

EUGÊNIO ROVARIS NETO

***E-BAYES* – SISTEMA ESPECIALISTA PARA A ANÁLISE DA
EVASÃO DISCENTE DE CURSOS DE GRADUAÇÃO NO ENSINO
SUPERIOR**

**Florianópolis – SC
2002**

**UNIVERSIDADE FEDERAL DE SANTA CATARINA
PROGRAMA DE PÓS-GRADUAÇÃO EM CIÊNCIA DA
COMPUTAÇÃO**

Eugênio Rovaris Neto

***E-BAYES* – SISTEMA ESPECIALISTA PARA A ANÁLISE DA
EVASÃO DISCENTE DE CURSOS DE GRADUAÇÃO NO ENSINO
SUPERIOR**

Dissertação submetida à Universidade Federal de Santa Catarina como parte dos requisitos para a obtenção do grau de Mestre em Ciência da Computação

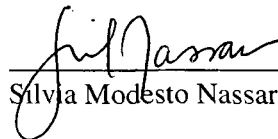
Profª Orientadora: Silvia Modesto Nassar

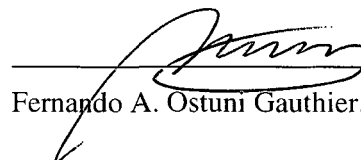
Florianópolis, Fevereiro de 2002

E-bayes – Sistema Especialista para a Análise da Evasão Discente de Cursos de Graduação no Ensino Superior

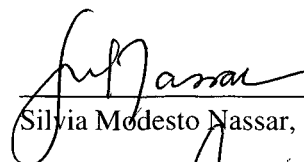
Eugênio Rovaris Neto


Esta Dissertação foi julgada adequada para a obtenção do título de Mestre em Ciência da Computação Área de Concentração (Sistemas de Conhecimento) e aprovada em sua forma final pelo Programa de Pós-Graduação em Ciência da Computação.

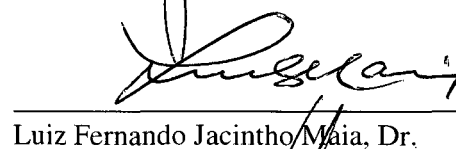

Sílvia Modesto Nassar, Dr^a (orientadora)

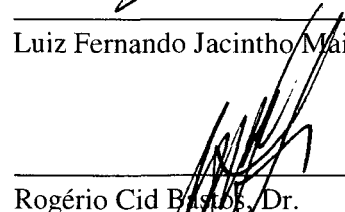

Fernando A. Ostuni Gauthier, Dr. (coordenador)

Banca Examinadora


Sílvia Modesto Nassar, Dr^a (presidente)


João Bosco da Mota Alves, Dr.


Luiz Fernando Jacintho Maia, Dr.


Rogério Cid Bastos, Dr.

A minha família Giovanna, Eugênio (5) e Arthur (2).

A meus inúmeros amigos, de Norte a Sul deste Brasil.

A meus mestres e patrocinadores.

*"As aparências para mente são de 4 tipos:
As coisas são o que parecem ser
Ou não são e nem parecem ser
Ou são e não parecem ser
Ou não são e mesmo assim parecem ser
Identificar corretamente todos esses casos
é a tarefa do homem sábio"*

Epictetus (Século II DC)

Agradecimentos

Primeiramente, agradeço a Deus pela proteção diária dada a minha família, em especial aos meus dois filhos. Agradeço aos professores Bosco e Gustavo pela iniciativa, incentivo e apoios logístico e pessoal. Com a professora Silvia, fico feliz pela sorte de conhecer e trabalhar com alguém como à senhora, minha orientadora.

Agradeço também aos meus alunos, professores, colegas de trabalho e a Universidade Luterana do Brasil, pelo qual cumprimento e agradeço a todos que a representam.

Dedico este trabalho a minha esposa Giovanna Ramos Strassburger e aos meus filhos Eugênio Strassburger Rovaris e Arthur Strassburger Rovaris, que sentiram minha falta nestes dois anos de ausência, mas que serão recompensatórios, lhes prometo.

Finalmente à Universidade Federal de Santa Catarina, o qual agradeço a todos os seus professores e dirigentes: bandeirantes, desbravadores da Amazônia que tiveram a honradez de atravessar esta nação, para dar o que há de mais valioso para um povo carente, isolado e esquecido: *conhecimento*.

Obrigado a todos.

Sumário

Lista de Figuras	I
Lista de Tabelas	III
Lista de Quadros	IV
Resumo	V
<i>Abstract</i>	VI
1.INTRODUÇÃO	01
1.1 Apresentação	02
1.2 Justificativa.....	04
1.2.1 Justificativa Teórica.....	04
1.2.2 Justificativa da Aplicação	05
1.3 Problema a ser Tratado.....	06
1.4 Objetivo Geral	06
1.5 Objetivo Específico	06
1.6 Estrutura do Trabalho	07
1.7 Limitações do Trabalho	08

2. ESTADO DA ARTE	10
2.1 Evolução da Inteligência Artificial	10
2.2 Os primeiros Sistemas Especialistas	13
2.3 Sistemas Especialistas Probabilísticos	15
3. FUNDAMENTAÇÃO TEÓRICA	18
3.1 Sistemas de Apoio à Decisão	18
3.2 <i>Data Mining</i>	23
3.2.1 Metodologia para o <i>Data Mining</i>	25
3.2.2 Motivação para a Utilização das Técnicas de <i>Data Mining</i>	27
3.3 Sistemas Especialistas Probabilísticos	28
3.3.1 Redes Bayesianas	30
3.3.2 Inferência Bayesiana	32
3.3.3 Teorema de <i>Bayes</i>	33
3.3.4 Independência Condicional.....	35
4. DESENVOLVIMENTO DO E-BAYES	37
4.1 Sistema Proposto	37
4.2 Recursos Computacionais	38
4.3 Metodologia.....	41
4.3.1 Metodologia aplicada no processo de <i>data mining</i>	41
4.4 O Conhecimento no <i>E-Bayes</i>	42
4.5 A Comunicação entre o <i>E-Bayes</i> e a Rede Gerada no <i>Netica</i>	44
4.5.1 O Ambiente <i>Netica</i>	46
4.5.2 Componentes no <i>Delphi</i> para o Ambiente <i>Netica</i>	47
5. TESTES, RESULTADOS OBTIDOS E DISCUSSÃO	52
5.1 Testes Realizados (Inferência Bayesiana)	52
5.1.1 Exemplo A - sem bolsa auxílio.....	53
5.1.2 Exemplo B - com bolsa de esportes.....	57
5.1.3 Exemplo C - com bolsa de monitoria	61
5.2 Resultados Obtidos e Discussão.....	65

6. CONCLUSÕES E TRABALHOS FUTUROS	70
6.1 Conclusões.....	70
6.2 Trabalhos Futuros.....	72

7. REFERÊNCIAS BIBLIOGRÁFICAS	74
--	-----------

APÊNDICES	78
A – Questionário Aplicado aos Alunos do Curso de Graduação.....	79
B – Rede Bayesiana Qualitativa <i>Data Mining</i>	84
C – Rede Bayesiana Qualitativa para o Protótipo.....	86
D – Rede Bayesiana Quantitativa para o Protótipo.....	88
E – Código Fonte do <i>E-BAYES</i> (Evasão <i>Bayes</i>).....	91
F - Código Fonte dos Componentes do <i>Netica</i> para o <i>Delphi</i>	99
G - Resumos Submetidos e Aceitos em Congressos	121

LISTA DE FIGURAS

Figura 1.1:	Situação do índice de evasão no curso em estudo.....	03
Figura 1.2:	Situação Acadêmica por sexo.	03
Figura 2.1:	A rede bayesiana do SISPAN.....	16
Figura 3.1:	Independência entre as categorias de sistemas de informação.....	19
Figura 3.2:	Visão geral de um SAD.....	20
Figura 3.3:	Visão geral de um SEP.....	20
Figura 3.4:	Motivação para o desenvolvimento deste trabalho	28
Figura 3.5:	Exemplo de Cálculo da Probabilidade Condicional	32
Figura 3.6:	Atualização bayesiana.....	32
Figura 3.7:	Retângulo amostral.....	33
Figura 3.8:	Independência Condicional	35
Figura 4.1:	Exemplo de informação disponível no SGBD.....	38
Figura 4.2:	Um dos primeiros modelos concebidos pelo <i>Netica</i> , para o <i>E-Bayes</i>	39
Figura 4.3:	Banco de Dados utilizado no protótipo.....	39
Figura 4.4:	Tela de Abertura do <i>E-Bayes</i>	40
Figura 4.5:	Situação atual do índice de evasão na história do curso em estudo	43
Figura 4.6:	Situação acadêmica por sexo	43
Figura 4.7:	Situação Acadêmica versus Transferência Interna.....	43
Figura 4.8:	Situação Acadêmica versus Transferência Externa	43
Figura 4.9:	Situação Acadêmica versus Índice de Evasão.....	43
Figura 4.10:	Situação Acadêmica versus Bolsa de Auxílio.....	43
Figura 4.11:	Apresenta a tela de cadastro do aluno e seu resumo bayesiano	45
Figura 4.12:	Apresenta a tela com a situação da rede bayesiana de cada aluno.	46
Figura 4.13:	Componentes do <i>Netica</i> , criados para o <i>Delphi</i> , acessando <i>Netica.dll</i> ..	49
Figura 4.14:	Componente do <i>Netica</i> para o <i>Delphi</i> : <i>TneticaEnviron</i>	50
Figura 4.15:	Componente do <i>Netica</i> para o <i>Delphi</i> : <i>TneticaNetwork</i>	51
Figura 4.16:	Componente do <i>Netica</i> para o <i>Delphi</i> : <i>TneticaNode</i>	52
Figura 5.1:	Resultado da consulta, do exemplo A, via <i>Netica</i>	53
Figura 5.2:	Resultado da consulta, do exemplo A, via Evasão <i>Bayes</i>	56
Figura 5.3:	Resultado da consulta, do exemplo B, via <i>Netica</i>	57
Figura 5.4:	Resultado da consulta, do exemplo B, via Evasão <i>Bayes</i>	60

Figura 5.5:	Resultado da consulta, do exemplo C, via <i>Netica</i>	61
Figura 5.6:	Resultado da consulta, do exemplo C, via Evasão <i>Bayes</i>	64
Figura 5.7:	Rede bayesiana final utilizada no Evasão <i>Bayes</i>	65
Figura 5.8:	Exemplo A: consulta no registro de um aluno pelo Evasão <i>Bayes</i>	66
Figura 5.9:	Exemplo A: consulta na rede bayesiana pelo E- <i>Bayes</i>	66
Figura 5.10:	Exemplo A: simulação	66
Figura 5.11:	Exemplo B: consulta no registro de um aluno pelo Evasão <i>Bayes</i>	67
Figura 5.12:	Exemplo B: consulta na rede bayesiana pelo E- <i>Bayes</i>	67
Figura 5.13:	Exemplo B: simulação	67
Figura 5.14:	Exemplo C: consulta no registro de um aluno pelo Evasão <i>Bayes</i>	68
Figura 5.15:	Exemplo C: consulta na rede bayesiana pelo E- <i>Bayes</i>	68
Figura 5.16:	Exemplo C: simulação	68
Figura 5.17:	Exemplo D: consulta no registro de um aluno pelo Evasão <i>Bayes</i>	69
Figura 5.18:	Exemplo D: consulta na rede bayesiana pelo E- <i>Bayes</i>	69
Figura 5.19:	Exemplo D: simulação	69
Figura 6.1:	Possível atualização com base no questionário aplicado	72
Figura 6.2:	Possível rede nova, com base no questionário aplicado	72

LISTA DE TABELAS

Tabela 5.1: Tabela das Estatísticas para o Cálculo das Probabilidades Condicionais.	54
Tabela 5.2: Tabela dos Δ para o Cálculo das Probabilidades Condicionais	55
Tabela 5.3: Tabela das $P(H_i)$ para o Cálculo das Probabilidades Condicionais	55
Tabela 5.4: Tabela responsável pela atualização bayesiana.....	56
Tabela 5.5: Tabela das Estatísticas para o Cálculo das Probabilidades Condicionais.	58
Tabela 5.6: Tabela dos Δ para o Cálculo das Probabilidades Condicionais	58
Tabela 5.7: Tabela das $P(H_i)$ para o Cálculo das Probabilidades Condicionais	58
Tabela 5.8: Tabela responsável pela atualização bayesiana.....	59
Tabela 5.9: Tabela das Estatísticas para o Cálculo das Probabilidades Condicionais.	62
Tabela 5.10: Tabela dos Δ para o Cálculo das Probabilidades Condicionais	62
Tabela 5.11: Tabela das $P(H_i)$ para o Cálculo das Probabilidades Condicionais	62
Tabela 5.12: Tabela responsável pela atualização bayesiana.....	63

LISTA DE QUADROS

Quadro I:	Possibilidades da situação acadêmica de um aluno evadido na IES estudada	02
Quadro II:	Características de um sistema de apoio à decisão <i>versus</i> SEP proposto.....	22
Quadro III:	Áreas de atuação do <i>data mining versus</i> objetivos deste trabalho	24
Quadro IV:	Comparação entre o conhecimento especializado	24
Quadro V:	Exemplos de regras representando o conhecimento incerto	28
Quadro VI:	Características de um SAD <i>versus</i> Evasão <i>Bayes versus Netica</i>	37
Quadro VII:	Metodologias aplicadas no processo de <i>data mining</i>	41
Quadro VIII:	Áreas de atuação do <i>data mining versus</i> Evasão <i>Bayes</i>	42
Quadro IX:	Principais propriedades do <i>TNeticaEnviron</i>	49
Quadro X:	Principais métodos do <i>TNeticaEnviron</i>	49
Quadro XI:	Principais propriedades do <i>TNeticaEnviron</i>	50
Quadro XII:	Principais métodos do <i>TNeticaEnviron</i>	50
Quadro XIII:	Principais propriedades do <i>TNeticaNode</i>	51
Quadro XIV:	Principais métodos do <i>TNeticaNode</i>	51

RESUMO

O presente trabalho tem por objetivo aplicar técnicas de *data mining* e conhecimentos de redes *bayesianas* para a construção de um Sistema Especialista Probabilístico - SEP que atue como um Sistema de Apoio à Decisão, capaz de simular e avaliar a situação da evasão dos alunos na graduação de uma Instituição de Ensino Superior – IES.

O trabalho, desenvolvido em seis etapas, constituiu-se na criação de um modelo piloto da rede bayesiana; no processo de *data mining* no Banco de Dados disponível aplicado para analisar o perfil do aluno; a proposta final da rede desenvolvida; a inclusão dos dados relevantes, numa base de dados intermediária, para que a mesma forneça informações da rede, de cada aluno individualmente; a criação de uma interface visual para o SEP e a validação do mesmo.

O *E-Bayes* (Evasão Bayes) é uma ferramenta, capaz de quantificar e qualificar esses dados, transformando-os em informação e gerando conhecimento capaz de apoiar no processo decisório de uma empresa desta modalidade, ou seja, um Sistema Especialista sobre a ótica de um SAD – Sistema de Apoio à Decisão.

Palavras-Chave: *Data mining*, SAD, Inteligência artificial e Evasão escolar.

ABSTRACT

The current work has the aim of applying data mining techniques and bayesian networks knowledge to building a Probabilistic Expert System – PES, that perform like a Support System Decision – SSD, able to simulate and evaluates the situation of the evasion of informatics students from Institute of Higher Education.

The work, developed in six stages, is constituted of creation of pilot model of bayesian network; in the process of data mining on the data base available applied to analysis of the profile from the student; the final proposal of the network developed; the inclusion of the relevant data, on a data base intermediated, that the same supply information from the network, from each individuality student; the building of a visual interface to the PES and the validate do the same.

The E-Bayes (Evasão Bayes) software, is a tool, able of quantitative and qualitative this data, transformation then in information and generating knowledge able of support on the decision process of the firm from this modality, like a Expert System with the SSD – System Support Decision optics.

Key Words: *Data mining, SSD, Artificial intelligence and Evasion scholar*

1. INTRODUÇÃO

A questão da evasão escolar universitária é discutida em muitos países. Quase todos realizam pesquisas sobre o tema e há uma acentuada preocupação com os índices de abandono dos cursos. No Brasil, como no resto do mundo, segundo (DAL MAS DIAS, 1997), a evasão de estudantes pode ser considerada como uma das características da educação brasileira, pois ocorre em todos os níveis de ensino; do primeiro grau aos cursos de pós-graduação; da escolarização regular ao ensino supletivo e não-convencional. Embora seja um fato incontestável, obras e pesquisas como (CECCON, 1997; FRIGOTTO, 1989; GIROUX, 1997; MORAIS, 1996) são mais dirigidas ao primeiro grau, carecendo-se de informações para os demais níveis.

Dentro dessa perspectiva, esta proposta de trabalho visa coletar e avaliar estatisticamente os reais problemas, ou mais prováveis, que possam vir a ocasionar a evasão escolar dos alunos do ensino superior.

Dentro deste enfoque, e também pelas características da situação econômica em que se encontra a Região Oeste do Pará, as dificuldades para se manter em Instituições de Ensino Privadas, são cada vez maiores. Desta forma, a evasão escolar no ensino superior pode estar principalmente relacionada, as dificuldades econômicas particulares dos alunos. Assim, busca-se através desta pesquisa, levantar as causas que realmente estão ocasionando o problema da evasão escolar na Instituição.

Indiferentes dos motivos que levaram o aluno a evadir, também, serão levados em consideração o levantamento de dados quanto ao grau de satisfação em relação ao curso, tanto para os alunos evadidos quanto para os alunos regularmente matriculados.

Como metodologia para o desenvolvimento deste trabalho, inclui-se a entrevista de campo com estes alunos evadidos, bem como a seleção e o processamento do material levantado. Anterior a esta fase está à análise e confecção do questionário.

É importante ressaltar que em muitas situações, o aluno evadido se sentirá muito mais à vontade ao ser entrevistado por um aluno (colega), do que por um (ex-)professor da área.

1.1 Apresentação

Para o desenvolvimento do trabalho, foram analisados todos os relatórios disponibilizados pelo SGBD – Sistema Gerenciador de Banco de Dados, desenvolvido em Oracle™ e pertencente a IES – Instituição de Ensino Superior em estudo. O qual centraliza as informações acadêmicas de todas as unidades do complexo educacional desta Instituição de Ensino Superior. O trabalho ficou limitado apenas a um curso de graduação.

Com base nos relatórios e tabelas disponibilizados (vide página 39), foi possível extrair informações essenciais para a revelação de informações estratégicas que se encontravam escondidas nessa grande massa de dados. A tecnologia que implementa essa busca de novas visões e relacionamentos, denomina-se *Data mining*.

Após o agrupamento e classificação dos dados foram montadas várias propostas de modelos da rede bayesiana, até que se chegou a um primeiro modelo proposto apenas para as informações encontradas no Banco de Dados da IES (Apêndice B). Na segunda etapa do trabalho, foram implementados novos nos relativos às informações extraídas de um questionário, de 64 perguntas (Apêndice A), aplicados aos alunos e ex-alunos do curso de graduação desta IES.

Sendo assim, este trabalho pretende retratar a situação real, ou mais provável, deste curso em estudo. Uma vez que todas as tabelas, que compõem os nós do sistema, contém dados estatísticos puros, reais e conferidos por (ROVARIS, 2000; ROVARIS & NASSAR, 2001), de acordo com as figuras 1.1 e 1.2. Validando assim os cálculos das probabilidades condicionais e conseqüentemente o raciocínio bayesiana das inferências solicitadas.

Para a empresa utilizada como modelo neste estudo, por evasão entenda-se os alunos que estiverem com a situação acadêmica, enquadrada no quadro I:

Quadro I: Possibilidades da situação acadêmica de um aluno evadido na IES estudada.

Situação	Explicação
Trancamento	Aluno que pretende garantir a vaga por um período de 2 anos
Cancelamento Total	Aluno que abre mão da vaga
Transferência Interna	Aluno que mudou de curso na mesma IES
Transferência Externa	Aluno que mudou de IES
Abandono	Aluno que abandonou o curso, sem formalizar a desistência do mesmo.
Formado	Egresso que não retornou ou não está cursando uma pós-graduação

Nao_Ev...	Evadir
33.808	66.192

Figura 1.1: Situação do índice de evasão no Curso em estudo. Fonte: (ROVARIS, 2000)

Situacao	Masculino	Feminino
Ativo	59.890	40.110
Trancamento	64.290	35.710
Cancelame...	58.820	41.180
Transf_Inter...	67.800	32.200
Transf_Ext...	90.910	9.090
Abandono	59.260	40.740
Formado	37.500	62.500

Figura 1.2: Situação Acadêmica por sexo. Fonte: (ROVARIS, 2000)

O desenvolvimento de um sistema especialista requer a participação de um engenheiro de conhecimento, que construirá o sistema em si, e de um especialista do domínio em estudo, que fornecerá as informações essenciais para montar a base de conhecimento.

No sistema aqui proposto, será trabalhado, como Engenheiro de Conhecimento e Especialista do domínio, o mesmo ator. Uma vez que ambos assumem a função de coordenação acadêmica do curso em estudo desde 1996. Facilitando assim a criação e montagem da *rede bayesiana*, outro fator importante é de que os fatos e regras foram extraídos de um Banco de Dados facilitando assim a função do especialista do domínio.

Por outro lado, como o sistema deseja integrar-se com os demais cursos e futuramente com as demais campi, foram ouvidas as sugestões dos demais colegas da área. A respeito do que eles esperam desta primeira proposta do E-BAYES.

Além de um especialista em inteligência artificial (engenheiro do conhecimento), foi preciso contar com o apoio de uma equipe multidisciplinar que servisse de apoio na compreensão e implementação deste projeto, como: um especialista da área educacional; uma especialista na área de banco de dados e sistemas de informação e por último um especialista na área de *software* básico e programação visual.

Os resultados aqui divulgados (estatísticas e probabilidades), bem como para os exemplos de inferência demonstrados foram utilizados dados **hipotéticos**, preservando assim a privacidade da empresa pesquisada. Uma vez que este trabalho objetiva aprovar o modelo e não apresentar os resultados efetivos da pesquisa de campo.

1.2 Justificativa

A seguir serão apresentadas as justificativas: teórica e da aplicação, que levaram a propor este trabalho.

1.2.1 Justificativa Teórica

Com a carência nas Instituições de Ensino Superior por um instrumento avaliador e sintomático nas condições de manutenção dos cursos existentes, ou na oferta de novos cursos, e que este instrumento contribua com o acompanhamento da evasão discente. Surge a necessidade da disponibilização de uma ferramenta inovadora, que apóie a alta diretoria na tomada de decisão e no suporte ao planejamento estratégico, com novas políticas que poderão ser tomadas frente às evidências encontradas.

A proposta da apresentação do protótipo de um sistema de consultas, do tipo SAD – Sistema de Apoio à Decisão, busca possibilitar a simulação do comportamento das informações contidas numa base de dados, oriunda ou não de um SGBD¹, em uma base de conhecimento probabilística², baseado no Teorema de *Bayes*.

Surge, então, a idéia de se trabalhar com Banco de Dados e Inteligência Artificial. Mais precisamente: mineração de dados (*data mining*) e técnicas de tratamento de incertezas (*redes bayesianas*) (KASABOV, 1998, p. 132). Uma rede bayesiana pode, ainda, ser montada a partir de informações extraídas de um especialista, de uma base de dados ou através de entrevistas (questionários).

¹ SGBD: Sistema Gerenciado de Banco de Dados

² SEP: Sistema Especialista Probabilístico

1.2.2 Justificativa da Aplicação

Desenvolver um trabalho inovador que efetivamente auxilie na compreensão dos problemas de evasão. Para tanto, faz-se necessário uma análise quantificadora dos principais motivos que levam o aluno a evadir em seu curso, bem como com o grau de satisfação que ele tem com o mesmo. Pois se entende que um aluno mesmo satisfeito pode ser levado a evadir.

Quanto à criação do SEP – Sistema Especialista Probabilístico (*redes bayesianas*), não haverá maiores problemas uma vez que o estado da arte, nesta área, está consolidado. O problema, esta em como resgatar numa base de dados, os registros resultantes de uma pesquisa de um SEP? Uma vez que ele retorna apenas valores probabilísticos (numéricos e percentuais).

Por exemplo, numa base de dados de uma faculdade, no futuro poderia-se testar a seguinte evidência: Qual curso *latu sensu* deve ser oferecido para o próximo ano? Assim, tendo em mãos uma *rede bayesiana* que demonstre o perfil do aluno, do egresso e do corpo docente graduado, pode-se encontrar um novo nicho de mercado, encontrar várias fatias da base de dados que são na verdade clientes em potencial. Ou seja, é possível obter o seguinte resultado: cliente classe **A** com 23 %, cliente classe **B** com 42 % e cliente classe **C** com 35 %, todos focados para o curso de especialização, por exemplo, Direito do Consumidor.

A questão é quem são essas pessoas classe **B**? Como resgata-las automaticamente, em um SGBD, a partir de um SEP? Como integrar um SEP, que seja dinâmico, a um SGBD, uma vez que para a atualização da base de conhecimento são necessários buscas e cálculos exaustivos numa base de dados?

Como resultado final desta proposta de trabalho, espera-se apresentar um protótipo capaz de acessar a base de dados em estudo e inferi-los sobre a rede bayesiana criada e alimentada com os dados estatísticos desta base de dados. Espera-se também além da análise individual e coletiva dos alunos, o resgate dos grupos de alunos selecionados de acordo com as evidências encontradas.

1.3 Problema a ser Tratado

Trabalhar com a análise da evasão do curso de graduação de uma IES, como projeto piloto, para trabalhos futuros interligar as demais redes bayesianas da unidade ou até mesmo dos demais *campi*.

1.4 Objetivo Geral

Apresentar um modelo de redes bayesianas para o Sistema Especialista Probabilístico, que retrate a situação acadêmica do curso de graduação tirado como amostra para este modelo. Rede esta que identifique e quantifique quais os principais motivos que levam um aluno a evadir-se de seu curso, uma vez que esta informação possibilitará a criação dos parâmetros de prevenção referentes ao problema "evasão escolar".

1.5 Objetivos Específicos

- Identificar motivos da evasão discente;
- Construir a rede bayesiana do Sistema Especialista Probabilístico, utilizando informações de um especialista;
- Aplicar técnicas de *data mining* no processo de alimentação da rede projetada;
- Refinar as probabilidades da rede bayesiana por meio de informações extraídas de uma base de dados;
- Desenvolver um Sistema de Apoio à Decisões que integre os módulos de inferência bayesiana do sistema especialista de evasão discente com o banco de dados disponibilizado;
- Avaliar o protótipo desenvolvido.

1.6 Estrutura do Trabalho

Este trabalho adota o seguinte formato:

No Capítulo um, mostra-se a introdução sobre a forma da apresentação deste trabalho, justificativa, problema a ser tratado, seguido dos objetivos geral e específico, bem como do resumo final da proposta.

No Capítulo dois é apresentado o estado da arte, iniciando-se com a evolução da Inteligência Artificial, abrangendo desde o período aristolético até a época contemporânea. Em seguida apresentam-se os primeiros sistemas especialistas que objetivavam trabalhar com tratamento de incertezas, até a chegada dos trabalhos que adotam, atualmente, como modelo de raciocínio probabilístico o Teorema de Bayes. Os sistemas especialistas Probabilísticos concluem esta etapa do trabalho.

O Capítulo três, relata a fundamentação teórica, onde serão abordados os temas que nortearão o desenvolvimento deste trabalho. Iniciando com os Sistemas de Apoio à Decisão, seguindo pela relevância no uso das técnicas de *data mining*, necessárias para adequação das informações contidas na base de dados, responsável pela construção da rede bayesiana.

O Capítulo quatro, apresenta o processo de desenvolvimento do *E-Bayes* (Evasão Bayes), bem como os recursos computacionais, metodologia, Conhecimento do *E-Bayes*, a sua comunicação com a biblioteca (DLL) e os arquivos (dne) do *Netica*, encerrando com a fase de testes. Enquanto que o Capítulo cinco apresentam algumas simulações, com seus resultados e discussões até o momento desenvolvido.

As conclusões parciais e os trabalhos futuros são apresentados no capítulo seis.

1.7 Limitações do Trabalho

O presente trabalho tem como objetivo aplicar técnicas de *data mining* e conhecimentos de redes *bayesianas* para simular um Sistema Especialista Probabilístico - SEP que avalie a situação da evasão dos alunos do curso de graduação de uma Instituição de Ensino Superior.

Motivado por tal problema, pretende-se mostrar a aplicação prática de técnicas de mineração de dados e uso de redes *bayesianas*. Simulando um sistema de consulta para estimar as chances que um aluno tem de permanecer até a conclusão de seu curso ou de evadir-se.

Neste experimento, a primeira etapa consistiu no processo de *data mining*, feito com base nas informações extraídas no Banco de Dados desta instituição. A segunda etapa foi à aplicação de um questionário com 61 perguntas, para analisar o perfil do aluno e obter as informações inexistentes na base de dados. Por último, o tratamento de todos dos dados, que foram processados e adaptados para o *Netica*, *software* apropriado para desenvolvimento de SEP.

As redes bayesianas facilitam a aquisição de conhecimento e tornam a representação do raciocínio probabilístico possível. O sistema proporciona uma gama excepcional de opções para visualização de estatísticas e nas previsões probabilísticas apoiadas na probabilidade condicional e no Teorema de Bayes, que são as bases do raciocínio com incerteza por aleatoriedade.

A pesquisa surgiu da necessidade de tentar compreender melhor a movimentação discente que evadi ou permanece cursando esta faculdade. Como produto final, será apresentada a rede bayesiana, que retrate o perfil dos alunos evadidos ou dos efetivamente matriculados, bem como a *interface* visual (*E-Bayes*), com acesso à biblioteca *DLL* do *Netica*. Permitindo assim, a análise individual de cada um desses alunos armazenados no Banco de Dados intermediário.

O *E-Bayes* é uma ferramenta, capaz de quantificar e qualificar esses dados, transformando-os em informação geradora de conhecimento capaz de dar suporte no processo decisório de uma empresa desta modalidade, ou seja, um SAD - Sistema de Apoio à Decisão.

2. ESTADO DA ARTE

Para (BARRETO, 1999), o que a Inteligência Artificial tem de mais recente é o seu nome, atribuído por *John McCarthy* em 1956 no *Darhmouth College*. Uma vez que o desejo humano de criar mecanismos que reproduzissem o comportamento humano de (GANASCIA, 1997, p. 97), datam desde 1837 com Babbage e sua máquina analítica, virando o século XX com Torres Y Quevedo, 1911, com sua máquina denominada *automata*, e assim por diante.

2.1 EVOLUÇÃO DA INTELIGÊNCIA ARTIFICIAL

Na busca de uma melhor compreensão sobre os estudos da Evolução da Inteligência Artificial, (BARRETO, 1999) a divide em épocas, (BASTOS & KELLER, 1998) e (COSTA, 1999) fazem o mesmo com a lógica e (WINSTON, 1988) com os programas e publicações mais significativos que influenciaram cada uma dessas fases. Nesta seção, somam-se ainda, as contribuições de (KOVÁCS, 1996) para as redes neurais e (KASABOV, 1998) sobre a lógica *fuzzy*. Estes trabalhos completos podem ser encontrados em (ROVARIS, 1999a, 1999b).

- **A época pré-histórica ou período Aristolético** (COSTA, 1999), inicia com os estudos sobre a lógica silogística ou clássica de Aristóteles (384-322 a.c.). Essencialmente esta lógica preocupa-se em derivar a verdade ou falsidade de um argumento filosófico. Esse tipo de lógica é ainda muito usado, por ser esta, a base de toda a argumentação “que fornece as leis e as regras formais” (BASTOS & KELLER, 1998). Nesta mesma época, outros filósofos e pensadores, como Teofrasta, Platão, Sócrates, Zenão de Eléia, e Protágoras, se destacaram nessa área em estudo. Em (COSTA, 1999), o autor defende este período como “época booleana”, atribuindo a Boole e aos outros autores supracitados, a influência e responsabilidade deles na evolução da lógica, em especial na álgebra booleana.
- **Época Antiga (1875-1943)**. O início desta fase é marcado pelo surgimento da lógica formal por Bertrand Russel (1872-1970) e Gödel. A lógica simbólica de Leibniz (1646-1716) passa a ser estudada por Carrol em 1817. Esta época se encerra com os trabalhos de McCulloch & Pitts

(MCCULLOCH, 1943 apud BARRETO, 1999, p. 8) em 1943, com a primeira proposta de um modelo formal para os neurônios. As limitações foram as grandes distâncias entre as conquistas da psicologia e da neuropsicologia.

- **A Época Romântica (1943-1956)**, tem em seu marco inicial a inserção oficial do termo “Inteligência Artificial” por John McCarthy (MCCARTHY, 1956 apud BARRETO, 1999, p. 8), na famosa reunião de *Darhmouth College*. Nesta fase, buscou-se simular a inteligência humana em situações pré-determinadas. Sua maior limitação estava nas capacidades computacionais.
- **A Época Barroca (1956-1969)**, foi à fase em que marcou a separação da IA Simbólica da Conexionista. Em 1957, Newell e Simon criam o GPS (Solucionador Geral de Problemas). Em seguida, (ROSENBLATT, 1958 apud BARRETO, 1999, p. 9) a partir dos trabalhos de McCulloch, cria uma rede de neurônios dispostos em camadas denominando-os *perceptrons*, derrubados em seguida, por Marvin Minsk e Seymour Papert, que argumentaram quanto às limitações básicas que os *perceptrons* apresentavam (MINSKY, 1969 apud BARRETO, 1999, p. 74).
- **A Época das Trevas (1969-1981)**, também chamada obscurantismo por (SCHILDT, 1989). Apesar do descrédito gerado na área da neurocomputação, os estudos no campo continuaram, ainda que englobadas em outras linhas de pesquisa, como processamento adaptativo de sinais, modelagem biológica e reconhecimento de padrões. Nesta mesma época há uma enxurrada de Sistemas Especialistas lançados no mercado, destacando-se o MYCIN (*Computer-Based Medical Consultations*) em 1976 de Edward Shortliffe, além do DENDRAL em 1969 de Edward Feingbaum, INTELLECT em 1977 de Larry R. Harris, LIFER em 1978 de Gary G. Hendrix, PROSPECTOR em 1978 de A.N. Campbell, R1(XCON) em 1982 da DEC e o CADUCEUS de Harry E. People em 1982.

