

**APENDIZADO DE MÁQUINA NO RECONHECIMENTO DE PADRÕES
LITOLÓGICOS**

Victor Ribeiro Carreira

Relatório apresentado ao Programa de Pós-graduação em Geofísica do Observatório Nacional, como parte dos requisitos necessários à obtenção do título de Doutor em Geofísica.

Orientador(a): Dr. Cosme F. Neto, Ponte

Rio de Janeiro
Setembro de 2017

Sumário

Lista de Figuras	2
Lista de Tabelas	3
1 Introdução	1
1.1 Redes Neuronais Artificiais	2
1.2 Redes com aprendizado não-supervisionado	4
1.2.1 A Rede de Kohonen	5
2 Contexto Geológico	7
3 Método Proposto e Objetivo	9
3.1 Objetivo	10
4 Dados de Perfilagem	11
4.1 Modelo proposto para gerar os dados sintéticos	14
4.2 Dado Real	17
5 Resultados e Discussões	19
5.1 Treinamento	21
5.2 Identificação	24
6 Conclusões	25
Referências Bibliográficas	26
Cronograma de Atividades	31

Listas de Figuras

1.1	Modelo esquemático de um neurônio de McCulloch-Pitts. Onde x_1, x_2, \dots, x_n são os <i>inputs</i> , w_1, w_2, \dots, w_n são os pesos, h é o treino, $g(x)$ é a função de ativação, e y é o <i>output</i>	3
1.2	Homúnculo de Penfield.	4
1.3	Neurônio e suas vizinhanças	6
2.1	Mapa geológico e de localização da área de estudo.	8
4.1	Exemplo de um dado público de uma perfilagem de poço composta realizada pela Petrobras, na Bacia do Paraná.	12
4.2	Modelo Simplificado baseado em MOHRIAK <i>et al.</i> (2008).	15
4.3	Dado de perfilagem sintético, T1. Aonde a porcentagem de CE indica a mistura de conglomerado com embasamento	16
4.4	Dado de perfilagem sintético, C1.	16
4.5	Dado de perfilagem sintético, C2.	17
4.6	Localização dos poços de trabalho.	18
5.1	Agrupamento de dados do poço T1.	19
5.2	Agrupamento de dados do poço C1.	20
5.3	Agrupamento de dados do poço C2.	21
5.4	Mapas auto-organizáveis e sua evolução temporal.	22
5.5	Teste de convergência da rede.	23
5.6	Dado de saída da rede para o poço de classificação C1.	24

Lista de Tabelas

4.1	Compilação de Perfis usados na inferência de litologia.	13
4.2	Compilação de Perfis usados na inferência de porosidade, permeabilidade.	14
1	Cronograma das atividades previstas para o primeiro biênio	32
2	Cronograma das atividades previstas para o segundo biênio	33

Capítulo 1

Introdução

O ser humano vem usando a sua habilidade de reconhecimento de padrões desde muito antes do início do processo civilizatório. Grupos de humanos paleolíticos já faziam registro dos padrões migratórios de certos grupos de cervídeos. Durante a aurora da revolução neolítica, nossa capacidade de reconhecimento de padrões foi direcionada para a agricultura com a criação de monumentos que registraram a mudança das estações ao longo do ano.

O cérebro humano evoluiu espantosamente. E no que se refere a quantidade de informação processada, o cérebro possui enorme vantagem em relação a quantidade de informação processada por um computador (HALL *et al.*, 2014). Este não para de funcionar somente porque algumas células morrem. Um computador, por sua vez, não funciona quando há degradação da sua unidade central de processamento (MAO, 1996).

O campo do aprendizado de máquina aborda a criação de programas computacionais que automaticamente melhorem a si mesmos através da experiência (LEVY, 1997; MICHIE *et al.*, 1994). Tanto a rede neuronal quanto a árvore de decisão despontam como estratégias de solução para a resolução de problemas de reconhecimento de padrões (MACKAY, 2005).

Uma rede neuronal artificial possui semelhanças com a rede neuronal natural, nesta o cômputo do cérebro é feito através de uma vasta quantidade de neurônios interconectados (FELDMAN *et al.*, 1988; POULTON, 2002). A comunicação entre essas células é realizada através de impulsos elétricos. Tais impulsos elétricos são transmitidos e recebidos por meio de sinapses nervosas entre axônios e dendritos. As sinapses são estruturas elementares e uma unidade funcional localizada entre dois neurônios (KROGH, 2008).

As Redes Neurais Artificiais são inspiradas em modelos sensoriais de processamento de tarefas realizadas pelo cérebro (HAGAN *et al.*, 1996). Uma RNA¹, por-

¹RNA: Rede Neural Artificial

tanto, pode ser criada através da aplicação de algoritmos matemáticos que imitem a tarefa realizada por um neurônio (NEDJAH *et al.*, 2016).

Já a árvore de decisão auxilia na predição da classe de um objeto em um estudo com base em um treinamento prévio. Ou seja, funciona como um algoritmo de aprendizado de máquina supervisionado que é basicamente aplicado em problemas de classificação (FREUND e MASON, 1999). Funciona tanto para variáveis categóricas quanto para variáveis dependentes. Nesse algoritmo, a população original é dividida em dois ou mais grupos de populações homogêneas (SIMARD *et al.*, 2000).

1.1 Redes Neuronais Artificiais

MCCULLOCH e PITTS (1943) redigem o trabalho pioneiro onde foi modelado um neurônio cuja resposta dependia do *input*² que provinha de outros neurônios e do peso utilizado. Já ROSENBLATT (1962) cria a teoria de convergência do *Perceptron* onde ele prova que modelos de neurônios possuem propriedades similares ao cérebro humano (KANAL, 2001). Neste sentido as redes neurais artificiais podem realizar performances sofisticadas no reconhecimento de padrões, mesmo se alguns neurônios forem destruídos (LEVY, 1997). MINSKY e PAPERT (1969) demonstraram que *Perceptrons* somente resolvem uma classe muito limitada de problemas que podem ser linearizados.

Os primeiros artigos sobre redes neurais em geofísica datam de 1989 e são focalizados basicamente em o que a RNA pode fazer com dados de natureza diferente. E como preparar o dado para inserir na rede para depois interpretá-lo. As redes neurais artificiais foram usualmente treinadas com dados sintéticos e depois testados em dados reais. Contudo, hoje é comum usar dados reais para treinar a rede (ADIBIFARD *et al.*, 2014). Embora, ambas as abordagens sejam aceitas. O foco a partir de 1995 até o presente relaciona-se a algumas aplicações específicas, tais como caracterização de reservatórios. E a ênfase tem sido dada mais na integração de dados associado a uma interpretação compreensiva, ao contrário de uma aplicação isolada (POULTON, 2002).

No problema específicos de poços, um passo importante é a identificação de topo e base de camadas que podem ser associadas com mudanças das propriedades petrofísicas (SALJOOGHI e HEZARKHANI, 2014). Algoritmos baseados em derivadas nas curvas de log não identificam camadas muito finas, ou ruído (ZHANG *et al.*, 1999). CHAKRAVARTHY *et al.* (1999) consegue através do uso da função radial localizar os limites de camadas em alta definição em dados de log de indução (HDIL). Já BENAOUEDA *et al.* (1999) consegue classificar tipos litológicos em poços

²Valor de entrada

parcialmente desmoronados através do uso da rede neuronal com propagação de erro e mudanças de classes a medida que prossegue a análise.

O neurônio de MCCULLOCH e PITTS (1943) propõe um limite binário para a criação de um modelo. Este neurônio artificial registra uma soma de pesos de n sinais de entrada, x_j , $j = 1, 2, 3, \dots, n$, e fornece um *output*³ de 1 caso esta soma esteja acima do limite u . Caso contrário o *output* é 0. Matematicamente essa relação pode ser descrita de acordo com a Eq. 1.1:

$$y = \theta \left(\sum_{j=1}^n w_j x_j - u \right) \quad (1.1)$$

Onde θ é o passo dado na posição 0, w_j é chamada sinapse-peso associado a um *j*esimo *input*. A título de simplificação a função limite⁴ u é considerada um outro peso $w_0 = -u$ anexado a um neurônio com um *input* constante $x_0 = 1$. Pesos positivos correspondem a uma sinapse **excitatória**, enquanto pesos negativos correspondem a uma sinapse **inibitória**. Este modelo contém uma série de simplificações que não refletem o verdadeiro comportamento dos neurônios biológicos (MAO, 1996).

Derivações do neurônio de MCCULLOCH e PITTS (1943) na escolha das funções de ativação. Uma função largamente utilizada é a função sigmóide, que exibe uma suavização dos *outputs* a medida que o valor da função diminui (MAO, 1996; MISRA e SAHA, 2010). Essa função de ativação pode ser expressa de acordo com a Eq. 1.2:

$$g(x) = 1/(1 + e^{-\beta x}) \quad (1.2)$$

Onde β é o parâmetro de inclinação. A Fig. 1.1 ilustra a sequência lógica da operação de uma RNA para um neurônio simples de McCulloch-Pitts.

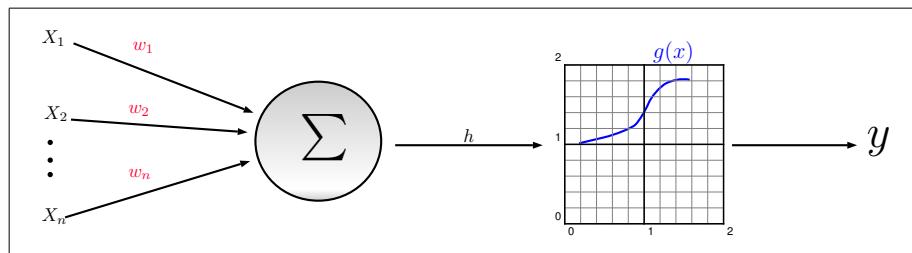


Figura 1.1: Modelo esquemático de um neurônio de McCulloch-Pitts. Onde x_1, x_2, \dots, x_n são os *inputs*, w_1, w_2, \dots, w_n são os pesos, h é o treino, $g(x)$ é a função de ativação, e y é o *output*.

³Valor de saída

⁴Genericamente chamada de função de ativação

Mais de 50 tipos de redes neurais artificiais tem sido criadas até o ano de 2014 (SALJOOGHI e HEZARKHANI, 2014).

1.2 Redes com aprendizado não-supervisionado

Nesta categoria de RNA's são apenas inseridos os valores de atributos de entrada. Os valores de saída são definidos pela própria rede que passa por um processo de treinamento não supervisionado. As redes que são submetidas a este tipo de treinamento são mais indicadas para tarefas aonde são exigidos agrupamento de dados (*Clustering*). Neste processo uma classe deve ser atribuída aos registros da rede observando-se apenas o comportamento de seus atributos, no caso em particular deste trabalho tratam-se de propriedades geofísicas.

