior a 65 e EOL maior ou igual a 585 e CONS estiver entre 4775 e 4830) então FA = 320, com um suporte absoluto de 3;
- Regra 44: Se (HID for inferior a 145 e EOL estiver entre 1165 e 1395 e CONS for inferior a 4775) então FA = 360, com um suporte absoluto de 3;
- Regra 121: Se (HID estiver entre 45 e 155 e EOL for maior ou igual a 1215 e CONS estiver entre 5215 e 5425) então FA = 560, com um suporte absoluto de 2;
- Regra 201: Se (PTEL for maior ou igual a 45 e HID for inferior a 175 e CONS for maior ou igual a 6145) então FA = 1030, com um suporte absoluto de 2;
- Regra 295: Se (HID for maior ou igual a 335 e EOL estiver entre 1825 e 1860) então FA = 1980, com um suporte absoluto de 2.

As regras foram validadas recorrendo à base de dados de 2015, onde estão presentes os atributos cuja previsão para o dia depois de amanhã é conhecida. Estes atributos são essenciais para a aplicação das regras e posterior classificação do atributo de decisão. Através deste processo, a produção média das centrais a Fio-de-água foi prevista para todos os dias de 2015, do qual se obteve um erro de 87,50%.

5.5.14 – Fio-de-água - Caso de estudo 2

Analisando ainda as previsões para a produção das centrais a Fio-de-água, são aplicados os códigos de equivalência do segundo método aos atributos de condição e ao atributo de decisão. Esta aplicação dos códigos de equivalência é feita através dos códigos presentes nas tabelas 5.5 e 5.7 do subcapítulo 5.3 e mantendo o mesmo código de equivalência da tabela 5.1 para o atributo de condição PTEL. Este processo de discretização permitiu elaborar a tabela 5.16. Esta tabela apresenta os dados necessários para a análise dos dois seguintes cenários, cenário 3 e cenário 4.

Tabela 5.16 – Base de dados relativa às Fio-de-água, aplicando códigos de equivalência do método 2

Objetos	Atributos				
	PTEL	HID	EOL	CONS	FA
1	10	300	2900	4900	1100
2	30	300	2900	5900	1300
3	40	300	2000	6000	1500
4	0	300	2900	5500	1700
5	0	300	2700	5300	1800
...
...
361	40	200	1700	5500	800
362	20	100	3000	5300	600
363	40	100	3100	6000	700
364	40	100	2500	6200	900
365	50	100	600	5700	1100

5.5.14.1 – Algoritmo de Indução Laplace

No cenário 3, tendo como base os dados da tabela 5.16, aplica-se o algoritmo de indução Laplace. Determinou-se assim um conjunto de regras exatas e aproximadas que representam todo o conhecimento da tabela inicial.

Estas regras perfazem um total de 159, sendo que estão divididas em 147 regras exatas e 12 regras aproximadas. Estas regras têm capacidade para classificar 22 classes diferentes, em que cada uma representa uma produção distinta.

Como exemplo foram retiradas cinco regras exatas, com um suporte absoluto de quatro e cinco objetos. Estas apresentam-se seguidamente na sua forma reescrita.

- Regra 10: Se (HID for inferior a 50 e EOL for maior ou igual a 550 e CONS estiver entre 4750 e 4950) então FA = 300, com um suporte absoluto de 5;
- Regra 24: Se (PTEL estiver entre 35 e 45 e HID for inferior a 150 e EOL estiver entre 950 e 1750 e CONS estiver entre 4750 e 5750) então FA = 400, com um suporte absoluto de 5;
- Regra 68: Se (PTEL for maior ou igual a 5 e HID for maior ou igual a 50 e EOL for maior ou igual a 3050 e CONS for inferior a 6050) então FA = 700, com um suporte absoluto de 5;
- Regra 73: Se (PTEL for maior ou igual a 35 e HID for maior ou igual a 250 e EOL for maior ou igual a 2850) então FA = 700, com um suporte absoluto de 4;

- Regra 114: Se (PTEL for maior ou igual a 15 e EOL estiver entre 1950 e 2250 e CONS estiver entre 6450 e 6650) então FA = 1800, com um suporte absoluto de 4.

Foram ainda selecionadas cinco das doze regras aproximadas, descritas na sua forma reescrita, para uma compreensão mais simples. Estas têm capacidade para classificar 8 classes distintas.

- Regra 148: Se (PTEL for maior ou igual a 35 e HID for maior ou igual a 150 e EOL estiver entre 750 e 850 e CONS for inferior a 5550) então FA = 900 ou FA = 1800, com um suporte absoluto de 2;
- Regra 149: Se (PTEL for maior ou igual a 45 e EOL estiver entre 1350 e 1450 e CONS estiver entre 5450 e 5550) então FA = 500 ou FA = 600, com um suporte absoluto de 3;
- Regra 150: Se (PTEL for inferior a 45 e HID for inferior a 150 EOL estiver entre 2150 e 2250) então FA = 200 ou FA = 400, com um suporte absoluto de 2;
- Regra 155: Se (EOL estiver entre 350 e 450 e CONS estiver entre 5850 e 5950) então FA = 400 ou FA = 800, com um suporte absoluto de 2;
- Regra 157: Se (PTEL for inferior a 55 e HID for inferior a 50 e EOL estiver entre 1050 e 1150) então FA = 300 ou FA = 400, com um suporte absoluto de 4;

Seguidamente, a validação de todas as regras foi feita introduzindo a base de dados de aprendizagem redefinida (tabela 5.16) e a base de dados de teste de 2015 (tabela 5.9). Aplicando as regras de indução geradas anteriormente é possível concluir que todo este processo é capaz de classificar todos os dias de 2015. Isto permite que nenhum dos dias fique por classificar. Efetuando o estudo do erro entre os valores previstos e os valores reais, é obtido um erro de 71,58%.

5.5.14.2 – Algoritmo de Indução Entropia

No cenário 4 é aplicado o algoritmo de indução Entropia aos dados da tabela 5.16. A aplicação deste algoritmo permitiu construir um conjunto de 215 regras exatas e 14 regras aproximadas. Estas 229 regras contêm todo o conhecimento da tabela inicial.

Estas regras permitem classificar 22 classes distintas, cada uma representando um valor de produção. A título de exemplo foram seleccionadas cinco regras exatas, com um suporte absoluto de quatro e cinco objetos. Reescritas seguidamente:

- Regra 8: Se (PTEL for maior ou igual a 45 e HID for inferior a 50 e EOL for maior ou igual a 650 e CONS estiver entre 4650 e 4950) então FA = 300, com um suporte absoluto de 4;
- Regra 28: Se (PTEL for maior ou igual a 45 e HID for inferior a 150 e EOL for maior ou igual a 1150 e CONS estiver entre 4850 e 5350) então FA = 400, com um suporte absoluto de 4;
- Regra 34: Se (PTEL for inferior a 45 e HID for inferior a 150 e EOL for inferior a 1450 e CONS for inferior a 4650) então FA = 400, com um suporte absoluto de 5;
- Regra 35: Se (EOL estiver entre 1250 e 1350 e CONS for inferior a 5150) então FA = 400, com um suporte absoluto de 5;
- Regra 91: Se (PTEL for maior ou igual a 35 e HID for maior ou igual a 250 e EOL for maior ou igual a 2850) então FA = 700, com um suporte absoluto de 4.

Como exemplo foram também seleccionadas cinco regras aproximadas capazes de classificar sete classes distintas. Estas são apresentadas na sua forma reescrita.

- Regra 216: Se (PTEL for maior ou igual a 35 e HID for maior ou igual a 150 e EOL estiver entre 750 e 850 e CONS for inferior a 5550) então FA = 900 ou FA = 1800, com um suporte absoluto de 2;
- Regra 217: Se (PTEL for maior ou igual a 45 e EOL estiver entre 1350 e 1450 e CONS estiver entre 5450 e 5550) então FA = 500 ou FA = 600, com um suporte absoluto de 3;
- Regra 224: Se (EOL estiver entre 350 e 450 e CONS estiver entre 5850 e 5950) então FA = 400 ou FA = 800, com um suporte absoluto de 2;
- Regra 225: Se (PTEL for inferior a 55 e EOL estiver entre 550 e 650 e CONS for maior ou igual a 5750) então FA = 300 ou FA = 400 ou FA = 500, com um suporte absoluto de 3;
- Regra 226: Se (PTEL for inferior a 55 e HID for inferior a 50 e EOL estiver entre 1050 e 1150) então FA = 300 ou FA = 400, com um suporte absoluto de 4.

As regras geradas são aplicadas na previsão da produção das centrais a Fio-de-água, para o ano 2015. Utilizando como *input* para o programa a base de dados de aprendizagem (tabela 5.16), a base de dados de teste onde se encontram os dados reais de 2015 (tabela 5.9) e as regras induzidas, foi possível classificar todos as produções médias para os 365 dias do ano. Com base nestes valores previstos foi feito um estudo do erro. O estudo permitiu concluir que o erro foi de 74,10%.

5.5.15 – Fio-de-água – Conclusões

Os diferentes cenários aplicados para a previsão da produção média proveniente das centrais a Fio-de-água permitiram apurar qual a metodologia mais adequada. Nas quatro metodologias testadas foram aplicados dois códigos de equivalência distintos e dois algoritmos de indução de regras. Tendo em conta a análise efetuada, conclui-se que o cenário 4 apresenta os melhores resultados, segundo o erro calculado. Este cenário utiliza o código de equivalência do método 2, com o algoritmo de indução Laplace. Para além de apresentar o menor erro, este cenário possui também o menor número de regras, resultando num esforço computacional em comparação com os restantes cenários.

5.5.16 – Bombagem - Caso de estudo 1

O presente caso de estudo tem como finalidade determinar qual a previsão da produção das centrais hidroelétricas com recurso à bombagem, para o dia depois de amanhã. O atributo de decisão que representa estes valores previstos é definido pelo acrónimo BOMB. Os atributos de decisão mantêm-se inalterados, tal como acontece nos restantes casos de estudo. Estes são o PTEL, HID, EOL e CONS, em que HID representa o valor da produção das centrais mini-hídricas, EOL representa o valor da produção das centrais eólicas, CONS o valor do consumo de eletricidade e o PTEL representa o valor do índice PTEL.

No caso estudo 1 foi aplicado o código de equivalência do primeiro método. Esta etapa é realizada de acordo com os códigos, presentes nas tabelas 5.2 e 5.4 do subcapítulo 5.3, permitindo a redefinição dos valores de cada um dos atributos em análise. O atributo PTEL utiliza o mesmo código de equivalência (tabela 5.1) para os dois casos de estudo.

Aplicando o método de discretização aos atributos em estudo, é construída a tabela 5.17 com os dados necessários para o *input* no *software*.

Tabela 5.17 – Base de dados relativa à Bombagem, aplicando códigos de equivalência do método 1

Objetos	Atributos				
	PTEL	HID	EOL	CONS	<i>BOMB</i>
1	10	330	2920	4860	420
2	30	310	2920	5930	120
3	40	300	1970	6000	140
4	0	310	2930	5540	210
5	0	320	2700	5290	330
...
...
361	40	150	1720	5470	110
362	20	140	2990	5290	390
363	40	150	3060	6010	150
364	40	140	2520	6160	260
365	50	130	570	5700	160

5.5.16.1 – Algoritmo de Indução Laplace

Recorrendo à base de dados discretizada, tabela 5.17, foi testada a aplicação do algoritmo de indução Laplace. Este permitiu gerar um conjunto de regras em que, o conjunto final das regras exatas contém todo o conhecimento da tabela inicial. Este processo foi definido como cenário 1, referente ao atributo de decisão Bombagem.

A aplicação deste algoritmo de indução gerou 182 regras, aptas para classificar 39 classes diferentes. Cada uma dessas classes representa uma produção distinta.

A título de exemplo foram retiradas cinco regras com um suporte de 4 e 8 objetos. As mesmas são reescritas de uma forma mais compreensível.

- Regra 1: Se (PTEL for inferior a 35 e HID estiver entre 335 e 345 e EOL for inferior a 2105 e CONS estiver entre 5515 e 6515) então BOMB = 0, com um suporte absoluto de 8;

- Regra 13: Se (PTEL estiver entre 35 e 55 e HID for maior ou igual a 35 e EOL estiver entre 345 e 435 e CONS for superior a 5745) então BOMB = 10, com um suporte absoluto de 4;
- Regra 24: Se (HID for inferior a 35 e EOL estiver entre 605 e 770 e CONS estiver entre 4790 e 5565) então BOMB = 30, com um suporte absoluto de 4;
- Regra 33: Se (PTEL for inferior a 65 e EOL for inferior a 565 e CONS estiver entre 4475 e 4535) então BOMB = 40, com um suporte absoluto de 4;
- Regra 99: Se (PTEL for maior ou igual a 35 e HID estiver entre 295 e 335 e EOL for inferior a 1975 e CONS for maior ou igual a 5995) então BOMB = 140, com um suporte absoluto de 4.

Seguidamente é feita a validação das regras, recorrendo ao método de validação *Single Pass*. Para a aplicação deste método é introduzida a base de dados de aprendizagem redefinida (tabela 5.17), a base de dados de teste onde se encontram os atributos de condição referentes ao ano 2015 (tabela 5.9) e as respetivas regras.

Este método de validação permite prever a produção média dos 365 dias de 2015. Através dos valores previstos e dos reais foi determinado o erro, este foi de 69,10%.

5.5.16.2 – Algoritmo de Indução Entropia

Neste cenário 2 aplica-se o algoritmo de indução Entropia com base nos dados da tabela 5.17. Aplicando este algoritmo são geradas 192 regras exatas que contêm todo o conhecimento da tabela inicial.

Estas regras permitem classificar 39 classes distintas, sendo que cada conjunto representa um valor de produção diferente. A título de exemplo são apresentadas cinco regras, cujo suporte é feito com três, quatro e seis objetos.

- Regra 1: Se (HID estiver entre 335 e 345 e EOL estiver entre 1375 e 2105 e CONS for inferior a 6505) então BOMB = 0, com um suporte absoluto de 6;
- Regra 27: Se (HID for inferior a 35 e EOL estiver entre 605 e 770 e CONS estiver entre 4790 e 5525) então BOMB = 30, com um suporte absoluto de 4;

- Regra 32: Se (HID for maior ou igual a 55 e EOL estiver entre 510 e 1315 e CONS estiver entre 4830 e 4930) então BOMB = 40, com um suporte absoluto de 3;
- Regra 56: Se (HID estiver entre 225 e 315 e EOL for inferior a 1585 e CONS estiver entre 5115 e 5455) então BOMB = 80, com um suporte absoluto de 3;
- Regra 96: Se (HID estiver entre 235 e 335 e EOL estiver entre 2305 e 2505 e CONS for maior ou igual a 5865) então BOMB = 130, com um suporte absoluto de 3.

As regras geradas foram validadas através da sua aplicação nos dados de 2015. Como *input* foi introduzida a base de dados de aprendizagem (tabela 5.17) e a base de dados de teste (tabela 5.9), do qual fazem parte os atributos cuja previsão para o dia depois de amanhã (D+2) é conhecida.

Com este processo é possível prever a produção média para o dia depois de amanhã, dos 365 dias do ano 2015. Através destes valores previstos foi feito um estudo do erro, recorrendo aos valores previstos e aos valores reais. Com este estudo determinou-se um erro de 72,60%.

5.5.17 – Bombagem - Caso de estudo 2

Neste caso de estudo 2, analisando ainda as previsões para a produção das centrais hidroelétricas com recurso à bombagem, foi aplicado um novo processo de discretização. Este baseou-se nos códigos de equivalência do segundo método. O processo foi aplicado a todos os valores de cada um dos atributos, através dos códigos de equivalência presentes nas tabelas 5.5 e 5.7 do subcapítulo 5.3. Construiu-se assim a tabela 5.18 que apresenta os dados necessários para análise dos dois seguintes cenários em estudo.

Tabela 5.18 – Base de dados relativa à Bombagem, aplicando códigos de equivalência do método 2

Objetos	Atributos				
	PTEL	HID	EOL	CONS	BOMB
1	10	300	2900	4900	400
2	30	300	2900	5900	100
3	40	300	2000	6000	100
4	0	300	2900	5500	200
5	0	300	2700	5300	300
...
...
361	40	200	1700	5500	100
362	20	100	3000	5300	400
363	40	100	3100	6000	100
364	40	100	2500	6200	300
365	50	100	600	5700	200

5.5.17.1 – Algoritmo de Indução Laplace

Com base nos dados da tabela 5.18, aplica-se o algoritmo de indução Laplace do *software* ROSE2. Construiu-se assim um conjunto de 85 regras em que, o conjunto final das regras exatas e das regras aproximadas contém todo o conhecimento da tabela inicial.

Estas regras estão divididas em 80 exatas e 5 aproximadas, com capacidade de classificar 5 classes diferentes.

Como exemplo são apresentadas na sua forma reescrita, cinco das regras exatas, cujo suporte é feito com mais de sete objetos.

- Regra 1: Se (HID for maior ou igual a 50 e EOL estiver entre 350 e 450 e CONS estiver entre 4850 e 5750) então BOMB = 0, com um suporte absoluto de 8;
- Regra 27: Se (HID estiver entre 50 e 250 e EOL estiver entre 950 e 1050) então BOMB = 100, com um suporte absoluto de 8;
- Regra 30: Se (PTEL for inferior a 55 e HID for inferior a 50 e EOL estiver entre 950 e 1350 e CONS estiver entre 4850 e 5650) então BOMB = 100, com um suporte absoluto de 8;
- Regra 38: Se (PTEL for maior ou igual 35 e HID for maior ou igual a 50 e EOL estiver entre 1450 e 1650 e CONS estiver entre 4850 e 5650) então BOMB = 100, com um suporte absoluto de 9;

- Regra 56: Se (PTEL estiver entre 35 e 55 e EOL estiver entre 950 e 1150 e CONS estiver entre 5050 e 5750) então BOMB = 100, com um suporte absoluto de 11;

Seguidamente são descritas na sua forma reescrita, as cinco regras aproximadas.

- Regra 81: Se (HID for inferior a 50 e EOL for inferior a 850 e CONS estiver entre 4950 e 5050) então BOMB = 0 ou BOMB = 200, com um suporte absoluto de 2;
- Regra 82: Se (PTEL for maior ou igual a 45 e EOL estiver entre 1350 e 1450 e CONS estiver entre 5450 e 5550) então BOMB = 0 ou BOMB = 200, com um suporte absoluto de 3;
- Regra 83: Se (HID for maior ou igual a 50 e EOL for inferior a 350 e CONS for inferior a 4650) então BOMB = 0 ou BOMB = 100, com um suporte absoluto de 3;
- Regra 84: Se (PTEL for inferior a 55 e EOL estiver entre 550 e 650 e CONS for maior ou igual a 5750) então BOMB = 0 ou BOMB = 100, com um suporte absoluto de 3;
- Regra 85: Se (PTEL for maior ou igual a 45 e HID for maior ou igual a 50 e CONS for inferior a 4550) então BOMB = 100 ou BOMB = 300, com um suporte absoluto de 2;

Aplicando as regras de indução geradas anteriormente é possível obter os resultados para este cenário 3. A validação destas regras é feita introduzindo a base de dados de aprendizagem redefinida (tabela 5.18), a base de dados de teste de 2015 (tabela 5.9) e as regras geradas.

