

UNIVERSIDADE FEDERAL DE GOIÁS  
ESCOLA DE ENGENHARIA ELÉTRICA E DE  
COMPUTAÇÃO

KLEBER PULLIG CORRÊA

**Estudo Empírico de Avaliação Inteligente  
Computadorizada de Questões Abertas  
Baseado em Colaboração**

Goiânia

2009

KLEBER PULLIG CORRÊA

**Estudo Empírico de Avaliação Inteligente  
Computadorizada de Questões Abertas  
Baseado em Colaboração**

Dissertação apresentada ao Programa de Pós-Graduação da Escola de Engenharia Elétrica e de Computação da Universidade Federal de Goiás, como parte dos requisitos para obtenção do título de Mestre no Programa de Mestrado em Engenharia Elétrica e de Computação.

**Área de Concentração:** Engenharia de Computação

**Linha de Pesquisa:** Sistemas Inteligentes

**Orientador:** Professor Paulo César Miranda Machado, Dr.

Goiânia

2009

**Dados Internacionais de Catalogação na Publicação (CIP)**  
**GPT/BC/UFG**

C823e Corrêa, Kleber Pullig.  
Estudo empírico de avaliação inteligente  
computadorizada de questões abertas baseado em  
colaboração [manuscrito] / Kleber Pullig Corrêa. - 2009  
xx, 106 f. : il. figs, tabs.

Orientador: Prof. Dr. Paulo César Miranda Machado.  
Dissertação (Mestrado) – Universidade Federal de  
Goiás, Escola de Engenharia elétrica e de Computação,  
2009.

Bibliografia.

Inclui lista de figuras, abreviaturas, siglas, quadros,  
equações e tabelas.

1. Computação – Redes neurais. 2. Sistemas de  
Avaliação Inteligente. 3. Inteligência Artificial. I. Título.

CDU: 004.89



Universidade Federal de Goiás  
Escola de Engenharia Elétrica e de Computação  
Coordenação do Programa de Pós-Graduação em  
Engenharia Elétrica e de Computação



## FOLHA DE APROVAÇÃO

# “Estudo Empírico de Avaliação Inteligente Computadorizada de Questões Abertas Baseado em Colaboração”

**KLEBER PULLIG CORRÊA**

Dissertação defendida e aprovada pela banca examinadora constituída pelos senhores:

---

Professor Dr. Paulo César Miranda Machado - Orientador (EEEC/UFG)

---

Professor Dr. Lauro Eugênio Guimarães Nalini - (UCG)

---

Professor PhD. Weber Martins - (EEEC/UFG)

---

Professor Dr. Rodrigo Pinto Lemos - (EEEC/UFG)

---

Professor Dr. Claudio Afonso Fleury - (UCG)



Todos os direitos reservados. É proibida a reprodução total ou parcial do trabalho sem autorização da universidade, do autor e do orientador.

### **Kleber Pullig Corrêa**

Bacharel em Ciência da Computação pelo Instituto Unificado de Ensino Superior Objetivo - IUESO (2004). Especialista em Tecnologia da Informação (2005), cursou o Mestrado na Escola de Engenharia Elétrica e de Computação (EEEC), na linha de pesquisa de Sistemas Inteligentes sob orientação do professor Dr. Paulo César Miranda Machado, com conclusão em Setembro de 2009. Atualmente é professor universitário e consultor na iniciativa privada.

*“Dedico os méritos desta conquista a minha esposa Alessandra, pelo apoio incondicional durante os anos dedicados a esta obra, provando mais uma vez seu inestimável valor de companheira e amiga.”*

## Agradecimentos

---

Agradecimentos especiais aos meus filhos: Ingryd, Evellyn e Juan, pelo carinho e compreensão nos dias de minha ausência.

Agradeço aos meus irmãos Alexandre e Danielle, que de alguma forma sempre se fizeram presente a cada etapa transposta.

Agradeço aos meus pais, Édine Calderam Corrêa (em memória) e Lúcia Maria Pullig Corrêa, que privando a si mesmos, legaram à seus filhos o acesso a educação.

Agradeço a meu diretor Dr. Carlos Alberto, pelo apoio, compreensão e exceções concedidas.

Agradeço ao professor Dr. Ademir Aparecido do Prado, pela cooperação e participação efetiva nesta pesquisa durante o processo de coleta de dados.

Agradeço ao amigo e professor PhD. Weber Martins, por sua postura irrepreensível e dedicação incondicional.

*“A mente que se abre a uma nova idéia jamais volta ao seu tamanho original.” (Albert Einstein)*

## Sumário

Lista de Figuras.....	xiv
Lista de Tabelas.....	xv
Lista de Quadros.....	xvi
Lista de Equações.....	xvii
Lista de Siglas e Abreviaturas.....	xviii
Resumo.....	xix
Abstract.....	xx
<b>Capítulo 01: Introdução.....</b>	<b>01</b>
1.1 Visão Geral.....	01
1.2 Justificativa.....	02
1.3 Objeto.....	05
1.3.1 Tema.....	05
1.3.2 Problema.....	06
1.3.3 Hipótese Básica.....	06
1.3.4 Hipóteses Secundárias.....	06
1.4 Principais Contribuições.....	07
1.5 Estrutura dos Capítulos.....	07
<b>Parte I - Fundamentação</b>	
<b>Capítulo 02: Avaliação da Aprendizagem.....</b>	<b>09</b>
2.1 Introdução.....	09

		11
2.2	Evolução Histórica.....	10
2.2.1	Breve Histórico.....	10
2.2.1.1	Corte Imperial.....	12
2.2.1.2	Rio de Janeiro (1865-1889).....	13
2.2.2	Avaliação e Poder.....	14
2.2.3	Avaliação e Exclusão.....	15
2.2.4	Avaliação e Fracasso Escolar.....	16
2.2.5	Erro no Processo de Aprendizagem.....	17
2.3	Avaliação.....	19
2.4	Funções da Avaliação.....	20
2.5	Tipos de Avaliação.....	23
2.5.1	Avaliação Somativa ou Certificadora.....	23
2.5.2	Avaliação Diagnóstica.....	24
2.5.3	Avaliação Formativa.....	25
2.6	Tipos de Questão.....	26
2.6.1	Visão Geral.....	26
2.6.2	Questões Abertas (dissertativas).....	27
2.6.3	Questões Fechadas (objetivas).....	30
2.6.3.1	Questões Dicotômicas.....	30
2.6.3.2	Questões de Evocação simples.....	32
2.6.3.3	Questões de Lacuna ou Complemento.....	32
2.6.3.4	Questões de Múltipla Escolha.....	34
2.6.3.5	Questões de Ordenação.....	36
2.6.3.6	Questões de Correspondência ou Pareamento.....	36
2.7	Divergência De Notas Entre Júris .....	38
2.8	Conclusão.....	40

<b>Capítulo 03: Redes Neurais Artificiais.....</b>	<b>42</b>
3.1 Introdução.....	42
3.2 Conceito.....	43
3.3 Breve Histórico.....	45
3.3.1 Marco Inicial.....	45
3.3.2 Perceptron.....	47
3.4 Modelos Booleanos.....	50
3.4.1 Neurônios RAM.....	50
3.4.2 Neurônios PLN.....	52
3.4.3 Neurônios GSN.....	53
3.5 Modelos Tradicionais de Redes Neurais.....	59
3.5.1 Mapas Auto Organizáveis.....	59
3.5.2 Muti Layer Perceptron.....	61
3.5.2.1 Conceito.....	61
3.5.2.2 Arquitetura.....	62
3.5.2.3 Treinamento.....	64
3.6 Desenvolvimento de Sistemas Conexionistas.....	68
3.7 Aplicações de Redes Neurais.....	70
<b>Capítulo 04: Sistema de Avaliação Inteligente Conexionista e Colaborativa.....</b>	<b>73</b>
4.1 Introdução.....	73
4.2 Conceitos Fundamentais.....	74
4.2.1 Definição de Termos.....	74
4.2.2 Interpolação Linear.....	75
4.2.3 Coeficiente de Correlação Linear de Pearson.....	76

4.3	Arquitetura do Sistema.....	78
4.3.1	Fase de Elaboração.....	78
4.3.2	Fase de Execução.....	79
4.3.3	Fase de Ranking ou Julgamento.....	81
4.3.4	Fase de Correção.....	82
4.3.5	Fase de Síntese de Notas.....	83
4.4	Trabalhos Correlatos.....	83

## Parte II – Validação Empírica

<b>Capítulo 05: Experimentos e Resultados.....</b>	<b>85</b>	
5.1	Introdução.....	85
5.2	Amostra.....	85
5.3	Metodologia.....	86
5.3.1	Método de Abordagem.....	86
5.3.2	Método de Procedimento.....	86
5.3.3	Técnicas.....	87
5.3.4	Variáveis.....	87
5.3.5	Material e Instrumentos.....	88
5.3.6	Delineamento Experimental.....	89
5.4	Resultados.....	89
5.4.1	Análise Descritiva.....	89
5.4.2	Análise Inferencial.....	91

	14
<b>Capítulo 06: Conclusão.....</b>	<b>93</b>
6.1 Introdução.....	93
6.2 Principais Contribuições.....	93
6.3 Discussão e Trabalhos Futuros.....	94
<b>Referências Bibliográficas.....</b>	<b>95</b>
Anexo I.....	105
Anexo II.....	106

## Lista de Figuras

---

Figura 2.1-	Distribuição Percentual de Frequência.....	30
Figura 3.1-	O Neurônio de McCulloch-Pitts.....	45
Figura 3.2-	Perceptron de Uma Camada Simples.....	47
Figura 3.3-	Estrutura do Neurônio RAM.....	51
Figura 3.4-	Exemplo de Arquitetura Piramidal.....	53
Figura 3.5-	Neurônio GSN.....	56
Figura 3.6-	Esquema da Rede SOM.....	59
Figura 3.7-	Processo de Redução de Vizinhança de uma Rede SOM.....	61
Figura 3.8-	MLP Típica com uma Camada.....	62
Figura 3.9-	Treinamento da Rede MLP.....	65
Figura 3.10-	Algoritmo <i>Backpropagation</i> .....	66
Figura 4.1-	Tela de Cadastro de Testes (questões e pesos).....	79
Figura 4.2-	Tela de Aplicação de Teste.....	80
Figura 4.3-	Tela de Registro de Respostas (etapa 01).....	80
Figura 4.4-	Tela de Registro de Ranking (etapa 02).....	81
Figura 4.5-	Tela de Correlação Tutor Humano.....	82
Figura 4.6-	Tela de Notas Inferidas pelo SAICO.....	83
Figura 4.7-	Tela de Síntese de Notas.....	84
Figura 5.1-	Grau de Satisfação: Software (SAICOweb).....	90
Figura 5.2-	Grau de Satisfação: Método de Avaliação.....	90
Figura 5.3-	Grau de Satisfação: Experiência Pedagógica.....	90
Figura 5.4-	Grau de Satisfação: Anonimato.....	91
Figura 5.5-	Análise de Desempenho.....	92

## Lista de Tabelas

---

Tabela 1- Tipos de Questões - McKenna e Bull (1999).....	27
Tabela 2- Interpretação do Valor do Coeficiente de Pearson.....	78

## Lista de Quadros

---

Quadro 1- Funções da Avaliação Como Medida.....	22
Quadro 2- Divergência de Notas entre Júris - Noizet e Caverni (1985).....	38
Quadro 3- Divergência de Notas entre Júris - Corretores e Provas.....	39
Quadro 4- Divergência de Notas entre Júris - Desvio.....	40

## Lista de Equações

---

Equação 1- Neurônio tipo RAN.....	52
Equação 2- Posição Média no Ranking.....	75
Equação 3- Correlação.....	77

## Lista de Siglas e Abreviaturas

---

SAICO	Sistema de Avaliação Inteligente Conexionista e Colaborativa
MLP	<i>Multilayer Perceptron</i>
QEA	Questionário de Estilo de Aprendizagem
VSLI	<i>Very Large-Scale Integration</i>
RNA	Rede Neural Artificial
XOR	“OU” Exclusivo
DARPA	<i>Defense Advanced Research Projects Agency</i>
ICNN	<i>International Conference on Neural Networks</i>
INNS	<i>International Neural Networks Society</i>
TNN	<i>Transactions on Neural Networks</i>
NC	<i>Neural Computation</i>
RAN	Neurônio Tipo RAN
PLM	<i>Probabilistic Logic Neuron</i>
GSN	<i>Goal Seek Neuron</i>
OCR	Reconhecimento Óptico de Caracteres
NQA	Número de Questões Avaliadas
NRR	Número de Respostas para <i>Ranking</i>
NT	Nota Tradicional
NI	Nota Inferida
PC	Ponto de Corte
PMR	Posição Média no <i>Ranking</i>
SOM	<i>Self Organized Map</i>

## Resumo

---

Esta pesquisa tem por objetivo submeter a testes empíricos em ambiente real, o **Sistema de Avaliação Inteligente Conexcionista e Colaborativa** idealizado por Martins (2004). Testado em ambiente de simulação por Guimarães (2004), observaram-se resultados consistentes e promissores enquanto tratando dados sintéticos. O sistema explora a Teoria de Aprendizagem de Kolb (1984), associando método de **correção colaborativa** e **inteligência artificial** (redes neurais do tipo perceptrons de múltiplas camadas MLP - *Multilayer Perceptron*), sendo capaz de aplicar e corrigir **avaliações computadorizadas** contendo inclusive **questões abertas (dissertativas)**.

Para viabilizar a coleta de dados em ambiente real, foram utilizados alunos matriculados em cursos de Engenharia da Universidade Federal de Goiás (UFG) no período de junho de 2008 a julho de 2009, tomando-se grupos heterogêneos compostos de homens e mulheres com idade média entre 16 e 26 anos e número mínimo de 18 alunos por avaliação. O software batizado de SAICOweb desenvolvido especificamente para este fim, permite aplicar **avaliações computadorizadas** de **questões abertas**, utilizando o método e as fases descritas na pesquisa de Guimarães (2004).

O Sistema SAICO foi validado empiricamente de forma satisfatória para uma única turma composta de 38 alunos (através do software SAICOweb em dois testes contendo 08 questões abertas cada). A análise dos dados, através da correlação (Coeficiente de Pearson,  $R_{xy}$ ), mostra forte correlação (0,81) entre notas atribuídas pelo SAICO e notas humanas (corrigidas pelo professor). As hipóteses secundárias foram confirmadas, expressas no grau de satisfação de professor e aluno ao avaliar os seguintes aspectos do experimento: Software SAICOweb, Método de Avaliação, Experiência Pedagógica e Anonimato. Pontos específicos foram identificados e sugeridos como trabalhos futuros, onde mais testes empíricos devem ser organizados para validar o sistema em diferentes ambientes.

**Palavras-chave:** Avaliação Computadorizada de Questões Abertas, Inteligência Artificial

Correção Colaborativa

## Abstract

---

This research submits to experimental tests in real-world environment, the Conexionist Collaborative Intelligent Evaluation System idealized by Martins (2004) and tested with simulations by Guimarães (2004), who have noticed consistent and promising results when dealing with synthetic datas. The system explores the Learning Theory by Kolb (1984), Peer Collaboration Evaluation and Artificial Intelligence (multilayer perceptrons neural networks), which are capable to apply and correct tests with open (free) questions.

To get a real-world data, some students from the course of Engineer registered at Universidade Federal de Goiás (UFG) on the period of june 2008 to july 2009 were taken in heterogeneous groups composed by men and women in age from 16 to 26 years old and at least 18 students per test. The software SAICOweb, developed specifically to this data collection, allows the computerized evaluations of open (free) questions, by using methods and stages described at Guimarães (2004) research.

SAICOweb system were experimental validited in a satisfactory way by an unique group composed by 38 students (by using two tests with 08 open questions each). The analisys of the results with Pearson correlation shows strong association (0,81) between grades given by SAICOweb and the human grades (corrected by professors). Another important result is the satisfaction of professor and students when evaluating Software SAICOweb, evaluation method, pedagogical experience and anonymity. Specific itens were identified and suggested as future updates, when more experimental tests should be designed to validade other aspects of the System.

**Key-words:** Computere's Evaluation of Open Questions, Artificial Intelligence, Collaborative Correction.

## Introdução

---

### 1.1 VISÃO GERAL

De acordo com o Ministério da Educação do Brasil, independente da estratégia pedagógica adotada, “*a verificação do rendimento escolar deve observar os seguintes critérios: avaliação contínua e cumulativa do desempenho do aluno, com prevalência dos aspectos qualitativos sobre os aspectos quantitativos e dos resultados ao longo do período*” (BRASIL, 1996). Com o principal interesse de promoção do aprendizado, a avaliação de desempenho deve verificar se o conhecimento (habilidade) foi adquirido pelo aluno.

Avaliar é tarefa de difícil execução e precisa garantir que o processo ensino-aprendizagem transcorra de forma objetiva e significativa na vida acadêmica do aluno (HAYDT, 2002).

O ritmo acelerado das transformações na era da Informática implica na necessidade do aperfeiçoamento da Educação frente à sociedade contemporânea. No contexto educacional, tal aperfeiçoamento visa ajustar a relação professor, aluno e conhecimento aos avanços tecnológicos, a partir de estratégias e métodos inovadores (LÉVY, 2004; CIVILETTI, RODRIGUES, LIMA, 2001).

A relação Educação-Informática evidencia os principais temas abordados nesta dissertação: avaliação computadorizada e questões abertas (dissertativas). Este trabalho testa empiricamente o Sistema de Avaliação Inteligente Conexional e Colaborativa - SAICO, concebido por Martins (2004) e testado em ambiente de simulação por Guimarães (2004).

O SAICO explora a Teoria de Aprendizagem de Kolb (1976), associando o método de correção automatizada baseada em colaboração e Inteligência Artificial (Redes Neurais do Tipo Perceptrons de Múltiplas Camadas MLP - *Multilayer Perceptron*), com o objetivo de aplicação e correção automatizada de avaliações computadorizadas contendo questões abertas.

A proposta deste trabalho é contribuir para o aprimoramento do processo de avaliação do aprendizado humano, destacando-se os seguintes objetivos específicos:

- Em ambiente real, avaliar a correlação entre notas atribuídas pelo tutor humano (professor) e o SAICO;
- Em ambiente real utilizando o SAICO, mensurar o grau de satisfação do professor e do aluno durante o processo de avaliação.

## 1.2 JUSTIFICATIVA

Métodos tradicionais de avaliação não se responsabilizam pela construção do conhecimento do aluno. Tais modelos permitem a “reprovação” em determinado conteúdo desde que ocorra a recuperação do aprendizado em outro. O aluno consegue “aprovação” no curso sem apresentar o conhecimento necessário de conteúdos não assimilados. A lacuna resultante limita o aprendizado futuro do aluno, sobretudo em matérias cumulativas (onde os primeiros conteúdos são pré-requisitos à compreensão de outros).

É possível diminuir os efeitos destes modelos avaliativos, privilegiando a utilização de avaliações diagnósticas e formativas (compostas por questões abertas), ou combinando ambas em intervalos reduzidos durante o curso (entre tópicos ou aulas). Tal estratégia permite ao professor e aluno melhor controle sobre a qualidade do ensino e aprendizado. Contudo, a eficiência desse procedimento aplicado a pequenos intervalos, seria contestada no meio educacional, dado que o uso de questões abertas é fator gerador de sobrecarga ao professor no momento da correção.

Para ilustrar o contexto descrito, analisaremos o seguinte exemplo: Um professor ministra quatro aulas por dia em quatro turmas compostas de 40 alunos cada. Caso opte por avaliá-los semanalmente utilizando testes contendo cinco questões abertas, ao final de cada bimestre, terá 1280 testes ou 6400 questões a serem corrigidas.

Obrigações do cotidiano docente, tais como: preparar planos de aula, buscar conhecimento atual, elaborar provas, trabalhos e corrigi-los, entre outras funções, direcionam o professor a eleger por conveniência o modelo avaliativo tradicional (mesmo ciente de que seu único propósito é medir e classificar ao término de cada módulo ou curso). Tal forma de avaliar normalmente emprega testes somativos, incapazes de diagnosticar, constituindo-se instrumento de mera classificação, conseqüentemente como ferramenta de exclusão.

A aplicação freqüente de testes formativos permite acompanhar a aprendizagem do aluno. Quando identificadas falhas no aprendizado, os exames fornecem subsídios para iniciar procedimentos corretivos antes do término do curso. Essa metodologia assegura que cada conjunto de tarefas foi assimilado antes de começar outras, induzindo o estudante a reformular seu método de estudo, melhorando-o gradualmente.

Pode-se afirmar que testes formativos ensinam, pois, ao estudar novamente o material, alunos elaboram e reafirmam conceitos baseados no processo de erro e acerto. Neste prisma, o teste apresenta-se como estratégia de ensino, verificando a extensão do aprendizado adquirido frente ao conteúdo ministrado.

Para atenuar a sobrecarga imposta ao docente e diminuir o tempo de correção durante o processo avaliativo, exames computadorizados contendo questões fechadas vêm sendo utilizados recentemente. Tais exames são aplicados e corrigidos por máquinas, com rapidez e exatidão. O processo de automação da aplicação de exames compostos de questões abertas é simples, mas como automatizar sua correção?

O trabalho apresentado nesta dissertação mostra que o uso do Sistema de Avaliação Inteligente Conexcionista e Colaborativa - SAICO automatiza com eficiência, a aplicação e correção de avaliações contendo questões abertas. Dentre os atributos que justificam esta pesquisa e o emprego do SAICO, destacam-se:

- Capacidade de aplicar avaliação computadorizada contendo questões abertas;
- Agilidade no processo de avaliação (sem limite máximo de alunos avaliados simultaneamente);
- Anonimato (controle das influências interpessoais entre professor e aluno durante todo o processo de aplicação e correção de exames);
- Agilidade na correção (a quantidade de alunos é fator secundário no tempo de correção);

- “*Feedback*” imediato ao aluno (o retorno rápido permite ao aluno reavaliar conceitos em tempo hábil de recuperação do conteúdo);
- Exercício de responsabilidade para o aluno (durante a utilização do módulo de correção colaborativa do SAICO);
- Experiência pedagógica de alto nível para o aluno (como parte ativa no processo de correção, oferta-se ao aluno a oportunidade de trabalhar e reelaborar seus conceitos sobre o conteúdo, momento em que a avaliação demonstra fortemente sua capacidade de ensinar) e,
- Possibilidade de situar o conhecimento do aluno em relação ao grupo (professor e aluno identificam falhas de aprendizado, a tempo de intervir e corrigi-las).

## 1.3 OBJETO

### 1.3.1 Tema

Avaliação somativa é atualmente o método mais utilizado para aferir o conhecimento adquirido pelo aluno ao término de cursos. Neste momento, a informação histórica é gerada podendo ou não ser aproveitada para melhorar turmas futuras.

A aquisição do conhecimento pelo aluno é um processo contínuo e deve ser avaliado regularmente, aplicando-se ações corretivas imediatas quando necessário. Durante o curso, mecanismos capazes de auxiliar o professor a obter tais informações, possibilitam identificar falhas de aprendizado a tempo de intervir adequadamente.

O assunto desta pesquisa, Sistema de Avaliação Inteligente Conexcionista e Colaborativa - SAICO é um método que sugere a necessidade de mudanças no atual modelo avaliativo, atuando como instrumento facilitador de tais adequações.