Uma rede com treinamento não supervisionado inspira-se no funcionamento do córtex cerebral. Neste modelo biológico, o organismo aprende a realizar alguma tarefa, por meio da identificação de padrões. Por exemplo, ao identificar uma música determinados padrões sonoros que compõe o conjunto harmonioso de notas precisam ser aprendidos antes de serem reconhecidos. Durante este processo, regiões específicas do cérebro vão sendo paulatinamente acionadas. Isto somente é possível, devido conexões específicas que são formadas entre os neurônios presentes no córtex, Fig. 1.2.

Os detalhes dos processos que regulam o córtex ainda não foram totalmente elucidados, contudo é seguro assumir que a primeira representação dos fenômenos que aprendemos, no mundo, pode ser representado por uma superfície topológica.

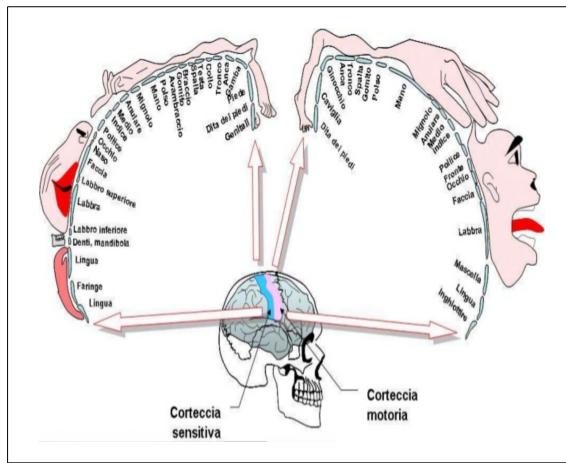


Figura 1.2: Homúnculo de Penfield.

Um cérebro que sofreu uma comossão grave perde a capacidade de acessar determinadas zonas do homúnulo responsáveis por atividades específicas. Contudo o cérebro tem a capacidade de destinar outras regiões para o controle destas ações que foram previamente perdidas.

Além de casos graves como um acidente o cérebro também perde a capacidade de aprendizado com o tempo. Em humanos, a capacidade de aprendizado vai da pequena infância até a puberdade. Após este período, o cérebro passa a reter o que fora aprendido. Sendo assim o aprendizado é uma função que depende, entre outras coisas, do tempo.

1.2.1 A Rede de Kohonen

A localização espacial de um neurônio da saída em um mapa topográfico corresponde a um domínio ou característica particular do dado retirado do espaço de entrada. E estas entradas são mapeadas de forma ordenada, a exemplo dos mapas cito-arquetais do córtex cerebral.

Neste processo de identificação de padrões a redundância torna-se impreverível, pois o neurônio da camada de saída que apresentar a maior resposta terá os seus pesos ajustados. Além disso, o peso dos neurônios vizinhos também serão ajustados em menor intensidade ao comparados com o neurônio vencedor.

Isto implica que os neurônios devem estar posicionados em um arranjo geométrico adequado. Esta teoria é baseada na suposição de que as células nervosas corticais estão organizadas anatomicamente em relação aos estímulos que recebem dos sensores aos quais estão ligadas (ARTERO, 2008).

Este modelo exige a definição de vizinhança entre neurônios de forma geométrica. Alguns arranjos são comumente utilizados, como por exemplo, os arranjos triangulares, hexagonal, retangulares, etc.

No caso de arranjos retangulares, diferentes vizinhanças de um neurônio $N_{i,j}$ podem ser configuradas em quartetos, diagonais e octetos.

A Fig. 1.3 ilustra o arranjo retangular e as vizinhanças, em quartetos, adotado neste trabalho.

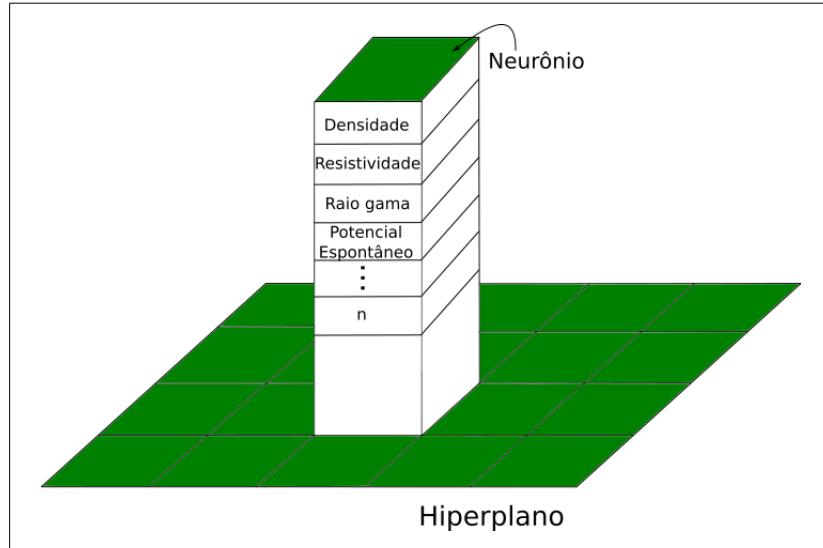


Figura 1.3: Neurônio e suas vizinhanças

O conceito de vizinhança representa uma competição pelo melhor aprendizado e o ajuste do vencedor e da sua vizinhança é um estímulo para que os neurônios ao redor do vencedor também melhorem.

Durante a etapa de treinamento é identificado o neurônio que tem os parâmetros de entrada mais parecidos com os valores dos pesos. Este procedimento é realizado via cálculo da distância euclidiana, Eq. 1.3, entre o parâmetro de entrada $x(t)$ e o peso $w_{i,j}$.

$$d(t) = \sum_{i=1}^n [x_i(t) - w_{i,j}(t)]^2 \quad (1.3)$$

Capítulo 2

Contexto Geológico

A Bacia do Paraná desenvolveu-se sobre uma área de escudo do continente Gondwana Sul e é composta por uma série de núcleos cratônicos, rodeados por vários cinturões móveis e cobertos por bacias molássicas, que foram desenvolvidas durante o ciclo termo-tectônico Brasiliano que se estendeu desde o neoproterozóico até o Ordoviciano. A deformação decorrente deste ciclo teve início entre 700 Ma e 650 Ma, sendo que a maior parte das intrusões de granitos que podemos observar na Bacia, situou-se dentro do limite entre o Proterozóico e o Paleozóico (cerca de 570 Ma) com resfriamento durante o Cambro-Ordoviciano entre 500 – 450 Ma (HAWKESWORTH *et al.*, 2000; ZALAN e WOLF, 1987).

O embasamento que circunda a Bacia do Paraná é dividido em: margem Leste/Sudeste, representado pelas faixas Dom Feliciano e Ribeira ,de idade Brasiliana e de direção NE-SW, separados por um núcleo cratônico designado Rio de La Plata/ Luiz Alves; margem Norte/Nordeste, representada pela faixa Uruaçu, de idade mesoproterozóica, de direção NW e por dois maciços arqueanos (Guaxupé e Goiás) remobilizados durante o ciclo Brasiliano; margem Oeste/Noroeste representada pela faixa de dobramentos Paraguai/Araguaia, também do ciclo Brasiliano, que delimita o extremo da borda Noroeste da Bacia (BORGHI, 2002; HAWKESWORTH *et al.*, 2000).

Dentre os principais grupos de estruturas, nota-se três grupos de lineamentos de direções preferenciais NW-SE, E-W e NE-SW, representando cada um evento termo-tectônico distinto. O conjunto de lineamentos NW-SE são os mais antigos e estão relacionados ao evento termo-tectônico do Transamazônico, e, as zonas de falhas geológicas associadas a este evento foram reativadas durante o rifteamento do Atlântico Sul, no Cretáceo. Os lineamentos E-W, tiveram início a partir do Triássico e são paralelos às zonas de fratura oceânica, sugerindo uma ligação com o desenvolvimento do Atlântico Sul. Os lineamentos NE-SW são derivados do evento tremo-tectônico Brasiliano e de seus cinturões móveis associados. Este último conjunto de lineamentos é isento de diques de basalto (MILANI e ZALAN, 1999).

O registro estratigráfico da Bacia do Paraná é formado por pacote sedimentar e magmático de espessura máxima em torno de 7000 m, que coincide geograficamente com o depocentro estrutural da sinéclise e com a calha do rio paraná (MILANI e RAMOS, 1998b). O registro estratigráfico da Bacia do Paraná é dividido em seis unidades de ampla escala ou supersequências (VAIL *et al.*, 1977) na forma de pacotes rochosos com intervalos temporais de algumas dezenas de milhões de anos de duração e envelopados por superfícies de discordância de caráter inter-regional: Rio Ivaí (Ordoviciano-Siluriano), Paraná (Devoniano), Gondwana I (Carbonífero-Eotriássico), Gondwana II (Meso a Neotriássico), Gondwana III (Neojurássico-Eocretáceo) e Bauru (Neocretáceo). As três primeiras supersequências são representadas por sucessões sedimentares que definem ciclos transgressivos e regressivos ligados às oscilações do nível relativo do mar, durante o Paleozóico, ao passo que as demais correspondem a pacotes de sedimentos continentais com rochas ígneas associadas. As unidades formais da litoestratigrafia, quais sejam os grupos, formações e membros comumente utilizados na descrição do arranjo espacial dos estratos da bacia, inserem-se como elementos particularizados neste arcabouço aloestratigráfico de escala regional (MILANI *et al.*, 2007).

O mapa geológico (Fig. 2.1) apresenta a extensão e os limites da Bacia do Paraná (BIZZI *et al.*, 2003).