Este processo permite classificar os dias depois de amanhã com a sua produção média, permitindo assim que nenhum dos dias de 2015 fique por classificar. Efetuando o estudo do erro entre os valores previstos e os valores reais, é possível determinar que o erro foi de 75,29%.

5.5.17.2 – Algoritmo de Indução Entropia

De acordo com a base de dados da tabela 5.18, é aplicado o algoritmo de indução Entropia. Este permitiu construir um conjunto de 160 regras, essenciais para aplicar neste cenário 4.

Das 160 regras geradas, 150 são regras exatas e 10 são regras aproximadas. Estas têm capacidade para classificar 5 classes distintas, cada uma representando um valor de produção. A título de exemplo são apresentadas as cinco regras cujo suporte é feito com mais de oito objetos.

- Regra 1: Se (HID for maior ou igual a 50 e EOL estiver entre 350 e 450 e CONS estiver entre 4650 e 5750) então BOMB = 0, com um suporte absoluto de 9;
- Regra 39: Se (PTEL estiver entre 35 e 45 e HID for maior ou igual a 150 e EOL estiver entre 950 e 2850 e CONS estiver entre 5350 e 5850) então BOMB = 100, com um suporte absoluto de 9;
- Regra 42: Se (HID for inferior a 150 e EOL estiver entre 950 e 1250 e CONS estiver entre 4850 e 5450) então BOMB = 100, com um suporte absoluto de 10;
- Regra 46: Se (PTEL estiver entre 35 e 55 e EOL estiver entre 950 e 1150 e CONS estiver entre 5050 e 5750) então BOMB = 100, com um suporte absoluto de 11;
- Regra 108: Se (PTEL estiver entre 35 e 45 e EOL estiver entre 850 e 1450 e CONS for inferior a 5050) então BOMB = 200, com um suporte absoluto de 9.

Das dez regras aproximadas são apresentadas na sua forma reescrita cinco regras que classificam quatro classes diferentes.

- Regra 151: Se (HID estiver entre 50 e 150 e EOL estiver entre 650 e 750 e CONS estiver entre 5450 e 5650) então BOMB = 100 ou BOMB = 200, com um suporte absoluto de 2;
- Regra 153: Se (PTEL for maior ou igual a 35 e HID for maior ou igual a 150 e EOL estiver entre 750 e 850 e CONS for inferior a 5550) então BOMB = 100 ou BOMB = 200, com um suporte absoluto de 2;
- Regra 154: Se (PTEL for maior ou igual a 45 e HID for maior ou igual a 50 e EOL estiver entre 1350 e 1450 e CONS for inferior a 5550) então BOMB = 0 ou BOMB = 200, com um suporte absoluto de 3;
- Regra 156: Se (PTEL for inferior a 55 e EOL estiver entre 550 e 650 e CONS for maior ou igual a 5750) então BOMB = 0 ou BOMB = 100, com um suporte absoluto de 3;
- Regra 160: Se (PTEL for maior ou igual a 45 e HID for maior ou igual a 50 e CONS for inferior a 4550) então BOMB = 100 ou BOMB = 300, com um suporte absoluto de 2.

As regras geradas são aplicadas na previsão da produção para o dia depois de amanhã. Utilizando como *input* para o programa a base de dados de aprendizagem (tabela 5.18), a base

de dados de teste onde se encontram os dados reais de 2015 (tabela 5.9) e as regras induzidas. Estes dados permitiram classificar a produção das centrais a Gás Natural para o ano 2015.

Este cenário 4 permite prever a produção para todos os dias, permitindo assim que dois dias antes já se consiga saber a produção média das centrais em questão. Através destes valores foi feito um estudo do erro, recorrendo aos valores previstos e aos valores reais, culminando num erro de 81,37%.

5.5.18 – Bombagem – Conclusões

A abordagem apresentada neste estudo pretendeu analisar qual a metodologia mais adequada para prever a produção média das centrais de Bombagem, para o dia D+2. Tendo em conta os resultados apresentados, conclui-se que o cenário 1 permite obter um menor erro, em comparação com os restantes cenários. Este cenário utiliza o código de equivalência do método 1, com o algoritmo de indução Laplace. Apesar do menor erro, este possui um número elevado de regras, resultando num esforço computacional superior.

5.5.19 – Importação - Caso de estudo 1

A seguinte análise tem como finalidade o estudo da previsão da produção de energia importada por Portugal, para o dia depois de amanhã. O atributo a prever é então o atributo de decisão da produção importada (Importação), definido pelo acrónimo IMP.

Neste caso de estudo 1 foi aplicado o código de equivalência do primeiro método. Esta etapa de discretização é realizada de acordo com os códigos, presentes nas tabelas 5.2 e 5.3 do subcapítulo 5.3. O atributo PTEL utiliza o mesmo código de equivalência (tabela 5.1) para os dois casos de estudo em análise.

Aplicando estas redefinições, é construída a tabela 5.19 com os dados necessários para o *input* no *software*. Estes dados são apresentados de forma parcial na tabela 5.19.

Tabela 5.19 – Base de dados relativa à Importação, aplicando códigos de equivalência do método 1

Objetos	Atributos				
	PTEL	HID	EOL	CONS	IMP
1	10	330	2920	4860	60
2	30	310	2920	5930	0
3	40	300	1970	6000	0
4	0	310	2930	5540	20
5	0	320	2700	5290	20
...
...
361	40	150	1720	5470	670
362	20	140	2990	5290	360
363	40	150	3060	6010	180
364	40	140	2520	6160	90
365	50	130	570	5700	470

5.5.19.1 – Algoritmo de Indução Laplace

Recorrendo aos dados da tabela 5.19, foi inicialmente testado a aplicação do algoritmo de indução Laplace. Este permitiu gerar um conjunto de regras exatas que contêm todo o conhecimento para classificar os dados em análise neste cenário 1, relativo à produção importada.

A aplicação deste algoritmo de indução gerou 267 regras, aptas para classificar 124 classes diferentes.

A título de exemplo foram retiradas cinco regras, apresentadas reescritas de uma forma mais compreensível.

- **Regra 1:** Se (PTEL for inferior a 35 e HID for inferior a 345 e EOL estiver entre 2515 e 2925 e CONS estiver entre 5825 e 6625) então $IMP = 0$, com um suporte absoluto de 9;
- **Regra 2:** Se (PTEL for inferior a 5 e HID for maior ou igual a 325 e CONS for inferior a 6495) então $IMP = 0$, com um suporte absoluto de 4;
- **Regra 3:** Se (HID for maior ou igual a 325 e EOL estiver entre 2125 e 2625 e CONS estiver entre 6355 e 6665) então $IMP = 0$, com um suporte absoluto de 8;

- Regra 8: Se (PTEL for inferior a 45 e HID for maior ou igual a 295 e EOL for inferior a 1825 e CONS estiver entre 5615 e 5945) então $IMP = 0$, com um suporte absoluto de 6;
- Regra 22: Se (PTEL for inferior a 25 e HID for maior ou igual a 305 e EOL estiver entre 2785 e 3240 e CONS estiver entre 5265 e 6515) então $IMP = 20$, com um suporte absoluto de 5.

As regras apresentadas como exemplo são maioritariamente relativas à classe 0. Isto ocorre uma vez que na base de dados utilizada há 50 objetos pertencentes a esta classe, provocando um aumento no número de regras geradas para classificar esta classe.

A validação das regras foi feita recorrendo ao método de validação *Single Pass*, onde é introduzido a base de dados de aprendizagem redefinida (tabela 5.19), a base de dados de teste onde se encontram os atributos de condição referentes ao ano 2015 (tabela 5.9) e as respetivas regras.

Este método de validação permite classificar o dia depois de amanhã de todos os objetos de 2015 com a produção média importada. Observando os valores previstos e os reais foi determinado o erro, sendo este de 90,58%.

5.5.19.2 – Algoritmo de Indução Entropia

No cenário 2 é analisada a viabilidade da aplicação do algoritmo de indução Entropia. Com base nos dados da tabela 5.19 foram geradas um conjunto de 284 regras. Estas são capazes de classificar 124 objetos distintos.

A título de exemplo são apresentadas as cinco regras, cujo suporte é feito com mais de cinco objetos.

- Regra 1: Se (HID estiver entre 325 e 345 e EOL estiver entre 2415 e 2925 e CONS estiver entre 5155 e 6665) então $IMP = 0$, com um suporte absoluto de 8;
- Regra 2: Se (PTEL for inferior a 35 e HID for maior ou igual a 325 e EOL estiver entre 2035 e 2245 e CONS for inferior a 6625) então $IMP = 0$, com um suporte absoluto de 6;

- **Regra 3:** Se (PTEL for maior ou igual a 35 e HID for maior ou igual a 335 e EOL for maior ou igual a 1105) então $IMP = 0$, com um suporte absoluto de 8;
- **Regra 6:** Se (HID for maior ou igual a 325 e EOL estiver entre 1105 e 1385) então $IMP = 0$, com um suporte absoluto de 5;
- **Regra 7:** Se (HID for maior ou igual a 325 e EOL estiver entre 1105 e 1385) então $IMP = 0$, com um suporte absoluto de 5.

As regras geradas foram validadas utilizando a base de dados de teste de 2015 (tabela 5.9). Com este processo é possível classificar os 365 objetos de 2015 com a previsão da importação para o dia depois de amanhã. O erro, calculado através dos valores previstos e reais, permitiu determinar um erro de 84,03% para este cenário.

5.5.20 – Importação - Caso de estudo 2

Neste caso de estudo 2 é analisado o processo de discretização dado pelos códigos de equivalência do segundo método. Este processo tem como finalidade a previsão da produção média importada para o dia depois de amanhã. Os códigos de equivalência foram aplicados aos atributos de condição (HID, EOL e CONS) e ao atributo de decisão (BOMB), segundo as redefinições presentes nas tabelas 5.5 e 5.6 do subcapítulo 5.3. A discretização do atributo de condição PTEL é mantida, segundo a tabela 5.1. Com base neste processo foi criada a tabela 5.20 que apresenta os dados necessários para a análise dos dois seguintes cenários em estudo.

Tabela 5.20 – Base de dados relativa à Importação, aplicando códigos de equivalência do método 2

Objetos	Atributos				
	PTEL	HID	EOL	CONS	IMP
1	10	300	2900	4900	100
2	30	300	2900	5900	0
3	40	300	2000	6000	0
4	0	300	2900	5500	0
5	0	300	2700	5300	0
...
...
361	40	200	1700	5500	700
362	20	100	3000	5300	400
363	40	100	3100	6000	200
364	40	100	2500	6200	100
365	50	100	600	5700	500

5.5.20.1 – Algoritmo de Indução Laplace

No presente cenário 3 é aplicado o algoritmo de indução Laplace à base de dados da tabela 5.20. Construiu-se assim um conjunto de 114 regras constituído por 102 regras exatas e 12 aproximadas. Estas regras contêm todo o conhecimento da tabela inicial e têm capacidade para classificar 16 classes diferentes.

Como exemplo foram retiradas cinco regras exatas cujo suporte é feito com mais de oito objetos. Estas apresentam-se seguidamente na sua forma reescrita.

- Regra 1: Se (HID for maior ou igual a 250 e EOL for inferior a 2750 e CONS estiver entre 6450 e 6650) então $IMP = 0$, com um suporte absoluto de 17;
- Regra 2: Se (PTEL estiver entre 5 e 35 e EOL estiver entre 2450 e 2950 e CONS for maior ou igual a 5250) então $IMP = 0$, com um suporte absoluto de 14;
- Regra 4: Se (PTEL estiver entre 15 e 45 e HID for maior ou igual a 250 e EOL estiver entre 1350 e 2150 e CONS estiver entre 5450 e 6550) então $IMP = 0$, com um suporte absoluto de 13;
- Regra 7: Se (HID for maior ou igual a 250 e EOL estiver entre 2250 e 2750 e CONS for inferior a 6350) então $IMP = 0$, com um suporte absoluto de 9;
- Regra 9: Se (HID for maior ou igual a 250 e EOL estiver entre 2350 e 3250 e CONS for maior ou igual a 6050) então $IMP = 0$, com um suporte absoluto de 16.

A título de exemplo foram ainda seleccionadas das dozes regras, cinco regras aproximadas capazes de classificar 11 classes diferentes. Estas são descritas na sua forma reescrita, para que sejam de mais fácil compreensão.

- Regra 103: Se (PTEL for maior ou igual a 35 e HID for maior ou igual a 150 e EOL estiver entre 750 e 850 e CONS for inferior a 5550) então $IMP = 0$ ou $IMP = 400$, com um suporte absoluto de 2;
- Regra 104: Se (PTEL for maior ou igual a 45 e EOL estiver entre 1350 e 1450 e CONS estiver entre 5450 e 5550) então $IMP = 200$ ou $IMP = 500$, com um suporte absoluto de 3;
- Regra 109: Se (PTEL for inferior a 55 e EOL estiver entre 550 e 650 e CONS for maior ou igual a 5750) então $IMP = 1200$ ou $IMP = 1300$ ou $IMP = 1400$, com um suporte absoluto de 3;

- Regra 110: Se (HID for inferior a 50 e EOL estiver entre 1050 e 1150 e CONS for inferior a 5350) então $IMP = 600$ ou $IMP = 700$, com um suporte absoluto de 2;
- Regra 111: Se (HID for inferior a 50 e EOL estiver entre 1050 e 1150 e CONS for maior ou igual a 5550) então $IMP = 400$ ou $IMP = 1000$, com um suporte absoluto de 2.

A validação das regras foi realizada aplicando estas mesmas regras à base de dados de teste de 2015 (tabela 5.9). Com este processo foi possível classificar todos os dias de 2015 com os valores da produção média importada, simulando a previsão para o dia D+2. Efetuando o estudo do erro entre os valores previstos e os valores reais foi possível obter um erro de 96,40%.

5.5.20.2 – Algoritmo de Indução Entropia

De acordo com a base de dados da tabela 5.20, é aplicado o algoritmo de indução Entropia. Este permitiu construir um conjunto de 211 regras em que o conjunto final das 197 regras exatas e 13 regras aproximadas contém todo o conhecimento da tabela inicial. Esta análise gerou a informação essencial para o cenário 4, com capacidade para classificar 16 classes distintas.

A título de exemplo são apresentadas cinco regras cujo suporte é feito com mais de sete objetos.

- Regra 1: Se (HID for maior ou igual a 250 e EOL for inferior a 3250 e CONS estiver entre 6450 e 6650) então $IMP = 0$, com um suporte absoluto de 20;
- Regra 2: Se (PTEL estiver entre 25 e 55 e HID for maior ou igual a 250 e EOL for inferior a 2250 e CONS for maior ou igual a 6150) então $IMP = 0$, com um suporte absoluto de 14;
- Regra 4: Se (HID for maior ou igual a 250 e EOL for maior ou igual a 2850 e CONS estiver entre 5150 e 5950) então $IMP = 0$, com um suporte absoluto de 8;
- Regra 6: Se (PTEL estiver entre 15 e 25 e EOL for inferior a 2350 e CONS estiver entre 5150 e 6450) então $IMP = 0$, com um suporte absoluto de 8;
- Regra 7: Se (HID for maior ou igual a 250 e EOL estiver entre 2350 e 2750) então $IMP = 0$, com um suporte absoluto de 16.

Como exemplo são ainda apresentadas cinco regras aproximadas, capazes de classificar nove classes distintas. Estas são apresentadas na sua forma reescrita.

- Regra 198: Se (PTEL for maior ou igual a 35 e HID for maior ou igual a 150 e EOL estiver entre 750 e 850 e CONS for inferior a 5550) então $IMP = 0$ ou $IMP = 400$, com um suporte absoluto de 2;
- Regra 199: Se (PTEL for maior ou igual a 45 e EOL estiver entre 1350 e 1450 e CONS estiver entre 5450 e 5550) então $IMP = 200$ ou $IMP = 500$, com um suporte absoluto de 3;
- Regra 200: Se (PTEL for inferior a 45 e HID for inferior a 150 e EOL estiver entre 2150 e 2250) então $IMP = 300$ ou $IMP = 700$, com um suporte absoluto de 2;
- Regra 205: Se (PTEL for inferior a 55 e EOL estiver entre 550 e 650 e CONS for maior ou igual a 5750) então $IMP = 1200$ ou $IMP = 1300$ ou $IMP = 1400$, com um suporte absoluto de 3;
- Regra 211: Se (PTEL for maior ou igual a 55 e HID for maior ou igual a 50 e EOL for inferior a 350 e CONS estiver entre 5550 e 5650) então $IMP = 1000$ ou $IMP = 1300$, com um suporte absoluto de 2;

As regras geradas são aplicadas para a previsão da produção importada, para o dia depois de amanhã, do ano 2015. Utilizando como *input* a base de dados de aprendizagem (tabela 5.20), a base de dados de teste onde se encontram os dados reais de 2015 (tabela 5.9) e as regras induzidas.

Com este cenário 4 todos os objetos de 2015 foram classificados com a sua importação. Através dos valores previstos neste cenário foi feito um estudo do erro, obtendo-se assim um erro de 84,21%.

5.5.21 – Importação – Conclusões

Os estudos realizados tiveram como finalidade analisar qual a metodologia mais adequada para prever a produção média importada, para o dia D+2. Tendo em conta os resultados apresentados conclui-se que o cenário 2 apresenta um menor erro. É de salientar que comparativamente com o cenário 4 a diferença do erro é apenas de 0,18%.

Este cenário 2 utiliza o código de equivalência do método 1, com o algoritmo de indução Entropia. Apesar do baixo valor de erro, este cenário é o que apresenta a maior quantidade de regras, resultando num esforço computacional acrescido.

5.5.22 – Exportação - Caso de estudo 1

A análise realizada neste caso de estudo pretende prever a produção exportada, para o dia depois de amanhã. O primeiro passo deste caso de estudo foi a discretização dos valores de cada um dos atributos de condição e do atributo de decisão. Este atributo de decisão é definido pelo acrónimo EXP e representa a produção exportada (Exportação). A etapa de discretização é realizada de acordo com os códigos de equivalência do primeiro método, presentes nas tabelas 5.2 e 5.3 do subcapítulo 5.3. O atributo PTEL, representando o índice PTEL, utiliza o mesmo código de equivalência (tabela 5.1) para os dois casos de estudo. Aplicando as redefinições aos atributos em estudo é construída a tabela 5.21.

Tabela 5.21 – Base de dados relativa à Exportação, aplicando códigos de equivalência do método 1

Objetos	Atributos				
	PTEL	HID	EOL	CONS	EXP
1	10	330	2920	4860	660
2	30	310	2920	5930	1280
3	40	300	1970	6000	760
4	0	310	2930	5540	1210
5	0	320	2700	5290	1130
...
...
361	40	150	1720	5470	20
362	20	140	2990	5290	40
363	40	150	3060	6010	1210
364	40	140	2520	6160	930
365	50	130	570	5700	100

5.5.22.1 – Algoritmo de Indução Laplace

O primeiro algoritmo a ser aplicado foi o de indução Laplace. De acordo com este algoritmo foram geradas um conjunto de 245 regras exatas, contendo todo o conhecimento da tabela 5.21. Este foi definido como o cenário 1, permitindo classificar 115 classes diferentes.