- **No Renascimento (1981-1987)**, os estudos sobre a memória associativa de (HOPFIELD, 1982) e em 1986 o trabalho de Rummelhart e McClelland do grupo PDP (Processamento Paralelo Distribuído), “fizeram com que a área de redes neurais tivesse um desenvolvimento explosivo com crescimento exponencial de pesquisas e convenções nesta mesma área” (KOVACS, 1996). Era o algoritmo *backpropagation* resolvendo limitações fundamentais para o treinamento de redes neurais complexas. Os Sistemas Especialistas se popularizaram, “houve o renascimento da IA, simbólica e conexionista” (BARRETO, 1999). Enquanto, segundo (WINSTON, 1982), surgem novos autores e publicações como Feigenbaum em 1981, em 1982 Parker e R. H. Winston, em 1983 Elaine Rich e em 1985 com Charwak & McDermot.
- **Na Época Contemporânea (a partir de 1988)**, embora a IA Simbólica e Conexionista tenham evoluído separadamente, ambas vem buscando o alargamento das aplicações em IA, como em pesquisas de campo de petróleo, bases de dados inteligentes, IA Híbrida ou Distribuída (BARRETO, 1999). Como metodologias e conquistas, a lógica difusa ou nebulosa hoje é usada largamente em indústrias para controle inteligente, bem como em produtos como o câmbio automático, em elevadores, na suspensão ativa, no ar condicionado, nas máquinas de lavar ou fotográficas e assim por diante, de acordo com (FUSTER-PARRA, 1996, p. 170) e (KASABOV, 1998). Ainda para (BARRETO, 1999, p. 9) “sistemas especialistas se torna tecnologia dominada”. Segundo (COSTA, 1999, p. 13) os estudos relativos à lógica alcançaram um alto grau de complexidade, tornando-se muito difícil estar completamente a par dos seus mais variados campos de estudo. No entanto citamos como exemplo os avanços da Lógica Paraconsistente, área esta, objeto de estudo da robótica pela inteligência artificial. Por último os sistemas híbridos, onde (SHAW & SIMÕES, 1999, p. 107) citam bons resultados em problemas mal definidos com sistemas usando controladores *neurofuzzy*.

A seguir, serão apresentados os primeiros sistemas especialistas, que de certa forma, cada um em seu tempo inspirou a necessidade do uso de técnicas mais modernas de projeto e implementação, com o raciocínio bayesiano, vistos no final desta próxima seção.

2.2 OS PRIMEIROS SISTEMAS ESPECIALISTAS

Até chegar na proposta deste trabalho, que trata da aplicação do Teorema de *Bayes*, desenvolvido segundo (FELDBAUM, 1973, p. 41) em 1784 por Thomas Bayes, logo após as suas contribuições na estatística, com a Probabilidade Condicional, em 1763 (PEARL, 1988, p. 17). Faz-se necessário uma breve retrospectiva dos primeiros sistemas especialistas que utilizaram a abordagem probabilística, destacando-se sua aplicabilidade nas mais variadas áreas do conhecimento.

Antes da explosão dos sistemas especialistas probabilísticos, baseados no Teorema de *Bayes*, muitos outros sistemas especialistas vinham sendo desenvolvidos, cada qual, seguindo modelos variados de raciocínio e abordagens da IA diferentes ou híbridas.

Os autores (WEISS & KULIKOWSKI, 1988), descrevem o CASNET, utilizado para o diagnóstico do glaucoma, como um sistema que não adotava o Teorema de *Bayes*, mas, uma rede causal-associativa (rede semântica). Um exemplo interessante da década de 70, que de acordo com os avanços técnico-científicos nestes últimos 30 anos, poderia hoje ser convertido para uma rede bayesiana, sem maiores problemas.

Ainda para (WEISS & KULIKOWSKI, 1988, p. 53), o INTERNST-I, antecessor do CADUCEUS, desenvolvido pela Universidade de *Pittsburgh*, foi um sistema especialista para o diagnóstico de medicina interna. Este, por sua vez representava o conhecimento por meio de uma rede causal (tratados como listas), onde através da manipulação destas listas causais, eram obtidas as pontuações (escores) para os diagnósticos mais prováveis. A pontuação nesta rede causal era calculada por meio de ponderações entre constatações (C) e diagnósticos (D), que se dava através das probabilidades inversas $P(D|C)$ ou diretas $P(C|D)$. Esta abordagem simples de probabilidade também é retratada por (HOLTZ, 1991). Distantes ainda dos avanços da probabilidade condicional e das redes bayesianas, mas retratando a época (1982) e os seus avanços, frente ao parágrafo anterior.

De acordo com (DUDA, 1979 apud RICH, 1988, p. 218), o PROSPECTOR, um SEP para exploração geológica, já utilizava o raciocínio bayesiano para calcular a probabilidade de que vários minérios estariam presentes, a partir da evidência coletada. Porém (WEISS &

KULIKOWSKI, 1988, p. 57), afirmam que o PROSPECTOR trabalhava com probabilidade, mas através de fatores ou graus de confiança, contradizendo-se aos autores anteriores. Como estes últimos autores detalham melhor o funcionamento do sistema, e em face da época em que foi desenvolvido, década de 70, é provável que naquele tempo não se trabalhasse com SEP e DAG - Gráfico Acíclico e Orientado (PEARL, 1995).

Os autores (GENARO, 1986), (CUNHA & RIBEIRO, 1987), (WINSTON, 1988), (RICH, 1988), (WEISS & KULIKOWSKI, 1988) apresentam ao longo de suas obras, o estado da arte, através de uma série de aplicações na área da Inteligência Artificial. Em função da época das suas respectivas obras, é dada uma maior evidência aos sistemas especialistas baseados em regras de produção, com ou sem fatores de confiança e também às redes semânticas.

A partir de participações em congressos, das dissertações de mestrados, teses de doutorados, cadernos de resumos, relatórios técnicos, anais e artigos mais recentes, foi possível levantar informações mais apropriadas e pertinentes ao tema deste trabalho, que trata do uso de redes bayesianas na concepção de sistemas especialistas probabilísticos, vistos a seguir.

2.3 SISTEMAS ESPECIALISTAS PROBABILÍSTICOS

Um dos primeiros sistemas especialistas da área médica que utilizou redes bayesianas, realizava o diagnóstico diferencial para tuberculose, bronquite e câncer pulmonar (LAURITZEN & SPIEGELHALTER, 1988).

A seguir, (LADEIRA, 2001), apresenta alguns sistemas e áreas de atuação, que foram concebidos baseados no paradigma probabilístico da inteligência artificial, ilustrando o estado da arte nas mais variadas áreas do conhecimento.

- Na área da educação, destaca-se o modelo de tutorial ANDES para Física, desenvolvido pelo *Learning R&D Center* da Universidade de Pittsburgh. <www.pitt.edu/~vanlehn/andes.html>;
- Para o comércio eletrônico, existe a proposta de trabalho com agentes móveis distribuídos, responsáveis pela supervisão das linhas de usuários em *sites web*, da *Manna Network Technologies* <www.mannanetwork.com>;
- O *SymText*, que trabalha com processamento de linguagem natural, na extração de informação dos textos. Desenvolvido pela Universidade de Utah. <<http://www.students.cs.byu.edu/~sbk/>>;
- O *Auto Fix*, um SEP³ de diagnóstico e reparo de automóveis, desenvolvido pela *knowledge industries* e utilizado na *Automotive Information Systems, Inc.* <www.PartsAmerica.com>;
- O TRACS - *Transport Reliability Assessment and Calculation System*, que busca a Confiabilidade (em nível de projeto e desenvolvimento) de veículos militares, utilizado pelo DERA - Ministério da Defesa Inglês <www.agenda.co.uk/tracs/index.html>;

³ SEP: Sistema Especialista Probabilístico.

O SEDIN foi objeto de estudo na dissertação desenvolvida na Universidade Federal de Santa Catarina por (STEIN, 2000), e que utiliza uma Base de Conhecimento Dinâmica. O sistema SEDIN foi concebido a partir do sistema especialista probabilístico SISPAN, o qual possui uma base de conhecimento estática. Ambos os sistemas oferecem um diagnóstico para avaliar o estado nutricional em crianças com até dois anos.

Um modelo híbrido difuso-probabilístico foi desenvolvido na Universidade Federal de Santa Catarina, por (BRIGNOLI, 2001). Apresentando uma proposta para a construção de Sistemas Especialistas referindo-se a uma modelagem híbrida, a qual utiliza a combinação de duas técnicas de Inteligência Computacional – Conjuntos Difusos e Redes Bayesianas. Esse trabalho propõe uma redistribuição de probabilidades das saídas de uma rede bayesiana, de forma a ajustar ou qualificar tais probabilidades em função da imprecisão em seus eventos, ou seja, eventos difusos.

Ao final desta seção, pôde-se perceber que no âmbito nacional essa área de pesquisa em inteligência artificial está apenas iniciando, uma vez que se encontram poucos trabalhos desenvolvidos, na sua maioria voltados para área médica, enquanto que no exterior as pesquisas já vêm sendo realizadas há vários anos e em várias áreas do conhecimento.

3 FUNDAMENTAÇÃO TEÓRICA

Faz-se necessário à apresentação da fundamentação teórica relativa aos temas que nortearão o desenvolvimento deste trabalho. Este capítulo inicia enfatizando a importância dos Sistemas de Apoio à Decisão – SAD nas organizações, seguindo pela relevância no uso das técnicas de *data mining*, necessárias para adequação das informações contidas numa base de dados. Estas informações tratadas possibilitam a construção de Sistemas Especialistas, neste caso Probabilístico, atuando como uma ferramenta de apoio no processo decisório de uma dada organização.

3.1 Sistemas de Apoio à Decisão

Com advento das redes de computadores e banco de dados integrados, somando-se aí um mercado cada vez mais competente e por isso competitivo, o que se pode oferecer hoje ao corpo estratégico de uma dada organização é uma margem mínima de erro no seu processo decisório.

Desta maneira, pode-se dizer, segundo (LAUDON & LAUDON, 2001, p. 27) que “*o sucesso nas organizações está em fornecer seus bens pelo menor custo e ou fornecer um serviço superior*”. Esta margem de manobra pode ser mais bem aproveitada com o uso de recursos computacionais que apoiem devidamente cada camada hierárquica de uma organização. Esta hierarquia implica em especialização, ou seja, necessidade de diferentes tipos de sistemas que integrados ofereçam o que uma empresa mais almeja, ser competitiva e inovadora, estando sempre um passo a frente dos seus possíveis concorrentes.

Assim sendo, (LAUDON & LAUDON, 2001, p.28-37) dividem os sistemas de informação em seis grandes categorias (figura 3.1) demonstrando a independência entre as categorias de sistemas de informação. Onde, os SPT fornecem informações para todos os demais sistemas, que por sua vez produzem informações para os outros sistemas. Outra característica destes sistemas é de que eles não são firmemente ligados entre si, na maioria das organizações.

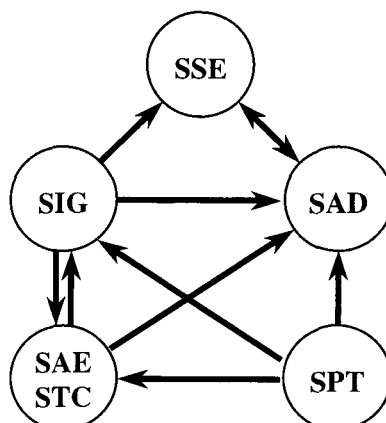


Figura 3.1: Independência entre as categorias de sistemas de informação. (LAUDON & LAUDON, 2001, p.37).

- **SPT - Sistemas de Processamento de Transações:** São sistemas empresariais básicos e que servem o nível operacional da organização, registrando as transações rotineiras diárias necessárias para a condução do negócio;
- **SAE - Sistemas de Automação de Escritórios e STC Sistemas de Trabalho do Conhecimento:** São sistemas que suprem as necessidades de informação no nível de conhecimento de uma determinada empresa. Assim como o SAE é destinado a pessoas que trabalham com dados, o STC, objetiva atender a uma clientela mais qualificada e que crie novos conhecimentos, a partir dos que já estão registrados em suas máquinas;
- **SIG - Sistemas de Informações Gerenciais e SAD - Sistemas de Apoio a Decisão:** Os SIG são sistemas que dão suporte ao nível gerencial de uma dada organização. Esta modalidade de sistema se alimenta basicamente do SPT e objetiva orientar a situação interna de uma empresa, demonstrando o desempenho atual com base nos registros históricos armazenados via SPT. Enquanto que um SAD, figura 3.2, embora seja também alimentado por um SPT, pode vir a ser alimentado por um SIG, mas também, como característica própria, podendo adicionar informações externas à empresa, incluindo-se aí o grau de satisfação de um dado cliente, o desempenho da concorrência ou ainda o perfil de um grupo de clientes, objetivando, i.e., segundo (AMARAL, 2001, p.

8) o *marketing* direto ao invés do *marketing* de massa⁴. Um SEP, portanto, pode se comportar conforme na figura 3.3 como um SAD.

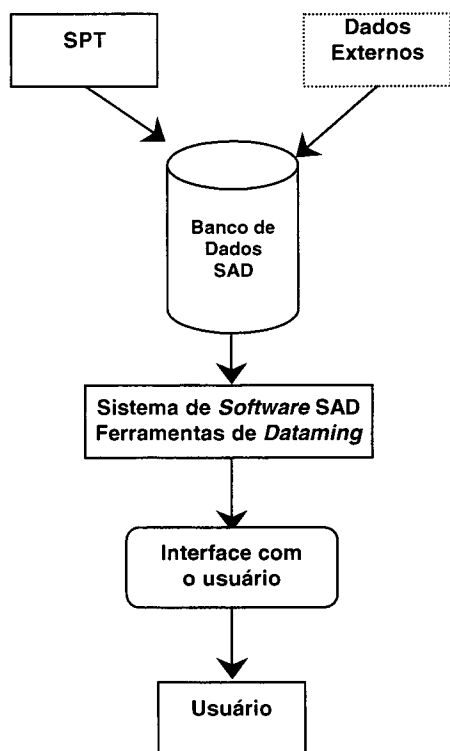


Figura 3.2: Visão geral de um SAD.

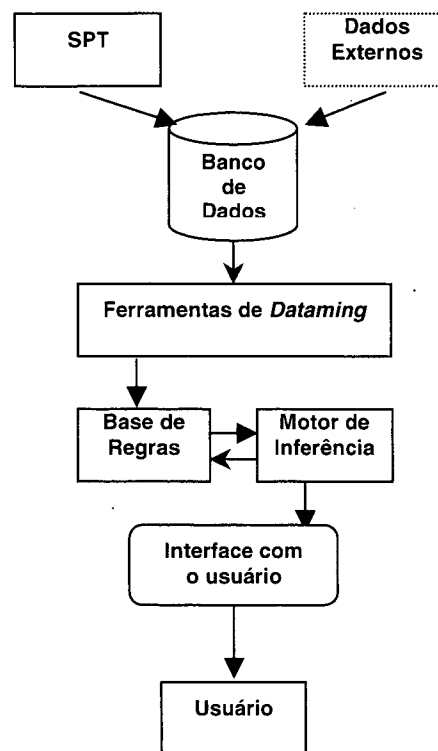


Figura 3.3: Visão geral de um SEP.

Na figura 3.2 é apresentada a visão geral de um SAD, sendo que os seus principais componentes são: O banco de dados, *software* (*data mining*) e a *interface* com o usuário. O banco de dados do SAD pode ser um pequeno banco de dados armazenados num PC ou uma grande *data warehouse* (LAUDON & LAUDON, 2001, p. 320).

Enquanto que na figura 3.3, apresenta-se à visão geral da construção de um SEP, sendo que os seus principais componentes são: A base de regras, o motor de inferência e a *interface* com o usuário (CUNHA & RIBEIRO, 1987). A base de regras pode ser alimentada via processo de *data mining* numa base de dados intermediária do SAD, que agrupe o SPT mais os dados externos. O motor de inferência pode ter raciocínio baseado em casos, regras de produção, *fuzzy* ou ainda probabilístico.

⁴ A vantagem do *marketing* direto sobre o *marketing* de massa, está exatamente em aumentar a taxa de resposta e reduzir os custos das campanhas, gerando de certa forma, maior lucro líquido para uma organização. Ainda em relação ao *marketing* direto pode-se entender a importância do *data mining*, bem como dos Sistemas Especialistas Probabilísticos, neste processo decisório quando se analisa o comportamento em prováveis situações.

- **SIE - Sistemas de Informações Executivas:** Este integra todos os anteriores, somando-se aí o uso de poderosas ferramentas gráficas que possibilitem a geração de resultados oriunda de diversas fontes (internas e externas) transferindo-as automaticamente para outros escritórios ou ambientes de reuniões. Uma característica desta modalidade de sistema esta em fornecer informações que respondam a questões relativos a aquisições, fusões ou simplesmente novos negócios que uma dada organização poderia estar inserida.

Quando (POLLONI, 2000, p. 79) diz que “...o gerente moderno precisa continuamente de informações a fim de obter sucesso em suas atividades...”, ele está sugerindo a necessidade do uso de ferramentas inovadoras que permitam a este executivo o acesso a informações estratégicas com um simples aperto de botão.

O modelo de raciocínio de um SEP adotado pelo *Netica* exemplifica bem esta questão; quando no momento em que se dá um “*click*” nas evidências, em cada um dos nós de interesse, passa-se assim a obter automaticamente e com o grau de precisão da rede bayesiana instalada, uma gama de informações, até então nunca imaginadas que fossem possíveis de se obter. Sobretudo quanto à ótica da inteligência artificial, por meio das técnicas de tratamento de incertezas. Incertezas estas que norteiam o processo decisório diário destas organizações modernas, citadas anteriormente.

Para (LAUDON & LAUDON, 2001), SAD são sistemas que apóiam os gerentes ou executivos no processo decisório e estratégico de uma determinada organização. Constituindo-se de modelos construídos com a finalidade de analisar uma grande massa de dados condensados num formulário ou numa tela, como é o caso da proposta deste trabalho.

De acordo com o quadro II, os Sistemas de Apoio à Decisão, devem apresentar as seguintes características:

Quadro II: Características de um Sistema de Apoio a Decisão *versus* SEP proposto.
(Laudon & Laudon, 2001 p. 31)

Características de um SAD	Sistema Proposto neste trabalho
Baixo número de informações	O processo de <i>data mining</i> , de levantamento estatístico e de inferência probabilística, reduzem significativamente a massa de dados.
Interatividade	Ferramenta gráfica com resposta automática e de fácil comunicação.
Simulação	Permite simulação individual (registro a registro) ou coletiva (todos os registros).
Permitir análise	A cada processo de interação, o sistema permite uma análise estatística ou probabilística de uma determinada situação.
Respostas para questões	A resposta pode estar expressa em termos estatísticos ou probabilísticos de acordo com as evidências marcadas. A decisão final é do usuário.
Profissionais e Gerência administrativa	A organização é quem definirá quem serão os usuários deste sistema

O Quadro II apresenta um comparativo entre as características de um SAD cuja fonte vem de (LAUDON & LAUDON, 2001 p. 31) *versus* a proposta deste trabalho que objetiva apresentar o protótipo de um Sistema Especialista Probabilístico, que tem a sua base de conhecimento alimentado via processo de *data mining* e que pretende ser utilizado como um Sistema de Apoio a Decisão.

Na sua obra, (POLLONI, 2000, p. 79) acrescenta ainda a necessidade do Sistema de Apoio a Decisões apresentar uma *interface* homem-máquina amigável, ou seja, um sistema de fácil manipulação e que apresente de forma simples e objetiva as informações estratégicas. Esta característica é contemplada pelo item “interatividade” supracitada.

Um Sistema de Apoio a Decisões deve interferir no processo decisório de uma empresa que pretenda entender o seu mercado atual, ou ainda, criar novos nichos de mercado. Isto ser torna possível quando identificamos um público alvo, sendo que por público alvo deve-se considerar: o ex-cliente, o cliente atual e o futuro cliente.

Para (LAUDON & LAUDON, 2001, p. 319) ferramentas de *data mining* devem ser fortemente recomendadas para o encontro de padrões escondidos em grandes massas de dados, com seus relacionamentos e regras de inferência a partir deles para prever um comportamento futuro capaz nos guiar num processo decisório. Neste contexto a criação de um sistema especialista por meio de uma ferramenta visual e algorítmica que trabalhe com raciocínio bayesiano pode ser adequada para esta problemática, pois ele permite simular situações preestabelecidas baseadas no conhecimento existente.

Sendo que segundo (IMON, 2001, p. 193) a mineração de dados pode ser usada para provar ou negar uma suposição a partir do processamento de exploração. Este processo deve envolver informações conhecidas, como ocorrem com as probabilidades condicionais.

Um sistema especialista, segundo (LAUDON & LAUDON, 2001, p. 303), “*deve capturar o conhecimento de um especialista*”. As organizações também possuem conhecimento que foram construídos ao longo dos anos, conhecimento este disponível em suas bases de dados e que podem ser capturados e armazenados utilizando-se de um modelo de raciocínio. Neste trabalho será apresentado o raciocínio bayesiano, que pode ser alimentado estatisticamente via processo de *data mining* e inferir probabilisticamente via Teorema de Bayes por meio de atualizações das suas probabilidades condicionais, de acordo com as evidências selecionadas (clicadas).

3.2 Data Mining

“*A história dos sistemas especialistas tem dado pouca importância ao aspecto da aquisição do conhecimento* (RABUSKE, 1995)”. Para a realização deste trabalho, serão adotadas duas técnicas de *data mining*, a seguir descritas.

Por *data mining*, entende-se como uma tecnologia utilizada para revelar informação estratégica escondida em grandes massas de dados. Ela pode ser utilizada em diversas áreas como: análise de riscos, *marketing* direcionado, controle de qualidade, probabilidade, estatística, análise de dados científicos, dentre outras. No quadro III, primeira coluna, (RODRIGUES, 2000), demonstra as áreas de atuação do *data mining*, enquanto que na segunda coluna, são apresentados os objetivos deste trabalho frente estas áreas de atuação.

Quadro III: Áreas de atuação do *data mining* versus objetivos deste trabalho (RODRIGUES, 2000).

Áreas de atuação do <i>data mining</i>	Áreas de aplicação que os resultados deste trabalho buscam alcançar
Análise de riscos	Pontos com riscos de evasão no curso
<i>Marketing</i> direcionado	Amenizar estes pontos de evasão
Controle de qualidade	Trabalhar a qualidade do curso em estudo
Estatística	Agrupar e classificar os pontos de riscos
Probabilidade	Predição e estimação destes pontos de riscos
Análise de dados científicos	Analisar e validar as estimativas encontradas

Para (AMARAL, 2001), hoje se faz necessário o uso de técnicas disponíveis, como a de *data mining*, para se resolver problemas que envolvam o excesso de dados “...desorganizados, inconsistentes, redundantes e pouco proveitosos para o processo de decisão...”. Ainda segundo o autor, os estatísticos e membros da comunidade que estudam os Sistemas de Apoio a Decisão - SAD adotaram o termo *data mining*, para exprimir o conhecimento extraído de uma base de dados a partir de uma ou mais técnicas de mineração de dados.

Enquanto que os membros da comunidade da inteligência artificial optaram pelo termo Descoberta de Conhecimento em Bancos de Dados - KDD, que objetivam a mesma busca por novos conhecimentos, quadro IV. Seja por definição de padrões, tendências, probabilidades ou fatos, porém com o uso do raciocínio artificial, no caso deste trabalho, o raciocínio probabilístico.

Quadro IV: Comparação, entre o conhecimento especializado humano e artificial (GENARO, 1986, p. 8).

Conhecimento Humano	Conhecimento Artificial
Perecível	Permanente
Difícil de transferir	Fácil de transferir
Difícil de documentar	Fácil de documentar
Imprevisível	Consistente
Caro	Razoável
Discriminatório	Imparcial
Individualizado	Social
Criativo	Sem inspiração
Adaptável	Inflexível
Enfoque amplo	Enfoque restrito
Baseado em senso comum	Técnico

O quadro IV apresenta um paralelo entre o conhecimento humano e o artificial, onde o autor demonstra que mesmo o conhecimento artificial apresentando muitas vantagens sobre

o conhecimento humano, ele não poderá substituí-lo em todas as situações devidas algumas limitações inerentes. Porém servindo em muitos casos como uma importante ferramenta de suporte à tomada de decisões.

Percebe-se assim, que os sistemas especialistas, por meio do conhecimento artificial facilitam e dinamizam determinados processos, porém o conhecimento humano ainda faz-se necessário.

Em ambos os casos, *data mining* e KDD, a descoberta de conhecimento, está relacionada ao suporte à decisão, o que implica na busca por padrões ainda não descobertos nos dados e que possam assim, gerar respostas o mais corretas possíveis para novos casos.

A busca por padrões para (IMON, 2001, p. 16) “... são o início de uma seqüência de atividades que levam, em última instância, à oportunidade corporativa ... nova e inovadora....”.

3.2.1 Metodologia para o *data mining*

Muitas são as técnicas disponíveis para a aplicação de *data mining*. Para a realização deste trabalho serão utilizados, de acordo com (RODRIGUES, 2000), as seguintes metodologias:

- ***Market Basket Analysis:***

Consiste em agrupar os elemento em categorias. No caso foram identificados e criados os conjuntos (nós), como Local de Origem, Situação Acadêmica, Sexo, etc.

Metodologia para a realização deste trabalho:

Estatística feita em cima da Base de Dados disponíveis, i.e., descobrir a cidade de origem do aluno; e a aplicação de questionários para a obtenção de informação complementar, i.e., saber se o aluno trabalha ou não;

- **Árvores de decisão:**

São ferramentas poderosas para classificação e prognóstico (representam regras/conhecimento).

Uma vez que existem inúmeros estudos sobre árvores de decisão e algoritmos de regras de indução e inferência descritas em (INMON, 2001), pode-se afirmar que na sua maioria esses são baseados em métodos evolutivos de modelos com base em probabilidade.

Ainda sobre as árvores de decisão, dado o seu problema de eliminação de dados em uma razão exponencial profunda, pode vir a ser substituída pelos grafos direcionados acíclicos. Uma vez que estes representam a dependência casual entre as variáveis e a força desta dependência é representada pela probabilidade condicional, que estão ligadas por cada grupo de nós pais e filhos da rede (PEARL, 1995, 2000).

Metodologia para a realização deste trabalho:

Adotou-se o modelo de grafo, denominado acíclico e orientado (DAG), este, utilizado nos modelos que implementam o uso de redes bayesianas. Uma vez que por meio de seus nós (variáveis de interesse) e arcos direcionados (ligações) tornam possível o raciocínio com incerteza por aleatoriedade.

Com o uso da biblioteca *Netica.dll* do *software Netica 1.12*, desenvolvida pela Universidade de Alberta e disponibilizada pela empresa canadense *Norsys Software Corporation* (NETICA, 2000), caracterizando assim, um Sistema Especialista Probabilístico. Esta biblioteca *DLL*, disponível em *C* foi adaptada ao *Borland Delphi 5.0*, criando-se assim componentes para abrir e tratar extensões *'dne'* do *Netica*, bem como analisar individualmente cada tupla do Banco de Dados ou ainda a toda Base de Dados.

3.2.2 Motivação para a utilização das técnicas de *data mining*

A motivação para se trabalhar com *data mining* se dá pelas seguintes razões:

- **Visualização de Dados:** Qualificar e organizar os dados a serem trabalhados, encontrando novas formas de visualizá-lo (natural e transparente). Esta visualização fica clara em termos de estatísticas e probabilidades demonstradas graficamente, com o uso do *Netica*.
- **Descoberta de novos conhecimentos:** Explicitar relacionamentos ocultos, padrões e correlações. Com a visualização dos nós e suas dependências, torna-se possível à extração, inferência e descoberta de novos conhecimentos.
- **Acurácia dos dados:** Decorre da necessidade de obter dados cada vez mais consistentes para processamento e análise futuras. Nas redes bayesianas, relacionamentos e nós fracos somos dispensáveis e até desnecessários. Todo sistema especialista deve ter o seu desempenho avaliado.

3.3 Sistemas Especialistas Probabilísticos

A Ciência da Computação, através da Inteligência Artificial, oferece a possibilidade de se trabalhar com os SEP - Sistemas Especialistas Probabilísticos (figura 3.4).

“Um sistema especialista é aquele que é projetado e desenvolvido para atender a uma aplicação determinada e limitada do conhecimento humano (CUNHA & RIBEIRO, 1987)”.

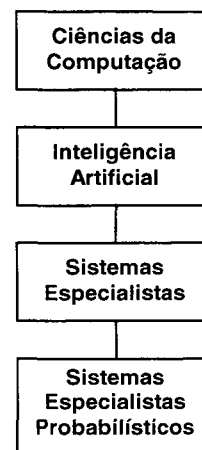


Figura 3.4 - Motivação

Os Sistemas Especialistas que trabalham com o tratamento de incertezas, têm em sua base de conhecimentos fatos e regras que representam o conhecimento do especialista num domínio de aplicação. Onde a cada regra é associado um grau de incerteza de acordo com o modelo de raciocínio adotado.

Para (ROVARIS, 2001b), no paradigma simbólico, o raciocínio é explicitado através das regras de produção do tipo SE-ENTÃO. Todavia, quando a incerteza num domínio de aplicação estiver presente durante a solução de um problema, deve-se adotar o paradigma numérico. Pois este trata a incerteza numericamente, através das inferências e combinações de evidências, que podem ser representadas pela teoria da probabilidade, ou seja, caso $P(A|B) = p$, significa que “SE B ENTÃO A com probabilidade p ”; pelos fatores de confiança através da teoria dos conjuntos; pela teoria dos conjuntos difusos de Zadeh, ou ainda utilizando-se ou da teoria da evidencia de Dempster-Schafer, de acordo com os exemplos apresentados no Quadro V.

Quadro V: Exemplos de regras representando o conhecimento incerto.

Representação da Incerteza	Exemplo de regra
Fator de Confiança (Grau de Certeza - nos moldes do Expert SINTA)	Se Sexo = M Então Evasão = Evadir CNF 70% Evasão = Não Evadir CNF 30%
Lógica Difusa (nos moldes do FUDGE)	Se Sexo = M Então Possibilidade de Evasão = Altíssima (<i>Membership out</i> = 0,7) Se Sexo = M Então Possibilidade de Não Evasão = Baixíssima (<i>Membership out</i> = 0,3)
Teoria da Probabilidade (nos moldes do Netica)	$P(E M) = 0.7$, ou seja, Se Sexo = Masculino Então Probabilidade de Evasão = 70 %
Teoria da Evidencia (nos moldes de Barreto, 1999)	$\theta = \{\text{Evadir, Não Evadir}\}$ Para Sexo = Masculino, surge uma crença de 0,7 de possibilidade de gerar evasão. Onde $m_1(\{\text{Evadir}\}) = 0.7$ e $m_2(\theta) = 0.3$

O conhecimento em um SEP, não precisa ser necessariamente totalmente extraído de um especialista, podendo também, ser fruto de uma busca exaustiva numa base de dados, ou utilizando-se ainda, de outras técnicas de *data mining*.

Quanto à representação deste conhecimento, a Inteligência Artificial, desvia-se voluntariamente da modelagem dos objetos, para se aproximar do comportamento do homem. Ou seja, (GANASCIA, 1997), quer dizer com isto que, para a construção de um sistema especialista que trate, por exemplo, da probabilidade da evasão discente, não é necessária a criação de um sistema que represente o aluno, mas sim, que represente e simule o seu comportamento.

A representação de incertezas em sistemas especialistas demanda uma criteriosa análise do esforço computacional e humano requerido. Se o domínio do sistema for abrangente, a quantidade de conhecimento necessária para representar o problema pode inviabilizar a elaboração do sistema (GONÇALO, 1999).

Os Sistemas Especialistas Probabilísticos, através de informações probabilísticas sobre fatos de um domínio chegam a uma nova conclusão a respeito de um novo fato, conclusão está associada a uma nova probabilidade (BARRETO, 1999). As probabilidades podem ser derivadas de dados estatísticos, de regras gerais ou de uma combinação de outras fontes.

Ainda para (RABUSKE, 1995), a aquisição de conhecimento objetiva a captura de dados, posteriormente transformados em formalismos de implementação.

Assim sendo, a aquisição de conhecimento do Especialista deve permitir chegar-se a uma Base de Conhecimento representada como uma rede bayesiana, contendo informações verossímeis relevantes e com relações causais (BARRETO, 1999). Sobre esta base são feitas inferências que levam a uma conclusão sobre o domínio em estudo.

3.3.1 Redes Bayesianas

De acordo com (PEARL, 1997), as Redes Bayesianas foram desenvolvidas nos anos 70, para modelar a leitura e compreensão em sistemas distribuídos, onde ambos, expectativa semântica e percentual de evidência, devem estar combinados, formando assim, uma interpretação coerente. A habilidade para coordenar inferências bidirecionais, preencheu a lacuna existente na tecnologia dos sistemas especialistas nos início dos anos 80, a partir daí, as Redes Bayesianas emergiram como uma forte ferramenta para a representação de incertezas⁵.