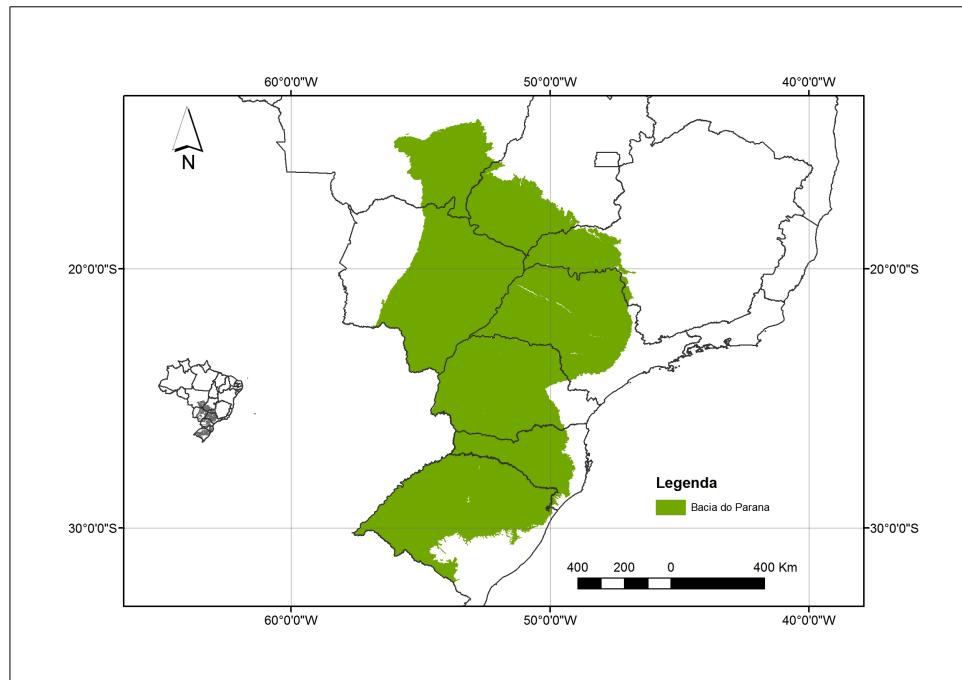


Figura 2.1: Mapa geológico e de localização da área de estudo.

Capítulo 3

Método Proposto e Objetivo

A parte operacional deste programa se divide em duas etapas: 1- Geração de dados sintéticos, 2- Treinamento e 3- Identificação. Cada uma destas etapas será realizada por um programa computacional específico, este programas vão funcionar de forma independente.

O primeiro programa tem por objetivo gerar dados sintéticos que devem simular os resultados obtidos num levantamento de um perfil composto.

O programa da etapa de Treinamento será alimentado com dados de perfilagens cujas respectivas fácies litológicas são conhecidas (inicialmente serão usados dados sintéticos e posteriormente dados reais). Este programa vai gerar um arquivo com os dados do treinamento, este arquivo será usado pelo programa de Operação. Esta é a fase de aprendizagem da rede.

O programa da etapa de identificação vai fazer a classificação, de forma autônoma, das facies litológicas em poços a partir de dados de perfilagem em poços nos quais a litologia é desconhecida. A aprendizagem da rede deve ocorrer de forma continuada, quanto mais informação temos sobre situações nas quais a litologia é conhecida mais bem preparada estará a rede em termos de aprendizagem. Este conceito de aprendizagem é acumulativo e isso ocorrerá através da atualização do arquivo com os dados de treinamento.

Durante a elaboração dos programas será necessário testar a sua eficiência. Estes testes serão realizados através de dados sintéticos que serão gerados por um terceiro programa, gerador de dados sintéticos. Este programa será alimentados com informações da literatura. Após os testes com dados sintéticos a metodologia será validada com dados reais, posteriormente depois de cumpridas todas estas etapas, a metodologia estará pronta para ser utilizada em situações reais.

É importante salientar que neste método o conhecimento do funcional geofísico que rege a relação entre a litologia e a propriedades físicas das rochas não é necessário durante o processo. O conceito de inteligência artificial que será utilizado prescinde do funcional geofísico, o aprendizado é feito através da identificação de padrões

recorrentes.

O ponto positivo desta metodologia é prescindir do funcional geofísico, que por vezes é desconhecido ou de alta complexidade o que exige uma modelagem matemática trabalhosa.

O ponto negativo é a necessidade de se ter muitos dados já analisados em situações conhecidas e variadas para a realização da etapa de treinamento. A etapa de treinamento tem um custo computacional alto.

3.1 Objetivo

O principal objetivo deste projeto é desenvolver um programa computacional do tipo “ machine learning ”, que será implementado na forma de uma Rede Neural Artificial (RNA) dentro do contexto da inteligência artificial. Este programa deve ter a capacidade de identificar, de forma autônoma, fácies litológicas a partir de dados de perfilagem de poços sem a necessidade do uso de um funcional geofísico.

É importante salientar que a metodologia que será desenvolvida neste projeto tem aplicação direta tanto na indústria de exploração mineral, quanto na de água, e na de petróleo e gás.

Capítulo 4

Dados de Perfilagem

As RNAs são capazes de reconhecer padrões (KONATÉ *et al.*, 2014; KUMAR *et al.*, 2015). E padrões muitas vezes são recorrentes no tocante a geologia (VAIL *et al.*, 1977).

Ciclos de deposição de siltes e argilas e areias muitas vezes são controlados pelas variações constantes das estações do ano (CRISTINA LOPES QUINTAS *et al.*, 1999; MILANI e RAMOS, 1998a; MILANI *et al.*, 2000) . Esse registro litológico se faz presente em dados de poços em todo o mundo (SCHERER e LAVINA, 2006).

Em uma perfilagem de poço composta são realizadas diversas medidas de propriedades físicas que ao serem analisadas, em conjunto, tornam possível ao geólogo identificar mudanças litológicas e consequentemente topos e bases de camadas de interesse (ARTUR e SOARES, 2008; FRANCA, ALMERIO & POTTER, 1991; ZALAN, 2007).

A Fig. 4.1 ilustra a disposição de um perfil de poço associado com topos e bases de rochas.

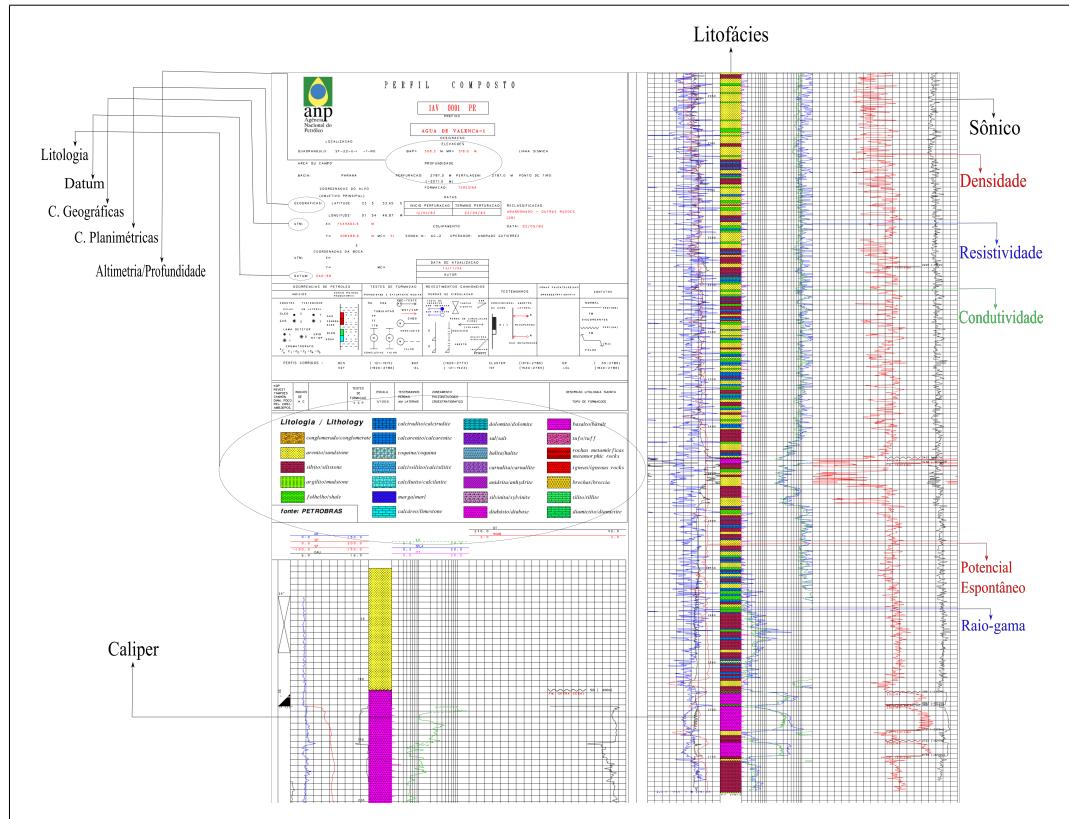


Figura 4.1: Exemplo de um dado público de uma perfilagem de poço composta realizada pela Petrobras, na Bacia do Paraná.

Entretanto não é toda a perfilagem de poço que contém o topo e base de camada. As RNAs se apresentam como uma solução para o problema de identificação litológica e dos topos e bases dessas camadas. Uma vez observado que a variação das propriedades físicas das rochas em subsuperfície variam obedecendo certos padrões (YAN *et al.*, 2014).

Em TELFORD e SHERIFF (1993), encontram-se variações das propriedades físicas dos principais grupos de rochas. A Tab. 4.1 e a Tab. 4.2 apresentam um compêndio desses principais valores.

Tabela 4.1: Compilação de Perfis usados na inferência de litologia.

Rocha	Densidade (g/cm^3)	Raios-Gama (Ci/g)	Potencial-Espontâneo (mV)
Conglomerado	2,50	—	—
Arenito	2,35	2,00 \leftrightarrow 4,00	—
Folhelho	2,40	—	—
Argilito	2,55	—	—
Siltito	2,21	—	—
Dolomita	2,70	8,00	—
Marga	2,50	—	—
Basalto	2,99	0,50	—
Diabásio	2,90	—	—
Lava	2,61	0,33	—
Granito	2,64	0,70 \leftrightarrow 4,80	—
Gabro	3,03	—	—
Peridotito	3,15	—	—
Quartzito	2,60	5,00	—
Xisto	2,64	—	—
Gnaisse	2,80	—	—
Serpentinito	2,78	—	—
Anfibolito	2,96	—	—
Eclogito	3,37	—	—
Mármore	2,75	—	—

Tabela 4.2: Compilação de Perfis usados na inferência de porosidade, permeabilidade.

