

**CLASSIFICAÇÃO DE LITOFAZES ATRAVÉS DA  
ANÁLISE AUTOMÁTICA DE PERFIS ELÉTRICOS DE  
POÇOS DE PETRÓLEO DA AMAZÔNIA**



**JOACIR MARQUES OLIVEIRA JÚNIOR**

**CLASSIFICAÇÃO DE LITOFAÇIES ATRAVÉS DA  
ANÁLISE AUTOMÁTICA DE PERFIS ELÉTRICOS DE  
POÇOS DE PETRÓLEO DA AMAZÔNIA**

Dissertação apresentada ao Programa de Pós-  
-Graduação em Informática do Instituto de  
Ciências Exatas da Universidade Federal do  
Amazonas como requisito parcial para a  
obtenção do grau de Mestre em Informática.

**ORIENTADOR: EULANDA M. DOS SANTOS E JOSÉ REGINALDO  
CARVALHO**

Campus Universitário Senador Arthur Virgílio Filho

Fevereiro de 2012

© 2012, Joacir Marques Oliveira Júnior.  
Todos os direitos reservados.

Oliveira Júnior, Joacir Marques

Classificação de Litofácies Através da Análise Automática de  
Perfis Elétricos de Poços de Petróleo da Amazônia / Joacir Marques  
Oliveira Júnior. — Campus Universitário Senador Arthur Virgílio  
Filho, 2012

xvi, 46 f. : il. ; 29cm

Dissertação (mestrado) — Universidade Federal do Amazonas  
Orientador: Eulanda M. dos Santos e José Reginaldo Carvalho

1. Computação — Teses. 2. Redes — Teses. I. Orientadora.  
II. Título.

## [Folha de Aprovação]

Quando a secretaria do Curso fornecer esta folha,  
ela deve ser digitalizada e armazenada no disco em formato gráfico.

Se você estiver usando o `pdflatex`,  
armazene o arquivo preferencialmente em formato PNG  
(o formato JPEG é pior neste caso).

Se você estiver usando o `latex` (não o `pdflatex`),  
terá que converter o arquivo gráfico para o formato EPS.

Em seguida, acrescente a opção `approval={nome do arquivo}`  
ao comando `\ppgccufm{img}`.

Se a imagem da folha de aprovação precisar ser ajustada, use:  
`approval=[ajuste] [escala] {nome do arquivo}`  
onde *ajuste* é uma distância para deslocar a imagem para baixo  
e *escala* é um fator de escala para a imagem. Por exemplo:  
`approval=[-2cm] [0.9] {nome do arquivo}`  
desloca a imagem 2cm para cima e a escala em 90%.



# Resumo

Dentro das várias etapas que são necessárias até o petróleo ser comercializado, a análise de perfis elétricos representa papel de grande importância para se estimar a capacidade produtiva de um poço. Um dos objetivos da análise visa a classificação de litofácies. Litofácies são unidades litológicas(rochas) que caracterizam o ambiente de formação e aspectos compositionais das rochas. Para que se forme um reservatório de petróleo, um conjunto de tipos de rochas precisa estar presente, sendo este um dos principais motivos para a classificação de litofácies. A análise hoje é semi-automatizada gerando gráficos de curvas dos quais geólogos com conhecimentos específicos, analisam e realizam a caracterização do reservatório. Esta dissertação de mestrado propõe realizar o reconhecimento automático dos padrões destes dados, demonstrar e comparar classificadores capazes de realizar análise automática de dados de perfis de poços de petróleo. Será ainda comparado o desempenho de classificadores individuais frente à combinação do mesmos através do voto majoritário. Por fim, iremos verificar se o treinamento de um poço pode ser aproveitado para outro por meio de classificadores individuais e combinados por voto majoritário. Para obter essas respostas, fizemos dois tipos de testes. No primeiro, treinamos e testamos os classificadores individualmente e combinados dentro do mesmo poço. Os resultados apresentados mostraram que o classificador Multi-layer Perceptron apresentou a melhor taxa de reconhecimento obtida no primeiro poço(w6) testado atingindo uma taxa de reconhecimento em 98.4029%. Nos demais poços, o classificador SVM atingiu os melhores resultados em 87.3762% para o poço w7 e 85.8889% para o poço w8. No segundo tipo de testes, treinamos em um poço e testamos em outro, simulando uma situação mais próxima do problema real que seria de calibrar os classificadores de uma reserva com um poço pioneiro e a partir daí replicar nos poços vizinho. Neste testes, a combinação de classificadores se mostrou a melhor solução em 4 das 6 combinações possíveis. Nas duas demais combinações, a combinação por voto majoritário foi a segunda melhor. Vale dizer ainda que na média simples, o sistema de votação majoritário, foi a melhor opção para classificar as litofácies. O classificador de regressão logística alcançou a melhor e a pior taxa de acerto atingindo 82.9238% para a combinação onde a base de treino foi para o poço w7 e os testes para o poço w6 e 50.8663% para a combinação w8 para treino e w7 para

testes. Nossos resultados indicam que combinar classificadores em um sistema de votação majoritário, mostra um melhor desempenho e uma melhor estabilidade dos resultados.

**Palavras-chave:** Visão Computacional, Perfilagem, Reconhecimento de Padrões, KNN, SVM, Litofácies.

# Abstract

Among several steps which are necessary in order to lead oil to be marketed, the analysis of well logs plays an important role to estimate the capacity of a well. One goal of this analysis is to classify lithofacies. Lithofacies are lithological units (rocks) that characterize the environment of formation and compositional aspects of the rocks. In order to compose an oil reservoir, a set of types of sedimentary rocks must be present, which is the major reason for the classification of lithofacies. The analysis is conducted in a today semi-automated process which generates graphs of curves used by geologists with specific knowledge to analyze and make the reservoir characterization. We propose in this work to carry out the automatic pattern recognition of well logs to demonstrate and compare classifiers able to perform automatic analysis of well logs data from oil wells obtained from the same and from different reservoirs with accuracy rate higher than or equal to the analysis of human experts.

**Keywords:** Visão Computacional, Perfilagem, Reconhecimento de Padrões, KNN, SVM, Litofácies.



# **Lista de Figuras**

2.1	Torre de Fracionamento.	7
2.2	Formação de petróleo.	9
2.3	Pré-Sal.	11
2.4	Sísmica marítima utilizando canhões de pressão.	12
2.5	Seção Sísmica.	12
2.6	Fluido de perfuração.	13
2.7	Diferentes tipos de broca.	14
2.8	Sonda terrestre.	15
2.9	Sonda marítima.	15
2.10	Perfis elétricos coletados durante a perfilagem de um poço.	21
2.11	Amostras de testemunho retirados de um poço. Apesar de ser a melhor representação do perfil de um poço, é um processo longo e caro, sendo feito portanto somente em alguns poços estratégicos.	22
2.12	Ferramentas de perfilagem utilizadas para extração das características de um poço.	23
2.13	Porosidade e permeabilidade.	24
2.14	Acumulações de Petróleo.	24
3.1	Ilustração sobre o funcionamento de SVM aplicado à classificação de rochas sedimentares.	27
3.2	Exemplo ilustrativo do funcionamento de kNN, bem como o processo de ajuste do parâmetro k.	28
3.3	Exemplo ilustrativo do funcionamento de Multilayer Perceptron.	29
3.4	Ilustração dos cinco Classificadores em conjunto por meio do Voto majoritário.	31
4.1	Amostras de rochas	37
4.2	Gráfico do desempenho dos classificadores treinados e testados no mesmo poço.	39
4.3	Gráfico do desempenho dos classificadores treinados em um poço e testado nos demais.	40

4.4	Box-plot para índice de acertos dos classificadores . . . . .	41
4.5	Desempenho Individual dos Classificadores e Combinados em Conjunto Utilizando Votação Majoritária. . . . .	42

# **Lista de Tabelas**

1.1	Distribuição da Produção de Petróleo e Gás Natural no Brasil por Estado (2011).	2
2.1	Principais Tipos de perfis elétricos e suas propriedades.	19
4.1	Exemplo de amostras da base de dados	36
4.2	Desempenho individual dos classificadores.	38
4.3	Desempenho de poço para poço dos classificadores.	39
4.4	Resumo das Estatísticas Descritivas	40



# Sumário

<b>Resumo</b>	<b>vii</b>
<b>Abstract</b>	<b>ix</b>
<b>Lista de Figuras</b>	<b>xi</b>
<b>Lista de Tabelas</b>	<b>xiii</b>
<b>1 Introdução</b>	<b>1</b>
1.1 Motivação . . . . .	2
1.2 Objetivos . . . . .	4
1.2.1 Geral . . . . .	4
1.2.2 Específicos . . . . .	4
1.3 Organização do Documento . . . . .	5
<b>2 Petróleo, o Ouro Negro</b>	<b>7</b>
2.1 Noções de Geologia de Petróleo . . . . .	8
2.1.1 Origem do Petróleo . . . . .	8
2.1.2 Migração do Petróleo . . . . .	9
2.2 Prospecção de Petróleo . . . . .	10
2.3 Métodos geológicos . . . . .	10
2.4 Métodos Sísmicos . . . . .	11
2.5 Perfuração . . . . .	13
2.6 Avaliação de Formações . . . . .	14
2.6.1 Testemunhagem . . . . .	16
2.6.2 Amostras de Calha . . . . .	17
2.6.3 Perfilagem Elétrica . . . . .	17
2.6.4 Fundamentos de Perfilagem Elétrica . . . . .	18

<b>3 Fundamentação Teórica e Trabalhos Relacionados</b>	<b>25</b>
3.1 Reconhecimento de Padrões . . . . .	25
3.2 Máquinas de Vetores de Suporte - SVM . . . . .	26
3.3 k-Vizinhos mais Próximos - kNN . . . . .	26
3.4 Multilayer Perceptron . . . . .	28
3.5 Regressão Logística . . . . .	29
3.6 Árvores de Decisão - J48 . . . . .	30
3.7 Conjunto de Classificadores . . . . .	30
3.7.1 Votação Majoritária . . . . .	30
3.8 Trabalhos Relacionados . . . . .	32
<b>4 Experimentos e Resultados</b>	<b>35</b>
4.1 Protocolo Experimental . . . . .	35
4.2 Ajuste de Parâmetros . . . . .	36
4.3 Normalização da Base . . . . .	36
4.4 Desempenho Individual e Combinado com Treino e Teste no Mesmo Poço .	38
4.5 Desempenho poço a poço dos classificadores. . . . .	38
4.6 Análise Estatística dos Resultados dos experimentos de Poço a Poço. . . .	40
4.7 Resultados . . . . .	42
<b>5 Conclusões</b>	<b>43</b>
<b>Referências Bibliográficas</b>	<b>45</b>

# **Capítulo 1**

## **Introdução**

A constatação de que a Amazônia é uma região petrolífera aconteceu no dia 13 de março de 1955 [Garcia, 2008]. A euforia tomou conta principalmente do povo do Amazonas quando jorrou petróleo no poço pioneiro batizado de NO-1-AZ, a uma profundidade de 2744 metros e distante 125 quilômetros de Manaus. Essa notícia logo ganhou cenário internacional. Na época acreditava-se que em apenas três anos o Brasil estaria livre da importação de combustíveis, tão grande era o otimismo dos engenheiros da Petrobras. A perspectiva de produção do poço pioneiro NO-1-AZ depois dos primeiros testes realizados era de 600 a 700 barris por dia. Essa perspectiva de produção era suficiente para estimular mais pesquisas e a abertura de novos poços na região. Em um momento em que a produção de petróleo declina mundialmente e seus preços disparam no mercado internacional, a indústria do petróleo brasileira cresce vigorosamente. Hoje o Brasil produz mais de 2,5 milhões de barris por dia e o estado do Amazonas é responsável por cerca de 107.496 barris diários (Tabela 1.1). Segundo o plano de negócios da Petrobras [2011], a empresa espera duplicar suas reservas provadas até 2020.

Apesar de um momento promissor da produção de petróleo nacional, a produção de petróleo ainda dispõe investimentos de cifras milionárias. Antes de um poço começar a produzir, existe um longo e caro processo que envolve o estudo e a análise de dados geológicos e geofísicos. Somente depois de uma exaustiva análise que favoreça a perspectiva positiva de produção em escala comercial é que os engenheiros decidem pela perfuração de um poço. Mesmo depois de perfurado, há ainda a necessidade de se confirmar a viabilidade econômica de exploração da reserva. Uma vez que se iniciou a perfuração de um poço, são descidas ainda durante a perfuração algumas ferramentas com o objetivo de avaliar melhor as características da formação. Essa avaliação é fundamental para justificar a continuidade da exploração da reserva. Se este for o poço pioneiro, provavelmente testes mais caros, porém mais precisos, serão realizados como a testemunhagem e amostra lateral. Esse último mais

**Tabela 1.1.** Distribuição da Produção de Petróleo e Gás Natural no Brasil por Estado (2011).

Estado	Petróleo (bbl/d)	Gás Natural (Mm <sup>3</sup> d)	Produção Total (boe/d)
Rio de Janeiro	1.551.597	25.360	1.711.114
Espírito Santo	319.518	11.901	394.373
Amazonas	35.110	11.508	107.496
Bahia	42.880	6.565	84.172
São Paulo	42.225	5.169	74.739
Rio Grande do Norte	58.639	1.701	69.338
Sergipe	42.365	2.428	57.640
Alagoas	5.198	1.539	14.878
Ceará	7.209	86	7.752
<b>Total geral</b>	<b>2.104.743</b>	<b>66.257</b>	<b>2.521.502</b>

barato que o primeiro.

Esta fase é denominada avaliação de formações. Trata-se de atividades e estudos que visam definir em termos qualitativos e quantitativos o potencial de uma jazida petrolífera, isto é, sua capacidade produtiva e a valoração de suas reservas de óleo e gás . Essa avaliação é baseada principalmente nos dados coletados na perfilagem de um poço aberto. A perfilagem é uma imagem visual em relação à profundidade, de uma ou mais características ou perfis das rochas perfuradas tais como resistividade elétrica, potencial eletroquímico natural, tempo de trânsito das ondas mecânicas, radioatividade natural ou induzida. Com perfilagem busca-se identificar propriedades das rochas e fluidos medindo informações elétricas, acústicas, termais e radioativas sendo possível extrair dados importantes das formações pelas quais o poço atravessa como a litologia (tipo de rocha), espessura, porosidade, presença de óleo, gás ou água. Dependendo da análise dos perfis e das informações complementares, pode se decidir pelo abandono do poço ou no caso de se provar interesse econômico, os testes de formação para se estimar a capacidade produtiva do poço são iniciados.