Assim as Redes Bayesianas podem ser definidas como um Grafo Acíclico e Orientado (DAG), onde os nós representam as variáveis de interesse (i.e., situação acadêmica, local de origem e tempo de curso) e as ligações⁶ representam a dependência casual entre as variáveis. A força desta dependência é representada pela probabilidade condicional, que estão ligadas por cada grupo de nós pais e filhos da rede (PEARL, 1995, 2000).

Para (SIMON, 1981; GAAG, 1996; PEARL, 1997; NASSAR, 1998), de uma forma geral, a estrutura de uma rede consiste de duas partes:

- **Parte Qualitativa:** Representa o modelo gráfico (grafo acíclico direcionado) onde as variáveis são: os nós com suas regras; e as relações de dependência condicional entre essas variáveis, que são os arcos direcionados³ (Apêndice C). Para (PEARL, 1997), a parte qualitativa representa assim, a dependência entre os nós.

⁵ *Uncertain* (incerto e necessitando de um adequado tratamento de incerteza)

⁶ e ³ Links (ligações)

- **Parte Quantitativa:** É o conjunto de probabilidades condicionais associadas aos arcos existentes no modelo gráfico e as probabilidades estimadas *à priori* das hipóteses diagnósticas (Apêndice D). Neste ponto, (SIMON, 1981), afirma que são necessárias as utilidades numéricas e estatísticas “iniciais”, para o cálculo das probabilidades “resultantes”. Com posse das estatísticas atribuídas e das probabilidades avaliadas, torna-se possível o cálculo do valor mais provável de cada utilidade e a escolha da melhor alternativa. Concluindo, segundo (PEARL, 1997), a parte quantitativa, através das tabelas de probabilidades condicionais, avalia em termos de probabilidades as dependências, representadas na parte qualitativa.

Para um melhor entendimento dos aspectos quantitativos do raciocínio probabilístico e do Teorema de Bayes, já citado anteriormente, é necessária uma breve introdução aos conceitos de probabilidade condicional. Dentro deste conceito, para dois eventos quaisquer A e B, sendo $P(A) > 0$, definimos a probabilidade condicional de B dado A, de acordo com (PEARL, 1988, p.17) pela (fórmula 3.1):

$$P(A | B) = \frac{P(A \cap B)}{P(B)} \quad (3.1)$$

Uma vez que a topologia da rede bayesiana é definida, basta especificar as probabilidades condicionais para os nodos que possuem dependências diretas e usá-las para computar qualquer outro valor de probabilidade.

O cálculo das probabilidades condicionais é realizado de acordo com o conhecimento prévio das informações contidas nas tabelas, ou seja, as suas estatísticas, conforme o exemplo da figura 3.5. Onde hipoteticamente, dos alunos inativos: 63,707 % destes trancaram a sua matrícula e 10,583 simplesmente abandonaram o curso, isto num universo de 33,808 % ativos contra 66,192 % de evadidos.

Evasao	Ativo	Tranca...	Cancel...	Transf I...	Transf ...	Abando...	Formado
Nao_Evadir	100.00	0.000	0.000	0.000	0.000	0.000	0.000
Evadir	0.000	63.707	6.052	8.868	8.463	10.583	2.327

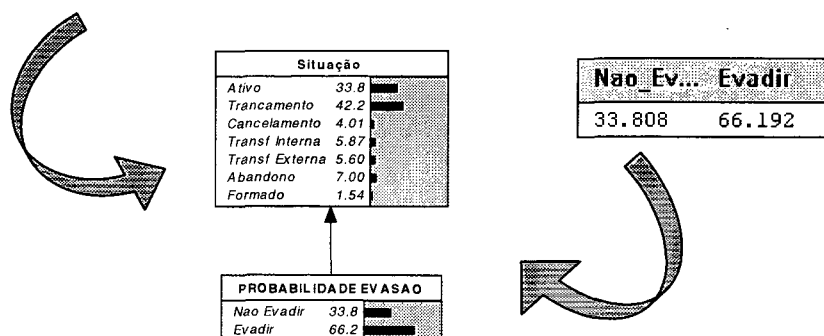


Figura 3.5: Exemplo de Cálculo da Probabilidade Condicional (ROVARIS & NASSAR, 2001, p. 110)

A *rede bayesiana* pode ser considerada uma base de conhecimento abstrata que contém uma grande variedade de composições diferentes, uma vez que, para (RUSSELL & NORVIG, 1995) ela representa a estrutura geral dos processos causais do domínio, em lugar de qualquer detalhe da população de indivíduos.

3.3.2 Inferência Bayesiana

Para Feldbaum em (GENARO, 1986, p. 13): “A *potência de um sistema especialista deriva do conhecimento que possui e não de particular formalismo ou esquema de inferência que empregue*”.

A inferência, mostrada na figura 3.6, é feita através da utilização da teoria da probabilidade, mais precisamente do Teorema de Bayes, atualizadas as probabilidades de acordo com as evidências selecionadas.

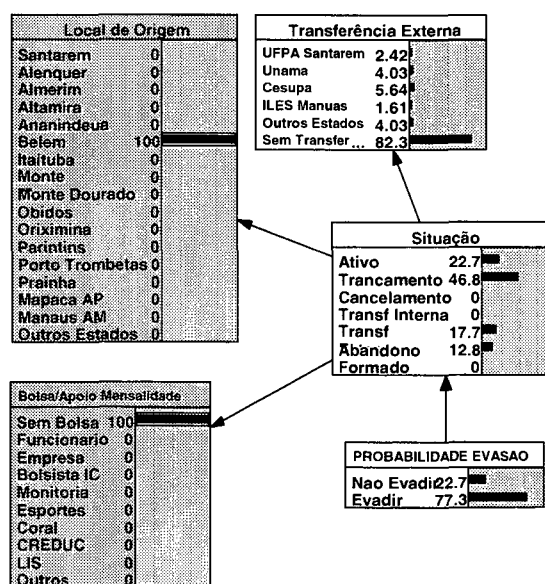


Figura 3.6: Atualização Bayesiana (ROVARIS, 2000)

De acordo com (LADEIRA, 2001), o Teorema de Bayes permite todos os tipos de inferências, seja de *diagnóstico*, *causal* ou *misto*.

As redes bayesianas facilitam a aquisição de conhecimento e tornam a representação do raciocínio probabilístico possível. O sistema proporciona uma gama excepcional de opções para visualização de estatísticas e nas previsões probabilísticas apoiadas na probabilidade condicional e no Teorema de Bayes, que são bases do raciocínio com incerteza por aleatoriedade (ROVARIS & NASSAR, 2001).

Há dois tipos de cálculos realizados por uma rede bayesiana: a *atualização* de crenças e a *revisão* de crenças. A atualização é o cálculo de probabilidades das variáveis aleatórias e a revisão refere-se à obtenção das probabilidades das hipóteses diagnosticas e a identificação da hipótese diagnostica com maior valor de probabilidade (NASSAR, 1998). Isto é possível quando no processo de seleção das evidências, de acordo com o sentido das setas, que efetuam o relacionamento entre os nós.

3.3.3 Teorema de Bayes

A partir dos diagramas da Figura 3.7, tem-se as seguintes situações:

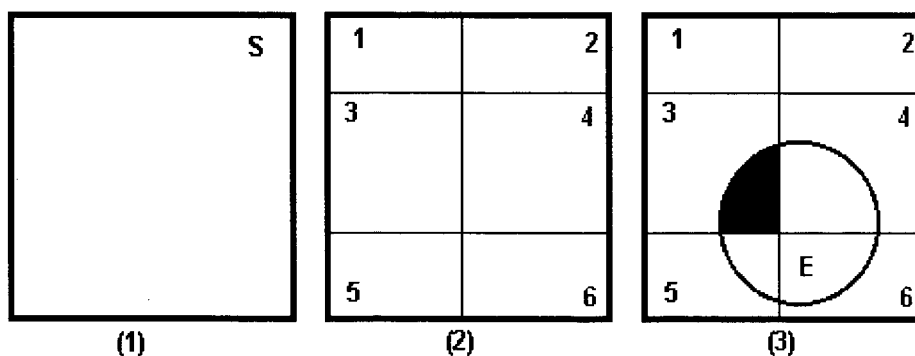


Figura 3.7: Retângulo Amostral (LADEIRA, 2001)

Na situação (1) tem-se um diagrama representando o espaço amostral S . Na situação (2) as hipóteses H_1, H_2, \dots, H_n que formam uma partição do espaço amostral S . Nesse caso as hipóteses H_1, H_2, \dots, H_n são mutuamente exclusivos e sua união é S . Na situação (3) tem-se

um evento qualquer E que possui intersecção com H_1, H_2, \dots, H_n Neste caso E pode ser representado como:

$$\begin{aligned}
 E &= \Sigma \cap E = (H_1 \cup H_2 \cup \dots \cup H_n) \cap E = \\
 &(H_1 \cap E) \cup (H_2 \cap E) \cup \dots \cup (H_n \cap E) \text{ (mutuamente exclusivos, onde tem-se)} \\
 P(E) &= P(H_1 \cap E) + P(H_2 \cap E) + \dots + P(H_n \cap E) \text{ (teorema da multiplicação, onde tem-se)} \\
 P(E) &= P(H_1) \cdot P(E/H_1) + P(H_2) \cdot P(E/H_2) + \dots + P(H_n) \cdot P(E/H_n) = \\
 &\sum_{j=1}^n P(H_j) \cdot P\left(\frac{E}{H_j}\right)
 \end{aligned}$$

Voltando a situação representada pelo diagrama (3) onde tem-se:

- $\{H_1, H_2, \dots, H_n\}$ uma partição do amostral S (domínio), isto é, $H_i \cap H_j = \phi$ para todo $i \neq j$, desde que $P(E) > 0$, ou seja, nenhum destes eventos tenha probabilidade nula (NASSAR, 1998);
- Onde E como um evento qualquer, com $P(H_1), P(H_2), \dots, P(H_n)$ conhecidos e $P(E/H_1), P(E/H_2), \dots, P(E/H_n)$ também conhecidos. Resultando no que segundo Feldbaum (1973, p. 41), Thomas Bayes em 1784, apresentou teorema que levou o seu nome, de acordo com a fórmula 3.2.

$$P(H_i | E) = \frac{P(H_i) \times P(E | H_i)}{\sum_{j=1}^k P(H_j) \times P(E | H_j)} \quad (3.2)$$

Este teorema torna possível a atualização dos graus de crenças, através da teoria estatística bayesiana, onde:

- $P(A)$ pode ser vista como a crença do agente na ocorrência de A (probabilidade subjetiva);
- $P(H_i|E)$: probabilidade condicional, ou *posteriori* da hipótese H_i ser verdadeira, dada à observação da evidência E ;

- $P(E|H_i)$: probabilidade da evidência **E** ser observada, dado que H_i seja verdadeira, ou seja, a *likelihood* (verossimilhança) de H_i causar **E**.
- $P(H)$: probabilidade *a priori* (incondicional) de que H_i seja verdadeira na ausência da evidência **E** ou de uma informação, como ocorre na situação (3) da figura 6.1, onde não tenho uma evidência **E** dividida por todo o retângulo amostral (6 partes).
- k : número de hipóteses possíveis.

A análise bayesiana seria, pois, uma teoria de decisão estatística para cálculo de probabilidade de uma proposição (problema), com base na probabilidade original e nas novas relevâncias e atualizações. Expresso em termos desta pesquisa é o conceito segundo o qual o valor preditivo de uma consulta feita, na rede de evasão discente, depende não apenas da sua sensibilidade e especificidade, mas também da probabilidade prévia (ou seja, da prevalência da evasão na população estudada).

3.3.4. Independência Condicional

Para (PEARL, 1988) muitos eventos reais são condicionalmente independentes de outros eventos, uma vez que não é necessário representar as probabilidades condicionais das suas combinações o que torna o problema mais tratável em computador.

Como é o caso da figura 3.8, onde **A** é irrelevante para a crença sobre **E** se eu conhecer **C**. Onde numa abordagem probabilística (quantitativa) teríamos $P(A,E,C) = P(A|E,C)P(E,C)$ e $P(A|E,C) = P(A|E)$, enquanto que numa abordagem gráfica (qualitativa), ficaríamos com $P(A,E,C)$ e **E separa A de C** (onde A torna-se desconexo de C ao removermos E).

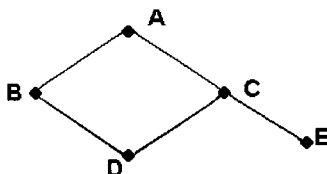


Figura 3.8: Independência Condicional (PEARL, 1988, p. 201)

A regra de cadeia (3.3) fica interessante com (PEARL, 1986), pois ele despreza as combinações relevantes e acelera o cálculo, levando-se em conta a dependência funcional.

$$\begin{aligned}
 P(a,b,c,d,e) &= P(e|a,b,c,d).P(d|a,b,c).P(c|a,b).P(b|a).P(a) & (3.3) \\
 &= P(e|c).P(d|b,c).P(c|a,b).P(b|a) P(a)
 \end{aligned}$$

Diferentemente das cadeias de *Markov* que não possuem direção, as Redes *Bayesianas* possuem um sentido, o que permite deduzir as independências funcionais. Por outro lado nas Redes Bayesianas não poderia haver ciclos, nem retroalimentação, pois os cálculos não funcionam.

4 DESENVOLVIMENTO DO *E-BAYES*

Este capítulo apresenta a proposta do sistema, os recursos computacionais, a metodologia e os procedimentos necessários à concepção do *E-Bayes* (Evasão Bayes).

4.1 Sistema Proposto

A proposta final do sistema ficou definida como sendo a criação de um Sistema Especialista Probabilístico – SEP. O *E-Bayes* é um sistema que utiliza o raciocínio bayesiano para realizar as inferências que tratam as incertezas. Incertezas estas, que levam o usuário ao diagnóstico dos motivos que fazem ou fizeram um acadêmico a evadir do seu curso de ensino superior. Classificando-o ainda como um Sistema de Apoio à Decisões - SAD, uma vez que este SEP é alimentado por dados estatísticos levantados dentro (*data mining*) e fora (entrevista) de uma base de dados. O *E-Bayes* é um SAD por ser uma ferramenta que fornece instrumentos que possam subsidiar interferências no processo decisório de uma empresa do ramo da educação.

O quadro VI confronta a proposta deste trabalho, com o seu produto final, o *E-Bayes*. Comparando-o ainda, com a funcionalidade do *Netica* e com as características de um SAD.

Quadro VI: Características de um SAD *versus* Evasão Bayes *versus* *Netica*.

Características de um SAD	Sistema Proposto neste trabalho	<i>Netica</i>	Evasão Bayes
Baixo número de informações	O processo de <i>data mining</i> , de levantamento estatístico e de inferência probabilística, reduzem significativamente a massa de dados.	As tabelas estatísticas precisam ser preenchidas manualmente	Lê qualquer arquivo com extensão (<i>dne</i>) e aplica as inferências via <i>Netica.dll</i>
Interatividade	Ferramenta gráfica com resposta automática e de fácil comunicação.	Não possui acesso a banco de dados e suas telas (nós e ligações) não possuem qualquer tipo de segurança	Acessa individualmente ou coletivamente todos os alunos e os nós não podem ser apagados ou desconfigurados.

Continua...

Continuação..

Características de um SAD	Sistema Proposto neste trabalho	Netica	Evasão Bayes
Simulação	Permite simulação individual (registro a registro) ou coletiva (todos os registros).	Não permite	Permite
Permitir análise	A cada processo de interação, o sistema permite uma análise estatística ou probabilística de uma determinada situação.	Permite coletivamente	Permite coletivamente ou individualmente
Respostas para questões	A resposta pode estar expressa em termos estatísticos ou probabilísticos de acordo com as evidências marcadas. A decisão final é do usuário.	Permite coletivamente	Permite coletivamente ou individualmente
Profissionais e Gerência administrativa	A organização é quem definirá quem serão os usuários deste sistema	Não possuem limitação de acesso	Pode ser implementada senha por acesso e por camada dos nós que poderão ou não ser exibidos.

4.2 Recursos Computacionais

Para a concepção do *E-Bayes* faz-se necessários os seguintes requisitos:

- O *E-Bayes* foi desenvolvido e roda em ambiente *windows* versão 95 ou superior. O dimensionamento da memória RAM dependerá do tamanho da rede e de sua respectiva base de dados. Para este caso, o sistema funcionou bem em máquinas Pentium com 64 MB;
- Acesso aos relatórios, figura 4.1, e ao SGBD objeto de estudo, no caso o cadastro de alunos;

```

-----
Aluno.: 991160018-6 EUGENIO ROVARIS NETO                               Sexo: M
Sit...: A Curso...: 20416 - INFORMATICA
Nasc...: 25/02/1970 Ender.: CEL. JOAQUIM BRAGA, 181/302
Bairro: CENTRO CEP: 68005-270 Cid: SANTAREM UF: PA
Fone...: 0915224957 Rg: 1936393 CPF: 67774253904 Ingr.: 18/02/1999
-----

```

Figura 4.1: Exemplo de informação disponível no SGBD

- Para a implementação da interface no *E-Bayes*, figura 4.4, optou-se pela linguagem de Programação *Borland Delphi 5.0*. O código fonte do *E-Bayes* (*Evasão Bayes*) está disponível no apêndice E;

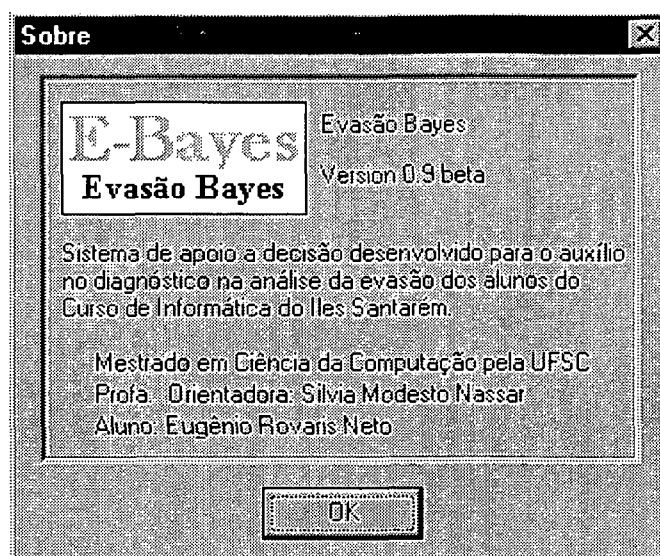


Figura 4.4: Tela de Abertura do *E-Bayes*

- À parte reservada aos cálculos e atualizações da rede bayesiana se deu com o acesso a biblioteca *DLL* do *Netica*. Para tanto foram criados componentes *Delphi*, responsáveis pela abertura da rede, compilação e manipulação (cálculos de inferência e atualizações) entre os nós de uma rede bayesiana gerada via *Netica*. O código fonte dos componentes e programas de acesso a *DLL* estão disponíveis no apêndice E;

4.3 Metodologia

A seguir serão apresentados os métodos e procedimentos necessários para a implementação do *E-Bayes* (Evasão Bayes).

4.3.1 Metodologias aplicadas no processo de *data mining*

Como visto anteriormente no Capítulo 3, muitas são as técnicas disponíveis para a aplicação de *data mining*. Para a execução deste trabalho foram utilizados, de acordo com (RODRIGUES, 2000), o quadro VII apresenta as seguintes metodologias:

Quadro VII: Metodologias aplicadas no processo de *data mining*

Metodologias	<i>Market Basket Analysis</i>	Árvores de decisão
Definição	Consiste em agrupar os elementos em categorias.	São ferramentas poderosas para classificação e prognóstico (representam regras/conhecimento).
Atividade	No caso foram identificados e criados os conjuntos (nós), como Local de Origem, Situação Acadêmica, Sexo, etc.	Procurou-se um software que atendesse a estas necessidades.
Metodologia para a realização deste trabalho	Estatística feita em cima da Base de Dados disponíveis, i.e., descobrir a cidade de origem do aluno; e a aplicação de questionários para a obtenção de informação complementar, i.e., saber se o aluno trabalha ou não;	Uso de Redes Bayesianas, através das probabilidades condicionais aplicadas sobre o conhecimento obtido, com o uso da biblioteca <i>Netica.dll</i> do software <i>Netica 1.12</i> , desenvolvida pela Universidade de Alberta e disponibilizada pela empresa canadense <i>Norsys Software Corporation (NETICA, 2000)</i> , caracterizando assim, um Sistema Especialista Probabilístico.
Refinamentos	Os SGBD geraram arquivos textos que foram abertos, lidos e calculados as suas estatísticas no formato as tabelas dos futuros nós, respeitando as suas ligações	Esta biblioteca <i>DLL</i> , disponível em C e C++ foi adaptada ao <i>Borland Delphi 5.0</i> , criando-se assim um componente para abrir e tratar extensões ' <i>dne</i> ' do <i>Netica</i> , bem como analisar individualmente cada tupla do Banco de Dados ou ainda a toda Base de Dados.
Produto Final	Tabelas de Estatística	rede bayesiana <EvasaoAlunos.dne> Que será lida pelo Evasão Bayes

O quadro VIII apresenta um comparativo entre as áreas de atuação do *data mining* (RODRIGUES, 2000) *versus* a proposta deste trabalho que se propõe a contribuir com todas as estas áreas, através do *software* Evasão Bayes.

Quadro VIII: Áreas de atuação do *data mining* *versus* Evasão Bayes

Áreas de atuação do <i>data mining</i>	Áreas de aplicação que os resultados deste trabalho buscam alcançar	Evasão Bayes
Análise de riscos	Pontos com riscos de evasão no curso	Apresenta claramente
<i>Marketing</i> direcionado	Amenizar estes pontos de evasão	Apresenta intrinsecamente
Controle de qualidade	Trabalhar a qualidade do curso em estudo	Apresenta claramente
Estatística	Agrupar e classificar os pontos de riscos	Apresenta claramente
Probabilidade	Predição e estimação destes pontos de riscos	Apresenta claramente
Análise de dados científicos	Analisar e validar as estimativas encontradas	Fica a cargo do usuário

4.4 O Conhecimento no Evasão Bayes

A seguir serão apresentados três exemplos que ilustrarão o funcionamento e o desempenho de uma rede bayesiana por meio de suas inferências, demonstrando a entrada, processamento (raciocínio) e a saída de uma inferência.

O conhecimento, como foi visto no capítulo 3 se dá pelo cálculo das probabilidades condicionais entre seus nós dependentes. Para tanto, assim como a base de regras expressa o conhecimento em um Sistema Especialista tradicional, nos SEP, ele é expresso por meio de tabelas (pertencentes a cada nó) e suas respectivas estatísticas, estas por sua vez, podem ser levantadas via processo de *data mining*.

As figuras 4.4 a 4.10, demonstram o conhecimento contido em cada tabela da rede criada. Esse conhecimento, no caso estatístico, é que gerara a probabilidade condicional de cada nó, e que posteriormente sofrerá as atualizações bayesianas de acordo com as evidências adotadas.

Nao_Ev... Evadir	
33.808	66.192

Figura 4.5: Situação atual do índice de evasão na história do curso em estudo

Situacao	Masculino Feminino	
	Ativo	59.890
Trancamento	64.290	35.710
Cancelame...	58.820	41.180
Transf_Inter...	67.800	32.200
Transf_Ext...	90.910	9.090
Abandono	59.260	40.740
Formado	37.500	62.500

Figura 4.6: Situação acadêmica por sexo

Situacao	Direito	Eng_Ag...	Letras	Pedago...	Sem_Tr...
Ativo	0.000	0.000	0.000	0.000	100.00
Trancamento	0.000	0.000	0.000	0.000	100.00
Cancelame...	0.000	0.000	0.000	0.000	100.00
Transf_Inter...	88.135	5.085	1.695	5.085	0.000
Transf_Ext...	0.000	0.000	0.000	0.000	100.00
Abandono	0.000	0.000	0.000	0.000	100.00
Formado	0.000	0.000	0.000	0.000	100.00

Figura 4.7: Situação Acadêmica versus Transferência Interna

Situacao	UFPA_S...	Unama	Cesupa	ILES M...	Outros_...	Sem_Tr...
Ativo	0.000	0.000	0.000	0.000	0.000	100.00
Trancamento	0.000	0.000	0.000	0.000	0.000	100.00
Cancelame...	0.000	0.000	0.000	0.000	0.000	100.00
Transf_Inter...	0.000	0.000	0.000	0.000	0.000	100.00
Transf_Ext...	13.630	22.730	31.820	9.090	22.730	0.000
Abandono	0.000	0.000	0.000	0.000	0.000	100.00
Formado	0.000	0.000	0.000	0.000	0.000	100.00

Figura 4.8: Situação Acadêmica versus Transferência Externa

Evasao	Ativo	Tranca...	Cancel...	Transf_I...	Transf_...	Abando...	Formado
Nao_Evadir	100.00	0.000	0.000	0.000	0.000	0.000	0.000
Evadir	0.000	63.707	6.052	8.868	8.463	10.583	2.327

Figura 4.9: Situação Acadêmica versus Índice de Evasão

Situacao	Sem_B..	Funcion.	Empresa	Bolsista.	Monitoria	Esportes	Coral	CREDOC	LIS	Outros
Ativo	72.680	1.546	1.546	1.033	3.608	10.309	1.546	6.186	1.546	0.000
Tranca...	88.000	0.000	0.000	0.000	0.000	12.000	0.000	0.000	0.000	0.000
Cancela...	100.00	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
Transf_I...	91.000	0.000	0.000	0.000	0.000	8.000	1.000	0.000	0.000	0.000
Transf_...	94.000	0.000	0.000	0.000	0.000	5.000	1.000	0.000	0.000	0.000
Abandono	100.00	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
Formado	62.500	12.500	0.000	0.000	25.000	0.000	0.000	0.000	0.000	0.000

Figura 4.10: Situação Acadêmica versus Bolsa de Auxílio

4.5 A Comunicação entre o E-Bayes e a rede gerada no Netica

Neste tópico, será apresentada a comunicação entre o E-Bayes e a rede bayesiana criada no Netica.

O Software Evasão Bayes (E-Bayes) quando executado, se alimenta de duas fontes:

- A base de dados <Alunos.DB>, onde se encontram os dados cadastrais de cada aluno, individualmente;
- A rede bayesiana <EvasaoAlunos.dne>, onde se encontram todos os nós com suas tabelas preenchidas pela estatística alimentadas via processo de data mining; suas ligações e suas probabilidades condicionais calculadas;

A figura 4.11 apresenta a guia <Dados do Aluno>, onde são apresentadas as informações referentes aos dados cadastrais de cada aluno. À direita da tela, são apresentados dois nós que resumem a situação atual do aluno em tela: A probabilidade deste evadir e de sua futura situação acadêmica (Ativo, Trancado, Cancelado, Transferência Interna, Transferência Externa, Formado ou ainda Abandono).

Figura 4.11: Apresenta a tela de cadastro do aluno e seu resumo bayesiano

A figura 4.12 apresenta a guia <rede bayesiana>, onde são apresentadas as evidências buscadas nos dados cadastrais de cada aluno. Nesta tela, os nós: LocalOrigem, Trabalho, Sexo, BolsaApoio, Faixa-Etária e Situação aparecem automaticamente setados (marcados), dado a busca realizada na base de dados, os demais nós ficam em aberto assumindo a atualização da rede bayesiana, baseada nas evidências já constatadas.

Por outro lado, nesta mesma tela, é possível alterar a todas as evidências já setadas (marcadas). Possibilitando assim a simulação da situação acadêmica e análise das novas respostas geradas pela rede bayesiana. A decisão final ficará sempre a cargo do usuário, que poderá se basear nas respostas emitidas pelo Evasão Bayes, um Sistema Especialista Probabilístico, atuando como um Sistema de Apoio à Decisões.

LocalOrigem	Transferencia Externa	Transferencia Interna	Apoio
Santarém: 100.00	UFPA-Santarém: 0.00	Direito: 0.00	Sem Bolsa: 0.00
Alenquer: 0.00	Unama: 0.00	Eng Agrícola: 0.00	Funcionario: 0.00
Almeirim: 0.00	Cesupa: 0.00	Letras: 0.00	Empresa: 0.00
Altamira: 0.00	ILES-Manaus: 0.00	Pedagogia: 0.00	Bolsa IC: 0.00
Ananindeua: 0.00	Outros Estados: 0.00	Sem Transferir: 100.00	Moradia: 100.00
Balim: 0.00	Sem Transferir: 100.00		Esportes: 0.00
Belém: 0.00			Conv: 0.00
Maritápolis: 0.00			CREBUC: 0.00
Maritápolis: 0.00			UIS: 0.00
Obidos: 0.00			Outros: 0.00
Oitombra: 0.00			
Parintins: 0.00			
Póço Trombete: 0.00			
Prainha: 0.00			
Mopoca AP: 0.00			
Maracá AM: 0.00			
Outros Estados: 0.00			

Situação	Evasão
Alivo: 53.64	Não Evadir: 63.64
Trancamento: 0.00	Evadir: 35.36
Cancelamento: 0.00	
Transf Interna: 0.00	
Transf Externa: 0.00	
Abandono: 0.00	
Formado: 35.36	

Trabalho	Sexo
Trabalha: 0.00	Masculino: 0.00
Não Trabalha: 100.00	Feminino: 100.00

Faixa Etária
Entre 18 e 19: 0.00
Entre 20 e 21: 100.00
Entre 22 e 23: 0.00
Entre 24 e 25: 0.00
Entre 26 e 30: 0.00
Entre 31 e 39: 0.00
Acima de 40: 0.00

Aluno: Geisicleia Coutinho Lima

E-Bayes
Evasão Bayes

Figura 4.12: Apresenta a tela com a situação da rede bayesiana de cada aluno.

4.5.1 O Ambiente *Netica*

O uso de redes *bayesianas*, através das probabilidades condicionais aplicadas sobre o conhecimento obtido, utilizado no E-Bayes, se dá com o uso da biblioteca *Netica.dll* do software *Netica 1.12*.

O *Netica* é um software desenvolvido pela Universidade de Alberta e disponibilizado pela empresa canadense *Norsys Software Corporation (NETICA, 2000)*. É uma ferramenta característica para o desenvolvimento de Sistemas Especialistas Probabilísticos, uma vez que ela implementa os algoritmos de inferência bayesiana.

O *Netica* apresenta como vantagens:

- Interface gráfica e de fácil compreensão;
- Ferramenta de fácil manipulação, para a criação das redes e no preenchimento dos seus nós da rede de crença e dos diagramas de influência;
- A sua rede pode ser incorporada dentro de outros documentos;
- Trabalha com a solução de diagramas de influência;
- Capacitado a aprender relações probabilísticas através de dados;
- Mantém diagramas complexos ordenados;
- Permite comentários;
- Variáveis que percam o interesse podem ser removidas sem alterações bruscas na estrutura e organização da rede.

O *Netica* apresenta como desvantagem:

- O ambiente gráfico não é seguro no que se refere às facilidades que o usuário tem em movimentar ou remover as ligações e o nós, em tempo de execução. Carecendo assim, da construção de uma *interface* amigável, mas que acesse a DLL onde estão contidas todos os procedimentos e funções responsáveis pela parte de manipulação e atualização dos cálculos probabilísticos.

4.5.2 Componentes no *Delphi* para o Ambiente *Netica*

Para construção dos componentes para o *Borland Delphi 5.0*, utilizou-se à biblioteca de API *Netica* em sua versão DLL, para *Windows*. A mesma foi baixada (*download*) diretamente do site da *Norsys Corporation (NETICA, 2000)*, e pôde ser utilizada com algumas restrições.

A API *Netica* é uma biblioteca de funções em C que trabalham com redes bayesianas e diagramas de influência. Ela contém funções para construir, aprender, modificar, transformar, salvar e carregar redes, sendo esta uma poderosa máquina de inferência. Podendo ainda ser utilizada em programas escritos em qualquer linguagem que suporte as chamadas das funções construídas em C.

As funções contidas na biblioteca API *Netica* estão muito bem documentadas no arquivo chamado “*Netica API Programmer's Library Reference Manual*”, que pode ser baixado do *site* da *Norsys* em (*NETICA, 2000*).

O primeiro passo para usar a biblioteca DLL no *Borland Delphi 5.0* é conhecer com detalhes o protótipo das funções contidas nela — O manual de referência da API *Netica* é a principal fonte sobre o assunto. De posse desta informação, é necessário declarar em uma *Unit* (módulo da linguagem *pascal*) os protótipos das funções, os tipos de dados e constantes.

Como ponto de partida analisou-se a *Unit* escrita por *Emile van Gerwen (GERWEN, 2001)*, que não está completa, pois não apresenta o protótipo de todas as funções e nem todas as constantes da biblioteca, mas economiza bastante tempo de codificação e ajuda a entender a interface de chamada as funções da DLL.

Na versão para C da API *Netica*, a biblioteca `<Netica.lib>` é acompanhada de um módulo de código fonte chamado `<NeticaEx.c>` que contém funções de alto nível, que expandem as funções contidas na biblioteca. Acompanhando este raciocínio criou-se também um módulo em *pascal* chamado `<NeticaEx.pas>`, no qual foram rescritas em *pascal* apenas as funções contidas no arquivo `<NeticaEx.c>` que eram necessárias ao projeto.

Finalmente decidiu-se pela criação de componentes *delphi* (figura 4.13) que representassem a lógica de funcionamento da biblioteca API *Netica*, e facilitassem a criação do projeto.

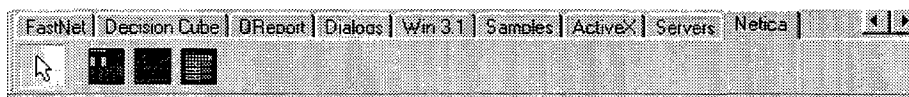


Figura 4.13: Componentes do *Netica*, criados para o *Delphi*, acessando o *Netica.dll*.

