



UNIVERSIDADE FEDERAL DE ALAGOAS
INSTITUTO DE COMPUTAÇÃO
PROGRAMA MULTIDISCIPLINAR DE PÓS-GRADUAÇÃO EM MODELAGEM
COMPUTACIONAL DE CONHECIMENTO

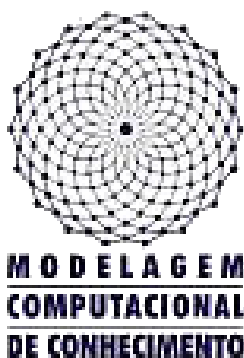
LEOPOLDO RAMOS DE OLIVEIRA

**UM MODELO PARA APOIAR A GESTÃO EDUCACIONAL DAS IES COM
DESCOBERTA DE CONHECIMENTO BASEADO NO PROCESSO DE
AUTOAVALIAÇÃO INSTITUCIONAL (SINAES)**

MACEIÓ
2012

LEOPOLDO RAMOS DE OLIVEIRA

**UM MODELO PARA APOIAR A GESTÃO EDUCACIONAL DAS IES COM
DESCOBERTA DE CONHECIMENTO BASEADO NO PROCESSO DE
AUTOAVALIAÇÃO INSTITUCIONAL (SINAES)**



Dissertação apresentada como requisito parcial
para obtenção do grau de mestre em
Modelagem Computacional de Conhecimento
pela Universidade Federal de Alagoas.

Orientador: **Dr. Patrick Henrique da Silva Brito**

MACEIÓ
2012

Catálogo na fonte
Universidade Federal de Alagoas
Biblioteca Central
Divisão de Tratamento Técnico
Bibliotecária: Fabiana Camargo dos Santos

048m Oliveira, Leopoldo Ramos de.

Um modelo para apoiar a gestão educacional das IES com descoberta de conhecimento baseado no processo de autoavaliação institucional (SINAES) / Leopoldo Ramos de Oliveira. – 2012.
126 f. : il.

Orientador: Patrick Henrique da Silva Brito.

Dissertação (Mestrado em Modelagem Computacional do Conhecimento) – Universidade Federal de Alagoas. Instituto de Computação. Maceió, 2012.

Bibliografia: f. 84-87.

Apêndices: f. 88-126.

1. Avaliação institucional. 2. Mineração de dados. 3. Apoio à tomada de decisão. 4. Processo de descoberta de conhecimento. 5. Instituições de ensino superior – Gestão. I. Título.

CDU: 004.89:37.07



Membros da Comissão Julgadora da Dissertação de Mestrado de Leopoldo Ramos De Oliveira, intitulada: “Um Modelo para Apoiar a Gestão Educacional das IES com Descoberta de Conhecimento Baseado no Processo de Autoavaliação Institucional (SINAES)”, apresentada ao Programa de Pós-Graduação em Modelagem Computacional de Conhecimento da Universidade Federal de Alagoas em 20 de dezembro de 2012, às 09h00min, na sala de aula do Mestrado em Modelagem Computacional de Conhecimento.

COMISSÃO JULGADORA

Prof. Dr. Patrick Henrique da Silva Brito

UFAL – Instituto de Computação

Orientador

Prof. Dr. Fábio Paraguaçu Duarte da Costa

UFAL – Instituto de Computação

Examinador

Prof. Dr. José Osman dos Santos

IFS – Instituto Federal de Sergipe

Examinador

Maceió, dezembro de 2012.

AGRADECIMENTOS

Agradeço a Deus pela saúde e coragem de concretizar este trabalho que representa uma realização pessoal.

Ao Instituto Federal de Sergipe (IFS) por acreditar no meu trabalho frente à CPA e em especial ao professor Alberto Aciole, que viabilizou a realização do seminário com o colégio de dirigentes para apresentação e discussão dos novos conhecimentos obtidos através desta pesquisa.

Ao professor Dr. Patrick Henrique Brito pela orientação, paciência, dedicação, competência e sugestões para enriquecer o presente trabalho de pesquisa.

Aos professores da banca, Prof. Dr. José Osman dos Santos e Prof. Dr. Fábio Paraguaçu Duarte da Costa, pela contribuição com comentários construtivos para o aprimoramento desta pesquisa.

A todos os professores do curso pelos exemplos e ensinamentos adquiridos e em Especial ao professor Dr. Fábio Paraguaçu pela viabilização deste MINTER e pela disposição em colaborar com críticas e sugestões ao longo do mestrado.

Aos colegas do mestrado Carlos Leopoldo, Dalton, Paulo Amaral, Sozzi, Acácio, Cristiane e em especial Fausto e Damião pelas discussões que muito contribuíram com este trabalho.

Aos meus amigos Adailton Novaes, Marcone, Rodolfo Cassaca, Lula, Ivonaldo Pacheco, Marcos César, Geilson, Jorge dos Santos e Welington Brito Pela amizade e palavras de incentivo para realizar este Mestrado.

Aos amigos Natanael Oliveira, Arnaldo Barreto pela amizade e participação efetiva nos momentos mais difíceis e conquistas da minha vida desde a época antes da graduação em matemática até o presente. Muito obrigado por tudo.

Ao amigo Welington Luz companheiro desde a época de república universitária que sempre repetia nos momentos turbulentos “depois da tempestade vem a bonança”. Muito obrigado por tudo.

Aos meus irmãos Paulo, Luiz e Francisca pela confiança e carinho a mim dedicados ao longo de minha vida, amo vocês.

Aos sobrinhos e familiares obrigado pelo incentivo.

Ao meu amigo e sobrinho Alex Paulo pelo incentivo e companheirismo de sempre.

Aos meus pais Valdir Ramos (in memorian) e Pedrina Maria de Oliveira pelo exemplo de vida que me deram e se cheguei até aqui devo a eles.

As minhas lindas filhas: Luiza com seu sorriso e bom humor que traz felicidade e Íngrid pelo apoio, companhia e carinho durante a realização desta jornada.

A minha amada esposa Valéria que me acompanha desde a graduação e esteve presente em todos os momentos difíceis e conquistas da minha vida (te amo!).

Enfim, se para concluir este trabalho tive que me ausentar da família, o mesmo deve servir de exemplo como paciência, perseverança e força de vontade para poder atingir metas na vida.

A todos, muito obrigado.

RESUMO

A **Autoavaliação institucional** é uma questão atual e de grande relevância no contexto da educação superior do Brasil. O problema desta pesquisa investigou como ajudar os gestores das IES a detectar as fragilidades na autoavaliação, tendo em vista o grande volume de dados oriundos dos formulários da avaliação do SINAES. A análise destas informações visa a apoiar a tomada de decisão por parte dos gestores. Na tentativa de responder tal questionamento, o objetivo precípua deste trabalho é elaborar um modelo para viabilizar a aplicação sistemática da política do Sistema Nacional de Avaliação da Educação Superior (SINAES). Na elaboração do referencial teórico desta pesquisa, procurou-se contextualizar a política definida pelo SINAES para implementar a autoavaliação e mostrar como os recursos da **Estatística** aliada ao **Processo de Descoberta de Conhecimento (KDD)** para coleta, análise e descobertas de novos conhecimentos podem contribuir no processo avaliativo das IES. A Metodologia utilizada consistiu na construção de instrumentos de avaliação para os discentes e gestores a fim de gerar as estatísticas e obter informações implícitas utilizando mineração de dados. A **mineração de dados** foi utilizada principalmente para obter a relevância dos atributos da base através do **software Rapid Miner**. O passo seguinte foi a elaboração da relevância dos atributos da base por um especialista em avaliação institucional, tendo como base documentos oficiais da IES. Os resultados obtidos nas três avaliações (discentes, gestores e especialista) foram consolidados num gráfico para análise das divergências e identificação de problemas para gerar um modelo de apoio à gestão visando a minimizar os pontos frágeis detectados. Em seguida foi realizado um seminário com o colégio de dirigentes da IES para avaliar os impactos na gestão face ao novo conhecimento obtido através da **Mineração dos dados** e estes consideraram os novos conhecimentos **importantes** como **ferramenta de apoio à tomada de decisão da IES**.

Palavras Chave: Avaliação institucional. Mineração de dados. Apoio à tomada de decisão. Processo de descoberta de conhecimento. Instituições de Ensino Superior – Gestão.

ABSTRACT

The Institutional Self-evaluation is a matter of great current relevance in the context of higher education in Brazil. The problem investigated in this research: How to help managers IES Detecting weaknesses in promoting self-assessment and corrective actions in the short, medium and long term. In attempting to answer this question the ultimate goal of this work is to develop a model to enable the systematic application of the policy of the National Assessment of Higher Education (SINAES). In developing the theoretical framework of this research sought to contextualize the policy defined by SINAES to implement self-evaluation and show how the combined resources of Statistical Process Knowledge Discovery (KDD) for collection, analysis and discovery of new knowledge can contribute to the process IES.A of evaluative methodology used was the construction of assessment tools for learners and managers to generate statistics and through this database to perform these mining and obtain the relevance of the attributes of a database by the software Rapid miner. The next step was the elaboration of the relevance of the attributes of the base by a specialist in institutional assessment based on official documents IES.Os results of the three assessments (students, managers and specialist) were consolidated in a graphic for analysis and identification of differences of problems to generate a model of recommendation to minimize the weaknesses detected. After that a seminar was held with the leaders of the college of IES to assess impacts in managing against the new knowledge obtained through data mining and new knowledge they considered important as a tool to support decision making of IES.

Keywords: Institutional assessment. Data Mining. Support decision making. Process of knowledge discovery. Higher Education Institutions - Management.

LISTA DE FIGURAS

Figura 1 - Etapas usuais de uma pesquisa quantitativa	30
Figura 2 - Valores de Z para alguns níveis de confiança	35
Figura 3 - Relação entre KDD Data Mining	37
Figura 4 - Interdisciplinaridade do processo KDD	38
Figura 5 - Etapas do processo de KDD.....	39
Figura 6 - Processo para apoio a decisão utilizando mineração de dados	45
Figura 7 - Etapas do processo de KDD.....	47
Figura 8 - Processo no Rapid Miner para construção da Árvore de Decisão .	54
Figura 9 - Acurácia do classificador sem os ajustes	54
Figura 10 - Processo no Rapid Miner para construção da Árvore de Decisão, obtenção dos pesos por relevância e a matriz de confusão	56
Figura 11 - Matriz de confusão para discentes	56
Figura 12 - Peso por relevância dos atributos da base dos discentes	57
Figura 13 - Matriz de confusão dos gestores	59
Figura 14 - Gráfico comparativo Discente X Gestor X Especialista	64
Figura 15 - Gráfico comparativo Especialista X Nota do Especialista	65
Figura 16 - Comparação entre os resultados peso discente x peso gestores x nota especialista.....	69
Figura 17 - Histograma	90

LISTA DE TABELAS

Tabela 1 - Valores de Z para alguns níveis de confiança.	35
Tabela 2 - Matriz de confusão para um problema de duas classes.	42
Tabela 3 - Avaliação geral de todos os cursos pelos discentes.	52
Tabela 4 - Medidas de tendência 53	53
Tabela 5 - Tabela do Rapid Miner. Matriz com pesos e normalização 58	58
Tabela 6 - Dados dos gestores..... 60	60
Tabela 7 - Dados do especialista 60	60
Tabela 8 - Matriz obtida pelo conhecimento do especialista 61	61
Tabela 9 - Comparação e relevância dos pesos obtidos 63	63
Tabela 10 - Ranking de Prioridades 66	66
Tabela 11 - Medidas de centralidade e dispersão (aluno, gestor, nota (especialista))..... 67	67
Tabela 12 - Exemplo de uma distribuição de frequência numérica por faixa . 88	88
Tabela 13 - Exemplo de uma distribuição de frequências numérica por valor 89	89
Tabela 14 - Exemplo de distribuição de frequência por categoria 89	89
Tabela 15 - Ferramentas que utilizam árvore de decisão 97	97
Tabela 16 - Dados consolidados 100	100
Tabela 17 - Qualificação docente 100	100
Tabela 18 - Carga horária das disciplinas 101	101
Tabela 19 - Carga horária total 102	102
Tabela 20 - Estrutura Curricular 103	103
Tabela 21 - Articulação de Conteúdos 104	104
Tabela 22 - Laboratórios de informática..... 105	105
Tabela 23 - Envolvimento do professor em pesquisa..... 106	106
Tabela 24 - Envolvimento do aluno em pesquisa..... 107	107
Tabela 25 - Qualificação docente a curto prazo 108	108
Tabela 26 - Qualificação docente a médio prazo 109	109
Tabela 27 - Qualificação docente a longo prazo 110	110
Tabela 28 - Laboratório de informática a curto prazo 111	111
Tabela 29 - Laboratório de informática a médio prazo 112	112
Tabela 30 - Laboratório de informática a longo prazo 113	113