Rocha	Resistividade (Ω/m)	Neutrão (API)	Velocidade (km/s)
Conglomerado	$2 \times 10^3 \leftrightarrow 10^4$	—	$1,80 \leftrightarrow 4,90$
Arenito	$1 \leftrightarrow 6,4 \times 10^8$	—	$4,00 \leftrightarrow 4,30$
Folhelho	$50 \leftrightarrow 10^7$	—	$2,15 \leftrightarrow 3,30$
Argilito	$10 \leftrightarrow 8 \times 10^2$	—	—
Siltito	$1 \leftrightarrow 100$	—	$4,00 \leftrightarrow 6,20$
Dolomita	$3,5 \times 10^2 \leftrightarrow 5 \times 10^3$	—	$5,70 \leftrightarrow 6,00$
Marga	$3 \leftrightarrow 70$	—	—
Basalto	$10 \leftrightarrow 1,3 \times 10^7$	—	$5,00 \leftrightarrow 5.80$
Diabásio	$20 \leftrightarrow 5 \times 10^7$	—	—
Granito Porfirítico (seco)	$1,3 \times 10^6$	—	5,80
Granito Porfirítico (úmido)	$4,5 \times 10^3$	—	$5,00 \leftrightarrow 5.60$
Gabro	$10^3 \leftrightarrow 10^6$	—	$5,00 \leftrightarrow 5.80$
Peridotito (seco)	$6,5 \times 10^3$	—	—
Peridotito (úmido)	3×10^3	—	—
Xisto	$20 \leftrightarrow 10^4$	—	—
Gnaisse (seco)	3×10^6	—	—
Gnaisse (úmido)	$6,8 \times 10^4$	—	—
Tufa (seca)	2×10^3	—	$1,80 \leftrightarrow 3,50$
Tufa (úmida)	10^5	—	—
Mármore	$10^2 \leftrightarrow 2,5 \times 10^8$	—	—

4.1 Modelo proposto para gerar os dados sintéticos

O modelo proposto para o teste da rede neuronal foi concebido com base em um modelo geológico esquemático proposto por MOHRIAK *et al.* (2008) *apud* (EIRAS, 1996). Esta simplificação reproduz, em uma bacia do tipo sinéclise, estruturas geológicas como Horts, Grábens, semi-grábens, falhas normais, reversas. E representa ainda processos halocinéticos referenciados por EIRAS (1996).

A Fig. 4.2 representa o modelo em tratado anteriormente. A caixa aumentativa evidencia a falha normal, representada no dado de poço por um contato não plano-paralelo.

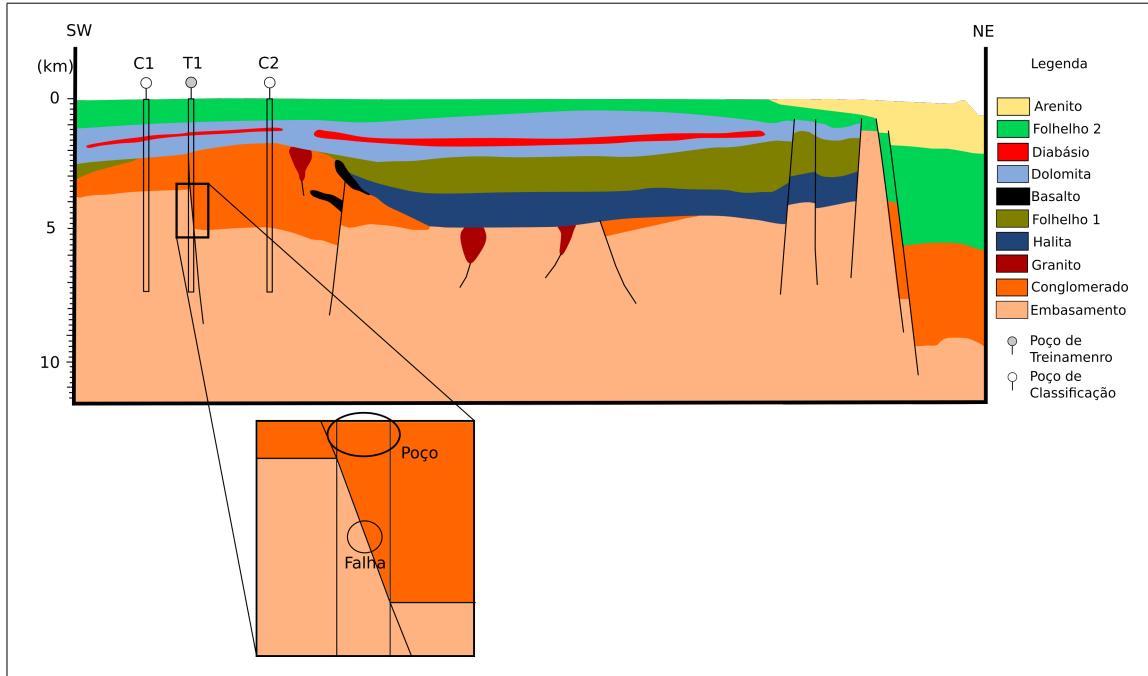


Figura 4.2: Modelo Simplificado baseado em MOHRIAK *et al.* (2008).

A partir da Fig. 4.2 foram gerados três poços, na parte SW do perfil, com profundidades de 7 km cada. Os três poços contém um conjunto com 4 dados de propriedades físicas que são densidade, raio-gama, resistividade e velocidade, respectivamente. Baseado em resultados publicados, na literatura geocientífica, anteriormente, e retirados de TELFORD e SHERIFF (1993) os poços simulam dados de *well logging*.

Os poços simulam diferentes padrões interpretativos usuais da ciência de perfilagem¹. O poço denominado T1² se localiza entre os poços C1 e C2 atravessando uma falha normal.

A Fig. 4.3 apresenta os dados do poço T1. As espessuras das camadas são de 800 m de embasamento, 2 km de uma mistura crescente entre conglomerado e embasamento, perfazendo um padrão sino nos dados de perfilagem, 2 km de conglomerado, 1 km de dolomita (pacote inferior), 300 m de diabásio, 400 m de dolomita (pacote superior), 600 m de folhelho 2. A falha foi representada por uma função linear da variação de profundidade por propriedade física de uma mistura crescente de conglomerado e embasamento.

¹Padrões usuais reconhecido por intérpretes geralmente associados a horizontes de interesse. Esses padrões são identificados como padrões sinos, sinos invertidos, serra e caixa.

²T1: poço escolhido para treinar a rede neuronal.

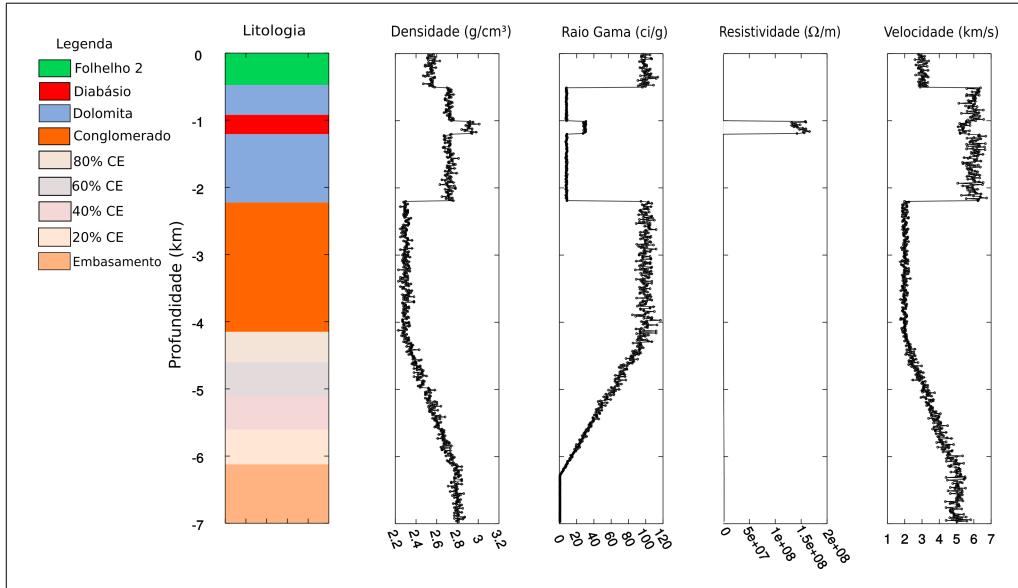


Figura 4.3: Dado de perfilagem sintético, T1. Aonde a porcentagem de CE indica a mistura de conglomerado com embasamento

O poço C1³, Fig. 4.4, possui as mesmas classes de rochas do poço T1. A escolha da posição dos poços quase que exclusivamente na parte SW do perfil se deu em virtude da localização do poço T1. Uma vez que espera-se da rede já treinada um reconhecimento das classes já estudadas. Os pacotes sedimentares apresentam espessuras de 2,8 km de embasamento, 1,6 km de conglomerado, 1 km de dolomita (segundo pacote), 200 m de diabásio, 500 m de dolomita (primeiro pacote) e 500 m de folhelho.

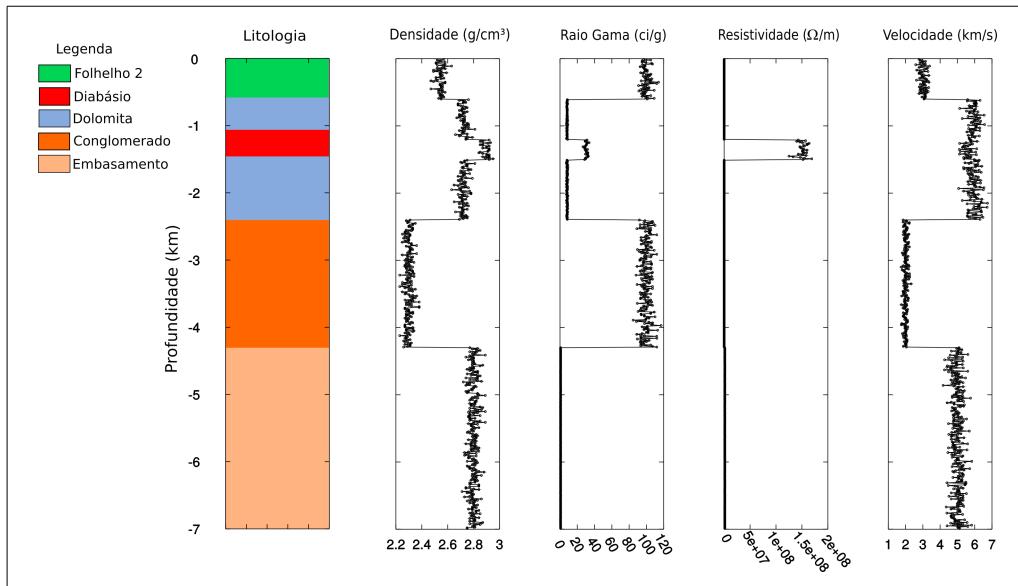


Figura 4.4: Dado de perfilagem sintético, C1.

³C1: Poço de classificação da rede neuronal número 1.

O poço C2⁴, Fig. 4.5, localiza-se em um alto estrutural, e apresenta espessura de 5 km de conglomerado. O embasamento possui uma espessura de 1,8 km. Os pacotes de folhelho 2, dolomita (pacote superior) e diabásio 500 m respectivamente. E o segundo pacote sedimentar de dolomita 1,6 km.