Na análise de perfis, há oito principais tipos de perfis elétricos: Raios Gama (GR), Neutrônico (NPHI), Indução (ILD), Sônico (DT), Densidade (RHOB), Caliper (Cal), Micro SFL (MSFL) e Índice Fotoelétrico (PEF). A descrição de cada perfil é mostrada na Tabela 2.1.

## 1.1 Motivação

A identificação de litofácies de um reservatório de petróleo é um processo intensivo que envolve o gasto de uma quantidade considerável de tempo por parte de um especialista experiente. O problema se torna muito mais difícil à medida que aumenta o número de perfis

(medidas de determinadas propriedades da formação geológica) simultâneos a serem analisados [Cunha, 2002].

Com base na análise destes perfis, poderá ser criado uma imagem a cada metro de várias características do poço como tipo de rocha, fluidos, características da formação, etc. É um processo que requer bastante habilidade e paciência dos geólogos e geofísicos e sua automatização trará por certo grandes ganhos para o estudo de reservatórios. Além disso, a caracterização de um reservatório é de fundamental importância para justificar investimentos da continuidade de exploração de um poço ou o seu abandono.

Muitos trabalhos já foram feitos com o intuito de analisar automaticamente estes perfis. No entanto, devido a heterogeneidade das diversas bacias sedimentares com características peculiares a elas próprias, o aprendizado em um estudo de uma bacia muitas vezes demonstrando taxas de acerto elevadas, quando levado a bacias diferentes, estes mesmos bons resultados não se repetem. O que muda então de uma bacia para outra ou mesmo de um poço para outro? O conhecimento do especialista geofísico é crucial nesse momento. Não somente o seu conhecimento acadêmico, é muito mais o seu conhecimento da região em estudo. É ele que terá condições portanto de avaliar se um poço possui características semelhantes, avaliando assim se os padrões verificados em um poço se repetem em um poço distinto. É nesse ponto que esse trabalho se destaca. Iremos trabalhar sempre com um poço chave analisado por um especialista geofísico. O aprendizado do mesmo será propagado a novos poços com características semelhantes quanto à formação geológica.

Outro diferencial neste trabalho que será explicado mais a frente é a combinação dos classificadores em voto majoritário. Foram testados 5 classificadores que individualmente mostraram menos estabilidade do que os mesmos 5 em conjunto. Sendo mais específico, quando combinados em conjunto, os classificadores se mantiveram em taxas de acerto elevadas para todos os cenários testados sendo que na média, a combinação atingiu a mais alta taxa de acerto em 79,45%. O classificador *logistic* por exemplo, atingiu 82.92% que foi a maior taxa de acerto, no entanto, também atingiu a taxa mais baixa em 50.87% o que diminui em muito sua confiabilidade mesmo tendo uma boa média em 73.24%.

O trabalho se encaixa dentro da linha Visão Computacional e Robótica, pois será baseado no estudo de percepção de padrões.

A classificação de perfis de poços de petróleo pode claramente ser tratada como um problema de reconhecimento automático de padrões. Não é surpresa a existência de trabalhos na literatura que realizam a classificação dos perfis de poços com ferramentas de reconhecimento de padrões Cunha [2002], Flexa et al. [2004], Wong et al. [2005] e An-nan & Lu [2009] sendo a maioria deles utilizando redes neurais e Máquinas de Vetores de Suporte (SVM).

Diante desse contexto, este trabalho aplica métodos de classificação ao problema de

identificação automática de litofácies utilizando dados de perfis de três poços de petróleo da Amazônia da mesma reserva. Como já foi dito, há muitos trabalhos publicados utilizando técnicas de reconhecimento de padrões. Neste trabalho, além de comparar o desempenho individual de cinco classificadores, é apresentado um sistema de votação majoritária onde escolhe-se a classe com maior número de votos dos classificadores a partir da combinação dos mesmos classificadores listados abaixo:

1. SVM - Support Vector Machine;
2. Multilayer Perceptron;
3. Logistic Regression;
4. k-Nearest Neighbor;
5. J48 Tree.

Este trabalho propõe através de uma das áreas de conhecimento da Informática, fazer o reconhecimento automático dos padrões destes dados e comparar os classificadores SVM, kNN, Multilayer Perceptron e Regressão Logística como classificadores capazes de realizar análise automática de dados de perfis de poços de petróleo da mesma reserva para a classificação de litofácies com índice de acerto igual ou superior à análise de especialistas humanos.

## 1.2 Objetivos

### 1.2.1 Geral

O objetivo geral deste documento é realizar uma análise comparativa entre métodos de classificação aplicados ao problema de classificação automática de litofácies utilizando dados de perfis de poços de petróleo, a fim de mostrar que métodos de classificação automática podem alcançar índice de acerto igual ou superior à análise de especialistas humanos e também identificar o método com maior taxa de classificação correta para o problema.

### 1.2.2 Específicos

Os objetivos específicos são:

- Comparar e combinar classificadores e apontar o melhor método para a classificação automática de litofácies utilizando dados de perfis de poços de petróleo divididos nos seguintes estudos:

- Comparar classificadores dentro do mesmo poço;
- Comparar classificadores dentro de poços diferentes;
- Comparar classificadores individuais e em conjuntos;
- Constatar se o treinamento em um poço pode ser aproveitado em um poço distinto com características semelhantes.

## 1.3 Organização do Documento

O restante deste documento está organizado da seguinte forma: no capítulo 2 são descritos vários conceitos importantes sobre a origem e exploração do petróleo. No capítulo 3 é descrita a fundamentação teórica necessária para a compreensão de reconhecimento de padrões, classificadores e conjunto de classificadores, bem como os trabalhos relacionados.

No capítulo 4, são apresentados a base de dados, normalização da base e descrição dos experimentos realizados com os classificadores através de testes individuais e em conjunto por meio de votação majoritária.

No capítulo 5, finalizamos com as conclusões deste trabalho.

Assim, no próximo capítulo, serão introduzidos os conceitos sobre a origem do petróleo e sua exploração.

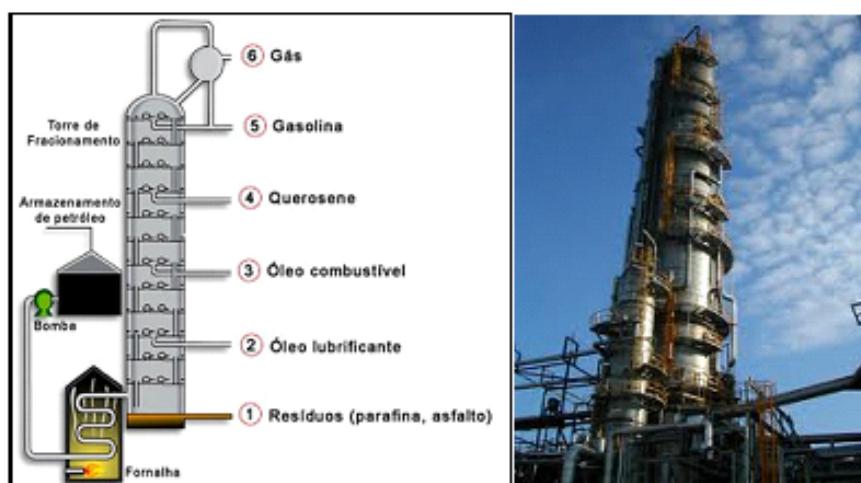


# Capítulo 2

## Petróleo, o Ouro Negro

Segundo Thomas et al. [2001], a palavra Petróleo vem do latim pedra(pedra) e oleum(óleo), o petróleo no estado líquido é uma substância oleosa, inflamável, menos densa que a água, com cheiro característico e cor variando entre o negro e o castanho claro. O petróleo é formado por centenas de compostos químicos, sendo os mais importantes para a geração de derivados, os hidrocarbonetos (moléculas de hidrogênio e carbono). Os derivados do petróleo como querosene, gasolina, lubrificante e outros, são obtidos de acordo com a faixa de ebulação por meio de uma torre de destilação ou fracionamento (Figura 2.1).

Através da torre de destilação (2.1), em uma escala crescente de temperatura, é possível separar cada composto de acordo com a Tabela 2.



**Figura 2.1.** Torre de Fracionamento.

Fração	Temperatura de ebulição (°C)	Composição aproximada	Usos
Gás residual e Gás liquefeito de petróleo - GLP	Até 40	$C_1 - C_2, C_3 - C_4$	gás combustível, gás combustível engarrafado, uso doméstico e industrial.
Gasolina	40 - 175	$C_5 - C_{10}$	combustível de automóvel, solvente.
Querosene	175 - 235	$C_{11} - C_{12}$	iluminação, combustível de aviões a jato.
Gasóleo leve	235 - 305	$C_{13} - C_{17}$	diesel, fornos.
Gasóleo pesado	305 - 400	$C_{18} - C_{25}$	combustível, matéria-prima para lubrificantes.
Lubrificantes	400 - 510	$C_{26} - C_{38}$	óleos lubrificantes.
Resíduo	Acima de 510	$C_{38}+$	asfalto, piche, impermeabilizantes.

As características do petróleo diferem para cada reservatório, podendo ser observados óleos mais viscosos, densos, pretos, com bastante gás ou pouco gás. Muitas vezes também é encontrado apenas gás em um reservatório. Entretanto, além de quantidades menos significativas de enxofre, nitrogênio, oxigênio e metais, uma grande quantidade de hidrogênio e carbono é sempre presente em torno de 11 a 14

## 2.1 Noções de Geologia de Petróleo

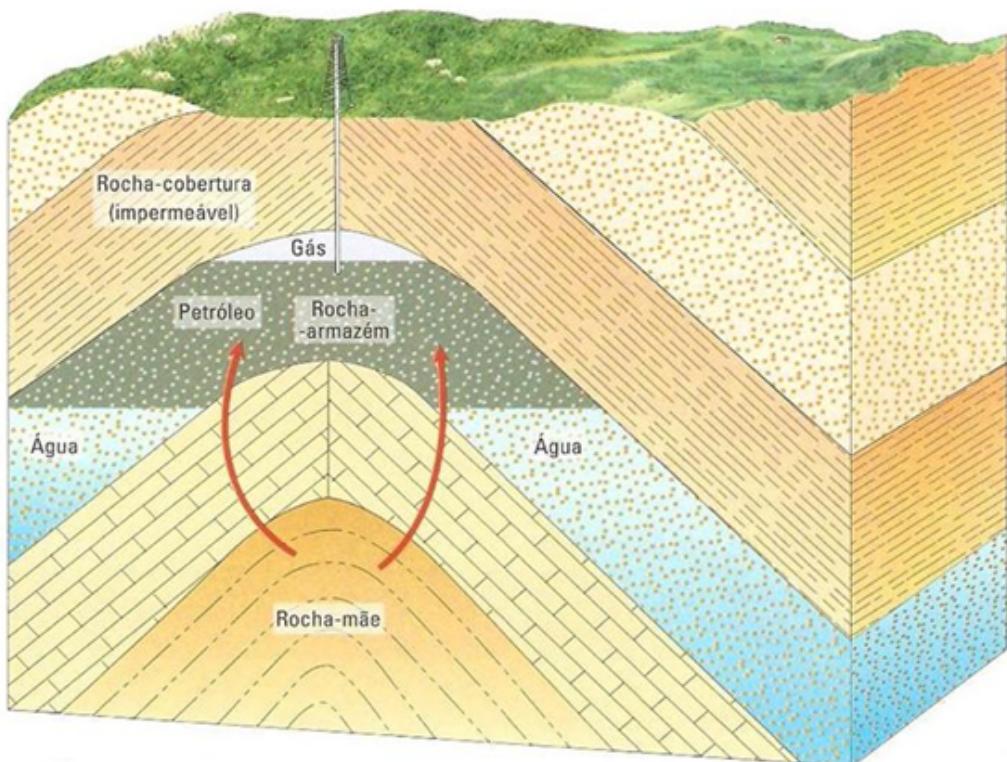
### 2.1.1 Origem do Petróleo

O petróleo tem origem na matéria orgânica depositada junto com os sedimentos. A matéria orgânica marinha é basicamente originada de microorganismos e algas que formam o fitoplanton e não pode sofrer processo de oxidação [Thomas et al., 2001]. Assim, para que óleo ou gás seja formado, há a necessidade de condições não oxidantes, geralmente encontrados em compostos de baixa densidade, não permitindo a ação de água circulante. Ainda segundo Thomas et al. [2001], a interação destes fatores - matéria orgânica, sedimento e condições termoquímicas apropriadas - é fundamental para o início da cadeia de processos que leva à formação do petróleo. O tipo de matéria orgânica irá determinar por exemplo o tipo de hidrocarboneto gerado, óleo ou gás. Por exemplo, matéria orgânica proveniente de vegetais lenhosos poderá dar origem à gás enquanto que matérias orgânicas provenientes de

fitoplâncton submetidas a condições térmicas adequadas, poderá vir a gerar a hidrocarboneto líquido.

### 2.1.2 Migração do Petróleo

Uma vez que o processo de geração de petróleo tenha ocorrido dentro das condições ideais para a formação hidrocarbonetos líquidos ou gasosos, para que um reservatório seja formado e possa ser extraído, é necessário ainda que este óleo ou gás migre para uma rocha com condições de armazenar este hidrocarboneto. Estas rochas são conhecidas como rochas reservatório ou rochas armazém. Os hidrocarbonetos tendem a migrar sempre para cima e irão mover-se nestas rochas até encontrar outra camada rochosa impermeável que interrompa esse fluxo. Estas rochas impermeáveis são frequentemente denominadas de rocha cobertura e dependendo de sua formação é dito que criou-se uma formação geológica conhecida como trapa(armadilha).



**Figura 2.2.** Formação de petróleo.

## 2.2 Prospecção de Petróleo

A descoberta de uma jazida de petróleo é um processo de um longo estudo por parte de geólogos e geofísicos. Hoje, várias ferramentas de TI auxiliam essa tarefa. Graças ao acúmulo de conhecimento por parte de geólogos e geofísicos e ferramentas de TI cada vez mais sofisticadas, a taxa de acerto para perfuração de poços de petróleo pode chegar a 100% como no caso da Bacia de Santos (<http://fatosedados.blogspetrobras.com.br/historico-monitor-fatos-e-dados/destaques-internacionais-16/>). É preciso explicar que esse é um caso excepcional do pré-sal. A camada de pré-sal é uma camada muito extensa e de baixíssima permeabilidade formando uma grande cobertura que impediu que óleo gerado a milhares de metros, fugisse e portanto ficou preso por milhares de anos. Essas reservas eram consideradas inatingíveis e de alto risco. Essa percepção mudou graças à evolução tecnológica e também do alto valor do petróleo que compensa enfim os custos de extração.