Tabela 31 - Envolvimento dos professores em pesquisa a curto prazo	114
Tabela 32 - Envolvimento dos professores em pesquisa a médio prazo	115
Tabela 33 - Envolvimento dos professores em pesquisa a longo prazo	116
Tabela 34 - Envolvimento dos alunos em pesquisa a curto prazo	117
Tabela 35 - Envolvimento dos alunos em pesquisa a médio prazo	118
Tabela 36 - Envolvimento dos alunos em pesquisa a longo prazo	119

LISTA DE GRÁFICOS

Gráfico 1 - Histograma dos dados consolidados da pesquisa	53
Gráfico 2 - Visão do colégio de dirigentes da IES quanto aos atributos analisados	77
Gráfico 3 - Questão 9	78
Gráfico 4 - Questão 10	79
Gráfico 5 - Questão 11	80
Gráfico 6 - Questão 12	81
Gráfico 7 - Qualificação Docente	101
Gráfico 8 - Carga horária das disciplinas	102
Gráfico 9 - Carga horária total	103
Gráfico 10 - Estrutura Curricular	104
Gráfico 11 - Articulação de Conteúdos.....	105
Gráfico 12 - Laboratórios de informática.....	106
Gráfico 13 - Envolvimento do professor em pesquisa	107
Gráfico 14 - Envolvimento do aluno em pesquisa	108
Gráfico 15 - Qualificação docente a curto prazo.....	109
Gráfico 16 - Qualificação docente a médio prazo.....	110
Gráfico 17 - Qualificação docente a longo prazo	111
Gráfico 18 - Laboratório de informática a curto prazo	112
Gráfico 19 - Laboratório de informática a médio prazo	113
Gráfico 20 - Laboratório de informática a longo prazo	114
Gráfico 21 - Envolvimento dos professores em pesquisa a curto prazo	115
Gráfico 22 - Envolvimento dos professores em pesquisa a médio prazo	116
Gráfico 23 - Envolvimento dos professores em pesquisa a longo prazo	117
Gráfico 24 - Envolvimento dos alunos em pesquisa a curto prazo	118
Gráfico 25 - Envolvimento dos alunos em pesquisa a médio prazo	119
Gráfico 26 - Envolvimento dos alunos em pesquisa a longo prazo	120

SUMÁRIO

1 INTRODUÇÃO	14
1.1 Motivação e Definição do Problema	16
1.1.1 Como viabilizar o SINAES?	16
1.2 Justificativa	17
1.3 Objetivo Geral	18
1.4 Objetivos Específicos	19
1.5 Hipóteses	19
1.6 Estrutura do Trabalho	20
2 FUNDAMENTAÇÃO TEÓRICA	22
2.1 Contextualização do SINAES	22
2.1.1 Concepções de avaliação do SINAES	23
2.1.2 Princípios fundamentais do SINAES	24
2.1.3 Avaliação das instituições de ensino superior (IES)	25
2.1.4 Dimensões da avaliação institucional	25
2.1.5 Avaliação interna ou autoavaliação	26
2.1.6 Requisitos da autoavaliação	27
2.1.7 Etapas da avaliação interna	28
2.2 Obtenção dos Indicadores (Estatísticas) da Autoavaliação	30
2.2.1 Planejamento da pesquisa	31
2.2.2 População e amostra	31
2.2.3 Variáveis quantitativas e qualitativas	32
2.2.4 A coleta dos dados	32
2.2.5 Elaboração dos questionários (instrumentos) para coleta de dados.	33
2.2.6 Procedimento estatístico para o cálculo do tamanho mínimo da amostra	34
2.3 Descoberta de Conhecimento e Mineração de Dados	36

2.3.1 Etapas do processo de KDD	38
2.3.2 Conceito de mineração de dados	40
2.3.3 Tarefas de mineração de dados	40
2.3.4 Abordagem geral para resolução de um problema de classificação.	41
2.3.5 A ferramenta Rapid Miner	42
2.3.6 Trabalhos relacionados	43
3 METODOLOGIA PROPOSTA PARA APOIO A DECISÃO DAS IES	45
3.1 Processos para Apoio a Decisão Utilizando Mineração de Dados	45
4 RESULTADOS.....	50
4.1 Diretrizes para Elaboração e Preenchimento dos Questionários Aplicados aos Discentes e Gestores do IFS.	50
4.1.1 Cálculo do tamanho mínimo da amostra obtida com os discentes dos cursos superiores do IFS	51
4.2 O Processo de Descoberta de Conhecimento (KDD) – Execução da Mineração de Dados.....	53
4.3 Avaliação dos Documentos Oficiais do IFS	61
4.4 Comparação da Relevância dos Pesos Obtidos na Avaliação dos Discentes, Gestores e Especialista do IFS	62
4.4.1 Diagnóstico Geral	64
4.4.2 Classificação de Fragilidades.....	65
4.4.3 Diagnóstico detalhado	66
5 INFLUÊNCIAS DA SOLUÇÃO PROPOSTA NO APOIO A TOMADA DE DECISÃO DOS GESTORES. UM ESTUDO DE CASO COM O COLÉGIO DE DIRIGENTES DA IES.....	75
6 CONCLUSÃO E TRABALHOS FUTUROS	82
6.1 Limitações da Pesquisa.....	83
REFERÊNCIAS BIBLIOGRÁFICAS	84
APÊNDICES	88

1 INTRODUÇÃO

A área de Avaliação institucional tem sido objeto de estudo nas instituições de ensino superior buscando o autoconhecimento para direcionar políticas visando à melhoria contínua da qualidade do ensino além do cumprimento da legislação em vigor. No entanto, uma das críticas mais comuns no processo de avaliação das IES é a utilização de instrumentos que mostram apenas uma visão parcial e fragmentada da realidade institucional. Esse fato demonstra a dificuldade para extrair informações implícitas a partir de um grande volume de dados obtidos no momento da autoavaliação.

Neste trabalho é proposto analisar como a estatística descritiva aliada a mineração de dados podem contribuir com o aperfeiçoamento da autoavaliação institucional no âmbito universitário brasileiro.

Utilizando-se dos recursos mencionados e realizando uma sensibilização¹ com toda a comunidade acadêmica² quanto à importância da participação dos mesmos na construção do processo de autoavaliação é possível coletar dados confiáveis e através destes detectar os pontos fracos, e apontar possíveis sugestões de melhorias pautadas em dados que reflitam os anseios da comunidade acadêmica (SINAES, 2004). Dessa forma, espera-se contribuir com as IES³ nas suas tomadas de decisão visando a oferecer serviços educacionais que atendam as necessidades dos atores institucionais⁴.

Para delinear nosso interesse na autoavaliação institucional faremos um recorte da mesma a partir da criação do sistema nacional de avaliação do ensino superior (SINAES) que implementou a avaliação institucional sistemática, controlada pelo governo federal a partir do ano de 2004.

O foco dessa pesquisa visa a estudar como a estatística e a mineração de dados podem ajudar no processo da autoavaliação, bem como na coleta, organização, análise, descoberta de novos conhecimentos e interpretação dos dados obtidos no processo avaliativo e, dessa forma, fornecer elementos para nutrir o planejamento estratégico das IES.

¹ Para mais detalhes sobre sensibilização, leia Seção 2.1.7

² Comunidade acadêmica é formada por discentes, docentes, gestores, técnicos administrativos e a sociedade civil organizada.

³ IES-Instituição de Ensino Superior.

⁴ Atores institucionais neste contexto indicam todos que participam da autoavaliação.

Para o desenvolvimento do presente trabalho foi realizado um estudo na área de avaliação institucional, mineração de dados e também a criação de um protótipo de um sistema onde serão armazenados os dados obtidos dos questionários respondidos via intranet da instituição de ensino alvo da pesquisa, contando com a participação de toda a comunidade acadêmica. Ressaltamos que contamos ainda com uma bibliografia que contempla o tema em estudo, além da nossa vivência profissional em uma instituição pública de ensino a 18 anos, dos quais quatro, como coordenador da comissão própria de avaliação (CPA).

Esta instituição é o antigo CEFET (Centro Federal de Educação Tecnológica) que passou a Instituto Federal de Sergipe (IFS) desde 31 de março de 2008.

No IFS várias ações têm sido tomadas no caminho da qualidade e excelência em ensino, pesquisa e extensão. Dentre as ações, podemos citar as alterações nos documentos oficiais da IES como PDI (Plano de desenvolvimento institucional), PPI (Plano pedagógico institucional) e PPC (Plano pedagógico de curso) além da autoavaliação institucional com foco nas dez dimensões estabelecidas pelo SINAES que serão discutidas no Capítulo 2 desta pesquisa.

Para conduzir o processo de autoavaliação ou avaliação interna o IFS constituiu a Comissão Própria de Avaliação (CPA) de acordo como art. 11 da lei 10.861 de 14 de abril de 2004.

Os membros dessa comissão são nomeados pelo dirigente máximo da IES e na sua composição é assegurada a participação de todos os segmentos da comunidade acadêmica (discentes, docentes, técnico administrativo, egressos e comunidade civil organizada) (SINAES, 2004).

De acordo com a lei brasileira é vedada uma composição que privilegie a maioria absoluta de um dos segmentos e a CPA é um órgão autônomo em relação a conselhos e colegiados existente na IES (SINAES, 2004).

Na solução proposta, a coleta dos dados da avaliação é conduzida através do preenchimento de questionários online através do protótipo de sistema desenvolvido junto ao setor de TI da IES. De posse dos dados é realizada a análise estatística e a mineração dos mesmos com o objetivo de descobrir novos conhecimentos da base e auxiliar os gestores no processo de tomada de decisão. Nesse sentido, a CPA realizou um seminário com a participação de todos os atores da avaliação interna como uma forma de divulgar os resultados da pesquisa apontando potencialidades e fragilidades e as respectivas sugestões de ações para neutralizar os pontos fracos

evidenciados e, dessa forma, começar sensibilizar a comunidade para continuidade da autoavaliação que de acordo com a legislação em vigor as IES devem realizar a avaliação interna anualmente (SINAES, 2004).

Para atingir os objetivos traçados se faz necessário utilizar ferramentas da estatística e a mineração dos dados obtidos para obter melhor conhecimento da realidade da IES.

A avaliação interna ou autoavaliação das IES fornecem a oportunidade de obtenção de diagnósticos para auxiliar no processo de tomada de decisão através dos seus pontos fortes e fracos evidenciados no processo avaliativo. A avaliação interna deverá proporcionar um momento de reflexão a toda comunidade acadêmica e deixar claro o compromisso das IES com ensino pesquisa e extensão cumprindo assim, o seu papel científico e social (BELLONI ET AL, DIAS SOBRINHO, 1995).

1.1 Motivação e Definição do Problema

1.1.1 Como viabilizar o SINAES?

A avaliação das IES é regulamentada pelo SINAES e deve ser executada por exigência legal do MEC. A mesma possui um grande número de atividades a serem desenvolvidas com a participação de toda comunidade acadêmica que é convidada para conhecer os objetivos da avaliação e responder de maneira voluntária aos questionários onde todos têm a oportunidade de avaliar e serem avaliados contribuindo com críticas e sugestões no sentido de apontar pontos fortes e frágeis para uma melhoria contínua das IES nos seus mais variados campos de atuação. Esse trabalho é coordenado pelas CPAs que possui na sua composição representante de todos os segmentos da comunidade acadêmica e da sociedade civil organizada. A CPA é um órgão autônomo em relação aos demais órgãos das IES (SINAES, 2004).