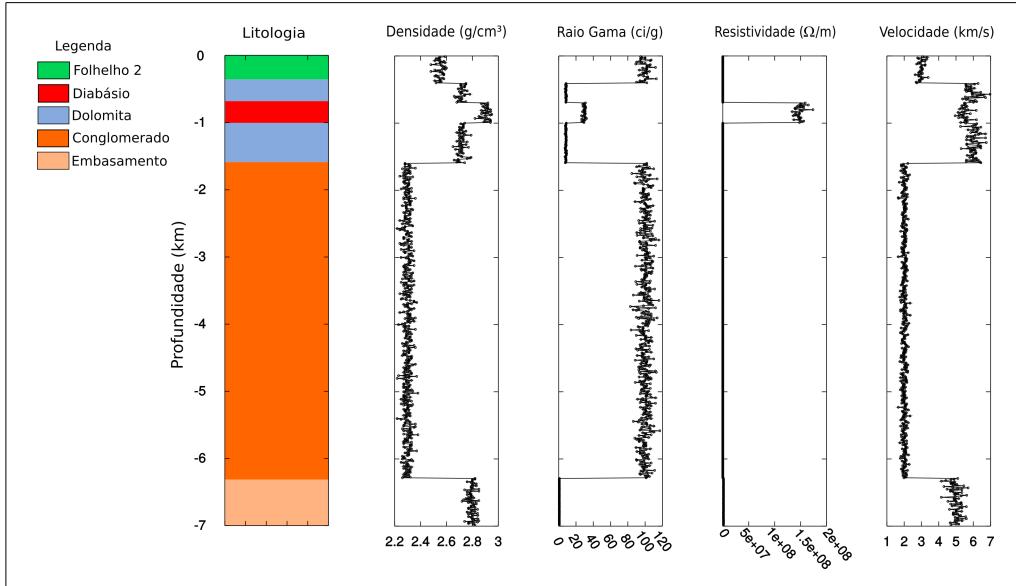


Figura 4.5: Dado de perfilagem sintético, C2.

4.2 Dado Real

A triagem dos dados públicos contemplaram 506 arquivos *.dlis, 113 arquitov *.lis, 118 dados adicionais, 125 perfis compostos digitalizados, 174 poços públicos, 120 arquivos *.agp, todos localizados na Bacia Sedimentar do Paraná.

O conjunto de dados *.lis e *.dlis estão sendo convertidos para arquivos em formato texto, que serão posteriormente concatenados com os arquivos *.agp afim de se obter o input da rede. Este processo ainda se encontra na fase inicial com cerca de 3% concluído.

A Fig. 4.6 mostra a localização e distribuição dos poços na Bacia do Paraná.

⁴C2: Poço de classificação da rede neuronal número 2.

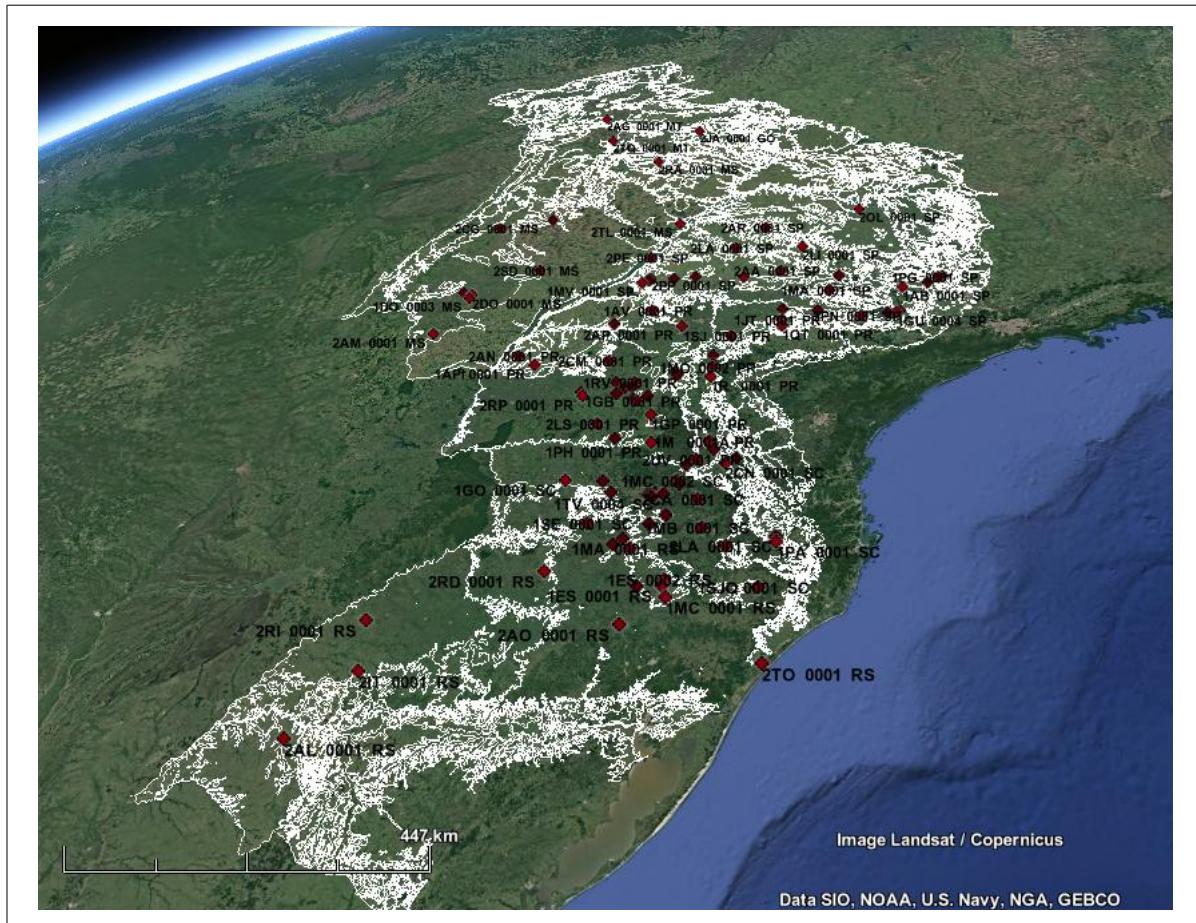


Figura 4.6: Localização dos poços de trabalho.

Capítulo 5

Resultados e Discussões

A Fig. 5.1 apresenta à variação das propriedades físicas analisadas por agrupamento de classes de rochas para o poço T1. Em destaque, de vermelho, o litotipo diabásio.

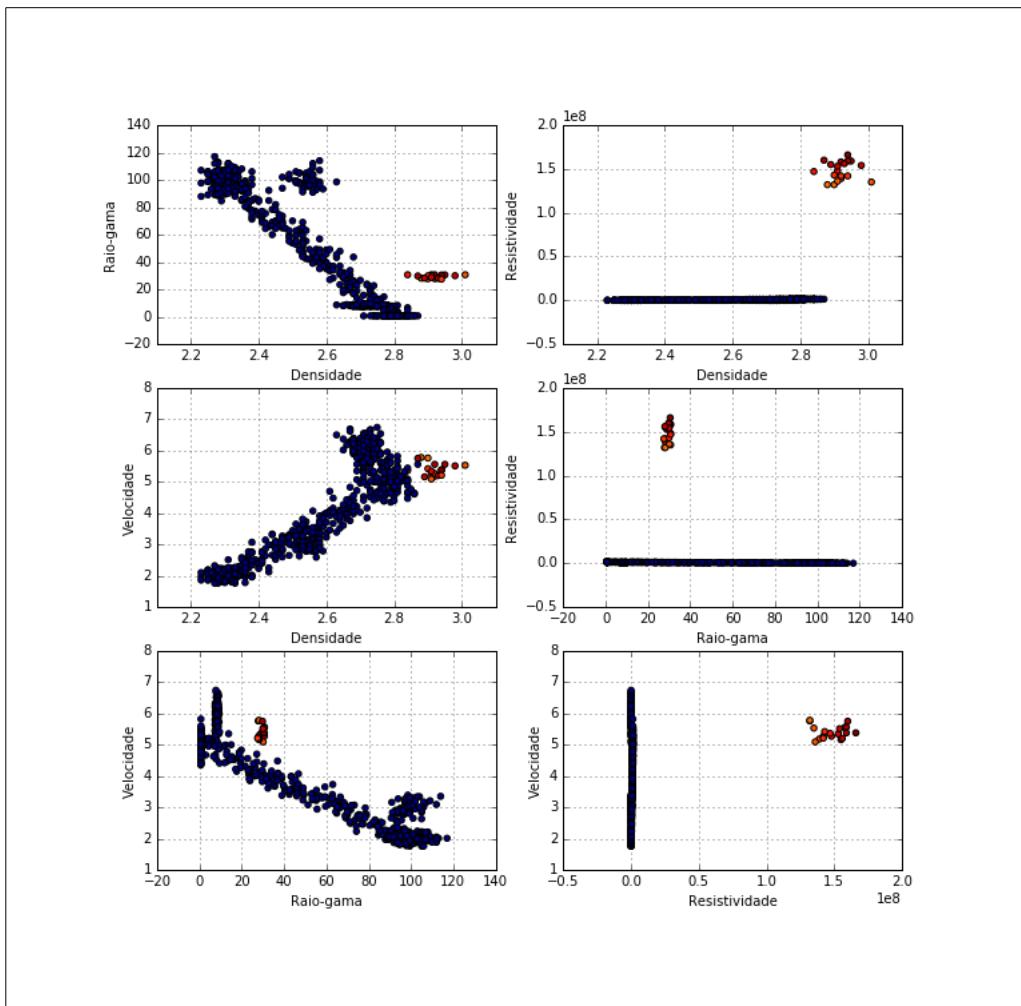


Figura 5.1: Agrupamento de dados do poço T1.

É perceptível o notável contraste de variação das propriedades físicas entre a

rocha de origem ígnea, em contraste com as propriedades físicas das demais rochas de origem sedimentar e metamórfica.

A Fig. 5.2 apresenta à variação das propriedades físicas analisadas por agrupamento de classes de rochas para o poço C1. Em destaque, de vermelho, o litotipo diabásio.

