

MINISTÉRIO DA DEFESA
EXÉRCITO BRASILEIRO
DEPARTAMENTO DE CIÊNCIA E TECNOLOGIA
INSTITUTO MILITAR DE ENGENHARIA
CURSO DE GRADUAÇÃO EM ENGENHARIA DE COMPUTAÇÃO

CARLOS ALBERTO VILLOTE DE OLIVEIRA

PAULO EDUARDO ALTHOFF

RENAN GEMIGNANI

EMPREGO DE REDES NEURAIIS PARA O POKER TEXAS HOLD'EM

RIO DE JANEIRO

2013

INSTITUTO MILITAR DE ENGENHARIA

CARLOS ALBERTO VILLOTE DE OLIVEIRA

PAULO EDUARDO ALTHOFF

RENAN GEMIGNANI

EMPREGO DE REDES NEURAIIS PARA O POKER TEXAS HOLD'EM

Projeto de Final de Curso apresentado ao
Curso de Engenharia de Computação do In-
stituto Militar (IME) de Engenharia, como
parte das exigências do IME.

Orientador: Prof. Marcos Veloso Peixoto –
D.C.

RIO DE JANEIRO

2013

c2013

INSTITUTO MILITAR DE ENGENHARIA

Praça General Tibúrcio, 80 – Praia Vermelha

Rio de Janeiro - RJ CEP: 22290- 270

Este exemplar é de propriedade do Instituto Militar de Engenharia, que poderá incluí-lo em base de dados, armazenar em computador, microfilmar ou adotar qualquer forma de arquivamento.

É permitida a menção, reprodução parcial ou integral e a transmissão entre bibliotecas deste trabalho, sem modificação de seu texto, em qualquer meio que esteja ou venha a ser fixado, para pesquisa acadêmica, comentários e citações, desde que sem finalidade comercial e que seja feita a referência bibliográfica completa.

Os conceitos expressos neste trabalho são de responsabilidade do(s) autor(es) e do(s) orientador(es).

----- Althoff, Paulo Eduardo/ Gemignani, Renan/ Oliveira, Carlos Alberto Villote de
Emprego de Redes Neurais para o Poker Texas Hold'em, no Rio de Janeiro /
Carlos Alberto Villote de Oliveira, Paulo Eduardo Althoff, Renan Genignani
- Rio de Janeiro : Instituto Militar de Engenharia, 2013.

n f. : il., graf., tab. : - cm.

Dissertação (graduação) - Instituto Militar de Engenharia, 2013.

1. Formação.

INSTITUTO MILITAR DE ENGENHARIA

CARLOS ALBERTO VILLOTE DE OLIVEIRA

PAULO EDUARDO ALTHOFF

RENAN GEMIGNANI

EMPREGO DE REDES NEURAIAS PARA O POKER TEXAS HOLD'EM

Projeto de Final de Curso apresentado ao Curso de Engenharia de Computação do Instituto Militar (IME) de Engenharia, como parte das exigências do IME.

Aprovado em __/__/____ pela seguinte Banca Examinadora:

Orientador: Marcos Veloso Peixoto – D.Sc., do IME

Alex Garcia, D.Sc., do IME

Ricardo Choren Noya, D.Sc., do IME

RIO DE JANEIRO

2013

“ A primeira coisa a entender é que você não entende”.

SOREN AABYE KIERKEGAARD

SUMÁRIO

SUMÁRIO	6
LISTA DE ILUSTRAÇÕES	8
LISTA DE TABELAS	10
1 INTRODUÇÃO	14
1.1 MOTIVAÇÃO	15
1.2 OBJETIVO	15
1.3 JUSTIFICATIVA	15
1.4 METODOLOGIA	16
1.5 ESTRUTURA	16
2 REDES NEURAIS	17
2.1 Histórico	17
2.2 Cérebro Humano	18
2.3 Componentes das Redes Neurais Computacionais	19
2.3.1 Conexões	20
2.3.2 Núcleo de Processamento	20
2.3.3 Funções de Limiar	20
2.3.4 Padrões	22
2.4 Classificações de Redes Neurais	22
2.5 Adaline	23
2.6 Backpropagation	23
3 O PÔQUER	25
3.1 O jogo	25
3.2 O baralho	25
3.3 Ranking de combinações de cartas	26
3.4 Ações	28
3.5 Tipos de pôquer	29
3.5.1 <i>Texas Hold'em</i>	29

4	ESTRUTURA DA SOLUÇÃO	31
4.1	Modelagem do oponente	31
4.2	Rede principal e classificação em perfis	32
5	SELEÇÃO DE JOGADORES PARA TREINAMENTO DA REDE	33
6	CONCLUSÃO	34
7	REFERÊNCIAS	35

LISTA DE ILUSTRAÇÕES

FIG. 2.2.1	Esquema de um neurônio	19
FIG. 2.3.2	Esquema de um neurônio artificial	19
FIG. 2.3.3	RNA na forma de grafo	20
FIG. 2.3.4	Função Linear	21
FIG. 2.3.5	Função Step	21
FIG. 2.3.6	Função Rampa	21
FIG. 2.3.7	Função Gaussiana	21
FIG. 2.3.8	Função Sigmoid	21
FIG. 2.4.9	Conexão feed-forward	22
FIG. 2.5.10	Topologia da rede Adaline	23
FIG. 2.6.11	Topologia Backpropagation	24
FIG. 3.2.1	naipe de copas	25
FIG. 3.3.2	Carta maior	26
FIG. 3.3.3	Par de dez	26
FIG. 3.3.4	Dois pares de sete com quatro	26
FIG. 3.3.5	Trinca de cinco	27
FIG. 3.3.6	Sequência de dez a ás	27
FIG. 3.3.7	<i>Flush</i> de espadas	27
FIG. 3.3.8	<i>Full House</i> de dama com dois	27
FIG. 3.3.9	Quadra de rei	27