Devido à grande quantidade de dados e informações com os quais as CPAs têm que lidar na coleta, tabulação, organização, análise e interpretação das dimensões avaliadas, torna-se necessário um amplo planejamento para utilização de recursos da estatística para auxiliar no referido trabalho, assim como, a Mineração de Dados para descobrir conhecimentos às vezes ocultos nos dados coletados.

Pelo exposto, está claro que a forma como cada IES vai lidar os dados da avaliação é um problema que necessita do auxílio de especialistas na área de tecnologia e estatística para evitar um viés nos resultados das pesquisas com a comunidade acadêmica e, dessa forma, interferir nos resultados obtidos na autoavaliação.

Tendo em vista o cenário apresentado, Como ajudar os gestores das instituições de ensino superior (IES) a detectar as fragilidades na autoavaliação e tentar minimizá-las a curto, médio e longo prazo?

1.2 Justificativa

O grande desafio do presente trabalho é construir um modelo que recomende ações para ajudar os gestores das IES a neutralizar as fragilidades detectadas na autoavaliação, posto que, a mesma é uma exigência legal do MEC e implica em credenciamento, credenciamento e autorização dos cursos superiores. Assim, o modelo em questão é uma ferramenta útil para as IES.

De acordo com a literatura sobre avaliação institucional há um consenso sobre as dificuldades encontradas por todas as IES em criar uma cultura de autoavaliação e quebrar estigmas de que avaliar normalmente está associado à punição.

Dentre as dificuldades encontradas durante o processo avaliativo, podemos citar:

- Grupos dentro das IES que por motivos às vezes políticos são do “contra” e não contribuem com a instituição;
- Resistência à avaliação;
- Dificuldade da CPA em sensibilizar toda a comunidade;
- Envolvimento de alguns servidores com a IES;
- Burocracia entre setores e dificuldade de articulação entre os mesmos;
- Investimento humano e financeiro no processo de autoavaliação;
- Sentimento de comprometimento com o desenvolvimento da IES;
- Receio de alguns atores⁵ da IES em participar da avaliação interna;

⁵ Atores neste contexto indicam comunidade acadêmica.

- Grande volume de dados, fruto da opinião dos diversos integrantes da IES (discentes, docentes e gestores);
- Consequente dificuldade de avaliação dos dados.

A construção de uma instituição de ensino superior deve ocorrer apoiada em discussões, ações e reflexões de acordo com suas características peculiares e, nesse contexto, um processo avaliativo busca viabilizar a interação construída pela própria comunidade acadêmica. Assim, a avaliação institucional contribuirá para que a IES repense sua missão de formar cidadãos críticos, reflexivos e preparados para o mercado de trabalho (SANCHES, 2004).

Face ao exposto, o sucesso da autoavaliação constitui um novo marco nas IES não apenas como cumprimento da exigência legal, mas na implementação de políticas que vão melhorar as Instituições de ensino superior, levando as mesmas a cumprirem com o seu papel social, conforme previsto nos seus documentos institucionais e oferecer serviços educacionais voltados às necessidades da comunidade acadêmica e, dessa forma, contribuir para o desenvolvimento da região onde está inserida.

1.3 Objetivo Geral

Uma instituição de ensino superior precisa constantemente repensar suas práticas, princípios, missão e o todo que a constitui, e para isto a autoavaliação é o caminho, pois tem o objetivo de romper visões centralizadoras e autoritárias (SANCHES, 2004).

O objetivo precípua deste trabalho é elaborar um modelo para viabilizar a aplicação sistemática da política do SINAES, de forma a minimizar os pontos frágeis das instituições de ensino superior e melhorar/manter as potencialidades detectadas, fazendo uso de um aparato tecnológico adequado.

Espera-se que este trabalho contribua para o aprimoramento do processo de autoavaliação nas instituições de ensino superior IES.

O processo de autoavaliação proporciona as IES o autoconhecimento, que deve ser evidenciado no momento da avaliação externa⁶, conduzida por especialistas designados pelo MEC conforme prevê o SINAES.

1.4 Objetivos Específicos

- Implantar um protótipo de sistema que possa ser utilizado para executar os modelos da autoavaliação, e gerar relatórios estatísticos;
- Utilizar algoritmos de mineração de dados, que possibilitem a geração de novas informações que apóiem a tomada de decisão dos gestores das IES com base nas diretrizes e exigências do SINAES;
- Avaliar o modelo de relatórios informativos propostos e ver o seu grau de importância das descobertas de informações não triviais, fruto da análise dos dados através de técnicas de mineração de dados e de acordo com o colégio de dirigentes da IES;
- Apoiar decisões dos gestores, em função da comparação dos critérios adotados pelos três atores institucionais (discentes, gestores e especialista) envolvidos no processo de autoavaliação.

1.5 Hipóteses

- Um modelo baseado em recursos da estatística para coleta, organização, análise e interpretação dos dados obtidos, possibilita a identificação de alguns indícios de problemas que recomende ações para ajudar os gestores das IES nas suas tomadas de decisões relacionadas aos resultados da autoavaliação prevista pelo SINAES;
- A utilização de mineração de dados pode ser extremamente relevante no domínio educacional para descoberta do conhecimento implícito, fruto da análise de dados relativos aos questionários da autoavaliação institucional prevista pelo SINAES;

⁶ Para entender avaliação externa leia secção 2.1.3.

- A comparação dos resultados da mineração de dados referentes aos atores institucionais pode indicar novos indícios de problemas que podem ser resolvidos pelos gestores das IES.

1.6 Estrutura do Trabalho

Esta pesquisa está estruturada em 6 Capítulos: O Capítulo 2 contempla a revisão da literatura sobre a contextualização do SINAES, a sua concepção de avaliação, princípios, avaliação interna, requisitos da avaliação, etapas da avaliação, obtenção das estatísticas da autoavaliação, planejamento da pesquisa, população e amostra, coleta dos dados, procedimentos para o cálculo do tamanho da amostra, processo de descoberta de conhecimento (KDD) e são apresentados trabalhos relacionados com o tema desta pesquisa.

No Capítulo 3 é apresentada a Metodologia proposta para apoio à decisão da IES através de uma descrição de como realizar cada etapa do presente trabalho. Também será criada uma matriz de referência através de um especialista em autoavaliação com base nos resultados das avaliações externas realizadas por comissões designadas pelo MEC na instituição pesquisada.

No Capítulo 4 são apresentados os resultados da mineração dos dados obtidos com a coleta através da autoavaliação do IFS e a Matriz de referência do especialista. Neste tópico, será feita uma discussão da metodologia proposta para identificação de fragilidades, baseada na análise dos dados da autoavaliação institucional a partir de comparações sistemáticas entre os dados dos discentes, gestores e do especialista do MEC. Durante a análise desses resultados, também são apresentadas discussões obtidas através da comparação das discrepâncias dos dados minerados, fruto da avaliação dos gestores, discentes e especialistas em autoavaliação, que tem como referência os resultados das avaliações externas realizadas por comissões designadas pelo INEP⁷. Todos esses dados foram relativos aos cursos superiores da referida IES.

O capítulo 5 apresenta a avaliação da metodologia proposta (resultados), que foi realizada em um cenário real, durante um seminário com o colégio de dirigentes

⁷ INEP – Instituto Nacional de Pesquisa Anísio Teixeira

da IES onde foi avaliado o impacto dos novos artefatos gerados fruto da mineração de dados, na decisão dos gestores a curto, médio e longo prazo.

No Capítulo 6 são apresentadas a conclusão, as contribuições, as perspectivas de trabalhos futuros e as limitações desta pesquisa.

2 FUNDAMENTAÇÃO TEÓRICA

2.1 Contextualização do SINAES

O SINAES (Sistema Nacional da Avaliação do Ensino Superior) foi construído a partir do trabalho da comissão especial de Avaliação (CEA), designada pela portaria da SESU em 28 de abril de 2003. Desse trabalho surgiu o relatório final do Sistema Nacional de Avaliação da Educação Superior que expressa o resultado da experiência em avaliação da educação desde 1980 até o amplo processo de avaliação institucional na década de 1990 através do PAIUB (programa nacional de avaliação institucional das universidades brasileiras) (CONAES, 2004).

De acordo com o artigo 6º da lei 10.861 Compete a CONAES (Comissão Nacional de avaliação da educação Superior) estabelecer diretrizes, critérios e estratégias para o processo de avaliação da educação superior.

A nova proposta possui como características fundamentais, a avaliação institucional, como centro do processo avaliativo e a integração de diversos instrumentos com uma visão global, buscando o respeito à identidade e a diversidade institucionais, possibilitando analisar a realidade e a missão de cada IES (CONAES, 2004).

O novo sistema de avaliação em tela abrange todas as IES (instituições de ensino superior) e ocorre de maneira permanente com finalidades construtiva e formativa. Esse sistema, devido à continuidade, visa desenvolver nas IES uma cultura de avaliação onde todos os atores institucionais são levados a participar do processo como sujeitos comprometidos com as mudanças em busca da qualidade da educação superior (CONAES, 2004).

A partir da lei nº 10861 de 14 de abril de 2004, o sistema nacional da avaliação do ensino superior (SINAES) foi instituído com o objetivo de promover uma melhora contínua na qualidade da educação superior, orientar a expansão da sua oferta e aprofundar os seus compromissos e responsabilidades sociais (SINAES, 2004).

O SINAES possui uma concepção de Avaliação que recupera as seguintes finalidades essenciais da avaliação:

“Ultrapassa a simples preocupação com desempenhos ou rendimentos estudantis, buscando os significados mais amplos da formação profissional;

Explicita a responsabilidade social da educação superior, especialmente quanto ao avanço da ciência, à formação da cidadania e ao aprofundamento dos valores democráticos;

Supera meras verificações e mensurações, destacando os significados das atividades institucionais, não apenas do ponto de vista acadêmico, mas também quanto aos impactos sociais, econômicos, culturais e políticos;

Aprofunda a ideia da responsabilidade social no desenvolvimento da IES, operando como processo de construção, com participação acadêmica e social, e não como instrumento de checagem e cobrança individual;

Valoriza a solidariedade e a cooperação e não a competitividade e o sucesso individual” (CONAES, 2004).

2.1.1 Concepções de avaliação do SINAES

O SINAES como parte de uma política de estado que visa à melhoria contínua da educação superior e possui na sua essência seguintes concepções de Avaliação:

- Avaliação como instrumento de política educacional: A partir da lei nº 9.394/96, de Diretrizes e Bases da Educação Nacional, em seu art.9º, inciso VI explicitou responsabilidade da união em “assegurar processo nacional de avaliação do rendimento escolar no ensino fundamental, médio e superior, em colaboração com sistemas de ensino, objetivando a definição de prioridades e a melhoria da qualidade do ensino”.

Para atender a lei mencionada o SINAES é fundamental no processo avaliativo, pois é uma forma das instituições de ensino superior públicas e privadas prestarem contas à sociedade das suas políticas educacionais no sentido de reformularem o projeto de desenvolvimento de cada instituição pautados por princípios de gestão democrática, responsabilidade social e o compromisso científico cultural da IES (CONAES, 2004).

- Avaliação institucional e efeitos regulatórios: A avaliação institucional é um processo desenvolvido pela comunidade acadêmica (professores, alunos, egressos, técnicos administrativos) e sociedade civil (parceiros

externos) com o objetivo de promover uma avaliação autônoma garantindo a imparcialidade e a qualidade acadêmica no ensino, pesquisa, extensão, e no cumprimento de sua pertinência e responsabilidade social de acordo com a missão da IES. A Regulação compreende o processo documental e a visita in loco por especialistas designados pelo MEC que farão a avaliação da IES com vistas ao credenciamento, autorização e reconhecimento de cursos. Do exposto, fica claro que a avaliação é um instrumento que a sociedade e o poder público dispõem para verificar a qualidade dos serviços educacionais oferecidos pelas IES em consonância com a sua missão acadêmica (CONAES, 2004).