É importante frisar que os componentes criados não implementam toda a funcionalidade oferecida pela biblioteca, e sim, apenas as funcionalidades requeridas pelo projeto que visa a implementação do *E-Bayes*. Os componentes criados serão apresentados em seguida, individualmente por página, para uma melhor compreensão. Enquanto que o código fonte dos componentes e programas de acesso a DLL estão disponíveis no apêndice E;

- ***TNeticaEnviron***

O componente não visual *TNeticaEnviron* (figura 4.14) representa o ambiente do *Netica*. Quando este objeto é criado no *Delphi*, cria um novo ambiente *Netica* através da função *NewNeticaEnviron_bn* da API *Netica*. Este componente possui uma lista interna que permite que se conecte a ele vários componentes do tipo *TNeticaNetwork*, passando a idéia que um ambiente *Netica* pode gerenciar várias redes.



Figura 4.14: Componente do *Netica* para o *Delphi*: *TNeticaEnviron*

Nos quadros IX e X, são descritos as principais propriedades e métodos deste componente, utilizados para a implementação do protótipo proposto neste trabalho.

Quadro IX: Principais propriedades do *TNeticaEnviron*

Propriedades	Descrição
<i>License</i>	Valor string que representa uma licença válida fornecida pela <i>Norsys Software Corporation</i> . Se a propriedade <i>license</i> for preenchida seu valor é repassado para a função da API <i>NewNeticaEnviron_bn</i> , desbloqueando as restrições que são imposta sem o uso da licença.
<i>NeticaVersion</i>	Propriedade <i>Read-Only</i> que apresenta o número da versão da biblioteca API <i>Netica</i> .
<i>NetworkCount</i>	Propriedade <i>Read-Only</i> que apresenta o número de objetos <i>TneticaNetwork</i> ligados ao componente.
<i>Enviroment</i>	Propriedade em tempo de execução que fornece um ponteiro para o ambiente <i>Netica</i> . Necessário para a chamada da maioria das funções da API.
<i>Networks</i>	Propriedade em tempo de execução que fornece um vetor dos objetos <i>TneticaNetwork</i> ligados ao componente.

Quadro X: Principais métodos do *TNeticaEnviron*

Métodos	Descrição
<i>ReportError</i>	Método público que encapsula a chamada à função da API <i>ReportError_bn</i> , para geração de um erro pelo programador no ambiente.
<i>AddNetwork</i>	Método Protegido que adiciona um objeto <i>TNeticaNetwork</i> à lista de redes.
<i>DeleteNetwork</i>	Método Protegido que remove um objeto <i>TNeticaNetwork</i> à lista de redes.
<i>NewNeticaEnviron</i>	Método protegido que encapsula a chamada à função da API <i>NewNeticaEnviron_bn</i> , usado na criação do objeto para a criação do ambiente <i>Netica</i> .
<i>CloseNetica</i>	Método protegido que encapsula a chamada à função da API <i>CloseNetica_bn</i> , usado na destruição do objeto para a fechar o ambiente <i>Netica</i> .

- **TNeticaNetwork**

O componente não visual *TNeticaNetwork* (figura 4.15), que representa uma rede bayesiana do *Netica*. Este objeto deve ser ligado ao componente **TNeticaEnviron** através de sua propriedade **NeticaEnviron**, esta ação dispara internamente o método **AddNetwork** do componente **TNeticaEnviron**.



Figura 4.15: Componente do *Netica* para o *Delphi*: *TNeticaNetwork*

O componente *TNeticaNetwork* foi construído apenas para carregar uma rede bayesiana diretamente de um arquivo criado previamente no ambiente visual do *Netica*. O nome do arquivo com o caminho inclusive deve ser informado através da propriedade *FileName*.

Este componente possui uma lista interna que permite que se conecte a ele vários componentes do tipo *TNeticaNode*, que é a representação visual dos nós de uma rede.

Nos quadros XI e XII, são descritos as principais propriedades e métodos deste componente, utilizados para a implementação do protótipo proposto neste trabalho.

Quadro XI: Principais Propriedades do *TNeticaEnviron*

Propriedades	Descrição
<i>Active</i>	Propriedade que ativa a rede, e conseqüentemente abre o arquivo definido na propriedade <i>FileName</i> .
<i>FileName</i>	Propriedade que deve ser preenchida com o nome do arquivo criado previamente no ambiente visual <i>Netica</i> .
<i>NeticaEnviron</i>	Propriedade que define o componente <i>TNeticaEnviron</i> ao qual o componente está ligado.
<i>NetName</i>	Propriedade <i>Read-Only</i> que apresenta o nome da rede.
<i>NodeCount</i>	Propriedade <i>Read-Only</i> que apresenta o número total de nós na rede.
<i>Nodes</i>	Propriedade em tempo de execução que fornece um vetor dos objetos <i>TNeticaNode</i> que estão ligados aos componentes.

Quadro XI: Principais Métodos do *TNeticaEnviron*

Métodos	Descrição
<i>Open</i>	Método público que encapsula a chamada à função da API ReadNet_bn , que carrega a rede do arquivo especificado.
<i>Update</i>	Método público que atualiza todos os nós da rede na tela.
<i>RetractAllFindings</i>	Método público que encapsula a chamada à função da API RetractAllFindings_bn .
<i>Compile</i>	Método privado que encapsula a chamada à função da API CompileNet_bn , que compila a rede bayesiana.
<i>NodeByName</i>	Método público que retorna um ponteiro para um determinado nó da rede, baseado no nome do nó.

- ***TNeticaNode***

O *TNeticaNode* (figura 4.16) é único componente *visual*, e representa os nós da rede bayesiana. Após ser ligado a um componente *TNeticaNetwork* através da propriedade *NeticaNetwork*, e ser definido o nome do nó que representa, através da propriedade *NodeName*, passa a representar visualmente o nó escolhido.



Figura 4.16: Componente do *Netica* para o *Delphi*: *TNeticaNode*

Nos quadros XIII e XIV, são descritos as principais propriedades e métodos deste componente, utilizados para a implementação do protótipo proposto neste trabalho.

Quadro XIII: Principais Propriedades do *TNeticaNode*.

Propriedades	Descrição
<i>NeticaNetwork</i>	Propriedade que define o objeto <i>TNeticaNetwork</i> ao qual o componente está ligado
<i>NodeName</i>	Propriedade que define o nome do nó que o componente representará visualmente.
<i>StateCount</i>	Propriedade <i>Read-Only</i> que apresenta o número total de estados no nó.

Quadro XIV: Principais Métodos do *TNeticaNode*.

Métodos	Descrição
<i>StateFinding</i>	Método público que encapsula a chamada à função da API <i>EnterFinding_bn</i> .
<i>IsStateFinding</i>	Método público que verifica se o nó está em <i>Finding</i> .
<i>GetStateBeliefs</i>	Método público que devolve o percentual de um determinado estado do nó
<i>UpdateStates</i>	Método público que atualiza os valores dos estados na tela

5. TESTES, RESULTADOS OBTIDOS E DISCUSSÃO

Neste capítulo, serão demonstrados três testes feitos no *E-BAYES* e comparados com os resultados gerados pelo *Netica*, através do funcionamento do seu algoritmo de Inferência Bayesiana que também será demonstrado passo a passo.

Logo após, serão apresentados os resultados e discussões gerados pelo *E-BAYES*, por meio de simulações que possíveis de serem feitas pelo usuário final.

5.1. Testes Realizados (Inferência Bayesiana)

Para tanto, será tomado como exemplo o caso de um aluno consultado na base de dados, analisado-o de três formas diferentes. Havendo assim, uma variação do nó *Bolsa/Apoio Mensalidade*, nas seguintes situações: **Sem bolsa auxílio, com bolsa de esportes** e por último, com **bolsa de monitoria**.

Enquanto que as figuras 5.1, 5.3 e 5.5 apresentarão os resultados das atualizações da rede bayesiana via *Netica* da *Norsys Co.*, as figuras 5.2, 5.4 e 5.6 demonstrarão os mesmos resultados, porém, com o uso do *Evasão Bayes*, como dito antes, desenvolvido ao longo deste trabalho.

Utilizando-se ainda, do acesso a *DLL* do *Netica*, nesta *DLL* encontram-se todas as funções e procedimentos referentes ao processo de manipulação de uma rede probabilística (Bayesiana).

5.1.1. Exemplo A – Aluno sem Bolsa Auxílio

A.1. Evidências para o indivíduo

- e1*: Oriundo de Santarém;
e2: Sem bolsa auxílio;
e3: Sexo Masculino;

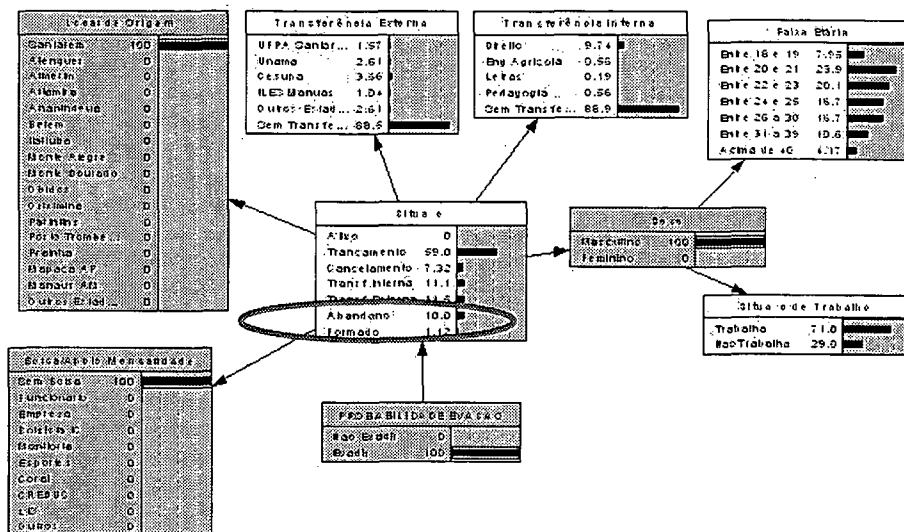


Figura 5.1. : Resultado da consulta, do exemplo A, via Netica.

A.2. Cálculo da probabilidade a posteriori para diagnóstico dá probabilidade de um aluno evadir ou não.

$$P(H_i | e1 \wedge e2 \wedge e3) = [P(e1 \wedge e2 \wedge e3)]^{-1} * P(H_i) * P(e1|H_i) * P(e2|H_i) * P(e3|H_i)$$

Sendo:

$P(e_i|H_i)$: Probabilidades Condicionais;
 H_i : Hipótese diagnostica à priori.

Pode ser reescrita:

$$P(H_i | e1 \wedge e2 \wedge e3) = \alpha * P(H_i) * \Delta$$

Sendo:

$$\lambda_i = P(e_i|H_i)$$

$$\alpha = [P(e1 \wedge e2 \wedge e3)]^{-1}$$

$$\Delta = \lambda_1 * \lambda_2 * \lambda_3$$

A.3 Probabilidades Condicionais

A tabela 5.1 foi alimentada a partir de um levantamento estatístico realizado sob as informações contidas no Banco de Dados e referentes aos alunos e ex-alunos do curso de objeto de estudo. A partir daí, extraí-se cada linha da tabela acima a partir dos nós envolvidos na consulta, mais especificamente, da coluna da evidência (*click*).

Tabela 5.1: Tabela das Estatísticas para o Cálculo das Probabilidades Condicionais

	<i>Ativo</i>	<i>Trancamento</i>	<i>Cancelamento</i>	<i>Transf. Externa</i>	<i>Transf. Interna</i>	<i>Abandono</i>	<i>Formado</i>
λ_1	0,8610	0,79592	1,0000	0,98276	0,77273	0,77778	1,0000
λ_2	0,7268	0,88000	1,0000	0,91000	0,94000	1,00000	0,6250
λ_3	0,5989	0,64290	0,5882	0,67800	0,90910	0,59260	0,3750

$$\text{Calculando } \Delta = \lambda_1 * \lambda_2 * \lambda_3$$

Para o preenchimento da tabela 5.2 multiplica-se o conteúdo de cada coluna da tabela 5.1 para obter o Δ , i.e. na tabela 5.1, coluna 2, onde $\text{Ativo} = 0.8610 * 0.7268 * 0.5989 = 0,37477652$ (tabela 5.2, coluna 2 - Ativo)

Tabela 4.2: Tabela dos Δ para o Cálculo das Probabilidades Condicionais

	<i>Ativo</i>	<i>Trancamento</i>	<i>Cancelamento</i>	<i>Transf. Externa</i>	<i>Transf. Interna</i>	<i>Abandono</i>	<i>Formado</i>
Δ	0,37477	0,45029	0,58820	0,60634	0,66033	0,46091	0,23437

$$\text{Calculando } \Delta * P(H_i)$$

Na linha 1 da tabela 5.3, $P(H_i)$ é colocado o valor extraído do nó **Situação** da rede bayesiana. Este valor será multiplicado por Δ (tabela 5.2) e colocado na linha 2 da tabela 5.3 em suas respectivas colunas, i.e. cálculo do valor de Trancamento = $0.45029333 * 0.6377073 = 0,28686972$ (tabela 5.3, linha 2 e coluna 3).

Tabela 5.3: Tabela das $P(H_i)$ do nó Situação para o Cálculo das Probabilidades Condicionais

	<i>Ativo</i>	<i>Trancamento</i>	<i>Cancelamento</i>	<i>Transf. Externa</i>	<i>Transf. Interna</i>	<i>Abandono</i>	<i>Formado</i>
$P(H_i)$	0,0000	0,637073	0,0605158	0,088678	0,0846322	0,105827	0,023274
$\Delta * P(H_i)$	0,0000	0,2868697	0,0355953	0,0537693	0,0558859	0,0487769	0,0054548

$$\text{Calculando } 1 = \alpha * \Delta * P(H_i)$$

$$1 = \alpha * \Delta * P(H_i)$$

$$1 = \alpha * (0,0 \ 0,28686972 \ 0,0355953935 \ 0,05376930803 \ 0,0558859856 \ 0,04877697 \ 0,00545484)$$

$$1 = \alpha * 0,48635221713$$

$$\alpha = \mathbf{2,056123041653}$$

$$\text{Calculando } P(H_i|e1 \wedge e2 \wedge e3)$$

$$P(H_i|e1 \wedge e2 \wedge e3) = \mathbf{2,056123041653} * (0,0 \ 0,2868697 \ 0,035595393 \ 0,0537693080 \ 0,055885985 \ 0,0487769 \ 0,0054548)$$

Nesta última tabela (5.4), a atualização bayesiana é encontrada, multiplicando-se os valores encontrados na tabela 5.3, cálculo da probabilidade condicional relacionada às ocorrências das evidências encontradas em $\Delta * P(H_i)$, multiplicados por α , que neste exemplo vale **2,056123041653**, em seguida na segunda linha da tabela dois, obtêm-se em termos de probabilidade, cada valor estimado no caso de uma possível evasão de um dado aluno, ou grupo de alunos com estas características.

Tabela 5.4: Tabela responsável pela atualização bayesiana, dada às evidências apresentadas e suas respectivas probabilidades condicionais calculadas.

	<i>Ativo</i>	<i>Trancamento</i>	<i>Cancelamento</i>	<i>Transf. Externa</i>	<i>Transf. Interna</i>	<i>Abandono</i>	<i>Formado</i>
$P(H_i e1 \wedge e2 \wedge e3)$	0	0.58984	0.073188	0.11056	0.11491	0.10029	0.011216
%	0	58,984	7,3188	11,056	11,491	10,029	1,1216

A.4 Conclusão do Exemplo A

Para o aluno analisado, dado às evidências, ele apresenta uma perspectiva razoável (58,98 %) de evadir por Trancamento (figura 5.2). Caso este aluno obtêm-se uma bolsa de esportes ou de monitoria, esta probabilidade se reverteria, como será mostrado nos Exemplo B e C, respectivamente.

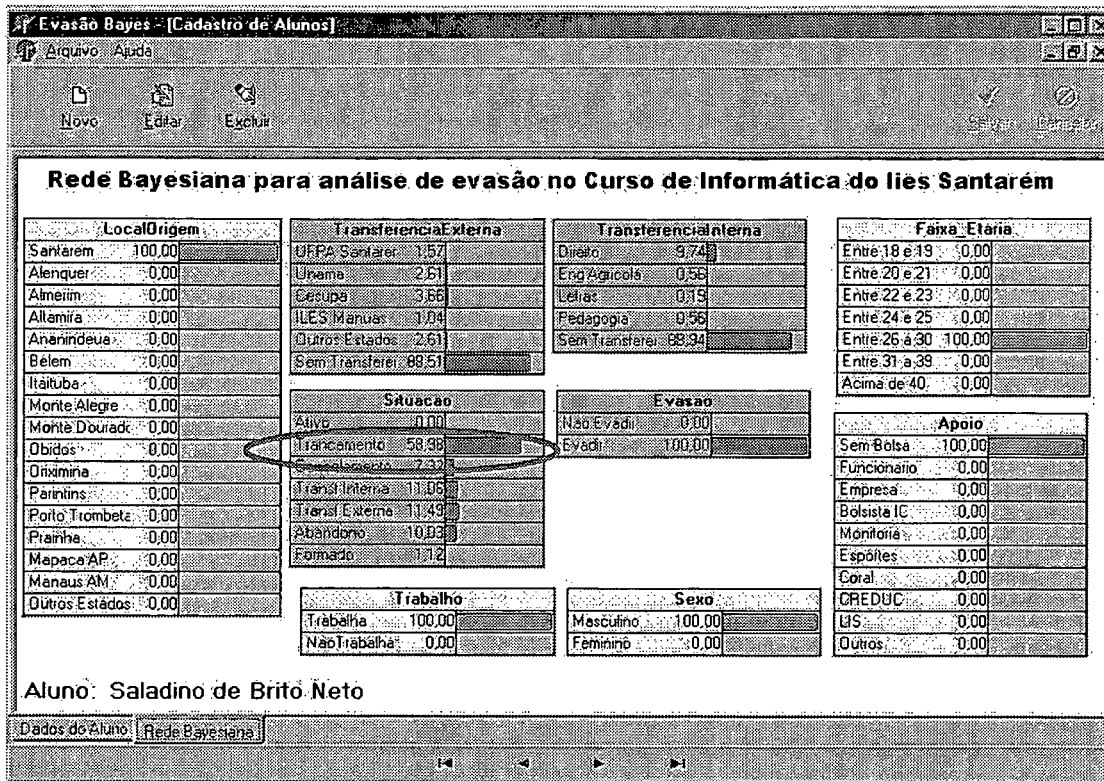


Figura 5.2. : Resultado da consulta, do exemplo A, via Evasão Bayes.

5.1.2 Exemplo B – Aluno com Bolsa Auxílio de Esportes

B.1. Evidências para indivíduo

- e1*: Oriundo de Santarém;
e2: Bolsa auxílio de Esportes;
e3: Sexo Masculino;

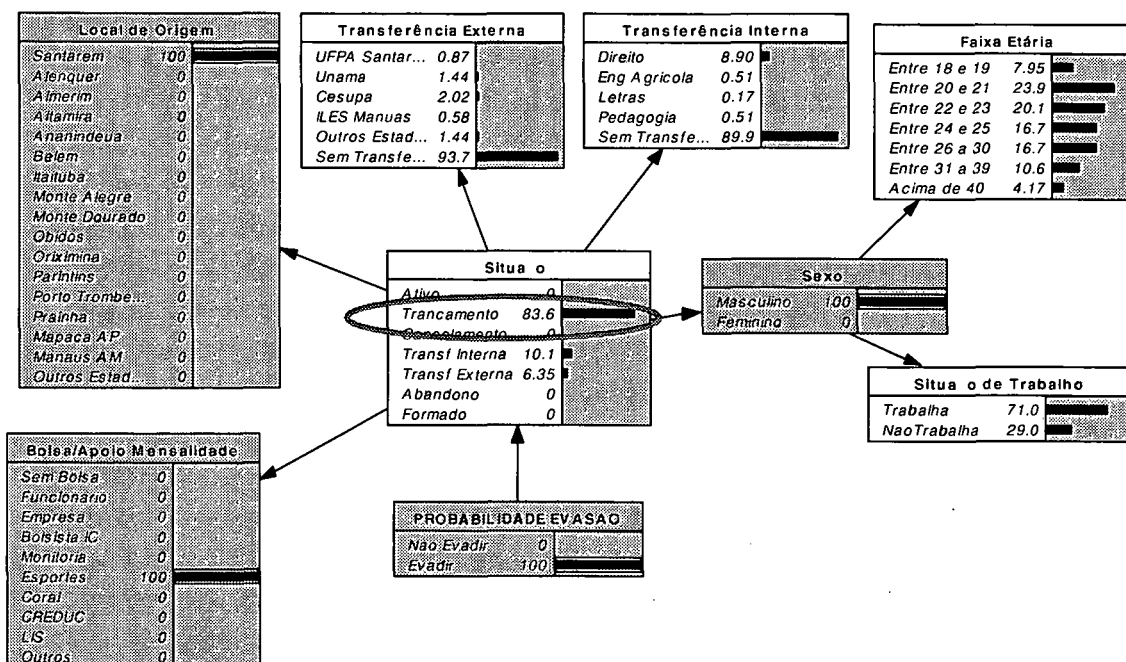


Figura 5.3: Resultado da consulta, do exemplo B, via Netica

B.2. Cálculo da probabilidade *a posteriori* para diagnóstico dá probabilidade de um aluno evadir ou não.

$$P(H_i | e1 \wedge e2 \wedge e3) = [P(e1 \wedge e2 \wedge e3)]^{-1} * P(H_i) * P(e1|H_i) * P(e2|H_i) * P(e3|H_i)$$

Sendo:

$P(e_i|H_i)$: Probabilidades Condicionais;

H_i : Hipótese diagnóstica *a priori*.

Pode ser reescrita:

$$P(H_i | e1 \wedge e2 \wedge e3) = \alpha * P(H_i) * \Delta$$

Sendo:

$$\lambda_i = P(e_i|H_i)$$

$$\alpha = [P(e1 \wedge e2 \wedge e3)]^{-1}$$

$$\Delta = \lambda_1 * \lambda_2 * \lambda_3$$

B.3. Probabilidades Condicionais

A tabela 5.5 foi alimentada a partir de um levantamento estatístico realizado sob as informações contidas no Banco de Dados e referentes aos alunos e ex-alunos do curso de objeto de estudo. A partir daí, extraí-se cada linha da tabela acima a partir dos nós envolvidos na consulta, mais especificamente, da coluna da evidência (*click*).

Tabela 5.5: Tabela das Estatísticas para o Cálculo das Probabilidades Condicionais

	<i>Ativo</i>	<i>Trancamento</i>	<i>Cancelamento</i>	<i>Transf. Externa</i>	<i>Transf. Interna</i>	<i>Abandono</i>	<i>Formado</i>
λ_1	0,86100	0,79592	1,0000	0,98276	0,77273	0,77778	1,0000
λ_2	0,10309	0,12000	0,0000	0,08000	0,05000	0,00000	0,0000
λ_3	0,59890	0,64290	0,5882	0,67800	0,90910	0,59260	0,3750

$$\text{Calculando } \Delta = \lambda_1 * \lambda_2 * \lambda_3$$

Para o preenchimento da tabela 5.6 multiplica-se o conteúdo de cada coluna da tabela 5.5 para obter o Δ , i.e. na tabela 5.5, coluna 2, onde *Ativo* = $0,8610 * 0,10309 * 0,5989 = 0,053158657461$ (tabela 5.6, coluna 2 - *Ativo*)

Tabela 5.6: Tabela dos Δ para o Cálculo das Probabilidades Condicionais

	<i>Ativo</i>	<i>Trancamento</i>	<i>Cancelamento</i>	<i>Transf. Externa</i>	<i>Transf. Interna</i>	<i>Abandono</i>	<i>Formado</i>
Δ	0,053158	0,061403	0,0000	0,053304	0,035124	0,0000	0,0000

$$\text{Calculando } \Delta * P(H_i)$$

Na linha 1 da tabela 5.7, $P(H_i)$ é colocado o valor extraído do nó **Situação** da rede bayesiana. Este valor será multiplicado por Δ (tabela 5.6) e colocado na linha 2 da tabela 5.7, em suas respectivas colunas, i.e. cálculo do valor de *Trancamento* = $0,06140363616 * 0,6377073 = 0,03911859869936$ (tabela 5.7, linha 2 e coluna 3).

Tabela 5.7: Tabela das $P(H_i)$ do nó Situação para o Cálculo das Probabilidades Condicionais

	<i>Ativo</i>	<i>Trancamento</i>	<i>Cancelamento</i>	<i>Transf. Externa</i>	<i>Transf. Interna</i>	<i>Abandono</i>	<i>Formado</i>
$P(H_i)$	0,0000000	0,637073	0,0605158	0,088678	0,0846322	0,105827	0,023274
$\Delta * P(H_i)$	0,0000000	0,0391185	0,0000000	0,0047269	0,0029726	0,0000000	0,0000000

$$\text{Calculando } 1 = \alpha * \Delta * P(H_i)$$

$$1 = \alpha * \Delta * P(H_i)$$

$$1 = \alpha * (0,000 \ 0,03911859869936 \ 0,000 \ 0,004726972135027 \ 0,002972658812927 \ 0,000)$$

$$1 = \alpha * 0,04681822964731$$

$$\alpha = \mathbf{21,3592014807}$$

$$\text{Calculando } P(H_i|e1 \wedge e2 \wedge e3)$$

$$P(H_i|e1 \wedge e2 \wedge e3) = 21,3592014807 * (0,0 \ 0,03911859869936 \ 0,0 \ 0,004726972135027 \ 0,002972658812927 \ 0,0 \ 0,0)$$

Nesta última tabela (5.8), a atualização bayesiana é encontrada, multiplicando-se os valores encontrados na tabela 5.7, cálculo da probabilidade condicional relacionada às ocorrências das evidências encontradas em $\Delta * P(H_i)$, multiplicados por α , que neste exemplo vale **21,3592014807**, em seguida na segunda linha da tabela dois, obtêm-se em termos de probabilidade, cada valor estimado no caso de uma possível evasão de um dado aluno, ou grupo de alunos com estas características.

Tabela 5.8: Tabela responsável pela atualização bayesiana, dada às evidências apresentadas e suas respectivas probabilidades condicionais calculadas.

	<i>Ativo</i>	<i>Trancamento</i>	<i>Cancelamento</i>	<i>Transf. Externa</i>	<i>Transf. Interna</i>	<i>Abandono</i>	<i>Formado</i>
$P(H_i e1 \wedge e2 \wedge e3)$	0,000	0.835540	0,0000000	0.100960	0.063494	0,000000	0,0000000
%	0	83,554	0	10,096	6,3494	0	0

B.4 Conclusão do Exemplo B

Para o aluno analisado, dado às evidências apresentadas, ele possui uma perspectiva forte (83,55%) de evadir por Trancamento (figura 5.4.). Porém, nenhum aluno que obteve bolsa de esportes conseguiu concluir o curso, vale lembrar que através de uma consulta simples no sistema, poderá ser observado que apenas 37,5% dos alunos com bolsa de esportes permaneceram no curso, enquanto que os demais 62,5% já evadiram. Caso este aluno obtê-se uma bolsa de monitoria, esta probabilidade se reverteria, como será mostrado nos Exemplo C.

Evasão Bayes - [Cadastro de Alunos]

Arquivo Ajuda

Novo Editar Excluir

Rede Bayesiana para análise de evasão no Curso de Informática do Iles Santarém

LocalOrigem	TransferenciaExterna	TransferenciaInterna	Faixa_Etaria
Santarém: 100,00	UFPA Santarém: 0,87	Direito: 8,90	Entre 18 e 19: 0,00
Alequer: 0,00	Unama: 1,44	Eng Agrícola: 0,51	Entre 20 e 21: 0,00
Almeirim: 0,00	Esuply: 2,02	Letras: 0,17	Entre 22 e 23: 0,00
Altamira: 0,00	ILES Maniás: 0,56	Pedagogia: 0,51	Entre 24 e 25: 0,00
Ananindeua: 0,00	Duitos Estados: 1,44	Sem Transferer: 89,90	Entre 26 a 30: 100,00
Belem: 0,00	Sem Transferer: 93,65		Entre 31 a 39: 0,00
Itaituba: 0,00			Acima de 40: 0,00
Monte Alegre: 0,00			
Monte Dourado: 0,00			
Obidos: 0,00			
Oriximina: 0,00			
Parintins: 0,00			
Porto Trombetas: 0,00			
Prainha: 0,00			
Mapaca AP: 0,00			
Manaus AM: 0,00			
Duitos Estados: 0,00			

Situacao	Evasao
Abixo: 0,00	Nao Evadi: 0,00
Trancamento: 83,52	Evadi: 100,00
Cancelamento: 0,00	
Transf Interna: 10,10	
Transf Externa: 6,25	
Abandono: 0,00	
Formado: 0,00	

Trabalho	Sexo
Trabalha: 100,00	Masculino: 100,00
NaoTrabalha: 0,00	Feminino: 0,00

Apoio
Sem Bolsa: 0,00
Funcionario: 0,00
Empresa: 0,00
Bolsista IG: 0,00
Monitoria: 0,00
Esportes: 100,00
Coral: 0,00
CREDUC: 0,00
LIS: 0,00
Outros: 0,00

Aluno: Saládino de Brito Neto

Dados do Aluno: Rede Bayesiana

Figura 5.4: Resultado da consulta, do exemplo B, via Evasão Bayes.

5.1.3 Exemplo C – Aluno com Bolsa Auxílio de Monitoria

C.1. Evidências para indivíduo

- $e1$: Oriundo de Santarém;
 $e2$: Bolsa auxílio de Monitoria;
 $e3$: Sexo Masculino;

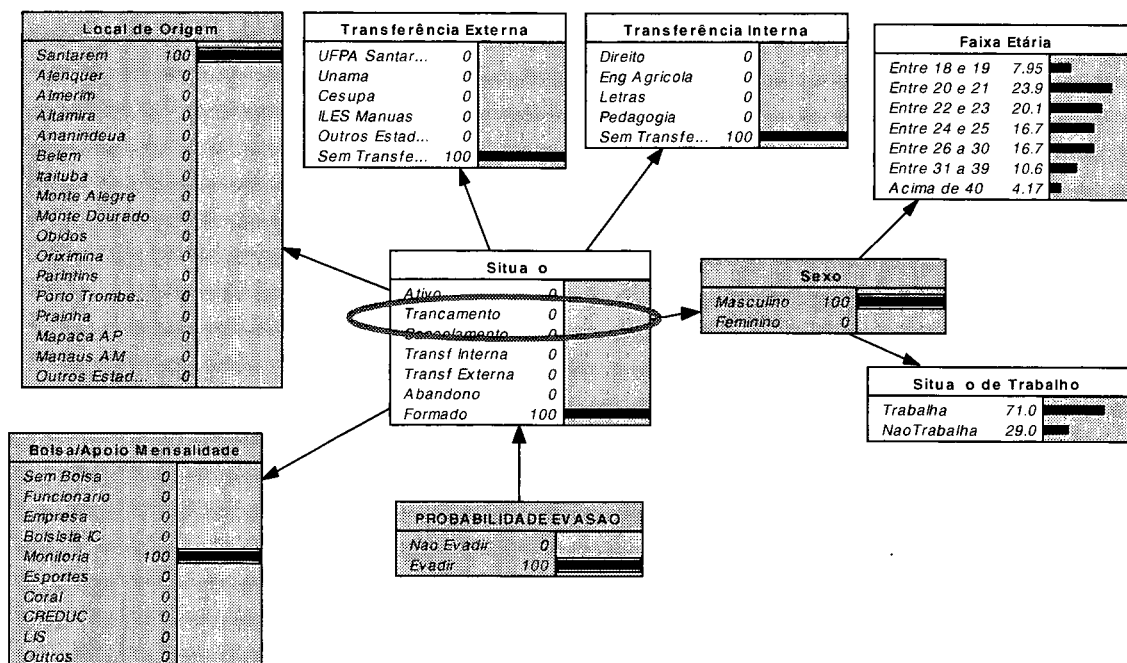


Figura 5.5 : Resultado da consulta, do exemplo C, via Netica

C.2. Cálculo da probabilidade *a posteriori* para diagnóstico da probabilidade de um aluno evadir ou não.

$$P(H_i | e1 \wedge e2 \wedge e3) = [P(e1 \wedge e2 \wedge e3)]^{-1} * P(H_i) * P(e1 | H_i) * P(e2 | H_i) * P(e3 | H_i)$$

Sendo:

$P(e_i | H_i)$: Probabilidades Condicionais;

H_i : Hipótese diagnóstica *a priori*.

Pode ser reescrita:

$$P(H_i | e1 \wedge e2 \wedge e3) = \alpha * P(H_i) * \Delta$$

Sendo:

$$\lambda_i = P(e_i | H_i)$$

$$\alpha = [P(e1 \wedge e2 \wedge e3)]^{-1}$$

$$\Delta = \lambda_1 * \lambda_2 * \lambda_3$$

$$\text{Calculando } 1 = \alpha * \Delta * P(H_i)$$

$$1 = \alpha * \Delta * P(H_i)$$

$$1 = \alpha * (0,000 \ 0,000 \ 0,000 \ 0,000 \ 0,000 \ 0,000 \ 0,0021819375)$$

$$1 = \alpha * 0,0021819375$$

$$\alpha = \mathbf{458,3082695998}$$

$$\text{Calculando } P(H_i|e1 \wedge e2 \wedge e3)$$

$$P(H_i|e1, e2, e3) = 458,3082695998 * (0,000 \ 0,000 \ 0,000 \ 0,000 \ 0,000 \ 0,000 \ 0,0021819375)$$

Nesta última tabela (5.12), a atualização bayesiana é encontrada, multiplicando-se os valores encontrados na tabela 5.11, cálculo da probabilidade condicional relacionada às ocorrências das evidências encontradas em $\Delta * P(H_i)$, multiplicados por α , que neste exemplo vale **458,3082695998**, em seguida na segunda linha da tabela dois, obtêm-se em termos de probabilidade, cada valor estimado no caso de uma possível evasão de um dado aluno, ou grupo de alunos com estas características.