FIG. 3.3.10	<i>Straight Flush</i>	28
FIG. 3.3.11	<i>Royal Straight Flush</i>	28
FIG. 4.2.1	Diagrama das classes utilizadas no aprendizado	32
FIG. 5.0.1	Diagrama de classe para as classes básicas	33

LISTA DE TABELAS

TAB. 4.1.1	Parâmetros de entrada da RN utilizada para modelar o oponente	31
TAB. 4.1.2	Parâmetros de saída da RN utilizada para modelar o oponente	31
TAB. 4.2.3	Parâmetros de entrada da RN utilizada para classificar o oponente para cada uma das jogadas em uma janela deslizando	32
TAB. 4.2.4	Parâmetros de saída da RN utilizada para classificar o oponente	32

LISTA DE SIGLAS

RESUMO

Palavras-chave: redes neurais, aprendizado de máquina, pôquer.

ABSTRACT

Keywords: Neural networks, machine learning, poker.

1 INTRODUÇÃO

O emprego de técnicas de aprendizado de máquina de modo a construir agentes autônomos que sejam capazes de jogar alguma forma de pôquer vem sendo estudado há algum tempo devido às características de informação imperfeita do jogo. A maior parte dos estudos na área advém do grupo de inteligência artificial aplicada ao pôquer da Universidade de Alberta. Uma parte dos estudos e implementações de software de pôquer faz uso de estratégias e contra-estratégias baseadas na teoria de equilíbrio de Nash [4], com algoritmos de avaliação sendo utilizados principalmente com o objetivo de selecionar a estratégia mais interessante para a situação atual dentre estas estratégias conforme resposta do oponente, de modo a se manter a menor quantidade de tempo possível em uma estratégia que se é percebida como não-ideal. Esta abordagem é baseada mais fortemente no fundamento matemático da teoria do jogo do que no melhoramento da estratégia em si.

Contudo, mesmo o software baseado em estratégias fixas se vale, em alguns casos, da utilização de algoritmos de aprendizado de máquina baseados em redes neurais para realizar ao menos a modelagem dos oponentes [3] de modo a prever as probabilidades de o oponente responder de certa forma dado seu histórico de jogadas, como no caso de uma das máquinas descritas em [1], em que a abordagem utilizada para construção do agente-fim não é descrita, contudo, redes neurais são usadas para prever as jogadas do oponente, iniciando-se com uma rede limpa que é consequentemente adaptada à medida que o oponente age. Uma desvantagem dessa utilização da rede neural é a possível infactibilidade de se treinar a rede satisfatoriamente em tempo real em um ambiente com agentes humanos. Mesmo com essa desvantagem, a modelagem de redes neurais mostra-se mais acurada do que uma modelagem baseada puramente em pesos relativos a cada tupla ação-mão possível como a mostrada em [2].

Uma abordagem em que o agente objetivo-fim do estudo é de fato baseado em redes neurais está descrita em [5], contudo sem tomar em consideração as especificidades do modo com que o oponente joga, levando a uma abordagem relativamente simples no que concerne a modelagem em si da rede, focando na evolução dos agentes predeterminados via combinação dos mesmos e utilização de heurísticas coevolucionárias como halls da fama para realizar a seleção dos agentes que mais se adaptassem, descartando modelos menos eficientes.

1.1 MOTIVAÇÃO

A utilização de redes neurais aplicadas a problemas de decisão em casos em que pode-se utilizar de grandes quantidades de tempo para realizar o treinamento da rede e no qual existe quantidade suficiente de dados para a realização do mesmo. O Texas Hold'em, em princípio, parece não se prestar exatamente ao modelo necessário para a utilização de redes neurais, uma vez que dada uma mesma entrada, não existe uma resposta 'correta', dado que a adequação de qualquer resposta, à é dependente da subsequente ação tomada pelo adversário. Deste modo, em qualquer modelo de RN utilizado no pôquer para propósitos distintos da mera modelagem de um oponente (dado seu histórico) para uso em outra aplicação, faz-se necessário utilizar alguma técnica de modo a prever as respostas mais prováveis do oponente a cada ação tomada, podendo assim julgar sua adequação, possibilitando assim uma análise mais acurada.

1.2 OBJETIVO

No presente trabalho, tenciona-se melhorar a maneira com que uma análise prévia do oponente pode ser utilizada como entrada adicional para uma rede neural de forma a conseguir classificá-lo em um dentre uma série de perfis distintos pré-gerados e treinados, também modelados utilizando redes neurais, informação esta que será utilizada como parâmetro de entrada pela rede de modo a conseguir gerar uma contra-estratégia para um determinado perfil de oponente.