Como exemplo foram retiradas cinco regras que apresentam um suporte absoluto de três objetos, escritas seguidamente na forma mais compreensível.

- Regra 1: Se (HID for inferior a 75 e EOL estiver entre 315 e 445 e CONS for inferior a 5850) então $EXP = 0$, com um suporte absoluto de 14;
- Regra 2: Se (HID for inferior a 55 e EOL for inferior a 945 e CONS for inferior a 5030) então $EXP = 0$, com um suporte absoluto de 18;
- Regra 6: Se (PTEL for inferior a 55 e HID for inferior a 45 e EOL for inferior 865 e CONS for maior ou igual a 5595) então $EXP = 0$, com um suporte absoluto de 9;
- Regra 7: Se (HID for inferior a 75 e EOL estiver entre 315 e 565 e CONS for inferior a 5425) então $EXP = 0$, com um suporte absoluto de 12;
- Regra 11: Se (HID for inferior a 35 e EOL estiver entre 835 e 1095 e CONS for inferior a 5480) então $EXP = 0$, com um suporte absoluto de 6.

A validação das regras é feita através do método *Single Pass*. Para a aplicação deste método é necessário introduzir a base de dados de aprendizagem redefinida (tabela 5.21), a base de dados de teste de 2015 (tabela 5.9) e as regras geradas. Estes dados de 2015 incluem os atributos cuja previsão para o dia depois de amanhã (D+2) é conhecida.

Através do processamento destes dados é possível classificar todos os dias de 2015, com uma antecedência de dois dias. Utilizando os valores previstos foi feito um estudo do erro entre os valores previstos e os valores reais, resultando num erro de 160,54%.

5.5.22.2 – Algoritmo de Indução Entropia

Na presente análise foi aplicado o algoritmo de indução Entropia à base de dados da tabela 5.21. Este permitiu gerar um conjunto de 266 regras exatas contendo todo o conhecimento da base de dados inicial.

A aplicação deste algoritmo à base de dados codificada pelo código de equivalência do método 1 foi definido como cenário 2. As regras geradas permitem classificar 115 conjuntos distintos, em que cada classe representa um valor diferente de produção exportada.

A título de exemplo foram selecionadas cinco regras com um suporte absoluto de três objetos. Estas foram reescritas para que sejam de mais fácil compreensão.

- Regra 1: Se (PTEL for inferior a 55 e HID for inferior a 75 e EOL for inferior a 785 e CONS for maior ou igual a 5525) então $EXP = 0$, com um suporte absoluto de 9;

- Regra 3: Se (HID for inferior a 55 e EOL for inferior a 985 e CONS for inferior a 5075) então $EXP = 0$, com um suporte absoluto de 20;
- Regra 4: Se (HID for inferior a 85 e EOL estiver entre 240 e 435 e CONS for maior ou igual a 5440) então $EXP = 0$, com um suporte absoluto de 6;
- Regra 7: Se (HID for inferior a 85 e EOL estiver entre 240 e 435 e CONS for maior ou igual a 5440) então $EXP = 0$, com um suporte absoluto de 13;
- Regra 17: Se (HID estiver entre 35 e 75 e EOL for inferior a 795 e CONS for maior ou igual a 5755) então $EXP = 0$, com um suporte absoluto de 7.

As regras foram validadas recorrendo ao método de validação *Single Pass*. Para isso, foram testadas aplicando-se aos atributos cuja previsão para o dia depois de amanhã é conhecida. Através deste processo todos os 365 dias foram previstos. Foi feito ainda um estudo do erro, entre os valores previstos e reais, que gerou um erro de 154,95%.

5.5.23 – Exportação - Caso de estudo 2

Neste caso de estudo 2 são aplicados os códigos de equivalência do segundo método aos atributos de condição (HID, EOL e CONS) e ao atributo de decisão (EXP). Esta discretização é aplicada recorrendo aos códigos de equivalência presentes nas tabelas 5.5 e 5.6 do subcapítulo 5.3. Este processo permite criar uma nova tabela, tabela 5.22, que apresenta os dados necessários de *input* para a análise a realizar nos dois seguintes cenários em estudo.

Tabela 5.22 – Base de dados relativa à Exportação, aplicando códigos de equivalência do método 2

Objetos	Atributos				
	PTEL	HID	EOL	CONS	EXP
1	10	300	2900	4900	700
2	30	300	2900	5900	1300
3	40	300	2000	6000	800
4	0	300	2900	5500	1200
5	0	300	2700	5300	1100
...
...
361	40	200	1700	5500	0
362	20	100	3000	5300	0
363	40	100	3100	6000	1200
364	40	100	2500	6200	900
365	50	100	600	5700	100

5.5.23.1 – Algoritmo de Indução Laplace

Neste cenário 3, aos dados da tabela 5.22 é aplicado ao algoritmo de indução Laplace. Construiu-se assim um conjunto de regras em que, o conjunto final das regras exatas e das regras aproximadas contém todo o conhecimento da tabela inicial.

Foram geradas 125 regras, divididas em 121 regras exatas e 4 aproximadas. As regras têm capacidade para classificar 20 classes diferentes.

Como exemplo são apresentadas cinco regras exatas cujo suporte é feito com quatro objetos.

- Regra 2: Se (HID for inferior a 150 e EOL estiver entre 450 e 1050 e CONS estiver entre 4550 e 5150) então $EXP = 0$, com um suporte absoluto de 18;
- Regra 3: Se (HID for inferior a 150 e EOL estiver entre 250 e 550 e CONS estiver entre 4750 e 5850) então $EXP = 0$, com um suporte absoluto de 25;
- Regra 7: Se (PTEL for maior ou igual a 45 e HID estiver entre 50 e 150 e EOL estiver entre 450 e 1050 e CONS for inferior a 5550) então $EXP = 0$, com um suporte absoluto de 15;
- Regra 14: Se (PTEL for maior ou igual a 35 e HID for inferior a 250 e EOL estiver entre 450 e 650 e CONS for inferior a 5650) então $EXP = 0$, com um suporte absoluto de 16;
- Regra 20: Se (HID estiver entre 50 e 150 e EOL for inferior a 1150 e CONS estiver entre 4850 e 5550) então $EXP = 0$, com um suporte absoluto de 22.

São ainda apresentadas as 4 regras aproximadas que foram geradas, descritas na sua forma reescrita.

- Regra 122: Se (PTEL for maior ou igual a 35 e HID for maior ou igual a 150 e EOL estiver entre 750 e 850 e CONS for inferior a 5550) então $EXP = 100$ ou $EXP = 800$, com um suporte absoluto de 2;
- Regra 123: Se (PTEL for maior ou igual a 45 e EOL estiver entre 1350 e 1450 e CONS estiver entre 5450 e 5550) então $EXP = 100$ ou $EXP = 200$, com um suporte absoluto de 3;
- Regra 124: Se (PTEL for maior ou igual a 55 e HID for inferior a 50 e EOL estiver entre 350 e 450 e CONS for maior ou igual a 5850) então $EXP = 0$ ou $EXP = 100$, com um suporte absoluto de 2;

- Regra 125: Se (PTEL for maior ou igual a 45 e EOL estiver entre 650 e 850 e CONS estiver entre 5150 e 5250) então $EXP = 0$ ou $EXP = 100$, com um suporte absoluto de 2.

Estas regras foram validadas recorrendo ao programa ROSE2 no qual foi introduzido a base de dados de aprendizagem redefinida (tabela 5.22), a base de dados de teste de 2015 (tabela 5.9) e as regras geradas.

Esta validação permitiu concluir que este cenário é capaz de classificar todos os dias de 2015 com a antecedência necessária. Efetuando o estudo do erro entre os valores previstos e os reais é possível determinar que o erro foi de 138,78%.

5.5.23.2 – Algoritmo de Indução Entropia

Neste cenário 4 o algoritmo de indução Entropia é aplicado à base de dados da tabela 5.22. A aplicação deste algoritmo gerou um conjunto de 198 regras, 193 regras exatas e 5 aproximadas. Este cenário permite classificar 20 classes, representando vinte valores distintos de produção média exportada.

A título de exemplo foram selecionadas cinco regras, apresentadas seguidamente na sua forma simplificada.

- Regra 4: Se (HID for inferior a 50 e EOL for maior ou igual a 650 e CONS estiver entre 4550 e 4750) então $ALB = 200$, com um suporte absoluto de 4;
- Regra 14: Se (HID for inferior a 50 e EOL for maior ou igual a 1350 e CONS estiver entre 5350 e 5550) então $ALB = 300$, com um suporte absoluto de 4;
- Regra 22: Se (PTEL for maior ou igual a 35 e HID for maior ou igual a 50 e EOL estiver entre 1150 e 1450 e CONS for inferior a 5350) então $ALB = 300$, com um suporte absoluto de 4;
- Regra 60: Se (PTEL for maior ou igual a 45 e HID for inferior a 50 e EOL estiver entre 650 e 750) então $ALB = 500$, com um suporte absoluto de 4;
- Regra 156: Se (PTEL estiver entre 35 e 55 e HID for maior ou igual a 150 e EOL estiver entre 250 e 1050 e CONS estiver entre 5050 e 5550) então $ALB = 1000$, com um suporte absoluto de 4.

As cinco regras aproximadas são apresentadas seguidamente na sua forma reescrita.

- Regra 194: Se (PTEL for maior ou igual a 35 e HID for maior ou igual a 150 e EOL estiver entre 750 e 850 e CONS for inferior a 5550) então $EXP = 100$ ou $EXP = 800$, com um suporte absoluto de 2;
- Regra 195: Se (PTEL for maior ou igual a 45 e EOL estiver entre 1350 e 1450 e CONS estiver entre 5450 e 5550) então $EXP = 100$ ou $EXP = 200$, com um suporte absoluto de 3;
- Regra 196: Se (PTEL for maior ou igual a 55 e HID for inferior a 150 e EOL estiver entre 650 e 750 e CONS estiver entre 5250 e 5650) então $EXP = 0$ ou $EXP = 100$, com um suporte absoluto de 2;
- Regra 197: Se (EOL estiver entre 750 e 850 e CONS estiver entre 5150 e 5250) então $EXP = 0$ ou $EXP = 100$, com um suporte absoluto de 2;
- Regra 198: Se (EOL estiver entre 350 e 450 e CONS estiver entre 5850 e 5950) então $EXP = 0$ ou $EXP = 100$, com um suporte absoluto de 2;

As regras geradas são aplicadas na previsão da produção média exportada para os dias depois de amanhã, do ano 2015. Este cenário permitiu prever com antecedência de dois dias, todos os valores para os dias de 2015. Efetuando o estudo do erro obteve-se um erro de 138,78%.

5.5.24 – Exportação – Conclusões

O presente estudo teve como finalidade analisar qual a metodologia mais adequada para prever a produção média a exportar no dia D+2. Tendo em conta a análise efetuada nos quatro cenários, conclui-se que o cenário 3 apresenta os melhores resultados. Esta conclusão é feita comparando o erro calculado nos vários cenários. Este cenário utiliza o código de equivalência do método 1, com o algoritmo de indução Laplace. Este não só apresenta o menor erro como também o menor número de regras, traduzindo-se num esforço computacional menor, em comparação com os restantes cenários.

5.5.25 – Core e Reduto

O *Core* e *Reduto*, definidos no subcapítulo 3.6, foram calculados para todas as bases de dados discretizadas, utilizando os dois códigos de equivalência distintos. O cálculo pode ser realizado

passo a passo eliminando cada atributo e verificando se a tabela de decisão continua consistente, no entanto o programa ROSE2 tem capacidade para efetuar este mesmo cálculo.

Determinando o *Core* e o Reduto para as bases de dados com o código de equivalência do método 1 foi possível concluir que os atributos de condição do *Core* e do Reduto eram os mesmos, significando que a redução de atributos estava no seu limite. Os atributos de condição que faziam parte do Reduto eram a HID, EOL e CONS, tirando para a previsão do Gás Natural, em que o Reduto era constituído apenas pelos atributos EOL e CONS. Os elementos que constituíam o *Core* eram os mesmos atributos pertencentes ao Reduto.

Para testar a fiabilidade do Reduto gerado, foi aplicado o algoritmo de indução Laplace ao Reduto de cada uma das bases de dados, ao qual já tinha sido aplicado o método 1 de discretização. Comparando o valor do erro entre as previsões utilizando a totalidade dos atributos de condição e utilizando apenas os atributos pertencentes ao Reduto, foi possível concluir que de um modo geral o erro era superior utilizando apenas o Reduto. Esta não demonstrou ser uma metodologia viável para aplicar na previsão destes dados.

Aplicando o mesmo processo às bases de dados discretizadas com o código de equivalência do método 2 foi possível concluir que para além do *Core* e Reduto serem os mesmos para os diferentes casos de estudo, todos os atributos de decisão são considerados o *Core* e o Reduto. Isto indica que todos os atributos são de elevada importância e nenhum deve ser removido de modo a manter a capacidade de classificação.

5.7 – Conclusão

Neste capítulo estudou-se e analisou-se a aplicação de vários cenários de modo a prever as produções das várias centrais, importações e exportações, para o dia depois de amanhã. Os casos de estudo expostos anteriormente, num total de 35 simulações, permitiram concluir que a metodologia aplicada tem capacidade para classificar os diferentes atributos de decisão para o dia $D + 2$ do ano 2015, recorrendo à base de dados de aprendizagem de 2014.

A análise extensa efetuada neste capítulo permitiu descrever de modo detalhado todo o processo inerente à aplicação da teoria dos conjuntos aproximativos, desde o tratamento de dados ao processo de discretização e à análise dos resultados provenientes do *software* ROSE2.

Ainda que os erros obtidos para os vários cenários não permitam determinar objetivamente qual o cenário ideal para aplicar de forma genérica em todas as previsões, analisando os erros e a quantidade de regras, é possível concluir que a utilização do código de equivalência do método 2 permite, de um modo geral, obter erros inferiores. A mesma análise permite também concluir que o método de indução de regras Laplace tem capacidade para classificar os objetos com um erro menor e utilizando o número de regras bastante inferior, comparativamente com o método de indução de regras Entropia.

Capítulo 6 – Conclusões Gerais e Sugestões para Trabalhos Futuros

6.1 – Conclusões Gerais

Nas últimas décadas o aumento significativo do consumo de energia elétrica e a importância da mesma no desenvolvimento económico e social dos países tem criado uma maior pressão para que esta energia elétrica chegue ao consumidor com a devida segurança e qualidade. De modo a garantir esses parâmetros de segurança e qualidade, diferentes procedimentos são efetuados.

No presente trabalho salienta-se a importância de um desses procedimentos, o *Day Ahead Congestion Forecast*. No seguimento dos objetivos deste procedimento, há a necessidade de efetuar o mesmo procedimento para o dia depois de amanhã, dia D+2. Este está no entanto condicionado pela escassez de informação prevista relativa à produção das várias centrais, para esse horizonte temporal.

Este projeto foi realizado com o intuito de desenvolver uma metodologia capaz de analisar a informação contida em bases de dados, de modo a que possa ser extraída e aplicada de uma maneira simples e objetiva no problema em questão. Foram assim aplicadas técnicas de *Data Mining* para gerar um conjunto de regras capaz de classificar novos objetos. Estas técnicas, com base na Teoria dos Conjuntos Aproximativos, revelaram-se capazes de classificar a produção média, para o dia D+2, dos vários tipos de centrais em análise e da Importação e Exportação. Destaca-se o sucesso da capacidade de previsão da produção média das centrais a Carvão.

Em síntese, com a realização deste projeto foi possível desenvolver uma metodologia capaz de fazer uma primeira abordagem ao problema do estudo da previsão das produções de energia para o dia D+2. Este apresenta-se como ponto de partida para novas abordagens ao problema em questão, servindo estes estudos de referência para novos trabalhos desenvolvidos nesta área.

Através do trabalho concretizado foi ainda possível a realização e apresentação de um artigo, intitulado “Coal Generation Profile for the Day 2 Ahead Congestion Forecast”. Este foi publicado no International Congress on Engineering - ICEUBI 2017, em colaboração com os meus orientadores Cristina Faustino Agreira e Rui Nóbrega Pestana.

6.2 – Sugestões Para Trabalhos Futuros

No presente projeto foi desenvolvida uma metodologia capaz de prever a produção de vários tipos de centrais, para o dia D+2. Este gerou um aglomerado de ideias, novas perspectivas de investigação e de melhoria da metodologia aplicada. Desta forma, nos pontos seguintes são apresentadas algumas sugestões para trabalhos futuros, que, decerto, serão um contributo valioso para uma melhor previsão do perfil de geração das centrais analisadas, para o dia D+2:

- A teoria aplicada recorre a bases de dados de aprendizagem para gerar as regras e posteriormente classificar novos objetos. Sugere-se para trabalho futuro, o aumento da base de dados de aprendizagem, incluindo assim dados de anos mais recentes, como por exemplo de 2015 e 2016, para a criação de uma base de dados de aprendizagem mais robusta e capaz de fazer face às situações controversas de produção de energia resultantes dos anos atípicos que têm ocorrido em termos de produção de energia;
- A utilização de mais atributos de condição de modo a tornar a regras geradas mais robustas. Atributos estes que possam aumentar a fiabilidade das regras para o horizonte temporal em análise. Estes podem ser dados relativos às previsões meteorológicas ou outros que tenham um impacto direto nas diferentes centrais e que seja ao mesmo tempo possível a sua previsão para o horizonte temporal em análise;
- No seguimento do trabalho realizado neste projeto é ainda proposto como trabalho futuro a previsão da produção, detalhada, para as 24 horas do dia D+2.

Com base neste projeto será ainda realizado um artigo para submeter ao 53rd International Universities Power Engineering Conference – UPEC 2018.

Referências Bibliográficas

- Agreira, C. I. F. (2010). *Técnicas de Data Mining Para Estudo e Análise da Segurança dos Sistemas Eléctricos de Energia*. Tese de Doutoramento em Engenharia Electrotécnica e de Computadores. Faculdade de Engenharia da Universidade do Porto. Porto.
- Ali, R., Siddiqi, M., Lee, S. (2015). *Rough Set-based Approaches for Discretization: A Compact Review*. Artif Intell Rev. Springer. DOI 10.1007/s10462-014-9426-2, pp. 235-263.
- Berry M. e Linoff G. (1997). *Data Mining Techniques, For Marketing, Sales, and Customer Relationship Management*. 2º Edição, Wiley Publishing, Inc. Indiana.
- Cios, K. J., Pedrycz, W., Swiniarski, R. W. (1998). *Data Mining Methods for Knowledge Discovery*. Springer Science+Business Media, LLC. New York, USA.
- ENTSO-E (2004). *Continental Europe Operation Handbook*. European Network of Transmission System Operators for Electricity. Acedido em: 23 de setembro de 2016, no Website do: ENTSO-E: www.entsoe.eu/publications/system-operations-reports/
- Fayyad, U., Piatetsky-Shapiro, G., Smyth, P. (1996). *The KDD Process for Extracting Useful Knowledge from Volumes of Data*. Communications of the ACM, Vol. 39, No. 11, pp. 29.
- Goebel, M. e Gruenwald, L. (1999). *A Survey of Data Mining and Knowledge Discovery Software Tools*. ACM SIGKDD Explorations, Vol. 1, No. 1.
- Grzymala-Busse, J. Stefanowski, J. (2001). *Three Discretization Methods for Rule Induction*. International Journal of Intelligent Systems, Vol. 16, pp. 29-38.
- Inmon, W. H. (2002). *Building the Data Warehouse*. 3º Edição, John Wiley & Sons, Inc.
- Lambert-Torres, G., Quintana, V. C., Alves da Silva, A. P., Borges da Silva, L. E. (1996a). *Classification of Power System Operation Point Using Rough Set Techniques*. IEEE International Conference on Systems, Man and Cybernetics, Vol. 3, pp. 1898-1903.