- Avaliações, participação e ética na educação superior: O SINAES prevê a participação de toda comunidade acadêmica no processo de autoavaliação, pois todos são convidados a terem a oportunidade de apontar pontos fracos e fortes das IES e dessa forma se envolver nas ações avaliativas, respeitando os papéis, competências profissionais, políticas, administrativas e éticas das diversas categorias envolvidas na avaliação. É papel da comunidade acadêmica sugerir e acompanhar as ações dos gestores para neutralizar os pontos frágeis das IES evidenciados na autoavaliação (CONAES, 2004).

2.1.2 Princípios fundamentais do SINAES

De acordo com o roteiro de autoavaliação, o SINAES possui os seguintes princípios, a saber (SINAES, 2004):

- A responsabilidade social da IES, buscando a qualidade da educação superior e respeitando a diversidade que existe no sistema educação superior;
- O respeito à identidade, à missão e à história das instituições, tendo em vista as realidades locais de cada instituição;
- A globalidade institucional, ou seja, a IES deve ser avaliada a partir de indicadores de qualidade percebidos em sua relação orgânica e não apenas de forma isolada.

- A continuidade do processo avaliativo como instrumento de política educacional não apenas para cumprir a legislação em vigor, mas para que cada instituição e o sistema de educação superior incorporem a avaliação como instrumento de gestão para auxiliar a tomada de decisão.

2.1.3 Avaliação das instituições de ensino superior (IES)

O SINAES integra três modalidades de instrumentos de avaliação em momentos distintos:

- A avaliação das instituições de ensino superior (AVALIES) dividida em duas etapas. A primeira é o processo de autoavaliação coordenado pela CPA de cada IES e a segunda é a avaliação Externa realizada por comissões enviadas pelo INEP atendendo as diretrizes traçadas pela CONAES.
- O segundo momento é Avaliação dos cursos de graduação (ACG) através de instrumentos e visitas in loco de comissões externas designadas pelo INEP.
- O terceiro momento é Avaliação do desempenho dos Estudantes (ENADE). Nesta avaliação, os estudantes no final do 1º ano e do último ano do curso realizam provas com o objetivo de avaliar a evolução do aluno durante o curso. A escolha dos alunos é realizada de forma aleatória e a CONAES define as áreas que participarão do exame (SINAES, 2004).

2.1.4 Dimensões da avaliação institucional

A lei nº 10.861/2004, no seu Art. 3º, estabelece as dimensões que devem ser o foco da avaliação institucional para garantir um processo avaliativo de acordo com a especificidade de cada IES.

As dimensões previstas na lei mencionada são listadas a seguir:

- A missão e o Plano de Desenvolvimento Institucional (PDI);
- A política para o ensino;

- A responsabilidade social da instituição;
- A comunicação com a sociedade;
- As políticas de pessoal;
- Organização e gestão da instituição;
- Infraestrutura física;
- Planejamento e avaliação;
- Políticas de atendimento aos estudantes;
- Sustentabilidade financeira;
- Outras dimensões – incluem outros itens considerados relevantes para a instituição, tendo em vista a compreensão e construção da sua identidade institucional, suas especificidades e sua missão.

Nesta pesquisa, o foco é discutir apenas a autoavaliação e o trabalho realizado pela CPA na coleta, discussão, análise, procedimentos, descoberta de novos conhecimentos, relativo aos cursos superiores, bem como estes resultados podem ser utilizados como ferramenta de apoio gestão da IES.

2.1.5 Avaliação interna ou autoavaliação

A avaliação interna ou autoavaliação é um processo que permiti que a IES construa o conhecimento sobre sua realidade com o objetivo de cumprir com qualidade o seu papel no contexto social onde está inserida. A análise e interpretação dos dados obtidos na avaliação visam a detectar potencialidades e fragilidades da IES e servir como referencial para a tomada de decisão no sentido de corrigir os pontos fracos da instituição e implementar políticas educacionais que atendam as demandas da comunidade. Esses dados também são utilizados para retroalimentar o planejamento estratégico da IES (SINAES, 2004).

“A avaliação é um processo cíclico, criativo, e renovador de análise, interpretação e síntese das dimensões que definem a IES” (SINAES, 2004).

Segundo Sanches e Raphael (2006)

“A Avaliação Institucional é, por sua essência, um processo global no qual se revisa o que foi planejado e se constrói, continuamente, a IES, mediante processo de autoconhecimento. Nessa análise contínua, deve-se considerar principalmente a missão da instituição e o contexto em que está inserida, bem como sua trajetória histórica. Não é um processo que, isolado, resolva

todos os problemas da instituição, mas que contribui para a sinalização dos mesmos e a efetiva reflexão sobre as alternativas mais eficazes.”

Quando o processo de avaliação é bem planejado, revela potencialidades e fragilidades da IES auxiliando a busca de novas direções e elementos para o planejamento estratégico da instituição (SANCHES, 2007).

Quando a IES é pública, a sociedade deve acompanhar como e onde estão sendo aplicados os recursos, pois, os mesmos são frutos dos impostos pagos por ela. Se a IES é privada, a comunidade está cada vez mais exigente quanto aos investimentos que ela faz. O processo de autoavaliação leva as IES a terem mais credibilidade e reconhecimento perante a sociedade que faz parte (SANCHES; RAPHAEL, 2006).

2.1.6 Requisitos da autoavaliação

Para que a autoavaliação alcance bons resultados é imprescindível que as seguintes condições sejam atendidas em sua totalidade (SINAES, 2004):

- Existência de uma equipe de coordenação com participação de todos os atores institucionais;
- Compromisso explícito dos gestores da IES em utilizar os resultados da autoavaliação através de ações decorridas com base nos pontos fracos evidenciados no processo avaliativo;
- Informações, coleta e dados sejam confiáveis, sendo necessária a utilização de recursos tecnológicos para dar credibilidade ao processo avaliativo.

Se algum dos itens mencionados acima não for atendido a avaliação pode ser comprometida. Por exemplo, se a IES coletou dados confiáveis e a CPA sugeriu medidas para neutralizar as fragilidades encontradas, mas os gestores não utilizaram os resultados de forma adequada ou se ignorar tais resultados, a eficácia da autoavaliação estará comprometida, podendo a IES ter dificuldades no momento da avaliação externa através de comissões designadas pelo INEP.

2.1.7 Etapas da avaliação interna

Os resultados da autoavaliação serão submetidos aos especialistas no momento da avaliação externa designada pelo MEC. assim, é preciso que esses dados reflitam a realidade da IES e que todas as ações decorridas face às avaliações internas estejam bem evidenciadas no relatório de autoavaliação e sejam amplamente divulgadas para comunidade. Recomendações de especialistas no momento da visita externa podem ajudar a IES a consolidar o seu processo de autoavaliação (SINAES, 2004).

Para garantir um processo de avaliação interna que atenda as exigências do MEC é necessário que ocorra as seguintes etapas:

- **Preparação:** De acordo com o art.11 da lei 10861/04, as IES devem constituir a Comissão Própria de Avaliação (CPA) com o objetivo de coordenar e articular o processo interno de avaliação e disponibilizar os resultados para comunidade acadêmica. Todas as CPAs são cadastradas no INEP representam o elo entre a IES e o INEP;
A composição da CPA deve ter representantes de toda comunidade acadêmica para dar legitimidade à comissão e, conseqüentemente ao processo de autoavaliação (SINAES, 2004).
- **Planejamento:** As CPAs devem elaborar um planejamento de todas as ações para a implantação da avaliação interna e todas as etapas devem ser amplamente discutidas com a comunidade. Esta avaliação leva em consideração as características das IES e todos os atores institucionais têm a oportunidade de participar do processo de maneira voluntária (SINAES, 2004);
- **Sensibilização:** A sensibilização tem por objetivo promover o envolvimento de toda comunidade e é realizada através de seminários, reuniões, palestras, entre outros. Esta etapa deve ocorrer tanto no início como na continuidade das avaliações, pois o processo de autoavaliação é contínuo e sempre haverá novos sujeitos participando do processo avaliativo com discentes, docentes ou técnico-administrativo;

- Desenvolvimento: No desenvolvimento da avaliação interna devem ser concretizadas as atividades planejadas, dentre elas destacamos (SINAES, 2004):
 - Reuniões para debates e implementação de estratégias de sensibilização da comunidade acadêmica;
 - Sistematização de ideias ou demandas fruto das discussões ocorridas nas reuniões;
 - Apresentação da proposta de avaliação interna e do SINAES;
 - Construção de instrumentos para coleta dos dados;
 - Definição da metodologia para análise e interpretação dos dados estatísticos obtidos no processo avaliativo;
 - Definição dos grupos e condições de trabalho: espaço físico, número de horas dedicadas à pesquisa entre outras;
 - Elaboração de relatórios e discussões dos resultados com a comunidade e publicação dos mesmos.
- Consolidação: Nesta etapa é feita a elaboração e divulgação do relatório final com análises de todas as dimensões avaliadas na IES previstas nos SINAES.
- Relatório: O relatório final de avaliação interna deve mostrar os resultados das discussões e análises dos dados obtidos bem como a avaliação dos cursos e desempenho dos estudantes quando estiverem disponíveis. Este relatório deve ser claro e apresentar sugestões de natureza administrativa pedagógica e técnico-científica a serem implementadas para neutralizar as fragilidades detectadas na autoavaliação (SINAES, 2004).
- Divulgação: O objetivo da divulgação é dar continuidade a avaliação interna e realizar a discussão pública dos resultados através de reuniões, seminários e outros. Nesse momento as ações oriundas do processo avaliativo devem ser evidenciadas para comunidade acadêmica e, assim, prestar contas com a sociedade sobre o futuro da IES (SINAES, 2004).
- Balanço Crítico: Ao final do processo avaliativo é necessário um balanço crítico de todas as dificuldades que ajudaram a implementar

ações futuras para melhorar a avaliação interna. É Através da avaliação interna, que as IES constroem o autoconhecimento institucional e cumprirão uma importante etapa para a avaliação externa prevista no SINAES (SINAES, 2004).

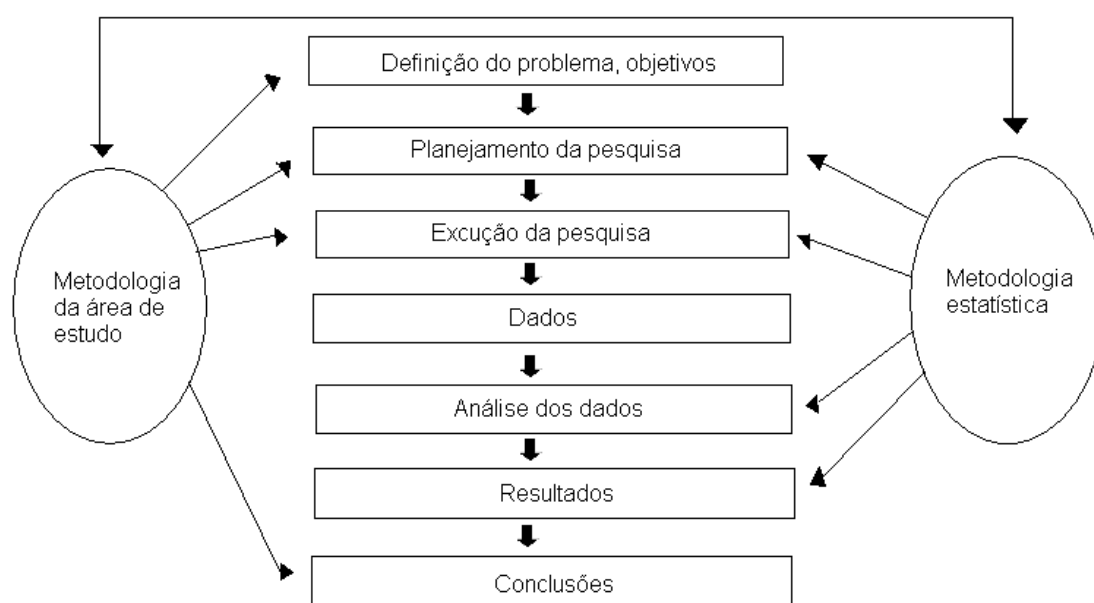
O Processo de avaliação institucional deve deixar claro para comunidade acadêmica que o mesmo não visa punição ou premiação dos atores participantes (GASPARETO, 1999).