### **1.3.2 Problema**

O problema de interesse do presente projeto está colocado na seguinte questão:

- Em ambiente real, o SAICO apresenta resultados (notas) semelhantes à correção do tutor humano (professor)?

### **1.3.3 Hipótese Básica**

Baseado na pesquisa desenvolvida por Guimarães (2004) em ambiente de simulação na Universidade Federal de Goiás - UFG acredita-se que:

- Se SAICO é empregado em ambiente real; então a avaliação computadorizada contendo questões abertas exibe resultado expresso em notas com nível semelhante ao obtido na correção por tutor humano (professor);

### **1.3.4 Hipóteses Secundárias**

Dentre as hipóteses decorrentes, supondo o fortalecimento da hipótese básica, destacam-se:

- Se SAICO for empregado em ambiente real; então ocorre satisfação do docente envolvido no processo de avaliação;

- Se SAICO for empregado em ambiente real; então ocorre satisfação do discente envolvido no processo de avaliação;

## 1.4 PRINCIPAIS CONTRIBUIÇÕES

As contribuições a serem alcançadas através deste trabalho podem ser resumidas em duas principais:

- A primeira é atestar o alto potencial do SAICO quando aplicado em ambiente real, demonstrado empiricamente.
- A segunda é a verificar o grau de aceitação do SAICO por aluno e professor.

## 1.5 ESTRUTURA DOS CAPÍTULOS

O presente trabalho compõe-se de duas partes principais. A primeira parte diz respeito à fundamentação teórica (Capítulos 2 e 3) e ao sistema estudado (Capítulo 4). Na segunda parte registramos a validação empírica (Capítulo 5) e principais resultados obtidos (Capítulos 6).

O Capítulo 2 apresenta uma evolução histórica da avaliação, enfatizando avaliação computadorizada contendo questões abertas (dissertativas).

O Capítulo 3 apresenta uma revisão sobre Redes Neurais Artificiais, destacando redes do tipo perceptrons de múltiplas camadas (*MLP - Multilayer Perceptron*).

O Capítulo 4 descreve o Sistema de Avaliação Inteligente Conexionista de Colaborativa - SAICO: conceitos básicos, estrutura de funcionamento e trabalhos correlatos.

O Capítulo 5 detalha a amostra, métodos utilizados e resultados obtidos.

No Capítulo 6 conclui-se o trabalho, discutindo os principais resultados e propondo trabalhos futuros.

## Parte I - Fundamentação

# CAPÍTULO 2

## Avaliação da Aprendizagem

---

### 2.1 INTRODUÇÃO

Pode-se definir avaliação como a comparação constante entre os resultados obtidos e metas previstas. Avaliação é o processo de determinar a extensão com que objetivos se realizam, combinando descrição e julgamento. O processo avaliativo é o mecanismo que recolhe informações sobre algo ou alguém, fornecendo dados a serem utilizados para valorar os resultados alcançados (TYLER, 1949; SCRIVEN, 1967; GUBA, LINCOLN, 1981; MIRAS, SOLÉ, 1992; DE KETETLE, 1993).

Avaliação impõe a emissão de julgamento com objetivos previamente definidos: idéia, obra, solução, método ou material, sendo necessário estabelecer critérios ou padrões para apreciar as tarefas propostas e valorar de forma quantitativa ou qualitativa os alunos, buscando a verificação do cumprimento dos objetivos educacionais (BLOOM, HASTINGS, MADAUS, 1971).

Noizet e Caverni (1985) definem avaliação como processo de verificação de objetivos, que compara a produtividade do aluno a um modelo pré-estabelecido. Cada aspecto contextual pode ser objeto de avaliação, constituindo a avaliação da aprendizagem apenas parte do sistema avaliativo educacional.

O processo avaliativo contribui para a eficácia do ensino, consistindo na observação e interpretação de seus efeitos, para orientar decisões. Ao avaliar aplicam-se métodos de investigação sobre os acontecimentos acadêmicos do aluno, para o diagnóstico de possíveis falhas de aprendizagem (NEVO, 1990; CARDINET, 1993; LUCKESI, 2005).

Modalidades de avaliação como: exames tradicionais, testes de inteligência padronizados, avaliação por normas, por critérios e avaliação formativa servem a dois propósitos distintos: um voltado aos interesses da administração escolar e outro que busca objetivos educativos e pedagógicos (avaliação somativa e por normas - medida; e avaliação formativa e por critérios - gestão).

A avaliação como medida (somativa) atende a demanda institucional de verificação do grau de absorção do conhecimento, enquanto a avaliação como gestão (formativa e diagnóstica) responde às exigências da aprendizagem (AFONSO, 2000; BONNIOL, VIAL, 2001).

## **2.2 EVOLUÇÃO HISTÓRICA**

### **2.2.1 Breve Histórico**

Os primeiros relatos de avaliação na história datam de 2205 A. C., na China, nos quais os oficiais, eram examinados a cada 03 anos pelo imperador com a finalidade de promovê-los ou demiti-los (OLIVEIRA, SANTOS, 2005).

O termo avaliação da aprendizagem surgiu especificamente nos Estados Unidos da América no século XIX, durante a criação do sistema de testagem por Horace Mann.

Nesse período Mann submeteu 12 relatórios anuais ao *Board of Education of the Commonwealth of Massachusetts*, apresentando registros dos problemas educacionais correntes com suportes empíricos. Tal acontecimento foi um marco na avaliação formal (WORTHEN E SANDERS, 1987).

Como solução dos problemas Mann propôs substituir os exames orais pelos exames escritos; utilizando poucas questões gerais em vez de um número maior de questões específicas e buscando padrões mais objetivos do alcance escolar, sendo instituída a prova escrita. Teve início grande discussão sobre a possibilidade de testar programas educacionais em larga escala, promovendo melhorias no padrão educacional tradicional da época (PINTO, 1996; DEPRESBITERES, 1997).

A partir das primeiras décadas do século XX, nos Estados Unidos o movimento dos testes educacionais por Robert Thorndike, aclamado o pai do movimento da testagem educacional. Tecnologias para mensurar e determinar habilidades prosperaram e a testagem então surge como principal meio avaliativo (RODRIGUES, 2007).

Para desestabilizar a idéia de mensuração no processo avaliativo, surgiram os primeiros escritos de Ralph Tyler. Sua concepção contrapunha a idéia que avaliação é sinônimo de aplicação de testes usando lápis e papel. No entanto, vale ressaltar que Tyler (1981) não descartava a importância da aplicação dos testes, apenas acreditava existir outras maneiras de reconhecimento das mudanças comportamentais.

Outro nome importante na história da avaliação - Bloom (1971), responsável pela criação da taxionomia de objetivos educacionais, enfatizou a aprendizagem para o domínio, tendo em vista um sistema coerente de ensino e avaliação.

Fundamentadas nas idéias de Bloom surgem os primeiros escritos sobre a importância da aprendizagem contínua, para promoção do desenvolvimento pleno do indivíduo.

No Brasil, teorias e práticas da avaliação da aprendizagem reproduzem influências marcantes das produções de autores norte-americanos como pode ser verificado através de legislação concernente ao assunto (um exemplo claro é a Lei 5692/71).

### **2.2.1.1 Corte Imperial**

A avaliação do início do século era voltada à mensuração comportamental. Testes para medir, classificar e selecionar alunos, eram utilizados sistematicamente buscando averiguar aptidões e habilidades dos alunos. O conceito de eficiência estava pautado na aplicação de métodos de disciplina, controle, vigilância e punição, no intuito de tornar o estudante obediente e submisso à estrutura educacional vigente (TEIXEIRA, NUNES, 2008).

Ainda segundo as autoras bons resultados nos exames e cumprimento de regras de comportamento e disciplina recebiam recompensas como: lugar distinto na sala de aula; espargimento de flores; vivas escolásticas e progressão à classe imediata. Infrações como: atrasos, ausência, desobediência, indecência ou desatenções eram reprimidas, pela aplicação de avaliações (provas, exames, testes e exercícios), além de exposição pública como: permanecer de joelhos; prisão por um a três dias; repreensão ou expulsão.

Neste período foram elaborados parâmetros “educacionais” onde conhecimento e habilidades do professor lhe conferiam ‘*status*’ de superioridade, justificando a forma de tratamento dispensado aos alunos, surgindo o modelo de “vigiar e punir” (GONDRA, 2005).

Investimentos governamentais eram dirigidos para a formação de comportamentos coletivos dominantes, forjando uma ordem social de hábitos higiênicos, civilizados e homogêneos. Neste âmbito, escolas assumem papel de poder disciplinador e de regulamentação, intervindo diretamente nas práticas ordinárias e cotidianas de controle do indivíduo (FOUCAULT, 1987).

#### **2.2.1.2 Rio de Janeiro (1865 - 1889)**

Analisando o Estatuto do Município da Corte do Rio de Janeiro no período compreendido entre os anos de 1865 a 1889, verificam-se sensíveis transformações no conjunto político e cultural da época. Foi criado o Estatuto das Escolas de Instrução Primária para controlar as atividades do cotidiano escolar e regularizar práticas pedagógicas, sanções disciplinares, exercícios e exames.

O documento mencionava aspectos de modelagem do corpo, da moral e de pensamento, como meio de condução à docilidade. Cria-se uma espécie de política de submissão dos movimentos físicos e intelectuais dos alunos e professores (SILVA, 2006).

O documento trata das obras e publicações adotadas, determinando o uso de livros designados pelo Inspetor Geral. A eleição da literatura da época era fruto do processo de controle político que se utilizava da estrutura educacional para doutrinar, moralizar e disciplinar a sociedade com base em preceitos divulgados de conhecimento conveniente à sociedade “civilizada”.

Livros escolares continham as propostas políticas existentes, mensagens morais, religiosas, civis, científicas ou patrióticas gerando condições favoráveis para o surgimento, permanência e disseminação de obras direcionadas. (TEIXEIRA, 2005).

### 2.2.2 Avaliação e Poder

O controle das relações entre pessoas é sinônimo de força e domínio. As relações de poder iniciadas dentro do ambiente familiar se estendem à instituição escolar naturalmente. Preconceitos e desigualdades fazem parte do cotidiano educacional onde, conhecimento é traduzido por poder, separando indivíduos em classes sociais distintas.

Considerando a disposição física, mecanismos disciplinares, organização hierárquica e a vigilância constante, comparamos a atual estrutura educacional a uma prisão. Tal analogia refere-se ao sistema penitenciário comum, onde reformatórios disciplinares primam pela readaptação e integração de “corpos dóceis” a sociedade (BORGES, 2005).

Ao exercer pressão sobre alunos para que estudem, aprendam, efetuem tarefas, respeitem as normas e obedeçam a ordens, impõe-se um sistema disciplinador, punitivo e com fortes características “normalizadoras” (SILVA, 2006).

Nas considerações de Foucault (1990), o modelo do dispositivo disciplinar é aplicado da seguinte forma: cria-se um espaço fechado e monitorado, onde indivíduos são inseridos, seus movimentos controlados e registrados por uma figura hierárquica que exerce seu poder para examinar e classificar.

Pais enviam crianças à escola para se habituarem a permanecer sentadas observando pontualmente o que é ordenado o que impede o surgimento de novas idéias, reproduz e legitima o sistema de desigualdade social, fabricando indivíduos conformados e subordinados frente a uma sociedade de valores essencialmente meritocráticos (KANT, 1996; DUBET, 1998).

No modelo que privilegia unicamente a atribuição de notas e classificação dos estudantes, o processo avaliativo é uma poderosa arma de controle, enquanto uma avaliação produtiva precisa ser dialógica, interativa, fazendo do indivíduo um ser melhor, mais criativo, autônomo, e participativo.

### **2.2.3 Avaliação e Exclusão**

Durante séculos o modelo educacional apresentou função seletiva, onde o ápice era o ingresso de cinco por cento dos alunos na universidade, implicando em completo desinteresse pelos outros. Tal processo de seleção criou a concepção que apenas indivíduos com habilidades raras estavam aptos a completar o ensino e ingressar no programa universitário com êxito, distinguindo os fracassados daqueles que foram bem sucedidos, atribuindo maior importância ao processo que aos resultados (SARMENTO, 1997).

A avaliação do aprendizado deve estar delimitada pela teoria e prática educativa que o requisitou como instrumento de verificação de resultados, ocorrendo dentro de uma modelagem teórica do mundo e da educação, traduzido em prática pedagógica (LUCKESI, 2000).

Para avaliar adequadamente é necessário utilizar uma pedagogia construtiva, que considere o conhecimento como processo em edificação, primando pelo crescimento individual do educando, cabendo a instituição escolar desenvolver no aluno características que lhe permitirão viver de forma eficiente e produtiva na sociedade.

A avaliação só assume papel inclusivo se acolhe o aluno dentro do processo educativo, diagnosticando o que pode ser feito para o seu crescimento.

O modelo avaliativo almejado desencadeia um processo de construção social, que visa conhecer, interpretar e transformar a instituição de ensino e seus alunos (SOBRINHO, 2003; LUCKESI, 2005).

#### **2.2.4 Avaliação e Fracasso Escolar**

Tentativas frustradas de aprimorar procedimentos escolares produziram formas renovadas de insucesso acadêmico ao longo dos anos. Medidas adotadas no sentido de ampliar o tempo médio de escolarização e universalizar o acesso ao ensino são apontadas por especialistas como insuficientes, sendo responsáveis pelos altos índices de repetência e evasão escolar registrados atualmente no Brasil (SOARES, 1981; LUDKE, MEDIANO, 1992; LUCKESI, 1995).

Na tentativa de potencializar o êxito do aluno e promover inclusão, instituem-se práticas educativas deficientes e desprovidas de pedagogia adequada. Essas medidas enfatizam métodos mecanizados de ensino e tornam o processo educativo desumano, desconsiderando as características individuais de absorção do conhecimento pelo aluno (ESTEBAN, 2001).

È insensatez pensar na Educação como meio único e suficiente de transformação na sociedade contemporânea. Se ideais são necessários para vivenciar essa prática, os mesmos são incapazes de gerar mudanças, servindo apenas como rota viável aos agentes transformadores. A educação formal em si é inapta para transformar, estando limitada à reprodução de valores moldados e preservados ao longo dos anos pelas classes dominantes (GADOTTI, 1984).

Interesses da sociedade são definidos pela minoria detentora do poder, o conhecimento transmitido e ensinado é manipulado de forma conveniente à manutenção dos valores pré-existentes, evidenciando a instituição escolar como instrumento de exclusão e domínio (MATTEI, 2008).

### **2.2.5 Erro no Processo de Aprendizagem**

Alvos inatingíveis no presente devem servir de inspiração para lutas futuras. Tais metas clamam por novas investidas, excitando a imaginação e oferecendo ao indivíduo motivação para alcançá-las (KAPPEL, 1960).

O erro está presente na busca e assimilação de novos conhecimentos. Reconhecer o erro é complexo, pois o mesmo assombra o homem desde os primórdios. A educação futura deve estar preparada para enfrentar esse problema (MORIN, 2003).

Sucessivos erros foram cometidos até serem construídas concepções mais elaboradas. Do ponto de vista da sobrevivência humana, por meio das relações opostas entre si, surge conceito de certo, errado, bom, e ruim. Tais “conceitos padrões” são frutos de escolhas e foram construídos a partir de experiências vivenciadas (TEIXEIRA, NUNES, 2008).

A palavra erro adquire um peso enorme junto ao processo de aprendizagem do aluno, sendo levada ao pé da letra como sinônimo de falta de: inteligência, capacidade, atenção, família e muitas outras relacionadas ao fracasso escolar. As pessoas se habitam de tal maneira a conceitos pré-estabelecidos por ignorar a existência de outras formas de tradução do mundo. Compreendê-lo por outro ângulo, impele indivíduos de assimilar e vivenciar novos conceitos (CARVALHO, 2004).

A concepção culposa atribuída ao erro conduz ao uso permanente do castigo para correção e direcionamento da aprendizagem, conferindo a avaliação, a função de classificação. Atitudes ameaçadoras garantem a propagação do medo, ansiedade e vergonha.

O erro como fonte de castigo impossibilita a função de suporte ao ensino para crescimento intelectual do aluno. O ambiente escolar, que supostamente deveria primar pela disseminação dos ideais de inclusão, assume prática contraditória, no qual a avaliação escolar absorve a tarefa de controle e punição, atendendo essencialmente à necessidade de selecionar para excluir (LUCKESI, 2005).

É necessária a compreensão do erro de forma dinâmica, contrariando o padrão para em seguida possibilitar novamente a conduta de conformidade como caminho para o aprendizado. Cada indivíduo apresenta um nível de desenvolvimento a ser respeitado, segundo Freire (2003) *“Ninguém é ignorante por completo, ou seja, ninguém é tão burro que não saiba nada, e ninguém é tão inteligente que conheça tudo”*.

A avaliação deve ser direcionada para que professores interfiram nos resultados de forma dinâmica e criativa, induzindo o aluno à reformulação de suas idéias e hipóteses, promovendo novas formas de aprendizagem. Processos de aprendizagem precisam quebrar paradigmas frente ao erro, vislumbrando-o de forma positiva, como indicador de assimilação do conhecimento.

As avaliações são expressões da síntese do conhecimento que o aluno atingiu. Caso o nível esteja insatisfatório, o conteúdo precisa ser redirecionado para a elaboração de novas formas de aprendizagem e sem medidas punitivas, mesmo que seja necessário retomar ao ponto de partida (TEIXEIRA, NUNES, 2008).

## 2.3 AVALIAÇÃO

De incultos a especialistas, indivíduos avaliam e manifestam opiniões constantemente. De igual modo: professores, alunos, diretores e administração escolar avaliam-se mutuamente a cada instante. Entretanto, o único aspecto formalmente reconhecido dentro da estrutura educacional é a avaliação do aluno pelo professor (LUCKESI, 2005).

Avaliar a aprendizagem é importante para professor e aluno. O processo avaliativo fornece informações que direcionam o docente quanto à representação e construção dos objetivos a serem atingidos, permitindo ao professor estimar o grau de conhecimento absorvido pelo aluno durante o curso.

Caso o ato da avaliação falhe em seu propósito, a comparação dos resultados alcançados com o plano de objetivos estabelecidos anteriormente pelo professor é comprometida. De acordo com Hadji (2001), *“os efeitos da avaliação determinarão a construção do destino escolar do aluno, conduzindo-o futuramente ao êxito ou fracasso acadêmico”*.

Ao abranger vários objetos e objetivos simultaneamente, o processo avaliativo exige tomada de decisões significativas no decorrer do curso, tornando-se ato de difícil execução prática. A falta de qualificação adequada por docentes reproduz a utilização de instrumentos avaliativos de modo equivocado, impossibilitando constatar se houve assimilação do novo conhecimento por parte dos discentes (BLOOM, HASTINGS, MADAUS, 1971).

O ato avaliativo deve apresentar parâmetros definidos, coerentes com conteúdos estudados e objetivos pré-estabelecidos. As técnicas (instrumentos) de verificação do conhecimento adquirido devem respeitar a capacidade individual do aluno, servindo de subsídio ao professor para determinar se o aprendizado está satisfatório ou precisa ser redirecionado (CIVILETTI, RODRIGUES, LIMA, 2001).

## 2.4 FUNÇÕES DA AVALIAÇÃO

Atribuição de notas aos estudantes é o propósito mais conhecido da avaliação no âmbito educacional. São utilizados números ou letras dentro de uma escala de valores representando o desempenho do estudante diante dos objetivos previstos pelo professor ou pela instituição de ensino. Bloom, Hastings, Madaus (1971) destacam outras aplicações comuns da avaliação:

- **Certificação de habilidades e capacidades** – o objetivo é garantir que naquele momento o aluno possui determinada habilidade ou capacidade.
- **Previsão de sucesso em cursos subseqüentes** - utilizada como orientação educacional, serve de parâmetro para prever o sucesso do aluno em cursos subseqüentes.
- **Definição do ponto inicial da instrução em um curso subseqüente** – fornece subsídios ao professor do ano subseqüente a fim de determinar o nível em que deve iniciar o ensino para a turma.
- **“Feedback” para os alunos** – situa o aluno sobre seu progresso durante o curso.
- **Teste de nível de um grupo** – o objetivo é determinar o nível de conhecimento do aluno.
- **Comparações dos resultados obtidos por grupos diferentes** – avalia resultados obtidos a partir de métodos e professores distintos com o objetivo de compará-los.

A avaliação indica se de fato estão ocorrendo mudanças no comportamento dos estudantes, bem como em que medida tais mudanças ocorrem. Caso a aprendizagem não seja eficiente, é possível programar mudanças para assegurar sua eficácia antes que seja tarde.

Um único teste utilizado pode servir para várias finalidades distintas. No intuito de potencializar sua eficiência, é necessário estipular regras e definir objetivos em termos de ações observáveis e plausíveis, informando aos professores quanto o ensino oferecido foi eficaz, contribuindo para a redefinição e o ordenamento dos conteúdos e adequação das estratégias de melhoria da aprendizagem.

De acordo com Natriello (1987) são quatro as funções básicas da avaliação: certificação, seleção, orientação e motivação. A certificação garante que o aluno atingiu determinado nível acadêmico. A seleção identifica os alunos aptos a prosseguirem os estudos. A orientação comunica ao aluno os resultados, permitindo o diagnóstico de falhas de aprendizagem para intervenção do avaliador. Os resultados da avaliação de modo geral asseguram a motivação e o empenho em novas atividades.

A função pedagógica da avaliação pode ser compreendida observando seus aspectos separadamente. Seguindo a definição de Pacheco (1994), o aspecto pessoal estimula o sucesso dos alunos na vida acadêmica. O aspecto didático, com fases de diagnóstico, visa a melhoria e a verificação dos resultados da avaliação. O aspecto curricular envolve a possibilidade de realizar adaptações curriculares diante das necessidades dos alunos. O aspecto educativo preocupa-se com a avaliação da qualidade da educação.

De modo geral, a fase do processo educativo onde ocorre a avaliação normalmente determina seu papel naquele momento. Durante a fase de planificação, a avaliação tem como função orientar este processo, enquanto que no decorrer da aprendizagem assume a função de regulação (LEMOS et al, 1993).

Vallejo (1979) afirma que as funções motivadoras e diagnósticas são as de maior importância da avaliação, a função diagnóstica liga-se a função motivadora visto que a avaliação reflete na motivação nos alunos. Tais considerações podem ser elucidadas pelo quadro 1 proposto pelo autor:

Quadro 1 - Funções da Avaliação Como Medida

<b>Funções da Avaliação como Medida</b>
<b>Clarifica os objetivos</b> <ul style="list-style-type: none"> <li>· Indica ao professor o que avaliar</li> <li>· Propõe objetivos realistas e avaliáveis</li> </ul>
<b>Identifica os Problemas</b> <ul style="list-style-type: none"> <li>· Qualidade da prova</li> <li>· Procedimento perante os erros individuais</li> <li>· Análise da avaliação</li> </ul>
<b>Motiva e Estimula Alunos e Professores</b> <ul style="list-style-type: none"> <li>· O que, como e quando o aluno estuda</li> <li>· O que e como o professor ensina</li> </ul>
<b>Sugere Novos Métodos</b> <ul style="list-style-type: none"> <li>· Exercícios</li> <li>· Material didático</li> </ul>
<b>Coordena Esforço</b> <ul style="list-style-type: none"> <li>· Impõe objetivos comuns</li> <li>· Facilita o intercâmbio de métodos e a colaboração na confecção da avaliação</li> </ul>
<b>Contribuição</b> <ul style="list-style-type: none"> <li>· Previsão de resultados futuros</li> <li>· Investigação (comparar métodos e grupos)</li> <li>· Supervisão dos professores</li> </ul>

## 2.5 TIPOS DE AVALIAÇÃO

A avaliação tem sido tradicionalmente concebida como algo que ocorre ao fim de um ciclo didático, mas no sentido atual deve ser vista como processo desenvolvido ao longo do curso, de forma contínua. Devemos entender a avaliação como juízo de qualidade sobre dados relevantes, tendo em vista a tomada de decisão. (LUCKESI, 1995).