- Lambert-Torres, G., Quintana, V. C., Alves da Silva, A. P., Borges da Silva, L. E. (1996b). *Knowledge-Base Reduction Based on Rough Set Techniques*. Canadian Conference on Electrical and Computer Engineering Vol. 1, pp. 278-281.
- Lambert-Torres, G. (2002). *Application of Rough Sets in Power System Control Center Data Mining*. IEEE Power Engineering Society Winter Meeting, Vol. 1, pp. 627-631.
- MIBEL (2016). *Operadores de Mercado*. Acedido em: 5 de junho de 2016, no Website do: Mercado Ibérico de Electricidade:
www.mibel.com/index.php?mod=pags&mem=detalle&relmenu=40&relcategoria=101&id=28.
- OMIP (2016a). *MIBEL*. Acedido em: 5 de junho de 2016, no Website do: Operador do Mercado Ibérico de Energia, Polo Português:
www.omip.pt/OMIP/MIBEL/tabid/72.aspx.
- OMIP (2016b). *Regras de Determinação de Índices de Electricidade*. Acedido em: 14 de junho de 2016, no Website do: Operador do Mercado Ibérico de Energia, Polo Português:
www.omip.pt/LinkClick.aspx?fileticket=4DVsRn-tLFk%3d.
- OMIP (2016c). *Índices de Electricidade do MIBEL – Regras*. Acedido em: 14 de junho de 2016, no Website do: Operador do Mercado Ibérico de Energia, Polo Português:
www.omip.eu/LinkClick.aspx?fileticket=oMaa_TqFcSY%3D.
- Panciatici, P. e Vrsecky, J. (2005). *Day Ahead Congestion Forecast for a Secure Operation of the European Transmission System*. 15th Power Systems Computation Conference (PSCC), Liège, Belgium.
- Pasten, C. e Santamarina, J. C. (2012). *Energy and Quality of Life*. Brown, S. P. A., Jefferson, M., *Energy Policy*, Elsevier. Georgia, USA.
- Pawlak, Z. (1982). *Rough Sets*. International Journal of Computer and Information Sciences, pp. 341-356.
- Pawlak, Z. (1991). *Rough Sets – Theoretical Aspects of Reasoning about Data*. Kluwer Academic Publishers.

- Pawlak, Z. (1995). *Rough Sets Approach to Knowledge-based Decision Support*. 14th European Conference on Operational Research, Israel.
- Pawlak, Z. (2002). *Rough Set theory and its applications*. Journal of Telecommunications and Information Technology Vol. 3, pp. 7-10.
- Pawlak, Z. e Skowron, A. (2006). *Rudiments of rough sets*. Information Sciences, Elsevier Inc.
- Pisicâ, I. e Postolache, P. (2010). *Rough Set Theory and Its Applications in Electrical Power Engineering. A Survey*. U.P.B. Sci. Bull, Series C, Vol. 72, 1^o Edição, pp. 211-222.
- Predki, B., Slowinski, R. Stefanowski, J., Susmaga, R. e Wilk Sz. (1998). *ROSE – Software Implementation of the Rough Set Theory In: L. Polkowski, A. Skowron*. Rough Sets and Current Trends in Computing, Vol. 1424, pp. 605-608.
- Predki, B. e Wilk Sz. (1999). *Rough Set Based Data Exploration Using ROSE System In: Z. W. Ras, A. Skowron, eds*. Foundations of Intelligent Systems, Lecture Notes in Artificial Intelligence, Vol. 1609, pp. 172-180.
- REN (2016). *Caracterização da Rede Nacional de Transporte para Efeitos de Acesso à Rede*. Acedido em: 13 de dezembro de 2016, no Website da: REN - Redes Energéticas Nacionais: www.centrodeinformacao.ren.pt/PT/publicacoes.
- Rissino, S. e Lambert-Torres, G. (2009). *Rough Set Theory – Fundamental Concepts, Principals, Data Extraction, and Applications*. Data Mining and Knowledge Discovery in Real Life Applications, pp. 35-58.
- Santos, F. S. e Azevedo, C. (2005). *Data Mining Descoberta de conhecimento em bases de dados*. FCA - Editora de Informática.
- Sassi, R. J. (2006). *Uma Arquitetura Híbrida para Descoberta de Conhecimento em Bases de Dados: Teoria dos Rough Sets e Redes Neurais Artificiais Mapas Auto-Organizáveis*. Tese de Doutorado em Engenharia Elétrica. Escola Politécnica da Universidade de São Paulo. Brasil.

Anexo I – Tabelas dos Códigos de Equivalência

Tabelas utilizadas para a codificação/discretização dos objetos em estudo

Tabela AI.1 – Código de equivalência do método 1 – Mini-hídrica, Eólica e Consumo

Mini-hídrica		Eólica		Consumo	
14.95 ≤ P < 24.95	20	134.95 ≤ P < 144.95	140	4044.95 ≤ P < 4054.95	4050
24.95 ≤ P < 34.95	30	164.95 ≤ P < 174.95	170	4454.95 ≤ P < 4464.95	4460
34.95 ≤ P < 44.95	40	194.95 ≤ P < 204.95	200	4464.95 ≤ P < 4474.95	4470
44.95 ≤ P < 54.95	50	204.95 ≤ P < 214.95	210	4474.95 ≤ P < 4484.95	4480
54.95 ≤ P < 64.95	60	224.95 ≤ P < 234.95	230	4494.95 ≤ P < 4504.95	4500
64.95 ≤ P < 74.95	70	244.95 ≤ P < 254.95	250	4504.95 ≤ P < 4514.95	4510
74.95 ≤ P < 84.95	80	254.95 ≤ P < 264.95	260	4514.95 ≤ P < 4524.95	4520
84.95 ≤ P < 94.95	90	274.95 ≤ P < 284.95	280	4524.95 ≤ P < 4534.95	4530
94.95 ≤ P < 104.95	100	284.95 ≤ P < 294.95	290	4534.95 ≤ P < 4544.95	4540
104.95 ≤ P < 114.95	110	294.95 ≤ P < 304.95	300	4544.95 ≤ P < 4554.95	4550
114.95 ≤ P < 124.95	120	304.95 ≤ P < 314.95	310	4554.95 ≤ P < 4564.95	4560
124.95 ≤ P < 134.95	130	314.95 ≤ P < 324.95	320	4564.95 ≤ P < 4574.95	4570
134.95 ≤ P < 144.95	140	324.95 ≤ P < 334.95	330	4574.95 ≤ P < 4584.95	4580
144.95 ≤ P < 154.95	150	334.95 ≤ P < 344.95	340	4584.95 ≤ P < 4594.95	4590
154.95 ≤ P < 164.95	160	344.95 ≤ P < 354.95	350	4594.95 ≤ P < 4604.95	4600
164.95 ≤ P < 174.95	170	354.95 ≤ P < 364.95	360	4614.95 ≤ P < 4624.95	4620
174.95 ≤ P < 184.95	180	364.95 ≤ P < 374.95	370	4634.95 ≤ P < 4644.95	4640
184.95 ≤ P < 194.95	190	374.95 ≤ P < 384.95	380	4654.95 ≤ P < 4664.95	4660
194.95 ≤ P < 204.95	200	384.95 ≤ P < 394.95	390	4664.95 ≤ P < 4674.95	4670
204.95 ≤ P < 214.95	210	394.95 ≤ P < 404.95	400	4674.95 ≤ P < 4684.95	4680
214.95 ≤ P < 224.95	220	404.95 ≤ P < 414.95	410	4684.95 ≤ P < 4694.95	4690
224.95 ≤ P < 234.95	230	424.95 ≤ P < 434.95	430	4694.95 ≤ P < 4704.95	4700
234.95 ≤ P < 244.95	240	434.95 ≤ P < 444.95	440	4714.95 ≤ P < 4724.95	4720
244.95 ≤ P < 254.95	250	444.95 ≤ P < 454.95	450	4724.95 ≤ P < 4734.95	4730
254.95 ≤ P < 264.95	260	464.95 ≤ P < 474.95	470	4734.95 ≤ P < 4744.95	4740
264.95 ≤ P < 274.95	270	474.95 ≤ P < 484.95	480	4744.95 ≤ P < 4754.95	4750
274.95 ≤ P < 284.95	280	484.95 ≤ P < 494.95	490	4764.95 ≤ P < 4774.95	4770
284.95 ≤ P < 294.95	290	494.95 ≤ P < 504.95	500	4774.95 ≤ P < 4784.95	4780
294.95 ≤ P < 304.95	300	514.95 ≤ P < 524.95	520	4794.95 ≤ P < 4804.95	4800
304.95 ≤ P < 314.95	310	534.95 ≤ P < 544.95	540	4804.95 ≤ P < 4814.95	4810
314.95 ≤ P < 324.95	320	544.95 ≤ P < 554.95	550	4814.95 ≤ P < 4824.95	4820
324.95 ≤ P < 334.95	330	554.95 ≤ P < 564.95	560	4834.95 ≤ P < 4844.95	4840
334.95 ≤ P < 344.95	340	564.95 ≤ P < 574.95	570	4844.95 ≤ P < 4854.95	4850
344.95 ≤ P < 354.95	350	574.95 ≤ P < 584.95	580	4854.95 ≤ P < 4864.95	4860
354.95 ≤ P < 364.95	360	584.95 ≤ P < 594.95	590	4874.95 ≤ P < 4884.95	4880
-----	----	594.95 ≤ P < 604.95	600	4884.95 ≤ P < 4894.95	4890
-----	----	604.95 ≤ P < 614.95	610	4894.95 ≤ P < 4904.95	4900
-----	----	614.95 ≤ P < 624.95	620	4914.95 ≤ P < 4924.95	4920
-----	----	624.95 ≤ P < 634.95	630	4934.95 ≤ P < 4944.95	4940
-----	----	634.95 ≤ P < 644.95	640	4944.95 ≤ P < 4954.95	4950
-----	----	644.95 ≤ P < 654.95	650	4954.95 ≤ P < 4964.95	4960
-----	----	654.95 ≤ P < 664.95	660	4974.95 ≤ P < 4984.95	4980

Tabela AI.1 – Código de equivalência do método 1 – Mini-hídrica, Eólica e Consumo (Cont).

Mini-hídrica		Eólica		Consumo	
-----	----	664.95 ≤ P < 674.95	670	5004.95 ≤ P < 5014.95	5010
-----	----	674.95 ≤ P < 684.95	680	5014.95 ≤ P < 5024.95	5020
-----	----	684.95 ≤ P < 694.95	690	5034.95 ≤ P < 5044.95	5040
-----	----	694.95 ≤ P < 704.95	700	5044.95 ≤ P < 5054.95	5050
-----	----	704.95 ≤ P < 714.95	710	5064.95 ≤ P < 5074.95	5070
-----	----	714.95 ≤ P < 724.95	720	5074.95 ≤ P < 5084.95	5080
-----	----	724.95 ≤ P < 734.95	730	5084.95 ≤ P < 5094.95	5090
-----	----	734.95 ≤ P < 744.95	740	5104.95 ≤ P < 5114.95	5110
-----	----	744.95 ≤ P < 754.95	750	5114.95 ≤ P < 5124.95	5120
-----	----	754.95 ≤ P < 764.95	760	5124.95 ≤ P < 5134.95	5130
-----	----	774.95 ≤ P < 784.95	780	5134.95 ≤ P < 5144.95	5140
-----	----	784.95 ≤ P < 794.95	790	5164.95 ≤ P < 5174.95	5170
-----	----	794.95 ≤ P < 804.95	800	5184.95 ≤ P < 5194.95	5190
-----	----	804.95 ≤ P < 814.95	810	5194.95 ≤ P < 5204.95	5200
-----	----	814.95 ≤ P < 824.95	820	5204.95 ≤ P < 5214.95	5210
-----	----	824.95 ≤ P < 834.95	830	5214.95 ≤ P < 5224.95	5220
-----	----	834.95 ≤ P < 844.95	840	5224.95 ≤ P < 5234.95	5230
-----	----	844.95 ≤ P < 854.95	850	5244.95 ≤ P < 5254.95	5250
-----	----	854.95 ≤ P < 864.95	860	5254.95 ≤ P < 5264.95	5260
-----	----	864.95 ≤ P < 874.95	870	5264.95 ≤ P < 5274.95	5270
-----	----	874.95 ≤ P < 884.95	880	5274.95 ≤ P < 5284.95	5280
-----	----	894.95 ≤ P < 904.95	900	5284.95 ≤ P < 5294.95	5290
-----	----	904.95 ≤ P < 914.95	910	5304.95 ≤ P < 5314.95	5310
-----	----	924.95 ≤ P < 934.95	930	5324.95 ≤ P < 5334.95	5330
-----	----	934.95 ≤ P < 944.95	940	5334.95 ≤ P < 5344.95	5340
-----	----	944.95 ≤ P < 954.95	950	5344.95 ≤ P < 5354.95	5350
-----	----	954.95 ≤ P < 964.95	960	5364.95 ≤ P < 5374.95	5370
-----	----	964.95 ≤ P < 974.95	970	5374.95 ≤ P < 5384.95	5380
-----	----	974.95 ≤ P < 984.95	980	5384.95 ≤ P < 5394.95	5390
-----	----	984.95 ≤ P < 994.95	990	5394.95 ≤ P < 5404.95	5400
-----	----	994.95 ≤ P < 1004.95	1000	5414.95 ≤ P < 5424.95	5420
-----	----	1014.95 ≤ P < 1024.95	1020	5424.95 ≤ P < 5434.95	5430
-----	----	1024.95 ≤ P < 1034.95	1030	5444.95 ≤ P < 5454.95	5450
-----	----	1044.95 ≤ P < 1054.95	1050	5454.95 ≤ P < 5464.95	5460
-----	----	1054.95 ≤ P < 1064.95	1060	5464.95 ≤ P < 5474.95	5470
-----	----	1084.95 ≤ P < 1094.95	1090	5484.95 ≤ P < 5494.95	5490
-----	----	1094.95 ≤ P < 1104.95	1100	5494.95 ≤ P < 5504.95	5500
-----	----	1104.95 ≤ P < 1114.95	1110	5504.95 ≤ P < 5514.95	5510
-----	----	1114.95 ≤ P < 1124.95	1120	5514.95 ≤ P < 5524.95	5520
-----	----	1124.95 ≤ P < 1134.95	1130	5524.95 ≤ P < 5534.95	5530
-----	----	1134.95 ≤ P < 1144.95	1140	5534.95 ≤ P < 5544.95	5540
-----	----	1144.95 ≤ P < 1154.95	1150	5544.95 ≤ P < 5554.95	5550

Tabela AI.1 – Código de equivalência do método 1 – Mini-hídrica, Eólica e Consumo (Cont).

Mini-hídrica		Eólica		Consumo	
-----	----	1154.95 ≤ P < 1164.95	1160	5554.95 ≤ P < 5564.95	5560
-----	----	1164.95 ≤ P < 1174.95	1170	5564.95 ≤ P < 5574.95	5570
-----	----	1174.95 ≤ P < 1184.95	1180	5574.95 ≤ P < 5584.95	5580
-----	----	1194.95 ≤ P < 1204.95	1200	5584.95 ≤ P < 5594.95	5590
-----	----	1204.95 ≤ P < 1214.95	1210	5594.95 ≤ P < 5604.95	5600
-----	----	1214.95 ≤ P < 1224.95	1220	5604.95 ≤ P < 5614.95	5610
-----	----	1234.95 ≤ P < 1244.95	1240	5614.95 ≤ P < 5624.95	5620
-----	----	1244.95 ≤ P < 1254.95	1250	5624.95 ≤ P < 5634.95	5630
-----	----	1264.95 ≤ P < 1274.95	1270	5634.95 ≤ P < 5644.95	5640
-----	----	1274.95 ≤ P < 1284.95	1280	5644.95 ≤ P < 5654.95	5650
-----	----	1284.95 ≤ P < 1294.95	1290	5654.95 ≤ P < 5664.95	5660
-----	----	1294.95 ≤ P < 1304.95	1300	5664.95 ≤ P < 5674.95	5670
-----	----	1304.95 ≤ P < 1314.95	1310	5674.95 ≤ P < 5684.95	5680
-----	----	1314.95 ≤ P < 1324.95	1320	5684.95 ≤ P < 5694.95	5690
-----	----	1334.95 ≤ P < 1344.95	1340	5694.95 ≤ P < 5704.95	5700
-----	----	1354.95 ≤ P < 1364.95	1360	5704.95 ≤ P < 5714.95	5710
-----	----	1364.95 ≤ P < 1374.95	1370	5724.95 ≤ P < 5734.95	5730
-----	----	1374.95 ≤ P < 1384.95	1380	5734.95 ≤ P < 5744.95	5740
-----	----	1384.95 ≤ P < 1394.95	1390	5744.95 ≤ P < 5754.95	5750
-----	----	1394.95 ≤ P < 1404.95	1400	5754.95 ≤ P < 5764.95	5760
-----	----	1404.95 ≤ P < 1414.95	1410	5764.95 ≤ P < 5774.95	5770
-----	----	1414.95 ≤ P < 1424.95	1420	5774.95 ≤ P < 5784.95	5780
-----	----	1424.95 ≤ P < 1434.95	1430	5784.95 ≤ P < 5794.95	5790
-----	----	1434.95 ≤ P < 1444.95	1440	5794.95 ≤ P < 5804.95	5800
-----	----	1454.95 ≤ P < 1464.95	1460	5804.95 ≤ P < 5814.95	5810
-----	----	1464.95 ≤ P < 1474.95	1470	5814.95 ≤ P < 5824.95	5820
-----	----	1474.95 ≤ P < 1484.95	1480	5824.95 ≤ P < 5834.95	5830
-----	----	1484.95 ≤ P < 1494.95	1490	5834.95 ≤ P < 5844.95	5840
-----	----	1494.95 ≤ P < 1504.95	1500	5854.95 ≤ P < 5864.95	5860
-----	----	1514.95 ≤ P < 1524.95	1520	5864.95 ≤ P < 5874.95	5870
-----	----	1524.95 ≤ P < 1534.95	1530	5874.95 ≤ P < 5884.95	5880
-----	----	1534.95 ≤ P < 1544.95	1540	5884.95 ≤ P < 5894.95	5890
-----	----	1544.95 ≤ P < 1554.95	1550	5894.95 ≤ P < 5904.95	5900
-----	----	1554.95 ≤ P < 1564.95	1560	5904.95 ≤ P < 5914.95	5910
-----	----	1564.95 ≤ P < 1574.95	1570	5924.95 ≤ P < 5934.95	5930
-----	----	1574.95 ≤ P < 1584.95	1580	5934.95 ≤ P < 5944.95	5940
-----	----	1584.95 ≤ P < 1594.95	1590	5944.95 ≤ P < 5954.95	5950
-----	----	1594.95 ≤ P < 1604.95	1600	5954.95 ≤ P < 5964.95	5960
-----	----	1604.95 ≤ P < 1614.95	1610	5974.95 ≤ P < 5984.95	5980
-----	----	1614.95 ≤ P < 1624.95	1620	5984.95 ≤ P < 5994.95	5990
-----	----	1644.95 ≤ P < 1654.95	1650	5994.95 ≤ P < 6004.95	6000

Tabela AI.1 – Código de equivalência do método 1 – Mini-hídrica, Eólica e Consumo (Cont).