1.3 JUSTIFICATIVA

Redes neurais tem se mostrado ferramentas úteis tanto para a modelagem de oponentes no pôquer quanto para a resolução de problemas de decisão, como já foi mostrado em trabalhos correlatos. Contudo, as implicações decorrentes da utilização de uma rede neural como

agente *stand-alone* capaz de jogar competitivamente ainda não foram exploradas. Apesar de a utilização de uma rede neural como estratégia de jogo estar condicionada aos jogadores cujos jogos forem usados para treinar a rede, estando deste modo sujeitos a certa exploração por parte de um oponente que jogasse de modo anormal quando comparado com os oponentes do jogador base nos jogos utilizados para testar a rede, A utilização de perfis de jogadores e a subsequente classificação dos oponentes dentre estes perfis pode ser utilizada de modo a diminuir o impacto desta potencial fraqueza;

1.4 METODOLOGIA

Primeiramente, realizaram-se estudos teóricos objetivando tanto o estudo das redes neurais quanto de outras abordagens de aprendizado de máquina, além do estudo do jogo do pôquer em si. Em seguida, foram construídas a topologia da rede a ser utilizada para modelar o oponente e a diagramação de classes do sistema como um todo.

1.5 ESTRUTURA

No capítulo 2, apresentamos a fundamentação teórica acerca de redes neurais mínima necessária para o entendimento do trabalho. No capítulo 3, apresentamos um breve resumo das regras do pôquer Texas Hold'em destinado ao leitor que não é familiarizado com o mesmo. No capítulo 4, apresentamos brevemente a topologia a ser utilizada pela rede de simulação do oponente, bem com. No capítulo 5, descreveremos a interface com o banco de dados de mãos bem como as heurísticas utilizadas para selecionar os jogadores utilizados para compor o conjunto de treinamento de cada perfil, bem como os jogadores utilizados para treinar a RN objetivo-fim.

2 REDES NEURAIS

Uma das mais promissoras técnicas de aprendizado de máquina atualmente, é a estrutura conhecida como rede neural, que consiste em formar uma rede de unidades de processamento que enviam mensagens umas as outras, tentando imitar o funcionamento do cérebro humano. Por esse motivo, é a técnica que mais auxilia no entendimento da inteligência humana, e que se espera obter uma inteligência artificial mais próxima daquela [7].

2.1 HISTÓRICO

As primeiras informações sobre neurocomputação surgiram em 1943, em artigos de McCulloch & Pitts, que inauguraram o estudo de redes neurais. Os autores fizeram uma analogia entre células nervosas vivas e o processo eletrônico, simulando o comportamento do neurônio natural, no qual o neurônio possuía uma saída que era uma função da soma de valor de suas diversas entradas. Eles introduziram o cálculo de pesos e unidades limiares.

Em 1948, o livro “Cybernetics” de Wiener descreve alguns conceitos sobre controle, comunicação e processamento estatístico de sinais.

Em 1950, por meio do livro “The Organization of Behavior”, Hebb descreveu um sistema de aprendizado por correlação dos neurônios que acabou dando origem a Regra de Aprendizagem de Hebb. No ano seguinte, Minsky construiu o primeiro neuro computador.

Em 1956, a Inteligência Artificial foi oficialmente criada por McCarthy, Minsky, Shannon, Newell & Simon, Samuel e outros. Em Dartmouth College nasceram os dois paradigmas da IA, a simbólica, que tenta simular o comportamento inteligente humano desconsiderando os mecanismos responsáveis por tal, e a conexionista, que acredita que construindo-se um sistema que simule a estrutura do cérebro, este sistema apresentará inteligência.

No ano seguinte, Rosenblatt criou o Perceptron, um modelo cognitivo que consistia de unidades sensoriais conectadas a uma única camada de neurônios de McCulloch & Pitts, demonstrando que se fossem acrescentadas sinapses ajustáveis, as redes neurais de McCulloch & Pitts podem ser treinadas para classificar padrões em classes linearmente separáveis, convergindo em um número limitado de passos.

Em 1969, Minsky e Papert publicam “Perceptron”, que expõe as limitações do modelo

de Rosenblatt, provando que tais redes não são capazes de resolver uma ampla classe de problemas devido às restrições de representam, tendo impacto devastador no interesse dos pesquisadores sobre redes neurais.

Na década de 1970 não houveram avanços devido a pouca credibilidade dos recentes estudos na área.

Em 1982, a introdução do modelo conexionista permitiu esclarecer boa parte das dúvidas até então existentes em relação ao processo dinâmico executado por certas redes neurais. Em meados da década de 1980 surgiu a descrição do algoritmo de treinamento backpropagation, que se baseia na retropropagação dos erros para realizar os ajustes de pesos das camadas intermediárias. A área de redes neurais teve um desenvolvimento explosivo após a publicação em 1986 do “Parallel Distributed Processing” editado por Rummerhart & McClelland, com a multiplicação exponencial de Journal’s, teses e paper’s científicos.

Na década de 1990 ocorreu a integração da pesquisa de redes neurais à corrente principal da inteligência artificial.

Nos anos 2000 as redes neurais são vistas como um das técnicas de aprendizado de máquina, e não necessariamente como a melhor técnica para muitos problemas.

2.2 CÉREBRO HUMANO

O cérebro humano é composto por unidades chamadas neurônios, cerca de 100 bilhões dessas unidades trabalham em paralelo, enviando sinais eletroquímicos uns aos outros. Apesar de cada neurônio trabalhar na faixa de menos de 1 kHz (uma taxa muito baixa se comparado a um computador moderno que trabalha em GHz), seu poder de processamento é muito elevado devido às múltiplas conexões entre eles e pela comunicação rápida, que permite que os 10^{11} neurônios trabalhem ao mesmo tempo.

A estrutura de um neurônio pode ser dividida em 3 partes: dendritos, corpo celular, axônio (FIG. 2.2.1).