Tabela 5.12: Tabela responsável pela atualização bayesiana, dada às evidências apresentadas e suas respectivas probabilidades condicionais calculadas.

	<i>Ativo</i>	<i>Trancamento</i>	<i>Cancelamento</i>	<i>Transf. Externa</i>	<i>Transf. Interna</i>	<i>Abandono</i>	<i>Formado</i>
$P(H_i e1, e2, e3)$	0,000	0,0000000	0,0000000	0,0000000	0,0000000	0,0000000	1,0000000
<i>%</i>	0	0	0	0	0	0	100

C.4 Conclusão do Exemplo C

Para o aluno analisado, dado às evidências apresentadas, o mesmo apresenta uma perspectiva excelente (zero %) de evadir por Trancamento (figura 5.6) ou por qualquer outro *status*, senão o de formado. Além disso, nenhum aluno que obteve bolsa de monitoria evadiu do curso, vale lembrar que através de uma consulta simples no sistema, poderá ser observado que apenas 76,0 % dos alunos com bolsa de monitoria permanecem no curso, enquanto que os demais 24,0% já evadiram, entretanto, sob o *status* de formado. Da mesma forma 100 % dos alunos com bolsa de iniciação científica permanecem no curso, enquanto que os demais: 0 % (zero) evadiram.

Evasão Bayes - [Cadastro de Alunos]

Arquivo Ajuda

Novo Editar Excluir

Fechar Cancelar

Rede Bayesiana para análise de evasão no Curso de Informática do Iles Santarém

LocalOrigem		TransferenciaExterna		TransferenciaInterna		Faixa Etaria	
Santarém	100,00	UFPA Santarém	0,00	Direito	0,00	Entre 18 e 19	0,00
Alenquer	0,00	Uirama	0,00	Eng Agrícola	0,00	Entre 20 e 21	0,00
Almerim	0,00	Cesupa	0,00	Letras	0,00	Entre 22 e 23	0,00
Allamira	0,00	ILES Manaus	0,00	Pedagogia	0,00	Entre 24 e 25	0,00
Ahanindeua	0,00	Outros Estados	0,00	Sem Transferência	100,00	Entre 26 a 30	100,00
Belém	0,00	Sem Transferência	100,00			Entre 31 a 39	0,00
Itaituba	0,00					Acima de 40	0,00
Monte Alegre	0,00						
Monte Dourado	0,00						
Obidos	0,00						
Oriximiná	0,00						
Parintins	0,00						
Porto Trombetas	0,00						
Prainha	0,00						
Mapaca AP	0,00						
Manaus AM	0,00						
Outros Estados	0,00						

Situação		Evasão	
Novo	0,00	Não Evade	0,00
Trancamento	0,00	Evadir	100,00
Cancelamento	0,00		
Transf Interna	0,00		
Transf Externa	0,00		
Abandono	0,00		
Formado	100,00		

Trabalho		Sexo	
Trabalha	100,00	Masculino	100,00
Não Trabalha	0,00	Feminino	0,00

Apoio	
Sem Bolsa	0,00
Funcionário	0,00
Empresa	0,00
Bolsista IC	0,00
Monitoria	100,00
Espótes	0,00
Coral	0,00
CREDUC	0,00
LIS	0,00
Outros	0,00

Aluno: Saladino de Brito Neto

Dados do Aluno Rede Bayesiana

Figura 5.6: Resultado da consulta, do exemplo C, via Evasão Bayes.

5.2. Resultados Obtidos e Discussão

Como resultados apresenta-se a rede bayesiana final (figura 5.7), utilizada como base para a implementação do protótipo do *software Evasão Bayes*. O mesmo passou por processos sucessivos de refinamento, conforme visto anteriormente.

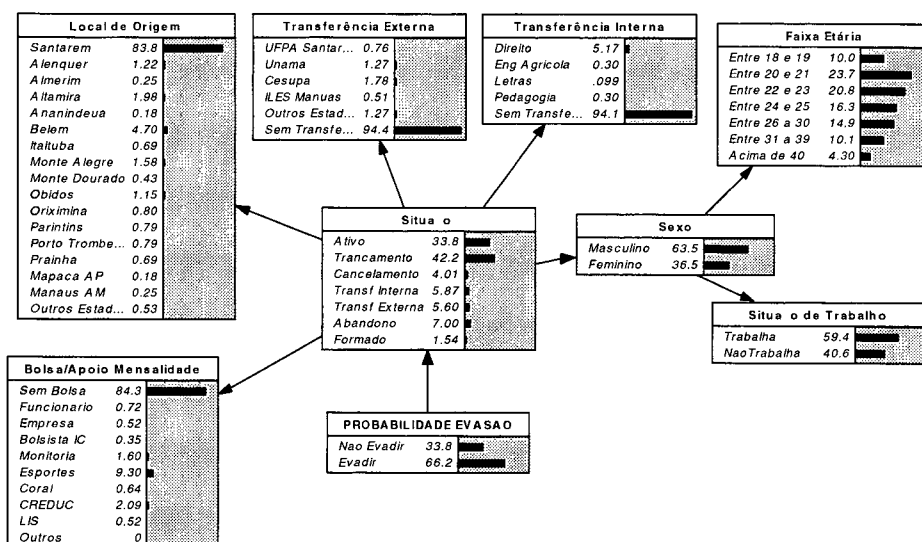


Figura 5.7 – Rede bayesiana final utilizada como base para a construção do *Evasão Bayes* (ROVARIS & NASSAR, 2001)

A seguir, um exemplo de consulta no *Evasão Bayes*. Sistema este, que permite a manutenção de registros dos alunos e a consulta individual dos seus registros (figura 5.8), além de uma visão particular da sua rede bayesiana (figura 5.9). Outro fator importante é a possibilidade que este aplicativo tem fazer simulações (figura 5.10) que permitam avaliar possíveis situações que possam vir a melhorar ou piorar as condições de permanência deste aluno no curso.

A figura 5.11 apresenta um exemplo de simulação para avaliar o comportamento do aluno através da rede bayesiana do mesmo aluno pelo *E-Bayes*. Neste caso, se a aluna trocasse a bolsa de monitoria (figura 5.12) por uma de iniciação científica, a mesma diminuiria a sua chance de evasão de 18,67 % para algo em torno de zero por cento, de acordo com os resultados demonstrados na figura 5.13.

Evasão Bayes - (Cadastro de Alunos)

Nome: Pietro Bezerra Macambira

Endereço: AV PLACIDO DE CASTRO 1700

Cidade: SANTARÉM PA

Estado: PA

CPF: 015231130

Ano: 2000

Curso: 010006

Sexo: Masculino Feminino

Idade: 20

Trabalho: Não Trabalha Trabalha

Trabalho: Trabalho Trabalho Trabalho

Bolsa / Apoio Monitoria: Sem Bolsa Monitoria Iniciação Científica Bolsista CTA CUB OPELUC LIS Outra

Evasão: Evasão (0.1867) Não Evasão (0.0000)

Figura 5.11: Exemplo B: consulta no registro de um aluno pelo Evasão Bayes

Rede Bayesiana para análise de evasão no Curso de Informática do Iles Santarém

Localização	Trabalho	Sexo	Evasão
Santarém: 1.0000	Trabalha: 1.0000	Masculino: 1.0000	Entre 18 e 19: 0.0000
Albuquerque: 0.0000	Não Trabalha: 0.0000	Feminino: 0.0000	Entre 20 e 21: 1.0000
Almeida: 0.0000			Entre 22 e 23: 0.0000
Alves: 0.0000			Entre 24 e 25: 0.0000
Ananias: 0.0000			Entre 26 e 27: 0.0000
Antônio: 0.0000			Entre 28 e 29: 0.0000
Araripe: 0.0000			Entre 30 e 31: 0.0000
Barcelos: 0.0000			Entre 32 e 33: 0.0000
Barcelos: 0.0000			Entre 34 e 35: 0.0000
Barcelos: 0.0000			Entre 36 e 37: 0.0000
Barcelos: 0.0000			Entre 38 e 39: 0.0000
Barcelos: 0.0000			Entre 40 e 41: 0.0000
Barcelos: 0.0000			Entre 42 e 43: 0.0000
Barcelos: 0.0000			Entre 44 e 45: 0.0000
Barcelos: 0.0000			Entre 46 e 47: 0.0000
Barcelos: 0.0000			Entre 48 e 49: 0.0000
Barcelos: 0.0000			Entre 50 e 51: 0.0000
Barcelos: 0.0000			Entre 52 e 53: 0.0000
Barcelos: 0.0000			Entre 54 e 55: 0.0000
Barcelos: 0.0000			Entre 56 e 57: 0.0000
Barcelos: 0.0000			Entre 58 e 59: 0.0000
Barcelos: 0.0000			Entre 60 e 61: 0.0000
Barcelos: 0.0000			Entre 62 e 63: 0.0000
Barcelos: 0.0000			Entre 64 e 65: 0.0000
Barcelos: 0.0000			Entre 66 e 67: 0.0000
Barcelos: 0.0000			Entre 68 e 69: 0.0000
Barcelos: 0.0000			Entre 70 e 71: 0.0000
Barcelos: 0.0000			Entre 72 e 73: 0.0000
Barcelos: 0.0000			Entre 74 e 75: 0.0000
Barcelos: 0.0000			Entre 76 e 77: 0.0000
Barcelos: 0.0000			Entre 78 e 79: 0.0000
Barcelos: 0.0000			Entre 80 e 81: 0.0000
Barcelos: 0.0000			Entre 82 e 83: 0.0000
Barcelos: 0.0000			Entre 84 e 85: 0.0000
Barcelos: 0.0000			Entre 86 e 87: 0.0000
Barcelos: 0.0000			Entre 88 e 89: 0.0000
Barcelos: 0.0000			Entre 90 e 91: 0.0000
Barcelos: 0.0000			Entre 92 e 93: 0.0000
Barcelos: 0.0000			Entre 94 e 95: 0.0000
Barcelos: 0.0000			Entre 96 e 97: 0.0000
Barcelos: 0.0000			Entre 98 e 99: 0.0000
Barcelos: 0.0000			Entre 100 e 101: 0.0000

Aluno: Pietro Bezerra Macambira

Figura 5.12: Exemplo B: consulta na rede bayesiana pelo E-Bayes.

Rede Bayesiana para análise de evasão no Curso de Informática do Iles Santarém

Localização	Trabalho	Sexo	Evasão
Santarém: 1.0000	Trabalha: 1.0000	Masculino: 1.0000	Entre 18 e 19: 0.0000
Albuquerque: 0.0000	Não Trabalha: 0.0000	Feminino: 0.0000	Entre 20 e 21: 1.0000
Almeida: 0.0000			Entre 22 e 23: 0.0000
Alves: 0.0000			Entre 24 e 25: 0.0000
Ananias: 0.0000			Entre 26 e 27: 0.0000
Antônio: 0.0000			Entre 28 e 29: 0.0000
Barcelos: 0.0000			Entre 30 e 31: 0.0000
Barcelos: 0.0000			Entre 32 e 33: 0.0000
Barcelos: 0.0000			Entre 34 e 35: 0.0000
Barcelos: 0.0000			Entre 36 e 37: 0.0000
Barcelos: 0.0000			Entre 38 e 39: 0.0000
Barcelos: 0.0000			Entre 40 e 41: 0.0000
Barcelos: 0.0000			Entre 42 e 43: 0.0000
Barcelos: 0.0000			Entre 44 e 45: 0.0000
Barcelos: 0.0000			Entre 46 e 47: 0.0000
Barcelos: 0.0000			Entre 48 e 49: 0.0000
Barcelos: 0.0000			Entre 50 e 51: 0.0000
Barcelos: 0.0000			Entre 52 e 53: 0.0000
Barcelos: 0.0000			Entre 54 e 55: 0.0000
Barcelos: 0.0000			Entre 56 e 57: 0.0000
Barcelos: 0.0000			Entre 58 e 59: 0.0000
Barcelos: 0.0000			Entre 60 e 61: 0.0000
Barcelos: 0.0000			Entre 62 e 63: 0.0000
Barcelos: 0.0000			Entre 64 e 65: 0.0000
Barcelos: 0.0000			Entre 66 e 67: 0.0000
Barcelos: 0.0000			Entre 68 e 69: 0.0000
Barcelos: 0.0000			Entre 70 e 71: 0.0000
Barcelos: 0.0000			Entre 72 e 73: 0.0000
Barcelos: 0.0000			Entre 74 e 75: 0.0000
Barcelos: 0.0000			Entre 76 e 77: 0.0000
Barcelos: 0.0000			Entre 78 e 79: 0.0000
Barcelos: 0.0000			Entre 80 e 81: 0.0000
Barcelos: 0.0000			Entre 82 e 83: 0.0000
Barcelos: 0.0000			Entre 84 e 85: 0.0000
Barcelos: 0.0000			Entre 86 e 87: 0.0000
Barcelos: 0.0000			Entre 88 e 89: 0.0000
Barcelos: 0.0000			Entre 90 e 91: 0.0000
Barcelos: 0.0000			Entre 92 e 93: 0.0000
Barcelos: 0.0000			Entre 94 e 95: 0.0000
Barcelos: 0.0000			Entre 96 e 97: 0.0000
Barcelos: 0.0000			Entre 98 e 99: 0.0000
Barcelos: 0.0000			Entre 100 e 101: 0.0000

Aluno: Pietro Bezerra Macambira

Figura 5.13: Exemplo B: simulação

Neste exemplo de consulta no Evasão Bayes (figura 5.14). Pode-se observar que no caso desta aluna, que é oriunda do município de Óbidos (figura 5.15), onde se ela fosse de Santarém (figura 5.16) a sua probabilidade de evadir aumentaria de 43,59 % para 71,76%.

Figura 5.14: Exemplo C: consulta no registro de um aluno pelo Evasão Bayes

Figura 5.15: Exemplo C: consulta na rede bayesiana pelo E-Bayes.

Figura 5.16: Exemplo C: simulação.

6 CONCLUSÕES E TRABALHOS FUTUROS

O presente trabalho teve como objetivo aplicar técnicas de *data mining* e conhecimentos de redes *bayesianas* para simular um Sistema Especialista Probabilístico - SEP que avaliasse a situação da evasão dos alunos do curso de graduação de uma IES, por meio de E-Bayes (Evasão Bayes), um *software* elaborado para este fim.

6.1 Conclusões

Os resultados obtidos até o momento com o *E-Bayes*, são satisfatórios de acordo com os testes realizados e demonstrados. Uma vez que ele consegue apresentar um modelo de rede bayesiana capaz de retratar a situação acadêmica atual do curso em estudo, bem como os motivos que podem levar um aluno a evadir ou permanecer em seu curso. Com este *software* esta sendo possível ainda, quantificar por meios de suas probabilidades e qualificar por meio de deduções (parte gráfica) quais os principais motivos que podem levar a um aluno abandonar o seu curso.

O *E-Bayes*, por ser um Sistema Especialista Probabilístico, se comunica com a rede bayesiana denominada <EvasaoAlunos.dne>, criada por um especialista no domínio da aplicação.

As técnicas de *data mining* utilizadas, embora consideradas simples e básicas por alguns autores, foram outro fator importante para o sucesso da elicitação destes dados internos e externos à base de dados da empresa. Pois os cálculos estatísticos e os nós (grupos) gerados se tornaram essenciais para a implementação da rede bayesiana.

Desta forma, foram estas informações extraídas a partir dessa base de dados que alimentaram a rede bayesiana, projetada para este fim, após sucessivos refinamentos até se chegar na proposta da rede final.

Não se espera que o *E-Bayes*, um Sistema Especialista Probabilístico, venha a substituir o especialista humano, mas sim que ele ofereça suporte ao mesmo. Auxiliando-o no processo decisório da sua organização, assumindo assim, uma característica que é clara dos Sistemas de Apoio à Decisão. Além do fato do mesmo ter sido alimentado por um processo de *data mining* realizado na base de dados da empresa, bem como, também por informações externas a esta base de dados. Sendo esta, uma outra característica dos Sistemas de Apoio à Decisões.

Um outro momento muito importante propiciado pelo *E-Bayes* está sendo o da sua possibilidade de simular novas situações coletivas ou individualmente (aluno-a-aluno), possibilitando desta maneira a avaliação do impacto gerado, dadas as novas evidências.

Observa-se ainda, que o *E-Bayes* (Evasão Bayes), é um Sistema Especialista Probabilístico que utiliza o Teorema da Probabilidade como método de raciocínio. Porém comportando-se como um Sistema de Apoio à Decisões, capaz de interferir no processo decisório de uma organização, no caso deste estudo, uma Instituição de Ensino Superior Particular.

O *E-Bayes* proporciona uma gama excepcional de opções para visualização de estatísticas e nas previsões probabilísticas apoiadas na probabilidade condicional e no Teorema de Bayes, que são as bases do raciocínio com incerteza por aleatoriedade.

Concluindo, o *E-Bayes* vem demonstrando ser uma ferramenta, capaz de quantificar e qualificar esses dados, transformando-os em informação e gerando conhecimento capaz de interferir, como dito anteriormente, no processo decisório de uma empresa desta modalidade, ou seja, um SAD - Sistema de Apoio à Decisão.

6.2 Trabalhos Futuros

Como trabalhos futuros, poderão ser implementadas novas redes ou atualizar a existente. Sendo que com base nas respostas obtidas no questionário aplicado, novos nós poderão ser criados referentes ao grau de satisfação do aluno, conhecimento que o mesmo tem sobre a estrutura da sua faculdade (figura 6.1) e dados pessoais como, por exemplo, à classe social que o mesmo está inserido (figura 6.2).

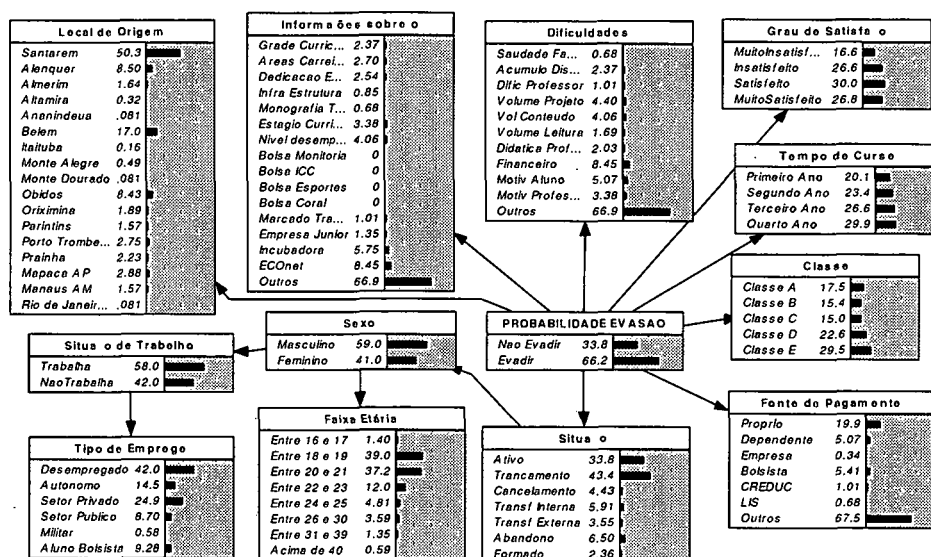


Figura 6.1: Possível atualização com base no questionário aplicado (dados hipotéticos)

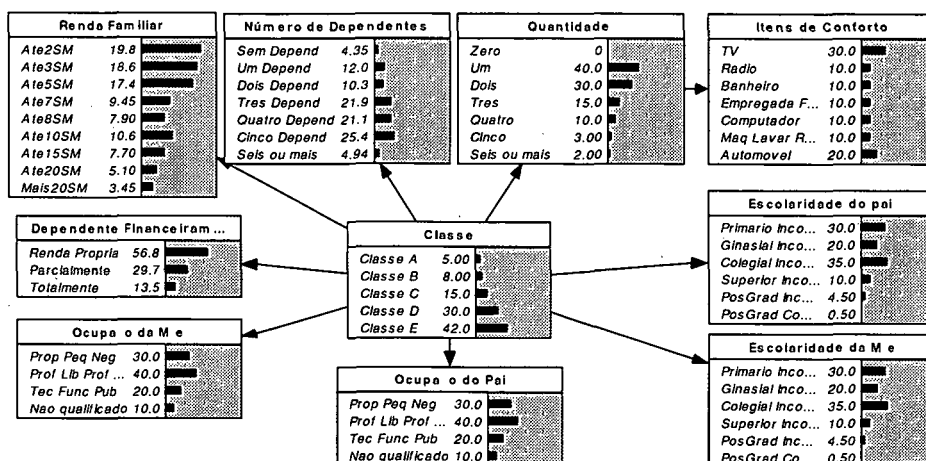


Figura 6.2: Possível rede nova, com base no questionário aplicado (dados hipotéticos)

Independente do modelo de rede a ser implementada ou atualizada, o primeiro passo deverá ser a busca pela total desvinculação, hoje existente, da biblioteca <Netica.dll>. Uma vez que o modelo atual implica em elevados custos com a licença de uso deste software.

Outras possibilidades de melhoramento deste trabalho seriam:

- A flexibilização do *software*, permitindo-o funcionar nos demais cursos, bem como, em outras unidades de ensino superior. Com as facilidades de instalação e configuração existentes em produtos comercializáveis;
- Aumentar a contribuição dos educadores, frente as suas necessidades perante o protótipo; isto possibilitaria inclusive a ampliação da base de dados contemplada por estes educadores;
- Uso da inteligência artificial híbrida, com o uso de multiagentes na busca de novos padrões na base de dados;
- Disponibilização desta ferramenta on-line e se possível via intranet.
- Adaptação do *E-BAYES* para que ele trabalhe com *data mining* dinâmico. Todavia, diferentemente do *data mining* estático, o dinâmico, segundo (INMON, 2001, p. 185) requer um planejamento e uma infra-estrutura tecnológica diferente, uma vez que precisa-se trabalhar “...com monitoração, gerenciamento de períodos de carregamento, gerenciamento de períodos de atualização e depuração...”. Finalmente, torna o processo de mineração de dados muito mais rigoroso e disciplinado.

7 REFERÊNCIAS BIBLIOGRÁFICAS

AMARAL, F.C.N. *Data mining - Técnicas e aplicações para o marketing direto*. São Paulo: Berkeley, 2001.

BARRETO, Jorge Muniz. **Inteligência artificial no limiar no século XXI – abordagem híbrida simbólica, conexcionista e evolucionária**. 2. ed. Florianópolis: J. M. Barreto, 1999. *Apud in* (MUCULLOCH, 1943), (ROSENBLATT, 1958), (MCCARTHY, 1956), (HOPFIELD, 1982), (MINSKY, 1969).

BASTOS, C. L.; KELLER V. **Aprendendo a lógica**. 6. ed. Petrópolis: Vozes, 1998.

BRIGNOLI, J.T. **Modelo híbrido difuso probabilístico**. Dissertação de Mestrado, Curso de Pós Graduação em Ciência da Computação, Universidade Federal de Santa Catarina, Florianópolis, 2001.

CECCON, C. *et. al.* **A vida na escola e a escola da vida**. 32. ed. Petrópolis: Vozes, 1997.

COSTA, N. C. A. da. *et al.* **Lógica paraconsistente aplicada**. São Paulo: Atlas. 1999.

CUNHA, H.; RIBEIRO, S. **Introdução aos sistemas especialistas**. Rio de Janeiro: LTC, 1987.

DAL MAS DIAS, Elaine. **A dúvida da continuidade dos estudos universitários: uma questão adolescente**. Taubaté: Cabral, 1997.

DUDA, R. O. *et al.* **A computer-based consultant for mineral exploration**. Technical report, SRI International, Sept. 1979.

FELDBAUM, A. **Principes theoriques des systemes asservis optimaux**. Tradução: V. Polonski. Moscou: Mir, 1973.

FRIGOTTO, G. **A Produtividade da escola improdutiva**. 3. ed. São Paulo: Cortez, 1989.

FUSTER-PARRA, Pilar. **A model for casual diagnostic reasoning. Extended inference modes and efficiency problems**. 1996. 225 f. Tese (Doctorado en Informàtica) – Departament de Ciències Matemàtiques e Informàtica, Universitat de les Illes Balears, Palma de Mallorca, Espanha.

GAAG, L. C., **Bayesian belief networks: Odds and ends**. The computer journal, vol. 39, n. 2, 1996, p. 97-113.

GANASCIA, J. G., **Inteligência artificial**. Tradução: Reginaldo Carmello Corrêa de Moraes. São Paulo: Ática, 1997. Título original: L'intelligence artificielle.

GENARO S., **Sistemas Especialistas: o conhecimento artificial**. Rio de Janeiro: LTC, 1986.

GERWEN, E.V. **Unit:** Pascal para o *Netica*. Disponível em : http://www.norsys.com/dl/delphi_Netica.txt. Acesso em: 10 de dezembro de 2001.

GIROUX, H. A. *Os professores como intelectuais*. Porto Alegre: Artes Médicas, 1997.

GONCALO, J. S. *et al.* **Aplicação de redes bayesianas na mensuração do status sócio-econômico do brasileiro**. INE, CTC, UFSC, 1999. *Apud in* (LAURITZEN e SPIEGELHALTER, 1988).

HOLTZ, F. **Sistemas especialistas: Programando em Turbo C**. Tradução Fernando Cabral. Rio de Janeiro: Campus, 1991. Título original: Expert systems.

HOPFIELD, J. **Neural networks and physical systems with emergent collective computational abilities**. Proceedings of the National Academy of Sciences 79 (1982), 2554-2558.

INMON, W.H. *et al.* **Data Wharehousing - Como transformar informações em oportunidades de negócios**. Tradução: Melissa Kassner. São Paulo: Berkeley, 2001.

KASABOV, N. K. **Foundations of neural networks, fuzzy systems, an knowledge engineering**. MIT Press. Cambridge. 2. ed. 1998.

KOVÁCS, Z. L. **Redes neurais artificiais: fundamentos e aplicações**. São Paulo: Collegium Cognitio, 1996.

KOEHLER, C. **Uma Abordagem Probabilística para Sistemas Especialistas**. Dissertação de Mestrado, Curso de Pós Graduação em Ciência da Computação, Universidade Federal de Santa Catarina, Florianópolis, 1998.

LADEIRA, M. *et al.* **Desenvolvimento de sistemas inteligentes baseados em raciocínio probabilístico**. In: I Jornada de Atualização em Inteligência Artificial. 1., 2001, Fortaleza. Anais XXI Congresso da Sociedade Brasileira de Computação. Fortaleza: UNIFOR, 2001. p. 143-186. *Apud in* (PEARL, 1986).

LAUDON, K. C.; LAUDON J.P. **Gerenciamento de sistemas de informação**. Tradução: Alexandre Oliveira. Rio de Janeiro: LTC, 2001.

(LAURITZEN e SPIEGELHALTER, 1988) LAURITZEN, S. L.; SPIEGELHALTER, D. J. **Local computations with probabilities on graphical structures and their applications to expert systems**. J. Royal Statistic. Soc., B, 50(2):154-227. 1988.

MCCARTHY, J.; SHANON, C. E. **Automata studies**. Princeton University Press, New Jersey, 1956.

MORAIS, R. *Sala de aula: Que espaço é esse?* 10. ed. Campinas: Papyrus, 1996.

MUCULLOCH, W.S.; PITTS, W. H. **A logical calculus of ideas immanent in nervous activity**. Bull of Mathematical Biophysics 5 (1943), 115-133.

MINSKY, M. L.; PAPERT, S. A. **Perceptrons: an introduction to computational geometry**. The MIT Press, 1969.

NASSAR, S. M. **A Estatística como apoio à inteligência artificial: sistemas especialistas probabilísticos**. In: Estatística e Informática: um processo interativo entre duas ciências. Trabalho apresentado no Concurso para Professor Titular, INE, CTC, UFSC, 1998.

NETICA. Norsys Corporation. Disponível em: <http://www.norsys.com.br>. Acesso em: 23 de março de 2000.

NEWELL, A.; SIMON, H. **GPS a program that simulates human thought**. In **computers and thought**. Eds. Academic Press, New York, 1963, p.279-93.

PEARL J. **A constraint-propagation and structing in belief networks**. In: Uncertainty in artificial intelligence, edited by Laveen Kanal and J. Lemmer, Amsterdam: North Holland, p. 357-370, 1986.

PEARL J. **Probabilistic reasoning in intelligent systems: networks of plausible inference**. Morgan Kaufmann. San Francisco, California. 1988.

PEARL J. **Bayesian networks**. UCLA Cognitive Systems Laboratory, Technical Report (R-216), Revision I. In M. Arbib (Ed.), *Handbook of Brain Theory and Neural Networks*, MIT Press, 149-153, 1995.

PEARL J. **Bayesian networks**. UCLA Cognitive Systems Laboratory, Technical Report (R-246), Revision I, July 1997. In MIT Encyclopedia of the Cognitive Sciences, Cambridge, MA, 1999.

PEARL J.; RUSSELL S. **Bayesian networks**. UCLA Cognitive Systems Laboratory, Technical Report (R-277), November 2000. To appear in M. Arbib (Ed.), *Handbook of Brain Theory and Neural Networks*, MIT Press, 2001.

POLLONI, E.G.F. **Sistemas de informação - estudo de viabilidade**. São Paulo: Futura, 2000.

RABUSKE, R.A. **Inteligência artificial**. Florianópolis: UFSC, 1995.

RICH, E. **Inteligência artificial**. Tradução Newton Vasconcelos. São Paulo: McGraw-Hill, 1988. *Apud in* (DUDA, 1979). Título Original: Artificial Intelligence.

RODRIGUES, J. A. F. **Data mining: conceitos básicos e aplicações**. In: I Jornada de Estudos em Computação de Piracicaba e Região. 1., 2000, Piracicaba. Anais da I Jornada de Estudos em Computação de Piracicaba e Região. Piracicaba: UNIMEP, 2000. p. 67-77.

ROSENBLATT, F. **The perceptron: a probabilistic model for information storage and organization in the brain**. Psychological Review 65. 1958. p. 386-408.

ROVARIS NETO, E. **Estudo sobre a evolução da inteligência artificial**. In: SEMANA DE INFORMÁTICA DO ILES SANTARÉM, 2., 1999, Santarém. Anais da II Semana de Informática do ILES Santarém. Santarém: ILES, 1999. p. 68-74.

ROVARIS NETO, E. **Inteligência artificial**. 1999. 24 f. Monografia (*Latu Sensu*) – Departamento de Informática - Universidade Luterana do Brasil, Canoas.

ROVARIS NETO, E.; et al. **Aplicação de redes bayesianas na análise de evasão dos alunos do curso de informática do ILES Santarém**. In: CONGRESSO DE CIÊNCIA E TECNOLOGIA DA AMAZÔNIA, 1., 2000, Santarém. Livro de Resumos – I Salão de Iniciação Científica do ILES Santarém. Santarém: ILES, 2000. p. 79.

ROVARIS NETO, E.; et al. **Redes bayesianas versus regras de inferência...** In: CONGRESSO DE CIÊNCIA E TECNOLOGIA DA AMAZÔNIA, 2., 2001, Santarém. Livro de Resumos–II Salão de Iniciação Científica do ILES Santarém. Santarém:ILES, 2001. p.105.

ROVARIS NETO, E.; et al. **Paradigma simbólico versus paradigma numérico...** In: CONGRESSO DE CIÊNCIA E TECNOLOGIA DA AMAZÔNIA, 2., 2001, Santarém. Livro de Resumos–II Salão de Iniciação Científica do ILES Santarém. Santarém: ILES, 2001. p.109.

ROVARIS NETO, E.; NASSAR, S.M. **Application of data mining with resources bayesian networks in the analysis of the evasion of information systems students from ILES Santarém**. In: CONGRESSO DE CIÊNCIA E TECNOLOGIA DA AMAZÔNIA, 2., 2001, Santarém. Livro de Resumos – II Salão de Iniciação Científica do ILES Santarém. Santarém: ILES, 2001. p. 110.

RUSSEL, S.; NORVIG, P. **Artificial intelligence: A modern approach**. New Jersey: Prentice-Hall, 1995.

SCHILDT, H. **Inteligência artificial: Utilizando linguagem C**. Tradução Cláudio Gaiger Silveira e Mônica Soares Rufino. São Paulo: McGraw-Hill, 1989. Título original: Artificial Intelligence using C.

SHAW, I. S.; SIMOES N. G. **Controle e Modelagem Fuzzy**. São Paulo: Edgard Blucher, 1999.

SIMON, H. A. **As ciências do artificial**. Tradução Luiz Moniz Pereira. Coimbra, Portugal: Arménio Amado, 1981. Título original: The sciences of the artificial.

STEIN, C. E. **Sistema Especialista Probabilístico: Base de Conhecimento Dinâmica**. Dissertação de Mestrado, Curso de Pós Graduação em Ciência da Computação, Universidade Federal de Santa Catarina, Florianópolis, 2000.

WEISS, S. M.; KULIKOWSKI, C. A. **Guia prático para projetar sistemas especialistas**. Tradução Carlos Octávio Pavel. Rio de Janeiro: LTC, 1988. Título original: A practical guide to designing expert systems.

WINSTON, P. H. **Inteligência Artificial**. Tradução Carlos Octávio Pavel. Rio de Janeiro: LTC, 1988. Título original: Artificial Intelligence.