Mini-hídrica		Eólica		Consumo	
-----	----	1674.95 ≤ P < 1684.95	1680	6004.95 ≤ P < 6014.95	6010
-----	----	1694.95 ≤ P < 1704.95	1700	6014.95 ≤ P < 6024.95	6020
-----	----	1704.95 ≤ P < 1714.95	1710	6024.95 ≤ P < 6034.95	6030
-----	----	1714.95 ≤ P < 1724.95	1720	6034.95 ≤ P < 6044.95	6040
-----	----	1724.95 ≤ P < 1734.95	1730	6044.95 ≤ P < 6054.95	6050
-----	----	1754.95 ≤ P < 1764.95	1760	6074.95 ≤ P < 6084.95	6080
-----	----	1764.95 ≤ P < 1774.95	1770	6094.95 ≤ P < 6104.95	6100
-----	----	1794.95 ≤ P < 1804.95	1800	6104.95 ≤ P < 6114.95	6110
-----	----	1804.95 ≤ P < 1814.95	1810	6114.95 ≤ P < 6124.95	6120
-----	----	1814.95 ≤ P < 1824.95	1820	6124.95 ≤ P < 6134.95	6130
-----	----	1824.95 ≤ P < 1834.95	1830	6134.95 ≤ P < 6144.95	6140
-----	----	1844.95 ≤ P < 1854.95	1850	6144.95 ≤ P < 6154.95	6150
-----	----	1864.95 ≤ P < 1874.95	1870	6154.95 ≤ P < 6164.95	6160
-----	----	1894.95 ≤ P < 1904.95	1900	6184.95 ≤ P < 6194.95	6190
-----	----	1954.95 ≤ P < 1964.95	1960	6194.95 ≤ P < 6204.95	6200
-----	----	1984.95 ≤ P < 1994.95	1990	6204.95 ≤ P < 6214.95	6210
-----	----	2004.95 ≤ P < 2014.95	2010	6234.95 ≤ P < 6244.95	6240
-----	----	2014.95 ≤ P < 2024.95	2020	6244.95 ≤ P < 6254.95	6250
-----	----	2024.95 ≤ P < 2034.95	2030	6274.95 ≤ P < 6284.95	6280
-----	----	2034.95 ≤ P < 2044.95	2040	6284.95 ≤ P < 6294.95	6290
-----	----	2044.95 ≤ P < 2054.95	2050	6344.95 ≤ P < 6354.95	6350
-----	----	2054.95 ≤ P < 2064.95	2060	6354.95 ≤ P < 6364.95	6360
-----	----	2084.95 ≤ P < 2094.95	2090	6364.95 ≤ P < 6374.95	6370
-----	----	2114.95 ≤ P < 2124.95	2120	6374.95 ≤ P < 6384.95	6380
-----	----	2124.95 ≤ P < 2134.95	2130	6384.95 ≤ P < 6394.95	6390
-----	----	2134.95 ≤ P < 2144.95	2140	6394.95 ≤ P < 6404.95	6400
-----	----	2154.95 ≤ P < 2164.95	2160	6404.95 ≤ P < 6414.95	6410
-----	----	2164.95 ≤ P < 2174.95	2170	6414.95 ≤ P < 6424.95	6420
-----	----	2184.95 ≤ P < 2194.95	2190	6424.95 ≤ P < 6434.95	6430
-----	----	2194.95 ≤ P < 2204.95	2200	6444.95 ≤ P < 6454.95	6450
-----	----	2224.95 ≤ P < 2234.95	2230	6464.95 ≤ P < 6474.95	6470
-----	----	2234.95 ≤ P < 2244.95	2240	6474.95 ≤ P < 6484.95	6480
-----	----	2244.95 ≤ P < 2254.95	2250	6484.95 ≤ P < 6494.95	6490
-----	----	2264.95 ≤ P < 2274.95	2270	6494.95 ≤ P < 6504.95	6500
-----	----	2284.95 ≤ P < 2294.95	2290	6504.95 ≤ P < 6514.95	6510
-----	----	2314.95 ≤ P < 2324.95	2320	6514.95 ≤ P < 6524.95	6520
-----	----	2324.95 ≤ P < 2334.95	2330	6534.95 ≤ P < 6544.95	6540
-----	----	2334.95 ≤ P < 2344.95	2340	6544.95 ≤ P < 6554.95	6550
-----	----	2344.95 ≤ P < 2354.95	2350	6554.95 ≤ P < 6564.95	6560
-----	----	2404.95 ≤ P < 2414.95	2410	6564.95 ≤ P < 6574.95	6570
-----	----	2414.95 ≤ P < 2424.95	2420	6604.95 ≤ P < 6614.95	6610

Tabela AI.1 – Código de equivalência do método 1 – Mini-hídrica, Eólica e Consumo (Cont).

Mini-hídrica		Eólica		Consumo	
-----	----	2434.95 ≤ P < 2444.95	2440	6614.95 ≤ P < 6624.95	6620
-----	----	2444.95 ≤ P < 2454.95	2450	6624.95 ≤ P < 6634.95	6630
-----	----	2494.95 ≤ P < 2504.95	2500	6634.95 ≤ P < 6644.95	6640
-----	----	2504.95 ≤ P < 2514.95	2510	6644.95 ≤ P < 6654.95	6650
-----	----	2514.95 ≤ P < 2524.95	2520	6674.95 ≤ P < 6684.95	6680
-----	----	2544.95 ≤ P < 2554.95	2550	6704.95 ≤ P < 6714.95	6710
-----	----	2564.95 ≤ P < 2574.95	2570	6724.95 ≤ P < 6734.95	6730
-----	----	2574.95 ≤ P < 2584.95	2580	6764.95 ≤ P < 6774.95	6770
-----	----	2584.95 ≤ P < 2594.95	2590	-----	----
-----	----	2654.95 ≤ P < 2664.95	2660	-----	----
-----	----	2674.95 ≤ P < 2684.95	2680	-----	----
-----	----	2684.95 ≤ P < 2694.95	2690	-----	----
-----	----	2694.95 ≤ P < 2704.95	2700	-----	----
-----	----	2714.95 ≤ P < 2724.95	2720	-----	----
-----	----	2724.95 ≤ P < 2734.95	2730	-----	----
-----	----	2734.95 ≤ P < 2744.95	2740	-----	----
-----	----	2754.95 ≤ P < 2764.95	2760	-----	----
-----	----	2764.95 ≤ P < 2774.95	2770	-----	----
-----	----	2794.95 ≤ P < 2804.95	2800	-----	----
-----	----	2824.95 ≤ P < 2834.95	2830	-----	----
-----	----	2844.95 ≤ P < 2854.95	2850	-----	----
-----	----	2864.95 ≤ P < 2874.95	2870	-----	----
-----	----	2874.95 ≤ P < 2884.95	2880	-----	----
-----	----	2884.95 ≤ P < 2894.95	2890	-----	----
-----	----	2914.95 ≤ P < 2924.95	2920	-----	----
-----	----	2924.95 ≤ P < 2934.95	2930	-----	----
-----	----	2984.95 ≤ P < 2994.95	2990	-----	----
-----	----	3024.95 ≤ P < 3034.95	3030	-----	----
-----	----	3054.95 ≤ P < 3064.95	3060	-----	----
-----	----	3064.95 ≤ P < 3074.95	3070	-----	----
-----	----	3074.95 ≤ P < 3084.95	3080	-----	----
-----	----	3094.95 ≤ P < 3104.95	3100	-----	----
-----	----	3164.95 ≤ P < 3174.95	3170	-----	----
-----	----	3304.95 ≤ P < 3314.95	3310	-----	----
-----	----	3324.95 ≤ P < 3334.95	3330	-----	----
-----	----	3334.95 ≤ P < 3344.95	3340	-----	----
-----	----	3414.95 ≤ P < 3424.95	3420	-----	----
-----	----	3524.95 ≤ P < 3534.95	3530	-----	----
-----	----	3544.95 ≤ P < 3554.95	3550	-----	----
-----	----	3584.95 ≤ P < 3594.95	3590	-----	----
-----	----	3674.95 ≤ P < 3684.95	3680	-----	----

Tabela AI.2 – Código de equivalência do método 1 – Carvão, Gás Natural, Importação e Exportação

Carvão		Gás Natural		Importação		Exportação	
174.95 ≤ P < 184.95	180	0 ≤ P < 4.95	0	0 ≤ P < 4.95	0	0 ≤ P < 4.95	0
184.95 ≤ P < 194.95	190	4.95 ≤ P < 14.95	10	4.95 ≤ P < 14.95	10	4.95 ≤ P < 14.95	10
204.95 ≤ P < 214.95	210	14.95 ≤ P < 24.95	20	14.95 ≤ P < 24.95	20	14.95 ≤ P < 24.95	20
224.95 ≤ P < 234.95	230	24.95 ≤ P < 34.95	30	24.95 ≤ P < 34.95	30	24.95 ≤ P < 34.95	30
234.95 ≤ P < 244.95	240	34.95 ≤ P < 44.95	40	34.95 ≤ P < 44.95	40	34.95 ≤ P < 44.95	40
244.95 ≤ P < 254.95	250	44.95 ≤ P < 54.95	50	44.95 ≤ P < 54.95	50	44.95 ≤ P < 54.95	50
254.95 ≤ P < 264.95	260	54.95 ≤ P < 64.95	60	54.95 ≤ P < 64.95	60	54.95 ≤ P < 64.95	60
264.95 ≤ P < 274.95	270	64.95 ≤ P < 74.95	70	64.95 ≤ P < 74.95	70	64.95 ≤ P < 74.95	70
274.95 ≤ P < 284.95	280	74.95 ≤ P < 84.95	80	74.95 ≤ P < 84.95	80	74.95 ≤ P < 84.95	80
284.95 ≤ P < 294.95	290	84.95 ≤ P < 94.95	90	84.95 ≤ P < 94.95	90	84.95 ≤ P < 94.95	90
294.95 ≤ P < 304.95	300	94.95 ≤ P < 104.95	100	94.95 ≤ P < 104.95	100	94.95 ≤ P < 104.95	100
304.95 ≤ P < 314.95	310	104.95 ≤ P < 114.95	110	104.95 ≤ P < 114.95	110	104.95 ≤ P < 114.95	110
314.95 ≤ P < 324.95	320	114.95 ≤ P < 124.95	120	114.95 ≤ P < 124.95	120	114.95 ≤ P < 124.95	120
324.95 ≤ P < 334.95	330	124.95 ≤ P < 134.95	130	124.95 ≤ P < 134.95	130	124.95 ≤ P < 134.95	130
354.95 ≤ P < 364.95	360	134.95 ≤ P < 144.95	140	134.95 ≤ P < 144.95	140	134.95 ≤ P < 144.95	140
374.95 ≤ P < 384.95	380	144.95 ≤ P < 154.95	150	144.95 ≤ P < 154.95	150	154.95 ≤ P < 164.95	160
424.95 ≤ P < 434.95	430	154.95 ≤ P < 164.95	160	154.95 ≤ P < 164.95	160	164.95 ≤ P < 174.95	170
434.95 ≤ P < 444.95	440	164.95 ≤ P < 174.95	170	164.95 ≤ P < 174.95	170	174.95 ≤ P < 184.95	180
454.95 ≤ P < 464.95	460	174.95 ≤ P < 184.95	180	174.95 ≤ P < 184.95	180	184.95 ≤ P < 194.95	190
464.95 ≤ P < 474.95	470	184.95 ≤ P < 194.95	190	184.95 ≤ P < 194.95	190	194.95 ≤ P < 204.95	200
474.95 ≤ P < 484.95	480	194.95 ≤ P < 204.95	200	194.95 ≤ P < 204.95	200	204.95 ≤ P < 214.95	210
484.95 ≤ P < 494.95	490	204.95 ≤ P < 214.95	210	204.95 ≤ P < 214.95	210	214.95 ≤ P < 224.95	220
494.95 ≤ P < 504.95	500	214.95 ≤ P < 224.95	220	214.95 ≤ P < 224.95	220	224.95 ≤ P < 234.95	230
504.95 ≤ P < 514.95	510	224.95 ≤ P < 234.95	230	224.95 ≤ P < 234.95	230	234.95 ≤ P < 244.95	240
534.95 ≤ P < 544.95	540	234.95 ≤ P < 244.95	240	234.95 ≤ P < 244.95	240	244.95 ≤ P < 254.95	250
544.95 ≤ P < 554.95	550	244.95 ≤ P < 254.95	250	244.95 ≤ P < 254.95	250	254.95 ≤ P < 264.95	260
554.95 ≤ P < 564.95	560	254.95 ≤ P < 264.95	260	254.95 ≤ P < 264.95	260	264.95 ≤ P < 274.95	270
584.95 ≤ P < 594.95	590	264.95 ≤ P < 274.95	270	264.95 ≤ P < 274.95	270	274.95 ≤ P < 284.95	280
614.95 ≤ P < 624.95	620	274.95 ≤ P < 284.95	280	274.95 ≤ P < 284.95	280	284.95 ≤ P < 294.95	290
664.95 ≤ P < 674.95	670	284.95 ≤ P < 294.95	290	294.95 ≤ P < 304.95	300	294.95 ≤ P < 304.95	300
684.95 ≤ P < 694.95	690	294.95 ≤ P < 304.95	300	304.95 ≤ P < 314.95	310	304.95 ≤ P < 314.95	310
704.95 ≤ P < 714.95	710	304.95 ≤ P < 314.95	310	314.95 ≤ P < 324.95	320	314.95 ≤ P < 324.95	320
714.95 ≤ P < 724.95	720	314.95 ≤ P < 324.95	320	324.95 ≤ P < 334.95	330	324.95 ≤ P < 334.95	330
734.95 ≤ P < 744.95	740	324.95 ≤ P < 334.95	330	334.95 ≤ P < 344.95	340	334.95 ≤ P < 344.95	340
764.95 ≤ P < 774.95	770	334.95 ≤ P < 344.95	340	344.95 ≤ P < 354.95	350	344.95 ≤ P < 354.95	350
774.95 ≤ P < 784.95	780	344.95 ≤ P < 354.95	350	354.95 ≤ P < 364.95	360	354.95 ≤ P < 364.95	360
794.95 ≤ P < 804.95	800	354.95 ≤ P < 364.95	360	364.95 ≤ P < 374.95	370	364.95 ≤ P < 374.95	370
804.95 ≤ P < 814.95	810	364.95 ≤ P < 374.95	370	374.95 ≤ P < 384.95	380	374.95 ≤ P < 384.95	380
814.95 ≤ P < 824.95	820	374.95 ≤ P < 384.95	380	384.95 ≤ P < 394.95	390	394.95 ≤ P < 404.95	400
824.95 ≤ P < 834.95	830	384.95 ≤ P < 394.95	390	394.95 ≤ P < 404.95	400	404.95 ≤ P < 414.95	410
834.95 ≤ P < 844.95	840	394.95 ≤ P < 404.95	400	414.95 ≤ P < 424.95	420	434.95 ≤ P < 444.95	440
844.95 ≤ P < 854.95	850	404.95 ≤ P < 414.95	410	424.95 ≤ P < 434.95	430	444.95 ≤ P < 454.95	450

Tabela AI.2 – Código de equivalência do método 1 – Carvão, Gás Natural, Importação e Exportação (Cont).