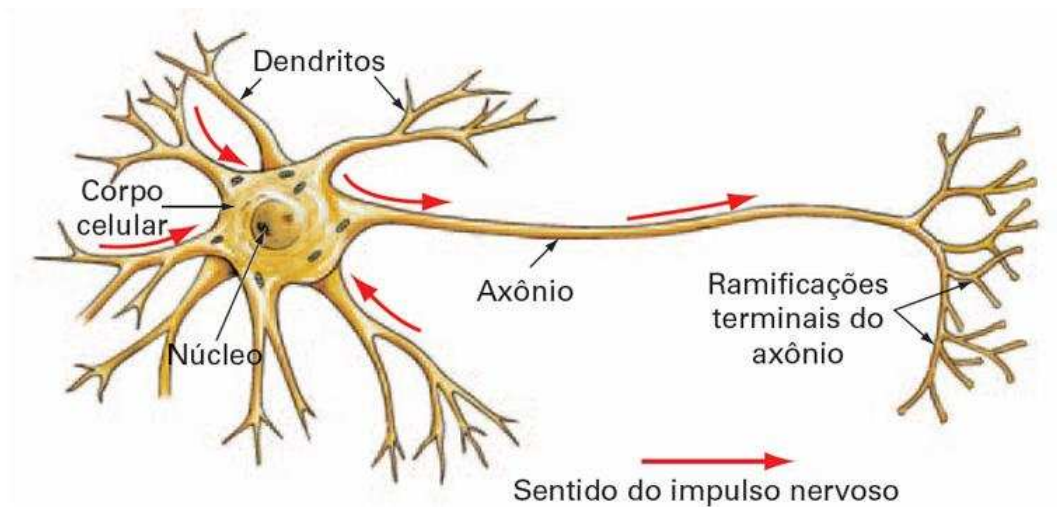


FIG. 2.2.1: Esquema de um neurônio

Os dendritos recebem sinais elétricos dos outros neurônios através das sinapses, que são impulsos nervosos (estímulos) enviados através de ativadores químicos. O corpo celular processa as informações, definindo pesos para as conexões e enviando uma resposta para outro neurônio através do axônio.

Conjuntos de redes de neurônios, são agrupados em áreas de responsabilidades como cerebelo, hipotálamo, córtex, etc. Aumentando a complexibilidade do sistema e intensificando o poder de processamento do cérebro, mas que foge do escopo das redes neurais tradicionais.

2.3 COMPONENTES DAS REDES NEURAIS COMPUTACIONAIS

As redes neurais computacionais, ou redes neurais artificiais (RNA), são feitas a partir de vários neurônios artificiais, que é um modelo eletrônico do neurônio biológico, como representado na figura .

FIG. 2.3.2: Esquema de um neurônio artificial

Podemos destrinchar esse funcionamento do neurônio artificial em 4 componentes principais, as conexões, o núcleo de processamento, as funções e os padrões.

2.3.1 CONEXÕES

Imaginando a RNA como um grafo monodirecionado, teríamos as conexões entre os nós fazendo o papel dos axiônios (ao enviar a informação de saída de um neurônio) e dos dendritos, (associando um peso a cada entrada). Esse peso nada mais é que um valor que será ajustado durante o treinamento da rede neural, podendo ser positivos (excitam a entrada) ou negativos (inibem a entrada). A informação percorre esse grafo por essas conexões para chegar aos núcleos de processamento para ser analisada.

FIG. 2.3.3: RNA na forma de grafo

2.3.2 NÚCLEO DE PROCESSAMENTO

O núcleo de processamento é a parte da rede neural onde é realizado o processamento das entradas vindas das conexões. Na analogia com o grafo, um núcleo de processamento seria um nó com várias entradas e uma única saída, obtida pelo processamento das informações de entrada, geralmente uma combinação linear destas ou uma análise comparativa dessa combinação. Essa saída só pode depender dos elementos de entrada, sem influência de dados externos e deve ser único.

As diversas funções que podem ser utilizadas para estabelecer essa transferência entre neurônios, são chamadas de funções de limiar e são abordadas no item a seguir.

2.3.3 FUNÇÕES DE LIMIAR

As funções de limiar, ou funções para transferência de sinais entre neurônios, são aquelas que determinam a intensidade e a forma que os valores são alterados para serem transmitidos entre eles. Destacam-se cinco tipos: a linear, a step, a em rampa, a gaussiana e a sigmoid.

A função linear é uma equação linear da forma $EQUAÇÃO(f=a.x+b)$, onde a é a inclinação da reta.

FIG. 2.3.4: Função Linear

A step, é uma equação que possui um resultado binário, ou seja, de apenas dois valores possíveis. Assim ela tem a forma EQUAÇÃO(pra x menor um valor, pra x maior outro). Essa função também é conhecida como hard-limiter. Uma desvantagem do uso dessa função é que quaisquer ajustes feitos nos pesos que alimentam os neurônios com o uso desta função terão efeitos não tão eficientes e limitados, devido a sua descontinuidade acentuada.

FIG. 2.3.5: Função Step

A função de limiar em rampa é semelhante à função step, mas a transição entre um valor e outro não é instantânea, devendo haver uma fase de transição entre 2 valores, atenuando o problema citado acima.

EQUAÇÃO(menor q $-y$ um valor, maior q y outro, entre eles uma reta) é o valor de saturação da função, ou seja, durante a transição o valor de $f(x)$ irá variar dentro do intervalo $(-y, y)$. Como mostra a figura FIG. 2.3.6.

FIG. 2.3.6: Função Rampa

A gaussiana é uma função que se baseia em um ponto médio e uma variância pré-definida, EQUAÇÃO $f(x) = e^{-(x^2/v)}$ onde v é a variância. Seu gráfico é apresentado na figura FIG. 2.3.7.