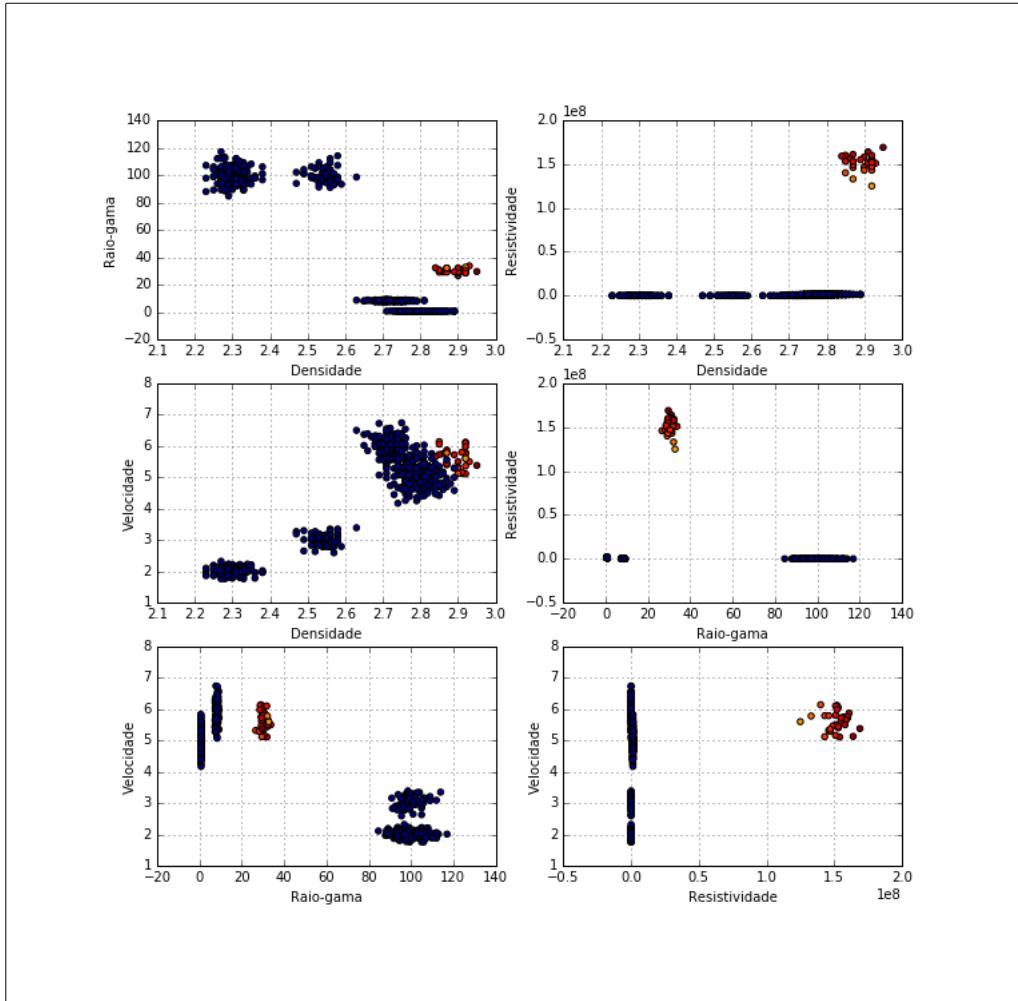


Figura 5.2: Agrupamento de dados do poço C1.

Neste caso, o agrupamento das classes de rochas é mais evidente, no gráfico de raio-gama por densidade, que evidencia os 5 litotipos distintamente. E, da mesma maneira, o gráfico de velocidade por densidade.

A Fig. 5.3 apresenta à variação das propriedades físicas analisadas por agrupamento de classes de rochas para o poço C2. Em destaque, de vermelho, o litotipo diabásio.

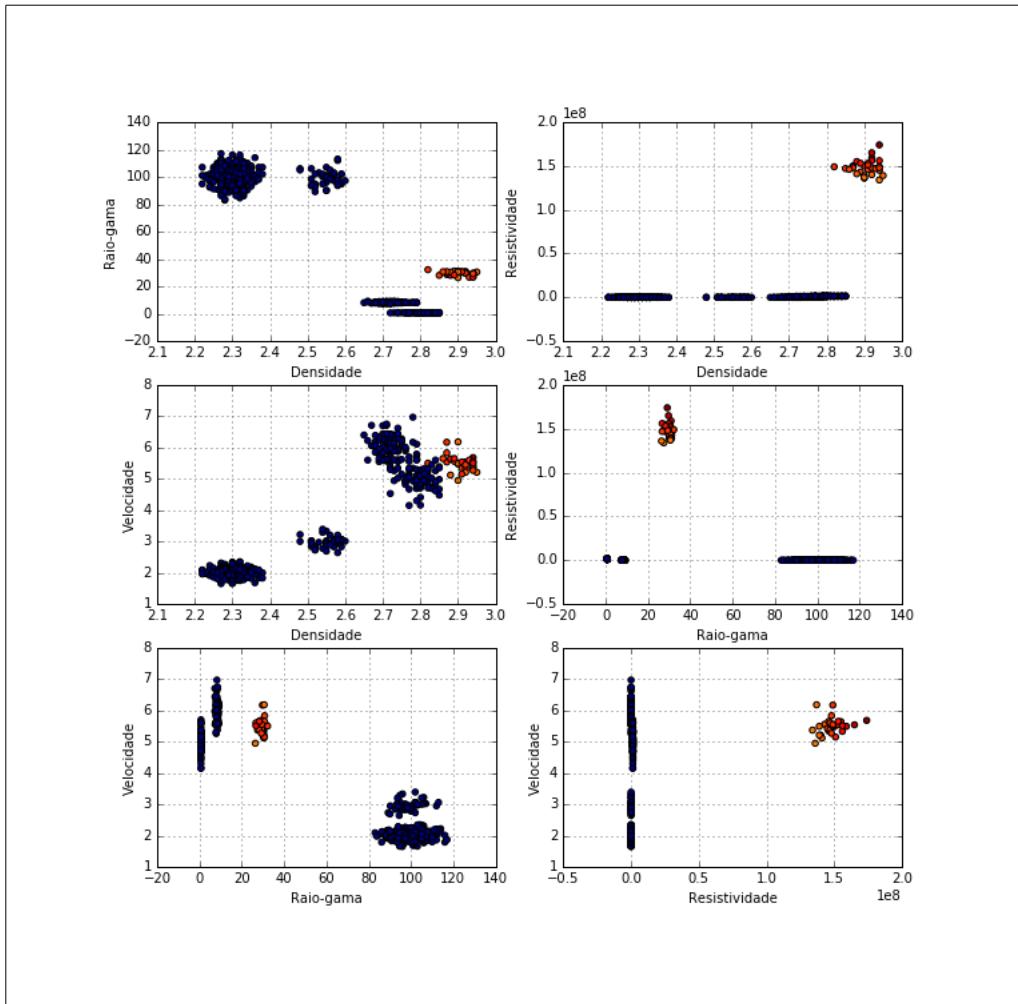


Figura 5.3: Agrupamento de dados do poço C2.

Na mesma forma, o agrupamento das classes de rochas é mais evidente, no gráfico de raio-gama por densidade, que evidencia os 5 litotipos distintamente. E, da mesma maneira, o gráfico de velocidade por densidade.

5.1 Treinamento

A etapa de treinamento consiste em um ajuste de pesos dos neurônios da rede. Nesta fase, é identificado o neurônio que tem os valores dos pesos mais parecidos com os parâmetros de entrada da rede.

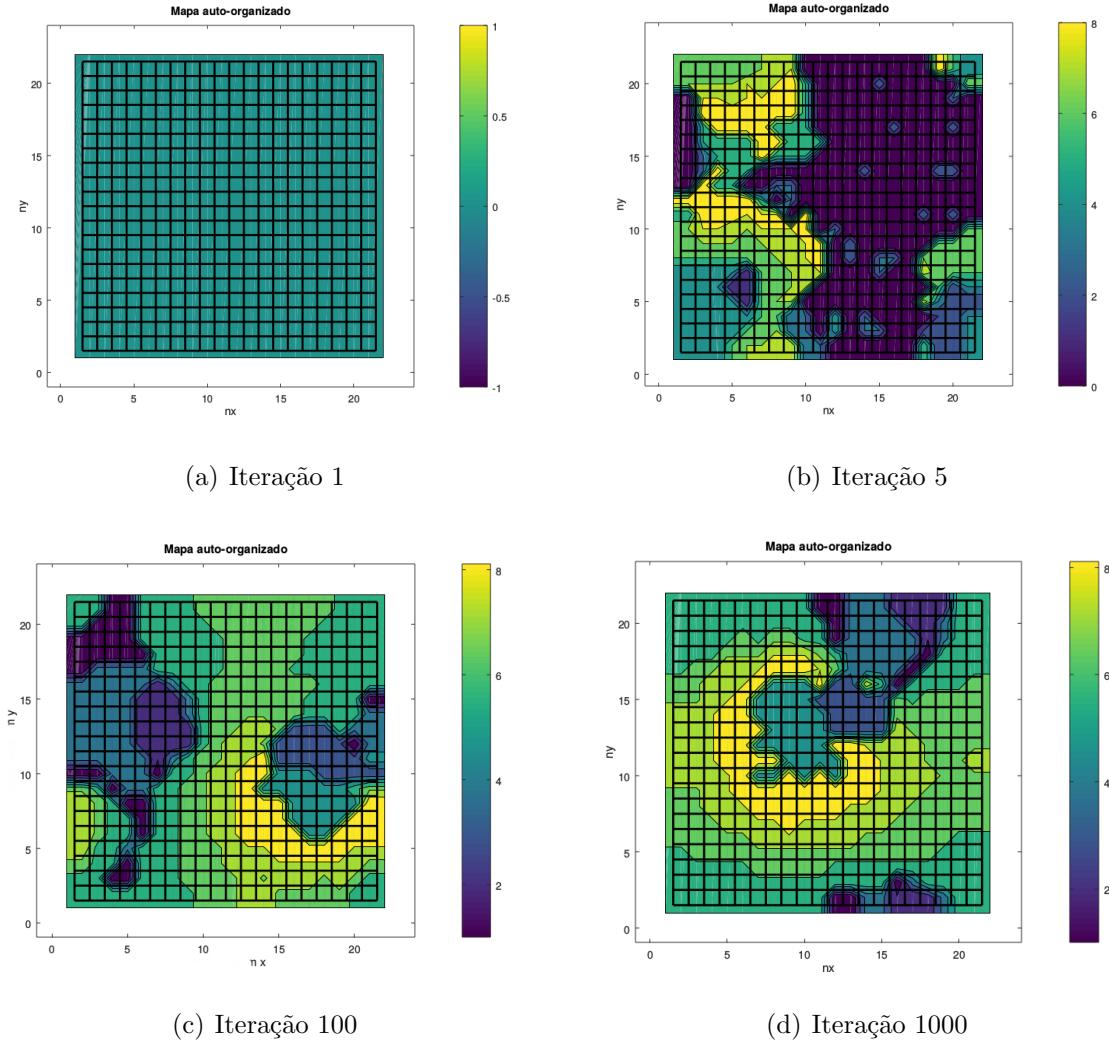


Figura 5.4: Mapas auto-organizáveis e sua evolução temporal.