Carvão		Gás Natural		Importação		Exportação	
854.95 ≤ P < 864.95	860	414.95 ≤ P < 424.95	420	434.95 ≤ P < 444.95	440	464.95 ≤ P < 474.95	470
874.95 ≤ P < 884.95	880	444.95 ≤ P < 454.95	450	444.95 ≤ P < 454.95	450	474.95 ≤ P < 484.95	480
884.95 ≤ P < 894.95	890	454.95 ≤ P < 464.95	460	454.95 ≤ P < 464.95	460	484.95 ≤ P < 494.95	490
894.95 ≤ P < 904.95	900	464.95 ≤ P < 474.95	470	464.95 ≤ P < 474.95	470	494.95 ≤ P < 504.95	500
904.95 ≤ P < 914.95	910	494.95 ≤ P < 504.95	500	474.95 ≤ P < 484.95	480	504.95 ≤ P < 514.95	510
914.95 ≤ P < 924.95	920	524.95 ≤ P < 534.95	530	484.95 ≤ P < 494.95	490	514.95 ≤ P < 524.95	520
924.95 ≤ P < 934.95	930	554.95 ≤ P < 564.95	560	494.95 ≤ P < 504.95	500	524.95 ≤ P < 534.95	530
984.95 ≤ P < 994.95	990	584.95 ≤ P < 594.95	590	504.95 ≤ P < 514.95	510	534.95 ≤ P < 544.95	540
994.95 ≤ P < 1004.95	1000	594.95 ≤ P < 604.95	600	514.95 ≤ P < 524.95	520	544.95 ≤ P < 554.95	550
1014.95 ≤ P < 1024.95	1020	614.95 ≤ P < 624.95	620	524.95 ≤ P < 534.95	530	554.95 ≤ P < 564.95	560
1044.95 ≤ P < 1054.95	1050	624.95 ≤ P < 634.95	630	534.95 ≤ P < 544.95	540	564.95 ≤ P < 574.95	570
1054.95 ≤ P < 1064.95	1060	644.95 ≤ P < 654.95	650	544.95 ≤ P < 554.95	550	584.95 ≤ P < 594.95	590
1094.95 ≤ P < 1104.95	1100	654.95 ≤ P < 664.95	660	554.95 ≤ P < 564.95	560	594.95 ≤ P < 604.95	600
1104.95 ≤ P < 1114.95	1110	684.95 ≤ P < 694.95	690	564.95 ≤ P < 574.95	570	614.95 ≤ P < 624.95	620
1114.95 ≤ P < 1124.95	1120	794.95 ≤ P < 804.95	800	574.95 ≤ P < 584.95	580	624.95 ≤ P < 634.95	630
1124.95 ≤ P < 1134.95	1130	844.95 ≤ P < 854.95	850	584.95 ≤ P < 594.95	590	634.95 ≤ P < 644.95	640
1154.95 ≤ P < 1164.95	1160	854.95 ≤ P < 864.95	860	594.95 ≤ P < 604.95	600	644.95 ≤ P < 654.95	650
1164.95 ≤ P < 1174.95	1170	864.95 ≤ P < 874.95	870	604.95 ≤ P < 614.95	610	654.95 ≤ P < 664.95	660
1174.95 ≤ P < 1184.95	1180	874.95 ≤ P < 884.95	880	614.95 ≤ P < 624.95	620	674.95 ≤ P < 684.95	680
1184.95 ≤ P < 1194.95	1190	914.95 ≤ P < 924.95	920	624.95 ≤ P < 634.95	630	684.95 ≤ P < 694.95	690
1204.95 ≤ P < 1214.95	1210	944.95 ≤ P < 954.95	950	634.95 ≤ P < 644.95	640	694.95 ≤ P < 704.95	700
1214.95 ≤ P < 1224.95	1220	954.95 ≤ P < 964.95	960	644.95 ≤ P < 654.95	650	704.95 ≤ P < 714.95	710
1234.95 ≤ P < 1244.95	1240	974.95 ≤ P < 984.95	980	654.95 ≤ P < 664.95	660	724.95 ≤ P < 734.95	730
1254.95 ≤ P < 1264.95	1260	984.95 ≤ P < 994.95	990	664.95 ≤ P < 674.95	670	754.95 ≤ P < 764.95	760
1274.95 ≤ P < 1284.95	1280	1004.95 ≤ P < 1014.95	1010	674.95 ≤ P < 684.95	680	764.95 ≤ P < 774.95	770
1294.95 ≤ P < 1304.95	1300	1074.95 ≤ P < 1084.95	1080	684.95 ≤ P < 694.95	690	804.95 ≤ P < 814.95	810
1304.95 ≤ P < 1314.95	1310	1084.95 ≤ P < 1094.95	1090	694.95 ≤ P < 704.95	700	814.95 ≤ P < 824.95	820
1324.95 ≤ P < 1334.95	1330	-----	----	704.95 ≤ P < 714.95	710	824.95 ≤ P < 834.95	830
1334.95 ≤ P < 1344.95	1340	-----	----	714.95 ≤ P < 724.95	720	834.95 ≤ P < 844.95	840
1344.95 ≤ P < 1354.95	1350	-----	----	724.95 ≤ P < 734.95	730	844.95 ≤ P < 854.95	850
1354.95 ≤ P < 1364.95	1360	-----	----	734.95 ≤ P < 744.95	740	854.95 ≤ P < 864.95	860
1364.95 ≤ P < 1374.95	1370	-----	----	754.95 ≤ P < 764.95	760	864.95 ≤ P < 874.95	870
1374.95 ≤ P < 1384.95	1380	-----	----	764.95 ≤ P < 774.95	770	874.95 ≤ P < 884.95	880
1384.95 ≤ P < 1394.95	1390	-----	----	774.95 ≤ P < 784.95	780	884.95 ≤ P < 894.95	890
1394.95 ≤ P < 1404.95	1400	-----	----	784.95 ≤ P < 794.95	790	904.95 ≤ P < 914.95	910
1404.95 ≤ P < 1414.95	1410	-----	----	794.95 ≤ P < 804.95	800	914.95 ≤ P < 924.95	920
1414.95 ≤ P < 1424.95	1420	-----	----	804.95 ≤ P < 814.95	810	924.95 ≤ P < 934.95	930
1424.95 ≤ P < 1434.95	1430	-----	----	814.95 ≤ P < 824.95	820	984.95 ≤ P < 994.95	990
1434.95 ≤ P < 1444.95	1440	-----	----	824.95 ≤ P < 834.95	830	994.95 ≤ P < 1004.95	1000
1444.95 ≤ P < 1454.95	1450	-----	----	844.95 ≤ P < 854.95	850	1004.95 ≤ P < 1014.95	1010
1454.95 ≤ P < 1464.95	1460	-----	----	854.95 ≤ P < 864.95	860	1014.95 ≤ P < 1024.95	1020
1464.95 ≤ P < 1474.95	1470	-----	----	864.95 ≤ P < 874.95	870	1034.95 ≤ P < 1044.95	1040

Tabela AI.2 – Código de equivalência do método 1 – Carvão, Gás Natural, Importação e Exportação (Cont).

Carvão		Gás Natural		Importação		Exportação	
1474.95 ≤ P < 1484.95	1480	-----	----	874.95 ≤ P < 884.95	880	1054.95 ≤ P < 1064.95	1060
1484.95 ≤ P < 1494.95	1490	-----	----	884.95 ≤ P < 894.95	890	1064.95 ≤ P < 1074.95	1070
1494.95 ≤ P < 1504.95	1500	-----	----	904.95 ≤ P < 914.95	910	1084.95 ≤ P < 1094.95	1090
1504.95 ≤ P < 1514.95	1510	-----	----	914.95 ≤ P < 924.95	920	1104.95 ≤ P < 1114.95	1110
1514.95 ≤ P < 1524.95	1520	-----	----	924.95 ≤ P < 934.95	930	1124.95 ≤ P < 1134.95	1130
1524.95 ≤ P < 1534.95	1530	-----	----	934.95 ≤ P < 944.95	940	1144.95 ≤ P < 1154.95	1150
1534.95 ≤ P < 1544.95	1540	-----	----	944.95 ≤ P < 954.95	950	1154.95 ≤ P < 1164.95	1160
1544.95 ≤ P < 1554.95	1550	-----	----	954.95 ≤ P < 964.95	960	1164.95 ≤ P < 1174.95	1170
1554.95 ≤ P < 1564.95	1560	-----	----	964.95 ≤ P < 974.95	970	1184.95 ≤ P < 1194.95	1190
1564.95 ≤ P < 1574.95	1570	-----	----	974.95 ≤ P < 984.95	980	1204.95 ≤ P < 1214.95	1210
1574.95 ≤ P < 1584.95	1580	-----	----	984.95 ≤ P < 994.95	990	1214.95 ≤ P < 1224.95	1220
1584.95 ≤ P < 1594.95	1590	-----	----	994.95 ≤ P < 1004.95	1000	1234.95 ≤ P < 1244.95	1240
1594.95 ≤ P < 1604.95	1600	-----	----	1014.95 ≤ P < 1024.95	1020	1254.95 ≤ P < 1264.95	1260
1604.95 ≤ P < 1614.95	1610	-----	----	1024.95 ≤ P < 1034.95	1030	1274.95 ≤ P < 1284.95	1280
1614.95 ≤ P < 1624.95	1620	-----	----	1044.95 ≤ P < 1054.95	1050	1294.95 ≤ P < 1304.95	1300
1624.95 ≤ P < 1634.95	1630	-----	----	1054.95 ≤ P < 1064.95	1060	1314.95 ≤ P < 1324.95	1320
1634.95 ≤ P < 1644.95	1640	-----	----	1064.95 ≤ P < 1074.95	1070	1364.95 ≤ P < 1374.95	1370
1644.95 ≤ P < 1654.95	1650	-----	----	1074.95 ≤ P < 1084.95	1080	1384.95 ≤ P < 1394.95	1390
1654.95 ≤ P < 1664.95	1660	-----	----	1094.95 ≤ P < 1104.95	1100	1404.95 ≤ P < 1414.95	1410
1664.95 ≤ P < 1674.95	1670	-----	----	1104.95 ≤ P < 1114.95	1110	1444.95 ≤ P < 1454.95	1450
1674.95 ≤ P < 1684.95	1680	-----	----	1124.95 ≤ P < 1134.95	1130	1464.95 ≤ P < 1474.95	1470
1694.95 ≤ P < 1704.95	1700	-----	----	1134.95 ≤ P < 1144.95	1140	1474.95 ≤ P < 1484.95	1480
1704.95 ≤ P < 1714.95	1710	-----	----	1144.95 ≤ P < 1154.95	1150	1504.95 ≤ P < 1514.95	1510
1714.95 ≤ P < 1724.95	1720	-----	----	1154.95 ≤ P < 1164.95	1160	1514.95 ≤ P < 1524.95	1520
1724.95 ≤ P < 1734.95	1730	-----	----	1164.95 ≤ P < 1174.95	1170	1564.95 ≤ P < 1574.95	1570
1734.95 ≤ P < 1744.95	1740	-----	----	1174.95 ≤ P < 1184.95	1180	1574.95 ≤ P < 1584.95	1580
1744.95 ≤ P < 1754.95	1750	-----	----	1204.95 ≤ P < 1214.95	1210	1604.95 ≤ P < 1614.95	1610
1754.95 ≤ P < 1764.95	1760	-----	----	1224.95 ≤ P < 1234.95	1230	1624.95 ≤ P < 1634.95	1630
-----	----	-----	----	1234.95 ≤ P < 1244.95	1240	1654.95 ≤ P < 1664.95	1660
-----	----	-----	----	1244.95 ≤ P < 1254.95	1250	1804.95 ≤ P < 1814.95	1810
-----	----	-----	----	1284.95 ≤ P < 1294.95	1290	1844.95 ≤ P < 1854.95	1850
-----	----	-----	----	1314.95 ≤ P < 1324.95	1320	1854.95 ≤ P < 1864.95	1860
-----	----	-----	----	1324.95 ≤ P < 1334.95	1330	-----	----
-----	----	-----	----	1334.95 ≤ P < 1344.95	1340	-----	----
-----	----	-----	----	1354.95 ≤ P < 1364.95	1360	-----	----
-----	----	-----	----	1364.95 ≤ P < 1374.95	1370	-----	----
-----	----	-----	----	1374.95 ≤ P < 1384.95	1380	-----	----
-----	----	-----	----	1394.95 ≤ P < 1404.95	1400	-----	----
-----	----	-----	----	1434.95 ≤ P < 1444.95	1440	-----	----
-----	----	-----	----	1514.95 ≤ P < 1524.95	1520	-----	----

Tabela AI.3 – Código de equivalência do método 1 – Albufeira, Fio-de-água e Bombagem

Albufeira		Fio-de-água		Bombagem	
114.95 ≤ P < 124.95	120	184.95 ≤ P < 194.95	190	0 ≤ P < 4.95	0
124.95 ≤ P < 134.95	130	194.95 ≤ P < 204.95	200	4.95 ≤ P < 14.95	10
134.95 ≤ P < 144.95	140	214.95 ≤ P < 224.95	220	14.95 ≤ P < 24.95	20
154.95 ≤ P < 164.95	160	224.95 ≤ P < 234.95	230	24.95 ≤ P < 34.95	30
174.95 ≤ P < 184.95	180	234.95 ≤ P < 244.95	240	34.95 ≤ P < 44.95	40
184.95 ≤ P < 194.95	190	244.95 ≤ P < 254.95	250	44.95 ≤ P < 54.95	50
194.95 ≤ P < 204.95	200	254.95 ≤ P < 264.95	260	54.95 ≤ P < 64.95	60
204.95 ≤ P < 214.95	210	274.95 ≤ P < 284.95	280	64.95 ≤ P < 74.95	70
214.95 ≤ P < 224.95	220	284.95 ≤ P < 294.95	290	74.95 ≤ P < 84.95	80
234.95 ≤ P < 244.95	240	294.95 ≤ P < 304.95	300	84.95 ≤ P < 94.95	90
244.95 ≤ P < 254.95	250	304.95 ≤ P < 314.95	310	94.95 ≤ P < 104.95	100
254.95 ≤ P < 264.95	260	314.95 ≤ P < 324.95	320	104.95 ≤ P < 114.95	110
264.95 ≤ P < 274.95	270	324.95 ≤ P < 334.95	330	114.95 ≤ P < 124.95	120
274.95 ≤ P < 284.95	280	334.95 ≤ P < 344.95	340	124.95 ≤ P < 134.95	130
284.95 ≤ P < 294.95	290	344.95 ≤ P < 354.95	350	134.95 ≤ P < 144.95	140
294.95 ≤ P < 304.95	300	354.95 ≤ P < 364.95	360	144.95 ≤ P < 154.95	150
304.95 ≤ P < 314.95	310	364.95 ≤ P < 374.95	370	154.95 ≤ P < 164.95	160
314.95 ≤ P < 324.95	320	374.95 ≤ P < 384.95	380	164.95 ≤ P < 174.95	170
324.95 ≤ P < 334.95	330	384.95 ≤ P < 394.95	390	174.95 ≤ P < 184.95	180
334.95 ≤ P < 344.95	340	394.95 ≤ P < 404.95	400	184.95 ≤ P < 194.95	190
344.95 ≤ P < 354.95	350	404.95 ≤ P < 414.95	410	194.95 ≤ P < 204.95	200
354.95 ≤ P < 364.95	360	414.95 ≤ P < 424.95	420	204.95 ≤ P < 214.95	210
364.95 ≤ P < 374.95	370	424.95 ≤ P < 434.95	430	214.95 ≤ P < 224.95	220
374.95 ≤ P < 384.95	380	434.95 ≤ P < 444.95	440	224.95 ≤ P < 234.95	230
384.95 ≤ P < 394.95	390	444.95 ≤ P < 454.95	450	234.95 ≤ P < 244.95	240
394.95 ≤ P < 404.95	400	454.95 ≤ P < 464.95	460	244.95 ≤ P < 254.95	250
404.95 ≤ P < 414.95	410	464.95 ≤ P < 474.95	470	254.95 ≤ P < 264.95	260
414.95 ≤ P < 424.95	420	474.95 ≤ P < 484.95	480	274.95 ≤ P < 284.95	280
424.95 ≤ P < 434.95	430	484.95 ≤ P < 494.95	490	284.95 ≤ P < 294.95	290
434.95 ≤ P < 444.95	440	494.95 ≤ P < 504.95	500	294.95 ≤ P < 304.95	300
444.95 ≤ P < 454.95	450	504.95 ≤ P < 514.95	510	304.95 ≤ P < 314.95	310
454.95 ≤ P < 464.95	460	514.95 ≤ P < 524.95	520	314.95 ≤ P < 324.95	320
464.95 ≤ P < 474.95	470	524.95 ≤ P < 534.95	530	324.95 ≤ P < 334.95	330
474.95 ≤ P < 484.95	480	534.95 ≤ P < 544.95	540	334.95 ≤ P < 344.95	340
484.95 ≤ P < 494.95	490	544.95 ≤ P < 554.95	550	344.95 ≤ P < 354.95	350
494.95 ≤ P < 504.95	500	554.95 ≤ P < 564.95	560	354.95 ≤ P < 364.95	360
504.95 ≤ P < 514.95	510	564.95 ≤ P < 574.95	570	384.95 ≤ P < 394.95	390
514.95 ≤ P < 524.95	520	574.95 ≤ P < 584.95	580	394.95 ≤ P < 404.95	400
524.95 ≤ P < 534.95	530	584.95 ≤ P < 594.95	590	414.95 ≤ P < 424.95	420
534.95 ≤ P < 544.95	540	594.95 ≤ P < 604.95	600	-----	-----
544.95 ≤ P < 554.95	550	604.95 ≤ P < 614.95	610	-----	-----
554.95 ≤ P < 564.95	560	614.95 ≤ P < 624.95	620	-----	-----

Tabela AI.3 – Código de equivalência do método 1 – Albufeira, Fio-de-água e Bombagem (Cont).

Albufeira		Fio-de-água		Bombagem	
564.95 ≤ P < 574.95	570	624.95 ≤ P < 634.95	630	-----	----
574.95 ≤ P < 584.95	580	634.95 ≤ P < 644.95	640	-----	----
594.95 ≤ P < 604.95	600	644.95 ≤ P < 654.95	650	-----	----
604.95 ≤ P < 614.95	610	654.95 ≤ P < 664.95	660	-----	----
614.95 ≤ P < 624.95	620	664.95 ≤ P < 674.95	670	-----	----
624.95 ≤ P < 634.95	630	674.95 ≤ P < 684.95	680	-----	----
634.95 ≤ P < 644.95	640	684.95 ≤ P < 694.95	690	-----	----
644.95 ≤ P < 654.95	650	694.95 ≤ P < 704.95	700	-----	----
654.95 ≤ P < 664.95	660	704.95 ≤ P < 714.95	710	-----	----
664.95 ≤ P < 674.95	670	714.95 ≤ P < 724.95	720	-----	----
674.95 ≤ P < 684.95	680	724.95 ≤ P < 734.95	730	-----	----
684.95 ≤ P < 694.95	690	734.95 ≤ P < 744.95	740	-----	----
694.95 ≤ P < 704.95	700	744.95 ≤ P < 754.95	750	-----	----
704.95 ≤ P < 714.95	710	764.95 ≤ P < 774.95	770	-----	----
714.95 ≤ P < 724.95	720	774.95 ≤ P < 784.95	780	-----	----
724.95 ≤ P < 734.95	730	794.95 ≤ P < 804.95	800	-----	----
734.95 ≤ P < 744.95	740	804.95 ≤ P < 814.95	810	-----	----
744.95 ≤ P < 754.95	750	814.95 ≤ P < 824.95	820	-----	----
754.95 ≤ P < 764.95	760	824.95 ≤ P < 834.95	830	-----	----
764.95 ≤ P < 774.95	770	834.95 ≤ P < 844.95	840	-----	----
774.95 ≤ P < 784.95	780	844.95 ≤ P < 854.95	850	-----	----
784.95 ≤ P < 794.95	790	884.95 ≤ P < 894.95	890	-----	----
794.95 ≤ P < 804.95	800	894.95 ≤ P < 904.95	900	-----	----
804.95 ≤ P < 814.95	810	904.95 ≤ P < 914.95	910	-----	----
814.95 ≤ P < 824.95	820	914.95 ≤ P < 924.95	920	-----	----
824.95 ≤ P < 834.95	830	924.95 ≤ P < 934.95	930	-----	----
834.95 ≤ P < 844.95	840	934.95 ≤ P < 944.95	940	-----	----
844.95 ≤ P < 854.95	850	944.95 ≤ P < 954.95	950	-----	----
854.95 ≤ P < 864.95	860	954.95 ≤ P < 964.95	960	-----	----
864.95 ≤ P < 874.95	870	964.95 ≤ P < 974.95	970	-----	----
874.95 ≤ P < 884.95	880	974.95 ≤ P < 984.95	980	-----	----
884.95 ≤ P < 894.95	890	994.95 ≤ P < 1004.95	1000	-----	----
894.95 ≤ P < 904.95	900	1004.95 ≤ P < 1014.95	1010	-----	----
904.95 ≤ P < 914.95	910	1024.95 ≤ P < 1034.95	1030	-----	----
914.95 ≤ P < 924.95	920	1034.95 ≤ P < 1044.95	1040	-----	----
924.95 ≤ P < 934.95	930	1044.95 ≤ P < 1054.95	1050	-----	----
934.95 ≤ P < 944.95	940	1054.95 ≤ P < 1064.95	1060	-----	----
944.95 ≤ P < 954.95	950	1074.95 ≤ P < 1084.95	1080	-----	----
954.95 ≤ P < 964.95	960	1084.95 ≤ P < 1094.95	1090	-----	----
964.95 ≤ P < 974.95	970	1094.95 ≤ P < 1104.95	1100	-----	----
974.95 ≤ P < 984.95	980	1104.95 ≤ P < 1114.95	1110	-----	----
984.95 ≤ P < 994.95	990	1124.95 ≤ P < 1134.95	1130	-----	----

Tabela AI.3 – Código de equivalência do método 1 – Albufeira, Fio-de-água e Bombagem (Cont).