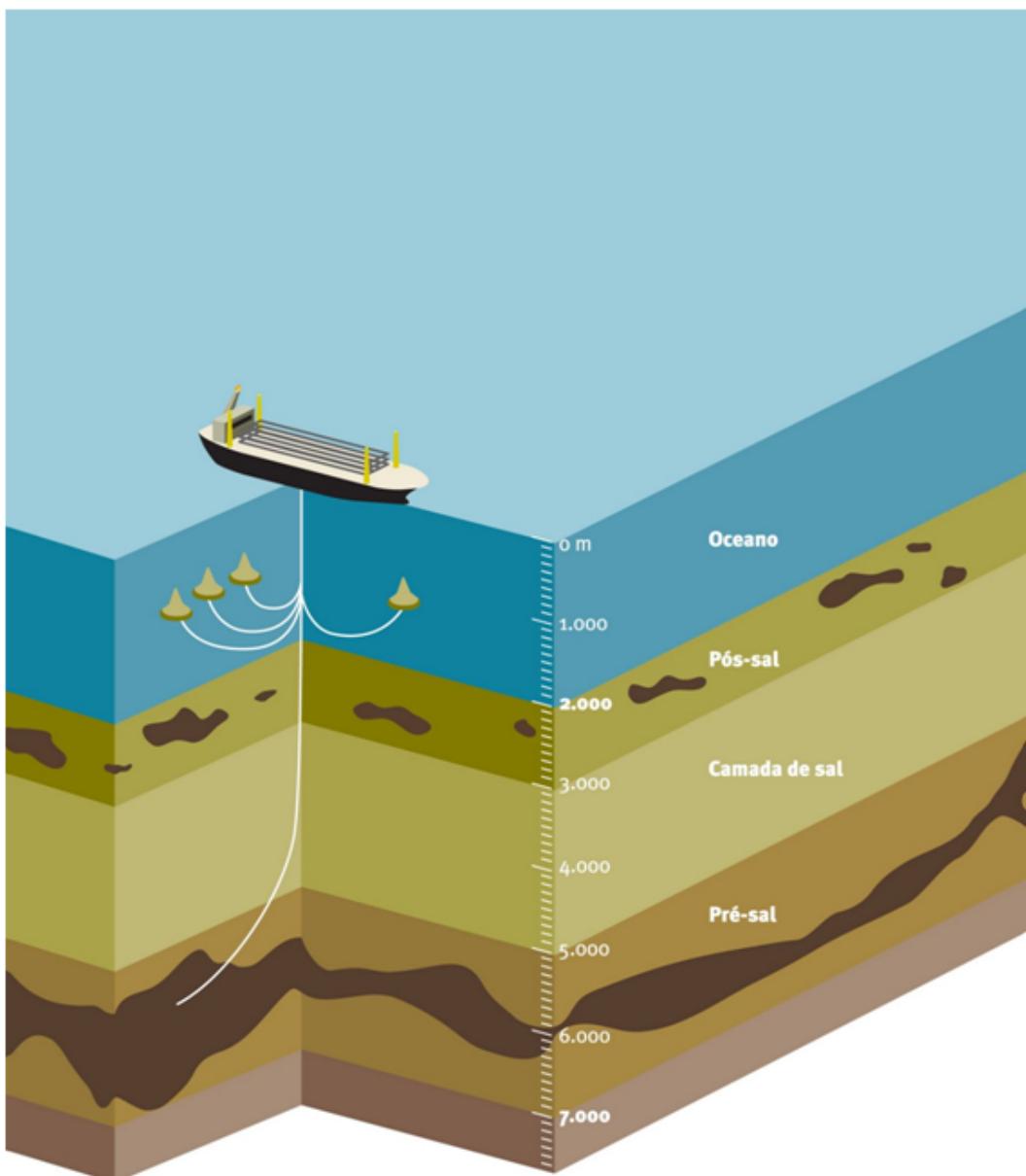
Achar reservas como essas, parte de um estudo conhecido como prospecção de petróleo que tem como objetivos fundamentais:

- Localizar dentro de uma bacia sedimentar as situações geológicas que tenham condição para acumulação de petróleo;
- Verificar qual, dentre estas situações, possui mais chance de conter petróleo.

Esse estudo será feito através de métodos geofísicos e geológicos. São investimentos que se tornam pequenos quando comparados ao alto custo de perfuração de um poço, justificando os mesmos.

## 2.3 Métodos geológicos

É o primeiro passo com o objetivo de se estudar se existem condições de formação de uma reserva em uma determinada região. O geólogo irá elaborar um mapa de geologia de superfície. Este mapa é feito através de rochas que podem ser observadas na superfície. Os mapas geológicos, que indicam as áreas potencialmente interessantes, são continuamente construídos e atualizados pelo exploracionistas. Nestes mapas, as áreas compostas por rochas ígneas e metamórficas são praticamente eliminadas, como também pequenas bacias com espessura sedimentar muito reduzida ou sem estruturas favoráveis à acumulação [Thomas et al., 2001].



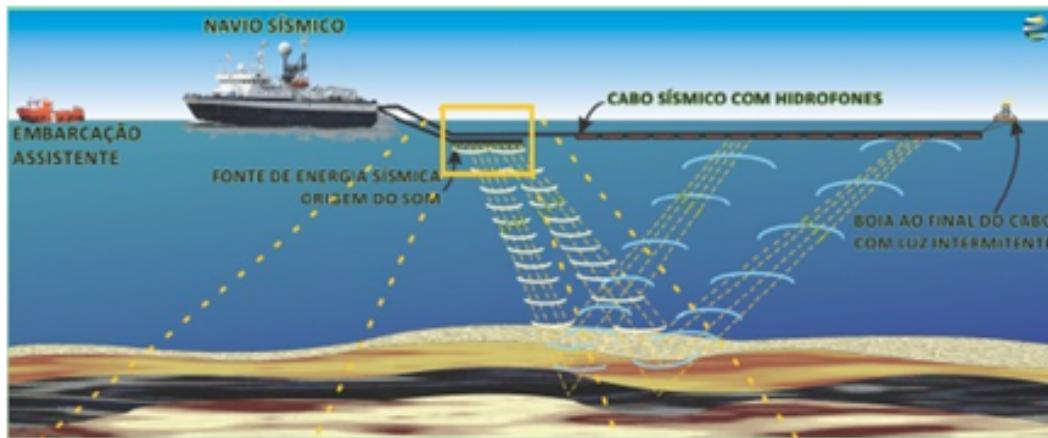
**Figura 2.3.** Pré-Sal.

## 2.4 Métodos Sísmicos

Os mais utilizados são os métodos sísmicos de reflexão. Ele fornece alta definição das feições geológicas em subsuperfície propícias à acumulação de hidrocarboneto, a um custo relativamente baixo.

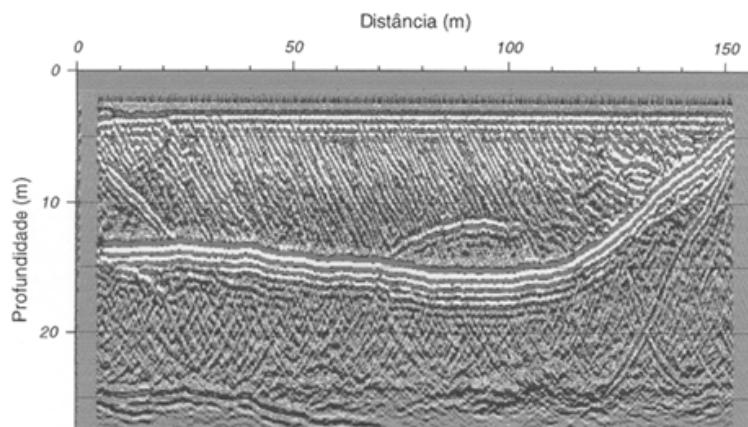
O levantamento sísmico inicia-se com geração de ondas através de fontes artificiais como dinamites e vibradores em terra e canhões de ar comprimido no levantamento marítimo (Figura 2.4). Cada uma dessas fontes irá emitir um sinal que irá refletir nas camadas

geológicas retornando o resultado desse sinal que irá variar de acordo com cada formação. Esses sinais trazem informações valiosas que serão captadas pelos geofones quando em terra e pelos hidrofones quando em água.



**Figura 2.4.** Sísmica marítima utilizando canhões de pressão.

O resultado desse processo será uma seção sísmica (Figura 2.5). Ela é composta por amostras dos valores das amplitudes, espaçadas lateralmente ou em função da distância entre estações receptoras ou em função da distância entre cabos, e verticalmente pela razão da amostragem temporal [Thomas et al., 2001].



**Figura 2.5.** Seção Sísmica.

O próximo passo será processar estes dados através de aplicativos específicos para a indústria do petróleo. O objetivo é eliminar ruídos e produzir imagens da subsuperfície de maneira que facilite a interpretação por geólogos e geofísicos. Estes profissionais interpretam estas imagens procurando situações que favoreçam a acumulação de hidrocarbonetos. Assim, diante destes dados, cabe a estes profissionais a responsabilidade de uma das decisões mais cruciais de uma empresa de petróleo: a perfuração de um poço.

## 2.5 Perfuração

A perfuração de um poço é a fase mais cara do processo de exploração de petróleo. Todo o estudo descrito anteriormente, tem por objetivo que a perfuração seja bem sucedida e que se ache reservatório de óleo ou gás que possam ser comercializados. A perfuração de um poço de petróleo é realizada através de uma sonda. Na perfuração rotativa, as rochas são atravessadas pela ação da rotação e peso aplicados a uma broca existente na extremidade de uma coluna de perfuração a qual consiste basicamente de comandos (tubos de paredes espessas) e tubos de perfuração (tubos de paredes finas). Os fragmentos das rochas são removidos continuamente através de um fluido de perfuração(Figura 2.6) ou simplesmente lama, uma vez que sua base principal é a argila[Thomas et al., 2001].



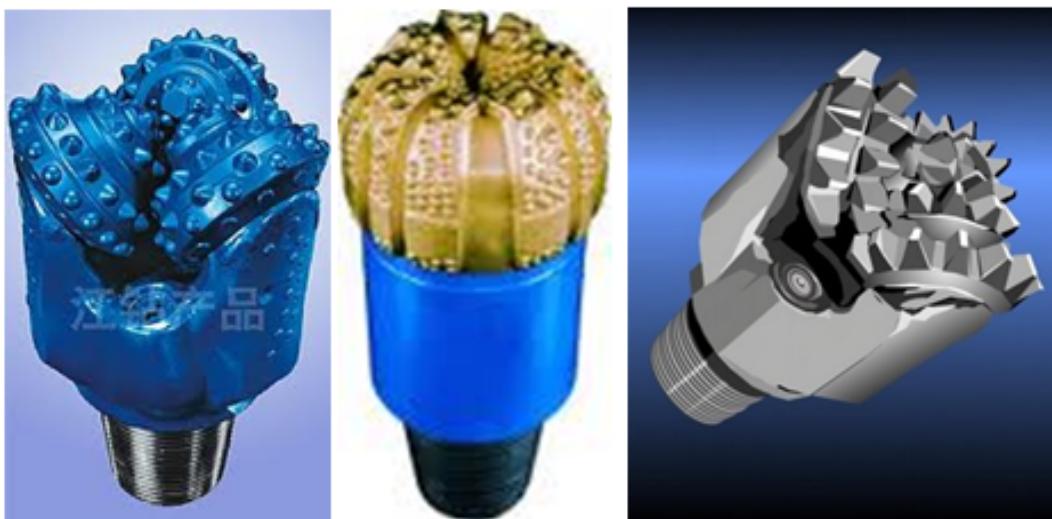
**Figura 2.6.** Fluido de perfuração.

Esses fluidos têm como função:

- Manter as pressões de formação sob controle;
- Carrear os cascalhos até a superfície;
- Manter a estabilidade mecânica do poço;
- Resfriar a broca;
- Transmitir força hidráulica até a broca;
- Manter os cascalhos em suspensão quando sem circulação;

- Entre outros.

O espaço entre os tubos de revestimento e as paredes do poço (anular) é cimentado com a finalidade de isolar as rochas atravessadas, permitindo então o avanço da perfuração com segurança. Sem esses cuidados, pode acontecer o desabamento da parede do poço. Poderíamos nos perguntar se antes da cimentação esse desabamento não ocorre. A resposta é sim, poderia acontecer, no entanto, como citado anteriormente, um dos objetivos do fluido é manter as pressões de formação sobre controle apenas provisoriamente claro, até que a cimentação ocorra. Após as operações de cimentação, a coluna de perfuração é novamente descida ao poço, tendo na sua extremidade uma nova broca de diâmetro menor do que a do revestimento para o prosseguimento da perfuração. Assim, um poço de petróleo é perfurado em diversas fases, caracterizadas pelos diferentes diâmetros das brocas(Figura 2.7).