Os mapas da Fig. 5.4 apresentam as zonas do hiperplano especializadas em identificar as classes de rochas. O código numérico 1 representa folhelho, 2 dolomita, 3 diabásio, 4 conglomerado, 5 embasamento, 6 mistura conglomerado/embasamento 75%, 7 mistura conglomerado/embasamento 50%, 8 mistura conglomerado/embasamento 25%.

A Fig. 5.6 apresenta o teste de convergência da rede neuronal.

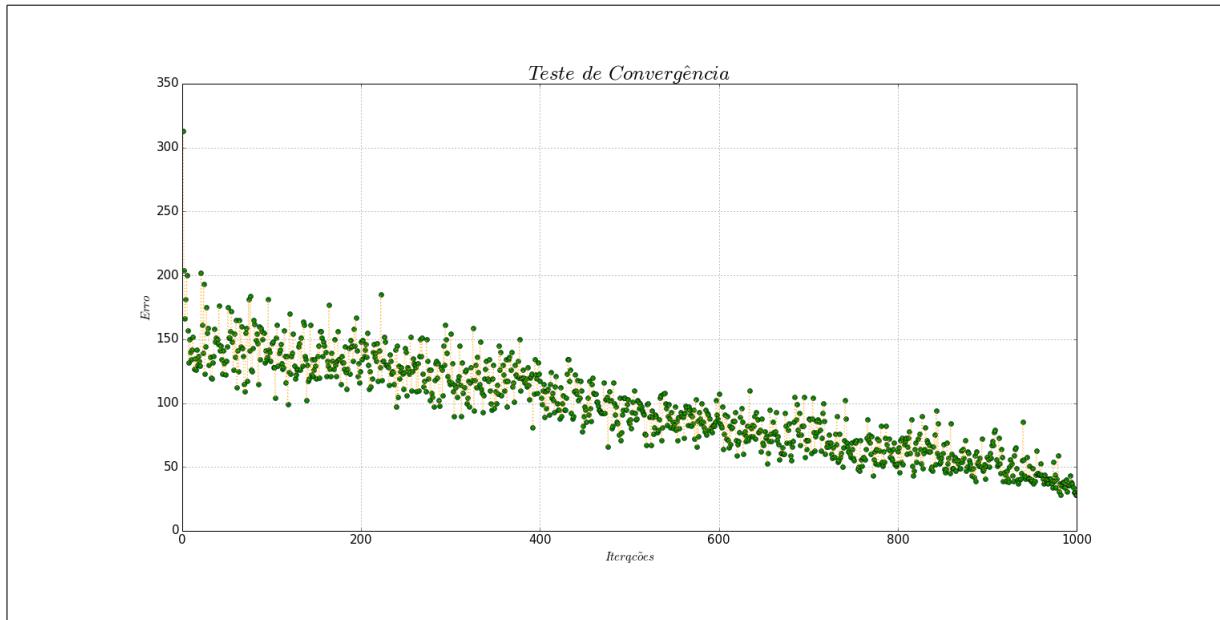


Figura 5.5: Teste de convergência da rede.

O teste de convergência mostra que a rede se encontra estabilizada em 1000 iterações. Isto significa ser inócuo aumentar a iteração afim de diminuir o erro.

5.2 Identificação

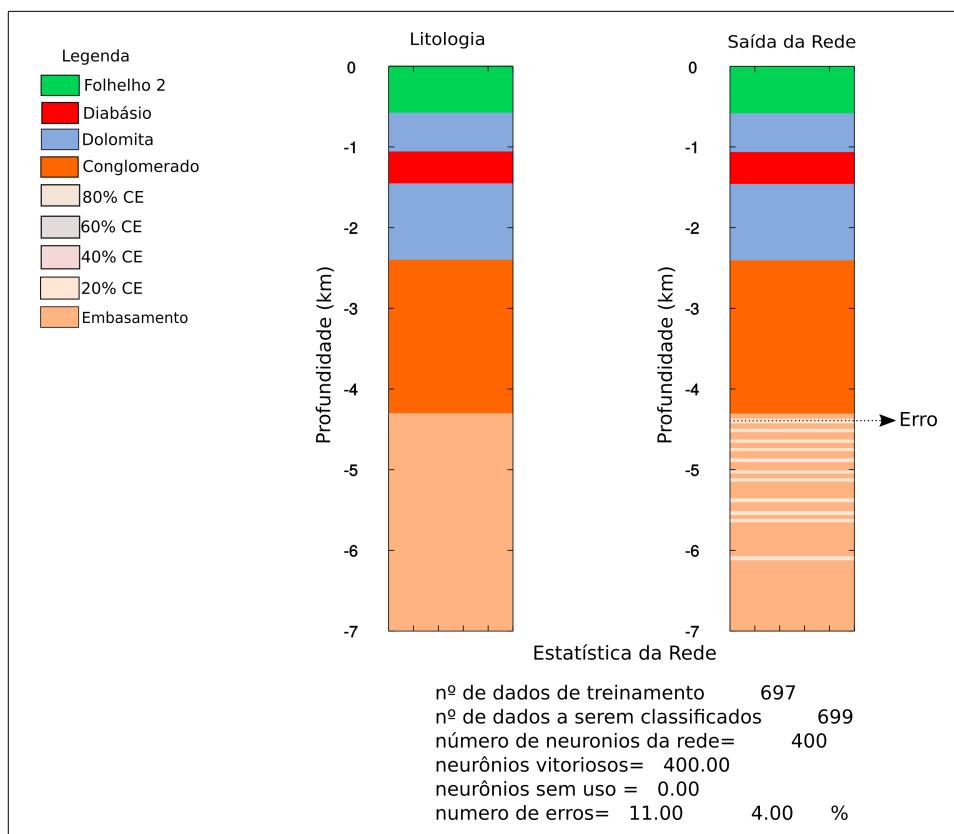


Figura 5.6: Dado de saída da rede para o poço de classificação C1.

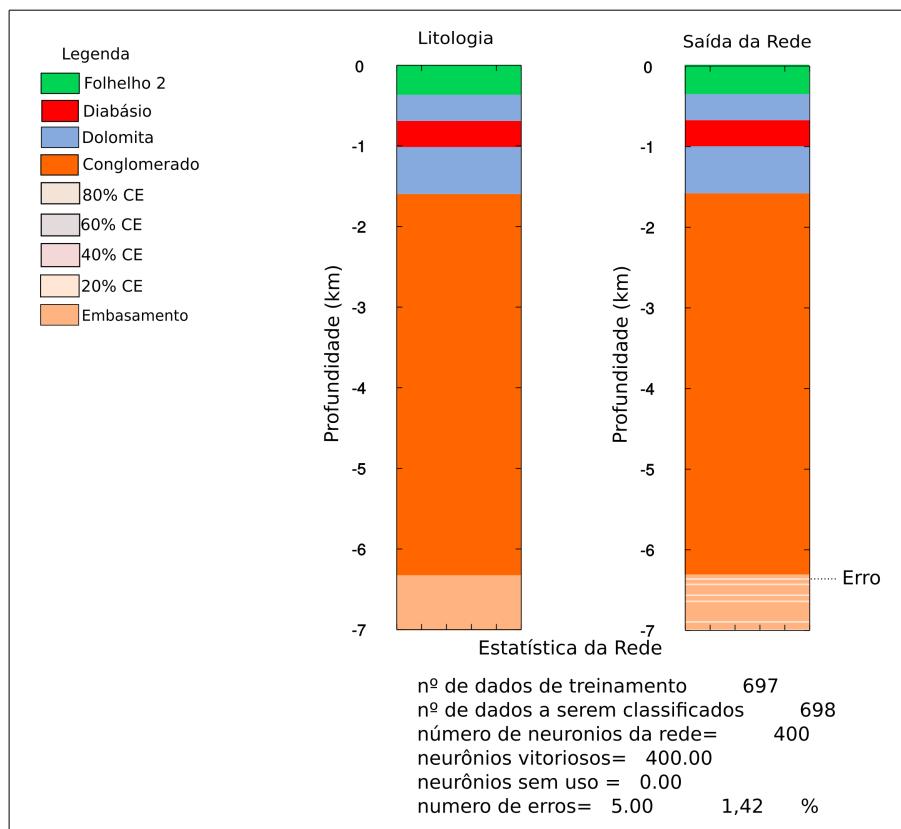


Figura 5.7: Dado de saída da rede para o poço de classificação C2.

Capítulo 6

Conclusões

Referências Bibliográficas

- ADIBIFARD, M., TABATABAEI-NEJAD, S., KHODAPANAH, E., 2014, “Artificial Neural Network (ANN) to estimate reservoir parameters in Naturally Fractured Reservoirs using well test data”, *Journal of Petroleum Science and Engineering*, v. 122, pp. 585–594. ISSN: 09204105. doi: 10.1016/j.petrol.2014.08.007. Disponível em: <<http://linkinghub.elsevier.com/retrieve/pii/S0920410514002563>>.
- ARTERO, A. O., 2008, *Inteligência Artificial Teórica e Prática*. 1st ed. São Paulo, Livraria da Física. ISBN: 9788578610296. Disponível em: <www.livrariadafisica.com.br>.
- ARTUR, P. C., SOARES, P. C., 2008, “Paleoestruturas e petróleo na Bacia do Paraná, Brasil”, *Brazilian Journal of Geology*, v. 32, n. 4, pp. 433–448.
- BENAOUEDA, D., WADGE, G., WHITMARSH, R. B., et al., 1999, “Inferring the lithology of borehole rocks by applying neural network classifiers to down-hole logs: An example from the Ocean Drilling Program”, *Geophysical Journal International*, v. 136, n. 2, pp. 477–491. ISSN: 0956540X. doi: 10.1046/j.1365-246X.1999.00746.x.
- BIZZI, A. L., SCHOBENHAUS, C., VIDOTTI, R. M., et al., 2003, *Geologia, tectônica e recursos minerais do Brasil: texto, mapas e SIG*. CPRM - Serviço Geológico do Brasil.
- BORGHI, L., 2002, “A Bacia do Paraná”, *Anuário do Instituto de Geociências - IGEO, Departamento de Geologia*.
- CHAKRAVARTHY, S., CHUNDURU, R., FANINI, O., et al., 1999, “Detection of layer boundaries from array induction tool responses using neural networks”, *69th Ann. Internat. Mtg*, pp. 140–143. doi: 10.1190/1.1820779.
- CRISTINA LOPES QUINTAS, M., SILVIA MARIA MANTOVANI, M., VICTOR ZALÁN, P., 1999, “Contribuição Ao Estudo Da Evolução Mecânica Da Bacia Do Paraná”, *Revista Brasileira de Geociências*, v. 29, n. 2, pp. 217–226.