Essa concepção de avaliação como processo contínuo e paralelo ao processo de ensino e aprendizagem, permite a utilização da avaliação através das seguintes modalidades: diagnóstica (no início do processo), formativa (no decorrer) e somativa (no final do processo).

### 2.5.1 Avaliação Somativa ou Certificadora

Esta modalidade avaliativa possui função classificatória, característica que torna comum sua utilização como forma de avaliar produtos, serviços ou processos. No âmbito acadêmico é usual sua aplicação ao final do curso, período letivo ou unidade de ensino para classificar os alunos de acordo com os níveis de aproveitamento. Normalmente os testes somativos são efetuados a cada quatro ou seis semanas para atribuir notas ao rendimento do aluno promovendo-o à série ou módulo subsequente (BLOOM, HASTINGS, MADDAUS, 1971).

Obter resultados satisfatórios ao utilizar essa modalidade de avaliação, requer que o docente determine criteriosamente os aspectos relevantes a serem valorados, estabelecendo critérios de representatividade e de importância do aprendizado.

Este tipo de avaliação possibilita verificar se objetivos estabelecidos foram alcançados pelos alunos, fornecendo dados para aperfeiçoar o processo de ensino-aprendizagem (HAYDT, 1997; HOPPER, 1998).

Ao utilizar modalidades avaliativas classificatórias para determinar o grau de conhecimento do aluno, desconsideram-se os aspectos subjetivos, lançando mão de preconceito em relação a outros meios de percepção da aprendizagem.

O modelo impõe injustificadamente, a classificação do aluno durante o processo de verificação do desempenho, deixando margem suficiente para o questionamento de sua adoção em detrimento de outros modelos avaliativos (TEIXEIRA, NUNES, 2008).

### **2.5.2 Avaliação Diagnóstica**

A principal função da avaliação diagnóstica é determinar a posição do aluno diante do conteúdo seqüencial de ensino a ser aplicado ao grupo (turma). A estratégia busca informações para homogeneizar o grupo antes de iniciar o curso, evitando o desinteresse dos alunos que possuem maior bagagem de conhecimento e frustração daqueles sem pré-requisitos mínimos (BLOOM, HASTINGS, MADAUS, 1971).

A avaliação diagnóstica está intimamente relacionada com a avaliação formativa e somativa, estando separada destas por dois propósitos: determinar a posição correta do aluno frente ao início da instrução e acompanhar possíveis deficiências (falhas) de aprendizagem durante a evolução do ensino.

A aplicação do método avaliativo diagnóstico segue uma seqüência lógica: valoração, descrição e classificação de determinado aspecto do conhecimento do aluno. Os esforços focam a instrução no intuito de localizar o ponto de partida mais adequado para o processo de ensino aprendizagem, utilizando como parâmetro informações obtidas a partir de testes somativos.

Outra função da avaliação diagnóstica é investigar razões de fracassos na prática educativa em qualquer momento do seu percurso (início, meio e fim), desta forma o teste serve para verificar como está o processo de construção do conhecimento e se os métodos estão surtindo resultados efetivos. A estratégia é válida na medida em que erros cometidos pela minoria do grupo são corrigidos individualmente e erros cometidos pela maioria considerados deficiências inerentes ao material ou processo de instrução.

### **2.5.3 Avaliação Formativa**

O objetivo da avaliação formativa é determinar o grau de domínio do aluno em relação a determinado conhecimento. Testes para fins de avaliação formativa são aplicados com maior frequência e enquanto a atividade é executada. O procedimento indica lacunas no conhecimento e auxiliam professor e aluno na recuperação específica do conteúdo, fornecendo subsídios a construção dos resultados parciais e para atingir objetivos pré-estabelecidos (WOLZ, 1997).

É comum professores disseminarem concepções equivocadas tais como: a) a avaliação formativa e a avaliação somativa distinguem-se através dos instrumentos utilizados; b) a avaliação formativa é subjetiva e a avaliação somativa é objetiva; e c) a avaliação formativa é toda e qualquer avaliação que se desenvolve dentro da sala de aula (BOAVIDA, 1996; FERNANDES et al, 1996). Daí a importância de esclarecer que o mesmo teste somativo aplicado para atribuição de notas, pode assumir função formativa quando utilizados em períodos curtos, ao final de pequenas unidades de ensino (BLOOM, HASTINGS, MADDAUS, 1971).

## 2.6 TIPOS DE QUESTÃO

### 2.6.1 Visão Geral

Conforme visto, as formas usuais de avaliar um grupo de alunos são: diagnóstica, formativa e somativa. Sua utilização requer conhecimento e habilidade por parte do docente durante o processo avaliativo.

O tipo de questão que irá compor o teste no momento de avaliar também deve ser cuidadosamente estudado, pois está diretamente relacionada ao tipo avaliativo e a metodologia de ensino adotada, que por sua vez responde pelo cumprimento ou não dos objetivos estabelecidos. Em termos gerais, o tipo de avaliação e de questão determina o resultado esperado pelo professor.

Segundo McKenna e Bull (1999), questões podem ser classificadas como: fechadas (objetivas) e abertas (dissertativas) dividindo-se em subtipos que apresentam características e funcionalidades próprias. Em avaliações contendo questões fechadas, o aluno responde a averiguação do professor de forma direta, sem argumentação, sendo mais adotada em virtude da agilidade e simplicidade durante o processo de correção. A Tabela 1 fornece uma visão geral da classificação de questões, detalhadas neste capítulo:

**Tabela 1:** Tipos de Questões - McKenna e Bull (1999)

Tipo	Sub-tipo
<b>Fechadas</b> (objetivas)	<ul style="list-style-type: none"> <li>- <b>dicotômica</b> (verdadeiro/falso ou certo/errado)</li> <li>- <b>evocação simples</b> (pergunta direta ou resposta curta)</li> <li>- <b>lacuna ou complemento</b></li> <li>- <b>múltipla escolha</b></li> <li>- <b>ordenação ou seriação</b></li> <li>- <b>conjugadas ou correspondência</b> (pareamento)</li> </ul>
<b>Abertas</b> (dissertativas)	<ul style="list-style-type: none"> <li>- <b>composição extensa</b> (crítica)</li> <li>- <b>composição curta</b> (resposta restrita)</li> </ul>

A utilização de questões fechadas requer maior conhecimento por parte do professor na elaboração da avaliação, podendo haver prejuízo dos resultados esperados quando: o aluno supõe na opção alternativa, mais que o previsto por quem elaborou a questão; sorte-azar; palpites e outros fatores de difícil controle, que podem comprometer a avaliação em sua totalidade (BUSH, 1999).

## 2.6.2 Questões Abertas (Dissertativas)

Questões abertas (dissertativas) normalmente são utilizadas quando se deseja avaliar o conhecimento dos alunos sobre assuntos ou temas onde respostas não podem ser medidas quantitativamente. Os avaliados respondem livremente sem se limitarem a escolher entre um rol de alternativas. Este tipo de avaliação exige que o aluno descreva com suas próprias palavras o que sabe a respeito do assunto abordado, descrevendo as relações que estabeleceu e os raciocínios que o levaram à compreensão.

**Vantagens:**

- É adequada para avaliar o raciocínio lógico, a capacidade de analisar, hierarquizar e sintetizar as idéias, a justificativa das opiniões, a clareza de expressão;
- Estimulam a cooperação;
- Permitem avaliar atitudes dos alunos em questões estruturadas;
- São úteis como primeira questão de determinado tema por deixarem o respondente mais à vontade;
- Cobrem pontos além das questões fechadas;
- Têm menor poder de influência nos respondentes que perguntas com alternativas previamente estabelecidas;
- Exigem menor tempo de elaboração;
- Proporcionam comentários, explicações e esclarecimentos significativos para interpretação e análise das perguntas com respostas fechadas;
- Evita-se o risco existente no caso das questões fechadas, do pesquisador deixar de relacionar alguma alternativa significativa no rol de opções;
- O aluno organiza os pensamentos forma lógica, com liberdade de resposta e as respostas não precisam necessariamente estar certas ou erradas, há graus de compreensão e precisão;

- É possível verificar a habilidade do aluno de organizar idéias, integrar princípios, expressar o próprio pensamento e criticar determinados sistemas;
- Força o aluno a responder o que sabe sobre a questão.

**Desvantagens:**

- São extremamente exaustivas frente à necessidade de serem submetidas ao processo de correção individual;
- Necessitam que sejam pré-estabelecidos critérios de valoração a fim de inibir à parcialidade do professor durante a compilação das notas;
- São onerosas e demoradas para serem analisadas se comparadas a outros tipos de questões;
- São influenciadas pela subjetividade do professor.

Estudo recente realizado com 270 alunos ingressantes dos cursos de administração, direito e psicologia de uma universidade particular do interior paulista, utilizou um questionário que investigava as estratégias de avaliação mais empregadas. Os resultados (Figura 2.4) evidenciaram que o tipo mais utilizado pelos docentes é a prova dissertativa individual e tal estratégia aparecia como a segunda mais citada pelos universitários como a forma que eles julgavam mais adequada para a avaliação de seu desempenho, estando atrás somente do trabalho dissertativo em grupo (OLIVEIRA, SANTOS 2005).

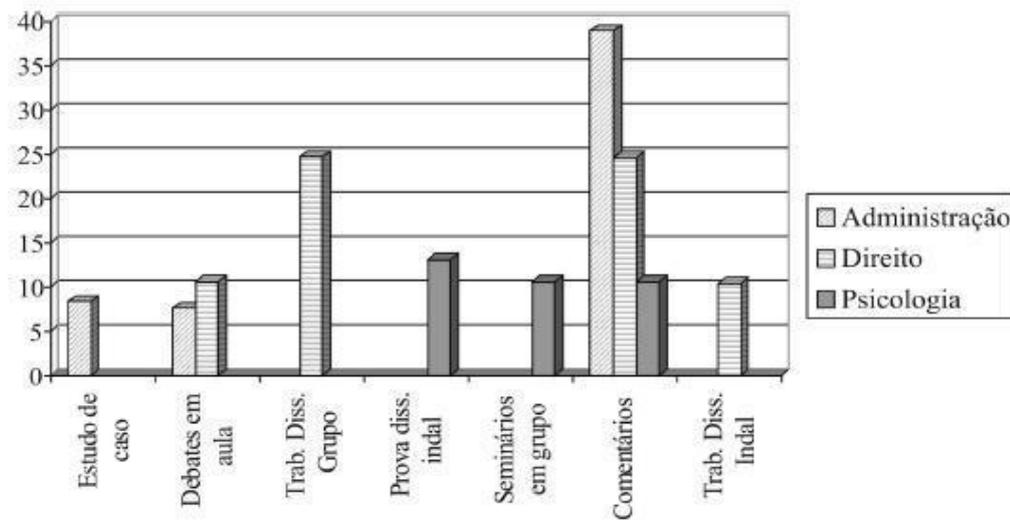


Figura 2.1 - Distribuição Percentual de Frequência

### 2.6.3 Questões Fechadas (Objetivas)

Provas objetivas satisfazem as características básicas da objetividade na avaliação, como brevidade da resposta e exatidão da correção. As respostas nestes tipos de prova são constituídas por um sinal, número, palavra ou frase.

#### 2.6.3.1 Questões Dicotômicas

Este tipo de questão exige que o aluno compare seu conhecimento ao fato, convenção, definição ou outro enunciado apresentado, semelhante a questões de múltipla escolha.

A elaboração dessas questões deve respeitar regras como: elaborar frases curtas; usar preferencialmente frases positivas; destacar frases negativas; evitar a inclusão de palavras ou expressões que lembram declarações falsas ou verdadeiras (nunca, nenhum, somente); evitar frases parcialmente certas ou parcialmente erradas (RODRIGUES, 2000).

Tais questões fornecem apenas duas alternativas de resposta: verdadeiro ou falso, sendo apenas uma a correta conforme o exemplo:

\*Leia cada uma das frases abaixo. Selecione V ou F para afirmativas verdadeiras ou falsas respectivamente.

(A) - Os lados opostos do paralelogramo são iguais (\_\_\_)

(B) - Triângulos isósceles têm três lados iguais (\_\_\_)

(C) - Um círculo pode ser definido como uma circunferência (\_\_\_)

**Vantagens:**

- Facilidade de elaboração, aplicação e automatização;
- Ajustável a todos os campos (áreas);
- Explora e estimula o raciocínio;
- Pode abranger grande quantidade de conteúdo.

**Desvantagens:**

- Fator sorte exerce um papel importante;
- Facilita a ocorrência de ambigüidades;
- Exige a elaboração de muitos itens;
- Geralmente, detalhes pequenos têm muita importância.

### 2.6.3.2 Questões de Evocação Simples

Este modelo de questão é muito usado em cursos do ensino fundamental. As questões são específicas, exigindo do aluno resposta direta, conforme observado no exemplo abaixo:

\*Quem descobriu o Brasil? \_\_\_\_\_

#### **Vantagens:**

- Facilidade de construção;
- Estimula a aprendizagem concreta;
- Aplicabilidade a qualquer tipo de conteúdo.

#### **Desvantagens:**

- Excesso de memorização de fatos;
- Risco de qualificação subjetiva;
- Possibilidade de não mensurar a compreensão global do conteúdo.

### 2.6.3.3 Questões de Lacuna ou Complemento

Em avaliações do tipo lacuna ou complemento, o aluno preenche espaços deixados em branco (vazio) em frases ou textos orientando-se a partir do seu conhecimento da matéria e contexto da questão.

Espera-se que o aluno saiba qual a palavra(s) mais adequada(s) para preencher as composições. O modelo permite ao professor criar o texto e assinalar no mínimo uma e no máximo três palavras a serem omitidas ao apresentar a questão para o aluno.

A construção de questões deste tipo deve respeitar as seguintes regras: evitar fornecer pistas para as respostas com a colocação de artigos ou preposições antes da lacuna; inserir por frase no máximo três lacunas para não comprometer a clareza da questão; garantir que conste nas lacunas apenas os elementos relevantes da sentença; elaborar as sentenças de forma clara, objetiva e precisa (BERCH, 1997).

Exemplo de questão do tipo lacuna ou complemento:

\* (a) O estado de Goiás faz parte da região \_\_\_\_\_.

\* (b) \_\_\_\_\_ é a capital do Tocantins.

**Vantagens:**

- Facilidade de elaboração, aplicação e automação;
- Evita o fator sorte.

**Desvantagens:**

- Estimula exageradamente a memorização;
- A relação entre os temas do conteúdo pode passar despercebida.

#### 2.6.3.4 Questões de Múltipla Escolha

Esta modalidade de questão é apropriada para testes de conhecimento de terminologia ou fatos específicos. Os respondentes optam por uma das alternativas, ou por determinado número permitido de opções. Ao elaborar perguntas de respostas múltiplas, o pesquisador se depara com dois aspectos essenciais a serem considerados: o número de alternativas oferecidas e sua posição (ordem de apresentação).

As alternativas devem cobrir todas as respostas possíveis, deixando apenas uma alternativa totalmente incompatível com as demais. É permitido incluir alternativas como: "Outros. Quais? \_\_\_\_\_" que ajudam a garantir a exclusão mútua. Para que sejam mutuamente exclusivas, cada respondente deverá identificar apenas uma opção que represente corretamente sua resposta, excluindo as demais.

As regras seguidas durante a criação são: apresentar todas as alternativas como possível resposta correta; evitar a inclusão de palavras ou expressões que sugiram declarações falsas ou verdadeiras (nunca, nenhum, somente, etc.); não usar a alternativa "todas as respostas acima" ou "nenhuma das respostas"; utilizar frases positivas e destacar frases negativas (BERCH, 1997).

Exemplo de questão de múltipla escolha:

\*Embora a continuidade da vida na Terra dependa substancialmente de todo o elenco de características que definem os sistemas vivos, duas dessas características assumem maior importância para a preservação da vida no planeta. São elas:

- A-( ) Composição química complexa e estado coloidal.
- B-( ) Elevado grau de organização e execução das funções vitais.
- C-( ) Manutenção da homeostase e alto nível de individualidade.
- D-( ) Consumo de energia e renovação contínua da matéria.
- E-( ) Capacidade de reprodução e hereditariedade.
- F-( ) Outras. Quais? \_\_\_\_\_

**Vantagens:**

- Impede a subjetividade (existe apenas uma opção correta);
- Facilidade de correção;
- Fácil de automatizar o processo;
- Facilidade de aplicação, processo e análise;
- Facilidade e rapidez de resposta;
- Apresentam pouca possibilidade de erros.

**Desvantagens:**

- Dificuldade em formular o número suficiente de questões para que o teste alcance critérios de validade e adequação;
- Exige muito tempo na preparação para garantir que todas as opções de respostas sejam oferecidas;

- O respondente pode ser influenciado pelas alternativas apresentadas;
- Exige que especialistas revisem as questões (devido ao seu alto grau de dificuldade durante o processo de construção).

### 2.6.3.5 Questões de Ordenação

Segundo Rodrigues (2000), este tipo de questão deve apresentar no mínimo três elementos para serem ordenados e evitar fornecer indícios de resposta, deixando claro o critério e o ponto de partida da ordenação desejada. Elaborar-se a lista de itens (alternativas) e o aluno registra a ordem seqüencial correta observando o enunciado, conforme exemplo abaixo:

\*Numere os fatos de 1 a 3, considerando sua seqüência cronológica:

- a) ( ) Proclamação da República
- b) ( ) Proclamação da Independência
- c) ( ) Eleições Diretas

### 2.6.3.6 Questões de Correspondência ou Pareamento

Permitem ao professor cobrir vários assuntos simultaneamente na mesma questão, no entanto são difíceis de elaborar.

O professor deve: inserir maior número de alternativas na coluna da direita (em relação à coluna da esquerda); evitar a possibilidade da exclusão de alternativas por eliminação; criar instruções explicando claramente a forma de associação; elaborar frases positivas; evitar construções gramaticais que indiquem a alternativa correta; manter a homogeneização quanto ao conteúdo e não incluir mais de um aspecto do conteúdo na mesma opção (BERCH, 1997).

Questões de correspondência apresentam duas colunas ao aluno, que deverá assinalar qual resposta da segunda coluna é correspondente à resposta da primeira, conforme o exemplo:

\*Numere a segunda coluna de acordo com a primeira, associando cada estado a sua capital:

- |        |                  |
|--------|------------------|
| (A) MT | ( ) Cuiabá       |
| (B) SC | ( ) Goiânia      |
| (C) GO | ( ) Porto Alegre |

## 2.7 DIVERGÊNCIA DE NOTAS ENTRE JÚRIS

Conforme observado neste capítulo, a avaliação ocupa papel de extrema importância dentro do contexto acadêmico do aluno. Autores como Noizet, Carverni (1985) e Moretti (2007) levantaram dados em pesquisas demonstrando a divergência existente entre médias atribuídas ao mesmo teste composto de questões abertas por júris (corretores humanos) distintos.

Noizet e Caverni (1985) na Academia D'aix - Marseille realizaram uma pesquisa utilizando treze júris diferentes, relatando que o desvio entre notas extremas destes corretores variavam de 1,3 a 3,9 pontos (para escala de notas de 0 a 20), ver Quadro 2.

Moretti (2007) a partir de dados da Universidade Federal de Santa Catarina e Universidade de Campinas utilizou-se de 52 júris na correção de dez testes, cujos resultados revelaram desvio entre 1,0 a 3,5 pontos (para escala de notas de 0 a 10), ver Quadro 3 e Quadro 4.