FIG. 2.3.7: Função Gaussiana

Neste trabalho, a função de limiar que será utilizada será uma variante da função em rampa, com a vantagem de diminuir mais ainda o problema da função step, deixando-a totalmente continua, o nome dessa função é sigmoid.

Ela permite uma transição gradual e não linear entre dois estados. Este tipo de função é utilizada também em outras áreas como sociologia e química.

Uma otimização que é muito utilizada no ramo das redes neurais ao se trabalhar com sigmoids, é mudar os limites da função para não trabalhar com números negativos, neste trabalho os valores da sigmoid variam entre 0 e 1, sendo aceito e não aceito respectivamente.

FIG. 2.3.8: Função Sigmoid

2.3.4 PADRÕES

Os padrões são os dados que alimentam a RNA. Podem ser valores numéricos ou imagens, caracteres etc que serão posteriormente transformados em números. A um determinado padrão de entrada corresponde um sinal de saída, sendo que o tipo do padrão de entrada pode ser diferente do padrão de saída. Nestes casos torna-se essencial um pré-processamento das informações, ou seja, a seleção e representação correta das características do padrão a ser aplicado pode afetar a performance da rede. Assim, para um bom desempenho da rede neural, deve-se optar pelo melhor conjunto de características para os padrões.

2.4 CLASSIFICAÇÕES DE REDES NEURAI

Não há um modo oficial de se classificar uma rede neural, mas existem classificações já consagradas quanto a 3 de suas características: tipo de aprendizado, tipo de entradas e forma de conexão.

Tipo de aprendizado identifica se na rede existe uma definição para o que se espera na saída, assim sendo utilizada para ajustar os pesos das entradas, neste caso a RNA é de aprendizado supervisionado, caso contrário é de auto-aprendizado.

Os tipos de entrada podem ser intervalar, onde as entradas são de forma contínua, ou seja, aceitam qualquer valor como entrada. Ou pode ser binária, só aceitando 0 ou 1.

Existem três tipos de forma de conexão, que é a forma como os neurônios estão conectados e como é a transmissão de dados. Na alimentação competitiva, os neurônios de uma mesma vizinhança recebem as mesmas entradas e competem entre si para ser a única saída.

Na alimentação à frente, a rede é dividida em camadas onde as saídas de uma servem como entradas para a próxima camada.

FIG. 2.4.9: Conexão feed-forward

Na retro-alimentação, os sinais de saída são utilizadas para alimentar camadas anteriores.

Neste trabalho será utilizado um modelo de de alimentação à frente, com aprendizado supervisionado e entrada intervalar, conhecido como backpropagation. Esse modelo é baseado

em outro modelo mais simples chamado adaline, e ambos os modelos estão explicitados nos itens subsequentes.

2.5 ADALINE

Adaline (Adaptative Linear Network), como seu nome diz utiliza um combinador adaptativo linear, onde a saída de cada neurônio é a combinação linear das entradas multiplicadas por um peso, seguindo a fórmula: FÓRMULA

A regra de adaptação de pesos utilizada no adaline pode ser o algoritmo perceptron ou o alpha-LMS, a diferença entre eles está na fórmula do peso e do erro, onde o perceptron utiliza as fórmulas:

$$P(t+1) = P(t) + \alpha(\varepsilon(t)/2)E(t)$$

$$P(t+1) = P(t) + \alpha(\text{epsulon}(t)/2)E(t)$$

$$\text{epsulon}(t) = d(t) - \text{flimiar}(S(t))$$

E o alpha-LMS utiliza:

$$P(t+1) = P(t) + \alpha(\text{epsulon}(t)E(t)/(|E(t)|^2))$$

$$\text{epsulon}(t) = d(t) - S(t)$$

Onde E é o vetor de entradas, P é o vetor de pesos, S é a saída, alpha é o coeficiente de aprendizado, epsulon o erro e d é a saída esperada.

A topologia de uma rede Adaline é baseada em camadas de neurônios, cuja saídas servem para alimentar os neurônios das camadas subsequentes, como na Figura .

FIG. 2.5.10: Topologia da rede Adaline

A rede adaline tem como vantagem sua simplicidade, pois a entrada não necessita de um pré-processamento complexo e são poucos parâmetros a serem ajustados, mas ela possui a limitação de conseguir apenas reconhecer padrões mais simples, necessitando de adaptações para análises mais complexas.

2.6 BACKPROPAGATION

O backpropagation, é um modelo composto por três ou mais camadas de adalines interconectados, que utilizam como função limiar a sigmoid. Por utilizar uma função contínua, o poder de classificação da rede é melhorado.

A atualização dos pesos neste modelo, é feita iniciando-se pelo erro da saída da última camada, e a partir daí são calculados os erros das camadas anteriores deste fato vem o nome backpropagation (retropropagação).

O erro da última saída é dado por $\epsilon_i(t) = S(t)(1-S(t))(d_i(t)-S(t))$ e com ele são ajustados os pesos da última camada por $P_i(t+1) = P_i(t) + \alpha \epsilon_i(t) E_n(t)$. À partir dele os erros das camadas anteriores podem ser calculados com $\epsilon_i(t) = E_n(t)(1-E_n(t)) \sum_k \epsilon_k p_{ik}(t)$, onde k são os neurônios conectados à seguir ao nó atual e p_{ik} é o peso associado a esta ligação. E com este erro recalcula-se os pesos das camadas intermediárias $P_i(t+1) = P_i(t) + \alpha \epsilon_i(t) E_n(t) + \mu(P_i(t) - P_i(t-1))$ onde μ é o momentum, coeficiente de aceleração do aprendizado.