- EIRAS, J. F., 1996, “Influência da tectônica do Arco de Carauari na sedimentação fanerozoica da Bacia do Solimões, norte do Brasil. In Congresso Brasileiro de Geologia, 39. Salvador, Bahia.” v. 1, pp. 52–53.
- FELDMAN, J. A., FANTY, M. A., GODDARD, N. H., 1988, “Computing With Structured Neural Networks.” *Computer*, v. 21, n. 3, pp. 91–103. ISSN: 00189162. doi: 10.1109/2.34.
- FRANCA, ALMERIO & POTTER, E., 1991. “Reservoir Potential of glacial deposits of the Itarare Group.pdf” . .
- FREUND, Y., MASON, L., 1999, “The alternating decision tree learning algorithm”, *International Conference on Machine Learning*, v. 99, pp. 124–133. ISSN: 14602431. doi: 10.1093/jxb/ern164.
- HAGAN, M. T., DEMUTH, H. B., BEALE, M. H., et al., 1996, “Neural Network Design”, p. 1012. Disponível em: <<http://books.google.ru/books?id=bUNJAAAACAAJ>>.
- HALL, P., DEAN, J., KABUL, I. K., et al., 2014, “An Overview of Machine Learning with SAS ® Enterprise Miner™”, , n. Rosenblatt 1958, pp. 1–24.
- HAWKESWORTH, C., GALLAGHER, K., KIRSTEIN, L., et al., 2000, “Tectonic controls on magmatism associated with continental break-up an example from the Paraná Etendeka Province”, *Earth and Planetary Science Letters*, v. 179 (jun.), pp. 335–349.
- KANAL, L. N., 2001, “Perceptrons”, *Encyclopedia of Computer Science*, pp. 11215–11218. Disponível em: <
- KONATÉ, A. A., PAN, H., KHAN, N., et al., 2014, “Prediction of porosity in crystalline rocks using artificial neural networks: An example from the Chinese Continental Scientific Drilling Main hole”, *Studia Geophysica et Geodaetica*, v. 59, n. 1, pp. 113–136. ISSN: 00393169. doi: 10.1007/s11200-013-0993-5.

- KROGH, A., 2008, “What are artificial neural networks?” *Nature biotechnology*, v. 26, n. 2, pp. 195–197. ISSN: 1546-1696. doi: 10.1038/nbt1386. Disponível em: <<http://www.ncbi.nlm.nih.gov/pubmed/18259176>>.
- KUMAR, R., AGGARWAL, R., SHARMA, J., 2015, “Comparison of regression and artificial neural network models for estimation of global solar radiations”, *Renewable and Sustainable Energy Reviews*, v. 52, pp. 1294–1299. ISSN: 13640321. doi: 10.1016/j.rser.2015.08.021. Disponível em: <<http://www.sciencedirect.com/science/article/pii/S1364032115008679>>.
- LEVY, S., 1997, “The Computer”, *Newsweek*, v. 130, n. 22, pp. 28. Disponível em: <<file:///D:/Dropbox/Whitfield/History/SemesterTwo/ChangingTimes/Research/EBSCOhost2.htm>>.
- MACKAY, D. J. C., 2005, *Information Theory, Inference, and Learning Algorithms* David J.C. MacKay, v. 100. ISBN: 9780521642989. doi: 10.1198/jasa.2005.s54. Disponível em: <<http://pubs.amstat.org/doi/abs/10.1198/jasa.2005.s54>> <http://www.cambridge.org/0521642981>>.
- MAO, J., 1996, “Why artificial neural networks?” *Communications*, v. 29, pp. 31–44. ISSN: 00189162. doi: 10.1109/2.485891. Disponível em: <http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=485891>.
- MCCULLOCH, W. S., PITTS, W., 1943, “A logical calculus of the ideas immanent in nervous activity”, *The Bulletin of Mathematical Biophysics*, v. 5, n. 4, pp. 115–133. ISSN: 00074985. doi: 10.1007/BF02478259.
- MICHIE, E. D., SPIEGELHALTER, D. J., TAYLOR, C. C., 1994, “Machine Learning , Neural and Statistical Classification”, *Technometrics*, v. 37, n. 4, pp. 459. ISSN: 00401706. doi: 10.2307/1269742. Disponível em: <<http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.94.3615>>.
- MILANI, E. J., RAMOS, V. A., 1998a, “Paleozoic orogenies in southwestern Gondwana and the subsidence cycles of the Parana Basin\rOrogenias paleozoicas no dominio sul-occidental do Gondwana e os ciclos de subsidencia da bacia do Parana”, *Revista Brasileira de Geociencias*, v. 28, n. 4, pp. 473–484.
- MILANI, E. J., BRANDÃO, J., ZALÁN, P. V., et al., 2000, “Petróleo na margem continental Brasileira: Geologia, exploração, resultados e perspectivas”,

- Revista Brasileira de Geofísica*, v. 18, n. 3, pp. 351–396. ISSN: 0102261X.
doi: 10.1590/S0102-261X2000000300012.
- MILANI, E. J., RAMOS, V. A., 1998b, “Orogenias paleozóicas no domínio sul-oeste do Gondwana e os ciclos de subsidência da Bacia do Paraná”, *Brazilian Journal of Geology*, v. 28, n. 4, pp. 473–484.
- MILANI, E. J., ZALAN, P. V., 1999, “An outline of the geology and petroleum systems of the Paleozoic interior basins of South America”, *Milani1999*, v. 22, pp. 199–205. Disponível em: <<http://www.episodes.co.in/www/backissues/223/199-205Milani.pdf>>.
- MILANI, E., SPADINI, A., TERRA, G., et al., 2007, *Boletim de geociências da Petrobras*, v. v. 15. Cenpes.
- MINSKY, M., PAPERT, S., 1969, *Perceptrons*. doi: 10.1016/j.neucom.2015.05.138.
- MISRA, J., SAHA, I., 2010, “Artificial neural networks in hardware: A survey of two decades of progress”, *Neurocomputing*, v. 74, n. 1-3, pp. 239–255. ISSN: 09252312. doi: 10.1016/j.neucom.2010.03.021. Disponível em: <<http://dx.doi.org/10.1016/j.neucom.2010.03.021>>.
- MOHRIAK, W., SZATMARI, P., ANJOS, S., 2008, *Sal: Geologia e Tectônica. Exemplos nas Bacias Brasileiras*. 1 ed. São Paulo, SP., Beca. ISBN: 978-85-87256-49-2.
- NEDJAH, N., DA SILVA, F. P., DE SÁ, A. O., et al., 2016, “A massively parallel pipelined reconfigurable design for M-PLN based neural networks for efficient image classification”, *Neurocomputing*, v. 183, pp. 39–55. ISSN: 18728286. doi: 10.1016/j.neucom.2015.05.138.
- POULTON, M. M., 2002, “Neural networks as an intelligence amplification tool: A review of applications”, *Geophysics*, v. 67, n. 3, pp. 979. ISSN: 0016-8033. doi: 10.1190/1.1484539.
- ROSENBLATT, F., 1962, “Principles of Neurodynamics. Perceptrons and the Theory of Brain Mechanisms.” *Archives of General Psychiatry*, v. 7, pp. 218–219. ISSN: 0003-990X. doi: 10.1001/archpsyc.1962.01720030064010.
- SALJOOGHI, B. S., HEZARKHANI, A., 2014, “Comparison of WAVENET and ANN for predicting the porosity obtained from well log data”, *Journal of Petroleum Science and Engineering*, v. 123, pp. 172–182. ISSN: 09204105. doi: 10.1016/j.petrol.2014.08.025.

- SCHERER, C. M. S., LAVINA, E. L. C., 2006, “Stratigraphic evolution of a fluvial-eolian succession: The example of the Upper Jurassic-Lower Cretaceous Guara and Botucatu formations, Parana Basin, Southernmost Brazil”, *Gondwana Research*, v. 9, n. 4, pp. 475–484. ISSN: 1342937X. doi: 10.1016/j.gr.2005.12.002.
- SIMARD, M., SAATCHI, S. S., DE GRANDI, G., 2000, “The use of decision tree and multiscale texture for classification of JERS-1 SAR data over tropical forest”, *IEEE Transactions on Geoscience and Remote Sensing*, v. 38, n. 5 I, pp. 2310–2321. ISSN: 01962892. doi: 10.1109/36.868888.
- TELFORD, W. M., SHERIFF, R. E., 1993, *Applied Geophysics*. Cambridge University Press.
- VAIL, P. R., MITCHUM, R. M., THOMPSON, S., 1977, *Seismic stratigraphy and global changes of sea level*. Seismic stratigraphy: applications to hydrocarbon exploration. APPG.
- YAN, Z., XIAODONG, Z., JIAOTONG, L., et al., 2014, “Lithology identification research based on self-organizing map of data mining method”, .
- ZALAN, P. V., WOLF, S., 1987, “Tectônica e Sedimentação da Bacia do Paraná”. In: *Simpósio sul-brasileiro de geologia, SBG, 3, Atas, Curitiba-PR.*, v. 1, pp. 441–477.
- ZALAN, P. V., 2007, “Evolução Fanerozóica das Bacias Sedimentares Brasileiras”. In: *Geologia da Plataforma Sul-Americana*, Petrobras, cap. 23, pp. 595–613, Rio de Janeiro.
- ZHANG, L., POULTON, M., ZHANG, Z., et al., 1999, “Fast forward modeling simulation of resistivity well logs using neural network”, *69th Ann. Internat. Mtg*, pp. 124–127. doi: 10.1190/1.1820734.

Cronograma de Atividades

Etapa	Meses																								
	01	02	03	04	05	06	07	08	09	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	
Pesquisa na Literatura	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	
Disciplinas		X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	
Formulação da Rede						X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	
Treino									X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	
Resultado																	X	X	X	X	X	X	X	X	
Artigo 1																							X	X	
Artigo 2																									
Tese																									

Tabela 1: Cronograma das atividades previstas para o primeiro biênio

Etapa	Meses											
	25	26	27	28	29	30	31	32	33	34	35	36
Pesquisa na Literatura	X	X	X	X	X	X						
Disciplinas												
Formulação da Rede												
Treino												
Resultado		X	X	X	X	X						
Artigo 1		X	X	X								
Artigo 2					X	X	X	X				
Tese									X	X	X	X

Tabela 2: Cronograma das atividades previstas para o segundo biênio