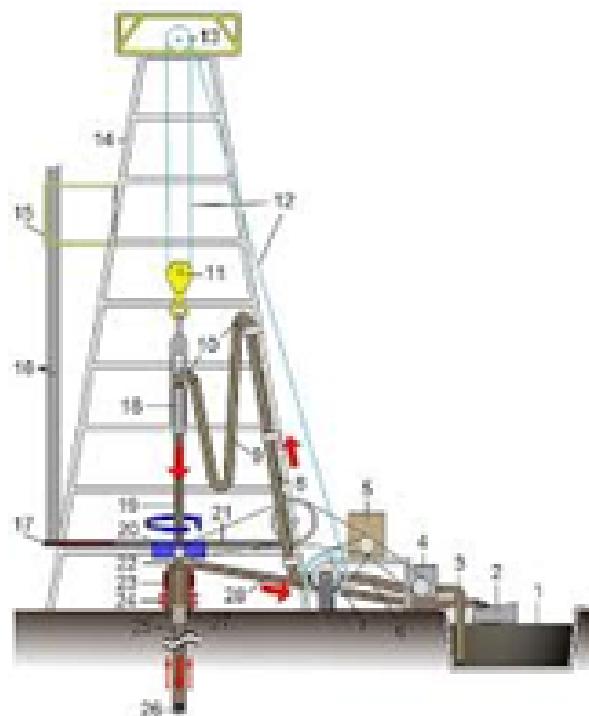


**Figura 2.7.** Diferentes tipos de broca.

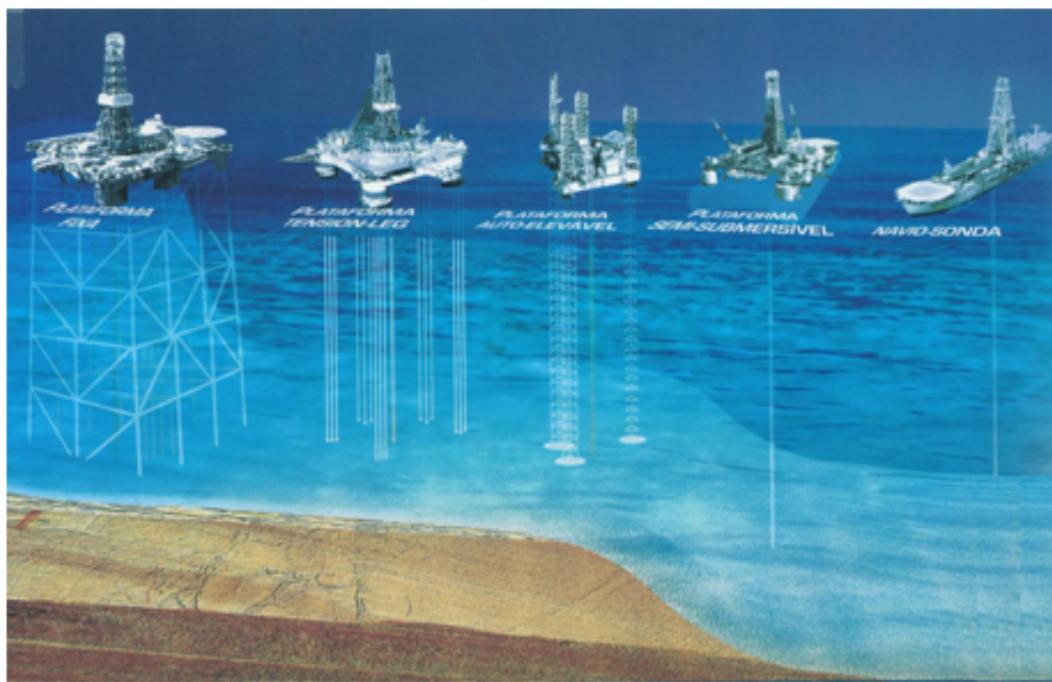
Para iniciar um poço de petróleo necessita-se selecionar um tipo de sonda ou plataforma compatível com as características gerais da região onde será perfurado o poço. Os serviços realizados nas áreas marítimas são denominados (Off Shore). Já os serviços realizados nas áreas terrestres são denominados (On Shore).

## 2.6 Avaliação de Formações

A avaliação de formações envolve atividades e estudos que visam definir em termos qualitativos e quantitativos o potencial de uma jazida petrolífera, isto é, sua capacidade produtiva e a valoração de suas reservas de óleo e gás [Thomas et al., 2001]. Essa avaliação é baseada principalmente nos dados coletados na perfilagem de um poço aberto, que segundo [Thomas



**Figura 2.8.** Sonda terrestre.



**Figura 2.9.** Sonda marítima.

et al., 2001] é uma imagem visual (Figura 2.10), em relação à profundidade, de uma ou mais características ou propriedades das rochas perfuradas tais como resistividade elétrica, po-

tencial eletroquímico natural, tempo de trânsito das ondas mecânicas, radioatividade natural ou induzida. Ou ainda, segundo An-nan & Lu [2009], perfilagem é o método de identificar propriedades das rochas e fluidos medindo informações elétricas, acústicas, termais e radioativas. Portanto, com a perfilagem é possível extrair dados importantes das formações pelas quais o poço atravessa como a litologia (tipo de rocha), espessura, porosidade, presença de óleo, gás ou água. Porém, a perfilagem sozinha não é suficiente para a análise. São necessárias informações auxiliares como as amostras de calhas que são resíduos das rochas trazidas à superfície e que ajudam na identificação das litofácies. Dependendo da análise dos perfis e das informações complementares, pode se decidir pelo abandono do poço ou no caso de se provar interesse econômico, os testes de formação para se estimar a capacidade produtiva do poço são iniciados.

### 2.6.1 Testemunhagem

Uma das técnicas utilizadas para extrair características do reservatório é a testemunhagem que é uma das mais antigas e ainda praticadas. Entretanto, testemunhar todos os poços em um campo muito grande pode ser economicamente inviável, além disso, o tempo consumido pode ser muito grande.

Segundo [Thomas et al., 2001], a testemunhagem é o processo de obtenção de uma amostra real de rocha de subsuperfície, chamado testemunho (ver Figura 2.11), com alterações mínimas nas propriedades naturais da rocha. Com a análise deste testemunho obtém-se informações importantes sobre geologia, engenharia de reservatórios, completação e perfuração, tais como litologia, textura, porosidade, permeabilidade, saturação de óleo e água.

O processo de obtenção de uma amostra de testemunho, como já foi dito, é caro e demorado. Geralmente ele é feito em poços chaves, como o poço pioneiro. Ele servirá também para confirmar os dados obtidos pela perfilagem elétrica. Os testemunhos são um volume cilíndrico de rocha que mantém os constituintes em sua estrutura original, como em profundidade, permitindo estudar os meios porosos onde estão armazenados os hidrocarbonetos. A partir deles ainda se tira amostras de lâminas para análise microscópica de granulometria onde se estuda a forma, tamanho, espaçamento dos grãos. Irá se verificar por exemplo a permeabilidade e porosidade da rocha.

Existem três métodos de retirada de testemunho:

- Testemunhagem com barrilete convencional: é o processo mais caro pois interrompe a perfuração. É necessário que se retire toda a coluna de perfuração e se coloque uma broca oca própria(barrilete) para criar o cilindro típico da rocha de testemunho. Depois sobe-se novamente essa broca já com a amostra e pode-se retornar com a broca de perfuração;

- Testemunhagem a cabo: Uma forma de se reduzir os custos da testemunhagem é substituir o barrilete convencional que é fixo a coluna de perfuração por um barrilete móvel preso a um cabo. Assim, o barrilete pode ser levantado até a superfície sem a necessidade de se retirar toda coluna;
- Amostra Lateral (Testemunhagem lateral): é utilizada para se testemunhar alguma formação já perfurada. É feita com cilindros ocos, presos a um cabo de aço e um canhão e são arremessados contra a parede do poço para retirada de amostras de rocha.

## 2.6.2 Amostras de Calha

A amostra de calha é utilizada juntamente com os perfis elétricos para ajudar na identificação das litofácies. Essa amostra contém os resíduos das formações rochosas, durante a perfuração, que permanecem na broca quando ela é levantada do interior do poço para a superfície.

## 2.6.3 Perfilagem Elétrica

Após a perfuração de uma fase do poço, são descidas várias ferramentas (Figura 2.6.3) com a finalidade de medir algumas características físicas que irão permitir a caracterização e a avaliação comercial[Thomas et al., 2001]. Essas características citadas por Thomas et al. [2001] são os perfis elétricos.

Com base na análise dos perfis é possível decidir quais intervalos de poços são de interesse econômico potencial para se executar os testes de formação. Se não houver intervalos de interesse, o poço é abandonado. Os testes de formação têm sido amplamente utilizados na indústria petrolífera para se estimar a capacidade produtiva do poço[Thomas et al., 2001]. Uma das mais importantes características que irá determinar o interesse de um intervalo é o tipo de rocha encontrado e suas características como porosidade e permeabilidade onde porosidade diz respeito a pequenos espaços dentro das rochas onde pode ser acumulado algum tipo de líquido ou gás. Permeabilidade é a capacidade que o conteúdo dos poros tem em se mover.

Na Figura 2.13 podemos ver alguns exemplos de rochas com boa porosidade e permeabilidade e também com porosidade ruim e permeabilidade ruim. Isso deve-se ao fato da existência de acumulações de petróleo depender das características e do arranjo de certos tipos de rochas sedimentares no subsolo. Conforme pode ser observado na Figura 2.14, basicamente, é preciso que existam rochas geradoras que contenham a matéria-prima que se transforma em petróleo e rochas reservatório, ou seja, aquelas que possuem espaços vazios, chamados poros, capazes de armazenar o petróleo. É necessário a ocorrência de rochas

selantes conhecidas como trapas, com baixa permeabilidade que não deixam o petróleo escapar. A ausência de qualquer um desses elementos impossibilita a existência de uma acumulação petrolífera.

Outro fator de influência é a localização de cada uma dessas rochas. Fica claro portanto, a importância da identificação das litofácies (tipos de rocha) na análise de um poço. É preciso ressaltar que a Figura 2.14 retrata de maneira simplista um reservatório de petróleo. No Brasil, por exemplo, conforme relatos colhidos por geólogos da Petrobras, dificilmente iremos encontrar rochas geradoras(folhelhos) perto de rochas reservatórios, como por exemplo arenito. Na maioria dos casos, as rochas geradoras estão a muitos quilômetros e gota a gota através de milhares de anos, foram se deslocando e se acumulando em um reservatório.

## 2.6.4 Fundamentos de Perfilagem Elétrica

A maioria dos perfis utiliza a resistividade como forma de avaliar o tipo sedimento ou fluido existente. A resistividade é simplesmente a capacidade, neste caso da rocha, em impedir ou deixar passar algum tipo de sinal. Então, podemos dizer que a intensidade de retorno desse sinal irá variar com o tipo rocha ou fluido existente. A porosidade da rocha poderá influenciar significativamente na variação do sinal uma vez que uma rocha mais porosa será capaz de conter mais água por exemplo, o que diminuiria a resistividade. Podemos dizer portanto, que a resistividade de uma rocha  $R_0$  varia na razão direta da resistividade da água e inversa da porosidade. A resistividade de uma rocha completamente saturada é dada por:

$$R_0 = \frac{aR_w}{\phi^m} \quad (2.1)$$

onde  $a$  é o coeficiente litológico,  $m$  é o coeficiente de cimentação e  $\phi$  o de porosidade. Lembrando que a cimentação como já falado, é a parede feita para revestir o poço impedindo que ele desmorona.

No caso da rocha conter qualquer tipo de fluido, como óleo, água ou gás, sua resistividade irá variar consideravelmente.

### 2.6.4.1 Tipos de perfis

Os principais tipos de perfis elétricos são: Raios Gama (GR), Neutrônico (NPHI), Indução (ILD), Sônico (DT), Densidade (RHOB), Caliper. A descrição de cada um deles é mostrada na Tabela 2.1.

Pela análise da Tabela 2.1, a principal propriedade encontrada a partir dos perfis é a porosidade de uma rocha definida pela relação entre o volume de espaços de uma rocha e pelo seu volume total. É justamente dentro desses espaços que poderemos encontrar óleo,

**Tabela 2.1.** Principais Tipos de perfis elétricos e suas propriedades.

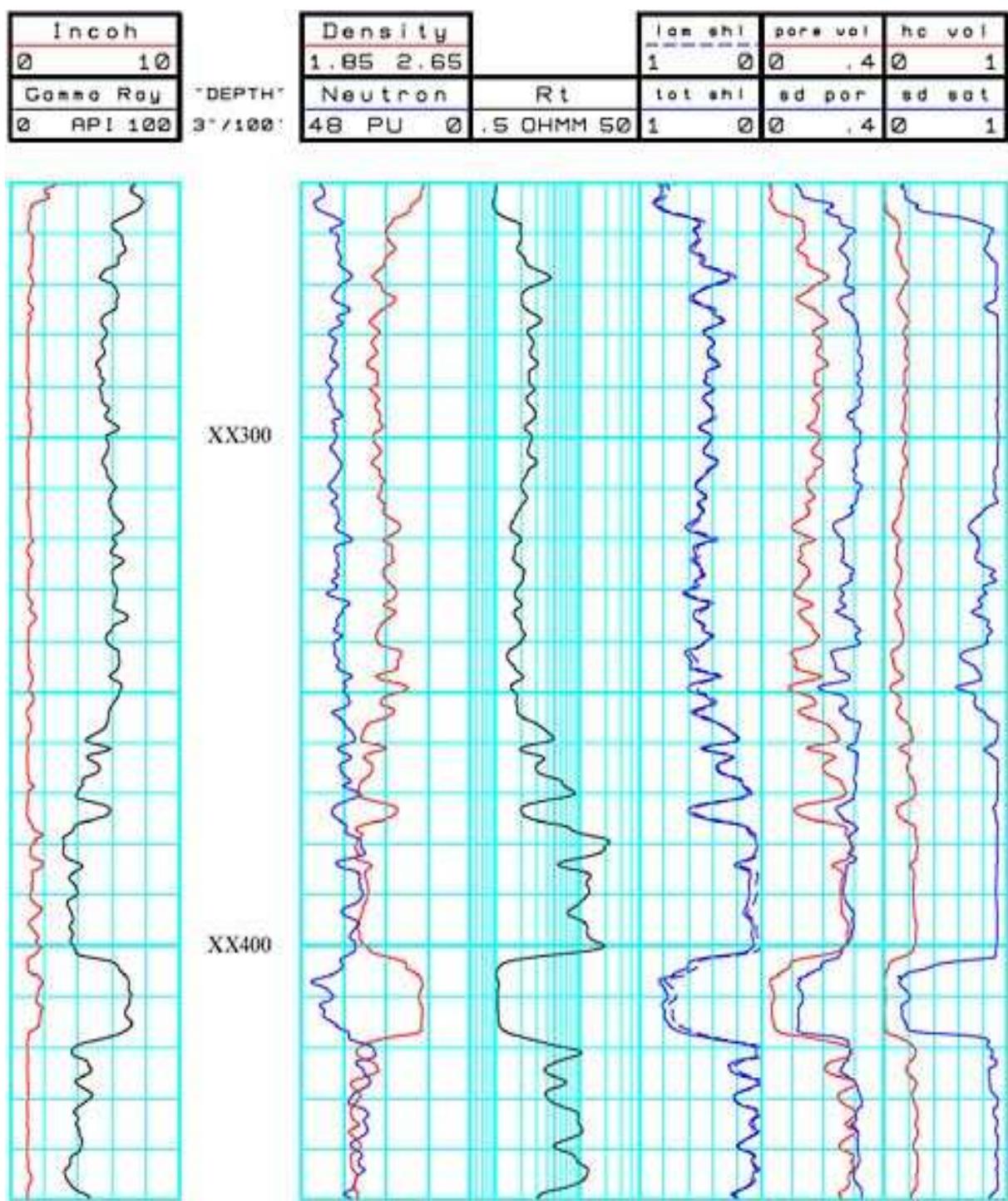
Perfil	Descrição
Raios Gama (GR)	Deteta a radioatividade total da formação geológica. Utilizado para a identificação da litologia, a identificação de minerais radioativos e para o cálculo do volume de argilas ou argilosidade;
Neutrônico (NPHI)	São utilizados para estimativas de porosidade, litologia e detecção de hidrocarbonetos leves ou gás;
Indução (ILD)	Fornece leitura aproximada da resistividade da rocha contendo hidrocarbonetos, através da medição de campos elétricos e magnéticos induzidos nas rochas;
Sônico (DT)	Mede a diferença nos tempos de trânsito de uma onda mecânica através das rochas - É utilizado para estimativas de porosidade, correlação poço a poço, estimativas do grau de compactação das rochas ou estimativas das constantes elásticas, detecção de fraturas e apoio à sísmica para a elaboração do sismograma sintético;
Densidade (RHOB)	Além da densidade das camadas, permite o cálculo da porosidade e a identificação das zonas de gás.
Caliper	Fornece o diâmetro do poço. É aplicado no cálculo do volume de cimento para tampões ou cimentação do revestimento, apoio a operações de teste de formação, controle de qualidade de perfis e indicações das condições do poço em um determinado intervalo
Potencial Espontâneo (SP)	É o registro da diferença de potencial entre um eletrodo móvel descido dentro do poço e outro fixo na superfície. Este perfil permite determinar as camadas permoporosas, calcular a argilosidade das rochas, determinar a resistividade da água da formação e auxiliar na correlação de informações com poços vizinhos.

gás ou água. Outra característica importante das rochas é sua permeabilidade definida pela medida da capacidade de circulação de um fluido através de uma rocha, sem alterar a sua estrutura interna. Portanto, para uma rocha ter uma boa permeabilidade, é necessário que seus poros não estejam isolados e nem sejam tão pequenos ao ponto de não permitir que estes fluidos circulem dentro da mesma. No entanto, a principal característica que pode ser identificada a partir dos perfis elétricos são as litofácies que atravessam o poço.