A topologia do backpropagation é um adaline expandido com camadas intermediárias como na Figura.

FIG. 2.6.11: Topologia Backpropagation

A vantagem do modelo é que ele abrange uma grande quantidade de padrões que podem ser utilizados, e por este motivo foi escolhido para este trabalho.

3 O PÔQUER

3.1 O JOGO

O pôquer é um jogo de cartas jogado por duas ou mais pessoas, no qual os jogadores com as cartas total ou parcialmente escondidas fazem apostas em um pote central e após o fim das apostas, o pote central é atribuído ao jogador que possuir o melhor conjunto de cartas dentre os que permaneceram na jogada, ou ao único que restou caso todos os outros tenham desistido. Pôquer é um banco de ensaio interessante para pesquisas em inteligência artificial, pois é um jogo de informações imperfeitas, onde os múltiplos agentes concorrentes têm de lidar com conhecimento probabilístico, avaliação de risco e a possível fraude, assim como são as decisões feitas no mundo real. A modelagem do oponente é um outro problema difícil em aplicações de tomada de decisão, sendo é essencial para alcançar alto desempenho no pôquer.

3.2 O BARALHO

O pôquer é jogado com um barulho comum, de 52 cartas, divididas em 4 grupos, chamados naipes. Os naipes são identificados por nomes e símbolos, são eles: ouro, copas, paus e espadas. No pôquer os naipes possuem o mesmo valor, ou seja, o naipe da carta nunca é critério de desempate no pôquer. Cada naipe possui 13 cartas em ordem crescente de valor: números de 2 a 10, o valete (J), a dama (Q), o rei (K) e o ás (A), como é ilustrado em FIG. 3.2.1



FIG. 3.2.1: naipe de copas

3.3 RANKING DE COMBINAÇÕES DE CARTAS

O ranking das combinações de 5 cartas no pôquer é exibido a seguir em ordem crescente de valor:

Carta maior: quando não há nenhuma combinação, para efeito de desempate, ganha o que possuir a carta de maior valor (Figura FIG. 3.3.2).



FIG. 3.3.2: Carta maior

Par: Um par de cartas, isto é, duas cartas iguais. Se mais de um jogador tiver um par, então ganha quem tiver o par de cartas de maior valor. Caso ambos tenham o mesmo par, então aquele que possuir a carta de maior valor fora do par é o vencedor. Essa carta de desempate é chamada *kicker* (Figura FIG. 3.3.3).



FIG. 3.3.3: Par de dez

Dois pares: Dois pares de cartas. Se mais de um jogador tiver um dos pares iguais, então ganha quem tiver o segundo par com o valor mais alto. Caso ambos tenham os mesmos dois pares, então aquele que tiver a quinta carta (*kicker*) de maior valor é o vencedor (Figura FIG. 3.3.4).



FIG. 3.3.4: Dois pares de sete com quatro

Trinca: Três cartas iguais. Se mais de um jogador tiver esse jogo, então ganha quem tiver a trinca com uma carta de maior valor (Figura FIG. 3.3.5).



FIG. 3.3.5: Trinca de cinco

Sequência (*Straight*): Cinco cartas em sequência. O ás pode ser a maior carta ao lado do rei ou a menor carta ao lado do 2. Se mais de um jogador tiver uma sequência, então ganha aquele que tiver a sequência com cartas de maior valor. Caso ambos tenham a mesma sequência, então é considerado empate (Figura FIG. 3.3.6).



FIG. 3.3.6: Sequência de dez a ás

Flush: Cinco cartas do mesmo naipe, mas não em sequência. Se mais de um jogador tiver um *flush*, então ganha aquele que tiver a carta de maior valor (Figura FIG. 3.3.7).



FIG. 3.3.7: *Flush* de espadas

Full House: Um par junto com uma trinca de cartas. Se mais de um jogador tiver um *full house*, então ganha aquele que tiver a trinca com cartas de maior valor (Figura FIG. 3.3.8).



FIG. 3.3.8: *Full House* de dama com dois

Quadra (*Four*): Quatro cartas iguais. Se mais de um jogador tiver uma quadra, então ganha aquele que tiver a quadra com uma carta de maior valor (Figura FIG. 3.3.9).



FIG. 3.3.9: Quadra de rei

Straight Flush: Sequência de cinco cartas do mesmo naipe. Se mais de um jogador tiver um *straight flush*, então ganha aquele que tiver a sequência com cartas de maior valor (Figura FIG. 3.3.10).



FIG. 3.3.10: *Straight Flush*

Royal Flush: As cinco cartas de maior valor do mesmo naipe. Um Royal Straight Flush é simplesmente um straight flush com um Ás como carta mais alta. É a combinação mais forte do pôquer (Figura FIG. 3.3.11).



FIG. 3.3.11: *Royal Straight Flush*

3.4 AÇÕES

No pôquer, os jogadores possuem disponíveis as seguintes ações: passar (*check*), desistir (*fold*), apostar (*bet*), pagar (*call*) e aumentar (*raise*).

Um jogador ao passar está apenas passando a vez a outro jogador, e não renunciando ao pote central. Se ainda não houver apostas na rodada atual, o jogador pode passar a vez, e se todos os jogadores passarem na rodada, então a rodada é considerada completa.

Um jogador pode desistir da jogada, renunciando o interesse ao pote central.