Quadro 2 - Divergência de Notas entre Júris - Noizet e Caverni (1985)

Corretores - 13 júris	Nota mais baixa	Nota mais alta	Desvio entre notas extremas
Filosofia	8,2	9,5	1,3
Filosofia	7,0	9,3	2,3
Física	8,7	11,4	2,7
Física	7,6	10,5	2,9
Ciências Naturais	7,7	11,6	3,9
Matemática	5,8	9,1	3,3

Quadro 3 - Divergência de Notas entre Júris - Moretti (2007)

Corretores e Provas										
	E 1	E 2	E 3	E 4	E 5	E 6	E 7	E 8	E 9	E 10
C01	3.00	3.00	4.50	3.00	1.00	0.00	2.50	4.00	3.00	3.00
C02	3.60	4.50	5.00	2.30	0.50	2.00	0.50	4.20	3.60	3.70
C03	2.00	4.00	5.00	0.00	0.50	0.00	0.00	4.50	4.00	2.50
C04	2.75	3.00	4.50	1.00	1.00	1.00	1.00	3.25	2.75	3.50
C05	2.00	3.00	5.00	0.00	1.50	0.50	0.50	4.50	3.50	3.50
C06	3.90	4.00	5.00	2.00	0.50	2.00	1.00	4.50	4.00	4.00
C07	3.50	4.00	4.50	1.50	0.00	0.50	0.00	3.00	3.50	3.50
C08	4.00	4.00	5.00	2.50	1.00	2.50	1.00	4.50	4.00	4.00
C09	3.25	3.50	5.00	2.50	0.50	2.00	0.50	4.50	3.75	4.00
C10	3.80	3.80	4.00	2.50	1.00	2.50	1.00	4.50	3.90	3.90
C11	4.00	4.00	5.00	1.00	0.50	1.50	0.50	4.00	4.00	4.00
C12	4.00	4.50	5.00	1.00	0.50	1.00	1.00	4.00	3.50	3.50
C13	3.50	3.50	5.00	1.00	0.50	1.50	0.50	4.00	3.50	3.50
C14	3.75	4.25	5.00	2.50	1.75	2.00	1.00	4.00	4.50	3.75
C15	3.00	4.00	5.00	0.50	2.00	0.00	0.00	4.00	3.50	3.50
C16	3.00	3.50	5.00	1.00	1.00	2.50	1.00	3.00	3.50	3.50
C17	3.00	3.50	5.00	0.50	0.50	2.00	0.50	4.50	3.00	3.00
C18	4.00	4.00	5.00	2.50	1.00	2.50	2.50	4.50	4.00	2.50
C19	4.00	4.50	5.00	2.50	1.00	2.50	1.00	4.50	4.00	4.00
C20	4.00	4.50	5.00	2.00	1.00	2.50	1.00	4.50	4.00	4.00
C21	3.00	3.50	5.00	0.00	0.50	1.00	0.50	4.00	3.00	3.50
C22	3.50	4.50	5.00	2.00	1.00	0.50	0.50	4.00	3.50	3.50
C23	3.50	4.50	5.00	2.50	0.50	2.50	1.00	4.50	4.00	4.00
C24	2.50	4.00	5.00	1.00	0.50	2.50	1.00	4.00	3.50	3.50
C25	3.50	4.50	5.00	2.00	1.50	2.50	1.00	4.50	4.00	4.00
C26	3.75	4.25	5.00	2.50	1.75	2.00	1.00	4.00	4.50	3.75
C27	4.00	4.00	5.00	2.00	1.00	2.00	0.50	4.50	4.00	4.00
C28	3.50	4.00	5.00	0.50	1.00	1.00	0.50	4.00	4.00	2.50
C29	3.00	3.00	4.50	0.50	1.00	1.00	0.50	3.50	3.50	3.50
C30	3.50	4.00	5.00	1.00	1.00	1.00	1.00	4.00	4.00	2.50
C31	4.00	4.50	5.00	2.50	1.00	2.50	1.50	4.00	3.50	3.50
C32	3.50	4.50	5.00	2.50	1.00	2.50	1.00	4.00	4.00	3.50
C33	3.50	4.00	5.00	2.00	1.00	2.50	1.00	4.00	4.00	4.00
C34	3.50	4.50	5.00	2.50	1.00	2.50	1.00	4.00	4.00	3.50
C35	4.00	4.00	5.00	2.50	1.00	2.50	1.00	4.00	3.50	4.00
C36	3.50	4.00	5.00	2.00	0.50	1.00	0.50	2.50	4.00	3.50
C37	3.50	4.00	4.90	2.20	2.70	2.20	2.50	3.80	3.50	4.00
C38	3.50	3.50	5.00	0.00	0.50	0.50	0.50	3.00	3.50	3.50
C39	3.00	3.00	4.00	1.00	1.00	1.00	1.00	3.50	3.00	2.50
C40	3.50	4.00	4.80	1.50	1.00	1.50	1.00	4.30	3.50	4.00
C41	3.50	4.00	5.00	1.50	0.00	3.50	0.50	4.50	3.50	3.50
C42	4.00	3.50	5.00	2.50	1.50	2.50	1.50	4.50	3.50	3.50
C43	3.00	3.50	5.00	1.50	1.50	2.00	2.00	3.00	3.50	3.50
C44	4.00	3.00	5.00	2.50	1.00	2.50	1.00	4.00	3.50	4.00
C45	4.00	2.50	4.50	0.50	0.50	0.50	0.50	3.00	3.50	3.50
C46	3.50	4.00	5.00	2.00	1.00	2.00	0.50	4.00	4.00	4.00
C47	3.50	4.00	5.00	2.50	1.00	2.00	1.50	4.50	4.00	4.00
C48	3.50	4.00	5.00	2.50	1.00	2.50	1.00	4.00	4.00	4.00
C49	4.00	4.00	5.00	1.50	0.50	2.50	0.50	4.00	3.50	3.50
C50	2.00	4.00	4.50	1.00	2.00	3.00	1.00	3.00	2.00	2.50
C51	2.50	4.00	5.00	1.00	3.00	1.00	1.50	4.50	3.50	2.50
C52	3.50	4.00	5.00	1.50	1.00	1.00	0.50	4.50	4.50	4.00

Quadro 4 - Divergência de Notas entre Júris - Desvio

Desvio entre Notas Extremas									
E1	E2	E3	E4	E5	E6	E7	E8	E9	E10
2,00	1,50	1,00	2,50	2,50	3,50	2,50	1,50	2,50	1,50

## 2.8 CONCLUSÃO

O processo avaliativo possibilita ao docente uma auto-avaliação da prática pedagógica, orientando-o quanto à modelagem e construção dos objetivos a serem atingidos e quanto ao grau de conhecimento absorvido pelo aluno, através da comparação constante entre resultados obtidos e metas definidas. Tal procedimento serve de subsídio para diagnosticar possíveis falhas de aprendizagem, permitindo intervir para melhoria do ensino e aprendizagem.

Quanto ao tipo de questão, pode-se afirmar que a utilização de questões abertas é benéfica ao aluno, por exigir que o aluno descreva com suas próprias palavras o que sabe do assunto abordado. Esse tipo de questão permite ao avaliado escrever sobre aspectos do texto, relações estabelecidas e raciocínios que o levaram a tal compreensão, aparecendo como a segunda forma mais citada pelos universitários na pesquisa descrita na seção 2.6.2, como a mais adequada para avaliação do desempenho.

Skinner (1982) afirma que aprendizagem é uma mudança de comportamento, a demonstração de que o indivíduo sabe algo que desconhecia anteriormente. A capacidade de agregar novos conhecimentos está diretamente ligada ao estilo de aprendizagem individual. Segundo Kolb (1976), algumas pessoas percebem melhor a informação de forma abstrata (pensamento) e de conceitos mentais (modelos conceituais), enquanto outras assimilam melhor a informação através dos sentidos vivenciando experiências concretas.

Uma avaliação de qualidade se desenvolve de forma contínua no decorrer do curso, combinando aspectos de caráter formativo e diagnóstico. Para tal, deve estar comprometida com os objetivos educacionais, levando em consideração os diversos estilos de aprendizagem.

A Educação do futuro precisa estar comprometida com a melhoria da aprendizagem, porém sem limitar os programas educativos ou o desenvolvimento de currículos escolares modernos. A melhoria de qualidade do trabalho pedagógico implica em aprimorar o sistema de apoio à aprendizagem do aluno, elegendo métodos para melhor aproveitamento de aluno e professor, estando direcionada a inclusão e a promoção da cidadania.

## CAPÍTULO 3

### Redes Neurais Artificiais

---

#### 3.1 INTRODUÇÃO

Há muito tempo o homem tenta elucidar os complexos mecanismos utilizados pela mente humana para pensar, falar, lembrar, sentir e aprender, na tentativa de modelar sua forma de operação. Uma estrutura disposta em rede chamada neurônio constitui a unidade básica, estrutural e funcional do sistema nervoso participando ativamente desta operação.

Redes Neurais Artificiais são modelos computacionais inspirados, no funcionamento do sistema nervoso animal, para solucionar problemas de inteligência artificial.

Embora a forma neuronal apresente muitas variações, o neurônio típico pode ser apresentado como sendo constituído por: dendritos, que são elementos receptores das informações (impulsos); o corpo celular, que constitui o centro metabólico, onde informações recebidas pelos dendritos são processadas; e o axônio, que transmite essas informações (impulsos) para os dendritos dos próximos neurônios (GUYTON, 1986).

Com o ressurgimento da tecnologia de Redes Neurais, novas pesquisas são concebidas na área de classificação de padrões. É constante a realização de trabalhos envolvendo reconhecimento de voz, visão, robótica e Inteligência Artificial utilizando respostas em tempo-real com dados complexos do mundo-real.

A maioria das pesquisas nesta área é motivada pelo desejo de entender e projetar Redes Neurais Artificiais baseadas em mecanismos de Redes Neurais Biológicas. Para tanto, é indispensável o aprendizado de novas aplicações de Inteligência Artificial, utilizando técnicas de projeto do VLSI (*Very Large-Scale Integration*) e disponibilidade de computadores de alta velocidade com grande quantidade de poder de processamento e memória.

## 3.2 CONCEITO

Redes Neurais Artificiais são ferramentas de Inteligência Artificial que possuem a capacidade de adaptação e aprendizado de determinada tarefa, ou comportamento, a partir de um conjunto de exemplos dados (OSÓRIO & BITTENCOURT, 2000). Tais ferramentas também podem ser definidas como algoritmos computacionais inspirados no processamento de informações do cérebro, que adquirem conhecimento através da experiência.

Grandes Redes Neurais Artificiais podem conter normalmente centenas ou milhares de unidades de processamento, enquanto o cérebro de um mamífero pode ter vários bilhões de neurônios. O desafio destas estruturas consiste em solucionar problemas de Inteligência Artificial, construindo sistemas com circuitos capazes de simular o funcionamento do cérebro, incluindo seu comportamento (aprender, errar e fazer descobertas).

Um grafo direcionado é um objeto geométrico que é representado por um conjunto de pontos chamados nós, ao longo de segmentos de linhas direcionadas entre eles. Uma Rede Neural é uma estrutura de processamento de informação distribuída na forma de um grafo direcionado, com algumas restrições e definições próprias.

De forma análoga, os nós deste grafo são chamados elementos de processamento. Suas arestas são conexões, que funcionam como caminhos de condução instantânea de sinais em uma única direção, de forma que seus elementos de processamento podem receber qualquer número de conexão de entrada.

Estas estruturas podem apresentar memória local e qualquer número de conexões de saída desde que os sinais nestas conexões sejam iguais. Em resumo, estes elementos têm uma única conexão de saída, dividindo-se em cópias para formar múltiplas conexões, desde que carreguem o mesmo sinal.

A única entrada permitida para a função de transferência (que cada elemento de processamento possui), são valores armazenados na memória local do elemento de processamento e valores atuais dos sinais de entrada nas conexões recebidas pelo elemento de processamento, enquanto os únicos valores de saída permitidos a partir da função de transferência são valores armazenados na memória local do elemento de processamento, e o sinal de saída do mesmo.

Funções de transferência podem operar continuamente ou esporadicamente. No primeiro caso, elementos estão sempre ativados, e a entrada "*activate*" chega através da conexão de um elemento de processamento agendado que também é parte da rede. No segundo caso, deve existir uma entrada chamada "*activate*" que ativa a função de transferência pelo sinal de entrada corrente e com valores da memória local, resultando num sinal de saída atualizado (ocasionalmente alterando valores da memória).

Sinais externos de entrada para uma Rede Neural chegam através de conexões originárias do mundo externo. Saídas da rede para o mundo externo são conexões que deixam a rede.

### 3.3 BREVE HISTÓRICO DE REDES NEURAIS

#### 3.3.1 Marco Inicial

O interesse em Redes Neurais teve início na década de 40, com o trabalho pioneiro de McCulloch, Pitts (1943). Warren McCulloch foi um psiquiatra e neuro-anatomista que estudou por vinte anos a representação de um evento no sistema nervoso. Walter Pitts foi um prodígio matemático, que firmou parceria com McCulloch em 1942 escrevendo um artigo que se tornou um clássico na área, recebendo a atenção da comunidade que estudava o modelo do neurônio.

As primeiras informações mencionadas sobre neuro-computação datam de 1943, com a publicação do trabalho de McCulloch e Pitts denominado: “*A logical Calculus of the Ideas Immanent in Nervous Activity*” cujo conteúdo sugeria a construção de uma máquina baseada ou inspirada no cérebro humano, descrevendo o modelo artificial de um neurônio e suas capacidades computacionais, sem se ater a técnicas de aprendizado (Figura 3.1).

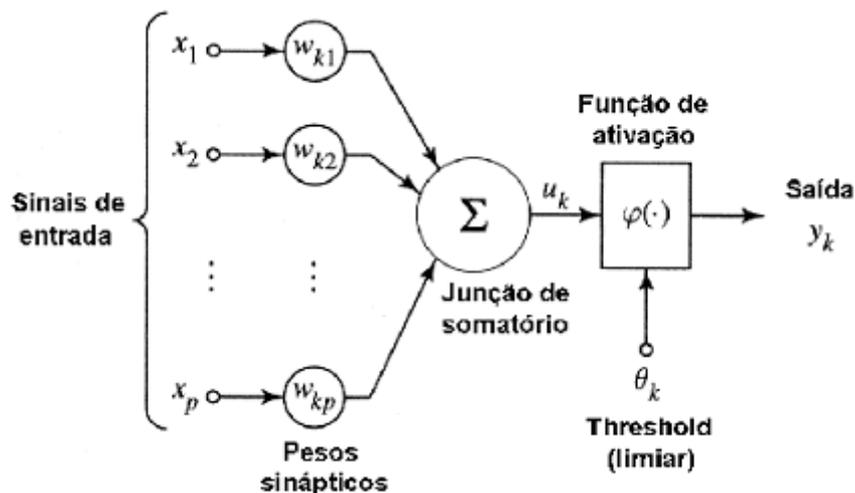


Figura 3.1 - O Neurônio de McCulloch-Pitts

Outros artigos e livros foram editados desde então, porém, por um longo período de tempo poucos foram os resultados reais de sucesso. Hebb (1949) escreveu o livro intitulado "*The Organization of Behavior*" (A Organização do Comportamento) que perseguia a idéia de que o condicionamento psicológico clássico está presente em qualquer parte dos animais pela propriedade dos neurônios individuais.

Suas idéias não eram completamente novas, mas Hebb foi o primeiro a propor uma lei de aprendizagem específica para as sinapses dos neurônios, demonstrando que o aprendizado das Redes Neurais está baseado no reforço das ligações sinápticas entre os nodos.

Este primeiro passo serviu de inspiração para outros pesquisadores seguirem a mesma linha de raciocínio e embora muito tenha sido estudado e publicado nos anos seguintes, tal proposta serviu mais como base para desenvolvimento posterior que para o próprio desenvolvimento. Estes primeiros estudos das Redes Neurais Biológicas formaram os fundamentos das Redes Neurais Artificiais (RNAs).

Em 1951 foi construído o primeiro neuro-computador por Minsky e Papert (1969), denominado Snark. O Snark operava com sucesso a partir de um ponto de partida técnico, ajustando seus pesos automaticamente. Apesar de nunca ter executado qualquer função de processamento de informação interessante, serviu de embasamento para as idéias de estruturas que o sucederam.

Em 1956 no "*Dartmouth College*" surgiram os dois paradigmas da Inteligência Artificial; denominados como simbólico e conexionista. A Inteligência Artificial Simbólica tentava simular o comportamento inteligente humano desconsiderando seus mecanismos. Já a Inteligência Artificial Conexionista acreditava que construindo um sistema que simulasse a estrutura do cérebro, este sistema apresentaria inteligência e seria capaz de aprender.

### 3.3.2 Perceptron

O primeiro neuro-computador a obter sucesso (*Mark I Perceptron*) surgiu em 1957, criado por Frank Rosenblatt, Charles Wightman e outros pesquisadores desta época. Devido ao pensamento inovador, a profundidade de seus estudos, e suas contribuições técnicas, muitos estudiosos vêem Rosenblatt como o fundador da neuro-computação. Seu interesse inicial para a criação do Perceptron era o reconhecimento de padrões, ou seja, seu modelo consistia basicamente em classificar padrões linearmente separáveis, utilizando nodos MCP com sinapses ajustáveis (Figura 3.2).

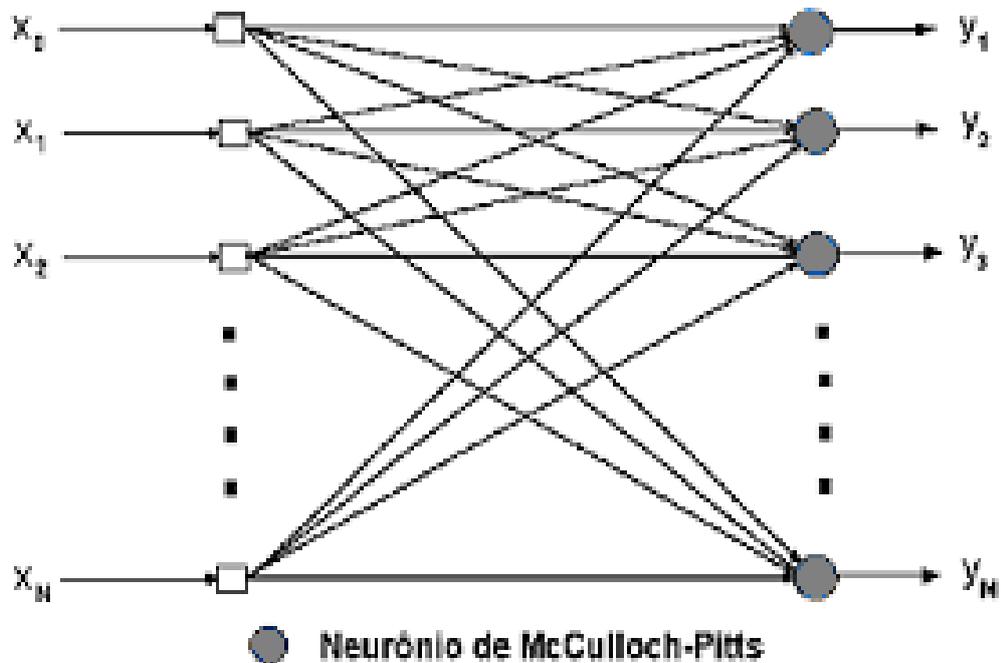


Figura 3.2 - Perceptron de Uma Camada Simples

Em 1960 Bernard Widrow, auxiliado por estudantes, desenvolveu um tipo de elemento de processamento de Redes Neurais denominado de Adaline (elemento linear adaptativo), equipado com uma poderosa lei de aprendizado, diferente do Perceptron, que ainda é utilizado atualmente. Widrow também fundou a primeira companhia de hardware de neuro-computadores e componentes (WIDROW, HOFF 1960).

A principal diferença entre o “Perceptron” de Rosenblatt e o “Adaline” de Widrow consiste no procedimento de treinamento. Widrow e seus estudantes propuseram uma das primeiras redes neurais com camadas treináveis e múltiplos elementos adaptativos, que foi chamada de “Madaline” (Múltiplo Adaline).

Depois da apresentação do perceptron na década de 1960, acreditava-se que Redes Neurais (perceptrons) estavam aptas a resolução de qualquer problema. Essa hipótese foi arrasada quando Minsky e Papert (1969) publicaram o livro *Perceptrons*, demonstrando que existiam limites para o perceptron com apenas uma camada intermediária e o mesmo não poderia nem mesmo resolver o simples problema do "ou exclusivo" (XOR).

Após a publicação da obra de Minsky e Papert, tratando das limitações dos computadores e estações de trabalhos na condução de experimentos, e por não haver suporte financeiro para conduzir projetos nesta área, as pesquisas com Redes Neurais ficaram esquecidas e os anos seguintes foram marcados pelo entusiasmo exagerado de vários pesquisadores, publicando obras declarando a impossibilidade de máquinas tão poderosas quanto o cérebro humano surgirem em tão curto espaço de tempo.

O fato abalou a credibilidade dos estudos desta área na época, tendo início um período de pesquisa silenciosa compreendido entre 1967 a 1982, havendo poucas publicações sobre o assunto. Entretanto, estes pesquisadores e os demais que surgiram no decorrer dos próximos treze anos conseguiram novamente estabelecer um campo concreto para o renascimento da área. Nos anos 80, muitos autores passaram a publicar propostas para exploração do desenvolvimento de Redes Neurais, bem como suas aplicações.

O fato mais importante deste período ocorreu quando Ira Skurnick, um administrador de programas da DARPA (*Defense Advanced Research Projects Agency*) decidiu ouvir os argumentos da neuro-computação e seus projetistas. Divergindo dos caminhos tradicionais de conhecimentos convencionais, fundou em 1983 um centro de pesquisas, abrindo as portas para a neuro-computação e elevando a DARPA ao status de líder mundial na área tecnológica.

Outro importante autor deste período foi John Hopfield, físico de reputação internacional que se interessou pela neuro-computação e escreveu artigos que ultrapassaram fronteiras, persuadindo centenas de cientistas, matemáticos e tecnólogos altamente qualificados a se unirem em torno desta área emergente. Estima-se que um terço dos pesquisadores aderiram à influência de Hopfield, mas apenas em 1986 houve um real crescimento das pesquisas em função da publicação do livro "*Parallel Distributed Processing*" (Processamento Distribuído Paralelo), editado por David Rumelhart e James McClelland.

Em 1987 em São Francisco, aconteceu a primeira conferência de Redes Neurais em tempos modernos, a ICNN (*International Conference on Neural Networks*), resultando na formação da *International Neural Networks Society* (INNS). A partir destes acontecimentos, várias universidades anunciaram a formação de institutos de pesquisa e programas de educação de neuro-computação. Em 1989 foi fundado o INNS journal, seguido do *Neural Computation* (NC) e do TNN *Transactions on Neural Networks* em 1990.

## 3.4 MODELOS BOOLEANOS

### 3.4.1 Neurônios RAM

Este modelo de neurônio é vantajoso por permitir sua implementação em hardware (ALEKSANDER, THOMAS, BOWDEN, 1984). Redes baseadas em neurônios deste tipo aceitam e produzem valores binários. Estes modelos de redes são conhecidos como modelos booleanos (ALEKSANDER, ALBROW, 1968; ALEKSANDER, MANDANI, 1968).

Um neurônio do tipo RAM possui:

- Um conjunto de valores ponderados que codificam suas entradas na forma de endereços sem sofrer alteração, sendo utilizados somente para rotular terminais de entrada do neurônio;
- Um conjunto de células que armazenam a informação aprendida;
- Um terminal de saída que informa o valor da célula endereçada pelos terminais de entrada;
- Um terminal de ensino onde a saída desejada é apresentada ao neurônio e um terminal de modo operação que indica se o neurônio está na fase de aprendizado ou de utilização.

A Figura 3.3 exibe a estrutura de um neurônio RAM.

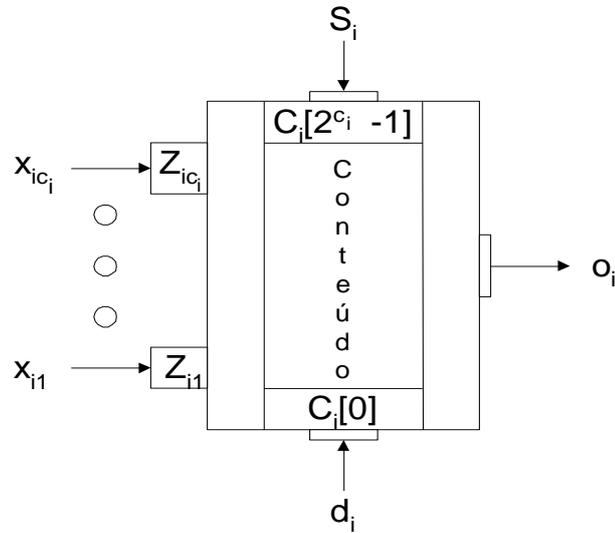


Figura 3.3 - Estrutura do Neurônio RAM

O neurônio do tipo RAM possui dois estados operacionais: o estado de aprendizado, onde neurônio recebe o valor desejado no terminal de ensino endereçando este valor em uma de suas células, e o estado de utilização, no qual o neurônio coloca o conteúdo de uma de suas células no terminal de saída. A função de saída é dada por:

$$O_i = \begin{cases} 1, & \text{se e somente se valor da célula endereçada é 1} \\ 0, & \text{se e somente se valor da célula endereçada é zero} \end{cases}$$

O endereço acessado por uma entrada  $x_i$  é definido por:

onde  $x_{ij}$  é o valor de entrada do  $j$ -ésimo terminal e  $z_{ij}$  é o endereço do  $j$ -ésimo terminal. Note a Equação 1.

$$a_{i_m} = \sum_{j=1}^{j=i} x_{ij} z_{ij}$$

Equação 1 - Neurônio tipo RAM

### 3.4.2 Neurônios PLN

Em 1987 foi criado um modelo de neurônio RAM por Aleksander e Kan (1987) nomeado PLN (*Probabilistic Logic Neuron*). O principal diferencial deste neurônio é a inclusão de um terceiro valor lógico (“u”) que pode ser armazenado em suas células. Quando uma célula com este valor é endereçada, produz uma saída aleatória no intervalo discreto fechado [0,1]. Este terceiro valor é um valor indefinido e significa que a célula endereçada pode possuir o valor 0 ou 1.