A análise dos perfis é sempre feita em conjunto para evitar erros de interpretação. Mudanças de valores significantes são sempre confirmadas em outro perfil. Por exemplo, a leitura de porosidade de uma rocha pelo perfil Sônico (DT) varia normalmente em torno de 9 a 27, sendo 27 uma rocha muito porosa com espaços internos onde poderia por exemplo existir óleo. Valores maiores que 27, poderiam indicar uma ruptura do poço e seria interes-

sante se checar o perfil Caliper. O Caliper é conhecido como perfil que mede a qualidade da perfilagem feita em um poço. Isto acontece porque ao se perfurar um poço, há a preocupação de que as paredes do mesmo não desabem. Para que isto não aconteça, a medida que o poço é perfurado, é injetado um fluido com densidade própria para que sustente a parede de um poço antes que se possa fazer a cimentação do mesmo. Caso a parede do poço desabe em algum ponto, irá criar uma zona de arrombamento comprometendo a leitura de outros perfis do poço. Perfis por exemplo que lêem a porosidade de uma rocha, ao realizar a leitura em uma zona onde existe um buraco provocado pela queda da parede do poço naquele local, irá indicar valores altos de porosidade, mas que na verdade indicam um buraco provocado por desabamento ou até mesmo uma caverna. O perfil Caliper irá confirmar essa informação.

A análise destes perfis é de responsabilidade do geofísico, que é o profissional que estuda as estruturas e os processos dinâmicos da Terra, além de analisar as forças físicas que atuam no planeta.



**Figura 2.10.** Perfis elétricos coletados durante a perfilagem de um poço.



**Figura 2.11.** Amostras de testemunho retirados de um poço. Apesar de ser a melhor representação do perfil de um poço, é um processo longo e caro, sendo feito portanto somente em alguns poços estratégicos.

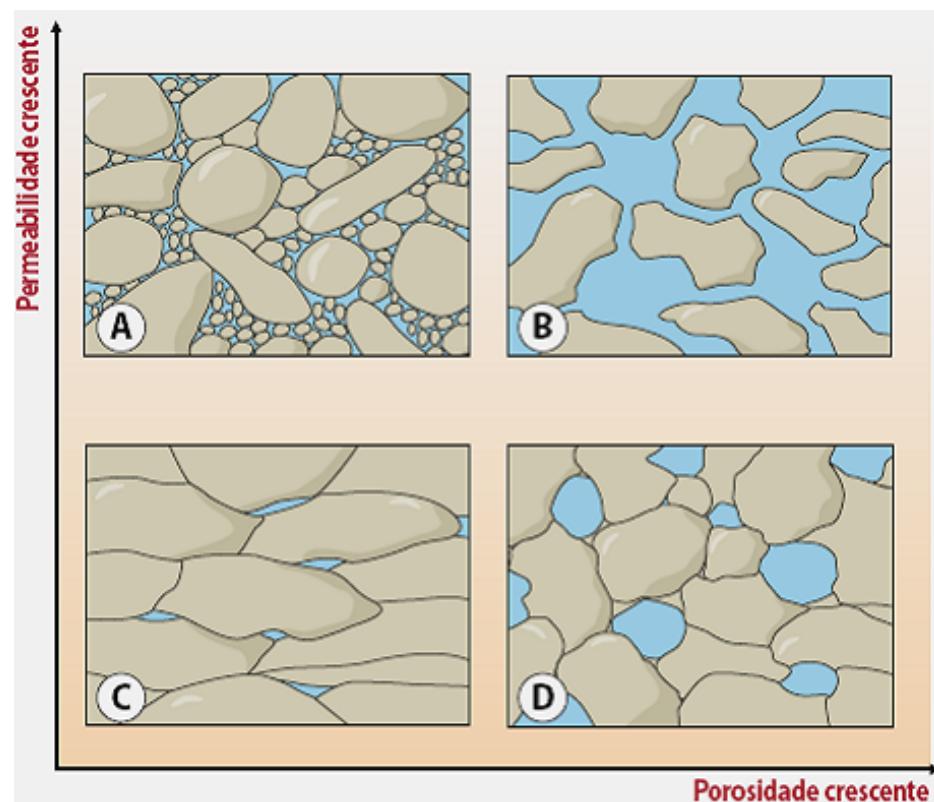


(a) Ferramentas de perfilagem sendo descidas no poço.

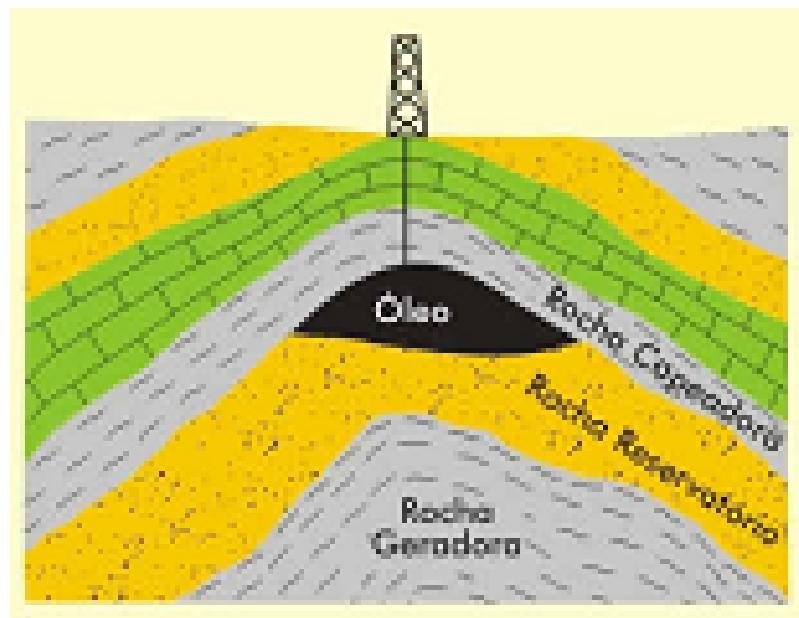


(b) Exemplos de ferramentas de perfilagem.

**Figura 2.12.** Ferramentas de perfilagem utilizadas para extração das características de um poço.



**Figura 2.13.** Porosidade e permeabilidade.



**Figura 2.14.** Acumulações de Petróleo.

# **Capítulo 3**

## **Fundamentação Teórica e Trabalhos Relacionados**

Conforme mencionado na introdução, neste trabalho serão utilizadas ferramentas de Reconhecimento de Padrões para a classificação automática de litofácies. Essas ferramentas são descritas em poucos detalhes na próxima seção. Em seguida, alguns trabalhos publicados na literatura relacionada ao objeto desta dissertação também são descritos e analisados.

### **3.1 Reconhecimento de Padrões**

O reconhecimento automático de padrões é normalmente dividido em três etapas [K.JAIN et al., 2000]: filtragem da entrada, extração/seleção de características e classificação. Na primeira etapa, dados irrelevantes, redundantes ou distorcidos, são eliminados. A segunda etapa, por sua vez, consiste na análise das características de entrada a fim de extrair e/ou selecionar características úteis para o processo de reconhecimento [WANG et al., 2003]. Por fim, na etapa de classificação, o objeto em análise é atribuído à uma das classes envolvidas no problema.

Nesta última etapa, métodos de classificação são investigados e comparados para que o melhor classificador para o problema seja identificado. São muitos os exemplos de classificadores que existem na literatura. Alguns dos mais utilizados são [K.JAIN et al., 2000]: classificadores estatísticos, redes neurais, árvores de decisão, SVM, kNN, etc.

Conforme mencionado anteriormente, cinco métodos de classificação são investigados neste trabalho: SVM, kNN, redes neurais do tipo Multylayer Perceptron, Árvores do tipo J48 Tree e Regressão Logística. Além de serem comparados individualmente, esses quatro classificadores são combinados em um conjunto de classificadores heterogêneos por meio de voto majoritário. Portanto, as próximas seções descrevem resumidamente o funcionamento

dos cinco métodos de classificação investigados, bem como o método de combinação de classificadores.

## 3.2 Máquinas de Vetores de Suporte - SVM

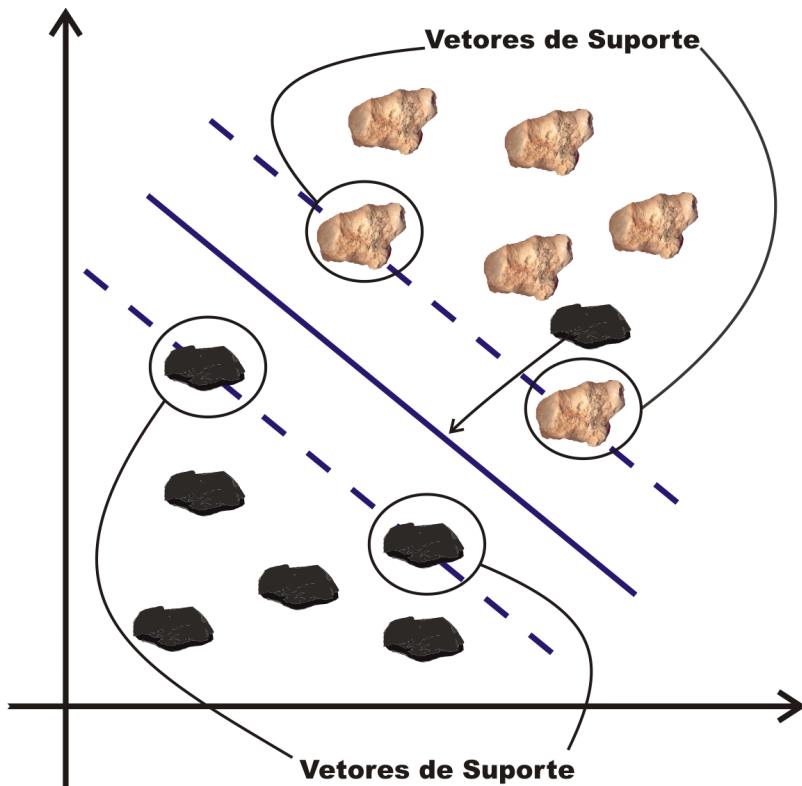
SVM é um tipo de classificador baseado na teoria da aprendizagem estatística proposto por Vapnik [1999] e outros pesquisadores. A teoria de Vapnik é baseada em uma forte fundamentação matemática para estimativa de dependências e prever o aprendizado a partir de conjuntos de dados finitos. O objetivo do SVM é minimizar tanto o erro de treinamento quanto a complexidade do modelo gerado durante o treinamento.

Basicamente, o funcionamento de SVM pode ser descrito da seguinte forma: dadas duas classes e um conjunto de instâncias de treinamento cujas amostras pertencem a essas classes, SVM constrói um hiperplano que divide o espaço de características em duas regiões, maximizando a margem de separação entre as mesmas. Esse hiperplano é conhecido como hiperplano de separação ótima. As amostras desconhecidas (exemplos de teste) são então mapeadas para esse mesmo espaço, e atribuídas a uma das classes [Alpaydim, 2010]. A Figura 3.1 ilustra o funcionamento de SVM. Dadas duas classes de rochas (calcaranito e folhelho, SVM define o hiperplano de separação ótima (reta não pontilhada) que maximiza a margem entre as duas classes (linhas pontilhadas). As amostras que encontram-se na fronteira (destacadas com círculos) são chamadas vetores de suporte e constituem os dados mais importantes para a tomada de decisão de SVM.

Aplicar o texto abaixo: Você precisa continuar... falta falar do valor de C e de funções kernel, pois você fala do ajuste desses parâmetros na próxima seção e o leitor precisa entender aqui, que parâmetros são esses. Para ajudar, veja o capítulo de livro do qual eu sou co-autora (documento resultante de um minicurso, que eu enviei a você), na página 16 há um texto sobre funções Kernel.

## 3.3 k-Vizinhos mais Próximos - kNN

A idéia de vizinho mais próximo (kNN) foi primeiro introduzida por Cover & Hart [1967], no qual a regra de decisão é designar uma amostra ainda não classificada para a mesma classificação do conjunto de pontos mais próximos já classificados. Os autores provaram que quando a quantidade de dados se aproxima do infinito, a classificação baseada nos vizinhos mais próximos é delimitada por duas vezes a taxa de erro assintótica como a regra de Bayes, independentemente da distância métrica aplicada.