Caso nenhuma aposta tenha sido feita ainda na rodada atual, um jogador pode apostar. Se um jogador apostar, o jogador adjacente (no sentido horário) tem três opções: desistir, pagar ou aumentar. A rodada só termina depois que todos os jogadores tenham igualado suas apostas, ou caso nenhum jogador pague a aposta feita por outro jogador, que se torna o vencedor da rodada.

Um jogador paga quando iguala a aposta feita por outro jogador na rodada.

Quando um jogador não só iguala a aposta de outro jogador, mas acrescenta um tanto mais, ele está aumentando a aposta. Todos os jogadores subsequentes deverão pagar o maior valor apostado, ou aumentar ainda mais, se quiserem continuar na jogada.

Quando a última aposta ou aumento na rodada final for igualada, ocorre a abertura de cartas. É nesse momento que se determina o vencedor do pote, uma vez que as combinações de cartas de cada jogador são mostradas uma a uma.

3.5 TIPOS DE PÔQUER

Existem diversas variantes do pôquer, dentre as quais pode-se categorizar livremente em: pôquer aberto, pôquer fechado e *community card poker*, dentre outras. O estilo de pôquer conhecido como Texas-Hold'ém é um tipo de *community card poker*, sendo atualmente o mais conhecido e jogado nos cassinos e torneios, e conseqüentemente com maior quantidade de estudo acerca do mesmo, fator pelo qual será utilizado como base no presente trabalho.

3.5.1 TEXAS HOLD'EM

No *Texas Hold'em* os jogadores recebem duas cartas fechadas e são dispostas cinco cartas fechadas sobre a mesa. Para vencer, o jogador deve combinar suas duas cartas fechadas com três das cinco abertas ou a melhor combinação, podendo usar as cinco cartas comunitárias. Ao contrário do pôquer aberto tradicional, as cartas abertas são as mesmas para todos os participantes. Existem um total de quatro rodadas de apostas. Antes do início das rodadas de apostas, os dois jogadores à esquerda do crupiê (*dealer*) devem fazer apostas obrigatórias, conhecidas como *Big Blind* e *Small Blind*. O *Small Blind* paga metade do *Big Blind*.

Na primeira rodada cada jogador recebe duas cartas fechadas e é feita a primeira rodada de apostas. O jogador à esquerda do *Big Blind* poderá desistir, pagar a aposta do *Big Blind* para entrar na jogada, ou aumentar a aposta. Os outros jogadores também optam por umas destas três ações.

Ao final da primeira rodada de apostas, o crupiê apresenta o *Flop*, isto é, abre as três primeiras cartas na mesa e inicia-se a segunda rodada de apostas pelo jogador à esquerda do crupiê.

Em seguida o crupiê apresenta o *Turn*, isto é, abre a quarta carta na mesa, e inicia-se a terceira rodada de apostas.

Na última rodada, o crupiê apresenta o *River*, isto é, abre a quinta e última carta fechada na mesa, e inicia-se a quarta e última rodada de apostas. Após o término da última rodada de apostas, os jogadores mostram seus jogos (*showdown*) a começar pelo último que apostou ou que aumentou alguma aposta. Somente este jogador tem a obrigação de mostrar seu jogo, os demais podem deixar suas cartas escondidas se perceberem que não vencerão.

Existem algumas variações no *Texas Hold'em* quanto ao limite máximo no valor de uma aposta. O *No Limit Hold'em* permite o chamado *all-in*, ou seja, apostar tudo que se tem. Entretanto, nenhuma aposta pode ser inferior ao valor apostado para entrar na partida.

4 ESTRUTURA DA SOLUÇÃO

4.1 MODELAGEM DO Oponente

Em primeiro lugar, apresentamos a topologia básica da rede neural básica utilizada para perfilar o oponente (assumindo uma mesa com tamanho máximo de 8 jogadores) nas tabelas 1 e 2. Note que para simplificar a modelagem matemática abordagem, foram ignorados alguns fatores que são de potencial influência sobre a ação do oponente, tal como a posição do mesmo na mesa, uma vez que a utilização da mesma implicaria em um encadeamento de previsões para prever as ações dos últimos jogadores na fila. Outra simplificação a ser realizada na classificação no decorrer de um jogo é assumir que as probabilidades de melhor mão estão identicamente distribuídas dentre os jogadores cujas mãos são desconhecidas e que estão remanescentes no jogo - deste modo, mesmo quando não se sabem as cartas do agente, pode-se utilizar alguma métrica de modo a estimar as ações do oponente. As saídas desta rede neural são representativas do perfil do jogador.

Parâmetros de entrada	Domínio
Probabilidade de a mão ser a melhor da mesa (π)	$0 \leq \pi \leq 1$
Inverso da quantidade de fichas possuída pelo jogador ($1/s$)	$0 \leq 1/s \leq 1$
Quantidade de jogadores ($n/8$)	$0 < n/8 \leq 1$
Inverso da quantidade de fichas possuída pelo i-ésimo oponente ($1/s_i$)	$0 \leq 1/s_i \leq 1$
Inverso da quantidade de fichas no pote ($1/p$)	$0 < 1/p < 1$
Fase do jogo (em ordem, pre-flop, flop, turn, river) ($\phi/4$)	$0 \leq \phi \leq 1$

TAB. 4.1.1: Parâmetros de entrada da RN utilizada para modelar o oponente

Parâmetros de entrada	Domínio
Fold?	$f_r \in \{0, 1\}$
Call?	$k_r \in \{0, 1\}$
Raise?	$c_r \in \{0, 1\}$
Fração do pote a ser apostada em raise	$0 \leq r/p \leq 1$