APÊNDICES

APÊNDICE A

QUESTIONÁRIO APLICADO AOS ALUNOS DO CURSO DE INFORMÁTICA

PERFIL DO ESTUDANTE DE INFORMÁTICA/SISTEMAS DE INFORMAÇÃO

A. DADOS PESSOAIS	A. DADOS SOBRE FREQUÊNCIA AO CURSO DE SISTEMAS DE INFORMAÇÃO/INFORMÁTICA
<p>1. Você está matriculado ? <input type="checkbox"/> Sim <input type="checkbox"/> Não</p> <p>02. Semestre e Ano de ingresso no curso: _____/_____</p> <p>03. Idade no ano de ingresso no curso: <input type="checkbox"/> 16 a 17 anos <input type="checkbox"/> 18 a 19 anos <input type="checkbox"/> 20 a 21 anos <input type="checkbox"/> 22 a 23 anos <input type="checkbox"/> 24 a 25 anos <input type="checkbox"/> 26 a 30 anos <input type="checkbox"/> 31 a 39 anos <input type="checkbox"/> Acima de 40 anos</p> <p>04. Idade no ano da entrevista (out/nov. 2000) <input type="checkbox"/> menos de 20 anos <input type="checkbox"/> 20 a 25 anos <input type="checkbox"/> mais de 25 anos</p> <p>05. Sexo <input type="checkbox"/> Masculino <input type="checkbox"/> Feminino</p> <p>06. Estado Civil <input type="checkbox"/> Solteiro(a) <input type="checkbox"/> Casado(a)/vivendo juntos <input type="checkbox"/> Viúvo(a)/separado</p> <p>07. Cidade e UF em que nasceu (família reside) : _____</p> <p>08. Escolaridade ao ingressar no curso de Informática/Sistemas de Informação <input type="checkbox"/> finalista do Ensino Médio (2º grau) <input type="checkbox"/> com o 2º grau concluído há mais de 1 ano <input type="checkbox"/> com o 2º grau concluído há mais de 2 anos <input type="checkbox"/> com outro curso superior iniciado. Qual/Onde? _____ <input type="checkbox"/> com outro curso superior concluído. Qual/Onde? _____</p> <p>09. Onde realizou o 2º grau ? (pode marcar mais de uma resposta) <input type="checkbox"/> escola particular. Qual/Onde? _____ <input type="checkbox"/> escola pública. Qual/Onde? _____ <input type="checkbox"/> supletivo. Qual/Onde? _____</p> <p>10. Se iniciou outro curso superior: <input type="checkbox"/> Formou-se <input type="checkbox"/> Abandonou <input type="checkbox"/> Trancou <input type="checkbox"/> Cancelou Total <input type="checkbox"/> Transferiu para o atual <input type="checkbox"/> cursa simultaneamente. Qual/Onde? _____</p> <p>11. Fez cursinho pré-vestibular ? <input type="checkbox"/> Não <input type="checkbox"/> Sim. Por quanto tempo/Qual/Onde ? _____</p> <p>12. Quantas vezes prestou vestibular antes de ingressar no curso de Informática/SI <input type="checkbox"/> Só uma, a de ingresso no curso <input type="checkbox"/> duas vezes <input type="checkbox"/> três ou mais vezes</p> <p>13. Se mais de uma vez, para qual curso ? <input type="checkbox"/> Informática/SI <input type="checkbox"/> outros. Quais/Onde ? _____</p>	<p>14. Porque escolheu o Curso de Informática/Sistemas de Informação ? <input type="checkbox"/> poucos candidatos, facilitando o acesso <input type="checkbox"/> poucos candidatos, facilitando a transferência para outros cursos desta unidade <input type="checkbox"/> poucos candidatos, facilitando a transferência para outros cursos de outras unidades <input type="checkbox"/> falta de opção, pois eu preferia o seguinte curso: _____ <input type="checkbox"/> grande interesse nessa área, é minha vocação e primeira opção <input type="checkbox"/> para ampliar minha cultura geral, desenvolver-me intelectualmente <input type="checkbox"/> para aumentar as chances de conseguir um bom emprego <input type="checkbox"/> para complementar minha formação (já tenho um curso superior) <input type="checkbox"/> outros. Quais ? _____</p> <p>15. Como soube da existência do curso no ILES de Santarém ? <input type="checkbox"/> pelo manual do candidato <input type="checkbox"/> por professores do ILES de Santarém <input type="checkbox"/> pela imprensa (jornais e TV) <input type="checkbox"/> pela internet <input type="checkbox"/> por cursinhos pré-vestibulares. Qual ? _____ <input type="checkbox"/> por colegas, parentes ou amigos que freqüentam/ trabalham no ILES. <input type="checkbox"/> outras fontes. Quais ? _____</p> <p>16. Que informações você tinha sobre curso antes de nele ingressar ? Grade curricular e disciplinas do curso <input type="checkbox"/> Sim <input type="checkbox"/> Não Áreas de especialização e carreiras oferecidas <input type="checkbox"/> Sim <input type="checkbox"/> Não Tempo de dedicação exigido <input type="checkbox"/> Sim <input type="checkbox"/> Não Disponibilidade de Hw, Sw e Biblioteca <input type="checkbox"/> Sim <input type="checkbox"/> Não Exigência de monografia de TCC <input type="checkbox"/> Sim <input type="checkbox"/> Não Exigência de Estágio Curricular <input type="checkbox"/> Sim <input type="checkbox"/> Não Nível de desempenho (notas) exigidas <input type="checkbox"/> Sim <input type="checkbox"/> Não Possibilidades de obter bolsas de Monitoria <input type="checkbox"/> Sim <input type="checkbox"/> Não Possibilidades de obter bolsas de IC <input type="checkbox"/> Sim <input type="checkbox"/> Não Possibilidades de obter bolsas de Esportes <input type="checkbox"/> Sim <input type="checkbox"/> Não Possibilidades de obter bolsas do Coral <input type="checkbox"/> Sim <input type="checkbox"/> Não Mercado de Trabalho <input type="checkbox"/> Sim <input type="checkbox"/> Não Empresa Júnior <input type="checkbox"/> Sim <input type="checkbox"/> Não Incubadora de empresas com base tecnológica <input type="checkbox"/> Sim <input type="checkbox"/> Não <i>ECOnet</i> <input type="checkbox"/> Sim <input type="checkbox"/> Não</p> <p>17. E HOJE ? grade curricular e disciplinas do curso <input type="checkbox"/> Sim <input type="checkbox"/> Não Áreas de especialização e carreiras oferecidas <input type="checkbox"/> Sim <input type="checkbox"/> Não Tempo de dedicação exigido <input type="checkbox"/> Sim <input type="checkbox"/> Não Disponibilidade de Hw, Sw e Biblioteca <input type="checkbox"/> Sim <input type="checkbox"/> Não Exigência de monografia de TCC <input type="checkbox"/> Sim <input type="checkbox"/> Não Exigência de Estágio Curricular <input type="checkbox"/> Sim <input type="checkbox"/> Não Nível de desempenho (notas) exigidas <input type="checkbox"/> Sim <input type="checkbox"/> Não Possibilidades de obter bolsas de Monitoria <input type="checkbox"/> Sim <input type="checkbox"/> Não Possibilidades de obter bolsas de IC <input type="checkbox"/> Sim <input type="checkbox"/> Não Possibilidades de obter bolsas de Esportes <input type="checkbox"/> Sim <input type="checkbox"/> Não Possibilidades de obter bolsas do Coral <input type="checkbox"/> Sim <input type="checkbox"/> Não Mercado de Trabalho <input type="checkbox"/> Sim <input type="checkbox"/> Não Empresa Júnior <input type="checkbox"/> Sim <input type="checkbox"/> Não Incubadora de empresas com base tecnológica <input type="checkbox"/> Sim <input type="checkbox"/> Não <i>ECOnet</i> <input type="checkbox"/> Sim <input type="checkbox"/> Não</p> <p>18. Você pretende continuar neste curso, sua intenção é: <input type="checkbox"/> não pretende continuar o curso <input type="checkbox"/> pretende trocar de curso nesta unidade. Qual curso ? _____ <input type="checkbox"/> pretende trocar de curso em outra unidade, nesta cidade. Qual curso? _____ <input type="checkbox"/> pretende trocar de curso em outra unidade e outra cidade. Qual curso/cidade ? _____ <input type="checkbox"/> obter o bacharelado e profissionalizar-se <input type="checkbox"/> obter o bacharelado e cursar uma pós graduação local <input type="checkbox"/> obter o bacharelado e cursar uma pós graduação em outro local</p> <p>19. Pretende especializar-se após a graduação ? <input type="checkbox"/> Não tenho idéia. <input type="checkbox"/> Sim, na área de: _____ <input type="checkbox"/> Não, Por que? _____</p> <p>20. Interrompeu o curso alguma vez? <input type="checkbox"/> Sim Quantas vezes? _____ <input type="checkbox"/> Não</p>

<p>21. Se interrompeu o curso. Motivos ?</p> <p><input type="checkbox"/> desilusão com o curso escolhido</p> <p><input type="checkbox"/> necessidade de trabalhar</p> <p><input type="checkbox"/> doença</p> <p><input type="checkbox"/> problemas pessoais ou familiares</p> <p><input type="checkbox"/> dificuldade de acompanhamento do curso</p> <p><input type="checkbox"/> outros. Quais? _____</p> <p>22. Foi reprovado em alguma disciplina durante o curso ?</p> <p><input type="checkbox"/> Sim Quantas vezes? _____</p> <p><input type="checkbox"/> Não</p>	<p>30. Você conhece e utiliza a biblioteca?</p> <p><input type="checkbox"/> conhece mas não utiliza. Por que ? _____</p> <p><input type="checkbox"/> conhece e utiliza, só quando o professor passa algum tipo de trabalho.</p> <p><input type="checkbox"/> conhece e utiliza, independentemente do professor passar algum tipo de trabalho.</p> <p><input type="checkbox"/> não conhece. Por que ? _____</p> <p>32. Você conhece e utiliza os PERIÓDICOS da biblioteca?</p> <p><input type="checkbox"/> não conhece. Por que ? _____</p> <p><input type="checkbox"/> conhece mas não utiliza. Por que ? _____</p> <p><input type="checkbox"/> conhece e utiliza, só quando o professor passa algum tipo de trabalho.</p> <p><input type="checkbox"/> conhece e utiliza, independentemente do professor passar algum tipo de trabalho. Neste caso: <input type="checkbox"/> Acha insuficiente. <input type="checkbox"/> Acha suficiente</p>																		
<p>23. Se foi, quais os motivos (pode assinalar mais de uma alternativa)</p> <p><input type="checkbox"/> acúmulo de disciplinas</p> <p><input type="checkbox"/> dificuldade com o conteúdo</p> <p><input type="checkbox"/> dificuldade com o professor</p> <p><input type="checkbox"/> volume de leituras a cumprir</p> <p><input type="checkbox"/> volume de projetos ou programas a cumprir</p> <p><input type="checkbox"/> falta de empenho, motivação ou interesse pessoal</p> <p><input type="checkbox"/> falta de empenho, motivação ou interesse do professor</p> <p><input type="checkbox"/> didática/conhecimentos do(a) professor(a)</p> <p><input type="checkbox"/> problemas de saúde ou familiares</p> <p><input type="checkbox"/> outros. Quais? _____</p>	<p>33. Você conhece e utiliza os laboratórios de Informática ?</p> <p><input type="checkbox"/> conhece mas não utiliza. Porque tem um computador em casa.</p> <p><input type="checkbox"/> conhece mas não utiliza. Porque tem um computador no escritório que pode utilizar para a realização dos trabalhos escolares.</p> <p><input type="checkbox"/> conhece mas não utiliza. Porque trabalha no horário mas tem um computador em casa.</p> <p><input type="checkbox"/> conhece mas não utiliza. Porque trabalha no horário mas tem NÃO um computador no escritório que possa utilizar para a realização dos trabalhos escolares.</p> <p><input type="checkbox"/> conhece e utiliza, só quando o professor passa algum tipo de trabalho.</p> <p><input type="checkbox"/> conhece e utiliza, independentemente do professor passar algum tipo de trabalho.</p> <p><input type="checkbox"/> não conhece. Por que ? _____</p>																		
<p>24. Quais os fatores que ajudaram o bom andamento de seus estudos no curso? (pode assinalar mais de uma alternativa)</p> <p><input type="checkbox"/> Ter escolhido um curso com o qual me identifiquei</p> <p><input type="checkbox"/> Coordenação de Curso</p> <p><input type="checkbox"/> Professores e funcionários</p> <p><input type="checkbox"/> Monitoria</p> <p><input type="checkbox"/> Amigos</p> <p><input type="checkbox"/> Laboratórios abertos pelos períodos da manhã e da tarde</p> <p><input type="checkbox"/> Biblioteca</p> <p><input type="checkbox"/> Estágio</p> <p><input type="checkbox"/> Cursos de Extensão</p> <p><input type="checkbox"/> Semanas Acadêmicas</p> <p><input type="checkbox"/> Jornada Acadêmicas</p> <p><input type="checkbox"/> Iniciação Científica</p> <p><input type="checkbox"/> outros. Quais? _____</p>	<p>34. Você participa das atividades promovidas pelo departamento ou instituição ?</p> <p><input type="checkbox"/> Sim</p> <p><input type="checkbox"/> Algumas. Quais ? _____</p> <p><input type="checkbox"/> Não. Por que ? _____</p> <p>35. Como você costuma organizar-se para estudar ? (pode indicar mais de uma alternativa)</p> <p><input type="checkbox"/> uso minhas anotações de aulas</p> <p><input type="checkbox"/> uso apostilas</p> <p><input type="checkbox"/> faço leituras recomendadas, em livros e revistas</p> <p><input type="checkbox"/> faço leituras em línguas estrangeiras</p> <p><input type="checkbox"/> consulto a internet</p> <p><input type="checkbox"/> consulto a biblioteca</p> <p><input type="checkbox"/> consulto colegas</p> <p><input type="checkbox"/> consulto professores</p> <p><input type="checkbox"/> consulto a monitoria</p> <p><input type="checkbox"/> estudo com amigos de classe</p> <p><input type="checkbox"/> estudo com grupos de estudo</p> <p><input type="checkbox"/> pago aulas particulares</p> <p><input type="checkbox"/> pago para fazerem meus trabalhos</p> <p><input type="checkbox"/> pago ou peço para entrar em uma equipe só com o meu nome (não enfrento obstáculos)</p> <p><input type="checkbox"/> outras : _____</p>																		
<p>25. Quais os fatores que prejudicaram, até o presente, o andamento de seus estudos no curso? (pode assinalar mais de uma alternativa)</p> <p><input type="checkbox"/> Ter escolhido um curso com o qual me identifiquei</p> <p><input type="checkbox"/> Coordenação de Curso</p> <p><input type="checkbox"/> Professores e funcionários</p> <p><input type="checkbox"/> Monitoria</p> <p><input type="checkbox"/> Amigos</p> <p><input type="checkbox"/> Laboratórios abertos pelos períodos da manhã e da tarde</p> <p><input type="checkbox"/> Biblioteca</p> <p><input type="checkbox"/> Estágio</p> <p><input type="checkbox"/> Cursos de Extensão</p> <p><input type="checkbox"/> Semanas Acadêmicas</p> <p><input type="checkbox"/> Jornada Acadêmicas</p> <p><input type="checkbox"/> Iniciação Científica</p> <p><input type="checkbox"/> outros. Quais? _____</p>	<p>36. Sobre o sistema de ensino adotado no curso, qual você prefere ? (enumere por ordem crescente de preferência)</p> <table border="0"> <thead> <tr> <th></th> <th style="text-align: right;">Sente falta?</th> </tr> </thead> <tbody> <tr> <td><input type="checkbox"/> aulas expositivas</td> <td style="text-align: right;"><input type="checkbox"/> S <input type="checkbox"/> N</td> </tr> <tr> <td><input type="checkbox"/> seminários, com apresentações pelos alunos</td> <td style="text-align: right;"><input type="checkbox"/> S <input type="checkbox"/> N</td> </tr> <tr> <td><input type="checkbox"/> seminários, com discussão de textos pela turma</td> <td style="text-align: right;"><input type="checkbox"/> S <input type="checkbox"/> N</td> </tr> <tr> <td><input type="checkbox"/> aulas ou palestras de convidados</td> <td style="text-align: right;"><input type="checkbox"/> S <input type="checkbox"/> N</td> </tr> <tr> <td><input type="checkbox"/> aulas práticas (pesquisa de campo, etc.)</td> <td style="text-align: right;"><input type="checkbox"/> S <input type="checkbox"/> N</td> </tr> <tr> <td><input type="checkbox"/> atividades extra-classe</td> <td style="text-align: right;"><input type="checkbox"/> S <input type="checkbox"/> N</td> </tr> <tr> <td><input type="checkbox"/> Aulas mistas (projutor, TV, laboratório, Retroprojutor)</td> <td style="text-align: right;"><input type="checkbox"/> S <input type="checkbox"/> N</td> </tr> <tr> <td><input type="checkbox"/> Aulas práticas, com aplicativos existentes no</td> <td style="text-align: right;"><input type="checkbox"/> S <input type="checkbox"/> N</td> </tr> </tbody> </table>		Sente falta?	<input type="checkbox"/> aulas expositivas	<input type="checkbox"/> S <input type="checkbox"/> N	<input type="checkbox"/> seminários, com apresentações pelos alunos	<input type="checkbox"/> S <input type="checkbox"/> N	<input type="checkbox"/> seminários, com discussão de textos pela turma	<input type="checkbox"/> S <input type="checkbox"/> N	<input type="checkbox"/> aulas ou palestras de convidados	<input type="checkbox"/> S <input type="checkbox"/> N	<input type="checkbox"/> aulas práticas (pesquisa de campo, etc.)	<input type="checkbox"/> S <input type="checkbox"/> N	<input type="checkbox"/> atividades extra-classe	<input type="checkbox"/> S <input type="checkbox"/> N	<input type="checkbox"/> Aulas mistas (projutor, TV, laboratório, Retroprojutor)	<input type="checkbox"/> S <input type="checkbox"/> N	<input type="checkbox"/> Aulas práticas, com aplicativos existentes no	<input type="checkbox"/> S <input type="checkbox"/> N
	Sente falta?																		
<input type="checkbox"/> aulas expositivas	<input type="checkbox"/> S <input type="checkbox"/> N																		
<input type="checkbox"/> seminários, com apresentações pelos alunos	<input type="checkbox"/> S <input type="checkbox"/> N																		
<input type="checkbox"/> seminários, com discussão de textos pela turma	<input type="checkbox"/> S <input type="checkbox"/> N																		
<input type="checkbox"/> aulas ou palestras de convidados	<input type="checkbox"/> S <input type="checkbox"/> N																		
<input type="checkbox"/> aulas práticas (pesquisa de campo, etc.)	<input type="checkbox"/> S <input type="checkbox"/> N																		
<input type="checkbox"/> atividades extra-classe	<input type="checkbox"/> S <input type="checkbox"/> N																		
<input type="checkbox"/> Aulas mistas (projutor, TV, laboratório, Retroprojutor)	<input type="checkbox"/> S <input type="checkbox"/> N																		
<input type="checkbox"/> Aulas práticas, com aplicativos existentes no	<input type="checkbox"/> S <input type="checkbox"/> N																		
<p>26. Você é ou foi um aluno bolsista?</p> <p><input type="checkbox"/> Sim, fui.</p> <p><input type="checkbox"/> Sim, sou.</p> <p><input type="checkbox"/> Não</p> <p>27. Se é ou foi bolsista. Que de bolsa tem ou teve ?</p> <p>_____</p>																			
<p>28. Se sim, de que tipo? (pode indicar mais de uma resposta)</p> <p><input type="checkbox"/> pesquisa bibliográfica para trabalho de curso</p> <p><input type="checkbox"/> projeto de pesquisa para a disciplina TCC ou Estágio</p> <p><input type="checkbox"/> pesquisa de campo para alguma disciplina ou ICC</p> <p><input type="checkbox"/> pesquisa por conta própria</p> <p><input type="checkbox"/> monitoria</p> <p><input type="checkbox"/> esportes</p> <p><input type="checkbox"/> coral</p> <p><input type="checkbox"/> outra. Qual: _____</p>																			
<p>29. Você conhece e utiliza a monitoria ?</p> <p><input type="checkbox"/> conhece mas não utiliza. Por que ? _____</p> <p><input type="checkbox"/> conhece e utiliza, principalmente em _____</p> <p><input type="checkbox"/> não conhece. Por que ? _____</p>																			

37. Grau de satisfação/insatisfação com diferentes aspectos do curso:

Aspectos	Excel	Bom	Regular	Insuficiente
Disponibilidade de livros e revistas na biblioteca (7.0000)	()	()	()	()
Cinco laboratórios (60 maquinas e 48 com internet)	()	()	()	()
Dedicação dos professores ao curso	()	()	()	()
Qualificação dos professores ao curso	()	()	()	()
Organização do currículo	()	()	()	()
Tamanhos das turmas	()	()	()	()
Duração das aulas	()	()	()	()
Disponibilidade dos professores para os alunos	()	()	()	()
Assiduidade dos professores	()	()	()	()
Clima de motivação no curso	()	()	()	()
Sistemas de avaliação	()	()	()	()
Oportunidade de alunos avaliarem	()	()	()	()
Oferta e variedade de disciplinas	()	()	()	()
Oferta de bolsas de iniciação científica	()	()	()	()
Oferta de bolsas de monitoria	()	()	()	()
Oferta de bolsas no esporte	()	()	()	()
Oferta de Estágios	()	()	()	()
Idéia da Empresa Júnior	()	()	()	()
Estímulo à pesquisa	()	()	()	()

38. Se você utiliza ou esta subordinado a um desses serviços ou coordenação, expresse a sua opinião.

Aspectos	Excel	Bom	Regular	Insuf	S/R
No ILES					
Direção Geral (Prof. Edgar)	()	()	()	()	()
Direção Acadêmica (Profa. Conceição)	()	()	()	()	()
Secretaria Geral (Sebastiana)	()	()	()	()	()
Biblioteca (Zélia)	()	()	()	()	()
Tesousaria (Dulfe)	()	()	()	()	()
No Departamento de Informática					
Coordenação de Curso (Prof. Eugênio)	()	()	()	()	()
Coordenação de Laboratórios (Prof. Ivo)	()	()	()	()	()
Coordenação de TCC (Prof. Jorge Ricardo)	()	()	()	()	()
Coordenação de Monitoria (Profa. Giovanna)	()	()	()	()	()
Coordenação de Extensão (Profa. Aletéia)	()	()	()	()	()
Coordenação de Pesquisa e ICC (Prof. Jorge Ricardo)	()	()	()	()	()
Coordenação de Estágio (a ser definido)	()	()	()	()	()
Coordenação da Empresa Júnior (Prof. Carlos)	()	()	()	()	()
Geral					
Xerox (nova administração)	()	()	()	()	()
Cantina (Zé Cangóia)	()	()	()	()	()
Cantina (Prédio novo)	()	()	()	()	()
Caixa automático (Banco do Brasil)	()	()	()	()	()
Estacionamento (Direção Geral)	()	()	()	()	()
Climatização nas salas de aula e nos laboratórios	()	()	()	()	()
Iluminação nas salas de aula e nos laboratórios	()	()	()	()	()

39. Como você se definiria em relação à vida universitária ?

- () interessado, mas não participa
 () muito interessado e participa sempre
 () pouco interessado, participa algumas, mas poucas vezes
 () nem um pouco interessado, tem muito mais que fazer
 () Sem opinião formada

40. Como você vê hoje curso em relação ao que você imaginava antes de nele ingressar? (marcar todas)

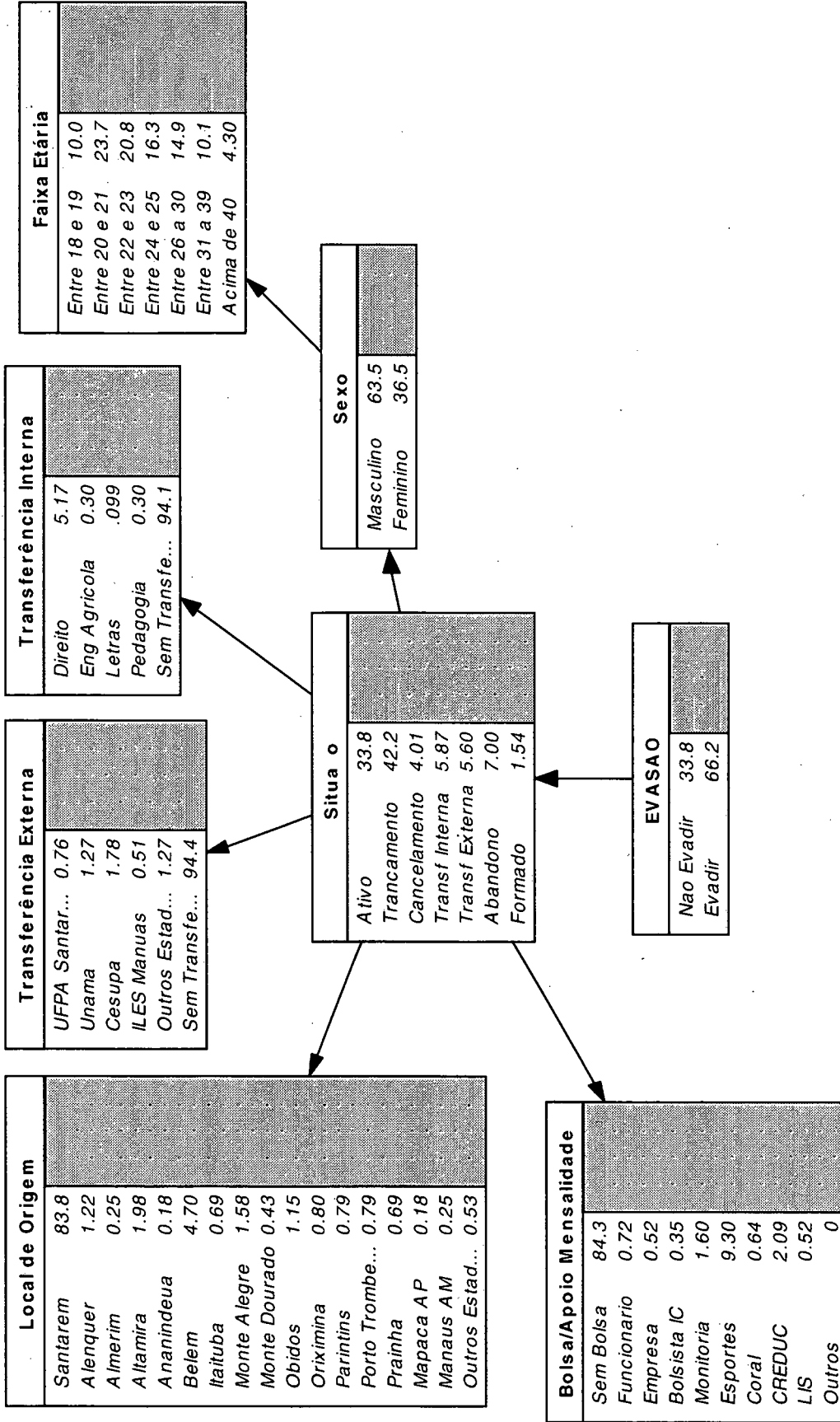
	SIM	NÃO
01. tem sido uma ótima experiência, supera em muito minhas expectativas	()	()
. a carga de estudos é muito maior do que esperava	()	()
. as pessoas ficam muito soltas, sem orientação	()	()
. o curso é muito teórico, sem ligação com a realidade	()	()
. os conteúdos das disciplinas não correspondem ao que eu esperava	()	()
. o curso é muito superficial e não proporciona os conhecimentos que serão necessários na vida profissional	()	()
. a qualidade das matérias e dos professores é pior do que eu esperava	()	()
. ainda não tem opinião formada	()	()
. o curso reprova pouco, não exigindo assim excelência acadêmica	()	()

41. O que mudou em suas expectativas desde que você entrou no seu curso ?

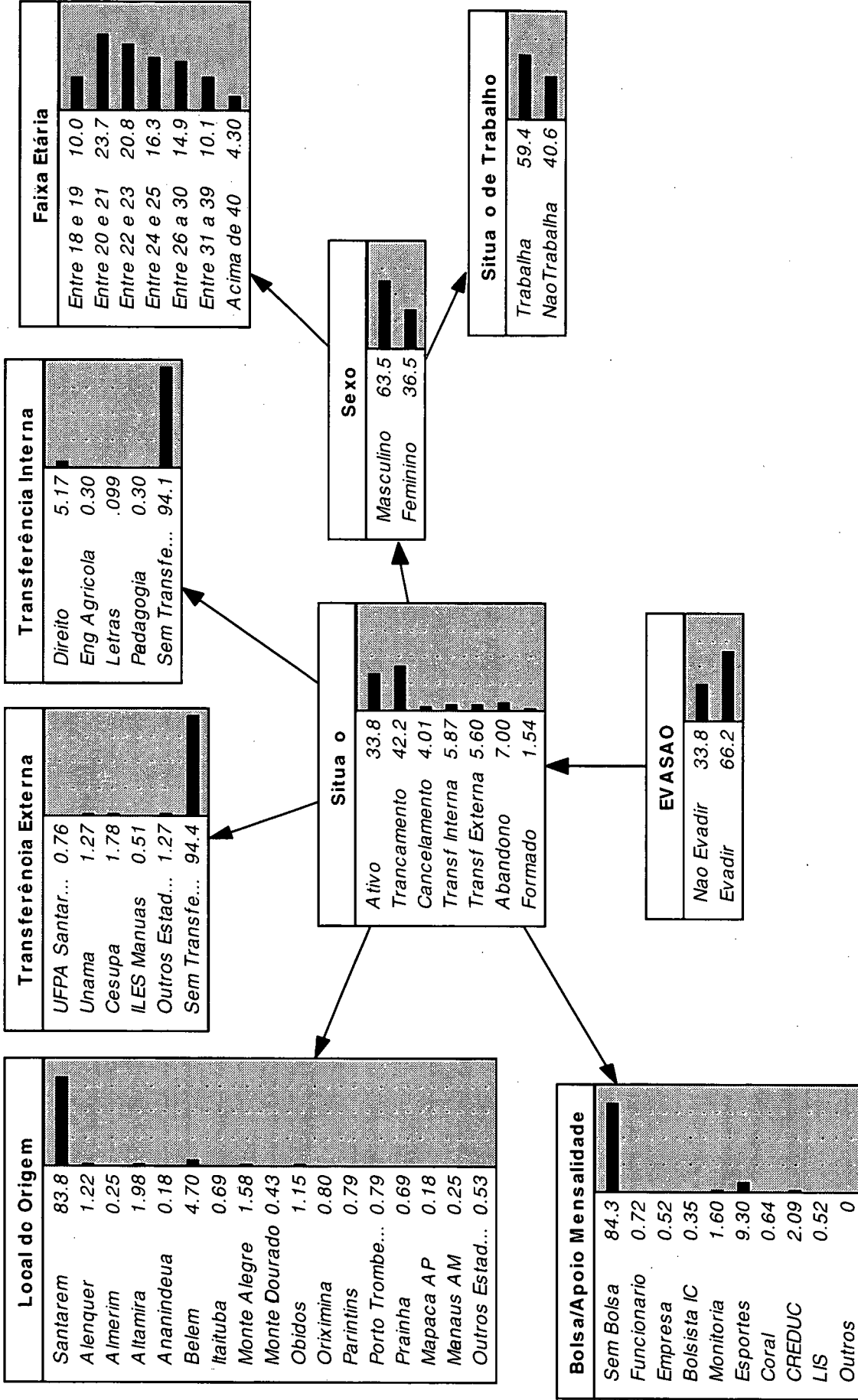
- A. Quanto à formação: _____
- B. Quanto à profissionalização: _____
- C. Quanto ao Mercado de Trabalho: _____