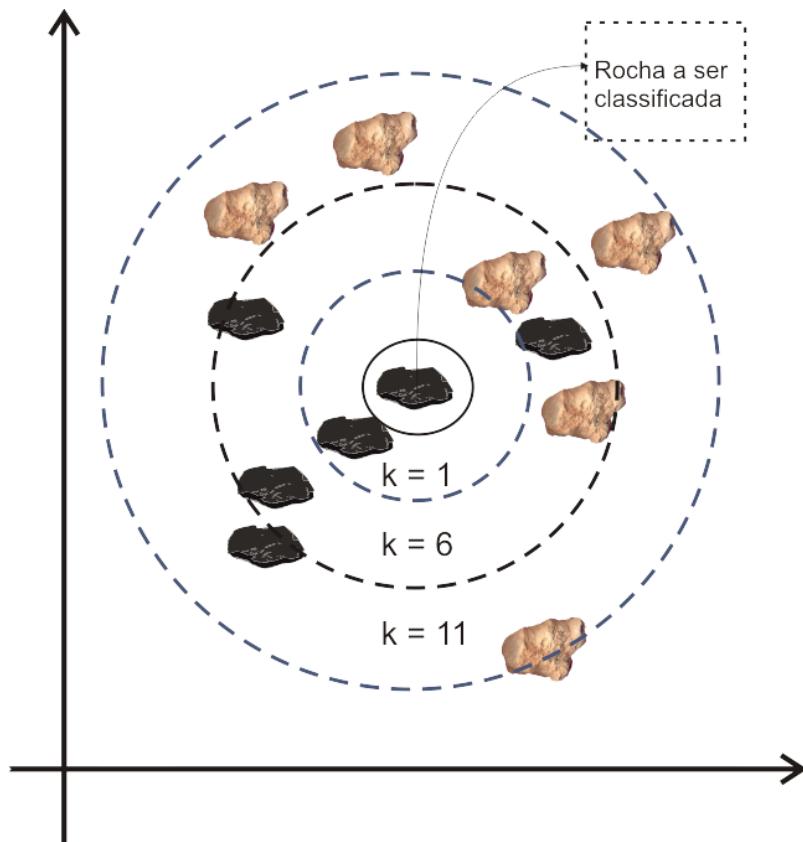
**Figura 3.1.** Ilustração sobre o funcionamento de SVM aplicado à classificação de rochas sedimentares.

Outra definição de kNN é dada pelos autores de Henke et al. [2011] onde KNN é uma técnica de classificação baseada em instâncias, isto é, consiste em atribuir a classe de cada elemento desconhecido (amostras de teste) a partir da classe majoritária obtida entre os seus vizinhos mais próximos identificados no conjunto de treinamento. A definição de vizinhança é feita segundo uma medida de similaridade que normalmente é uma medida de distância calculada no espaço de características. Uma das formas mais utilizadas para medir a distância é calculada pela distância a Euclidiana que é descrita pela seguinte fórmula:

$$d(x, y) = \sqrt{\sum_{i=1}^n (x_i - y_i)^2} \quad (3.1)$$

Uma maneira de ilustrar o método é proposto na Figura 3.2. A instância mais ao centro da figura é utilizada para demonstrar como novas instâncias são classificadas, trata-se de uma instância da classe folhelho. Na figura, quando  $k=1$ , a instância será classificada corretamente. Aumentando o valor de  $k$  para 6, novamente a instância será classificada corretamente, pois, dentre suas seis instâncias mais próximas, quatro instâncias são da classe folhelho e apenas duas são da classe anidrita, isto é, folhelho é a classe majoritária. Por fim,

se o valor atribuído a  $k$  for 11, a instância será classificada incorretamente, pois, a classe majoritária (seis votos) será a classe anidrita. Portanto, o classificador kNN precisa do ajuste do parâmetro  $k$ , que define a quantidade de vizinhos que contribuirão para a definição das classes das amostras desconhecidas.



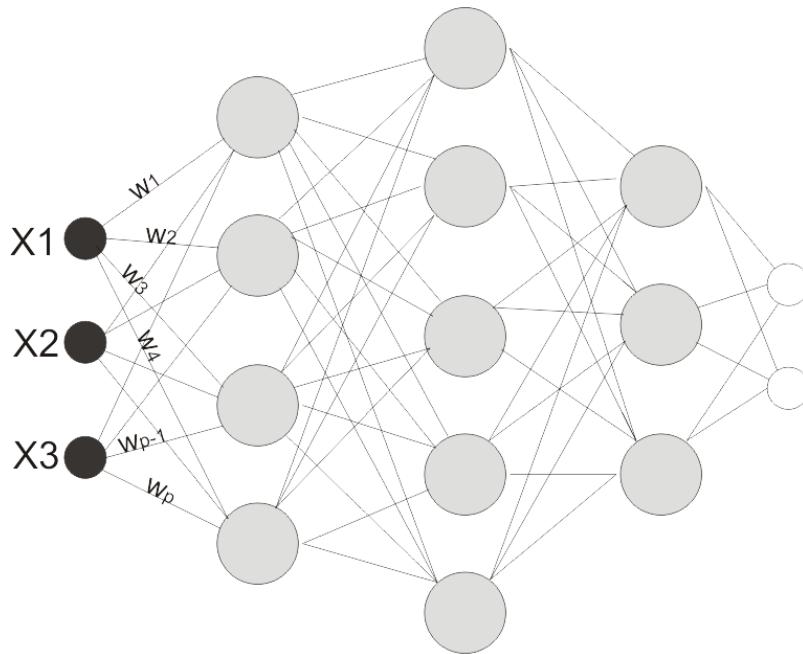
**Figura 3.2.** Exemplo ilustrativo do funcionamento de kNN, bem como o processo de ajuste do parâmetro  $k$ .

## 3.4 Multilayer Perceptron

*Multilayer Perceptron* é uma rede neural com as seguintes características principais:

- Múltiplas entradas, múltiplas saídas e múltiplas camadas;
- Entradas e saídas podem ser tanto analógicas quanto digitais;
- Não há restrição de separabilidade entre classes.

Redes Neurais Artificiais (RNA) são técnicas computacionais que propõem um modelo matemático que simula a rede de neurônios de um cérebro humano. Uma RNA procura por



**Figura 3.3.** Exemplo ilustrativo do funcionamento de Multilayer Perceptron.

relacionamentos, constrói modelos automaticamente e os corrige de modo a diminuir seu próprio erro [Tafner et al., 1995].

Segundo Cunha [2002], o neurônio artificial foi projetado para imitar as características de um neurônio biológico. O neurônio artificial (ver Figura 3.5) possui várias entradas ( $X_1, X_2, \dots, X_p$ ), que podem ser estímulos do sistema ou saídas de outros neurônios. Cada entrada é multiplicada por um peso correspondente ( $W_1, W_2, \dots, W_p$ ), gerando entradas ponderadas. Logo após, todas as entradas ponderadas são somadas e o valor resultante da soma será comparado com um valor limite para ativação do neurônio (função de ativação). Caso o valor da soma alcance o valor limite para ativação do neurônio, ele se ativará, caso contrário, ficará inativo. A saída no neurônio ( $Y$ ) é o valor processado pela função de ativação.

Dentre os principais parâmetros que devem ser ajustados durante o treinamento de uma RNA do tipo Multilayer Perceptron, destacam-se: quantidade de camadas escondidas, quantidade de neurônios em cada camada escondida, tipo de função de ativação, etc.

## 3.5 Regressão Logística

Regressão Logística em estatística é um tipo de análise de regressão usada para prever o resultado de uma variável categórica dependente, com base em uma ou mais variáveis de precisão. As probabilidades que descrevem o possível resultado de um único ensaio são modeladas como uma função de variáveis explanatórias, utilizando uma função logística.

## 3.6 Árvores de Decisão - J48

Outro classificador que testamos nesse trabalho foi o algoritmo J48 que permite a criação de modelos de decisão em árvore. Seu algorítmo utiliza uma tecnologia greedy para induzir árvores de decisão para posterior classificação. O modelo de árvore de decisão é construído pela análise dos dados de treino e o modelo utilizado para classificar dados ainda não classificados. O J48 gera árvores de decisão, em que cada nó da árvore avalia a existência ou significância de cada atributo individual. As árvores de decisão são construídas do topo para a base, através da escolha do atributo mais apropriado para cada situação. Uma vez escolhido o atributo, os dados de treino são divididos em sub-grupos, correspondendo aos diferentes valores dos atributos e o processo é repetido para cada sub-grupo até que uma grande parte dos atributos em cada sub-grupo pertençam a uma única classe. A indução por árvore de decisão é um algoritmo que habitualmente aprende um conjunto de regras com elevada acuidade.

## 3.7 Conjunto de Classificadores

O objetivo de combinar classificadores é aumentar a taxa de acerto em relação a classificadores utilizados individualmente. Até mesmo juntar um único tipo de classificador com parâmetros diferentes pode improver melhores taxas como descrito por Altinçay [2007] onde o autor demonstra que um conjunto de classificadores kNN obtém melhores taxas quando comparado ao mesmo classificador sozinho. Outra motivação para se utilizar conjuntos de classificadores é dada por dos Santos [2008]: conjunto de classificadores tentam superar um classificador robusto e bem parametrizado, por classificadores relativamente simples e combinados. A vantagem citada pela autora vai além de se obter melhores taxas de acerto. Simplificar a parametrização de classificadores complexos como o SVM é um grande ganho. Podemos citar, por exemplo, que os melhores parâmetros de um classificador, identificados em um reservatório, podem ser diferentes em outros reservatórios, e nem sempre geólogo especializados em perfis terão familiaridade com a parametrização de classificadores.

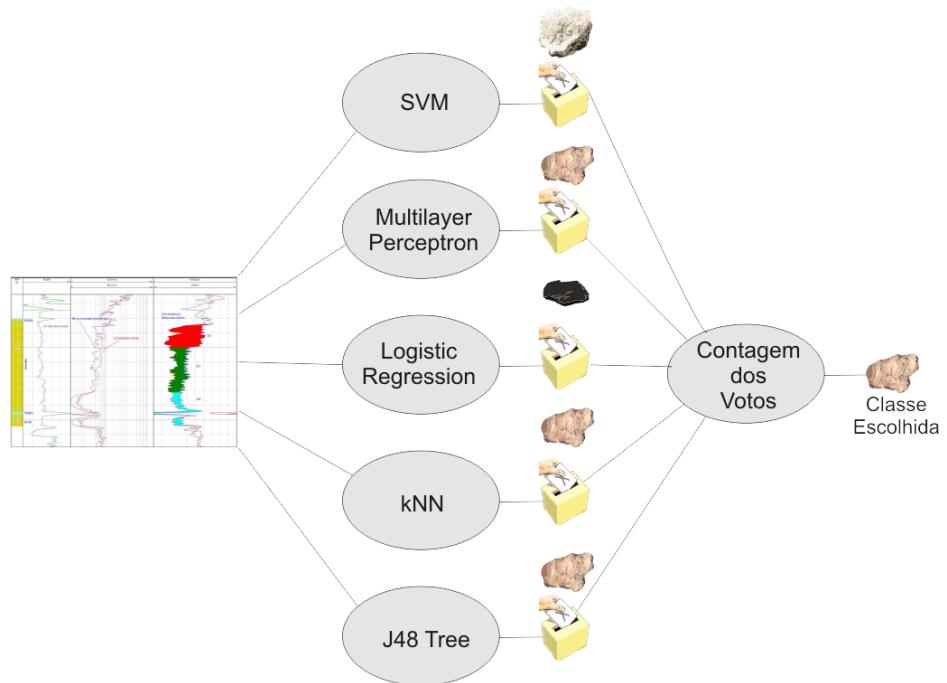
### 3.7.1 Votação Majoritária

Conjuntos de classificadores podem ser gerados de diversas formas, tais como: manipulando os dados de treinamento, os atributos dos dados, as classes do problemas e os membros do conjunto. Esta última estratégia possibilita a criação de conjuntos de classificadores heterogêneos, por meio do uso de diferentes métodos de classificação, diferentes arquiteturas ou parâmetros de um mesmo classificador. Neste trabalho, cinco tipos diferentes de métodos de

classificação são combinados em um conjunto de classificadores heterogêneos, por meio de voto majoritário, conforme descrito a seguir.

Existem várias estratégias para se agrupar classificadores em conjunto. A mais simples e mais popular é a da votação majoritária [dos Santos, 2008]. Nesse método, escolhe-se a classe com maior número de votos dentre os classificadores membros. Seu funcionamento é ilustrado na Figura 4.5. Os classificadores que compõem o conjunto investigado neste trabalho são:

1. SVM;
2. Multilayer Perceptron;
3. Logistic Regression;
4. kNN;
5. J48 Tree.



**Figura 3.4.** Ilustração dos cinco Classificadores em conjunto por meio do Voto majoritário.

A expectativa é que elevadas taxas de classificação sejam obtidas com a automatização do processo de classificação. Isso traria uma economia enorme de tempo e subsídios preciosos para justificar um projeto de exploração de um poço de petróleo ou ainda o cancelamento do mesmo.

Algumas tentativas de automotização desse processo têm sido publicadas na literatura. Na próxima seção são discutidos alguns desses trabalhos.

### 3.8 Trabalhos Relacionados

Cunha [2002] utiliza além dos perfis, testemunhos (Figura 2.11) de poços aplicando redes neurais para a classificação de litofácies. Em seu trabalho, Cunha utilizou uma base de 8 poços de petróleo e atingiu uma taxa de acerto de 80% e relata que uma solução para o problema só foi possível após a incorporação de uma estratégia para agrupamento prévio das litofácies e tratamento de padrões problemáticos (regiões de conhecimento incerto nos conjuntos de treinamento e de teste). A autora relata que a principal diferença de seu trabalho em relação a trabalhos prévios utilizando redes neurais está na maior quantidade de dados. No treinamento foi utilizado um simulador de Redes Neurais, o SNNS (*Stuttgart Neural Network Simulator*) e a arquitetura utilizada nos experimentos foi Multilayer Perceptron com algoritmo de treinamento Backpropagation. Nesse trabalho, apesar de terem sido utilizados dados de 08 poços, todos pertenciam à mesma reserva e, portanto, não há como afirmar qual o comportamento do seu classificador treinado quando aplicado a uma reserva diferente.

Flexa et al. [2004] propuseram uma metodologia baseada na técnica da análise discriminante para a identificação de litologias nos perfis de poços de duas reservas. Os autores pesquisaram a associação entre as variações das propriedades físicas e os tipos litológicos em profundidade. Os perfis geofísicos utilizados nesse trabalho foram raios gama (GR), sônico (DT), porosidade neutrônica (PHIN), densidade (RHOB) e resistividade (ILD). Segundo Flexa et al. [2004], a análise discriminante é uma técnica da estatística multivariada com grande eficiência na solução de problemas geofísicos. A técnica é utilizada nos casos em que se deseja realizar a separação estatística entre dois ou mais grupos de dados, que devem ser definidos previamente pela caracterização de suas variáveis discriminantes. Entretanto, no trabalho de Flexa et al. [2004], a técnica de análise discriminante não é comparada com outros classificadores. Os autores consideram que bons resultados foram obtidos, porém o desempenho real não foi apresentado.