A função probabilística que produz a saída do neurônio pode gerar zeros ou uns com qualquer probabilidade, sendo mais comum a construção de números com mesma probabilidade. Esta função tem sua saída definida por:

$$O_i = \begin{cases} 0, & \text{se e somente se valor da célula endereçada é zero} \\ 1, & \text{se e somente se valor da célula endereçada é 1} \\ \text{Ale}(0,1), & \text{se e somente se valor da célula endereçada é “u”} \end{cases}$$

Onde, Ale(0,1) é uma função aleatória que produz zeros e uns com a mesma probabilidade.

As redes com neurônios do tipo RAM simplificadas utilizam somente uma camada de neurônios (com ou sem retroalimentação). Normalmente a arquitetura da rede é parcialmente conectada, em função do crescimento exponencial da memória utilizada pelos neurônios RAM.

As redes que utilizam o neurônio PLN possuem normalmente arquitetura piramidal. Esta pirâmide, ou árvore, é uma estrutura hierárquica composta de camadas neuronais. Cada neurônio de determinada camada é conectado a um único neurônio da camada posterior, de acordo com o exemplo piramidal visto na Figura 3.4.

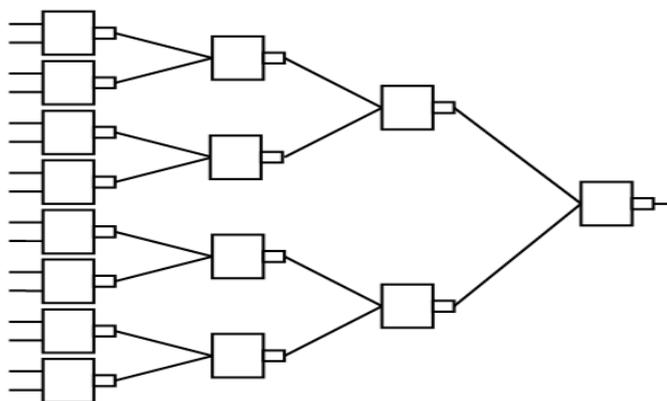


Figura 3.4 - Exemplo de Arquitetura Piramidal

### 3.4.3 Neurônios GSN (Goal Seek Neuron)

O neurônio do tipo GSN foi projetado para vencer algumas das limitações dos neurônios booleanos citados acima, sem perder de vista a produção de resultados precisos. Estas limitações dizem respeito ao crescimento exponencial da memória utilizada pela rede (FILHO, 1990).

Redes com neurônios GSN viabilizam máxima utilização da capacidade de memória juntamente com a representação e recuperação de padrões para qualquer problema. Assim como o PLN, o GSN é baseado no elemento de memória de acesso aleatório, onde endereços de entrada armazenam células contendo valores do conjunto  $\{0,1,u\}$ , onde  $u$  é valor indefinido.

Em contraste com o PLN, o GSN recebe e produz todos os três valores, ao invés de ter que fazer uma escolha arbitrária quando uma célula com valor indefinido é acessada. Conseqüentemente o GSN pode endereçar um conjunto de células ao invés de um único endereço. Desse modo pode endereçar qualquer quantidade de endereços: de uma célula até todas as células. O GSN possui três diferentes estados de operação:

- Estado de validação: o neurônio valida a possibilidade de aprender sem destruir os valores armazenados;
- Estado de aprendizagem: o neurônio seleciona um endereço e armazena a saída desejada;
- Estado de reconhecimento: o neurônio faz a computação do resultado baseado na entrada recebida.

A arquitetura do GSN consiste em:

- Um conjunto de terminais que representam a entrada para o neurônio;
- Um conjunto de valores ponderados que codificam a entrada na forma de endereços;

- Um conjunto de células que armazenam conteúdo ou informação aprendida;
- Um terminal de saída que transmite a saída do neurônio;
- Um terminal de ensino que identifica a saída desejada a ser aprendida no estado de aprendizado;
- Um conjunto de terminais de estado que indicam o estado do neurônio.

O espaço de endereçamento de um GSN é formado por todas as combinações binárias das entradas do neurônio. O modelo GSN recebe e gera valores do conjunto  $T = \{0, 1, u\}$ , onde a presença de valores indefinidos nos terminais de entrada identifica um conjunto de possíveis endereços, ao invés de um único endereço. O conjunto dos possíveis endereços é denominado conjunto endereçado, e o conjunto dos conteúdos assim endereçados são designados conteúdo endereçado.

A área onde o objetivo é procurado pode variar de um simples endereço (no caso onde todas as entradas são valores binários) até todo o conjunto de endereços (no caso onde todas as entradas são valores indefinidos). O processamento é feito pelo exame do conteúdo endereçado e a geração do melhor valor para o objetivo procurado.

O conjunto endereçado é determinado pelo processamento dos endereços físicos, dados pelos valores definidos nos terminais de entrada, e pela pesquisa de todas as combinações de endereços para os valores indefinidos nos terminais de entrada. O neurônio GSN possui a forma da Figura 3.5.

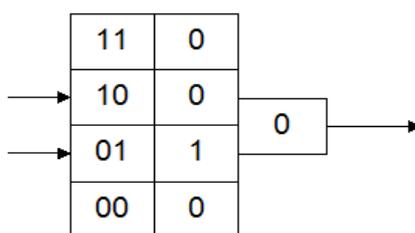


Figura 3.5 - Neurônio GSN

A estrutura típica de uma pirâmide GSN é semelhante a uma árvore binária. Cada neurônio da pirâmide recebe duas entradas externas podendo assumir valores do conjunto  $T=\{0,1,u\}$ . A forma como o neurônio irá responder depende do estado em que se encontra.

**Estado de Validação:** quando uma rede GSN está no estado de validação, o objetivo é produzir um valor indefinido (“u”) em sua saída. Um valor indefinido, neste estado, indica que a rede pode aprender qualquer saída desejada. Uma saída definida (0 ou 1) indica que a rede somente poderá aprender o padrão, se a saída desejada para o mesmo for igual à saída produzida pela rede. Isto é, caso um padrão tenha como saída desejada 1 e a rede tenha produzido 0 na validação do padrão, ou vice versa, temos um caso onde a rede não pode aprender o padrão apresentado. Esta situação indica que o padrão possui uma estrutura que contradiz a estrutura dos padrões anteriormente aprendidos pela rede, sendo talvez um falso exemplo.

No **estado de validação**, quando o conteúdo endereçado contém um valor indefinido, ou quando existe uma mistura de zeros (0) e uns (1) a saída é indefinida. A saída é zero (0) quando todos os conteúdos endereçados são zeros (0) e a saída é um (1) quando todos os conteúdos endereçados são uns (1).

Este modo de operação propaga caminhos indefinidos através da rede, procurando células não utilizadas ou com valores conflitantes. Uma célula não utilizada (com o valor indefinido “u”) no conteúdo endereçado pode ser usada para armazenar a saída desejada (0 ou 1). Se existem valores conflitantes no conteúdo endereçado, uma célula que armazena o bit de saída desejado é selecionada sendo usada no mapeamento, sem perda de informação previamente armazenada.

O objetivo no estado de validação é produzir um valor indefinido (“u”). Como todas as células do conteúdo endereçado no exemplo possuem o valor indefinido (“u”), uma célula é escolhida aleatoriamente para produzir a saída do neurônio.

**Estado de Aprendizagem:** quando uma rede está no estado de aprendizado, o objetivo de cada neurônio é aprender a saída desejada com um endereço que já armazene um valor apropriado, de outro modo, uma célula com valor indefinido do conteúdo endereçado, é usada para armazenar a saída desejada.

Cada neurônio tentará fazer associação da saída desejada com uma célula existente no conjunto endereçado. Se o neurônio falhar, ele escolherá um valor indefinido armazenando o valor da saída desejada. Existindo um número de possíveis escolhas, então uma decisão aleatória é feita. Isto ocorre porque é necessário escolher uma célula particular para representar o valor de saída, pois, o endereço da célula escolhida é passado de maneira retroativa para suas conexões de entrada (como o valor de saída desejado para os níveis anteriores), desta forma, cada neurônio recebe sua saída desejada aprendendo e produzindo a saída para níveis anteriores.

Redes GSN são extremamente sensíveis à qualidade dos dados utilizados em seu treinamento. Padrões com informações contraditórias a padrões armazenados na rede, não podem ser aprendidos e são descartados.

**Estado de Reconhecimento:** Nesse estado a meta é produzir o valor binário de maior ocorrência no conjunto endereçado. Se o número de uns (1) no conteúdo endereçado é maior que número de zeros (0), então o neurônio produz um (1). Se número de zeros é maior, então o neurônio produz zero (0). Se o número de zeros e uns é igual, então neurônio produz valor indefinido. Assim, mesmo se o conteúdo endereçado contém somente um único valor 1, e valores restantes são indefinidos, o neurônio irá produzir 1. A razão para adoção deste esquema é minimizar a propagação de valores indefinidos.

A fase de reconhecimento consiste na submissão de um padrão aos neurônios de entrada e na computação do valor de saída caracterizando a presença de uma classe particular de padrões.

Um GSN no estado de reconhecimento procura pelo valor binário de saída com maior probabilidade de ocorrência no conjunto endereçado. Neste estágio, a rede trabalha “*feedforward*” e cada pirâmide é pesquisada indiretamente para produzir a saída mais provável, associada ao padrão de entrada. Uma pirâmide produzindo um valor indefinido, pode ser visto como rejeição para o padrão associado.

Dependendo do tipo de problema, o comportamento dos neurônios pode ser alterado conforme a necessidade específica. O objetivo no estado de reconhecimento pode ser modificado produzindo valor indefinido quando ambos os valores binários são encontrados no conteúdo endereçável.

Redes formadas por este tipo de GSN rejeitarão mais padrões que redes compostas pelos GSNs propostos originalmente. Outra política levaria em conta se o número de uns no conteúdo endereçado é maior que número de zeros e valores indefinidos. Neste caso, o neurônio produz o valor 1; se o número de zeros superar o número de uns e de valores indefinidos, então o neurônio produz zero; se o número de valores indefinidos é maior que número de zeros e de uns, o neurônio produz um valor indefinido. Adotando este esquema, mais valores indefinidos são propagados.

## 3.5 MODELOS TRADICIONAIS DE REDES NEURAIAS

### 3.5.1 Mapas Auto Organizáveis

Este modelo foi desenvolvido por Teuvo Kohonen (1989) na década de 80 com forte inspiração neurofisiológica. Os mapas auto-organizáveis se inspiram no córtex cerebral. Neste modelo de rede neural os neurônios formam estruturas topológicas ordenadas (uma matriz bi-dimensional na maioria das vezes) e estão ligados com variáveis de entrada. Na Figura 3.6 temos um exemplo de uma rede SOM.

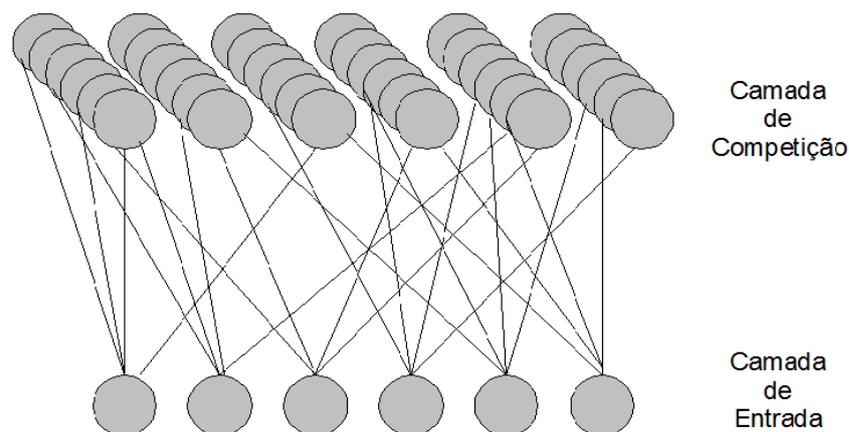


Figura 3.6 - Esquema da Rede SOM

Todos os pesos das conexões são inicializados aleatoriamente com um pequeno valor. A taxa de aprendizado decresce durante o treinamento sofrendo grandes alterações no início e pequenos ajustes no final.

O funcionamento de uma rede SOM resume-se em:

- Um padrão é apresentado à camada de entrada e transmitido aos demais neurônios de saída;
- Um processo de competição entre neurônios de saída indica um neurônio vencedor que obtém o direito de representar o padrão de entrada;
- O neurônio vencedor, e seus vizinhos dentro de um raio pré-determinado, atualizam seus pesos para que a resposta seja próxima do padrão de entrada.

O treinamento de redes SOM ocorre em duas fases:

- Fase de ordenação: que implica na ordenação topológica dos vetores de pesos;
- Fase de convergência: pela qual é feito um ajuste fino do mapa com menor taxa de aprendizado e com o raio de vizinhança envolvendo um ou nenhum vizinho.

A Figura 3.7 mostra o processo de redução da região de vizinhança durante o treinamento de uma rede SOM.

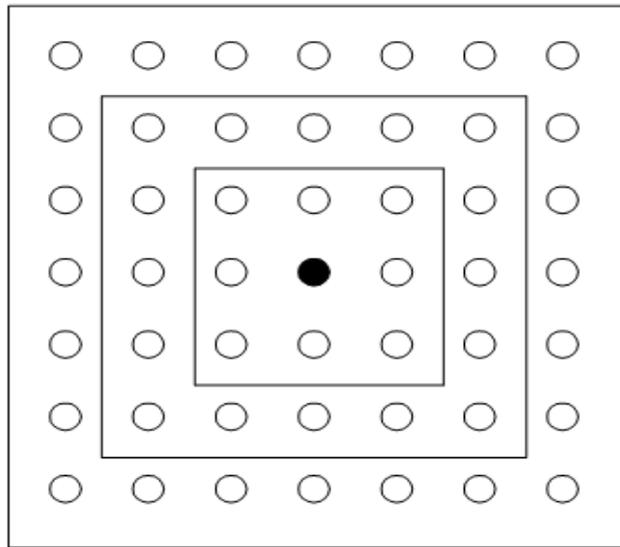


Figura 3.7 - Processo de Redução de Vizinhaça de uma Rede SOM

As redes SOM são utilizadas em vários tipos de aplicações, podendo destacar: diagnóstico financeiro, mineração de dados, controle de processos químicos e reconhecimento de voz (KOHONEN, 1989).

### 3.5.2 Multi Layer Perceptron - MLP

#### 3.5.2.1 Conceito

Como visto anteriormente, o perceptron é uma Rede Neural utilizada para classificação de dados linearmente separáveis, constituídos de um único neurônio com pesos sinápticos ajustáveis. Ainda que seu autor Rosenblatt a descreva como uma rede de três níveis, a topologia do perceptron é de apenas uma camada, com um conjunto de entradas e a camada de neurônios. O neurônio realiza o processamento das entradas (soma dos sinais de entrada ponderados pelos respectivos pesos) e, através de uma função de transferência, produz uma saída.

Matematicamente foi definido que redes de uma camada são incapazes de solucionar problemas não linearmente separáveis (MINSKY, PAPERT, 1969). Alguns autores por desacreditarem na possibilidade de construção de um método de treinamento para redes com mais de uma camada, concluíram que redes neurais seriam sempre suscetíveis a essa limitação.

Ao contrariar essa afirmação Rumelhart (1986) apresentou um modelo de estrutura com maior capacidade de representação possibilitando treinar eficientemente redes com camadas intermediárias, para solução de problemas mais complexos, resultando no modelo de Rede Neural Artificial MLP - Perceptron Multicamadas.

### 3.5.2.2 Arquitetura

As redes MLP possuem um poder computacional significativamente maior que os modelos sem camadas intermediárias, tendo cada camada uma função específica. A Figura 3.8 apresenta a arquitetura típica de uma rede MLP.

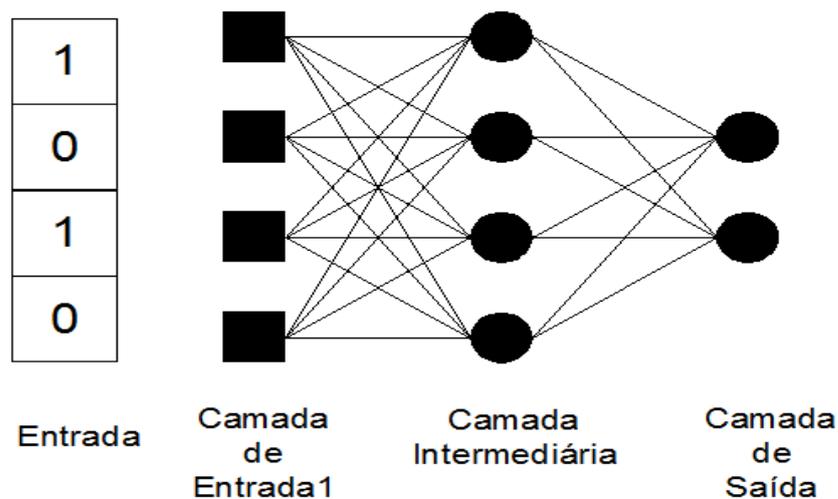


Figura 3.8 - MLP Típica com uma Camada

Nas redes **MLP** o processamento de cada neurônio é definido pela combinação do processamento realizado pelos neurônios conectados a ele resultantes da camada anterior. Tipicamente, para uma rede com duas camadas intermediárias o seguinte processamento é feito em cada uma das camadas.

- Primeira camada intermediária: os neurônios desta camada traçam retas no espaço de padrões de treinamento;
- Segunda camada intermediária: os neurônios formam regiões convexas pela combinação das retas traçadas pelos neurônios da camada anterior conectados a ele, onde o número de lados é definido pelo número de unidades a eles conectado;
- Camada de saída: os neurônios fazem a combinação das regiões convexas definidas pelos neurônios a ele conectados, ou seja, recebem estímulos da camada intermediária construindo o padrão de resposta;
- A combinação de três fatores distintos somados a habilidade de aprender através de treinamento fornecem a rede **MLP** maior poder computacional;
- Possuem uma ou mais camadas de neurônios ocultos, que não fazem parte nem da entrada e nem da saída. Estes neurônios capacitam a rede a aprender tarefas complexas, extraindo características significativas dos padrões de entrada;

- Demonstra alto grau de conectividade, determinado pelas sinapses da rede. Uma modificação na conectividade requer mudança nas conexões ou seus pesos;
- Cada neurônio da rede inclui uma função de ativação não linear (a mais comum é a função logística sigmóide).

Havendo conexões apropriadas entre unidades de entrada e um conjunto suficientemente grande de unidades intermediárias, encontra a representação que produzirá o mapeamento correto da entrada para a saída através das unidades intermediárias. Cybenko (1988) provou que para aproximar qualquer função contínua, apenas uma camada intermediária é suficiente.

### **3.5.2.3 Treinamento**

Algoritmos para treinamento de redes MLP são normalmente supervisionados. Existem vários algoritmos de treinamento, porém o mais conhecido é o back-propagation (RUMELHART, HINTON, WILLIAMS 1986).

O back-propagation utiliza par (entradas, saída desejada) e um mecanismo de correção de erros que ajusta os pesos da rede com base na saída obtida/saída desejada. Este algoritmo é baseado na regra delta proposta por Widrow e Hoff (1960), também conhecida como regra delta generalizada em que o ajuste dos pesos se baseia no método de gradiente descendente (RUMELHART, HINTON, WILLIAMS 1986).

A regra delta generalizada exige que funções de ativação utilizadas pelos neurônios sejam contínuas, diferenciáveis e geralmente não decrescentes. A aprendizagem consiste de um passo para frente (propagação ou “*forward*”) e um passo para trás (retropropagação ou “*backward*”) por meio das diferentes camadas da rede. Para frente propagam-se os estímulos, enquanto os erros são retropropagados. A Figura 3.9 mostra este processo.

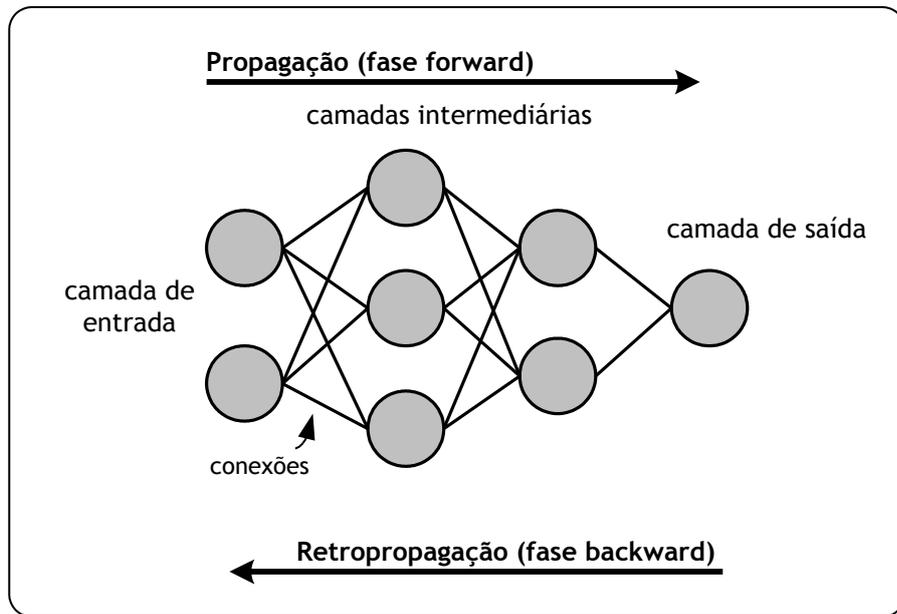


Figura 3.9: Treinamento da Rede MLP

As variáveis do conjunto de entrada, durante a propagação, são apresentadas na camada de entrada da rede (sensores). Cada entrada é ponderada pelo peso respectivo e aplicado ao neurônio ou neurônios da camada subsequente, produzindo determinado efeito. Esta situação se repete nas camadas subsequentes até a camada de saída onde é produzida uma representação da rede (resposta da rede) para padrões aplicados na camada de entrada, ocorrendo tanto na fase de treinamento, como na fase de utilização da rede (treinada).

Na fase de treinamento, a resposta da rede é comparada ao valor esperado (supervisão). Caso valores sejam diferentes, o algoritmo de treinamento *backpropagation* inicia a retroação. A resposta da rede é subtraída do valor de resposta desejada, produzindo um sinal de erro que retorna através da rede, no sentido contrário realizado na propagação, vindo daí o nome de retropropagação de erro (*backpropagation*).

Na retroação, cada peso sináptico é reajustado para que a resposta da rede se aproxime do valor desejado (convergência). O algoritmo *backpropagation* é um algoritmo supervisionado que utiliza pares de entrada e saída para ajustar os pesos da rede pela correção dos erros. O algoritmo *backpropagation* é apresentado na Figura 3.10.