SITUAÇÃO DE TRABALHO	SITUAÇÃO FAMILIAR (ABA - ABIPEME)																																																																																																							
<p>42. Qual a sua situação de trabalho ?</p> <p><input type="checkbox"/> nunca trabalhou</p> <p><input type="checkbox"/> não trabalha atualmente</p> <p><input type="checkbox"/> trabalha em tempo parcial</p> <p><input type="checkbox"/> trabalha eventualmente (bicos, serviços prestados, etc.)</p> <p><input type="checkbox"/> trabalha em tempo integral, inclusive aos sábados.</p> <p>43. Trabalhando ou não, com que recursos você se sustenta?</p> <p><input type="checkbox"/> recursos próprios</p> <p><input type="checkbox"/> empresa</p> <p><input type="checkbox"/> dos pais/família</p> <p><input type="checkbox"/> do conjuge</p> <p><input type="checkbox"/> bolsa de estudo/estágio</p> <p><input type="checkbox"/> Crédito Educativo (CREDOC/LIS)</p> <p><input type="checkbox"/> outros _____</p> <p>44. Se trabalha, qual a sua atual ocupação ?</p> <p><input type="checkbox"/> setor público</p> <p><input type="checkbox"/> setor privado</p> <p><input type="checkbox"/> autônomo</p> <p><input type="checkbox"/> militar</p> <p><input type="checkbox"/> aluno bolsista</p> <p><input type="checkbox"/> outras: _____</p> <p>45. Sua entrada neste Instituto, no Curso de Informática/Sistemas de Informação, já alterou sua situação de trabalho ?</p> <p><input type="checkbox"/> Sim De que maneira: _____</p> <p><input type="checkbox"/> Não</p>	<p>47. Escolaridade dos pais</p> <table border="0"> <tr> <td></td> <td>PAI</td> <td>MÃE</td> </tr> <tr> <td>1. analfabeto/primário incompleto</td> <td><input type="checkbox"/></td> <td><input type="checkbox"/></td> </tr> <tr> <td>2. primário completo/ginasial incompleto</td> <td><input type="checkbox"/></td> <td><input type="checkbox"/></td> </tr> <tr> <td>3. ginasial completo/colegial incompleto</td> <td><input type="checkbox"/></td> <td><input type="checkbox"/></td> </tr> <tr> <td>4. colegial completo/superior incompleto</td> <td><input type="checkbox"/></td> <td><input type="checkbox"/></td> </tr> <tr> <td>5. superior completo/pós incompleta</td> <td><input type="checkbox"/></td> <td><input type="checkbox"/></td> </tr> <tr> <td>6. pós-graduação completa</td> <td><input type="checkbox"/></td> <td><input type="checkbox"/></td> </tr> </table> <p>48. Ocupação dos Pais (a ultima, caso sejam falecidos/aposentados ou inválidos)</p> <table border="0"> <tr> <td></td> <td>PAI</td> <td>MÃE</td> </tr> <tr> <td>1. proprietário de pequenos negócios</td> <td><input type="checkbox"/></td> <td><input type="checkbox"/></td> </tr> <tr> <td>2. profissional liberal, professor universitário</td> <td><input type="checkbox"/></td> <td><input type="checkbox"/></td> </tr> <tr> <td>3. técnico, bancário, prof. de 1º ou 2º grau, func. público ou privado não graduado</td> <td><input type="checkbox"/></td> <td><input type="checkbox"/></td> </tr> <tr> <td>4. operário, trabalho não qualificado</td> <td><input type="checkbox"/></td> <td><input type="checkbox"/></td> </tr> <tr> <td>5. não sabe</td> <td><input type="checkbox"/></td> <td><input type="checkbox"/></td> </tr> </table> <p>49. Número de dependentes do chefe da família</p> <p><input type="checkbox"/> sem dependentes</p> <p><input type="checkbox"/> um dependente</p> <p><input type="checkbox"/> dois dependentes</p> <p><input type="checkbox"/> três dependentes</p> <p><input type="checkbox"/> quatro dependentes</p> <p><input type="checkbox"/> cinco dependentes</p> <p><input type="checkbox"/> seis ou mais dependentes</p> <p>50. Itens de conforto</p> <table border="0"> <tr> <td></td> <td>0</td> <td>1</td> <td>2</td> <td>3</td> <td>4</td> <td>5</td> <td>6 ou +</td> </tr> <tr> <td>TV</td> <td><input type="checkbox"/></td> <td><input type="checkbox"/></td> <td><input type="checkbox"/></td> <td><input type="checkbox"/></td> <td><input type="checkbox"/></td> <td><input type="checkbox"/></td> <td><input type="checkbox"/></td> </tr> <tr> <td>rádio</td> <td><input type="checkbox"/></td> <td><input type="checkbox"/></td> <td><input type="checkbox"/></td> <td><input type="checkbox"/></td> <td><input type="checkbox"/></td> <td><input type="checkbox"/></td> <td><input type="checkbox"/></td> </tr> <tr> <td>banheiro</td> <td><input type="checkbox"/></td> <td><input type="checkbox"/></td> <td><input type="checkbox"/></td> <td><input type="checkbox"/></td> <td><input type="checkbox"/></td> <td><input type="checkbox"/></td> <td><input type="checkbox"/></td> </tr> <tr> <td>empregada fixa</td> <td><input type="checkbox"/></td> <td><input type="checkbox"/></td> <td><input type="checkbox"/></td> <td><input type="checkbox"/></td> <td><input type="checkbox"/></td> <td><input type="checkbox"/></td> <td><input type="checkbox"/></td> </tr> <tr> <td>computador</td> <td><input type="checkbox"/></td> <td><input type="checkbox"/></td> <td><input type="checkbox"/></td> <td><input type="checkbox"/></td> <td><input type="checkbox"/></td> <td><input type="checkbox"/></td> <td><input type="checkbox"/></td> </tr> <tr> <td>maquina de lavar roupa</td> <td><input type="checkbox"/></td> <td><input type="checkbox"/></td> <td><input type="checkbox"/></td> <td><input type="checkbox"/></td> <td><input type="checkbox"/></td> <td><input type="checkbox"/></td> <td><input type="checkbox"/></td> </tr> <tr> <td>automóvel</td> <td><input type="checkbox"/></td> <td><input type="checkbox"/></td> <td><input type="checkbox"/></td> <td><input type="checkbox"/></td> <td><input type="checkbox"/></td> <td><input type="checkbox"/></td> <td><input type="checkbox"/></td> </tr> </table>		PAI	MÃE	1. analfabeto/primário incompleto	<input type="checkbox"/>	<input type="checkbox"/>	2. primário completo/ginasial incompleto	<input type="checkbox"/>	<input type="checkbox"/>	3. ginasial completo/colegial incompleto	<input type="checkbox"/>	<input type="checkbox"/>	4. colegial completo/superior incompleto	<input type="checkbox"/>	<input type="checkbox"/>	5. superior completo/pós incompleta	<input type="checkbox"/>	<input type="checkbox"/>	6. pós-graduação completa	<input type="checkbox"/>	<input type="checkbox"/>		PAI	MÃE	1. proprietário de pequenos negócios	<input type="checkbox"/>	<input type="checkbox"/>	2. profissional liberal, professor universitário	<input type="checkbox"/>	<input type="checkbox"/>	3. técnico, bancário, prof. de 1º ou 2º grau, func. público ou privado não graduado	<input type="checkbox"/>	<input type="checkbox"/>	4. operário, trabalho não qualificado	<input type="checkbox"/>	<input type="checkbox"/>	5. não sabe	<input type="checkbox"/>	<input type="checkbox"/>		0	1	2	3	4	5	6 ou +	TV	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	rádio	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	banheiro	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	empregada fixa	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	computador	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	maquina de lavar roupa	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	automóvel	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
	PAI	MÃE																																																																																																						
1. analfabeto/primário incompleto	<input type="checkbox"/>	<input type="checkbox"/>																																																																																																						
2. primário completo/ginasial incompleto	<input type="checkbox"/>	<input type="checkbox"/>																																																																																																						
3. ginasial completo/colegial incompleto	<input type="checkbox"/>	<input type="checkbox"/>																																																																																																						
4. colegial completo/superior incompleto	<input type="checkbox"/>	<input type="checkbox"/>																																																																																																						
5. superior completo/pós incompleta	<input type="checkbox"/>	<input type="checkbox"/>																																																																																																						
6. pós-graduação completa	<input type="checkbox"/>	<input type="checkbox"/>																																																																																																						
	PAI	MÃE																																																																																																						
1. proprietário de pequenos negócios	<input type="checkbox"/>	<input type="checkbox"/>																																																																																																						
2. profissional liberal, professor universitário	<input type="checkbox"/>	<input type="checkbox"/>																																																																																																						
3. técnico, bancário, prof. de 1º ou 2º grau, func. público ou privado não graduado	<input type="checkbox"/>	<input type="checkbox"/>																																																																																																						
4. operário, trabalho não qualificado	<input type="checkbox"/>	<input type="checkbox"/>																																																																																																						
5. não sabe	<input type="checkbox"/>	<input type="checkbox"/>																																																																																																						
	0	1	2	3	4	5	6 ou +																																																																																																	
TV	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>																																																																																																	
rádio	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>																																																																																																	
banheiro	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>																																																																																																	
empregada fixa	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>																																																																																																	
computador	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>																																																																																																	
maquina de lavar roupa	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>																																																																																																	
automóvel	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>																																																																																																	
<p style="text-align: center;">SITUAÇÃO DE MORADIA</p> <p>46. Residência</p> <p><input type="checkbox"/> casa dos pais</p> <p><input type="checkbox"/> casa de parentes ou amigos da família/cedida</p> <p><input type="checkbox"/> república / casa de estudantes, alojamento/ quarto/ aluguel</p> <p><input type="checkbox"/> residência própria (só com cônjuge e/ou filhos)</p>																																																																																																								
<p style="text-align: center;">EXAME NACIONAL DE CURSO - ENC</p> <p>53. Você é a favor do Exame Nacional de Curso?</p> <p><input type="checkbox"/> Sim</p> <p><input type="checkbox"/> Não. Por que ? _____</p> <p>54. Se sim, você é a favor que a sua nota individual saia no seu diploma ?</p> <p><input type="checkbox"/> Sim <input type="checkbox"/> Não</p> <p>55. Se sim, você é a favor que a nota do seu curso (media geral dos seus colegas) saia no seu diploma ?</p> <p><input type="checkbox"/> Sim <input type="checkbox"/> Não</p> <p>56. Se sim, você é a favor de simulados para a preparação do mesmo ?</p> <p><input type="checkbox"/> Sim <input type="checkbox"/> Não</p> <p>57. Se sim, você é a favor de aulas de reforço, na busca de um melhor desempenho ?</p> <p><input type="checkbox"/> Sim <input type="checkbox"/> Não</p> <p>58. Você tem medo de representar a nossa faculdade num provão, se considera bem preparado, dentro do que o instituto lhe oferece e não do que você, por outros motivos, efetivamente aproveita.</p> <p><input type="checkbox"/> Sim. Por que ? _____</p> <p><input type="checkbox"/> Não</p>	<p>51. Renda Mensal familiar</p> <p><input type="checkbox"/> até 2 salários mínimos</p> <p><input type="checkbox"/> até 3 salários mínimos</p> <p><input type="checkbox"/> até 5 salários mínimos</p> <p><input type="checkbox"/> até 7 salários mínimos</p> <p><input type="checkbox"/> até 8 salários mínimos</p> <p><input type="checkbox"/> até 10 salários mínimos</p> <p><input type="checkbox"/> até 15 salários mínimos</p> <p><input type="checkbox"/> até 20 salários mínimos</p> <p><input type="checkbox"/> acima de 20 salários mínimos</p> <p>52. dependente financeiramente da família ou de outras pessoas para cursar a faculdade ?</p> <p><input type="checkbox"/> não, tem renda própria (trabalha ou tem outras rendas pessoais)</p> <p><input type="checkbox"/> sim, parcialmente</p> <p><input type="checkbox"/> sim, totalmente</p>																																																																																																							
<p style="text-align: center;">SATISFAÇÃO</p> <p>59. Como esta o seu grau de satisfação com o nosso curso ?</p> <p><input type="checkbox"/> muito insatisfeito</p> <p><input type="checkbox"/> insatisfeito</p> <p><input type="checkbox"/> satisfeito</p> <p><input type="checkbox"/> muito satisfeito</p> <p>60. Você pretende fazer a rematricula aqui, no próximo semestre ?</p> <p><input type="checkbox"/> Sim <input type="checkbox"/> Não</p> <p>61. Se não. Qual a sua MAIOR dificuldade ?</p> <p><input type="checkbox"/> Saudades da família. Local de origem: _____</p> <p><input type="checkbox"/> Financeiro</p> <p><input type="checkbox"/> Vocacional</p> <p><input type="checkbox"/> Outros. Quais ? _____</p>	<p style="text-align: center;">OUTRAS OBSERVAÇÕES, SUGESTOES QUE DESEJAM APRESENTAR</p>																																																																																																							

APÊNDICE B
REDE BAYESIANA QUALITATIVA (*DATA MINING*)
(DADOS HIPOTÉTICOS)



APÊNDICE C
REDE BAYESIANA QUALITATIVA PARA O PROTÓTIPO
(DADOS HIPOTÉTICOS)



APÊNDICE D
REDE BAYESIANA QUANTITATIVA PARA O PROTÓTIPO
(DADOS HIPOTÉTICOS)

NETICA (GB) 1.12 WIN, (C) 1990-1998 NORSYS SOFTWARE CORP.**NETICA OPERATING WITHOUT A PASSWORD; THERE ARE SOME LIMITATIONS.**

COMMAND LINE: "C:\EVASAO\EVASAOALUNOS.DNE"

COMPILED TO 8 CLIQUES, WITH TOTAL STATE SIZE (INCLUDING SEPSETS) OF 351.

PROBABILITY OF NEW EVIDENCE = 100 %, OF ALL EVIDENCE = 100 %.

REPORT NETWORK-----
EVASAOALUNOSDB_ULBRA LAST MODIFIED 16/JAN/2002 20:42 H

9 NODES (NOT INCLUDING CONSTANTS)
 0 DECISION NODES
 0 UTILITY NODES
 0 CONSTANT NODES

 8 LINKS (NOT DISCONNECTED)
 0 DISCONNECTED LINKS
 0 TIME DELAY LINKS
 0 DIRECTED CYCLES (WITHOUT DELAYS)
 0 LOOPS (DISREGARDING LINK DIRECTIONS)
 1 SEPARATE NETWORKS (IGNORING CONSTANT NODES)

 314 CONDITIONAL PROBABILITIES TOTAL
 0 DECISION CONDITIONS

CASE NAME: <NONE>

0 FINDINGS NODES (NOT INCLUDING CONSTANTS)
 0 NEGATIVE OR LIKELIHOOD FINDINGS NODES

REPORT JUNCTON TREE

 CLIQUE [JOINED TO] SIZE MEMBER NODES (* MEANS HOME)
 0 [1 2 5 6 7] 014 (*EVASÃO, *SITUACAO)
 1 [0] 119 (SITUACAO, *LOCALORIGEM)
 2 [0 3 4] 014 (*SEXO, SITUACAO)
 3 [2] 004 (*TRABALHO, SEXO)
 4 [2] 014 (*FAIXA_ETARIA, SEXO)
 5 [0] 070 (*APOIO, SITUACAO)
 6 [0] 042 (*TRANSFERENCIAEXTERNA, SITUACAO)
 7 [0] 035 (*TRANSFERENCIAINTERNA, SITUACAO)

SUM OF CLIQUE SIZES = 312(WITH SEPSETS = 351)
-----**REPORT ELIMINATION ORDER**

 TRABALHO
 EVASÃO
 FAIXA_ETARIA
 SEXO
 TRANSFERENCIAINTERNA
 TRANSFERENCIAEXTERNA
 APOIO
 SITUACAO
 LOCALORIGEM

REPORT BELIEF**TRANSFERÊNCIA EXTERNA**

UFPA_SANTAREM	0.0076355
UNAMA	0.0127330
CESUPA	0.0178250
ILES_MANUAS	0.0050922
OUTROS_ESTADOS	0.0127330
SEM_TRANSFERENCIA	0.9439800

LOCAL DE ORIGEM

SANTAREM	0.8376400
ALENQUER	0.0122240
ALMERIM	0.0025461
ALTAMIRA	0.0198160
ANANINDEUA	0.0018087
BELEM	0.0469730
ITAITUBA	0.0068959
MONTE_ALEGRE	0.0158420
MONTE_DOURADO	0.0043013
OBIDOS	0.0115360
ORIXIMINA	0.0080174
PARINTINS	0.0079187
PORTO_TROMBETAS	0.0079187
PRAINHA	0.0068959
MAPACA_AP	0.0018087
MANAUS_AM	0.0025461
OUTROS_ESTADOS	0.0053132

BOLSA/APOIO MENSALIDADE

SEM_BOLSA	0.8426100
FUNCIONARIO	0.0071524
EMPRESA	0.0052267
BOLSISTA_IC	0.0034924
MONITORIA	0.0160490
ESPORTES	0.0929520
CORAL	0.0063739
CREDUC	0.0209140
LIS	0.0052267
OUTROS	0.0000000

SEXO

MASCULINO	0.6351600
FEMININO	0.3648400

SITUAÇÃO

ATIVO	0.3380800
TRANCAMENTO	0.4216900
CANCELAMENTO	0.0400570
TRANSF_INTERNA	0.0586980
TRANSF_EXTERNA	0.0560200
ABANDONO	0.0700490
FORMADO	0.0154060

EVASAO

NAO_EVADIR	0.3380800
EVADIR	0.6619200

TRANSFERÊNCIA INTERNA

DIREITO	0.0517330
ENG_AGRICOLA	0.0029847
LETRAS	0.0009949
PEDAGOGIA	0.0029847
SEM_TRANSFERENCIA	0.9413000

FAIXA ETÁRIA

ENTRE_18_E_19	0.100270
ENTRE_20_E_21	0.236860
ENTRE_22_E_23	0.208060
ENTRE_24_E_25	0.162720
ENTRE_26_A_30	0.148500
ENTRE_31_A_39	0.100530
ACIMA_DE_40	0.043049

SITUAÇÃO DE TRABALHO

TRABALHA	0.593860
NAOTRABALHA	0.406140

APÊNDICE E
CÓDIGO FONTE DO E-BAYES (EVASÃO BAYES)

{ PROGRAMA PRINCIPAL }

PROGRAM EVASAOBAYES;

USES

Forms,

CadAlunos in '**CadAlunos.pas**'; {*Manutenção do cadastro de alunos e da Rede Bayuesiana*}

main in '**Main.pas**'; {*Programa principal - Menus e Navegação*}

Sobre in '**Sobre.pas**'; {*Identificação e versão do programa*}

{**\$R *.RES**}

BEGIN

Application.Initialize; *{Inicializa a aplicação}*

Application.Title := 'Evasão Bayes'; *{Nome da aplicação}*

Application.CreateForm (TfrmMDIMain, frmMDIMain); *{Formulários utilizados}*

Application.Run; *{Executa a aplicação}*

END.

{ UNIT CAD ALUNOS }**Unit CadAlunos;****interface****uses**

Windows, Messages, SysUtils, Classes, Graphics, Controls, Forms, Dialogs, Buttons, ExtCtrls, ImgList, BCtrls, Db, DBTables, StdCtrls, Mask, NeticaComponents, ComCtrls;

type

```
TfrmCadAlunos = class(TForm)
  Panel1: TPanel; // Menu de Botões
    sbNovo: TSpeedButton;
    sbEditar: TSpeedButton;
    sbExcluir: TSpeedButton;
    sbSalvar: TSpeedButton;
    sbCancelar: TSpeedButton;
  Panel2: TPanel; // Acesso ao BD dos Alunos
    DBNavigator1: TDBNavigator;
    tbAlunos: TTable;
    dsAlunos: TDataSource;
    tbAlunosMatricula: TStringField;
    tbAlunosNome: TStringField;
    tbAlunosSexo: TStringField;
    tbAlunosSituacao: TStringField;
    tbAlunosCurso: TStringField;
    tbAlunosData_Nasc: TDateField;
    tbAlunosEndereco: TStringField;
    tbAlunosBairro: TStringField;
    tbAlunosCEP: TStringField;
    tbAlunosCidade: TStringField;
    tbAlunosUF: TStringField;
    tbAlunosFone: TStringField;
    tbAlunosAno_Ingresso: TSmallintField;
    tbAlunosSemestre: TSmallintField;
    tbAlunosTrabalha: TBooleanField;
    tbAlunosBolsa: TSmallintField;
  NeticaEnviron1: TNeticaEnviron; // Acesso ao EvasaoAlunos.dne via netica.dll
  PageControl1: TPageControl;
  TabSheet1: TTabSheet;
  TabSheet2: TTabSheet;
  ScrollBox1: TScrollBox;
  Label1: TLabel;
  Label2: TLabel;
  Label3: TLabel;
  Label4: TLabel;
  Label5: TLabel;
  Label7: TLabel;
  Label8: TLabel;
  Label9: TLabel;
```

```

Bevel1: TBevel;
DBEdit1: TDBEdit;
DBEdit2: TDBEdit;
DBEdit3: TDBEdit;
DBEdit4: TDBEdit;
DBEdit5: TDBEdit;
DBEdit7: TDBEdit;
DBEdit8: TDBEdit;
DBEdit9: TDBEdit;
DBRadioGroup1: TDBRadioGroup;
GroupBox1: TGroupBox;
Label6: TLabel;
Label10: TLabel;
DBEdit6: TDBEdit;
DBEdit10: TDBEdit;
DBRadioGroup2: TDBRadioGroup;
DBRadioGroup3: TDBRadioGroup;
DBRadioGroup4: TDBRadioGroup;
ntkEvasao: TNeticaNetwork; // Busca a rede criada no netica (EvasaoAlunos.dne)
ScrollBox2: TScrollBox;
// Busca cada nó da rede criada no netica EvasaoAlunos.dne, via netica.dll
ndEvasao: TNeticaNode;
ndSituacao: TNeticaNode;
ndTranfInterna: TNeticaNode;
ndTranfExterna: TNeticaNode;
ndApoio: TNeticaNode;
ndLocalOrigem: TNeticaNode;
ndSexo: TNeticaNode;
ndFaixaEtaria: TNeticaNode;
ndTrabalho: TNeticaNode;
NeticaNode10: TNeticaNode;
NeticaNode11: TNeticaNode;
Label11: TLabel;
DBText1: TDBText;
Label12: TLabel;
procedure dsAlunosStateChange(Sender: TObject);
procedure FormCreate(Sender: TObject);
procedure FormClose(Sender: TObject; var Action: TCloseAction);
procedure sbNovoClick(Sender: TObject);
procedure sbEditarClick(Sender: TObject);
procedure sbSalvarClick(Sender: TObject);
procedure sbCancelarClick(Sender: TObject);
procedure dsAlunosDataChange(Sender: TObject; Field: TField);
private
  { Private declarations }
  procedure SetCidade;
  procedure SetFaixaEtaria;
  procedure SetTrabalho;
  procedure SetSexo;
  procedure SetApoio;

```

```

    procedure SetSituacao;
    public
    { Public declarations }
    end;
var
    frmCadAlunos: TfrmCadAlunos;
implementation

uses main;

{$R *.DFM}
// Cidades disponíveis no sistema
Const Cidades: array[0..15] of string[20] = ('SANTAREM',
                                             'ALENQUER',
                                             'ALMERIM',
                                             'ALTAMIRA',
                                             'ANANINDEUA',
                                             'BELEM',
                                             'ITAITUBA',
                                             'MONTE ALEGRE',
                                             'MONTE DOURADO',
                                             'OBIDOS',
                                             'ORIXIMINA',
                                             'PARINTINS',
                                             'PORTO TROMBETAS',
                                             'PRAINHA',
                                             'MACAPA',
                                             'MANAUS');

// Permite Manutenção no Cadastro do Aluno
procedure TfrmCadAlunos.dsAlunosStateChange(Sender: TObject);
begin
    sbNovo.Enabled := (dsAlunos.State = dsBrowse);
    sbEditar.Enabled := (dsAlunos.State = dsBrowse);
    sbExcluir.Enabled := (dsAlunos.State = dsBrowse);
    sbSalvar.Enabled := (dsAlunos.State = dsInsert) OR (dsAlunos.State = dsEdit);
    sbCancelar.Enabled := (dsAlunos.State = dsInsert) OR (dsAlunos.State = dsEdit);
end;

// Habilita a rede bayesiana do aluno, puxando do EvasaoAlunos.dne, via netica.dll
procedure TfrmCadAlunos.FormCreate(Sender: TObject);
begin
    dsAlunos.DataSet.Open;
    if not ntkEvasao.Active then
        ntkEvasao.Active := true;
end;

// Fecha BD Alunos ao sair do sistema
procedure TfrmCadAlunos.FormClose(Sender: TObject; var Action: TCloseAction);
begin

```

```

dsAlunos.DataSet.Close;
frmMDIMain.mnuArquivoAlunos.Enabled := True;
Action := caFree;
end;

```

// Insere alunos no BD

```

procedure TfrmCadAlunos.sbNovoClick(Sender: TObject);
begin
  dsAlunos.DataSet.Append;
end;

```

// Edita alunos no BD

```

procedure TfrmCadAlunos.sbEditarClick(Sender: TObject);
begin
  dsAlunos.DataSet.Edit;
end;

```

// Salva alunos no BD

```

procedure TfrmCadAlunos.sbSalvarClick(Sender: TObject);
begin
  dsAlunos.DataSet.Post;
end;

```

// Cancela uma manutenção no alunos no BD

```

procedure TfrmCadAlunos.sbCancelarClick(Sender: TObject);
begin
  dsAlunos.DataSet.Cancel;
end;

```

// Atualiza a rede de acordo com as informações contidas no BD dos alunos

```

procedure TfrmCadAlunos.dsAlunosDataChange(Sender: TObject; Field: TField);
begin
  if dsAlunos.State = dsBrowse then
    begin
      if not ntkEvasao.Active then
        ntkEvasao.Active := true;
    end;
end;

```

// Reseta a rede após uma manutenção do cadastro de um dado aluno ou na troca de //alunos no sistema

```

  ntkEvasao.RetractAllFindings;
// Define valores para o aluno corrente
  SetCidade;
  SetFaixaEtaria;
  SetTrabalho;
  SetSexo;
  SetApoio;
  SetSituacao;
// Atualiza a rede após um click
  ntkEvasao.UpDate;
end;
end;

```

// Carrega a cidade do DB dos Alunos e seta (click) na rede bayesiana do Delphi

```

procedure TfrmCadAlunos.SetCidade;
var i, state: integer;
    cidade: String[20];
begin
    state := 16;
    cidade := UpperCase(Trim(tbAlunosCidade.AsString));
    for i := 0 to 15 do
        if cidade = cidades[i] then
            state := i;
    // Atualiza a rede via {StateFinding}
    ndLocalOrigem.StateFinding(state);
end;

```

// Carrega a data de nascimento do aluno na DB dos Alunos, calcula a sua faixa etária e //seta (click) na rede bayesiana do Delphi

```

procedure TfrmCadAlunos.SetFaixaEtaria;
var idade: integer;
    hoje: TDateTime;
    d, m, a1, a2: word;
    faixa: integer;
begin
    hoje := Date();
    DecodeDate(hoje, a2, m, d);
    DecodeDate(tbAlunosData_Nasc.AsDateTime, a1, m, d);
    idade := a2 - a1;
    if idade < 20 then
        faixa := 0
    else if idade < 22 then
        faixa := 1
    else if idade < 24 then
        faixa := 2
    else if idade < 26 then
        faixa := 3
    else if idade < 31 then
        faixa := 4
    else if idade < 40 then
        faixa := 5
    else
        faixa := 6;
    // Atualiza a rede via {StateFinding}
    ndFaixaEtaria.StateFinding( faixa );
end;

```

// Carrega a situação de trabalho do aluno no DB dos Alunos e seta (click) na rede bayesiana do Delphi

```

procedure TfrmCadAlunos.SetTrabalho;
begin
    if tbAlunosTrabalha.AsBoolean then
// Atualiza a rede via {StateFinding}

```



```

    ndTrabalho.StateFinding(0)
else
// Atualiza a rede via {StateFinding}
    ndTrabalho.StateFinding(1);
end;

// Carrega o sexo do aluno no DB dos Alunos e seta (click) na rede bayesiana do Delphi
procedure TfrmCadAlunos.SetSexo;
begin
    if tbAlunosSexo.AsString = 'M' then
// Atualiza a rede via {StateFinding}
        ndSexo.StateFinding(0)
    else
// Atualiza a rede via {StateFinding}
        ndSexo.StateFinding(1);
    end;

// Carrega a situação de Apoio do aluno no DB dos Alunos e seta (click) na rede bayesiana do Delphi
procedure TfrmCadAlunos.SetApoio;
var state: integer;
begin
    state := tbAlunosBolsa.AsInteger;
// Atualiza a rede via {StateFinding}
    ndApoio.StateFinding( state );
end;

// Carrega a Situação Acadêmica do aluno no DB dos Alunos e seta (click) na rede bayesiana do Delphi
procedure TfrmCadAlunos.SetSituacao;
var state: integer;
begin
    state := tbAlunosSituacao.AsInteger - 1;
    if state > 0 then
// Atualiza a rede via {StateFinding}
        ndSituacao.StateFinding( state );
    end;

end.

```

APÊNDICE F

CÓDIGO FONTE DOS COMPONENTES DELPHI

```
(*
* NeticaEx.pas - Versão em português
* Extras e Exemplo (external) Código Fonte (Delphi 4.x ou superior)
* para Netica - API, versão 2.00 ou superior.
*
* Baseado na biblioteca NeticaEx.c distribuída gratuitamente pela Norsys Corp.
*
* O software neste arquivo pode ser copiado, modificado, e/ou incluído em
* trabalhos derivados sem ônus ou obrigação.
* O autor, ou a Norsys, não dão garantias de suporte para este software.
*)
```

```
unit NeticaEx;
```

```
interface
```

```
uses NeticaConvert, Sysutils, Dialogs;
```

```
var env: p_environ_ns;
```

```
function NodeNamed(node_name: PChar; net: p_net_bn): p_node_bn;
procedure SetNodeStateNames(node: p_node_bn; names: array of PChar);
procedure SetNodeProbs(node: p_node_bn; names: array of PChar; probs: array of prob_bn);
function GetNodeBelief(node_name, state_name: PChar; net: p_net_bn): prob_bn;
procedure EnterFinding(node_name, state_name: PChar; net: p_net_bn);
procedure EnterFindingNot(node_name, state_name: PChar; net: p_net_bn);
procedure RetractNodeFindings(node_name: PChar; net: p_net_bn);
function GetNodeFinding(node_name: PChar; net: p_net_bn): state_bn;
Function IsStateFinding(node_name, state_name: PChar; net: p_net_bn): boolean;
```

```
implementation
```

```
Const ARR_SIZE = 20;
```

```
{** NodeNamed *****Como NodeNamed_bn, mas gera um erro se o nome não existir.*****}
```

```
function NodeNamed(node_name: PChar; net: p_net_bn): p_node_bn;
var node: p_node_bn;
    net_name: PChar;
begin
    node := NodeNamed_bn (node_name, net);
    if (node = Nil) then
        begin
            net_name := GetNetName_bn(net);
            ReportError_ns(env, 0, ERROR_ERR,
                PChar(format('NodeNamed: There is no node named "%s" in network
"%s"',
                [node_name, net_name]))));
        end;
end;
```

```
Result := node;
end;
```

```
{** SetNodeStateNames***** Este procedimento encapsula várias
chamadas para SetNodeStateName_bn em uma. Por exemplo para dar nomes aos 3 estados no
nó 'Weather' você pode chamar: SetNodeStateNames (Weather, ['sunny', 'cloudy', 'rainy']);
*****}
```

```
procedure SetNodeStateNames(node: p_node_bn; names: array of PChar);
var state, numstates: state_bn;
begin
  numstates := GetNodeNumberStates_bn(node);
  for state := 0 to (numstates - 1) do
    SetNodeStateName_bn(node, state, names[state]);
  end;
```

```
{** SetNodeProbs
*****}

Esta rotina parece ser uma versão mais conveniente (mas mais lenta) da
SetNodeProbs_bn. Seu primeiro argumento é o nó cujas probabilidades estamos
definindo. Seguido pelos nomes dos estados "condicionados" de seus pais como
Strings. Finalmente vem uma lista de doubles, representando as probabilidades
para cada estado do nó.
*****}
```

```
procedure SetNodeProbs(node: p_node_bn; names: array of PChar; probs: array of prob_bn);
var parent_states: array[0..ARR_SIZE] of state_bn;
  statename: PChar;
  numstates: state_bn;
  parents: p_nodelist_bn;
  pn, numparents: integer;
begin
  numstates := GetNodeNumberStates_bn(node);
  parents := GetNodeParents_bn(node);
  numparents := LengthNodeList_bn(parents);

  if (numstates > ARR_SIZE) or (numparents > ARR_SIZE) then
    begin
      reportError_ns(env, 0, XXX_ERR, 'SetNodeProbs: Array size defined too small');
    end;

  for pn := 0 to (numparents - 1) do
    begin
      statename := names[pn];
      if statename[1] = '*' then
        parent_states[pn] := EVERY_STATE
      else
        parent_states[pn] := StateNamed_bn(statename, NthNode_bn (parents, pn));
    end;
```

```

SetNodeProbs_bn(node, @parent_states[0], @probs[0]);
end;

```

```

{*** GetNodeBelief ***** Como o GetNodeBeliefs_bn, mas são passado nomes
(strings) e retorna-se o belief para um determinado Estado.*****}

```

```

function GetNodeBelief(node_name, state_name: PChar; net: p_net_bn): prob_bn;
var node: p_node_bn;
    state: state_bn;
    probs: p_prob_bn;
    prob: prob_bn;
    i: integer;
begin
    node := NodeNamed(node_name, net);
    state := StateNamed_bn(state_name, node);
    probs := GetNodeBeliefs_bn(node);

    i := 0;
    prob := probs^;
    while i < state do
        begin
            inc(probs, 1);
            prob := probs^;
            inc(i);
        end;

    Result := prob;
end;

```

```

{*** EnterFinding ***** Como o EnterFinding_bn, mas são passado nomes
(strings)*****}

```

```

procedure EnterFinding(node_name, state_name: PChar; net: p_net_bn);
var node: p_node_bn;
    state: state_bn;
begin
    node := NodeNamed(node_name, net);
    state := StateNamed_bn(state_name, node);
    EnterFinding_bn(node, state);
end;

```

```

{*** EnterFindingNot ***** Como o EnterFindingNot_bn, mas são passado nomes
(strings)*****}

```

```

procedure EnterFindingNot(node_name, state_name: PChar; net: p_net_bn);
var node: p_node_bn;
    state: state_bn;
begin
    node := NodeNamed(node_name, net);

```

```

state := StateNamed_bn(state_name, node);
EnterFindingNot_bn(node, state);
end;

{*** RetractNodeFindings ***** Como o RetractNodeFindings_bn, mas são passado
nomes (strings)*****}

procedure RetractNodeFindings(node_name: PChar; net: p_net_bn);
var node: p_node_bn;
begin
node := NodeNamed(node_name, net);
RetractNodeFindings_bn(node);
end;

{*** GetNodeFinding *****Como o GetNodeFinding_bn, mas são passado nomes (strings)
*****}

function GetNodeFinding(node_name: PChar; net: p_net_bn): state_bn;
var node: p_node_bn;
begin
node := NodeNamed(node_name, net);
Result := GetNodeFinding_bn(node);
end;

{*** IsStateFinding ***** Verifica se um estado está em finding.*****}

Function IsStateFinding(node_name, state_name: PChar; net: p_net_bn): boolean;
var state: state_bn;
node: p_node_bn;
begin
node := NodeNamed(node_name, net);
state := StateNamed_bn(state_name, node);
result := GetNodeFinding(node_name, net) = state
end;

end.