An-nan & Lu [2009] propuseram um trabalho de identificação de litofácies a partir de dados de perfilagem utilizando DE-SVM com uma taxa de acerto de 90%. DE é o acrônimo de diferença evolutiva, um algoritmo baseado em algoritmos genéticos. Os autores observam que SVM tem sido usado com sucesso para identificação de litologias, mesmo com uma amostra pequena de dados, mas que no entanto, SVM não fornece um método de seleção para o melhor parâmetro de kernel  $\sigma$  e o fator de penalização  $C$ , que podem afetar a precisão de aprendizagem e previsão. O artigo apresentou um método de otimização baseada

na diferença aritmética-evolutiva (DE) para selecionar os parâmetros acima, chegando assim ao melhor kernel e parâmetro de penalização C para o problema de identificação de litologia. O melhor valor encontrado para C, foi de 263.2 e para o parâmetro de kernel  $\sigma$  foi 11.5. Porém, a base utilizada era composta por dados de três poços adjacentes, não sendo possível tirar conclusões sobre a eficiência do método para reservas diferentes. Além disso, como o objetivo principal dos autores foi desenvolver um método para encontrar os melhores parâmetros para SVM, nesse trabalho não houve comparações entre classificadores como kNN e Redes Neurais.

Wong et al. [2005] investigaram o desempenho do classificador SVM para a caracterização de reservatórios como permeabilidade e porosidade. Estas características são fundamentais para a localização de hidrocarbonetos, pois conforme já discutido, sem uma boa porosidade não há como a rocha conter fluidos e sem uma permeabilidade adequada, o fluido não tem como escoar entre os poros da rocha, podendo assim, ser retirado. Nos experimentos, o classificador SVM teve uma taxa de erro de 2.6% enquanto o classificador BPNN (*Backpropagation Neural Networks*) atingiu uma taxa de erro maior em 7.4%.

Al-Anazi & Gates [2010a] comentam sobre a demanda de novas técnicas de predição de litologia em reservatórios que sejam mais eficientes que os métodos estatísticos. Os autores citam a aplicação de técnicas como redes neurais e mais atualmente, a aplicação de SVM com bons resultados mesmo em reservatórios heterogêneos. Eles compararam então estes três métodos e revelam melhores resultados com SVM frente à análise estatística discriminante e redes neurais. Os experimentos foram realizados a partir de dois poços provenientes de reservatórios distintos. O autor relata que o desempenho do classificador SVM variou de acordo com o valor do parâmetro de penalidade C. O método de validação cruzada foi utilizado para encontrar o melhor parâmetro. Em um outro trabalho mais recente, Al-Anazi & Gates [2010b] aplicam SVM em um reservatório bastante heterogêneo com dados de três poços para classificar eletrofácies e predizer distribuição de permeabilidade e novamente compararam o desempenho de SVM com redes neurais e com a técnica de análise discriminante, e novamente confirmam a superioridade de SVM frente a estas duas outras técnicas.

Existem muitos trabalhos sobre classificação de litofácies utilizando o reconhecimento de padrões. Os mais recentes têm demonstrado a superioridade do classificador SVM frente a classificadores como redes neurais e análise estatística discriminante ([Al-Anazi & Gates, 2010a], [Wong et al., 2005]). Os trabalhos em sua maioria têm utilizado uma base muito pequena, quase sempre de somente um poço, impossibilitando conclusões sobre o comportamento dos classificadores no caso de serem utilizados em um novo poço. A exceção foram os trabalhos de Al-Anazi & Gates [2010a] e Flexa et al. [2004], sendo que neste último não ficou claro se os testes foram executados em bases separadas ou se as bases dos dois

reservatórios foram trabalhadas como uma só base. O autor também não revelou o real desempenho da sua classificação. Já no trabalho de Al-Anazi & Gates [2010a], não foram realizados experimentos para cada base separadamente e não foi verificado se o aprendizado de um classificador a partir de dados de uma reserva apresenta desempenho similar quando aplicado a uma outra reserva. Outro ponto observado é que nenhum trabalho comparou o desempenho do classificador kNN para o problema de classificação de litofácies a partir de dados de perfis elétricos.

O presente trabalho propõe responder as perguntas que foram deixadas pelos trabalhos anteriores utilizando uma base de dados de poços heterogêneos. A partir dessa base, serão comparados o desempenho dos classificadores SVM, kNN, Multilayer Perceptron e Regressão Logistica em diversos contextos como quando no mesmo poço e em poços diferentes e se o desempenho do aprendizado com dados do mesmo poço é maior do que o desempenho do aprendizado que utilize dados de poços diferentes em sua base de treinamento. É ainda estudado neste trabalho a combinação destes classificadores por meio de voto majoritário nos mesmos contextos abordados por classificadores individualmente. Na classificação por votação majoritária, escolhe-se a classe com maior número de votos dos classificadores. Os resultados obtidos com nossos experimentos são descritos no próximo capítulo.

# **Capítulo 4**

## **Experimentos e Resultados**

Este capítulo descreve os experimentos e os resultados obtidos. Inicialmente, porém, descrevemos os detalhes do protocolo experimental utilizado nos experimentos. Em seguida, os classificadores Multilayer Perceptron, SVM, kNN, Árvore (J48) e Logistic Regression são comparados quanto ao seu desempenho individual. Por fim, os classificadores são combinados em um sistema de votação majoritária, cujo o resultado será apresentado no final.

Todos os passos acima são repetidos em dois contextos diferentes. No primeiro, o treino e teste é realizado dentro do mesmo poço. No segundo contexto, que simula uma situação real de aplicação da classificação automática, uma base de um poço é utilizada para treinamento e depois o teste é feito em uma base distinta.

A base foi fornecida pela Petrobras e é constituída de 03 poços da mesma reserva e classificada em litofácies por um geólogo especialista. Serão então comparados os classificadores que apresentam melhor desempenho em um poço classificado por litofácies. Com a base de treinamento de um poço, iremos comparar o desempenho dos classificadores na classificação de um poço diferente, porém dentro da mesma reserva. Será respondido assim, se o classificador com melhor desempenho em um poço, generaliza seu conhecimento quando o aprendizado for medido em dados de um poço diferente da mesma reserva. Será necessário ainda comparar os mesmos resultados utilizando dados de mais de um poço.

Por fim, os classificadores são combinados em um sistema de votação majoritária, cujo o resultado será apresentado no final.

### **4.1 Protocolo Experimental**

Os experimentos foram realizados com uma base de dados composta por 03 poços, identificados por w6, w7 e w8, onde:

**Tabela 4.1.** Exemplo de amostras da base de dados

nº	Depth	Cal	DT	GR	ILD	MSFL	NPHI	Pef	RhOB
1	2435,0471	6,7227	56,2182	106,4375	27,6483	35,1158	10,8416	3,168	2,6629
2	2435,1995	6,7227	57,9455	146,5	25,7482	53,9805	12,8258	3,1602	2,6624
278	2477,262	6,7099	71,9797	75,8699	0,6001	1,6795	11,4636	2,9233	2,5086
283	2478,0239	6,582	68,1109	63,4688	5,0932	7,0145	13,3006	2,5742	2,4555

- w6 - constituído por 815 leituras de perfis;
- w7 - constituído por 808 leituras de perfis;
- w8 - constituído por 900 leituras de perfis;

Os perfis lidos para cada registro foram DT(Sônico), GR (Raios Gama), LLD e RHOB.

## 4.2 Ajuste de Parâmetros

Foi realizada uma etapa de ajuste de parâmetros, onde no caso do classificador SVM, o kernel com melhor desempenho foi o Puk. O parâmetro de regularização C teve no valor 110, a sua melhor taxa de sucesso encontrada em 89.312%. Para o classificador *Multilayer Perceptron*, o parâmetro *learning rate* foi ajustado para 0,17 e o parâmetro *momentum* foi ajustado para 0,1 atingindo sua melhor taxa de acerto em 87,715%. Para o classificador *Simple Logistic* não houve ajuste de parâmetros. Para essa fase de ajuste, foram usados os dados do poço w6 e os melhores parâmetros encontrados foram repetidos para os demais poços. Os dados estão divididos em quatro classes de litofácies: calcarenito, anidrita, folhelho e halita, e foram rotulados por um especialista geofísico. A Tabela 4.1 mostra um exemplo das amostras da base de dados.

## 4.3 Normalização da Base

Como pode ser observado na tabela 4.1, os valores obtidos para cada perfil estão em escalas diferentes. Essa não normalização dos dados pode influenciar negativamente o desempenho dos classificadores, especialmente kNN, que é baseado em medida de distância. Diante desse fato, um processo de normalização dos dados foi realizado. Aksoy & Haralick [2000] apresentam um estudo comparativo entre seis procedimentos de normalização que possibilitam que os dados fiquem dentro de um intervalo [0,1]. Neste trabalho foram utilizados os dois primeiros procedimentos investigados em Aksoy & Haralick [2000]. No primeiro procedimento de pré-processamento é feito um escalonamento linear normalizando os dados, sendo que é dado um limite inferior  $l$  e um limite superior  $u$  para uma característica  $x$ ,

$$\tilde{x} = \frac{x - l}{u - l}, \quad (4.1)$$

resultando em  $\tilde{x}$  dentro de um intervalo  $[0,1]$ .

No segundo procedimento de pré-processamento os dados são padronizados, consistindo em transformar uma característica  $x$  para uma variável aleatória com média zero e variância igual a 1 como:

$$\tilde{x} \cong \frac{\mu - l}{\sigma}, \quad (4.2)$$

onde  $\mu$  e  $\sigma^2$  são a média da amostra e a variância igual a 1 da característica  $x$ , respec-



(a) Folhelho



(b) Calcareito



(c) Anidrita



(d) Halita

**Figura 4.1.** Amostras de rochas

tivamente.

## 4.4 Desempenho Individual e Combinado com Treino e Teste no Mesmo Poço

No passo seguinte, testamos os classificadores individualmente para cada poço. O classificador Multilayer Perceptron atingiu a melhor taxa de acerto em 98.4029% para o poço w6. Para os demais poços, o classificador SVM obteve os melhores resultados sendo 87.3762% para o poço w7 e 85.8889% para o poço w8. Os piores classificadores foram o kNN e Regressão Logística que além de apresentarem as piores médias de taxas de acerto para todos os poço, atingiram os piores resultados individualmente para o poço w7 com taxas de acerto de 59.0346% para kNN e 58.5396% para o classificador de Regressão Logística. É interessante notar que mesmo com dois classificadores com taxas de acerto tão baixas, os valores para o sistema de votação majoritária atingiu ainda 78.9603%. Todos estes valores podem ser vistos em detalhe na tabela 4.2 e no gráfico representado Figura 4.2.

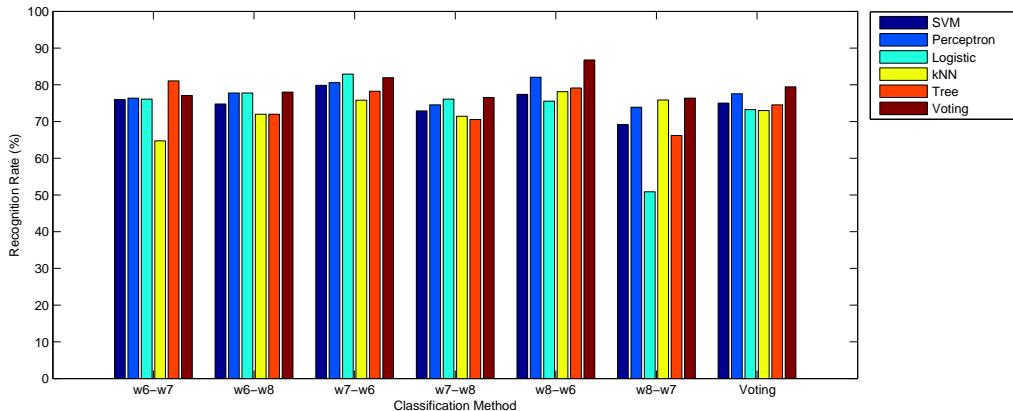
**Tabela 4.2.** Desempenho individual dos classificadores.

Well	SVM	kNN	Perceptron	Logistic	Tree	Voting
w6	89.68%	89.68%	<b>98.40%</b>	95.94%	93.61%	97.54%
w7	<b>87.38%</b>	59.03%	85.02%	78.96%	58.53%	78.96%
w8	<b>85.89%</b>	73.89%	82.67%	79.22%	83.78%	83.22%
Média	87.68%	74.20%	<b>88.70%</b>	87.71%	78.64%	87.17%

No sistema de votação majoritária, o conjunto de classificadores não apresentou melhor desempenho em nenhum dos poços, no entanto é importante observar que o conjunto de classificadores apresentaram taxas muito próximas dos melhores classificadores para os poços w6 e w8 e na média das classificações. No poço w7, os classificadores kNN e Tree prejudicaram o desempenho do sistema de votação, embora como já dito antes, sua taxa de acerto atingiu 78.9603% que representa 21.4207% a mais que o pior classificador.

## 4.5 Desempenho poço a poço dos classificadores.

Nesta sub-seção testamos o desempenho de cada classificador treinando em um poço e testando nos demais. O mesmo se repetiu para os classificadores combinados por votação majoritária. O desempenho dos mesmos pode ser visto na Tabela 4.3 e no Gráfico 4.3.



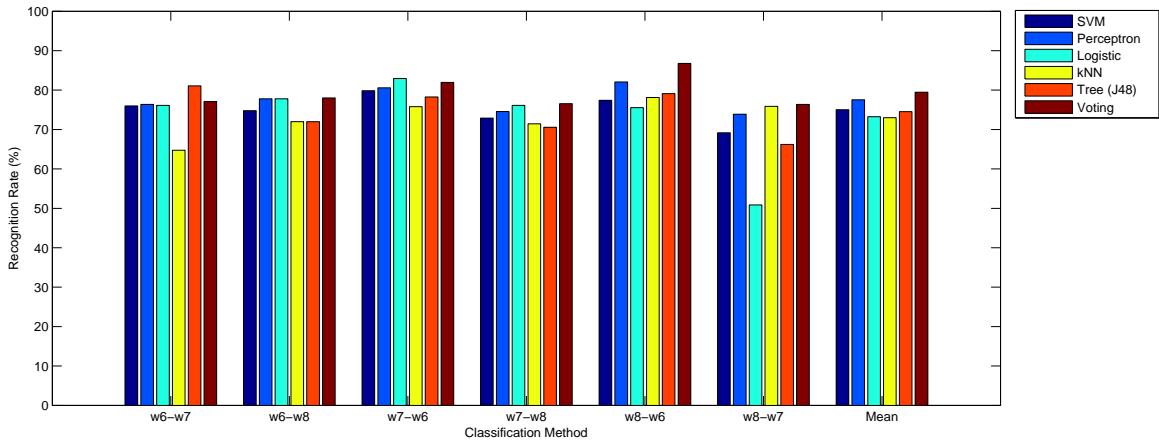
**Figura 4.2.** Gráfico do desempenho dos classificadores treinados e testados no mesmo poço.