TAB. 4.1.2: Parâmetros de saída da RN utilizada para modelar o oponente

4.2 REDE PRINCIPAL E CLASSIFICAÇÃO EM PERFIS

A topologia da rede principal é similar, contudo, não há necessidade de se estimar a probabilidade de haver a melhor mão, uma vez que presumimos que se sabem as cartas na mão. Para cada potencial oponente (até um máximo de 7) serão adicionadas as informações de perfil adquiridas nos perfis que melhor se encaixarem às informações da tabela 4.1. As jogadas utilizadas para classificar o perfil seguirão um esquema de janela deslizante, de modo a evitar que um oponente que troque de estratégia com frequência seja constantemente classificado de forma incorreta. Deste modo, também descrevemos, uma outra rede de modo a classificar os jogadores em perfis, com sua topologia descrita nas tabelas 4.2.1 e 4.2.2.

Parâmetros de entrada	Domínio
Probabilidade de a mão ser a melhor da mesa (π)	$0 \leq \pi \leq 1$
Inverso da quantidade de fichas possuída pelo jogador ($1/s$)	$0 \leq 1/s \leq 1$
Quantidade de jogadores ($n/8$)	$0 < n/8 \leq 1$
Inverso da quantidade de fichas possuída pelo i-ésimo oponente ($1/s_i$)	$0 \leq 1/s_i \leq 1$
Inverso da quantidade de fichas no pote ($1/p$)	$0 < 1/p < 1$
Fase do jogo (em ordem, pre-flop, flop, turn, river) ($\phi/4$)	$0 \leq \phi \leq 1$
Oponente deu fold?	$f_r \in \{0, 1\}$
Oponente deu call?	$k_r \in \{0, 1\}$
Oponente deu raise?	$c_r \in \{0, 1\}$
Fração do pote que foi apostada em raise	$0 \leq r/p \leq 1$

TAB. 4.2.3: Parâmetros de entrada da RN utilizada para classificar o oponente para cada uma das jogadas em uma janela deslizante

Parâmetros de entrada	Domínio
Melhor perfil (i)	$0 < i/n \leq 1$

TAB. 4.2.4: Parâmetros de saída da RN utilizada para classificar o oponente

Para representar a estrutura de classes da solução, utilizaram-se os seguintes diagramas de classe apresentados nas figuras, descrevendo as classes utilizadas na API de aprendizado e na interface com o banco de dados de mãos.

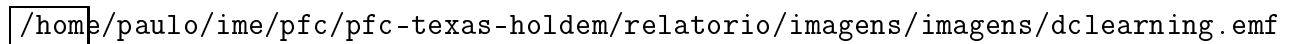
/home/paulo/ime/pfc/pfc-texas-holdem/relatorio/imagens/imagens/dclearning.emf

FIG. 4.2.1: Diagrama das classes utilizadas no aprendizado

5 SELEÇÃO DE JOGADORES PARA TREINAMENTO DA REDE

Para realizar o treinamento da rede, serão usados um ou mais banco de dados de mãos de jogos já existentes, que serão unificados em uma única estrutura de dados apresentada na figura 5.1 e inseridos em um banco de dados MySQL. Para realizar a interface com o banco de dados, utilizar-se-á a estrutura de classes descritas na figura 5.2.

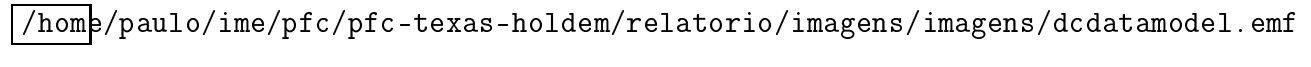
/home/paulo/ime/pfc/pfc-texas-holdem/relatorio/imagens/imagens/dcdatamodel.emf

FIG. 5.0.1: Diagrama de classe para as classes básicas

Uma vez obtidos os dados, faz-se necessário utilizar alguma heurística para selecionar os melhores jogadores. Utilizaremos como métrica inicial a quantidade de fichas adquiridas pelo jogador no decorrer do tempo visto. Também é importante seccionar os jogadores em diferentes perfis para treinar diferentes modelagens de oponente.

6 CONCLUSÃO

Da pesquisa bibliográfica realizada, concluímos que é factível a utilização de redes neurais para a simulação e análise de oponentes no jogo de pôquer, dando a entender que é promissora a possibilidade de utilização de uma rede neural como agente *stand-alone* do jogo.

7 REFERÊNCIAS

- [1] Salim, Rohwer. **Poker Opponent Modeling** - Indiana University, 2006.
- [2] Billings, Papp et al. **Opponent Modeling in Poker** - AAAI, University of Alberta, Edmonton, Canada, 1998.
- [3] Davidson, Billings, Schaeffer, Szafron. **Improved Opponent Modeling in Poker** - University of Alberta, Edmonton, Canada, 2000.
- [4] Johnson. **Robust Strategies and Counter-Strategies: Building a Champion Level Computer Poker Player** - University of Alberta, 2007.
- [5] Nicolai, Hilderman. **No-Limit Texas Hold'em Poker Agents Created with Evolutionary Neural Networks** - IEEE Symposium on Computational Intelligence and Games, 2009.
- [6] B. Beattie, G. Nicolai, D. Gerhard, and R. J. Hilderman. **Pattern classification in no-limit poker: A head-start evolutionary approach** - Canadian Conference on AI, 204–215, 2007.
- [7] Russell S., Norvig P. - **Artificial Intelligence: A Modern Approach (3rd Edition)**, 20??.