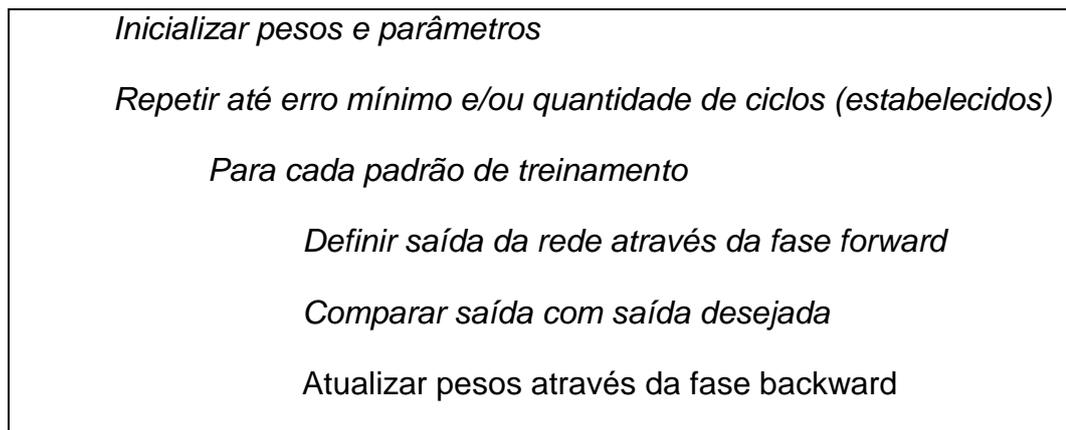


Figura 3.10 - Algoritmo *Backpropagation*

É importante a definição de alguns parâmetros estratégicos no treinamento da rede, sendo necessário estabelecer objetivos do treinamento bem como a forma de utilização e apresentação do conjunto de dados.

Definir o objetivo significa estabelecer até onde o treinamento será realizado, ou seja, o ponto de parada do algoritmo de treinamento. O treinamento deve ser finalizado ao estabelecer valores de performance a serem atingidos, quantidade de ciclos de treinamento ou, ainda, a combinação dos dois fatores. O valor de performance depende da expectativa de utilização da rede e forma de utilização do conjunto de dados, estabelecendo como a amostra será dividida para compor conjuntos de treinamento e teste da rede.

O conjunto de treinamento é o conjunto de dados utilizado para o treinamento da rede, enquanto o conjunto de validação (uma subdivisão do conjunto de treinamento), é o utilizado na verificação da eficiência da rede quanto a capacidade de generalização durante o treinamento, podendo ser empregado como critério de parada do treinamento.

Após o encerramento do treinamento, o conjunto de teste é utilizado para verificar o desempenho da rede sob condições reais de utilização. Nas situações onde a representatividade da amostra é pequena, alguns pesquisadores utilizam toda a amostra obtida como conjunto de treinamento. Nestas situações, o teste da rede é observado durante sua utilização.

A apresentação dos dados à rede durante o treinamento acontece de forma individual com retropropagação, sendo realizada após a apresentação de cada um ou por lote onde o conjunto de exemplos do treinamento são apresentados antes de iniciar a fase de retropropagação para ajuste de pesos. Pela simplicidade e fácil implementação, as redes MLP são as mais utilizadas atualmente em várias aplicações, dentre as quais destacamos: diagnóstico médico, previsão de ações na bolsa, verificação de assinaturas, códigos postais e outras (BURKE, ROSEN, GOODMAN, 1995).

### 3.6 DESENVOLVIMENTO DE SISTEMAS CONEXIONISTAS

A metodologia para desenvolvimento de sistemas com redes neurais artificiais tem início com análise cuidadosa do problema para minimizar ambigüidades e erros nos dados. A seguir, realiza-se a coleta de dados relativos ao problema. Os dados coletados devem ser significativos cobrindo amplamente o domínio do problema sem se ater apenas as operações normais ou rotineiras, mas também as exceções e condições nos limites do domínio do problema.

Após a coleta, os dados são separados em conjuntos de treinamento e de testes. Ao determinar estes conjuntos, eles são colocados em ordem aleatória para prevenção de tendências associadas à ordem de apresentação dos dados. Além disso, pode ser necessário pré-processar estes dados através de normalizações, escalonamentos e conversões de formato adequando-os à utilização na rede.

A próxima etapa é a definição da configuração da rede, que pode ser dividida em três etapas (HAYKIN, 2001):

- Seleção do paradigma neural apropriado à aplicação;
- Determinação da topologia da rede a ser utilizada (o número de camadas, o número de unidades em cada camada, etc.);
- Determinação de parâmetros do algoritmo de treinamento e funções de ativação (passo de grande impacto no desempenho do sistema resultante).

Normalmente estas escolhas são feitas de forma empírica, exigindo experiência dos projetistas. O próximo passo é o treinamento da rede. Nesta fase, seguindo o algoritmo de treinamento escolhido, serão ajustados os pesos das conexões. É importante considerar alguns aspectos, tais como: a inicialização da rede, o modo e o tempo de treinamento. Uma boa escolha dos valores iniciais dos pesos da rede diminui o tempo necessário para o treinamento. Normalmente, os valores iniciais dos pesos da rede são números aleatórios.

Quanto ao modo de treinamento, o ciclo pode ser completado a cada dado do conjunto (individual) ou após a apresentação de todos os dados do conjunto (*batch*). O modo individual aloca menor espaço no armazenamento de dados, sendo menos susceptível ao problema de mínimos locais. Por outro lado, no modo em lote se tem melhor estimativa do vetor gradiente, o que torna o treinamento mais estável. A eficiência relativa dos dois modos de treinamento depende do problema a ser tratado.

Quanto ao tempo de treinamento, vários fatores influenciam sua duração, sendo necessário utilizar algum critério de parada. O critério de parada do algoritmo backpropagation não é bem definido, podendo ser considerado um número máximo de ciclos, taxa de erro médio, capacidade de generalização da rede ou a combinação destes critérios. Em determinado instante do treinamento, a aprendizagem começa a degenerar e a rede se especializa no conjunto de dados do treinamento, perdendo capacidade de generalização. Este ponto pode ser outro critério para parada do treinamento.

O treinamento deve ser interrompido quando a rede apresentar boa capacidade de generalização e quando a taxa de erro for suficientemente pequena, ou seja, menor que o erro admissível. Assim, é necessário encontrar um ponto ótimo de parada com erro mínimo e capacidade de generalização máxima.

Com a conclusão do treinamento, pode-se realizar o teste da rede. Durante esta fase o conjunto de teste é utilizado para determinar a performance da rede com dados que não foram previamente utilizados. Em alguns casos, o teste é realizado durante as primeiras utilizações da rede. A performance da rede, medida nesta fase, é uma boa indicação de sua performance real.

### **3.7 APLICAÇÕES COM REDES NEURAIAS**

São muitas as aplicações de Redes Neurais em nosso cotidiano. No prognóstico de mercados financeiros, grupos de investimento utilizam redes neurais para analisar partes do mercado financeiro e auxiliarem em suas seleções e previsões. O reconhecimento óptico de caracteres (OCR) é outro tipo de aplicação que faz uso eficiente de Redes Neurais Artificiais.

Algumas aplicações bem sucedidas das técnicas de Redes Neurais Artificiais são: controle de processos industriais, aplicações climáticas e identificação de fraude no sistema de gerenciamento de cartões de crédito. O banco americano chamado Mellon Bank utilizou-se desta tecnologia para criar um sistema de detecção de fraudes de cartão de crédito, onde seus gastos com a implantação do sistema foram cobertos em seis meses pelo departamento de controle de fraudes.

A partir dessa experiência, vários outros bancos começaram a utilizar sistemas baseados em redes neurais para controlar fraudes via cartão de crédito. Estes sistemas têm capacidade de reconhecer uso fraudulento a partir de padrões criados no passado, com maior precisão que outros sistemas.

Outro exemplo da utilização de Redes Neurais é no auxílio de tomada de decisões durante diagnóstico médico. Em seu aprendizado, submete-se uma série de diagnósticos de pacientes, com diferentes características, variados sintomas, bem como seus resultados de exames e os diagnósticos médicos para cada enfermidade. Então quando apresentados os dados de um novo paciente, e seus respectivos sintomas, a rede fornecerá um diagnóstico para os novos casos, baseado em experiências anteriores. Isto essencialmente criará um sistema com o conhecimento de vários médicos, e fornecerá um diagnóstico inicial em tempo real ao profissional, atuando como uma ferramenta auxiliar ao diagnóstico médico.

Outros exemplos de aplicações de Redes Neurais de Computadores:

- Análise e processamento de sinais;
- Controle de processos;
- Robótica;
- Navegação de robôs;
- Classificação de dados;
- Reconhecimento de padrões em linhas de montagem;
- Filtros contra ruídos eletrônicos;
- Análise de imagens;
- Análise de voz;
- Avaliação de crédito;

- Simulações de circuitos eletrônicos;
- Reconhecimento de voz;
- Reconhecimento de imagem (Íris do olho e outros).

## CAPÍTULO 4

### Sistema de Avaliação Inteligente Conexionista e Colaborativa

---

#### 4.1 INTRODUÇÃO

No Capítulo 2 verificamos o desenvolvimento de diferentes métodos em Educação visando facilitar o processo de avaliação. Em particular, avaliações subjetivas (dissertativas compostas por questões abertas) não apresentam vantagens em relação ao tradicional método lápis e papel; pois, no final do processo, o professor necessita corrigir todas as avaliações.

A proposta deste trabalho (registrada no Capítulo 1) é apresentar e testar empiricamente uma ferramenta que permita a aplicação de avaliações subjetivas, desenvolvida a partir do conceito do SAICO, acelerando assim significativamente o processo de correção de questões abertas e contribuindo para o aprendizado do aluno.

Em linhas gerais, o Sistema de Avaliação Inteligente Conexionista e Colaborativa - SAICO combina a idéia de ranking (ordenação crescente ou decrescente) com a habilidade das redes neurais em descobrir mapeamentos ocultos. O ranking é construído através da participação dos próprios alunos (que vivenciam uma experiência interessante e oportuna de aprendizagem) enquanto a Rede Neural aponta as melhores posições dignas de correção pelo tutor humano (professor).

## 4.2 CONCEITOS FUNDAMENTAIS

### 4.2.1 Definição de termos

Dentre os termos necessários à compreensão da proposta apresentada, destacamos:

- Número de questões da avaliação (NQA): variável determinada pelo professor, informando a quantidade de questões contidas na avaliação.
- Número de respostas para ranking (NRR): quantidade de respostas apresentadas simultaneamente aos alunos, após a realização da avaliação, para a construção de um ranking (julgamento). Variável fixada em cinco respostas neste experimento.
- Nota tradicional (NT): variável que contém a nota do aluno, atingida pela correção tradicional de um tutor humano (professor). NT serve como referencial de comparação para a nota inferida pelo sistema (NI).
- Nota Inferida (NI): variável que contém a nota do aluno gerada pelo sistema, proveniente de todo o processo de ordenação (realizada pelos alunos), correção (apontada pela rede neural e realizada pelo professor) e síntese (realizada por um algoritmo de interpolação).

- **Ponto de Corte (PC):** posição no ranking de notas finais, indicada pela rede neural, que otimiza a correlação entre notas geradas através da correção tradicional e notas inferidas pelo sistema.
- **Posição média no ranking (PMR):** Após o julgamento, o sistema calculará a posição média no ranking (PMR) de cada “resposta” em relação às demais. Sinteticamente, PMR é a variável que determina o “grau de acerto” da resposta de cada aluno.

O valor da Posição Média no Ranking (PMR) é calculado conforme a Equação 2 (ver abaixo).

$$\text{PMR} = \frac{\mathbf{N1 \times P1 + N2 \times P2 + \dots + Nk \times Pk}}{\mathbf{N1 + N2 + \dots + Nk}}$$

Equação 2 - Posição Média no Ranking

Onde, N1 = Número de vezes que a questão aparece em 1º no Ranking;

P = Peso da Posição no Ranking (1º = 1, 2º = 2, ... ,k-ésimo = k);

k = Quantidade de respostas para ranking (NRR).

#### 4.2.2 Interpolação Linear

Interpolação é um procedimento de estimação do valor de um atributo em locais não amostrados, a partir de pontos amostrados na mesma área ou região.

Seja uma tabela com valores de  $x$  e  $y$  associados. Para se obter o valor de  $y$  para um valor específico de  $x$  não constante na tabela, é necessário explicitar uma função relacionando tais variáveis. As funções mais utilizadas para modelar estas relações são obtidas através de métodos de interpolação. Dentre os métodos para interpolar a partir de um conjunto de dados, destacamos: interpolação linear e quadrática, polinômios interpoladores (Lagrange, Newton e Gregory-Newton) e interpolação por Splines.

Como o sistema proposto sintetiza as notas de toda turma, baseado na correção executada pelo professor de algumas respostas selecionadas pela Rede Neural (e no ranking construído pelos próprios alunos), temos a necessidade de interpolar (e talvez extrapolar) de modo a atingir as respostas de todos os alunos envolvidos.

### **4.2.3 Coeficiente de Correlação Linear de Pearson**

Instrumentos de medida são válidos quando medem efetivamente a característica para a qual foram concebidos. Ao indagar se um instrumento é válido, o que se quer saber se este realmente mede, de forma efetiva a variável de interesse.

O termo correlação significa relação em dois sentidos (co + relação) sendo usado em estatística para designar a força que mantém unido dois conjuntos de valores (COSTA 1998).

Para medir a tensão elétrica de uma bateria, pode-se, por exemplo, ligar uma lâmpada em seus pólos e, de acordo com a claridade obtida, exprimir se a tensão está forte ou fraca. Tal medição é imprecisa, pois temos conceitos distintos sobre a classificação de forte e fraca. Entretanto, nas mesmas condições, se em lugar de uma lâmpada for colocado um medidor de tensão (voltímetro), a tensão exata da bateria será definida.

Para medir tensões elétricas, um medidor de tensão é um instrumento mais válido que uma lâmpada. A qualidade de um experimento está diretamente relacionada aos conceitos de validade interna e externa. Então, a questão não é somente efetuar a medição, mas medir precisamente.

Uma das maneiras possíveis de avalia-se validade é através de um recurso estatístico denominado correlação. Há várias maneiras de medir correlação; uma delas - a correlação linear simples - resulta do cálculo de um coeficiente (Pearson) que varia de -1 a +1. Assim, dado um instrumento cuja validade seja desconhecida, basta efetuar com este e simultaneamente com outro instrumento cuja validade seja conhecida as mesmas medições e calcular em seguida, o coeficiente de correlação. Quanto mais próximo de +1 for o coeficiente de Pearson, mais válido o instrumento. Para calcular a correlação utiliza-se a fórmula apresentada na Equação 3:

$$r = \frac{\sum (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\left(\sum (x_i - \bar{x})^2\right)\left(\sum (y_i - \bar{y})^2\right)}}$$

Equação 3 - Correlação

Quanto maior o valor absoluto de r, mais forte a associação. A Tabela 2 fornece diretrizes para a descrição da correlação em relação a seu valor numérico. As interpretações (afirmações na área específica do problema em questão), entretanto, dependem particularmente de cada contexto.

Tabela 2 - Interpretação do Valor do Coeficiente de Pearson

Valor absoluto de r	Interpretação
0	Correlação Inexistente
0,01 a 0,25	Correlação Fraca
0,26 a 0,50	Correlação Média
0,50 a 0,75	Correlação Forte
0,76 a 0,100	Correlação Perfeita

### 4.3 ARQUITETURA DO SISTEMA

O sistema proposto (Sistema de Avaliação Inteligente Conexionista Baseado em Colaboração - SAICO) é, em termos sintéticos, uma ferramenta que utiliza a avaliação concebida pelo professor, aplica-a aos alunos e corrige as questões automaticamente (baseando-se em poucas correções do professor).

Para facilitar a compreensão do sistema, o processo é dividido em cinco fases: elaboração, execução, ranking, correção e síntese. O professor participa da elaboração e correção, enquanto alunos participam da execução e ranking. A Rede Neural é responsável pela indicação das provas a serem corrigidas pelo professor, enquanto o algoritmo de interpolação realiza a síntese de notas, na última fase.

#### 4.3.1 Fase de Elaboração

Nesta fase, o professor elabora a avaliação, fornecendo os enunciados de cada questão, o peso e a resposta esperada (“gabarito”). Em seguida elege o número de respostas (NRR) a serem apresentadas aos alunos na fase de ranking, neste experimento fixado ao número determinado de cinco respostas (Figura 4.1).

>> Cadastro de Prova/Questões Usuário: Weber Martins

Nome:  Maior ranking:   
 Autor:  Menor ranking:   
 Disciplina:  Inter ranking:

**ufgsabprim**

Questões já adicionadas

Número	Enunciado	Resposta	Alternativas	Peso	Aberta
1	A fim de estabelecer uma leitura de estudo, quais aspectos devem ser considerados para resultar em uma leitura proveitosa? (mínimo 04 aspectos)	Atenção, intenção, reflexão, espírito crítico, análise e síntese.		2,0	Sim <input type="radio"/>
2	A leitura informativa engloba várias fases ou etapas, como leitura de reconhecimento, exploratória, seletiva, reflexiva, crítica e interpretativa. Defina leitura seletiva.	Leitura seletiva é a leitura que resulta na seleção das informações mais importantes relacionadas com o problema em questão, eliminando o supérfluo e concentrando-se nas informações verdadeiramente pertinentes ao problema.		2,0	Sim <input type="radio"/>
3	Conceitue resumo.	O resumo é a apresentação concisa e seletiva do texto, destacando-se os elementos de maior interesse e importância, isto é, as principais idéias do autor da obra.		2,0	Sim <input type="radio"/>
4	De acordo com as autoras Marconi e Lakatos, defina seminário e sua finalidade no âmbito educacional?	Seminário é uma técnica de estudo que inclui pesquisa, discussão e debate. Cujas finalidades são pesquisar e ensinar a pesquisar.		2,0	Sim <input type="radio"/>

---

Pesquisa de Questões - Parâmetros

Disciplina:  Professor:  Palavra chave:

Enunciado	Peso	Aberta
		<input type="button" value="Add"/>

Figura 4.1 - Tela de Cadastro de Testes (questões e pesos)

### 4.3.2 Fase de Execução

Nesta fase, aplica-se a avaliação. O professor seleciona o teste previamente cadastrado, determina o tempo limite disponibilizado para o aluno responder as duas etapas que compõem o mesmo, e em seguida liberando-o (Figura 4.2). Os alunos, individualmente ou em grupo (a critério do professor), acessam a ferramenta para registrar suas respostas (Figura 4.3). Neste experimento optou-se pelo acesso simultâneo e individual.

>> Aplicação de provas Usuário: Weber Martin

Prova:

Turma:

Primeira etapa:  minutos

Segunda etapa:  minutos

39 item(ns)

ID	Data	Prova	Turma	1ª Etapa (Min)	2ª Etapa (Min)
1 17766	04/12/2009 19:45	testedefesa	Engenharia Sanitaria	15	10
2 16079	07/04/2009 19:04	ufgsecseg	Engenharia Civil - 6P - TurmaB	30	20
3 15211	07/04/2009 18:02	ufgsecseg	Engenharia Civil - 6P - TurmaA	30	25
4 15015	06/04/2009 15:19	testesisa	Engenharia Sanitaria	15	15
5 14769	31/03/2009 20:01	testeufg	Engenharia Sanitaria	2	10
6 14029	31/03/2009 16:01	ufgteraseg	Metodologia Cientifica - Weber - Terça 02	25	20
7 13333	31/03/2009 13:56	ufgteraseg	Metodologia Cientifica - Weber - Terça 01	30	20
8 13091	30/03/2009 14:25	testeufg	Engenharia Sanitaria	15	15
9 12441	28/03/2009 08:07	ufgsabseg	Metodologia Cientifica - Weber Sábado 28/03	30	25
10 10828	24/03/2009 19:45	ufgcbprim	Engenharia Civil - 6P - TurmaB	30	30
11 9729	24/03/2009 18:15	ufgcbprim	Engenharia Civil - 6P - TurmaA	30	30
12 8951	24/03/2009 16:46	ufgterbprim	Metodologia Cientifica - Weber - Terça 02	30	30
13 8245	24/03/2009 14:16	ufgteraprim	Metodologia Cientifica - Weber - Terça 01	30	30
14 7469	21/03/2009 08:06	ufgsabprim	Metodoloia Cientifica - Weber - Sábado 01	45	30
15 7017	20/03/2009 14:20	testeufg	Engenharia Sanitaria	60	45

Figura 4.2 - Tela de Aplicação do Teste

>> Resolução de Prova

Prova: testedefesa

**Tempo restante:**

**1 a) Defina esforços internos.**

Resposta:

Você tem 352 caracteres restantes.

**2 a) Conceitue resumo.**

Resposta:

Você tem 386 caracteres restantes.

Figura 4.3 - Tela de Registro de Respostas (etapa 01)

### 4.3.3 Fase de Ranking ou Julgamento

Na fase de ranking (ou julgamento), cada aluno realiza a ordenação (estabelece o ranking) das respostas selecionadas pelo sistema (com base no gabarito informado previamente pelo professor). Cinco respostas são apresentadas aos alunos para serem julgadas (Figura 4.4).

É importante ressaltar que, a autoria das respostas analisadas está oculta. O aluno julga sem saber quem são os autores das respostas, eliminando a possibilidade de envolvimento emocional entre colegas durante a fase de julgamento.

>> Resolução de Prova

Prova: testedefesa

**Tempo restante:**

**Defina esforços internos.**

Resposta: São resultantes de tensões atuantes ao longo de uma determinada seção transversal.

limpar

1 a) Resposta aluno 01

4 b) Resposta aluno 02

2 c) Resposta aluno 09

--- d) Resposta aluno 05

e) Resposta aluno 03

Rascunho: 1-4-2-5-3 ok

Prosseguir

Figura 4.4 – Tela de Registro do Ranking (Etapa 02)

### 4.3.4 Fase de Correção

Na fase de correção, o professor (ou outra pessoa com conhecimentos adequados) corrige pessoalmente as respostas selecionadas pela rede neural. É importante observar que, nem sempre, tais respostas serão propriedade de um mesmo aluno. A correção de algumas respostas é necessária para que o sistema tenha pontos de apoio (notas referentes aos pontos de corte) para sustentar a inferência das notas finais (NI). Ressalvando que nem sempre a melhor nota de uma questão atinge o valor máximo, cabendo ao professor decidir quanto vale cada resposta em particular. A Rede Neural aponta os pontos de apoio mais importantes para o sucesso da síntese das notas por meio de interpolação (Figura 4.5).

>> Teste Extrato - Correção - ufgecseg Usuário: Ademir Aparecido do Prado

Questão 1 - Como são chamados e qual o significado físico dos coeficientes delta  $d_{i0}$  das equações (de equilíbrio) do Método das Forças?  
 "São chamados de termos ou fatores de carga e representam os deslocamentos nas direções  $i$  provocados pela solicitação externa."

Posição	Num.	MR	Nota	Resposta
1	16199	4,333	8,8	Fatores de carga, significa um deslocamento em $i$ , produzido pelo carregamento aplicado na estrutura.
2	16165	3,400	6,7	Tais coeficientes das equações são chamados de fatores de carga e representam os deslocamentos nas direções $i$ provocados por ações externas.
3	16403	1,125	3,4	É o deslocamento causado pela configuração de cargas

Questão 2 - Como são chamados e qual o significado físico dos coeficientes delta  $d_{ij}$  das equações (de equilíbrio) do Método das Forças?  
 "São chamados de coeficientes de flexibilidade e representam a relação entre o deslocamento na direção  $i$  provocado por uma força na direção  $j$ ."