**Tabela 4.3.** Desempenho de poço para poço dos classificadores.

Well	SVM	kNN	Perceptron	Logistic	Tree	Voting
w6-w7	75.99%	64.73%	76.36%	76.11%	81.06%	77.10%
w6-w8	74.78%	72.00%	77.78%	77.78%	72.00%	78.00%
w7-w6	79.85%	75.79%	80.59%	82.92%	78.26%	81.94%
w7-w8	72.89%	71.44%	74.56%	76.11%	70.56%	76.56%
w8-w6	77.39%	78.13%	82.06%	75.55%	79.11%	86.75%
w8-w7	69.18%	75.87%	73.89%	50.87%	66.21%	76.36%
Média	75.01%	72.99%	77.56%	73.24%	74.53%	79.45%

Diferente dos experimentos feitos onde treino e teste pertenciam ao mesmo poço, o sistema de votação majoritária apresentou taxas de acerto superiores aos demais classificadores em 4 dos 6 testes realizados. Nos demais, ele apresentou a segunda melhor taxa de acerto, além de na média ter sido também a melhor. É importante ressaltar que estes testes simulam o que acontece no mundo real da indústria petrolífera. Isso se explica porque em bacias diferentes, a leitura de valores similares dos perfis pode apresentar classificações distintas. Por isso é importante o conhecimento da reserva por um geofísico para que este profissional possa fazer uma primeira classificação de um poço dando um padrão que possa ser aprendido pelos classificadores. Assim, a automatização deste processo, precisaria sempre de um poço onde seria feito o primeiro estudo pelo geofísico. Este poço serviria assim para calibrar os classificadores e a partir deste, classificar poços vizinhos com características similares pouparindo o profissional de classificar todos os poços e limitando-se apenas a poços chaves.

Na próxima seção, iniciamos uma análise estatística destes dados.



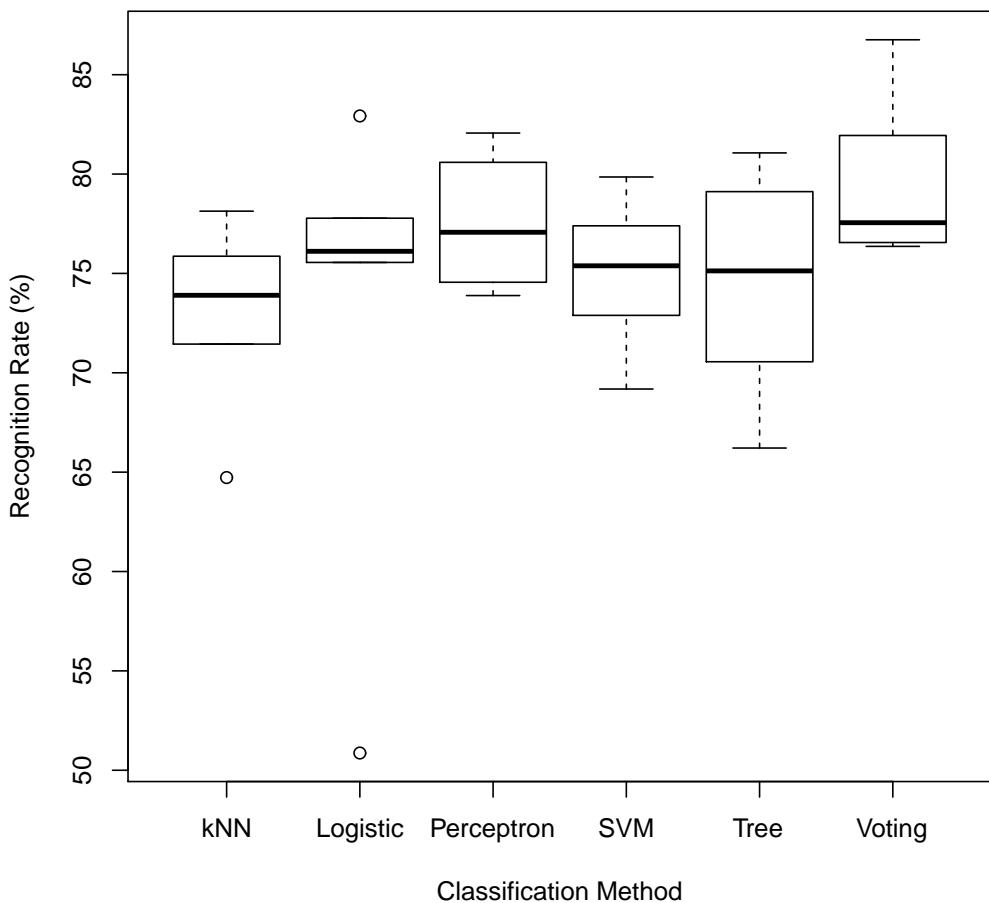
**Figura 4.3.** Gráfico do desempenho dos classificadores treinados em um poço e testado nos demais.

**Tabela 4.4.** Resumo das Estatísticas Descritivas

Class	Média	Mediana	D.Padrão	Mínimo	Máximo	1.Quartil	2.Quartil	Coef.Assim
KNN	72.99%	73.90%	4.78	64.73	78.13%	71.58%	75.85%	-0.0863
Logistic	73.22%	76.11%	11.28	50.87	82.92%	75.69%	77.36%	0.4940
Perceptron	77.54%	77.07%	3.27	73.89	82.06%	75.01%	79.89%	0.1556
SVM	75.01%	75.38%	3.70	69.18	79.85%	73.36%	77.04%	-0.0999
Tree	74.53%	75.13%	5.81	66.21	81.06%	70.92%	78.90%	-0.0545
Votação	79.45%	77.55%	4.13	76.36	86.76%	76.69%	80.96%	0.5962

## 4.6 Análise Estatística dos Resultados dos experimentos de Poço a Poço.

São apresentados na Tabela 4.4 algumas estatísticas como média, mediana, desvio padrão, valor mínimo e máximo, e coeficiente de assimetria. Estes valores são referentes as taxas de acerto obtidos por cada um dos classificadores. Verificamos nessa tabela que o classificador Logistic Regression apresentou o menor índice de acertos em 50.87% caracterizado como um outlier no Gráfico 4.4, porém possui a menor variabilidade em relação aos valores centrais localizados entre o primeiro e terceiro quartil. O classificador Logistic Regression possui também o maior desvio padrão demonstrando uma grande variação da taxa de acerto. Os classificadores combinados em conjunto por meio de um sistema de votação majoritária apresentou mediana superior aos demais, mas similar ao classificador Perceptron, com 50% dos dados acima de 77%, indicando maior índice de acertos atingindo 86.76%, e seus valores centrais localizados entre 76% e 81%. O classificador kNN possui a menor mediana dentre os classificadores e também apresenta valores extremos. O classificador Perceptron possui valores centrais entre 75% e 79%, indicando margem de acertos um pouco inferior aos classificadores combinados por meio de votação. O classificador Tree apresentou uma



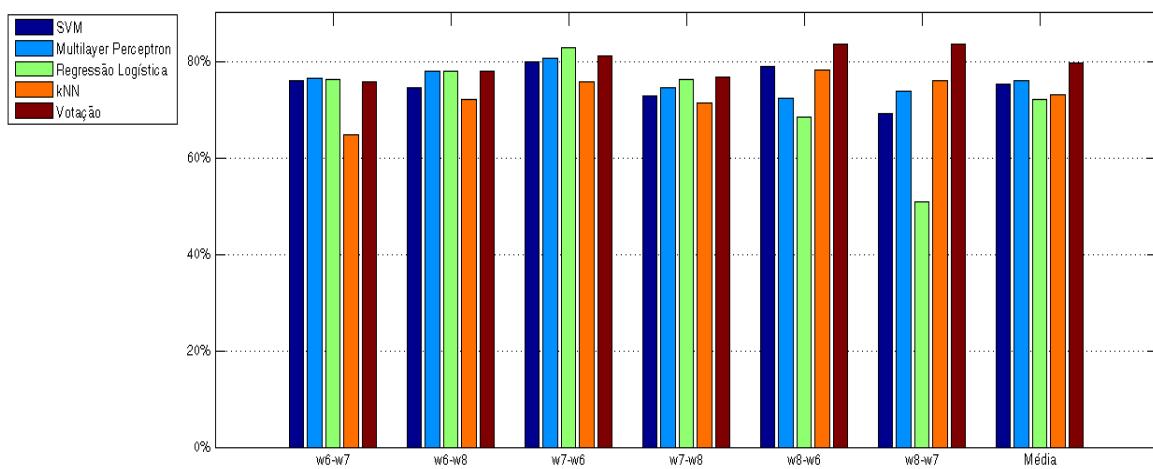
**Figura 4.4.** Box-plot para índice de acertos dos classificadores

maior variabilidade em relação aos valores centrais. Quanto à assimetria, os classificadores Logistic, Perceptron e Votação apresentaram assimetria positiva, isto é, cauda longa à direita. Os demais classificadores apresentaram assimetria negativa.

Segundo o teste de Friedman ?, teste estatístico não paramétrico que faz uma comparação entre os métodos(classificadores) e testa a hipótese nula ( $H_0$ ), de que não há diferença entre os métodos, contra a hipótese alternativa ( $H_1$ ) de que pelo menos um dos métodos difere dos demais, observamos que o sistema de Votação e o classificador Perceptron não apresentam diferença significativa, no entanto, Votação difere dos demais. O p-valor com resultado igual a 0.0053 sugere que devemos rejeitar a hipótese de igualdade entre os métodos.

## 4.7 Resultados

A combinação de classificadores por voto majoritário, se mostrou inferior a classificadores individuais quando a base de treinamento e de testes pertencem ao mesmo poço. Porém, quando usamos uma base de treinamento de um poço para classificar dados de poços distintos, a combinação de classificadores apresentou desempenho superior a todos os classificadores individuais para os testes entre os poços w8-w6 e w8-w7. Na média dos testes realizados entre todos os poços, a combinação de classificadores mais uma vez apresentou superioridade.



**Figura 4.5.** Desempenho Individual dos Classificadores e Combinados em Conjunto Utilizando Votação Majoritária.

# **Capítulo 5**

## **Conclusões**

Neste trabalho foram comparados os classificadores Multilayer Perceptron, SVM, kNN, Árvore (J48) e Logistic Regression para o problema de classificação de litofácies por meio de perfis elétricos. Depois de comparados individualmente, os classificadores foram combinados em um sistema de votação majoritária. Simulamos duas situações hipotéticas. Na primeira, treinamos e testamos utilizando dados no mesmo poço. Para este cenário, os classificadores Perceptron e SVM apresentaram o melhor desempenho. O sistema de votos majoritário não se mostrou como melhor opção, porém foi mais estável que classificadores como Tree J48 e kNN. No segundo cenário, testamos os classificadores em um ambiente que reflete melhor a realidade, onde usamos um poço para o treinamento e o teste foi realizado em um poço distinto. Nesse ambiente, o conjunto de classificadores utilizando voto majoritário, se mostrou a melhor opção entre 4 de 6 testes e nos dois testes restantes, teve o segundo melhor desempenho. Nos testes estatísticos, confirmamos seu melhor desempenho quando analisando todos os dados provando que classificadores combinados incrementam as taxas de sucesso além de dar mais estabilidade à classificação.



# Referências Bibliográficas

- Aksoy, S. & Haralick, R. M. (2000). Effects of feature normalization on image retrieval. *IAPR International Conference on Pattern Recognition*, 22:563 – 582.
- Al-Anazi, A. & Gates, I. (2010a). On the capability of support vector machines to classify lithology from well logs. *Natural Resources Research*, 19:125 – 139.
- Al-Anazi, A. & Gates, I. (2010b). A support vector machine algorithm to classify lithofacies and model permeability in heterogeneous reservoirs. *Engineering Geology*, 114(3–4):267 – 277.
- Altinçay, H. (2007). Ensembling evidential k-nearest neighbor classifiers through multi-modal perturbation.
- An-nan, J. & Lu, J. (2009). Studying the lithology identification method from well logs based on de-svm. Em *Control and Decision Conference, 2009. CCDC '09. Chinese*, pp. 2314 –2318.
- Cover, T. & Hart, P. (1967). Nearest neighbor pattern classification. *IEEE Transactions on Information Theory*, 13(1):21 –27.
- Cunha, E. S. (2002). Identificação de litofácies de poços de petróleo utilizando um método baseado em redes neurais artificiais. Dissertação de mestrado, Universidade Federal de Campina Grande.
- dos Santos, E. M. (2008). Static and dynamic overproduction and selection of classifier ensembles with genetic algorithms.
- Flexa, R. T.; Andrade, A. & Carrasquilla, A. (2004). Identificação de litotipos nos perfis de poço do campo de namorado (bacia de campos, brasil) e do lago de maracaibo (venezuela) aplicando estatística multivariada. *Revista Brasileira de Geociências*, 34:571–578.
- Garcia, E. N. (2008). *A PETROBRAS NA AMAZÔNIA: a história da pesquisa de combustíveis fósseis*. Norma Editora, Manaus, 1 edição.

- Henke, M.; Santos, C.; Nunan, E.; Feitosa, E.; dos Santos, E. & Souto., E. (2011). *Aprendizagem de Máquina para Segurança em Redes de Computadores: Métodos e Aplicações*.
- K.JAIN, A.; Duin, R. P. W. & Mao, J. (2000). Statistical pattern recognition. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 22:4–37.
- Petrobras (2011). Conheça nosso plano de negócios 2011-2015.
- Thomas, J. E.; Triggia, A. A. & Correia, C. A. (2001). *Fundamentos de Engenharia de Petróleo*. Interciênciac.
- Vapnik, V. N. (1999). *The Nature of Statistical Learning Theory (First Edition)*. Springer, 2 edição.
- WANG; Xuechuan; PALIWAL & K., K. (2003). Feature extraction and dimensionality reduction algorithms and their applications in vowel recognition. *Pattern Recognition*, 33:2429–2439.
- Wong, K.; Ong, Y.; Gedeon, T. & Fung, C. (2005). Reservoir characterization using support vector machines. Em *Computational Intelligence for Modelling, Control and Automation, 2005 and International Conference on Intelligent Agents, Web Technologies and Internet Commerce, International Conference on*, volume 2, pp. 354 –359.