Albufeira		Fio-de-água		Bombagem	
994.95 ≤ P < 1004.95	1000	1134.95 ≤ P < 1144.95	1140	-----	----
1004.95 ≤ P < 1014.95	1010	1154.95 ≤ P < 1164.95	1160	-----	----
1014.95 ≤ P < 1024.95	1020	1194.95 ≤ P < 1204.95	1200	-----	----
1024.95 ≤ P < 1034.95	1030	1244.95 ≤ P < 1254.95	1250	-----	----
1034.95 ≤ P < 1044.95	1040	1264.95 ≤ P < 1274.95	1270	-----	----
1054.95 ≤ P < 1064.95	1060	1284.95 ≤ P < 1294.95	1290	-----	----
1064.95 ≤ P < 1074.95	1070	1294.95 ≤ P < 1304.95	1300	-----	----
1074.95 ≤ P < 1084.95	1080	1334.95 ≤ P < 1344.95	1340	-----	----
1094.95 ≤ P < 1104.95	1100	1344.95 ≤ P < 1354.95	1350	-----	----
1104.95 ≤ P < 1114.95	1110	1354.95 ≤ P < 1364.95	1360	-----	----
1114.95 ≤ P < 1124.95	1120	1364.95 ≤ P < 1374.95	1370	-----	----
1124.95 ≤ P < 1134.95	1130	1394.95 ≤ P < 1404.95	1400	-----	----
1134.95 ≤ P < 1144.95	1140	1414.95 ≤ P < 1424.95	1420	-----	----
1144.95 ≤ P < 1154.95	1150	1444.95 ≤ P < 1454.95	1450	-----	----
1174.95 ≤ P < 1184.95	1180	1454.95 ≤ P < 1464.95	1460	-----	----
1184.95 ≤ P < 1194.95	1190	1474.95 ≤ P < 1484.95	1480	-----	----
1194.95 ≤ P < 1204.95	1200	1484.95 ≤ P < 1494.95	1490	-----	----
1204.95 ≤ P < 1214.95	1210	1494.95 ≤ P < 1504.95	1500	-----	----
1254.95 ≤ P < 1264.95	1260	1514.95 ≤ P < 1524.95	1520	-----	----
1274.95 ≤ P < 1284.95	1280	1524.95 ≤ P < 1534.95	1530	-----	----
1284.95 ≤ P < 1294.95	1290	1544.95 ≤ P < 1554.95	1550	-----	----
1294.95 ≤ P < 1304.95	1300	1604.95 ≤ P < 1614.95	1610	-----	----
1304.95 ≤ P < 1314.95	1310	1614.95 ≤ P < 1624.95	1620	-----	----
1324.95 ≤ P < 1334.95	1330	1624.95 ≤ P < 1634.95	1630	-----	----
1344.95 ≤ P < 1354.95	1350	1664.95 ≤ P < 1674.95	1670	-----	----
1354.95 ≤ P < 1364.95	1360	1674.95 ≤ P < 1684.95	1680	-----	----
1364.95 ≤ P < 1374.95	1370	1684.95 ≤ P < 1694.95	1690	-----	----
1374.95 ≤ P < 1384.95	1380	1704.95 ≤ P < 1714.95	1710	-----	----
1384.95 ≤ P < 1394.95	1390	1714.95 ≤ P < 1724.95	1720	-----	----
1394.95 ≤ P < 1404.95	1400	1724.95 ≤ P < 1734.95	1730	-----	----
1404.95 ≤ P < 1414.95	1410	1734.95 ≤ P < 1744.95	1740	-----	----
1414.95 ≤ P < 1424.95	1420	1744.95 ≤ P < 1754.95	1750	-----	----
1434.95 ≤ P < 1444.95	1440	1754.95 ≤ P < 1764.95	1760	-----	----
1454.95 ≤ P < 1464.95	1460	1764.95 ≤ P < 1774.95	1770	-----	----
1474.95 ≤ P < 1484.95	1480	1774.95 ≤ P < 1784.95	1780	-----	----
1484.95 ≤ P < 1494.95	1490	1784.95 ≤ P < 1794.95	1790	-----	----
1494.95 ≤ P < 1504.95	1500	1794.95 ≤ P < 1804.95	1800	-----	----
1514.95 ≤ P < 1524.95	1520	1814.95 ≤ P < 1824.95	1820	-----	----
1524.95 ≤ P < 1534.95	1530	1824.95 ≤ P < 1834.95	1830	-----	----
1584.95 ≤ P < 1594.95	1590	1834.95 ≤ P < 1844.95	1840	-----	----
1594.95 ≤ P < 1604.95	1600	1844.95 ≤ P < 1854.95	1850	-----	----
1614.95 ≤ P < 1624.95	1620	1854.95 ≤ P < 1864.95	1860	-----	----

Tabela AI.3 – Código de equivalência do método 1 – Albufeira, Fio-de-água e Bombagem (Cont).

Albufeira		Fio-de-água		Bombagem	
1624.95 ≤ P < 1634.95	1630	1864.95 ≤ P < 1874.95	1870	-----	----
1644.95 ≤ P < 1654.95	1650	1874.95 ≤ P < 1884.95	1880	-----	----
1664.95 ≤ P < 1674.95	1670	1884.95 ≤ P < 1894.95	1890	-----	----
1674.95 ≤ P < 1684.95	1680	1894.95 ≤ P < 1904.95	1900	-----	----
1724.95 ≤ P < 1734.95	1730	1904.95 ≤ P < 1914.95	1910	-----	----
1744.95 ≤ P < 1754.95	1750	1914.95 ≤ P < 1924.95	1920	-----	----
1834.95 ≤ P < 1844.95	1840	1924.95 ≤ P < 1934.95	1930	-----	----
1854.95 ≤ P < 1864.95	1860	1934.95 ≤ P < 1944.95	1940	-----	----
1894.95 ≤ P < 1904.95	1900	1964.95 ≤ P < 1974.95	1970	-----	----
1984.95 ≤ P < 1994.95	1990	1974.95 ≤ P < 1984.95	1980	-----	----
2024.95 ≤ P < 2034.95	2030	1994.95 ≤ P < 2004.95	2000	-----	----
2154.95 ≤ P < 2164.95	2160	2004.95 ≤ P < 2014.95	2010	-----	----
-----	----	2014.95 ≤ P < 2024.95	2020	-----	----
-----	----	2024.95 ≤ P < 2034.95	2030	-----	----
-----	----	2034.95 ≤ P < 2044.95	2040	-----	----
-----	----	2044.95 ≤ P < 2054.95	2050	-----	----
-----	----	2054.95 ≤ P < 2064.95	2060	-----	----
-----	----	2074.95 ≤ P < 2084.95	2080	-----	----
-----	----	2094.95 ≤ P < 2104.95	2100	-----	----
-----	----	2104.95 ≤ P < 2114.95	2110	-----	----
-----	----	2114.95 ≤ P < 2124.95	2120	-----	----
-----	----	2124.95 ≤ P < 2134.95	2130	-----	----
-----	----	2134.95 ≤ P < 2144.95	2140	-----	----
-----	----	2144.95 ≤ P < 2154.95	2150	-----	----
-----	----	2154.95 ≤ P < 2164.95	2160	-----	----
-----	----	2164.95 ≤ P < 2174.95	2170	-----	----
-----	----	2214.95 ≤ P < 2224.95	2220	-----	----
-----	----	2234.95 ≤ P < 2244.95	2240	-----	----
-----	----	2254.95 ≤ P < 2264.95	2260	-----	----

Tabela AI.4 – Código de equivalência do método 2 – Mini-hídrica, Eólica e Consumo

Mini-hídrica		Eólica		Consumo	
$0 \leq P < 49.95$	0	$49.95 \leq P < 149.95$	100	$4049.95 \leq P < 4149.95$	4100
$49.95 \leq P < 149.95$	100	$149.95 \leq P < 249.95$	200	$4149.95 \leq P < 4249.95$	4200
$149.95 \leq P < 249.95$	200	$249.95 \leq P < 349.95$	300	$4249.95 \leq P < 4349.95$	4300
$249.95 \leq P < 349.95$	300	$349.95 \leq P < 449.95$	400	$4349.95 \leq P < 4449.95$	4400
$349.95 \leq P < 449.95$	400	$449.95 \leq P < 549.95$	500	$4449.95 \leq P < 4549.95$	4500
-----	----	$549.95 \leq P < 649.95$	600	$4549.95 \leq P < 4649.95$	4600
-----	----	$649.95 \leq P < 749.95$	700	$4649.95 \leq P < 4749.95$	4700
-----	----	$749.95 \leq P < 849.95$	800	$4749.95 \leq P < 4849.95$	4800
-----	----	$849.95 \leq P < 949.95$	900	$4849.95 \leq P < 4949.95$	4900
-----	----	$949.95 \leq P < 1049.95$	1000	$4949.95 \leq P < 5049.95$	5000
-----	----	$1049.95 \leq P < 1149.95$	1100	$5049.95 \leq P < 5149.95$	5100
-----	----	$1149.95 \leq P < 1249.95$	1200	$5149.95 \leq P < 5249.95$	5200
-----	----	$1249.95 \leq P < 1349.95$	1300	$5249.95 \leq P < 5349.95$	5300
-----	----	$1349.95 \leq P < 1449.95$	1400	$5349.95 \leq P < 5449.95$	5400
-----	----	$1449.95 \leq P < 1549.95$	1500	$5449.95 \leq P < 5549.95$	5500
-----	----	$1549.95 \leq P < 1649.95$	1600	$5549.95 \leq P < 5649.95$	5600
-----	----	$1649.95 \leq P < 1749.95$	1700	$5649.95 \leq P < 5749.95$	5700
-----	----	$1749.95 \leq P < 1849.95$	1800	$5749.95 \leq P < 5849.95$	5800
-----	----	$1849.95 \leq P < 1949.95$	1900	$5849.95 \leq P < 5949.95$	5900
-----	----	$1949.95 \leq P < 2049.95$	2000	$5949.95 \leq P < 6049.95$	6000
-----	----	$2049.95 \leq P < 2149.95$	2100	$6049.95 \leq P < 6149.95$	6100
-----	----	$2149.95 \leq P < 2249.95$	2200	$6149.95 \leq P < 6249.95$	6200
-----	----	$2249.95 \leq P < 2349.95$	2300	$6249.95 \leq P < 6349.95$	6300
-----	----	$2349.95 \leq P < 2449.95$	2400	$6349.95 \leq P < 6449.95$	6400
-----	----	$2449.95 \leq P < 2549.95$	2500	$6449.95 \leq P < 6549.95$	6500
-----	----	$2549.95 \leq P < 2649.95$	2600	$6549.95 \leq P < 6649.95$	6600
-----	----	$2649.95 \leq P < 2749.95$	2700	$6649.95 \leq P < 6749.95$	6700
-----	----	$2749.95 \leq P < 2849.95$	2800	$6749.95 \leq P < 6849.95$	6800
-----	----	$2849.95 \leq P < 2949.95$	2900	-----	----
-----	----	$2949.95 \leq P < 3049.95$	3000	-----	----
-----	----	$3049.95 \leq P < 3149.95$	3100	-----	----
-----	----	$3149.95 \leq P < 3249.95$	3200	-----	----
-----	----	$3249.95 \leq P < 3349.95$	3300	-----	----
-----	----	$3349.95 \leq P < 3449.95$	3400	-----	----
-----	----	$3449.95 \leq P < 3549.95$	3500	-----	----
-----	----	$3549.95 \leq P < 3649.95$	3600	-----	----
-----	----	$3649.95 \leq P < 3749.95$	3700	-----	----

Tabela AI.5 – Código de equivalência do método 2 – Carvão, Gás Natural, Importação e Exportação

Carvão		Gás Natural		Importação		Exportação	
149.95 ≤ P < 249.95	200	0 ≤ P < 49.95	0	0 ≤ P < 49.95	0	0 ≤ P < 49.95	0
249.95 ≤ P < 349.95	300	49.95 ≤ P < 149.95	100	49.95 ≤ P < 149.95	100	49.95 ≤ P < 149.95	100
349.95 ≤ P < 449.95	400	149.95 ≤ P < 249.95	200	149.95 ≤ P < 249.95	200	149.95 ≤ P < 249.95	200
449.95 ≤ P < 549.95	500	249.95 ≤ P < 349.95	300	249.95 ≤ P < 349.95	300	249.95 ≤ P < 349.95	300
549.95 ≤ P < 649.95	600	349.95 ≤ P < 449.95	400	349.95 ≤ P < 449.95	400	349.95 ≤ P < 449.95	400
649.95 ≤ P < 749.95	700	449.95 ≤ P < 549.95	500	449.95 ≤ P < 549.95	500	449.95 ≤ P < 549.95	500
749.95 ≤ P < 849.95	800	549.95 ≤ P < 649.95	600	549.95 ≤ P < 649.95	600	549.95 ≤ P < 649.95	600
849.95 ≤ P < 949.95	900	649.95 ≤ P < 749.95	700	649.95 ≤ P < 749.95	700	649.95 ≤ P < 749.95	700
949.95 ≤ P < 1049.95	1000	749.95 ≤ P < 849.95	800	749.95 ≤ P < 849.95	800	749.95 ≤ P < 849.95	800
1049.95 ≤ P < 1149.95	1100	849.95 ≤ P < 949.95	900	849.95 ≤ P < 949.95	900	849.95 ≤ P < 949.95	900
1149.95 ≤ P < 1249.95	1200	949.95 ≤ P < 1049.95	1000	949.95 ≤ P < 1049.95	1000	949.95 ≤ P < 1049.95	1000
1249.95 ≤ P < 1349.95	1300	1049.95 ≤ P < 1149.95	1100	1049.95 ≤ P < 1149.95	1100	1049.95 ≤ P < 1149.95	1100
1349.95 ≤ P < 1449.95	1400	-----	----	1149.95 ≤ P < 1249.95	1200	1149.95 ≤ P < 1249.95	1200
1449.95 ≤ P < 1549.95	1500	-----	----	1249.95 ≤ P < 1349.95	1300	1249.95 ≤ P < 1349.95	1300
1549.95 ≤ P < 1649.95	1600	-----	----	1349.95 ≤ P < 1449.95	1400	1349.95 ≤ P < 1449.95	1400
1649.95 ≤ P < 1749.95	1700	-----	----	1449.95 ≤ P < 1549.95	1500	1449.95 ≤ P < 1549.95	1500
1749.95 ≤ P < 1849.95	1800	-----	----	-----	----	1549.95 ≤ P < 1649.95	1600
-----	----	-----	----	-----	----	1649.95 ≤ P < 1749.95	1700
-----	----	-----	----	-----	----	1749.95 ≤ P < 1849.95	1800
-----	----	-----	----	-----	----	1849.95 ≤ P < 1949.95	1900

Tabela AI.6 – Código de equivalência do método 2 – Albufeira, Fio-de-água e Bombagem

Albufeira		Fio-de-água		Bombagem	
$49.95 \leq P < 149.95$	100	$149.95 \leq P < 249.95$	200	$0 \leq P < 49.95$	0
$149.95 \leq P < 249.95$	200	$249.95 \leq P < 349.95$	300	$49.95 \leq P < 149.95$	100
$249.95 \leq P < 349.95$	300	$349.95 \leq P < 449.95$	400	$149.95 \leq P < 249.95$	200
$349.95 \leq P < 449.95$	400	$449.95 \leq P < 549.95$	500	$249.95 \leq P < 349.95$	300
$449.95 \leq P < 549.95$	500	$549.95 \leq P < 649.95$	600	$349.95 \leq P < 449.95$	400
$549.95 \leq P < 649.95$	600	$649.95 \leq P < 749.95$	700	-----	----
$649.95 \leq P < 749.95$	700	$749.95 \leq P < 849.95$	800	-----	----
$749.95 \leq P < 849.95$	800	$849.95 \leq P < 949.95$	900	-----	----
$849.95 \leq P < 949.95$	900	$949.95 \leq P < 1049.95$	1000	-----	----
$949.95 \leq P < 1049.95$	1000	$1049.95 \leq P < 1149.95$	1100	-----	----
$1049.95 \leq P < 1149.95$	1100	$1149.95 \leq P < 1249.95$	1200	-----	----
$1149.95 \leq P < 1249.95$	1200	$1249.95 \leq P < 1349.95$	1300	-----	----
$1249.95 \leq P < 1349.95$	1300	$1349.95 \leq P < 1449.95$	1400	-----	----
$1349.95 \leq P < 1449.95$	1400	$1449.95 \leq P < 1549.95$	1500	-----	----
$1449.95 \leq P < 1549.95$	1500	$1549.95 \leq P < 1649.95$	1600	-----	----
$1549.95 \leq P < 1649.95$	1600	$1649.95 \leq P < 1749.95$	1700	-----	----
$1649.95 \leq P < 1749.95$	1700	$1749.95 \leq P < 1849.95$	1800	-----	----
$1749.95 \leq P < 1849.95$	1800	$1849.95 \leq P < 1949.95$	1900	-----	----
$1849.95 \leq P < 1949.95$	1900	$1949.95 \leq P < 2049.95$	2000	-----	----
$1949.95 \leq P < 2049.95$	2000	$2049.95 \leq P < 2149.95$	2100	-----	----
$2049.95 \leq P < 2149.95$	2100	$2149.95 \leq P < 2249.95$	2200	-----	----
$2149.95 \leq P < 2249.95$	2200	$2249.95 \leq P < 2349.95$	2300	-----	----

Anexo II – Artigo ICEUBI 2017

Artigo realizado para o International Congress on Engineering – ICEUBI 2017

Coal Generation Profile for the Day 2 Ahead Congestion Forecast

A. F. dos Santos Martins - afdsmartins@gmail.com

IPC/ISEC

R. Pestana - rui.pestana@ren.pt/rpestana@deea.isel.pt

REN - Rede Eléctrica Nacional, S.A.

ISEL - Instituto Superior de Engenharia de Lisboa

C. I. Faustino Agreira - cif@isec.pt

IPC/ISEC

Abstract

The Day Ahead Congestion Forecast (DACF) arises as a common procedure for power flow forecasts in the network. It is performed on a daily basis (today-for-tomorrow), by all the European Transmission System Operators (TSOs). Currently the System Operator creates 24 savecases for the DACF, based on the results from the daily market of the Iberian Electricity Market (MIBEL), providing it to the ENTSO-E. This is essential in order to identify critical situations in the transmission network. However, the necessity of knowing the procedure for the day after tomorrow (D+2) brings added value, anticipating problems that may occur in the future, affecting the power supply. Since the interconnection capacity strongly depends on the energy generation profile, creating the cases for the Day 2 Ahead Congestion Forecast (D2ACF) makes it possible to calculate the interconnection capacity for the DACF with more accuracy. The developed methodology uses the Rough Sets Theory to search databases for meaningful decision rules in order to acquire knowledge to classify the energy generation profile of the Ordinary Production Regime (PRO) for the day after tomorrow (D+2).

Keywords

Data Mining; Rough Sets Theory; Rule Generation; Congestion Forecast

Coal Generation Profile Forecast for the Day 2 Ahead Congestion Forecast

CT16 - Mathematical Methods in Engineering

Introduction

Over the years, the increase of energy consumption in Portugal and its dependency from petroleum derivatives, that are effecting the environment and whose prices are unstable and increasing, has made it a necessity to invest in other forms of energy, namely, renewable energies. These changes together with the liberalisation of the electricity market have brought major changes and new challenges to the System Operator, whose objective is to ensure the continuity of the supply in accordance with the required safety and quality.