2.2 Obtenção dos Indicadores (Estatísticas) da Autoavaliação

O trabalho proposto sugere a utilização da estatística descritiva para coleta, organização e análise dos dados da autoavaliação. Segundo Barbetta (2007) nas pesquisas científicas precisamos coletar dados e a partir dos mesmos obter e generalizar conclusões. Para que os resultados da pesquisa sejam confiáveis é necessário que a coleta e a análise sejam feitas de forma criteriosa e objetiva.

Ainda Segundo o autor, as principais etapas de uma pesquisa que envolve levantamento e análise de dados são mostradas na Figura 2.

Figura 1 - Etapas usuais de uma pesquisa quantitativa



Fonte: Barbetta, 2007.

As técnicas estatísticas normalmente são aplicadas no momento da análise dos dados, no entanto, a metodologia estatística deve ser utilizada em todas as etapas da pesquisa, pois para obter informações confiáveis os dados devem ser coletados de maneira adequada, ou seja, a qualidade da informação depende da qualidade dos dados coletados (Barbetta, 2007).

Foram objeto do presente estudo os seguintes instrumentos (questionários) utilizados na avaliação interna do IFS.

- Avaliação dos cursos superiores da IES pelo corpo discente;
- Avaliação dos cursos superiores da IES pelos gestores;
- Avaliação dos cursos superiores pelo Especialista em autoavaliação;
- Avaliação pelo colégio de dirigentes da IES quanto aos novos conhecimentos obtidos após a mineração dos dados (MD) e a comparação das divergências entre a avaliação dos discentes, gestores e o especialista em avaliação institucional.

2.2.1 Planejamento da pesquisa

Como o problema da pesquisa está bem definido e os seus objetivos são claros, os instrumentos utilizados (questionários) foram elaborados em discussão com a comunidade acadêmica que pode participar com sugestões e críticas com o objetivo de tornar os questionários abrangentes, com rigor técnico, mas de fácil entendimento de todos os segmentos da IES que participaram da pesquisa conforme diretrizes traçadas pela CONAES. No Capítulo 4 será discutida a metodologia utilizada na criação dos instrumentos da autoavaliação.

2.2.2 População e amostra

Segundo Freund e Simon (2000) se um conjunto de dados possui todas as observações possíveis de um determinado fenômeno, o mesmo pode ser chamado de população e, um conjunto que representa uma parte (subconjunto não vazio) da população é referenciado como amostra. Quando a população em estudo é grande trabalha-se com amostras devido ao custo, tempo de execução da pesquisa e confiabilidade dos resultados obtidos. Um exemplo comum da utilização de amostras

para descrever a tendência da população é o caso das pesquisas eleitorais onde as amostras são definidas de acordo com critérios estatísticos e têm boa confiabilidade.

De acordo com Barbetta (2007) população acessível ou simplesmente população é o conjunto que contém todos os elementos que interessam ao estudo de um determinado fenômeno e que podem ser observados quanto às variáveis (características) que se pretende levantar.

2.2.3 Variáveis quantitativas e qualitativas

Segundo Sullivan e Michael (2006), a Estatística pode ser vista Como a ciência que se ocupa da coleta, organização, análise e interpretação de informações. Nestas observações coletam-se dados através de medidas ou contagem. Neste contexto, uma Variável é uma característica que pode ser mensurável. Ainda de acordo com os autores, uma variável é dita discreta quando pode assumir apenas um conjunto finito de valores e contínua quando pode assumir todo e qualquer valor real num intervalo. Como exemplo de variáveis discretas e contínuas pode-se citar número de celulares vendidos por uma empresa e o salário dos funcionários desta, respectivamente.

2.2.4 A coleta dos dados

Para coletar os dados foi utilizada uma amostra aleatória simples com discentes e gestores do IFS com o objetivo de obter dados capazes de gerar conhecimento sobre o todo institucional conforme prevê a política do SINAES.

Segundo Stevenson (2001) uma amostra é dita aleatória quando cada elemento da população possui a mesma chance de participar da amostra. Ou seja, uma Amostra de uma população discreta é uma amostra cuja probabilidade de Extrair qualquer um dos **N** elementos numa única prova é igual a $1/N$. De modo análogo, grupos com a mesma quantidade de elementos têm a mesma chance de serem incluídos na amostra.

O presente trabalho lida com uma amostra aleatória simples cujos dados foram obtidos a partir de uma população discreta.

2.2.5 Elaboração dos questionários (instrumentos) para coleta de dados.

De acordo com (Barbetta, 2007) construção dos questionários é uma etapa fundamental do processo avaliativo e deve ser executada com rigor tendo em vista a população a ser estudada e os objetivos da pesquisa. Outro aspecto importante é que os questionários sejam construídos para facilitar a análise dos dados, e mesmo abrangentes, não sejam extensos para não perder a confiabilidade das respostas obtidas.

Ainda de acordo com autor na elaboração dos instrumentos (questionários) da pesquisa é importante o planejamento de como as utilizar as respostas para atender aos questionamentos da pesquisa e os seguintes procedimentos devem ser observados:

- Separar as características (variáveis a serem levantadas) que interessam ao estudo da pesquisa;
- Fazer uma revisão bibliográfica com o objetivo de verificar as possíveis formas de mensurar as variáveis em estudo;
- Estabelecer a maneira de mensuração das variáveis de acordo com o objetivo da pesquisa;
- Elaborar perguntas para cada variável a ser observada;
- Verificar se a pergunta está suficiente clara através de um pré-teste respondido por elementos da população em estudo. Nesta etapa também é importante observar se a pergunta não está induzindo a alguma resposta ou se não é óbvia.

Segundo Richardson (1989) após a elaboração preliminar dos questionários que serão utilizados para coleta de dados devem ser realizadas as seguintes etapas:

- Sensibilizar todas as pessoas envolvidas no processo avaliativo mostrando objetivo da pesquisa;
- Discutir como responder as perguntas do instrumento de avaliação visando obter respostas conscientes contribuindo para obtenção de dados confiáveis;
- Realizar o pré-teste da pesquisa utilizando uma amostra reduzida dos participantes do processo para corrigir eventuais falhas e assim evitar vieses contidos no instrumento de avaliação.

2.2.6 Procedimento estatístico para o cálculo do tamanho mínimo da amostra

Na realização das avaliações internas (autoavaliação) é importante a utilização de amostras para coletar os dados da pesquisa. Esse fato demonstra a importância dos recursos da Estatística para evitar vies no levantamento de dados obtidos a partir de uma amostra.

Segundo Carbonari et al (2007) os instrumentos utilizados na coleta de dados serão válidos para uma amostragem quando:

- O objetivo e o que se quer medir na pesquisa está bastante claro;
- A população a ser estudada está bem definida.

O primeiro passo do pesquisador para calcular o tamanho da amostra é especificar o quanto ele admite errar na avaliação dos parâmetros em análise. Esse erro é chamado de Erro Amostral e é definido Como a diferença entre uma estatística e o parâmetro que se quer estimar (BARBETTA, 2007).

Para coletar os dados consideramos uma amostra aleatória, que de acordo com o teorema do Limite Central a distribuição da média amostral tende para uma distribuição normal à medida que o tamanho n da amostra tende ao infinito. Segundo (Spiegel, 2004) a distribuição amostral das médias tende para uma distribuição normal para $n > 30$.

De acordo com Barbetta et al (2004), Stevenson (2001) o cálculo para determinar o tamanho n da amostra quando o parâmetro de interesse é a média pode ser determinado através da fórmula:

$$n \geq \frac{z^2 \sigma^2}{E_o^2} \quad \text{Inequação 1}$$

Considerando:

E_o - erro amostral máximo tolerado;

Z - é o nível de confiança desejado para amostras normalmente distribuídas;

σ^2 - é o desvio padrão e n_o é o menor n que satisfaz a Inequação 1.

Ainda de acordo com autores mencionados acima a dificuldade operacional para calcular o tamanho da amostra é o valor da variância populacional que normalmente não é conhecida. Nesse caso o valor da variância pode ser avaliado por

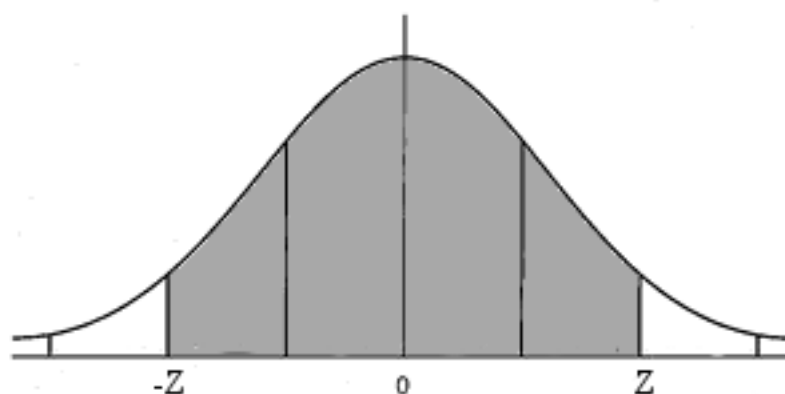
meio de estudos anteriores ou pela experiência do engenheiro do conhecimento. Outra possibilidade para contornar essa dificuldade é o cálculo da variância através de uma amostra piloto.

De acordo com Barbetta et al (2004) se a população for infinita o tamanho da amostra é dado por $n=n_0$. Caso o tamanho N da população seja conhecido e não seja muito grande, o tamanho da amostra é dado por:

$$n = \frac{N \cdot n_0}{N + n_0} \quad \text{Equação 1}$$

Segundo Rumsey (2009), Barbetta (2007) e Spiegel (2004), o nível de confiança desejado nas pesquisas pode ser obtido conforme mostra a Figura 2.

Figura 2 - Valores de Z para alguns níveis de confiança



Fonte: Spiegel, 2004.

Tabela 1 - Valores de Z para alguns níveis de confiança.

Área	0,800	0,900	0,950	0,980	0,990	0,995	0,998
Z	1,282	1,645	1,960	2,326	2,576	2,807	3,090

Fonte: Spiegel, 2004.

É comum nas pesquisas adotar o nível de confiança de 95% obtendo assim, $Z = 1,96$, conforme tabela1.

2.3 Descoberta de Conhecimento e Mineração de Dados

Nas duas últimas décadas tem ocorrido um aumento significativo Na quantidade de dados armazenados em meio eletrônico. No entanto, muitas informações úteis que podem ajudar os gestores de organizações que geram e /ou coletam dados são de difícil acesso e compreensão, e dessa forma, não os auxiliam na tomada de decisão (MARTINHAGO, 2005).