```

unit NeticaEnviron;

interface**uses**

Windows, Messages, SysUtils, Classes, Graphics, Controls, Forms, Dialogs, StdCtrls, NeticaConvert, NeticaEx;

type

```
ENeticaError = class(Exception)
end;
```

```
TNetworkList = class(TObject)
```

```
private
```

```
  { Private declarations }
```

```
  FList: TList;
```

```
  function GetCount: Integer;
```

```
protected
```

```
  { Protected declarations }
```

```
public
```

```
  { Public declarations }
```

```
  constructor Create;
```

```
  Function Add(Network: Pointer):integer;
```

```
  Function Network(Index: Integer): Pointer;
```

```
  Procedure Delete(Index: Integer);
```

```
  destructor Destroy; override;
```

```
  //Propiedades
```

```
  property Count:Integer read GetCount;
```

```
published
```

```
  { Published declarations }
```

```
end;
```

```
TNeticaEnviron = class(TComponent)
```

```
private
```

```
  { Private declarations }
```

```
  FEnv: p_environ_ns;
```

```
  Flicense: String;
```

```
  FActive: Boolean;
```

```
  FNetworks: TNetworkList;
```

```
  FMessages: TMemo;
```

```
  procedure Setlicense(const Value: String);
```

```
  procedure SetActive(const Value: Boolean);
```

```
  function GetNetwork(index: Integer): Pointer;
```

```
  function GetNetworkCount: Integer;
```

```
protected
```

```
  { Protected declarations }
```

```
  procedure NewNeticaEnviron;
```

```
  procedure CloseNetica;
```

```
public
```

```
  { Public declarations }
```

```

constructor Create(AOwner: TComponent); override;
destructor Destroy; override;
function AddNetwork(Network: Pointer): Integer;
procedure DeleteNetwork(Index: Integer);

// Propriedades
property Enviroment: p_environ_ns read FEnv;
property Networks[index: Integer]: Pointer read GetNetwork;

published
{ Published declarations }
property license : String read Flicense write Setlicense;
property Active: Boolean read FActive write SetActive;
property Messages: TMemo read FMessages write FMessages;
property NetworkCount: Integer Read GetNetworkCount;
end;

procedure Register;

implementation

procedure Register;
begin
  RegisterComponents('Netica', [TNeticaEnviron]);
end;

{ TNeticaEnviroment }

function TNeticaEnviron.AddNetwork(Network: Pointer): Integer;
begin
  Result := FNetworks.Add(Network);
end;

procedure TNeticaEnviron.CloseNetica;
var x, res: Integer;
    mesg: array[0..MESG_LEN_ns - 1] of Char;
begin

{ TODO : Liberar memoria usada(s) pela(s) rede(s) }
{ // Libera recursos usados pela rede
if net <> nil then
  FreeNet_bn(net);}

// Inicializa variavel para mensagem
for x := 0 to MESG_LEN_ns - 1 do mesg[x] := ' ';

// Fecha Ambiente Netica
res := CloseNetica_bn(FEnv, @mesg);

{ FMessages.Add( Trim(mesg) );}

```



```

if FMessages <> Nil then
  FMessages.Lines.Add( Trim(mesg) );

// Se ocorreu um erro...
if res < 0 then
  begin
    ENeticaError.Create('Error closing Netica!');
  end;
end;

constructor TNeticaEnviron.Create(AOwner: TComponent);
begin
  inherited Create(AOwner);
  FLicense := "";
  FActive := False;
  FMessages := Nil;
  FNetworks := TNetworkList.Create;
end;

procedure TNeticaEnviron.DeleteNetwork(Index: Integer);
begin
  FNetworks.Delete(Index);
end;

destructor TNeticaEnviron.Destroy;
begin
  CloseNetica;
  FNetworks.Destroy;
  inherited;
end;

function TNeticaEnviron.GetNetwork(index: Integer): Pointer;
begin
  Result := FNetworks.Network(index);
end;

function TNeticaEnviron.GetNetworkCount: Integer;
begin
  Result := FNetworks.Count;
end;

procedure TNeticaEnviron.NewNeticaEnviron;
var x, res: Integer;
    mesg: array[0..MSG_LEN_ns - 1] of Char;
begin
  // Inicializa variavel para mensagem
  for x := 0 to MSG_LEN_ns - 1 do mesg[x] := ' ';

  // Cria um novo ambiente Netica
  if FLicense = " then

```

```

    FEnv := NewNeticaEnviron_bn( Nil )
else
    FEnv := NewNeticaEnviron_bn( PChar( Flicense ) );

// Inicializa o Netica
res := InitNetica_bn( @FEnv, @mesg );

{ FMessages.Add( Trim( mesg ) ); }
if FMessages <> Nil then
    FMessages.Lines.Add( Trim( mesg ) );

// Se ocorreu um erro...
if res < 0 then
    begin
        ENeticaError.Create( 'Error initiating Netica!' );
    end;
end;

procedure TNeticaEnviron.SetActive( const Value: Boolean );
begin

    if Value = true then
        begin
            { Messages.Clear; }
            NewNeticaEnviron;
        end
    else
        CloseNetica;

    FActive := Value;
end;

procedure TNeticaEnviron.Setlicense( const Value: String );
begin
    Flicense := Value;
end;

{ TNetworkList }

function TNetworkList.Add( Network: Pointer ): Integer;
begin
    Result := FList.Add( Network );
end;

constructor TNetworkList.Create;
begin
    FList := TList.Create;
end;
procedure TNetworkList.Delete( Index: Integer );

```

```
begin
  FList.Delete(Index);
end;

destructor TNetworkList.Destroy;
begin
  FList.Free;
  inherited;
end;

function TNetworkList.GetCount: Integer;
begin
  Result := FList.Count;
end;

function TNetworkList.Network(Index: Integer): Pointer;
begin
  Result := FList.Items[Index];
end;
end.
```

unit NeticaEnviron;

interface

uses

Windows, Messages, SysUtils, Classes, Graphics, Controls, Forms, Dialogs,
StdCtrls, NeticaConvert, NeticaEx;

type

ENeticaError = class(Exception)
end;

TNetworkList = class(TObject)

private

{ Private declarations }

FList: TList;

function GetCount: Integer;

protected

{ Protected declarations }

public

{ Public declarations }

constructor Create;

Function Add(Network: Pointer):integer;

Function Network(Index: Integer): Pointer;

Procedure Delete(Index: Integer);

destructor Destroy; override;

//Propiedades

property Count:Integer read GetCount;

published

{ Published declarations }

end;

TNeticaEnviron = class(TComponent)

private

{ Private declarations }

FEnv: p_environ_ns;

Flicense: String;

FActive: Boolean;

FNetworks: TNetworkList;

FMessages: TMemo;

procedure Setlicense(const Value: String);

procedure SetActive(const Value: Boolean);

function GetNetwork(index: Integer): Pointer;

function GetNetworkCount: Integer;

protected

{ Protected declarations }

procedure NewNeticaEnviron;

procedure CloseNetica;

public

```

{ Public declarations }
constructor Create(AOwner: TComponent); override;
destructor Destroy; override;
function AddNetwork(Network: Pointer): Integer;
procedure DeleteNetwork(Index: Integer);

// Propriedades
property Enviroment: p_environ_ns read FEnv;
property Networks[index: Integer]: Pointer read GetNetwork;

published
{ Published declarations }
property license : String read Flicense write Setlicense;
property Active: Boolean read FActive write SetActive;
property Messages: TMemo read FMessages write FMessages;
property NetworkCount: Integer read GetNetworkCount;
end;

procedure Register;

implementation

procedure Register;
begin
  RegisterComponents('Netica', [TNeticaEnviron]);
end;

{ TNeticaEnviroment }

function TNeticaEnviron.AddNetwork(Network: Pointer): Integer;
begin
  Result := FNetworks.Add(Network);
end;

procedure TNeticaEnviron.CloseNetica;
var x, res: Integer;
    mesg: array[0..MESG_LEN_ns - 1] of Char;
begin

{ TODO : Liberar memoria usada(s) pela(s) rede(s) }
{ // Libera recursos usados pela rede
if net <> nil then
  FreeNet_bn(net);}

// Inicializa variavel para mensagem
for x := 0 to MESG_LEN_ns - 1 do mesg[x] := ' ';

// Fecha Ambiente Netica
res := CloseNetica_bn(FEnv, @mesg);

```

```

{ FMessages.Add( Trim(mesg) );}
if FMessages <> Nil then
  FMessages.Lines.Add( Trim(mesg) );

// Se ocorreu um erro...
if res < 0 then
  begin
    ENeticaError.Create('Error closing Netica!');
  end;
end;

constructor TNeticaEnviron.Create(AOwner: TComponent);
begin
  inherited Create(AOwner);
  Flicense := '';
  FActive := False;
  FMessages := Nil;
  FNetworks := TNetworkList.Create;
end;

procedure TNeticaEnviron.DeleteNetwork(Index: Integer);
begin
  FNetworks.Delete(Index);
end;

destructor TNeticaEnviron.Destroy;
begin
  CloseNetica;
  FNetworks.Destroy;
  inherited;
end;

function TNeticaEnviron.GetNetwork(index: Integer): Pointer;
begin
  Result := FNetworks.Network(index);
end;

function TNeticaEnviron.GetNetworkCount: Integer;
begin
  Result := FNetworks.Count;
end;

procedure TNeticaEnviron.NewNeticaEnviron;
var x, res: Integer;
    mesg: array[0..MESG_LEN_ns - 1] of Char;
begin
  // Inicializa variavel para mensagem
  for x := 0 to MESG_LEN_ns - 1 do mesg[x] := ' ';

  // Cria um novo ambiente Netica

```

```

if Flicense = " then
  FEnv := NewNeticaEnviron_bn( Nil )
else
  FEnv := NewNeticaEnviron_bn( PChar( Flicense ) );

// Inicializa o Netica
res := InitNetica_bn( @FEnv, @mesg );

{ FMessages.Add( Trim( mesg ) ); }
if FMessages <> Nil then
  FMessages.Lines.Add( Trim( mesg ) );

// Se ocorreu um erro...
if res < 0 then
  begin
    ENeticaError.Create( 'Error initiating Netica!' );
  end;
end;

procedure TNeticaEnviron.SetActive( const Value: Boolean );
begin

  if Value = true then
    begin
      { Messages.Clear; }
      NewNeticaEnviron;
    end
  else
    CloseNetica;

  FActive := Value;
end;

procedure TNeticaEnviron.Setlicense( const Value: String );
begin
  Flicense := Value;
end;

{ TNetworkList }

function TNetworkList.Add( Network: Pointer ): Integer;
begin
  Result := FList.Add( Network );
end;

constructor TNetworkList.Create;
begin
  FList := TList.Create;
end;

```

```
procedure TNetworkList.Delete(Index: Integer);

begin

    FList.Delete(Index);

end;

destructor TNetworkList.Destroy;

begin

    FList.Free;

    inherited;

end;

function TNetworkList.GetCount: Integer;

begin

    Result := FList.Count;

end;

function TNetworkList.Network(Index: Integer): Pointer;

begin

    Result := FList.Items[Index];

end;

end.
```


**// Unit de Conversão para utilizar funções das redes bayesianas do netica em
// programas do delphi**

Unit NeticaConvert;

interface

const

NeticaDLL = 'Netica.dll';

const

MESG_LEN_ns = 256;

NAME_MAX_ns = 30;

const

//

// query

QUERY_ns = -1;

//

// errdanger_ns

//

NOTHING_ERR = 1;

REPORT_ERR = 2;

NOTICE_ERR = 3;

WARNING_ERR = 4;

ERROR_ERR = 5;

XXX_ERR = 6;

//

// nodekind_bn

//

NATURE_NODE = 0;

ASSUME_NODE = 1;

DECISION_NODE = 2;

UTILITY_NODE = 3;

DISCONNECTED_NODE = 4;

//

// state_bn

//

WILDCARD_STATE = -5;

IMPOSS_STATE = -4;

UNDEF_STATE = -3;

//

// checking_ns

//

NO_CHECKING = 23;

QUICK_CHECK = 24;

REGULAR_CHECK = 25;

COMPLETE_CHECK = 26;

//

// bool_ns

//

```

FALSE_ns      = 0;
TRUE_ns       = 1;

//
//
EVERY_STATE   = -5;
NO_FINDING    = -8;
//
// caseposn_bn
//
FIRST_CASE    = -15;
NEXT_CASE     = -14;
NO_MORE_CASES = -13;

type
level_bn      = Double;
long          = Integer;
int           = Integer;
p_environ_ns  = pointer;
p_p_environ_ns = pointer;
p_net_bn      = pointer;
errdanger_ns  = Integer;
p_report_ns   = pointer;
p_node_bn     = pointer;
state_bn      = long;
p_state_bn    = pointer;
p_PChar       = ^PChar;
prob_bn       = Single; // float in "C"
p_prob_bn     = ^prob_bn;
checking_ns   = int;
p_nodelist_bn = pointer;
bool_ns       = int;
p_stream_ns   = pointer;
caseposn_bn   = long;
p_caseposn_bn = ^caseposn_bn;
p_long        = ^long;
p_Double      = ^Double;
{double       = Double fortunately}
//
// Administration
//
function NewNeticaEnviron_bn(license : PChar) : p_environ_ns; stdcall; external
NeticaDLL;
function InitNetica_bn(env : p_p_environ_ns; mesg : PChar) : int; stdcall; external
NeticaDLL;
function CloseNetica_bn(env : p_environ_ns; mesg : PChar) : int; stdcall; external
NeticaDLL;
function NewNet_bn(name : PChar; env : p_environ_ns) : p_net_bn; stdcall; external
NeticaDLL;

```

```

procedure FreeNet_bn(net : p_net_bn); stdcall; external NeticaDLL;
function GetNeticaVersion_bn(env : p_environ_ns; version : p_PChar) : int; stdcall; external
NeticaDLL;
function ArgumentChecking_ns(setting : checking_ns; env : p_environ_ns) : checking_ns;
stdcall; external NeticaDLL;
function GetNetName_bn(net : p_net_bn) : PChar; stdcall; external NeticaDLL;
procedure CompileNet_bn(net : p_net_bn); stdcall; external NeticaDLL;
procedure WriteNet_bn(net : p_net_bn; netfile : p_stream_ns); stdcall; external NeticaDLL;
function ReadNet_bn(netfile : p_stream_ns; read_visual : bool_ns) : p_net_bn; stdcall;
external NeticaDLL;
function FileNamed_ns(filename : PChar; env : p_environ_ns) : p_stream_ns; stdcall;
external NeticaDLL;
//
// Errors
//
function GetError_ns(env : p_environ_ns; danger : errdanger_ns; after : p_report_ns) :
p_report_ns; stdcall; external NeticaDLL;
function ErrorMessage_ns(error : p_report_ns) : PChar; stdcall; external NeticaDLL;
function ErrorNumber_ns(error : p_report_ns) : int; stdcall; external NeticaDLL;
function ErrorDanger_ns(error : p_report_ns) : errdanger_ns; stdcall; external NeticaDLL;
procedure ClearError_ns(error : p_report_ns); stdcall; external NeticaDLL;
procedure ClearErrors_ns(env : p_environ_ns; danger : errdanger_ns); stdcall; external
NeticaDLL;

{ Acrescentado por Enoque Calvino Melo Alves }
function ReportError_ns(env : p_environ_ns; number: integer; danger : errdanger_ns; const
mesg: PChar): p_report_ns; stdcall; external NeticaDLL;

//
// Node operations
//
function NewNode_bn(name : PChar; num_status : Integer; net : p_net_bn) : p_node_bn;
stdcall; external NeticaDLL;
procedure SetNodeStateName_bn(node : p_node_bn; state : state_bn; state_name : PChar);
stdcall; external NeticaDLL;
procedure SetNodeProbs_bn(node : p_node_bn;
parent_states : pointer;
probs : pointer {p_prob_bn}); stdcall; external NeticaDLL;
function GetNodeBeliefs_bn(node : p_node_bn) : p_prob_bn; stdcall; external NeticaDLL;
function NodeNamed_bn(name : PChar; net : p_net_bn) : p_node_bn; stdcall; external
NeticaDLL;
function GetNodeLevel_bn(node : p_node_bn; state : state_bn) : level_bn; stdcall; external
NeticaDLL;

{ Acrescentado por Enoque Calvino Melo Alves }
procedure SetNodeTitle_bn(node: p_node_bn; const name: PChar); stdcall; external
NeticaDLL;
function GetNodeTitle_bn(const node: p_node_bn): PChar; stdcall; external NeticaDLL;
function StateNamed_bn(const name: PChar; node: p_node_bn): state_bn; stdcall; external
NeticaDLL;

```

```

function GetNodeStateName_bn(const node: p_node_bn; state: state_bn): PChar; stdcall;
external NeticaDLL;

//
// Node lists
//
function LengthNodeList_bn(nodes : p_nodelist_bn) : int; stdcall; external NeticaDLL;
function GetNetNodes_bn(net : p_net_bn) : p_nodelist_bn; stdcall; external NeticaDLL;
function NthNode_bn(nodes : p_nodelist_bn; index : int) : p_node_bn; stdcall; external
NeticaDLL;
function GetNodeParents_bn(node : p_node_bn) : p_nodelist_bn; stdcall; external
NeticaDLL;
function GetNodeChildren_bn(node : p_node_bn) : p_nodelist_bn; stdcall; external
NeticaDLL;
function GetNodeName_bn(node : p_node_bn) : PChar; stdcall; external NeticaDLL;
function GetNodeNumberStates_bn(node : p_node_bn) : int; stdcall; external NeticaDLL;
function DuplicateNodes_bn(nodes : p_nodelist_bn; new_net : p_net_bn) : p_nodelist_bn;
stdcall; external NeticaDLL;
function DupNodeList_bn(nodes : p_nodelist_bn) : p_nodelist_bn; stdcall; external
NeticaDLL;
procedure FreeNodeList_bn(nodes : p_nodelist_bn); stdcall; external NeticaDLL;
//
// Link operations
//
function AddLink_bn(parent : p_node_bn; child : p_node_bn) : int; stdcall; external
NeticaDLL;
procedure DeleteLink_bn(link_index : int; child : p_node_bn); stdcall; external NeticaDll;
//
// Evidence
//
procedure RetractAllFindings_bn(net : p_net_bn); stdcall; external NeticaDLL;
procedure EnterFinding_bn(node : p_node_bn; state : state_bn); stdcall; external NeticaDLL;
procedure EnterNodeValue_bn(node : p_node_bn; value : Double); stdcall; external
NeticaDLL;
function GetNodeFinding_bn(node : p_node_bn) : state_bn; stdcall; external NeticaDLL;

{ Acrescentado por Enoque Calvino Melo Alves }
procedure EnterFindingNot_bn(node: p_node_bn; state: state_bn); stdcall; external
NeticaDLL;
procedure RetractNodeFindings_bn(node: p_node_bn); stdcall; external NeticaDLL;

//
// Joint probabilities
//
function RandomCase_bn(nodes : p_nodelist_bn; method : int; num : Double) : int; stdcall;
external NeticaDLL;
procedure ReadCase_bn(var case_posn : {p_}caseposn_bn; casefile : p_stream_ns; nodes :
p_nodelist_bn; ID_num : p_long; freq : p_Double); stdcall; external NeticaDLL;
function WriteCase_bn(nodes : p_nodelist_bn; casefile : p_stream_ns; ID_num : long; freq :
Double) : caseposn_bn; stdcall; external NeticaDLL;

```

```
function CaseProbability_bn(net : p_net_bn) : Double; stdcall; external NeticaDLL;  
implementation
```

```
end.
```

UNIT NETICANETWORK;**INTERFACE****USES**

Windows, Messages, SysUtils, Classes, Graphics, Controls, Forms, Dialogs,
NeticaEnviron;

TYPE

```
TNeticaNetwork = class(TComponent)
private
  FNeticaEnviron: TNeticaEnviron;
  FIndex: Integer;
  FFileName: String;
  procedure SetNeticaEnviron(const Value: TNeticaEnviron);
  procedure SetFileName(const Value: String);
  { Private declarations }
protected
  { Protected declarations }
public
  { Public declarations }
  constructor Create(AOwner: TComponent); override;
  destructor Destroy; override;
published
  { Published declarations }
  property NeticaEnviron: TNeticaEnviron read FNeticaEnviron write SetNeticaEnviron;
  property FileName: String read FFileName write SetFileName;
end;
```

```
procedure Register;
```

```
implementation
```

```
procedure Register;
begin
  RegisterComponents('Netica', [TNeticaNetwork]);
end;
```

```
{ TNeticaNetwork }
```

```
constructor TNeticaNetwork.Create(AOwner: TComponent);
begin
  inherited Create(AOwner);
  FNeticaEnviron := Nil;
end;
```

```
destructor TNeticaNetwork.Destroy;
begin
  if FNeticaEnviron <> Nil then
    begin
```

```
    FNeticaEnviron.DeleteNetwork(FIndex);
    FNeticaEnviron := Nil;
  end;
  inherited;
end;

procedure TNeticaNetwork.SetFileName(const Value: String);
begin
  FFileName := Value;
end;

procedure TNeticaNetwork.SetNeticaEnviron(const Value: TNeticaEnviron);
begin
  if Value <> FNeticaEnviron then
    if Value = Nil then
      begin
        FNeticaEnviron.DeleteNetwork(FIndex);
        FNeticaEnviron := Value;
      end
    else
      begin
        FNeticaEnviron := Value;
        FIndex := FNeticaEnviron.AddNetwork(Self);
      end;
    end;
end;

end.
```

APÊNDICE G
RESUMOS ACEITOS EM CONGRESSOS

Evento: I Congresso de Ciência e Tecnologia da Amazônia
Realização: 07 a 10 de novembro de 2000
Local: Instituto Luterano de Ensino Superior de Santarém
URL: //santarem.ulbra.br/ccta2000

APLICAÇÃO DE REDES *BAYESIANAS* NA ANÁLISE DA EVASÃO DE ALUNOS DO CURSO DE INFORMÁTICA

Eugênio Rovaris Neto⁽¹⁾
Giovanna Ramos Strassburger⁽²⁾
Roberta Márcia Riker Maduro⁽³⁾

O presente trabalho tem como objetivo aplicar os conhecimentos de redes *bayesianas* para simular um Sistema Especialista Probabilístico (SEP) que avalie a situação da evasão dos alunos do Curso de Informática do ILES de Santarém, com base nas informações extraídas de questionários aplicados aos alunos matriculados, bem como aos ex-alunos deste curso. Informações como: *classe social, dificuldades, fonte de pagamento, semestre em que se encontra matriculado tipo de emprego, grau de satisfação, afinidade com a área, faixa etária, sexo e número de filhos* tornam-se variáveis importantes, que se relacionam entre si no auxílio aos cálculos das Probabilidades Condicionais e do Raciocínio *Bayesiano*. A pesquisa surgiu da necessidade de tentar compreender melhor a movimentação discente que evadi ou permanece cursando esta faculdade. Motivados por tal problema, pretende-se com este trabalho mostrar a aplicação prática das redes *bayesianas*, simulando um sistema de consulta para estimar as chances que um aluno tem de permanecer até a conclusão de seu curso ou de evadi-lo. Uma vez que as redes bayesianas facilitam a aquisição de conhecimento e tornam a representação do raciocínio probabilístico possível. Até o momento detectou-se, por exemplo, que apenas 33,808% dos alunos estão matriculados, 26,273% trancaram a sua matrícula, 1,833% cancelaram totalmente o curso, 12,220% optaram pela transferência interna para outros cursos da instituição enquanto que 6,110% dos ex-alunos mudaram para outras instituições de ensino superior através da transferência externa, além disso 19,756% dos alunos simplesmente abandonaram o curso, perfazendo assim um total de 66,192% de evasão.

Palavras-Chaves: Inteligência Artificial, Redes *Bayesianas* e Evasão.
Órgão Financiador: ULBRA

(1) Orientador e pesquisador do curso de Informática do ILES Santarém.
(2) Colaboradora e pesquisadora do curso de Informática do ILES Santarém.
(3) Acadêmica do 7º semestre do curso de Informática do ILES Santarém.

Evento: II Congresso Científico do Centro Universitário Luterano de Palmas
Realização: 16 a 18 de maio de 2001
Local: Centro Universitário Luterano de Palmas
URL: <http://www.ulbra-to.br>

REDES BAYESIANAS APLICADAS NO DESENVOLVIMENTO DE UM SISTEMA ESPECIALISTA PARA INFERÊNCIA DO DIAGNÓSTICO DE DOENÇAS INFECCIOSAS NO OESTE DO PARÁ

SANTOS, Benedita Neves dos¹; SOUZA, Lucélia Cristina Machado de¹ ;
 FROHLICH, Ricardo¹; ROVARIS, Eugênio Neto²

(INTRODUÇÃO) Em virtude da grande ocorrência de doenças infecciosas na região do Oeste do Pará, e da dificuldade de diagnosticá-las devido ao grande número de sintomas semelhantes entre elas, surgiu a idéia de desenvolver um sistema especialista (SE) que simulasse a atuação de peritos humanos. No entanto no paradigma simbólico, o conhecimento do especialista é disposto numa base de conhecimento, onde as inferências são representadas através de regras do tipo *se-então*. No caso de diagnóstico médico os sinais e sintomas que o médico coleta apresentam vários problemas de incertezas, tais como: a inexatidão dos relatos do paciente e a gravidade de cada sintoma, entre outros. Surge então, uma outra alternativa, a aplicação de *redes bayesianas*, onde se torna possível a inserção de fatores de incerteza, permitindo assim uma inferência mais próxima da realidade humana, uma vez que a rede expressa relações casuais e valores de probabilidade refletindo a força da relação. Este processo se faz pelo uso da probabilidade condicional do Teorema de Bayes. (METODOLOGIA) Foram analisados dois softwares específicos para cada abordagem. O *shell do expert sinta* na tentativa de construir uma base de conhecimento baseada em regras de produção com fatores de confiança e o *norsys netica*, voltado para o paradigma numérico de representação de incertezas, através das *probabilidades condicionais* e do *raciocínio bayesiano*. (RESULTADOS E DISCUSSÃO) Os resultados obtidos no *Expert Sinta* não foram satisfatórios devido aos poucos recursos oferecidos pela ferramenta neste tipo de problema, além da dificuldade da parte dos peritos humanos em expressar o conhecimento através de regras de produção. Já com a utilização do *norsys netica*, se obteve resultados satisfatórios, a partir da utilização de dados precisos fornecidos por profissional da área médica. Além disso, da parte do especialista humano, houve uma maior aceitação e compreensão da lógica de funcionamento do paradigma numérico-probabilista. (CONCLUSÃO) O desenvolvimento de SE Probabilísticos para o diagnóstico de doenças infecciosas é de grande importância e traz resultados significativos, visto que torna o diagnóstico mais preciso e rápido, facilitando a relação médico-paciente.

¹ Acadêmicos do Curso de Informática do ILES Santarém – celia@netsan.com.br

² Prof. Pesquisador do Depto de Informática, ILES Santarém – rovaris@ta.p.com.br

Evento: II Congresso de Ciência e Tecnologia da Amazônia
Realização: 01 a 05 de outubro de 2001
Local: Instituto Luterano de Ensino Superior de Santarém
URL: //santarem.ulbra.br/ccta

APLICAÇÃO DE *data mining* COM RECURSOS DE redes *bayesianas* NA ANÁLISE DA EVASÃO DE ALUNOS DO CURSO DE INFORMÁTICA DO ILES SANTARÉM

Eugênio Rovaris Neto⁷
Silvia Modesto Nassar⁸
eugenio@santarem.ulbra.br

O presente trabalho tem como objetivo aplicar técnicas de *data mining* e conhecimentos de redes *bayesianas* para simular um Sistema Especialista Probabilístico - SEP que avalie a situação da evasão dos alunos do Curso de Informática do Instituto Luterano de Ensino Superior de Santarém. Motivados por tal problema, pretende-se mostrar a aplicação prática de técnicas de mineração de dados e uso de redes *bayesianas*. Simulando um sistema de consulta para estimar as chances que um aluno tem de permanecer até a conclusão de seu curso ou de evadi-lo. As etapas envolvidas foram à mineração de dados numa base de dados existente e numa nova base de dados criada para armazenar o perfil deste aluno. Por último, foram processados e adaptados num *software* apropriado para desenvolvimento de SEP. As redes bayesianas facilitam a aquisição de conhecimento e tornam a representação do raciocínio probabilístico possível. O sistema proporciona uma gama excepcional de opções para visualização de estatísticas e nas previsões probabilísticas apoiadas na probabilidade condicional e no Teorema de Bayes, que são bases do raciocínio com incerteza. A pesquisa surgiu da necessidade de tentar compreender melhor a movimentação discente que evadi ou permanece cursando esta faculdade.

Palavra-chave: *Data Mining*, Redes Bayesianas, Evasão
Órgão financiador: PROICT/ULBRA, UFSC

⁷ Aluno do Curso Mestrado em Ciências da Computação - CTC-INE-UFSC

⁸ Orientadora do Curso de Pós-Graduação em Ciências da Computação - CTC-INE-UFSC

Evento: II Congresso de Ciência e Tecnologia da Amazônia
Realização: 01 a 05 de outubro de 2001
Local: Instituto Luterano de Ensino Superior de Santarém
URL: //santarem.ulbra.br/ccta

PARADIGMA SIMBÓLICO *versus* PARADIGMA NUMÉRICO, NA AVALIAÇÃO DE COMPORTAMENTO DO PROTÓTIPO DE UM SISTEMA ESPECIALISTA PARA O AUXÍLIO NO DIAGNÓSTICO DAS DST – *doenças sexualmente transmissíveis*

Suelen Cristina Alves Cruz⁹
Bruno Henrique Jares¹
Clebson de Souza Feitosa¹
Eugênio Rovaris Neto¹⁰
Silvia Modesto Nassar¹¹
eugenio@santarem.ulbra.br

Este trabalho apresenta um estudo comparativo entre os paradigmas simbólicos e numéricos, na representação de incertezas. Enquanto o paradigma simbólico infere as incertezas através de regras de produção com fatores de confiança, o paradigma numérico trabalha a incerteza numericamente, através das inferências e combinações de evidências, que podem ser representadas pela teoria das probabilidades. O problema escolhido para aplicarmos estes dois paradigmas foi à construção de um sistema especialista que auxilie no diagnóstico de doenças sexualmente transmissíveis - DST. Para o paradigma simbólico adotou-se um aplicativo que trabalha com regras de produção. No paradigma numérico, optou-se por um outro que trabalhe com a teoria das probabilidades. Algumas das doenças que podem ser diagnosticadas nestes dois protótipos são: *Vaginose Bacteriana, Candidíase, Tricomaníase, Cancro Duro, Gonorréia e Sífilis*. A representação de incerteza no paradigma simbólico é respondida pelo usuário, durante a execução do sistema. Enquanto que no paradigma numérico é gerada automaticamente, através do cálculo das probabilidades condicionais e das atualizações na *Rede Bayesiana*. O que classifica este último, como mais confiável e de melhor comportamento, indicado assim, para os problemas de diagnóstico nos sistemas especialistas.

Palavra-chave: Paradigma Simbólico, Paradigma Simbólico, DST
Órgão financiador: PROICT/ULBRA

⁹ Acadêmicos do Curso de Sistemas de Informação do ILES Santarém

¹⁰ Professor do Curso de Sistemas de Informação do ILES Santarém

¹¹ Orientadora do Curso de Pós-Graduação em Ciências da Computação - CTC-INE-UFSC

Evento: II Congresso de Ciência e Tecnologia da Amazônia
Realização: 01 a 05 de outubro de 2001
Local: Instituto Luterano de Ensino Superior de Santarém
URL: //santarem.ulbra.br/ccta

REDES BAYESIANAS *versus* REGRAS DE INFERÊNCIA, NA IDENTIFICAÇÃO DA RAÇA DO POVO BRASILEIRO

Adriana Pereira Silva¹²
Luciano Barbosa Oliveira¹
Eugênio Rovaris Neto¹³
Silvia Modesto Nassar¹⁴
eugenio@santarem.ulbra.br

Enquanto que as regras de inferência representam as exceções no raciocínio do especialista. O que implica na necessidade de criação uma base de conhecimento com uma pequena quantidade de exceções. A teoria das probabilidades, através das redes bayesianas, trata de uma estrutura rigorosa de representação de eventos aleatórios. Para a análise do comportamento de um conjunto de regras de inferência, foi adotado o *shell expert SINTA*, que trabalha com regras de produção. Quanto a análise do comportamento do raciocínio bayesiano, optou-se pelo aplicativo *Netica* que trabalha com a teoria da probabilidade. Para confronto destes dois modelos, construiu-se o protótipo para um sistema especialista de classificação. O modelo implementado com regras de inferência gerou uma base de conhecimento com 06 regras e 07 variáveis. Enquanto que o modelo Bayesiano gerou 03 nós naturais com 12 variáveis de estados. Deve-se evitar o uso de sistemas que tenham como saídas: cálculos de probabilidade, uma vez que os usuários finais ainda não estão adaptados com este novo paradigma de saída de um sistema. Com exceção dos problemas de diagnósticos, onde esta abordagem se adequou como nenhuma outra. Já os modelos de regras de inferência, são mais indicados para este tipo de problema: classificação simples, poucas regras e preferencialmente sem uso dos fatores de confiança.

Palavra-chave: Regras de Inferência, Redes Bayesianas, Raça.
Órgão financiador: ULBRA

¹² Acadêmicos do Curso de Sistemas de Informação do ILES Santarém

¹³ Professor do Curso de Sistemas de Informação do ILES Santarém

¹⁴ Orientadora do Curso de Pós-Graduação em Ciências da Computação - CTC-INE-UFSC