In order to achieve this objective, many procedures have to be fulfilled. Each day, the System Operator informs the *Iberian Electricity Market* - MIBEL about the interconnection capacity between Portugal and Spain for the day ahead [1]. The System Operator uses a tool called *Technical Verification of Programming* - VTP, which uses as input the current snapshots of the grid (24 *savecases*), the generation profile of the Ordinary Regime Production - PRO (based on the information from the MIBEL), the unviability's of the grid, the load forecast and the wind, solar and small hydropower forecast. This tool allows them to create 24 *savecases* forecasting what will happen on the day after (D+1). This is essential in order to identify critical situations and to posteriorly make the necessary changes in order to maintain the security and proper functioning of the grid.

The necessity of knowing the procedure for the day after tomorrow (D+2) brings added value, anticipating problems that may occur in the future, affecting the safety, quality and reliability of the electrical grid. Despite the available information regarding the inputs necessary for the day after tomorrow, the MIBEL doesn't provide an important input, the generation profile of the PRO for the day (D+2). Consequently, this doesn't allow the System Operator to forecast the 24 *savecases* necessary for that day. The PRO is defined as the electricity generated from traditional non-renewable sources and large hydroelectric power plants [2]. This excludes all the production of electricity subject to special regulatory regime (including cogeneration, minigeneration and microgeneration) and production of electricity from endogenous resources, whether renewable or non-renewable. The PRO is grouped in two types of power plants, large hydroelectric and thermal. Currently, the thermal consists on all the power plants powered by fossil fuels such as coal and natural gas. On the other hand, the large hydroelectric is formed by the run-of-river hydropower and the reservoir hydropower.

The MIBEL resulted from the cooperation between the Portuguese and Spanish Governments with the aim of promoting the integration of both countries' electrical systems [3]. As integration factors of MIBEL, two divisions were created, responsible for the management of the organised markets:

- i. OMIE (Spanish division), responsible for the management of the day and intraday market;
- ii. OMIP (Portuguese division), responsible for the management of the derivative markets.

The OMIE manages the wholesale market. This division doesn't have the economic information for the day after tomorrow (D+2), for this reason the OMIP is used. The OMIP manages the future market that has the objective of establishing long-term contracts for the purchase and sale of electricity. In order to establish these contracts electricity indexes were conceived based on hour marginal prices formed on the daily market of the Spot Market managed by OMIE, to the Spanish and Portuguese zones [4]. For the Portuguese system the PTEL Base Index is adopted in order to know the price expectation for the future market, making it an indispensable variable for the generation profile forecast.

This paper is devoted to a classification technique used for the forecast of the generation profile of power plants running on coal, using the Rough Set Theory. Due to the amount of different types of power plants, this paper focuses its research only on the forecast of the energy production using coal. However, the study has also been done for the forecast of the energy production using the remaining PRO power plants (e.g. CCGT NG, large hydroelectric). The Rough Set Theory, proposed

by Pawlak in 1982, is a technique used for knowledge discovery in databases and experimental data sets. It is a new mathematical tool used to deal with data classification problems. Rough Set Theory deals with identifying structural relationships between data found in a data table called information table. This information table is made up of columns that represent attributes, rows that represent objects and each row represents a piece of information about the corresponding object [5]. Rough Set Theory constitutes a framework for inducing minimal decision rules that can be used to perform classification tasks. The main goal of the analysis using this theory is to search large databases for meaningful decision rules and then acquire knowledge [6]. An information table, can contain two types of attributes, designated as the decision attribute and the condition attribute. With these attributes it is possible to determine decision rules that specify the decisions (actions) that must be taken when specific conditions (indicated by condition attributes) are satisfied [7].

The approach present in this paper aims to define the attributes necessary to form a robust database capable of providing data to be analysed using the Rough Set Theory. This technique allows the evaluation of the attributes and eliminates, if necessary, redundant attributes. The developed methodology produces a classification for the average daily power, for the coal power plants. Due to the range of energy production of these power plants, the classification is done on a scale varying between 200 MW and 1800 MW, with an increment of 100 MW.

Rough Set Theory

The Rough Set Theory begins with an information table, used to represent the data that will be used. The idea of rough set consists of the approximation of a set by a pair of sets, called lower and upper approximation of this set. There are five regions (or sets) of interest: \underline{RX} and \overline{RX} , and $POS_R(X)$, $BN_R(X)$ and $NEG_R(X)$ [5]. These sets are defined as shown below.

Let a set $X \subseteq U$, R be an equivalence relation and a knowledge base $K = (U, R)$. Two subsets can be associated:

- i. R-lower: $\underline{RX} = \cup \{Y \in U/R : Y \subseteq X\}$
- ii. R-upper: $\overline{RX} = \cup \{Y \in U/R : Y \cap X \neq \emptyset\}$

It means that the elements that belong to \underline{RX} set can certainly be classified as elements of X ; while the elements that belong to \overline{RX} set can possibly be classified as elements of X . In the same way, $POS_R(X)$, $BN_R(X)$ and $NEG_R(X)$ are defined below:

- iii. $POS_R(X) = \underline{RX} \Rightarrow$ certainly member of X
- iv. $NEG_R(X) = U - \overline{RX} \Rightarrow$ certainly non-member of X
- v. $BN_R(X) = \overline{RX} - \underline{RX} \Rightarrow$ possibly member of X

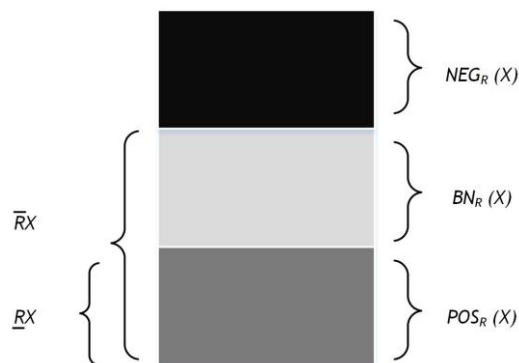


Figure 1 - Definition of R-approximation sets and R-regions

Another important aspect of this theory are the concepts *reduct* and *core*. These concepts are important in the knowledge of base reduction. A *reduct* is the minimal subset of attributes that provide the same quality of classification as the set of all attributes. If the information table has

more than one *reduct*, the intersection of all of them is called the core which is the collection of the most relevant attributes in the table [5, 7, 8].

Application Example

As an application of the Rough Set Theory for the development of a set of rules capable of predicting the energy production of coal power plants, for the day after tomorrow (D+2), we consider a set of data as shown in Table 1. For the sake of clarity, we adopt a reduced number of cases in this example, using only 20 objects, each representing one day. However, the procedures are the same for the complete database. The database used in this paper is from the year 2014.

Table 1-Energy forecast database

No.	PTEL INDEX [€/MWh]	SMALL HYDRO [MW]	WIND [MW]	CONSUMPTION [MW]	COAL [MW]
1	25.0	313	2921	5930	816
2	4.7	307	2928	5539	180
3	12.8	321	2723	6281	264
4	27.0	340	2680	6468	931
5	9.5	346	2042	6483	466
6	16.2	353	1848	6124	384
7	33.7	325	985	6034	592
8	35.7	290	247	5642	725
9	39.0	141	684	4474	1264
10	48.2	113	1440	5539	1447
11	47.4	101	1311	4905	1381
12	41.6	93	2200	5637	1304
13	60.4	62	411	5806	1750
14	51.2	52	336	4569	1753
15	36.1	47	1561	4524	1023
16	56.6	42	942	5790	1709
17	49.5	32	1426	5226	1558
18	54.9	67	2882	5771	1122
19	56.4	51	401	4593	1236
20	45.3	247	2690	5633	1540

Table 1 contains the attributes chosen, represented by the set {PTEL INDEX, SMALL HYDRO, WIND, CONSUMPTION} and the decision attribute, COAL. The PTEL INDEX attribute represents the baseload price index for the Portuguese zone, which corresponds to the arithmetic mean of the hourly marginal prices of the Portuguese system for the 24 (23 or 25) hours of the day [4]. SMALL HYDRO and WIND represent the average daily power of the renewable energies, small hydropower and wind energy. CONSUMPTION represents the average daily power consumption. The decision attribute COAL, represents the average daily power of the coal power plants. These attributes were chosen based on the relevant and available data for the day after tomorrow. The objects for this database were selected based on the COAL attribute. Each object chosen represents a different class, in a total of 17 classes. The remaining 3 objects were included in order to exemplify what occurs when one class has more than one object representing it.

One of the most important operations during the pre-processing phase is to perform the discretization of the attribute values. In this step numerical values are converted into intervals. After testing different discretization techniques, the ones that contributed to better results are represented in two range attribute coding tables, Table 2 and Table 3.

Table 2-Discretization - Range code for PTEL INDEX attribute

PTEL INDEX [€/MWh]	
$0 \leq \text{Attribute} < 4.95$	0
$4.95 \leq \text{Attribute} < 14.95$	10
$14.95 \leq \text{Attribute} < 24.95$	20
$24.95 \leq \text{Attribute} < 34.95$	30
$34.95 \leq \text{Attribute} < 44.95$	40
$44.95 \leq \text{Attribute} < 54.95$	50
$54.95 \leq \text{Attribute} < 64.95$	60
$64.95 \leq \text{Attribute} < 74.95$	70
$74.95 \leq \text{Attribute} < 84.95$	80

Table 3-Discretization - Range attribute code for SMALL HYDRO, WIND, COAL and CONSUMPTION

SMALL HYDRO, WIND, COAL [MW]		CONSUMPTION [MW]	
0 ≤ Attribute < 49.95	0	4449.95 ≤ Attribute < 4549.95	4500
49.95 ≤ Attribute < 149.95	100	4549.95 ≤ Attribute < 4649.95	4600
149.5 ≤ Attribute < 249.95	200	4649.95 ≤ Attribute < 4749.95	4700
249.5 ≤ Attribute < 349.95	300	4749.95 ≤ Attribute < 4849.95	4800
349.5 ≤ Attribute < 449.95	400	4849.95 ≤ Attribute < 4949.95	4900
649.5 ≤ Attribute < 749.95	700	5149.95 ≤ Attribute < 5249.95	5200
849.5 ≤ Attribute < 949.95	900	5449.95 ≤ Attribute < 5549.95	5500
949.95 ≤ Attribute < 1049.95	1000	5549.95 ≤ Attribute < 5649.95	5600
1249.95 ≤ Attribute < 1349.95	1300	5749.95 ≤ Attribute < 5849.95	5800
1349.95 ≤ Attribute < 1449.95	1400	5849.95 ≤ Attribute < 5949.95	5900
1549.95 ≤ Attribute < 1649.95	1600	5949.95 ≤ Attribute < 6049.95	6000
1749.95 ≤ Attribute < 1849.95	1800	6049.95 ≤ Attribute < 6149.95	6100
1949.95 ≤ Attribute < 2049.95	2000	6149.95 ≤ Attribute < 6249.95	6200
2149.95 ≤ Attribute < 2249.95	2200	6249.95 ≤ Attribute < 6349.95	6300
2649.95 ≤ Attribute < 2749.95	2700	6349.95 ≤ Attribute < 6449.95	6400
2849.95 ≤ Attribute < 2949.95	2900	6449.95 ≤ Attribute < 6549.95	6500

Using these redefinitions for each example of Table 1, Table 4 arises rewritten using the coding systems described above.

Table 4-Energy forecast database with range values

No.	PTEL INDEX [€/MWh]	SMALL HYDRO [MW]	WIND [MW]	CONSUMPTION [MW]	COAL [MW]
1	30	300	2900	5900	800
2	0	300	2900	5500	200
3	10	300	2700	6300	300
4	30	300	2700	6500	900
5	10	300	2000	6500	500
6	20	400	1800	6100	400
7	30	300	1000	6000	600
8	40	300	200	5600	700
9	40	100	700	4500	1300
10	50	100	1400	5500	1400
11	50	100	1300	4900	1400
12	40	100	2200	5600	1300
13	60	100	400	5800	1800
14	50	100	300	4600	1800
15	40	0	1600	4500	1000
16	60	0	900	5800	1700
17	50	0	1400	5200	1600
18	50	100	2900	5800	1100
19	60	100	400	4600	1200
20	50	200	2700	5600	1500

After this step, it is necessary to verify if any attribute can be eliminated by repetition. Analyzing Table 4, it can be verified that the attributes are different for all examples, which doesn't permit any attribute elimination [5].

Using the ROSE2 software and the information present in Table 4 we can define the *reduct* and the *core*. This software was developed by the Laboratory of Intelligent Decision Support Systems to implement the Rough Set Theory [9, 10]. In order to minimize the computational complexity while maintaining the same quality of classification, ROSE2 calculates the following *reducts* {PTEL INDEX, SMALL HYDRO, CONSUMPTION} and {WIND, CONSUMPTION}. The intersection of these two *reducts* gives us the *core* which is the attribute {CONSUMPTION}. The core is the subset of attributes whose removal affects the classification power of attributes. Core can be an empty set. This would mean that the software couldn't find which attributes are more or less important and all attributes would have to be considered [11].

The *reduct* chosen was the {PTEL INDEX, SMALL HYDRO, CONSUMPTION}, due to the importance of the effect of the attribute PTEL INDEX in the energy production. Table 5 shows the *reduct* of the database used.

Table 5-Reduct of the energy forecast database

No.	PTEL INDEX [€/MWh]	SMALL HYDRO [MW]	CONSUMPTION [MW]	COAL [MW]
1	30	300	5900	800
2	0	300	5500	200
3	10	300	6300	300
4	30	300	6500	900
5	10	300	6500	500
6	20	400	6100	400
7	30	300	6000	600
8	40	300	5600	700
9	40	100	4500	1300
10	50	100	5500	1400
11	50	100	4900	1400
12	40	100	5600	1300
13	60	100	5800	1800
14	50	100	4600	1800
15	40	0	4500	1000
16	60	0	5800	1700
17	50	0	5200	1600
18	50	100	5800	1100
19	60	100	4600	1200
20	50	200	5600	1500

The information table present in Table 5 gives us the possibility to initiate the learning process, using ROSE2 software. The information table was entered into an input file that was computed using the MODLEM algorithm, expressing the knowledge existent in the form of decision rules. The rules generated by this algorithm use a syntax of conditions in the following form: *attribute* ≤ *value* or *attribute* ≥ *value*. A decision rule is an assertion of the form IF - THEN. The IF side is known as the *antecedent* (or premise) and the THEN side is the *consequent*. The antecedent consists of one or more conditional attributes, and the consequent is a single decision attribute. With the database used, 18 rules were generated. There were two rules generated for the class of 1800 and one rule for each of the other classes. The rules generated are all exact, meaning that each rule only has one decision attribute. Rule 1 and 13 are presented as an example:

rule 1. (PTEL INDEX < 5) => (COAL = 200); [1, 1, 100.00%, 100.00%]

rule 13. (PTEL INDEX >= 45) & (SMALL HYDRO >= 50) & (CONSUMPTION in [4750, 5550]) => (COAL = 1400); [2, 2, 100.00%, 100.00%]

Rule 1 can be interpreted as following:

- If the attribute PTEL INDEX is under the value of 5 (*antecedent*), the average daily power of the coal power plants will be of 200 MW (*consequent or decision*).

The following 4 parameters between brackets describe the absolute cover, absolute support, relative support and the level of discrimination or confidence.

Rule 13 can be interpreted as following:

- If the attributes PTEL INDEX is greater than or equal to 45 and SMALL HYDRO is greater than or equal to 50 and CONSUMPTION is between 4750 and 5550 (*antecedents*), the average daily power of the coal power plants will be of 1400 MW (*consequent or decision*).

The descriptive parameters can be interpreted as following. The absolute cover is the number of learning examples that match the condition part of the rule (*antecedents*). In rule 13 there are two, object 10 and 11. The absolute support is the number of learning examples that match both the condition and the decision parts of the rule. This value is equal to the number of absolute cover, which indicates that the data is consistent. The relative support is the absolute cover divided by the number of examples in the decision class indicated in the decision part of the rule. The level of

discrimination of 100% indicates that the rule was induced from consistent data. It is calculated by dividing the absolute support by the absolute cover.

Conclusions

The generation profile forecast of the PRO for the day after tomorrow is a recent and complex topic. The technique presented in this paper allows the development of a database of rules based on the information collected from the past years, the knowledge discovery provided by the Rough Set Theory and the ROSE2 software. Given these rules and the information already forecasted for the day after tomorrow, we intend to be capable of classifying each day after tomorrow with the correct average daily power of the coal power plants and of the remaining PRO power plants.

The study presents a systematic approach to transform the information present in large databases in a reduced and useful set of rules. Using the complete database of the 365 days of 2014, we were capable of generating 114 rules. Applying these rules to the year 2015, the coal generation profile was forecasted for the 365 days, with an error rate of 21.2%. The application of this theory revealed its importance in this field of study, allowing the creation of a method capable of forecasting the coal generation profile for the day after tomorrow.

References

- [1] Silva, M. A. T.: “Previsão da capacidade de transporte das linhas aéreas para o mercado diário do MIBEL”, Instituto Superior de Engenharia de Lisboa, Lisbon, Portugal, 2016.
- [2] REN. *The Electricity Sector*. [Online]. Available: <https://www.ren.pt/>. [Accessed: 10-Jul-2017].
- [3] Iberian Electricity Market. *MIBEL*. <http://www.mibel.com> [Accessed: 20-Jul-2017].
- [4] OMIP. *Notice 06/2006 - Rules for Determining Electricity Indexes*. [Online]. Available: <http://www.omip.pt/>. [Accessed: 17-May-2016].
- [5] Pawlak, Z.: *Rough Sets - Theoretical Aspects of Reasoning about Data*. Kluwer Academic Publishers. 1991.
- [6] Agreira, C. I. F., Ferreira, C. M. M., Pinto, J. A. D. and Barbosa, F. P. M.: “Steady - State Contingency Classification Using the Rough Set Theory” in *UPEC 2003, 38th Int. Universities’ Power Engineering Conf.*, Thessaloniki, Greece, 1-13 Sept. 2003.
- [7] Rissino, S. and Lambert-Torres, G.: *Rough Set Theory - Fundamental Concepts, Principals, Data Extraction, and Applications*. [Online]. Available: <https://www.intechopen.com>. [Accessed: 05-May-2017].
- [8] Díaz-Martínez, Z., Sánchez-Arellano, A. and Segovia-Vargas M. J.: “Prediction of financial crises by means of rough sets and decision trees” *INNOVAR. Revista de Ciencias Administrativas y Sociales*, Vol. 21 n ° 39 (2011), ISSN:0121-5051, pp. 83-100.
- [9] Predki, B.; Slowinski, R.; Stefanowski, J.; Susmaga, R. and Wilk, Sz.: “ROSE - Software Implementation of the Rough Set Theory In: L.Polkowski, A.Skowron” *Rough Sets and Current Trends in Computing*, Vol. 1424 (1998), ISBN 978-3-540-64655-6, pp. 605-608.
- [10] Predki, B. and Wilk, Sz.: “Rough Set Based Data Exploration Using ROSE System. In: Z.W.Ras, A.Skowron, eds.” *Foundations of Intelligent Systems, Lecture Notes in Artificial Intelligence*, Vol. 1609 (1999), ISBN 978-3-540-65965-5 / ISSN 1611-3349, pp. 172-180.
- [11] Vashist, R. and Garg, M.L.: “Reduct Based Rule Generation for Medical Application Using Rough Sets” *International Journal of Advanced Research in Computer Science*, Vol. 3 n ° 5 (2012), ISSN: 0976-5697, pp. 297-301.