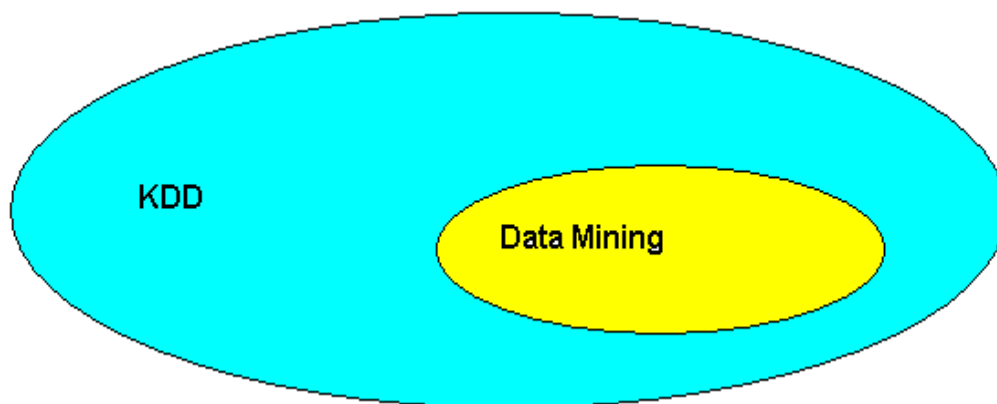
De acordo com Martinhago (2005) as organizações precisam ter formas de acesso ao grande volume de informações escondidas em seus sistemas transacionais, pois, às vezes, informações importantes não são percebidas nos mesmos. O conhecimento de toda a base de dados auxiliará no processo de tomada de decisão sendo uma maneira eficiente das organizações se manterem competitivas no mercado. Assim, espera-se que as técnicas e ferramentas de análise de dados automatizadas possam ser utilizadas na descoberta de conhecimento e tomadas de decisão. Nesse contexto, está a descoberta de conhecimento em banco de dados (Knowledge Discovery in Databases-KDD), sendo a Mineração de dados (Data Mining) a principal etapa.

Em um banco de dados pode haver além das informações triviais de domínio dos usuários diversos, padrões e comportamentos que não são de domínio dos mesmos. Para descobrir informações não triviais (implícitas) foi criado o KDD, com o objetivo de transformar os dados em informações úteis para tomada de decisão (FAYYAD, 1996).

O processo KDD analisa conjuntos de dados buscando a descoberta de padrões e tendências, e a Mineração de Dados é a sua principal etapa, pois através de algoritmos concebidos para este fim, informações implícitas são descobertas FAYYAD et al (1996).

Segundo Carvalho (2002) o KDD descreve todo o processo de extração de conhecimento em uma base de dados e o termo Mineração de Dados (MD) refere-se à etapa principal deste processo. A Figura 3, mostra como pode ser visualizada a relação entre o KDD e a MD. Muitas vezes esses conceitos são interpretados equivocadamente, como sinônimos.

Figura 3 - Relação entre KDD Data Mining



Fonte: Carvalho, 2002.

Segundo Fayyad et al (1996), o processo de descoberta é uma ferramenta importante para descobrir conhecimento, e deve ser utilizado, preferencialmente, onde se especula a existência de conhecimento novo. Novos conhecimentos geram diferencial competitivo, e são os alvos mais procurados por empresas que trabalham com conhecimento/pesquisa. De acordo com Adriaans e Zantinge (1996), a descoberta de conhecimento é interdisciplinar e pode ser realizada através da combinação de ferramentas diversas: Estatística, matemática, banco de dados, aprendizado de máquina, sistemas especialistas, reconhecimento de padrões, etc. O KDD combina essas ferramentas com o objetivo de extrair conhecimento e torná-lo útil para tomada de decisão. A Figura 4 mostra como KDD pode se relacionar com as áreas mencionadas.

Figura 4 - Interdisciplinaridade do processo KDD



Fonte: Adaptado de Adrians e Zantinge, 1996.

2.3.1 Etapas do processo de KDD

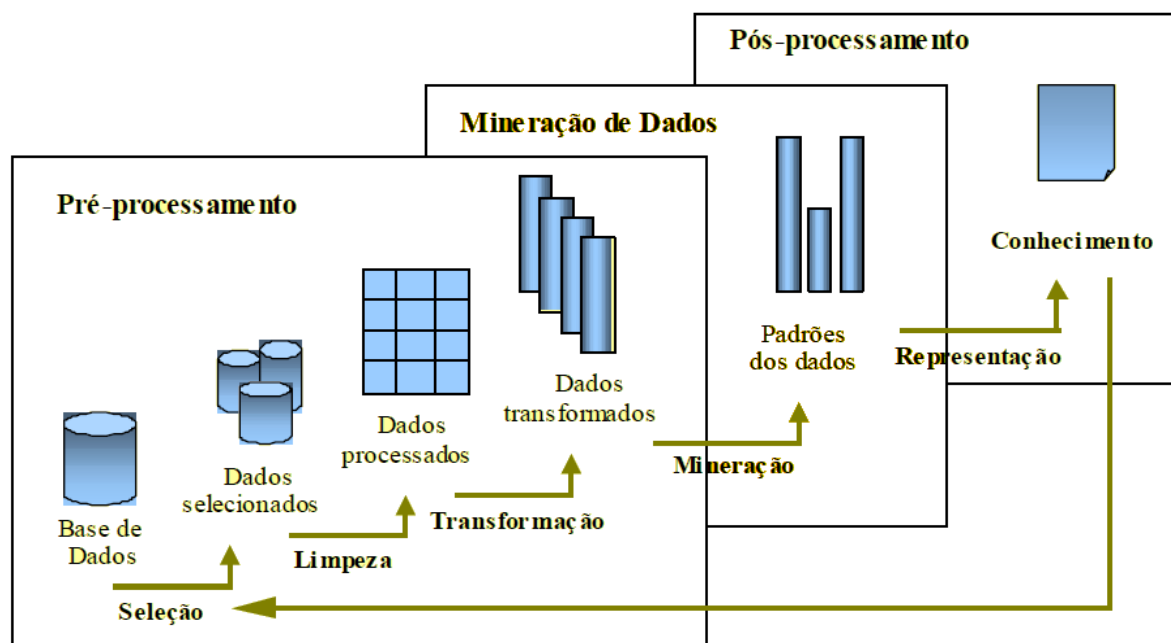
Esse processo transforma os dados em informações, que podem ser analisadas gerando conhecimento útil na tomada de decisões. O processo de KDD é composto pelas atividades observadas na figura 5.

No processo de KDD, o primeiro passo é ter conhecimento sobre o domínio de aplicação, identificar o problema e ter bem definidos os objetivos que se quer atingir. No início do processo os dados ainda são brutos e ao final ocorre a extração do conhecimento da base em estudo (MARTINHAGO, 2005).

As etapas do KDD consistem em:

- A Seleção de dados: consiste na identificação das informações que serão trabalhadas na base de dados necessárias para gerar a descoberta de conhecimento (GOLDSCHMIDT; PASSOS, 2005).

Figura 5 - Etapas do processo de KDD



Fonte: FAYYAD, 1996.

- Pré-Processamento e limpeza: Nesta etapa é realizada uma avaliação dos dados com o objetivo de detectar inconsistência das informações; Como exemplos, dados faltantes ou duplicados (OLIVEIRA, 2000).

Segundo Kumar, Steinbach e Tan (2006) a finalidade do pré-processamento é a transformação dos dados brutos de entrada em um formato que permita análises posteriores. Ainda de acordo com os autores, devido à forma como os dados são coletados e armazenados o pré-processamento pode ser o passo mais demorado e trabalhoso no processo geral de descoberta de conhecimento, tendo em vista, que nesta etapa, ocorre a limpeza dos dados para remoção de ruídos e observações duplicadas, bem como, a seleção dos registros e características identificados como importantes para mineração dos dados.

- Transformação dos dados: Nesta etapa é importante a participação de um especialista na área visando dar mais confiabilidade ao processo (OLIVEIRA, 2000); Neste trabalho não foi necessário realizar transformação, pois os dados já estavam normalizados;
- A Mineração de dados: é a principal etapa do processo de KDD; Neste momento vários experimentos são realizados com algoritmos para extrair padrões dos dados analisados ou gerar regras que

descrevam o comportamento da base de dados em estudo (BERRY e LINOFF,1997). Esses padrões podem ser extraídos através de técnicas de classificação, agrupamento ou sumarização (FAYYAD, 1996);

- Interpretação e avaliação dos resultados: Esta etapa consiste em validar o conhecimento extraído da base de dados através da visualização, análise e interpretação dos dados. Em geral, o engenheiro do conhecimento e o especialista do domínio de aplicação analisam os resultados abstraídos para filtrar as informações que serão úteis para tomada de decisão (GOLDSCHIMIDT; PASSOS, 2005).

Caso os resultados não sejam satisfatórios, é necessário retornar a etapa de mineração de dados ou qualquer outra anterior. O conhecimento só é extraído após a avaliação e validação dos resultados por um especialista no problema analisado (MARTINHAGO, 2005).

2.3.2 Conceito de mineração de dados

O KDD é um processo completo de descoberta de conhecimento, enquanto a mineração de dados é a sua principal etapa na qual são aplicados algoritmos para produzir padrões sobre uma base de dados (FAYYAD ET al. 1996).

“A Mineração de Dados é o processo de descoberta automática de informações úteis em grandes depósitos de dados” (KUMAR, STEINBACK e TAN, 2006).

Segundo Fayyad et al (1996) a Mineração de dados possui tarefas como a predição e a descrição. A predição utiliza as variáveis contidas na base para prever valores desconhecidos ou valores futuros enquanto a descrição busca padrões e tendências que descrevam os dados para interpretação.

2.3.3 Tarefas de mineração de dados

Segundo Navega (2002) Durante a Mineração de Dados são aplicados diversos algoritmos que processam os dados na base buscando informações importantes, e mesmo existindo algoritmos capazes de identificar dados novos é

necessária a participação de especialistas para conduzir o processo de extração do conhecimento da base.

As tarefas de mineração de dados utilizadas em diversos domínios são a Classificação, estimativa ou regressão, associação, segmentação ou *clustering* e sumarização.

Dentre as tarefas de mineração de dados é utilizada neste trabalho a indução de pesos por relevância dos atributos, além do algoritmo de classificação C4.5 para a geração da árvore de decisão. A definição de pesos por relevância é importante para apoiar a tomada de decisão gerencial, uma vez que no domínio empregado neste trabalho pode indicar divergências de critérios e prioridades. Já a árvore de decisão pode ser utilizada para representar o conhecimento subliminar dos dados e servir de instrumento para facilitar a compreensão do estado atual do domínio. Mais detalhes sobre a utilização da mineração de dados como apoio a gestão educacional são apresentados no Capítulo 4.

2.3.4 Abordagem geral para resolução de um problema de classificação.

De acordo com Kumar, Steinbach e Tan (2006).

As técnicas de classificação são utilizadas para construção de modelos de classificação a partir de um conjunto de treinamento no qual as instâncias tenham definido o rótulo das classes. Destacam-se as seguintes técnicas:

- Classificadores de árvores de decisão;
- Classificadores baseados em regras;
- Redes Neurais;
- Máquinas de vetor suporte;
- Classificadores Bayes simples.

Cada técnica utilizada requer um algoritmo de aprendizagem que identifique um modelo mais adequado ao relacionamento entre as variáveis (atributos) da base e o Rótulo das classes do conjunto de dados de entrada (treinamento). O principal objetivo do algoritmo de aprendizagem é construir modelos que realizem a predição dos rótulos das classes de registros ainda não conhecidos.

Ainda de acordo com os mesmos autores, a avaliação do desempenho de um modelo de classificação pode ser visualizada através de uma tabela chamada Matriz

de Confusão a qual contém o total de previsões corretas e erradas feitas pelo modelo.

A tabela 2 mostra um exemplo matriz de confusão para um problema de classificação binária.

Tabela 2 - Matriz de confusão para um problema de duas classes.

		Classe prevista	
		Classe = 1	Classe = 0
Classe Real	Classe =1	f_{11}	f_{10}
	Classe = 0	f_{01}	f_{00}

Fonte: Kumar, Steinbach e Tan 2006.