Posição	Num.	MR	Nota	Resposta
1	16182	4,600	90,0	são chamados de coeficientes de flexibilidade e representam a relação entre o deslocamento na direção $i$ provocado por uma carga na direção $j$ .
2	16148	3,000	8,0	Os coeficientes delta $d_{ij}$ são denominados coeficientes de flexibilidade. O significado físico é a representação da relação entre o deslocamento provocado por $u$ na direção $i$ .
3	16267	1,400	2,0	Representa o deslocamento real em um ponto qualquer, resultante de uma força e direção arbitrária aplicada em certo ponto.

Questão 3 - Defina Grau de Hiperestaticidade.  
 "É o número de vínculos que excedem ao necessário para garantir o equilíbrio estático."

Posição	Num.	MR	Nota	Resposta
1	16148	4,371	10,0	É o número de vínculos além do necessário para que se obtenha o equilíbrio estático na estrutura.
2	16114	3,000	8,0	É o número de vínculos que excedem ao necessário para garantir o equilíbrio estático da força na direção $j$ .
3	16233	1,667	3,0	Grau de hiperestaticidade é as reações (vínculos) a mais que estão proporcionando que a estrutura seja hiperestática. Grau 2, por exemplo significa que ela tem duas reações a mais que acaba não deixando ela isostática.

Figura 4.5 – Tela de Correção (tutor humano)

### 4.3.5 Fase de Síntese de Notas

O processo de síntese tem início pelo cálculo da posição média da resposta no ranking (PMR). Obtidas tais posições, o sistema calcula as notas inferidas (NI), aplicando interpolação (e/ou extrapolação), a partir da PMR da resposta e as notas conferidas pelo professor na fase de correção (Figura 4.6). Em seguida, em avaliações contendo mais de uma questão, agrupam-se as respostas pertencentes a cada aluno, somando as notas das respostas (calculadas nesta fase), obtendo assim a nota final da avaliação (Figura 4.7).

>> Teste Extrato - Correção - ufgecseg Usuário: Ademir Aparecido do Prado

Questão 1 - Como são chamados e qual o significado físico dos coeficientes delta di0 das equações (de equilíbrio) do Método das Forças?  
 "São chamados de termos ou fatores de carga e representam os deslocamentos nas direções i provocados pela solicitação externa."

Posição	Num.	MR	Nota	Resposta
1	16199	4,333	8.8	Fatores de carga, significa um deslocamento em i, produzido pelo carregamento aplicado na estrutura.
2	16335	4,333	8.8	Representam os deslocamentos nas direções "i" provocados por solicitações externas. São chamados de termos (ou fatores) de carga.
3	16386	4,200	8.5	São conhecidos como termos ou fatores de carga e representam os deslocamentos na regiões i provocadas por solicitações externas.
4	16182	4,167	8,425	di0 são chamados de termos ou fatores de carga e representam os deslocamentos nas direções i provocados por uma força atuante(externa),.
5	16216	4,000	8.05	Os coeficientes di0 das equações de equilíbrio do método das forças são os chamados fatores de carga, que são os deslocamentos sofridos na direção i devido ao carregamento.
6	16369	4,000	8.05	Esses coeficientes são chamados termos de carga ou fatores de carga. Eles não representam flexibilidade, podendo representar deslocamento ou rotação (deslocamento angular). São a única parte da equação dependente do carregamento no método das forças.
7	16114	3,667	7.3	São chamados de Termo ou Fator de Carga e representam os deslocamentos resultantes da atuação das cargas externas nas direções "i".
8	16148	3,500	6,925	Os coeficientes di0 também são chamados de fatores de carga, representando os deslocamentos na direção i provocados pela solicitação externa.
9	16165	3,400	6.7	Tais coeficientes das equações são chamados de fatores de carga e representam os deslocamentos nas direções i provocados por ações externas.
10	16284	3,400	6.7	Fatores de carga. Representam os deslocamentos nas direções i provocados pelas forças atuantes..
11	16250	3,000	6.11978	São chamados de fatores de carga e representam os deslocamentos provocados pela solicitação externa.
12	16318	2,857	5,91256	Termo ou fator de carga : são os deslocamentos provocados pelas cargas externas.
13	16080	2,200	4,95934	São fatores de carga que apresentam os deslocamentos na direção i, do carregamento 0.
14	16097	2,000	4,66923	É CHAMADO DE DESLOCAMENTO , O "1" SIGNIFICA A DIREÇÃO DESSE DESLOCAMENTO E O "0" SIGNIFICA QUE O CARREGAMENTO É O ORIGINAL..
15	16233	2,000	4,66923	São chamados de deslocamento. O i indica a direção do deslocamento e O indica o que é provocado pela força atuante.
16	16352	2,000	4,66923	SÃO CHAMADOS DE TERMOS DE CARGA.ELES REPRESENTAM FLEXIBILIDADE,SIGNIFICANDO DESLOCAMENTO OU ROTAÇÃO.
17	16131	1,500	3,94396	Representam os deslocamentos provocados por forças.
18	16267	1,500	3,94396	Representa o deslocamento real em um ponto, resultante da aplicação de uma força qualquer no mesmo ponto.
19	16403	1,125	3,4	É o deslocamento causado pela configuração de cargas
		0,000	0.0	Sem resposta

Figura 4.6 – Tela de Notas Inferidas pelo SAICO

>> Teste Extrato - ufgecseg Usuário:

	Q1			Q2			Q3			Q4			Q5			Q6			Q7			Q8			Média R	Média N
	P1	R1	N1	P2	R2	N2	P3	R3	N3	P4	R4	N4	P5	R5	N5	P6	R6	N6	P7	R7	N7	P8	R8	N8	MR	MN
1	1.5	2,200	4,959	1.5	4,167	67,792	1.5	4,000	9,273	1.0	3,857	8,971	1.0	3,875	5,000	1.0	3,667	10,000	1.0	4,333	10,000	1.5	3,750	7,125	3,731	16,769
2	1.5	2,000	4,669	1.5	3,400	28,500	1.5	2,000	4,250	1.0	2,250	8,000	1.0	1,571	0,000	1.0	2,333	9,048	1.0	2,250	6,250	1.5	3,600	7,020	2,426	8,996
3	1.5	3,667	7,300	1.5	3,200	18,250	1.5	3,000	8,000	1.0	4,500	10,000	1.0	3,333	5,000	1.0	2,333	9,048	1.0	3,500	10,000	1.5	3,250	6,775	3,348	9,454
4	1.5	1,500	3,944	1.5	1,667	3,000	1.5	3,000	8,000	1.0	1,333	8,000	1.0	1,667	0,333	1.0	2,750	9,643	1.0	1,500	1,136	1.5	3,429	6,900	2,106	5,188
5	1.5	3,500	6,925	1.5	3,000	8,000	1.5	4,571	10,000	1.0	1,667	8,000	1.0	5,000	5,000	1.0	2,500	9,286	1.0	3,500	10,000	1.5	2,000	6,500	3,217	7,942
6	1.5	3,400	6,700	1.5	3,333	25,083	1.5	3,000	8,000	1.0	3,500	8,400	1.0	3,000	5,000	1.0	4,286	10,000	1.0	3,500	10,000	1.5	2,444	6,544	3,308	10,289
7	1.5	4,167	8,425	1.5	4,600	90,000	1.5	2,500	6,125	1.0	2,600	8,000	1.0	3,571	5,000	1.0	3,125	10,000	1.0	3,500	10,000	1.5	3,000	6,600	3,383	19,973
8	1.5	4,333	8,800	1.5	1,500	2,375	1.5	2,000	4,250	1.0	1,857	8,000	1.0	3,143	5,000	1.0	2,250	8,929	1.0	1,500	1,136	1.5	3,000	6,600	2,448	5,610

Figura 4.7 – Tela de Síntese de Notas

#### 4.4 TRABALHOS CORRELATOS

A revisão da literatura mostra outras tentativas utilizando-se de métodos e abordagens distintas no enfrentamento automático de questões abertas e dissertações.

Ainda que pesquisas como a de [Burstein, Chodorow, Leacock 2004], [Yang 2005], [Haberman 2004], [Holland and Hoskens 2003] e [Steinhart 2001], tenham apresentado resultados considerados satisfatórios entre as correções automatizadas e correções efetuadas por tutores humanos, não atingiram níveis de correlação suficiente para credenciá-las para aplicação imediata.

## Parte II - Validação Empírica

### CAPÍTULO 5

#### Experimentos e Resultados

---

##### 5.1 INTRODUÇÃO

Este capítulo descreverá a amostra, material e instrumentos utilizados na experimentação, bem como o delineamento experimental.

Posteriormente são analisados os resultados dos dados empíricos obtidos na situação experimental específica proposta neste trabalho. Tais resultados evidenciam as características amostrais (análise descritiva) e verificam possíveis generalizações (análise inferencial).

##### 5.2 AMOSTRA

A amostra selecionada foi composta por 38 alunos matriculados no sétimo período do curso de Engenharia Civil da Universidade Federal de Goiás - UFG, cuja frequência às aulas foi igual ou superior a 70%. Trata-se de grupo heterogêneo composto de homens e mulheres com idade média entre 16 e 26 anos.

## 5.3 METODOLOGIA

### 5.3.1 Método de Abordagem

Tendo em vista que nossa hipótese básica está operacionalmente definida e foram coletados dados em pesquisa de campo, adotou-se o método hipotético-dedutivo. O método visa definir e explicar racionalmente relações entre as variáveis observadas na pesquisa científica.

Em uma investigação científica (hipotético-dedutivo), usualmente segue-se os passos abaixo descritos:

- Um problema é identificado;
- Uma solução (hipótese, passível de ser testada) é formulada;
- Esta hipótese é testada.

### 5.3.2 Método de Procedimento

Os métodos de procedimento aplicados neste projeto são citados e justificados a seguir:

- **Método estatístico:** amostra grande (grupo mínimo de 38 alunos por avaliação);
- **Método comparativo:** análise do grupo, comparando os resultados (notas) resultantes da avaliação corrigida por tutor humano ao do sistema SAICO.

### 5.3.3 Técnicas

As técnicas empregadas neste projeto são:

- **Pesquisa bibliográfica:** visando aperfeiçoar a teoria de base e revisão bibliográfica;
- **Construção de protótipo:** O software SAICOweb (ferramenta de coleta de dados). Capaz de aplicar e corrigir avaliações computadorizadas contendo questões abertas, implementando o método e as fases descritas no SAICO proposto por Guimarães (2004);
- **Pesquisa documental:** Realizada através de dados dos participantes (frequência dos alunos, notas obtidas no sistema tradicional de avaliação e SAICO);
- **Formulário:** utilizado para medir o nível de satisfação de aluno e professor.

### 5.3.4 Variáveis

Dentre as variáveis que foram manipuladas nesta pesquisa, destacam-se:

- **Variável independente:**  
**X1** = Uso do Sistema de Avaliação Inteligente Conexcionista e Colaborativa - SAICO em ambiente real de sala de aula (nível nominal: sim/não).
- **Variáveis dependentes:**  
**Y1** = Correlação entre notas: SAICO e tutor humano (nível ordinal: baixo, médio, alto);

**Y2** = Grau de absorção do conhecimento assimilado pelos alunos (nível racional: média final da nota obtida);

**Y3** = Grau de satisfação do aluno (nível ordinal: muito baixo, baixo, médio, alto, muito alto);

**Y4** = Grau de satisfação do professor (nível ordinal: muito baixo, baixo, médio, alto, muito alto).

- **Variável de controle:**

**C1** = frequência às aulas (percentagem em relação ao curso completo, usando apenas os alunos que atingirem um mínimo de 70%).

### 5.3.5 **Material e Instrumentos**

Para observação e registro dos valores das variáveis foi tomado o depoimento do professor participante do experimento, desenvolvido um questionário e um website (software) como instrumentos de coleta de dados, descritos abaixo:

- Questionário de **satisfação do aluno** (Apêndice I).
- **Depoimento do professor** participante do experimento (Apêndice II).

Para aplicação de testes utilizando questões abertas, foi desenvolvido o website **Software SAICOWeb** que aplica testes automatizados contendo questões abertas, implementando as fases do Sistema de Avaliação Inteligente Conexcionista e Colaborativa - SAICO de acordo com o Capítulo 4.

### **5.3.6 Delineamento Experimental**

O experimento consistiu na aplicação de testes utilizando o software SAICOWeb no primeiro semestre de 2009. Utilizando as instalações do laboratório de informática da Universidade Federal de Goiás, os 38 alunos participantes responderam a dois testes compostos de oito questões abertas cada (referente ao conteúdo da disciplina de Análise Estrutural II), com intervalo de quinze dias entre a realização dos testes.

Cada aluno dispunha de terminal individual de computador para registrar suas respostas (sem consulta) diretamente no software SAICOWeb. Encerrado o teste, os alunos responderam ao formulário (questionário de satisfação do aluno) apresentado no Apêndice I e ao final do processo de coleta de dados, o professor participante do experimento apresentou seu depoimento a cerca da experiência vivenciada (depoimento do professor) detalhado no Apêndice II.

## **5.4 RESULTADOS**

Neste trabalho, foram considerados válidos os dois testes realizados para coleta de dados, conforme descrito na seção 5.3.6.

### **5.4.1 Análise Descritiva**

Análise do Grau de satisfação dos alunos em relação ao software SAICOWeb utilizado na coleta de dados. Na Figura 5.1, observou-se o alto índice de satisfação, acima de 80% de aceitação entre os participantes.



Figura 5.1 - Grau de Satisfação: Software (SAICOWeb)

Análise do Grau de satisfação dos alunos em relação ao método de avaliação. Na Figura 5.2, observou-se o alto índice de satisfação, acima de 80% de aprovação entre os participantes.

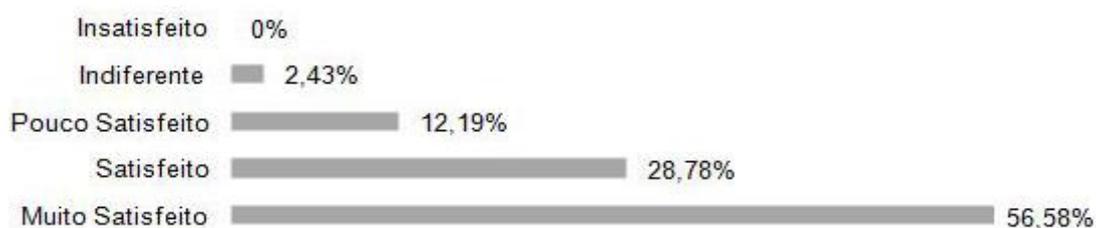


Figura 5.2 - Grau de Satisfação: Método de Avaliação

Análise do Grau de satisfação dos alunos em relação à experiência pedagógica vivenciada durante a fase de ranking (julgamento). Na Figura 5.3, observou-se novamente o índice acima de 85% de satisfação entre os participantes.

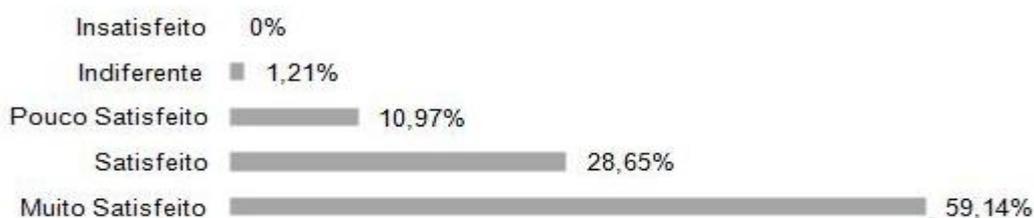


Figura 5.3 - Grau de Satisfação: Experiência Pedagógica

Análise do Grau de satisfação dos alunos em relação ao aspecto de anonimato entre aluno e professor no processo avaliativo. Na Figura 5.4, observou-se mais uma vez o alto índice de satisfação, acima de 85% de aceitação entre os participantes.

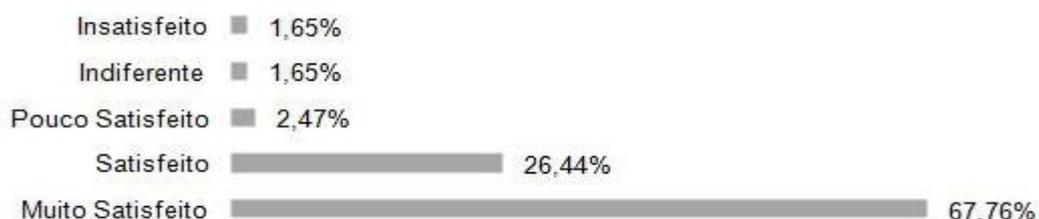


Figura 5.4 - Grau de Satisfação: Anonimato

Análise do Grau de satisfação do professor participante do experimento foi tomada baseando-se em seu depoimento pessoal a cerca da experiência vivenciada durante a pesquisa (apresentado no Anexo II). O relato permite concluir como muito alto o índice de satisfação por parte do docente (o resultado não foi apresentado em forma de gráfico por se tratar do parecer de um único professor).

#### 5.4.2 Análise Inferencial

Comparando os resultados (notas) obtidos pelos alunos, através da correlação (Coeficiente de Pearson,  $R_{xy}$ ), observa-se forte correlação (0,81) entre notas atribuídas pelo SAICO e notas corrigidas por tutor humano (referência). O sistema manteve mais de 80% das diferenças entre notas inferidas e notas-referência abaixo de 2,0 pontos (ver Figura 5.5). Tais diferenças são comuns mesmo entre docentes ou bancas de avaliação conforme descrito na seção 2.7.

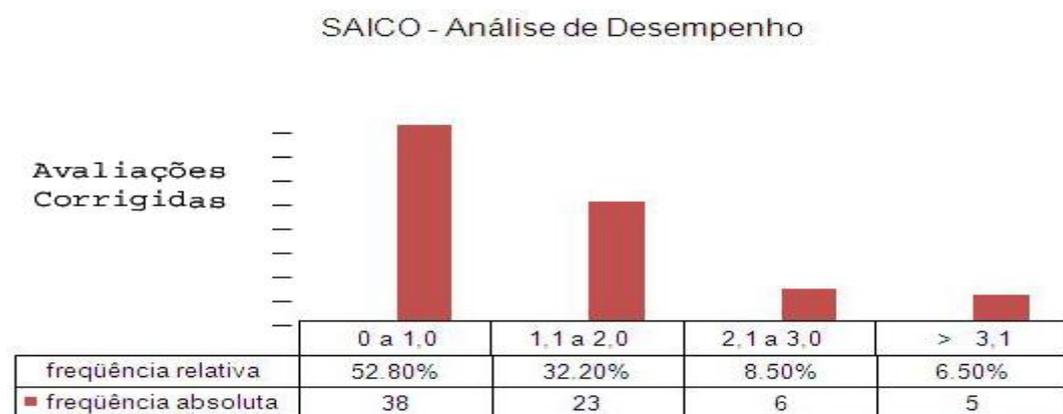


Figura 5.5 - Análise de desempenho

## CAPÍTULO 6

### Conclusão

---

#### 6.1 INTRODUÇÃO

Em conclusão, o sistema SAICO foi validado empiricamente de forma satisfatória para uma única turma composta de 38 alunos. Os resultados confirmam hipótese básica e secundárias apresentadas na seção 1.3.3.

Outros testes empíricos devem ser organizados para validar o sistema em diferentes ambientes. Entretanto, o sucesso no tratamento automático de avaliações com questões abertas (dissertativas), usando colaboração de alunos e professor, evidencia que tal método de correção ganha o merecido espaço no processo de ensino-aprendizagem.

#### 6.2 PRINCIPAIS CONTRIBUIÇÕES

As contribuições deste trabalho podem ser resumidas em duas principais:

A primeira é o alto potencial do sistema SAICO quando aplicado em ambiente real, demonstrado a partir dos resultados satisfatórios obtidos (correlação 0,81).

A segunda é a excelente aceitação do sistema SAICO por aluno e professor, observado nos altos índices apresentados no formulário “grau de satisfação do aluno” e no “depoimento do professor” participante do experimento.

### 6.3 DISCUSSÃO E TRABALHOS FUTUROS

A análise de dados evidencia como sugestão de trabalhos futuros, os seguintes fatores para aprimoramento do sistema SAICO:

- Quando a pior resposta apontada pela Rede Neural para receber correção do professor merecer pontuação zero, é mais indicado descartar este resultado e tomar a pior resposta imediata registrada para uma nova correção, evitando possível tendência das notas inferidas pelo SAICO por interpolação linear para o extremo menor (zero).
- Adotar Rede Neural com maior número de saídas no intuito de obter maior número de posições “Ponto de Corte” aproxima os pontos de referência quando efetuada a interpolação linear. Tais medidas permitem promover maior uniformidade entre as notas inferidas pelo SAICO e notas corrigidas pelo tutor humano (professor).
- Não permitir a utilização de questões abertas de resposta restrita, pois tais questões tendem a enfraquecer ou anular o efeito da fase colaborativa do processo, operação de extrema importância na obtenção de resultados fidedignos durante a fase de Síntese de Notas.

## REFERÊNCIAS BIBLIOGRÁFICAS

- (AFONSO 2000) AFONSO, A. J. Avaliação Educacional: regulação e emancipação - para uma sociologia das políticas avaliativas contemporâneas. São Paulo: Cortez Editora, 2000.
- (ALEKSANDER, ALBROW 1968) ALEKSANDER, I & ALBROW, R. C.. Adaptative logic circuits. Computer J. Maio 1968.
- (ALEKSANDER, KAN 1987) KAN. K. W.: ALEKSANDER. I. A Probabilistic Logical Neural Network for Associative Learning. In Proc. Of IEEE 1'st Anual Conference on Neural Networks, San Diego, 1987.
- (ALEKSANDER, MANDANI 1968) ALEKSANDER, I & MANDANI, E. H. Microcircuit learning net improved recognition by means of a pattern feddback. Eletronics Letters. Vol. 4 1968.
- (ALEKSANDER, THOMAS, BOWDEN, 1984) ALEKSANDER, I, THOMAS, W. V. and BOWDEN, P. A. Wisard: a radical step forward in image recognition. Sensor Review, vol 4, nr. 3 1984.
- (AUSUBEL 1978) AUSUBEL, D. at All Educational Psychology: A Cognitive View (2nd Ed.). New York: Holt, Rinehart & Winston, 1978.
- (AUSUBEL 2001) AUSUBEL, NOVAK Artigo: "Teoria de Aprendizagem significante e Mapas Conceituais", REVISTA CESGRANRIO, setembro de 2001.
- (BERCH 1997) BERCH, Magda. Avaliação pedagógica como fator para a construção de Estratégias de ensino em ambientes de ensino e aprendizagem computadorizados. 1997. Exame de qualificação - Programa de Pós- Graduação em Computação, Universidade Federal Do Rio Grande do Sul, Porto Alegre (1997).
- (BLOOM, HASTINGS, MADAUS 1971) BLOOM, B., HASTINGS e MADAUS . Handbook on Formative and Sumative Evaluation of Student Learning. New York: McGraw-Hill Book Company. (Manual de Avaliação Formativa e Sumativa do Aprendizado Escolar. S. Paulo: Livraria Pioneira Editora (1971).