De acordo com os dados apresentados na Matriz de Confusão da tabela 2, o número de previsões corretas executadas pelo modelo é dado é pela soma ($f_{11} + f_{00}$) e de previsões incorretas por ($f_{10} + f_{01}$). O Total de previsões é calculado pela soma ($f_{11} + f_{10} + f_{01} + f_{00}$). A precisão do modelo é definida pelas relações:

$$\text{PRECISÃO} = \frac{f_{11} + f_{00}}{f_{11} + f_{10} + f_{01} + f_{00}} \quad \text{Equação 2}$$

De modo análogo a taxa de erro é dada pela relação:

$$\text{TAXA DE ERRO} = \frac{f_{10} + f_{01}}{f_{11} + f_{10} + f_{01} + f_{00}} \quad \text{Equação 3}$$

Pelo exposto, os algoritmos utilizados para a classificação procuram alcançar a maior precisão, ou seja, a menor taxa de erro.

2.3.5 A ferramenta Rapid Miner

O software Rapid Miner versão 5.2 é um ambiente de trabalho com algoritmos de aprendizagem de máquina. Essa ferramenta fornece recursos de importação de algoritmos de classificação de dados, algoritmos de agrupamento entre outros de grande utilidade no processo de descoberta de conhecimento. Dentre os vários

operadores desta ferramenta são utilizados neste trabalho: Decision Tree (árvore de decisão), apply model (aplicação de modelo), performance (classificação), weight by relief (peso por relevância). Maiores detalhes sobre esses operadores consultar apêndice **B** desta pesquisa.

2.3.6 Trabalhos relacionados

O artigo apresentado por Quintão Lima et al (2006) analisa os impactos causados pelo sistema de informação desenvolvido para o setor de autoavaliação (CPA) da Faculdade Metodista Granbery. Os dados obtidos através dos instrumentos de pesquisa são armazenados no sistema permitindo a geração de relatórios e possibilitando uma melhor análise da realidade da instituição, além de apresentar informações não triviais através da Mineração de Dados. Esse trabalho busca auxiliar a CPA da IES a otimizar o processo de autoavaliação e não visa à comparação entre as avaliações dos participantes do processo avaliativo.

A pesquisa realizada por Martinhago (2005) teve como objetivo principal delinear o perfil do candidato ao processo de seleção para o ensino superior da Universidade federal do Paraná, campus Curitiba, e através das informações, fornecer subsídios para as instituições de ensino superior na tomada de decisões. Para isto utilizou a Mineração de Dados através dos algoritmos de classificação J48, J48 e J48.PART implementado na ferramenta WEKA (*Waikato Environment for Knowledge Analysis*). A base de dados foi obtida através do questionário sócio educacional preenchido pelos candidatos no momento da inscrição. Esta pesquisa apesar de fornecer uma ferramenta de apoio à gestão das IES, não lida com a percepção da comunidade acadêmica em relação aos cursos superiores da instituição pesquisada.

A pesquisa de Sanches (2007) teve como objetivo estudar a importância de articulação entre o processo de avaliação institucional e os projetos pedagógicos de cursos superiores visando aprimorá-los a médio e longo prazo. A metodologia adotada utilizou-se de pesquisa qualitativa buscando conhecer e interpretar a realidade pesquisada adotando como modo de investigação o estudo de caso. Esta pesquisa procura compreender a avaliação institucional como instrumento de

melhoria da qualidade da educação superior vinculada aos projetos pedagógicos. Portanto, não lida com dados quantitativos em busca de novos conhecimentos.

A pesquisa realizada por Bertelli (2004) procurou mostrar como os resultados obtidos na avaliação institucional e avaliação externa foram imprescindíveis para orientar a IES na tomada de decisão visando à melhora da qualidade institucional. Os dados da avaliação interna foram obtidos através de questionários eletrônicos respondidos pela comunidade acadêmica e não teve nenhum tratamento através da mineração de dados. Apenas houve comparações entre médias da avaliação docente e discente no período de 2001 a 2003.

A pesquisa de Santos (2010) teve como objetivo fazer uma reflexão sobre a questão da avaliação institucional no ensino superior, tomando como base a teoria da psicanálise com foco na resistência da comunidade acadêmica em participar do processo avaliativo. Este trabalho não lida com análise quantitativa de dados.

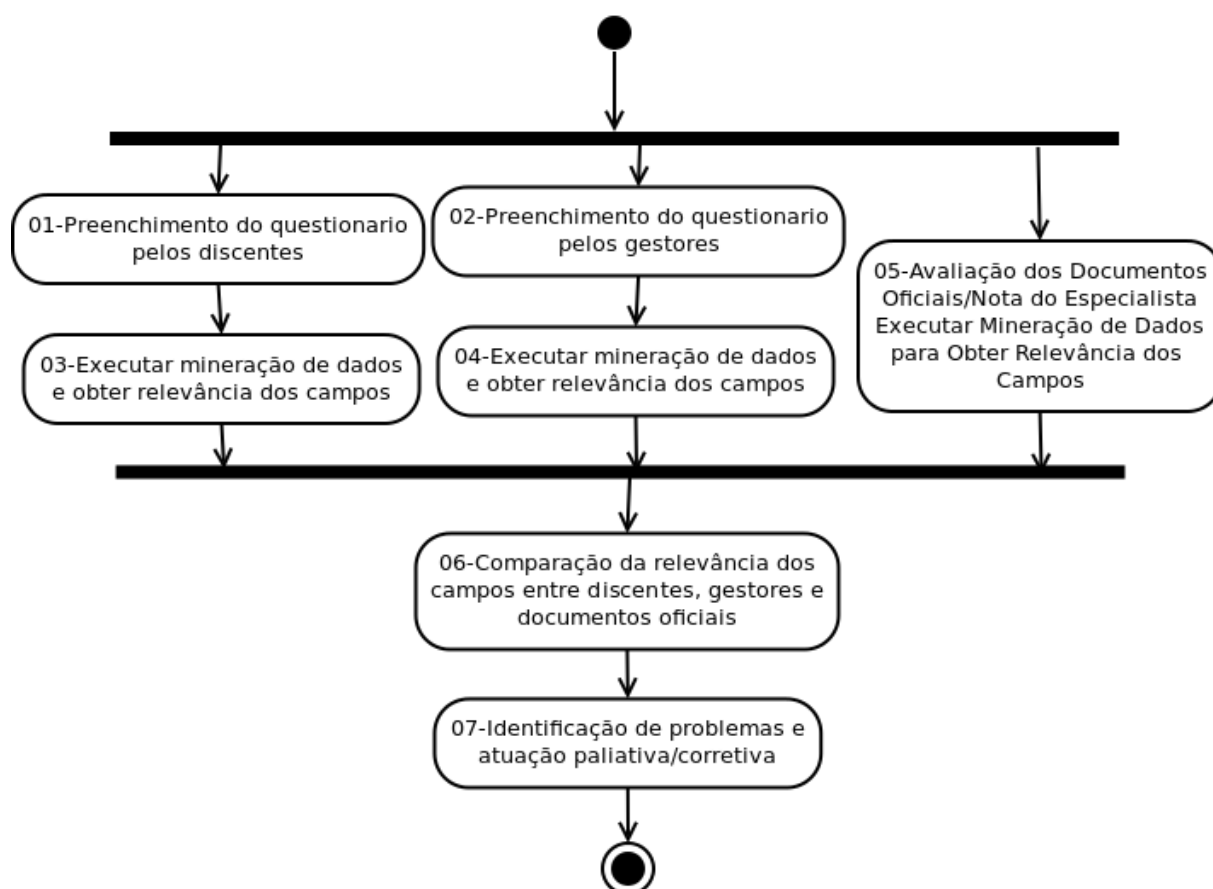
Face aos trabalhos pesquisados, a presente pesquisa busca contribuir no cenário da educação superior brasileira como uma ferramenta de apoio a gestão educacional baseada na descoberta de conhecimentos “escondidos” nas bases dos dados coletados no processo de avaliação interna conduzidos pela CPA e a comparação destes com as avaliações externas das IES, evidenciando as discrepâncias destas avaliações, e apontando caminhos para que os gestores possam analisá-los e implementar políticas para corrigir as fragilidades detectadas no processo avaliativo a curto, médio e longo prazo.

3 METODOLOGIA PROPOSTA PARA APOIO A DECISÃO DAS IES

3.1 Processos para Apoio a Decisão Utilizando Mineração de Dados

A Figura 6 apresenta uma visão geral da metodologia proposta para apoio à decisão da IES utilizando a notação de diagrama de atividades UML.

Figura 6 - Processo para apoio a decisão utilizando mineração de dados



Fonte: Autor, 2012.

ETAPA 01: Preenchimento do questionário pelos discentes

Nesta etapa serão elaborados os instrumentos⁸ de avaliação (questionários), realizada a sensibilização dos discentes quanto à importância da participação dos mesmos no processo de autoavaliação⁹, contribuindo assim para obtenção dos

⁸ Mais detalhes sobre os procedimentos quanto à elaboração dos instrumentos da autoavaliação leia Secção 2.2.5.

⁹ Mais detalhes sobre o que é autoavaliação leia a Secção 2.1.5.

dados para diagnóstico dos cursos da IES. Após a conclusão desta ação ocorrerá avaliação interna conduzida pela CPA visando coletar dados que reflitam a realidade do IFS. A autoavaliação será conduzida conforme prevê o SINAES¹⁰.

A coleta dos dados realizar-se-á utilizando uma amostra aleatória simples¹¹ cujo cálculo para determinar o tamanho da referida amostra será feito de acordo com os critérios estabelecidos na seção 2.2.6 desta pesquisa.

A população desta pesquisa é formada por todos os discentes devidamente matriculados nos cursos superiores do IFS. Também será realizada uma pesquisa com os gestores dos cursos (coordenadores, pedagogos e pró-reitores ligados aos cursos superiores).

ETAPA 02: Preenchimento do questionário pelos gestores

Esta etapa ocorrerá em paralelo com a primeira e quinta etapa e será elaborado na mesma o instrumento de avaliação dos gestores, bem como, a sensibilização dos mesmos, visando a obter dados que reflitam a realidade da IES, e dessa forma, contribuir para o diagnóstico da mesma. Durante esta ação também serão coletados os dados da avaliação interna dos gestores conduzida pela CPA de acordo com o SINAES. A coleta dos dados e a elaboração dos instrumentos de avaliação serão executadas de acordo com a explanação da etapa 01.

ETAPA 03: Executar mineração de dados e obter relevância dos campos da avaliação dos discentes

Nesta etapa realizar-se-á o processo de descoberta de conhecimento (KDD¹²), no qual será realizada a mineração dos dados obtidos através da avaliação dos discentes, quanto aos cursos superiores da IES na etapa 01. Após a mineração dos dados serão obtidos uma tabela com os pesos (relevância¹³) variando de zero (0) a um (1), uma matriz de confusão¹⁴ e uma árvore de decisão com as regras relativas aos dados das questões da avaliação interna. A mineração será realizada através do software Rapid Miner¹⁵ versão 5.2.

¹⁰ Mais detalhes sobre o que é SINAES leia a seção 2.1

¹¹ Mais detalhes sobre amostra aleatória simples leia seção 2.2.4

¹² Mais detalhes sobre o KDD leia seção 2.3

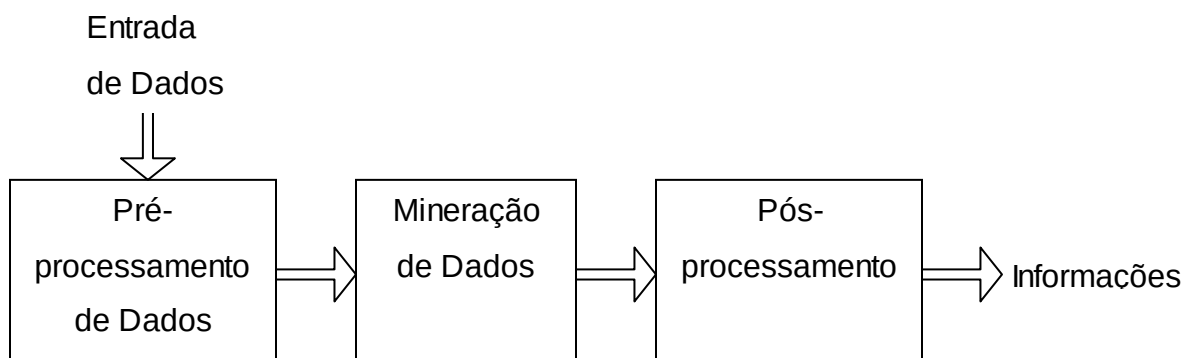
¹³ Mais detalhes sobre pesos por relevância leia Apêndice B.