- (BOAVIDA) BOAVIDA, Joaquim. Concepções e práticas de avaliação das aprendizagens de professoras do 1.º ciclo do ensino básico: Três estudos de caso. Tese de mestrado em ciências da educação (Avaliação em educação) não publicada (1996).
- (BONHAM 1992) BONHAM, L. A. Learning styles use in need of perspective. *Lifetime Learning*, v. 11 (5), 1988 apud DRUMMOND, R. J. & STODDARD, A. H. *Learning style and personality type. Perceptual and Motor Skills*, 1992.
- (BONNIOL, VIAL 2001) BONNIOL, J-J ; VIAL, M. Modelos de Avaliação : textos fundamentais. Porto Alegre: Artmed, 2001.
- (BORGES 2005) BORGES, Juliano. Escola e Disciplina: uma abordagem foucaultiana. *Revista Urutágua*, nº 05, Dez/Jan/Fev/Mar, DCS/UEM. Maringá: Paraná: 2005.
- (BRASIL 1996) (Código Civil - Brasil - Lei 9.394, de 20 Dez 1996, art. 24, inciso V - a)
- (BURKE, ROSEN, GOODMAN 1995) BURKE H., ROSEN, D. and GOODMAN P. Comparating the prediction accuracy of artificial neural networks and other statistical models for breast cancer survival. *Neural Information Processing Systems*. MIT Press.(1995).
- (BUSH 1999) BUSH, Martin, paper: “Alternative Marking Schemes for On-Line Multiple Choice Tests”, CAA Centre - South Bank University (School of Computing, Information Systems e Mathematics), on line - pdf, London - United kingdom, Jan - Dec, 1999, disponível em <http://www.caacentre.ac.uk/dldocs/BUSHMARK.pdf>. Acesso em 10 Dez 2008.
- (CARDINET 1993) CARDINET, J. Avaliar é Medir? Rio Tinto: Edições Asa. Julho de 1993.
- (CARDOSO 1998) CARDOSO, S. M. V. e JANDL, P. Jr. Estilos de Aprendizagem: aprendendo a aprender, Direito, USF, 1998.
- (CARDOSO 2001) CARDOSO, Rodrigo Ferrugem, UFRGS ,Brasil, Artigo: “Sistema interativo para gerência de questões e aplicação de avaliações na Web (AvalWeb)” Julho 2001
- (CARVALHO 2004) CARVALHO, Isabel Cristina de Moura. Educação Ambiental: A Formação do Sujeito Ecológico. São Paulo: Cortez, 2004).

- (CIVILETTI, RODRIGUES, LIMA 2001) CIVILETTI, RODRIGUES e LIMA, Campus Virtual da Universidade Gama Filho (UGF), Artigo “Educação Democrática”, Fonte: SECOM On Line - Secretaria de Comunicação do Estado do Tocantins, 30 de Agosto 2001
- (COSTA 1998) COSTA, Sergio Francisco. Introdução Ilustrada à Estatística 3ª edição. Capítulo 13 Correção Linear Simples (1998).
- (DE KETELE 1993) DE KETELE, J.M. L.Evaluation Conjuguée en Paradigmes. Revue Française de Pédagogie (1993).
- (DEPRESBITERIS 1997) DEPRESBITERIS, L. Avaliação da aprendizagem: revendo conceitos e posições. Em C. P. Sousa (Org.). Avaliação do rendimento escolar 6. ed. Campinas: Papyrus (1997).
- (DERRY 1989) DERRY, H. How students learn. Educational Leadership, v. 46 (4), p.4-10, 1988/89 apud DRUMMOND, R. J. & STODDARD, A. H. Learning style and personality type. Perceptual and Motor Skills, 1992.
- (DUBET 1998) DUBET, François. A formação dos indivíduos: a desinstitucionalização. Contemporaneidade & Educação. Rio de Janeiro, ano 3, n. 3, março 1998.
- (EHRMAN, OXFORD 1992) EHRMAN, M. & OXFORD, R. Adult language learning styles and strategies in an intensive training setting. Modern Language Journal, ,1990 apud DRUMMOND, R. J. & STODDARD, A. H. Learning style and personality type. Perceptual and Motor Skills, 1992.
- (ESTEBAN 2001) ESTEBAN, Maria Tereza. O Que Sabe Quem Erra- Reflexões Sobre Avaliação Escolar e o Fracasso Escolar. Rio de Janeiro: DP&A, 2001.
- (ESTEBAN 2003) ESTEBAN, Maria Teresa (org.). Escola, currículo e Avaliação. São Paulo. Cortez, 2003.
- (ESTEBAN 2005) ESTEBAN, Maria Teresa (org.). Escola, currículo e Avaliação. São Paulo. Cortez, 2005.
- (ESTEBAN 2008) ESTEBAN, Maria Teresa Revista Portuguesa de Educação, p. 5-31 CIED - Universidade do Minho.(2008)
- (FERNANDES 1993) FERNANDES, D. Avaliação na Escola Básica Obrigatória: Fundamentos para uma Mudança de Práticas.I.I.E.(1993).

- (FERNANDES 2005) FERNANDES, Domingos. Avaliação das Aprendizagens: Desafios às Teorias, Práticas e Políticas. Cacém: Texto Editores (2005).
- (FERNANDES et.al 1996) FERNANDES, Domingos; NEVES, Anabela; CAMPOS, Cristina & LALANDA, Joaquim. Das Concepções, Práticas e Organização da Avaliação das Aprendizagens à Formação de Professores. (Relatório do 1.º ano do Projecto PI/12/94 financiado pelo Instituto de Inovação Educacional) (1996).
- (FILHO 1990) FILHO, E. 1990. "Investigation of boolean neural network based on a novel Goal-Seeking-Neuron", Tese de Doutorado, Electronic Engineering Laboratories, University of Kent, UK (1990).
- (FOUCAULT 1987) FOUCAULT, Michael. Vigiar e Punir: Nascimento da Prisão. Petrópolis, Rio de Janeiro: Vozes, 1987.
- (FOUCAULT 1990) FOUCAULT, M. Vigiar e punir. 7ª ed., Petrópolis, Vozes (1990).
- (FREIRE 2003) FREIRE, Paulo. Segunda Carta do Direito e do dever de Mudar o Mundo. In: A Importância do ato de Ler. São Paulo: Cortez, 2003.
- (GADOTTI, 1984) GADOTTI, Moacir, Educação e poder: Introdução à Pedagogia do Conflito. 5 ed. São Paulo: Cortez -Autores Associados. 1984.
- (GADOTTI, 2003) GADOTTI, Moacir, Educação e poder: Introdução à Pedagogia do Conflito. São Paulo: Cortez , 2003.
- (GONDRA 2005) GONDRA, José Paul-Michel Foucault - Uma caixa de ferramentas para a história da educação? In: FARIA FILHO, Luciano (org.). Pensadores sociais e História da Educação. Belo Horizonte: Autêntica, 2005.
- (GUBA, LINCOLN 1981) GUBA, E. & LINCOLN, Y. Effective Evaluation. San Francisco: Jossey-Bass.(1981).
- (GUIMARÃES, 2004) GUIMARÃES, Fábio Teixeira. Avaliação Inteligente Conexionista Baseado em Colaboração. Dissertação (Mestrado) - Universidade Federal de Goiás, Escola de Engenharia Elétrica e de Computação, 2004.
- (GUYTON, 1986) GUYTON, Arthur C. Tratado de Fisiologia Médica. Rio de Janeiro: Editora Guanabara, 1986.

- (HAYDT 1997) HAYDT, Regina Cazaux - Avaliação do Processo Ensino-Aprendizagem. Editora Ática - São Paulo, 1997.
- (HAYDT 2002) HAYDT, Regina Cazaux - Avaliação do Processo Ensino-Aprendizagem. Editora Ática - 6<sup>o</sup> Edição, São Paulo, 2002.
- (HAYKIN 2001) HAYKIN, Simon. Redes Neurais. Princípios e prática. Porto Alegre, RS: Bookman, 2001.
- (HEBB 1949) HEBB, D. O.. *The organization of behavior*. New York: Wiley, 1949.
- (HOPFIELD) HOPFIELD, John, "Parallel Distributed Processing" (Processamento Distribuído Paralelo) editado por David Rumelhart e James McClelland.
- (HOPPER 1998) HOPPER, Mary. "Assessment in WWW-Based Learning Systems: Opportunities and Challenges". *Journal of Universal Computer Science*, Cambridge, v. 4, n. 4.
- (KANT 1996) KANT, Immanuel. Sobre a Pedagogia. Tradução de Francisco Cock Fontanella. Piracicaba: Editora Unimep, 1996.
- (KAPPEL, 1960) KAPPEL, F.R. Vitality in a business enterprise. Nova York: MacGraw-Hill, 1960.
- (KOHONEN 1989) KOHONEN, Teuvo, The Self-Organizing Map, Proceedings of the IEEE. (1989).
- (LEMOS 1993) LEMOS, V. O Critério do Sucesso - Técnicas de Avaliação da Aprendizagem. Porto: Texto Editora. 5<sup>a</sup> ed. (1993).
- (LEMOS et.al 1993) LEMOS, V.; Neves, A.; Campos, C.; Conceição, J. & Alaiz, V. (1993). A Nova Avaliação da Aprendizagem. O Direito ao Sucesso, 2<sup>a</sup> ed. Lisboa: Texto Editora.
- (LÉVY 2004) LÉVY, Pierre - As Tecnologias da Inteligência - O futuro do pensamento na era da informática. Editora 34 - 13 Edição, São Paulo, 2004.
- (LUCKESI 1995) LUCKESI, Cipriano C. Avaliação qualitativa. 5.ed. Campinas: Autores Associados, 1995. (Coleção Polêmicas do Nosso Tempo).

- (LUCKESI 2000) LUCKESI, Cipriano Carlos. A avaliação da aprendizagem escolar. 10ed., São Paulo. Cortez, 2000.
- (LUCKESI 2001) LUCKESI, website: [www.luckesi.com.br](http://www.luckesi.com.br). Entrevista concedida à revista Nova Escola sobre Avaliação da Aprendizagem. Novembro de 2001. acesso em 10 janeiro 2008.
- (LUCKESI 2003) LUCKESI, Carlos Cipriano - A Avaliação da Aprendizagem na Escola - Reelaborando conceitos e recriando a prática - Ed. Malabares Comunicações e Eventos, São Paulo, 2003.
- (LUCKESI 2005) LUCKESI, Carlos Cipriano - A Avaliação da Aprendizagem Escolar. São Paulo, 2005.
- (LUDKE, MEDIANO 1992) LÜDKE, Menga, MEDIANO, Zélia (Coord.). Avaliação na escola de 1º grau: uma análise sociológica. Campinas: Papyrus, 1992.
- (MARCONI, LAKATOS 2005) MARCONI, Marina de Andrade; LAKATOS, Eva Maria. Fundamentos de Metodologia Científica, 6º ed., São Paulo: Editora Atlas, 2005.
- (MARTINS 2004) MARTINS, Weber. et al. "A Novel Hybrid Intelligent Tutoring System and Its Use of Psychological Profiles and Learning Styles. Lecture Notes on Computer Science, v.3220, 2004.
- (MATTEI, 2008) MATTEI, Giovana. O professor e aluno com altas habilidades e superdotação: relações de saber e poder que permeiam o ensino, Revista "Educação Especial" n. 31, Santa Maria: 2008.
- (MCCULLOCH, PITTS 1943) MCCULLOCH, W. S.; PITTS, W. H. A logical calculus of the ideas immanent in neural nets, Bulletin of Mathematical Biophysics, v.5, 1943.
- (MCKENNA, BULL 1999) MCKENNA, Collen; BULL, Joanna, artigo: "Designing Effective Objective Test Questions: an Introductory Workshop, CAA Centre - Loughborough University, on line - pdf, Leicestershire - Reino Unido, Jun - Dez, 1999, disponível em <http://www.caacentre.ac.uk/dldocs/otghdout.pdf>. Acesso em 10 Jan 2009.
- (MINSKY 1969) MINSKY, M. L.; PAPERT, S. S. Perceptrons: na introduction to computational geometry, MIT Press, Cambridge, MA 1969.

- (MIRAS, SOLÉ 1992) MIRAS, N., Solé, I. La evaluación del aprendizaje y la evaluación en el proceso de enseñanza e aprendizaje. Desarrollo psicológico y educación II. Psicología de la Educación. Madrid. Alianza.(1992).
- (MORETTI 2007) MORETTI, Mércles Thadeu. Estudo das divergências entre corretores na avaliação de provas em matemática. Revista da Avaliação da Educação Superior (Campinas), vol.12 no.2 Sorocaba June 2007. Disponível em [http://www.scielo.br/scielo.php?script=sci\\_arttext](http://www.scielo.br/scielo.php?script=sci_arttext), acesso em setembro de 2009.
- (MORIN 2003) MORIN, Edgar. Os sete saberes necessários à educação do futuro, 8ª ed. São Paulo: Unesco/Cortez, 2003.
- (NATRIELLO 1987) NATRIELLO, G. The Impact of Evaluation Processes on Students. Educational Psychologist. (1987).
- (NEVO 1990) NEVO, D. Role of the Evaluator. In H. Walber & G. Haertel (Ed.), The International Encycloppedia of Educational Evaluation. Oxford: Pergamon Press. (1990).
- (NOIZET, CAVERNI 1985) NOIZET, G. & CAVERNI, J. Psicologia da Avaliação Escolar. Coimbra: Coimbra Editora (1985).
- (NOVA ESCOLA 1992) Revista NOVA ESCOLA. Como trabalhar o erro. Ano VII, nº60, set/1992. São Paulo, Abril Cultural, Fundação Vitor Civita.
- (OLIVEIRA, SANTOS 2005) OLIVEIRA, Katya Luciane de e SANTOS, Acácia Aparecida Angeli dos. Avaliação da aprendizagem na univesidade. Psicol. esc. educ., jun. 2005, vol.9, no.1, p.37-46, ISSN 1413-8557.
- (OSÓRIO, BITTENCOURT, 2000) OSÓRIO, F. ; BITTENCOURT, J. R. Sistemas inteligentes baseados em RNAs aplicados ao processamento de imagens. In: Workshop de Inteligência Artificial, Santa Cruz do Sul: UNISC, 2000.
- (PACHECO 1994) PACHECO, J. A Avaliação dos Alunos na Perspectiva da Reforma. Porto: Porto Editora, (1994).
- (PERRENOUD 1978) PERRENOUD, P. Das Diferenças Culturais às Desigualdades Escolares: a Avaliação e a Norma num Ensino Diferenciado. Análise Psicológica (1978).
- (PERRENOUD 1982) PERRENOUD, P. Não Mexam na Minha Avaliação! Para uma Abordagem Sistémica da Mudança Pedagógica. In A. Estrela e A. Nóvoa (Eds) . Avaliações em Educação: Novas Perspectivas (1982).

- (PERRENOUD 1999) PERRENOUD, Philippe. Avaliação: da excelência à regulação das aprendizagens - entre duas lógicas. Porto Alegre: ArtMed, 1999.
- (PERRENOUD 2000) PERRENOUD, Ph. Construindo competências. Revista Nova Escola, edição nº 135. São Paulo, SP: Abril Cultural, 2000.
- (PINTO 1996) PINTO, A. L. G. Avaliação formal e informal da aprendizagem em sala de aula - Psico-USF (1996).
- (POTTER 1998) POTTER, Donna J. Evaluation Methods Used in Web-based Instruction and the Online Course, Taming the Electronic Frontier. Disponível por WWW em [http://mason.gmu.edu/~dpotter1/1djp\\_611.html](http://mason.gmu.edu/~dpotter1/1djp_611.html).
- (RODRIGUES 2000) RODRIGUES, Alessandra P. O processo Avaliação de Ensino-Aprendizagem em Ensino à Distância: trabalho individual I. Porto Alegre: PPGC da UFRGS, 2000.
- (RODRIGUES 2007) RODRIGUES, M. M. M. Avaliação Educacional Sistêmica na Perspectiva dos Testes de Desempenho e seus Resultados. Estudo do Sistema Nacional de Avaliação do Ensino Básico (SAEB). Tese. Brasília - DF. Disponível em [http://repositorio.bce.unb.br/bitstream/10482/1622/1/Tese Doc Final Margarida.pdf](http://repositorio.bce.unb.br/bitstream/10482/1622/1/Tese%20Doc%20Final%20Margarida.pdf). Acesso em 16 de setembro de 2009.
- (RUMELHART, HINTON, WILLIAMS 1986) RUMELHART, D. E.; HINTON, G. E.; WILLIAMS, R. J. Learning internal representation by error propagation. Em Parallel distributed processing: explorations in the microstructures of cognition. MIT Press, 1986.
- (SARMENTO 1997) SARMENTO, Diva Chaves (org.). O Discurso e a Prática da Avaliação na Escola. Juiz de Fora: Pontes, 1997.
- (SCRIVEN 1967) SCRIVEN, M. (1967). The Methodology of Evaluation. In R. Tyler, R.M. Gagné e M. Sciven (Eds). Perspectives of Curriculum Evaluation, AERA Monograph Series on Curriculum Evaluation (1). Chicago: Rand Mac Nally.
- (SILVA 2006) SILVA, Rosemaria J. V. Regular e disciplinar: Análise de dispositivos disciplinadores da educação escolar na corte imperial UERJ/CNPQ/NEPHE, IV Congresso Brasileiro da História de Educação - 05 a 08 de novembro de 2006.
- (SKINNER 1972) SKINNER, Burrhus Frederic. Tecnologia do Ensino - Ed. Da Universidade de São Paulo, São Paulo, 1972.

- (SKINNER 1982) SKINNER, B.F. Sobre o Behaviorismo. Cultrix São Paulo, 1982.
- (SOARES 1981) SOARES, Magda. B. Avaliação educacional e clientela escolar. In: PATTO, M.H.S. Introdução à Psicologia Escolar. São Paulo: T.A . Queiróz, 1981.
- (SOBRINHO 2003) SOBRINHO, José Dias. Avaliação Democrática e Construção da Cidadania. In. Avaliação: Políticas Educacionais e Reformas da Educação Superior. São Paulo: Cortez, 2003.
- (TEIXEIRA, 2005) TEIXEIRA, Giselle Baptista. Caminhos do saber escolarizado: Produção, controle e circulação de livros nas escolas primárias da Corte Imperial. Rio de Janeiro: UERJ, Monografia, 2005.
- (TEIXEIRA, 2007) TEIXEIRA, Josele. Da preminência dos lugares a expulsão Escolar. Do Aspergimento de Flores à Repreensão Pública. (Re) Pensando a Avaliação da Aprendizagem. Rio de Janeiro: UERJ, 2007. Monografia de Conclusão de Curso de Pedagogia.
- (TEIXEIRA, NUNES 2008) TEIXEIRA, Josele; NUNES, Liliane. Avaliação Escolar: da Teoria à Prática, Rio de Janeiro: Wak Editora: 2008.
- (TYLER 1949) TYLER, R. (1949). Basic Principles of Curriculum and Instruction. Chicago: University of Chicago.
- (TYLER 1981) TYLER, R. Princípios básicos de currículo e ensino. Porto Alegre: Globo, 1981.
- (VALLEJO 1979) VALLEJO, P. (1979). Manual de Avaliação Escolar. Coimbra: Almedina.
- (WIDROW, HOFF 1960) WIDROW, B. , HOFF M. E “Adaptive switching circuits,” in 1960 WESCON Convention Record, New York, 1960.
- (WILLIS 1996) WILLIS, Barry. Evaluation for Distance Educators. 1996. Englewood Cliffs, NJ: Educational Technology Publications. Disponível por em <http://www.uidaho.edu/evo/dist4.html> acesso em 20 de julho de 2009.

- (WOLZ 1997)                      WOLZ, Ursula et al. Computer-Mediated Communication in Collaborative Educational Settings ITICSE Bulletin, New York, v.29, n.3, Trabalho apresentado no ITICSE, 1997.
- (WORTHEN, SANDERS, 1987)                      WORTHEN, B. e SANDERS, J. (1987). Educational evaluation. NY: Longman.

## ANEXO I

## (Questionário de Satisfação do Aluno)

ITEM	ASPECTO	GRAU DE SATISFAÇÃO				
						
SOFTWARE (SAICO)	NAVEGAÇÃO (É Simples?)	<input type="radio"/>				
	APARÊNCIA (É agradável?)	<input type="radio"/>				
	DESEMPENHO (Atende as expectativas?)	<input type="radio"/>				
	SEQUENCIA DE TELAS (É intuitivo?)	<input type="radio"/>				
	USABILIDADE (É confortável?)	<input type="radio"/>				
MÉTODO DE AVALIAÇÃO (se comparado ao método tradicional)	AVALIAÇÃO FORMATIVA (Acredita ser uma forma mais justa de ser avaliado?)	<input type="radio"/>				
	CONFORTO (Sente-se confortável com o método de avaliação?)	<input type="radio"/>				
	ATRATIVIDADE (O método de avaliação demonstra-se interessante a você?)	<input type="radio"/>				
	“FEEDBACK” (Correção rápida - 24 hrs, ajuda o aluno a identificar deficiências no aprendizado?)	<input type="radio"/>				
	QUESTÕES ABERTAS (Automatização do processo é um aspecto facilitador na sua opinião?)	<input type="radio"/>				
EXPERIÊNCIA PEDAGÓGICA	AVALIAÇÃO (O processo permite agregar conhecimento sobre o conteúdo?)	<input type="radio"/>				
	PROCESSO (Permite re-elaborar conceitos a respeito do conteúdo?)	<input type="radio"/>				
	RANKING (Permite reafirmar seu conhecimento sobre o conteúdo?)	<input type="radio"/>				
	EXPERIÊNCIA (Permite ao aluno situar-se em relação ao real nível de absorção do conteúdo?)	<input type="radio"/>				
ANONIMATO	“RANKING” (Controla influências interpessoais entre alunos durante a segunda etapa?)	<input type="radio"/>				
	CORREÇÃO (Controla influências interpessoais entre professor-aluno?)	<input type="radio"/>				
	AUDITORIA (Evita o constrangimento do aluno ao questionar sua nota?)	<input type="radio"/>				
Comentário:						

## ANEXO II

### (Depoimento do Professor)

Devo confessar que aceitar participar deste projeto foi um ato de extrema coragem de minha parte, considerando que as inúmeras etapas previstas durante o processo de coleta e validação empírica exigiriam horas diárias de planejamento e dedicação de minha parte.

No entanto, o tema já havia me chamado a atenção pela sua proposta inovadora, prendendo-me ainda mais após o início dos trabalhos em ambiente real (sala de aula). O método simplifica e facilita de forma extremamente eficaz o processo de avaliação, viabilizando a utilização de questões abertas a grandes grupos de alunos.

Portanto, é com grande liberdade que registro a minha satisfação incondicional ao método avaliativo aqui descrito e testado por mim pessoalmente nesta pesquisa.

Goiânia, 30 de Junho de 2009.

Professor Dr. Ademir Aparecido do Prado (UFG)