¹⁴ Mais detalhes sobre matriz de confusão leia Apêndice B.

¹⁵ Mais detalhes sobre a ferramenta Rapid Miner leia seção 2.3.5

A figura 7 mostra as etapas que serão realizadas desde o pré-processamento até o pós-processamento dos dados obtidos através da pesquisa realizada com os gestores e discentes do IFS.

Figura 7 - Etapas do processo de KDD



Fonte: Adaptado de KUMAR, STEINBACK e TAN, 2009.

A seleção dos dados será feita através da base da Avaliação Institucional do Instituto Federal de Sergipe do ano de 2012. Os dados foram coletados de uma base de dados Postgresql e analisados a partir do Rapid Miner. Devido à grande quantidade de dados que são coletados no momento da avaliação institucional, o presente trabalho tem como objetivo avaliar os dados relativos aos cursos superiores do IFS.

Assim, a base de dados escolhida para participação neste processo se refere à avaliação dos discentes e gestores quanto aos cursos superiores, cujos instrumentos (questionários) de avaliação possuem respostas com notas variando de 1 a 5, considerando:

- 1 - Desconheço;
- 2 - Ruim;
- 3 - Regular;
- 4 - Bom;
- 5 - Muito Bom.

De posse da escolha do domínio a ser abordado (Avaliação Institucional, mais precisamente avaliação dos discentes e gestores quanto aos cursos superiores no Instituto Federal de Sergipe) é obtida a base de dados. Será executada a limpeza dos dados, para a definição de campos (atributos) relevantes ao processo de

Descoberta de Conhecimento e conversão do rótulo dos mesmos, da modalidade de questões numeradas para atributos nominais, ficando assim mais significativos.

Também foi criado um atributo (Rótulo da classe) chamado Avaliação Curso para atuar como classe na base de conhecimento. Esse atributo classificará o Curso em: Ruim, Regular, Bom, Muito Bom. O nome dos cursos que participarão da avaliação será retirado da base de dados, pois os mesmos não têm impactos significativos na avaliação dos cursos, ou seja, a percepção dos alunos é convergente independente do curso.

A base resultante desse processo contém 321 registros rotulados de acordo com os seguintes atributos:

- Estrutura Curricular;
- Carga Horária das Disciplinas;
- Carga Horária Total;
- Laboratórios Informática;
- Número de Docentes;
- Qualificação Docente;
- Promoção de Extensão;
- Envolvimento Prof/Pesquisa;
- Envolvimento Alunos/Pesquisa;
- Articulação de Conteúdos;
- Laboratórios Específicos;
- Aulas Práticas de campo;
- Curso e Mercado de Trabalho;
- Contribuição Disciplinas para Formação;
- Avaliação Curso (Classe).

ETAPA 04: Executar mineração de dados e obter relevância dos campos da avaliação dos gestores

Nesta etapa será realizada a mineração dos dados obtidos através da avaliação dos gestores quanto aos cursos da IES na etapa 02. Após a mineração dos dados serão obtidos uma tabela com os pesos (relevância) variando de zero (0) a um (1), uma matriz de confusão e uma árvore de decisão com as regras relativas aos atributos das questões da avaliação interna conforme descrito na etapa 03. Também será necessário normalizar os dados relativos aos pesos obtidos para cada

variável (questões do instrumento de avaliação dos discentes e gestores) após a mineração visando obter uma melhor interpretação dos mesmos para realizar uma análise comparativa destes dados com os relativos aos discentes e especialistas.

ETAPA 05: Avaliação dos documentos oficiais da IES (Especialista).

Consiste na elaboração de diretrizes de referência baseadas nos documentos oficiais da IES, nas normas estabelecidas pelos SINAES e nos resultados das avaliações realizadas por comissões externas do MEC quanto aos cursos do IFS. Estas diretrizes são notas atribuídas às questões do instrumento de avaliação utilizado para os gestores e discentes. Estas notas são atribuídas de acordo com as avaliações externas (dos cursos superiores) já realizadas no IFS numa escala de 1 a 5. Estes dados serão minerados realizando-se o mesmo processo descrito nas etapas 03 e 04 visando uma comparação entre os resultados obtidos na avaliação Gestores, Discentes e Especialistas.

ETAPA 06: Comparação da relevância (pesos) dos campos (questões) entre discentes, gestores e documentos oficiais (Especialista).

Nesta atividade será realizada uma comparação dos campos obtidos na Mineração dos Dados dos discentes e gestores obtidos nas etapas 03 e 04 com os dados obtidos na etapa 05 do Especialista, visando a uma comparação das tabelas com pesos por relevância das avaliações já mencionadas. Para tornar claro esta comparação serão construídos gráficos no sistema cartesiano ortogonal Atributo x Pesos (Discentes, Gestores e Especialista), Atributo x Nota (Especialista) x Peso (Especialista) para uma análise do comportamento dos resultados que serão obtidos no processo avaliativo.

ETAPA 07: Identificação de problemas e ações paliativas corretivas

Nesta atividade será realizada a identificação dos pontos fracos detectados na avaliação interna e as respectivas medidas corretivas/paliativas para minimizar as fragilidades detectadas no processo de autoavaliação. Dessa forma, espera-se que as ações que serão propostas possam contribuir no processo de Apoio a tomada de decisão das IES.

4 RESULTADOS

Tendo em vista os objetivos traçados para realização desta pesquisa e a metodologia adotada, neste capítulo são apresentados os resultados das seguintes atividades:

- Diretrizes para elaboração e preenchimento dos questionários aplicados aos discentes, gestores, especialista e o cálculo do tamanho da amostra utilizada na pesquisa realizada com os discentes;
- Discussão dos experimentos realizados através do processo de descoberta de conhecimento na base de dados estudada utilizando o software RAPIDMINER;
- Dificuldades encontradas durante as experiências na mineração dos dados e análise dos resultados encontrados, bem como, os conhecimentos não triviais descobertos através da Mineração de Dados;
- Comparação da relevância (pesos) dos campos (questões) entre discentes, gestores e especialista (documentos oficiais);
- Identificação de indícios de problemas na avaliação dos discentes e gestores e sugestão de ações para corrigir pontos frágeis evidenciados na avaliação interna;
- Resultado da pesquisa realizada durante seminário com o colégio de dirigentes da IES a cerca do impacto na gestão quanto ao conhecimento gerado através da Mineração de Dados.

4.1 Diretrizes para Elaboração e Preenchimento dos Questionários Aplicados aos Discentes e Gestores do IFS.

As diretrizes utilizadas na elaboração dos instrumentos de avaliação (questionários) e a sensibilização dos discentes e gestores que participaram da autoavaliação foram realizadas de acordo com a secção 2.2.5 desta pesquisa.

4.1.1 Cálculo do tamanho mínimo da amostra obtida com os discentes dos cursos superiores do IFS

O procedimento de amostragem adotado foi o de amostragem aleatória simples, considerando a população como o conjunto formado por todos os alunos devidamente matriculados nos cursos superiores do IFS (sede) no ano de 2012.

Para o cálculo do tamanho mínimo da amostra foi utilizada as relações discutidas na secção 2.2.6.

Para utilizar as fórmulas descritas na secção acima, consideramos:

- $N = 898$ (nº de elementos da população dos discentes)
- $E_0 = 0,05$ (erro amostral)
- $Z = 1,96$ (para um nível de confiança de 95% conforme secção 2.2.6)

O desvio padrão utilizado foi $\sigma = 0,56$ obtido em uma amostra piloto¹⁶ e n_0 é o menor natural que satisfaz a inequação abaixo conforme a secção 2.2.6.

$$n \geq \frac{z^2 \cdot \sigma^2}{E_0^2} = \frac{1,96^2 \cdot 0,56^2}{0,05^2} = 481,89$$

Assim, temos $n_0 = 481$,

Fazendo a correção devido à população ser pequena, vem:

$$n = \frac{N \cdot N_0}{N + N_0} = \frac{898 \times 481}{898 + 481} = 313,68$$

$\therefore n = 313$ (representa o número mínimo de participantes da amostra)

A coleta dos dados da autoavaliação contou com a participação real de 321 discentes, portanto, o tamanho mínimo da amostra foi alcançado no processo avaliativo, e dessa forma tem-se uma amostra estatisticamente significativa da população em estudo.

¹⁶ Para entender amostra piloto leia secção 2.2.6.

Para realizar a coleta dos dados relativos aos gestores, a população considerada foi formada pelos coordenadores dos cursos superiores, pedagogos e pró-reitores ligados ao ensino superior. A amostra teve a participação de 13 dos 14 previstos nesta população, ou seja, 92,85%. Quanto aos dados relativos ao Especialista foram obtidos de acordo com as avaliações externas da IES contida em seus documentos oficiais.

Após a consolidação dos dados da avaliação dos discentes foi gerada as Tabelas 3 e 4 contendo uma distribuição de frequências, as medidas de tendência central e o desvio padrão, respectivamente. Também foi construído um histograma com o objetivo de dar uma visão dos dados consolidados e apresentar alguns recursos estatísticos disponíveis no protótipo do sistema desenvolvido para coletar os dados da autoavaliação relativos ao presente trabalho.

Tabela 3 - Avaliação geral de todos os cursos pelos discentes.

Conceitos	Frequência	%	% Acumulada
Desconheço	4	1,2	1,2
Fraco	20	6,2	7,5
Regular	65	20,2	27,7
Bom	176	54,8	82,6
Muito Bom	56	17,4	100,0
Total	321	100,0	

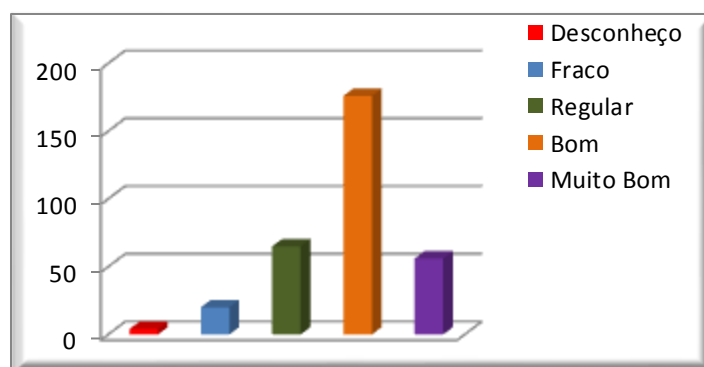
Fonte: Autor, 2012.

Tabela 4 - Medidas de tendência central e dispersão

Média	3,81
Mediana	4,00
Moda	4,00
Desvio Padrão	0,84

Fonte: Autor, 2012.

Gráfico 1 - Histograma dos dados consolidados da pesquisa



Fonte: Autor, 2012.

4.2 O Processo de Descoberta de Conhecimento (KDD) – Execução da Mineração de Dados

Durante o processo de descoberta de conhecimento seguiu-se as etapas descritas secção 2.3.1 desta pesquisa.

Etapa 01: Pré-processamento

Nesta etapa foi necessária a retirada das questões (4, 5, 6, 16 e 17) relativas à avaliação do curso conforme questionário dos discentes em anexo uma vez que não influenciavam na avaliação do curso (ou seja, seleção de dados), bem como, das tuplas incompletas (Questões com respostas em branco). Optou-se pelo descarte ao invés do ajuste devido à amostra ser reduzida em face de a população ser pequena, e dessa forma, eventuais ruídos (atributos vazios) teriam impactos significativos nos resultados. Este fato ficou claro quando a mineração dos dados foi realizada sem os ajustes descritos acima e a acurácia do classificador foi de apenas 78,20% Conforme descrito na análise dos dados da Figura 9 apresentada no processamento.

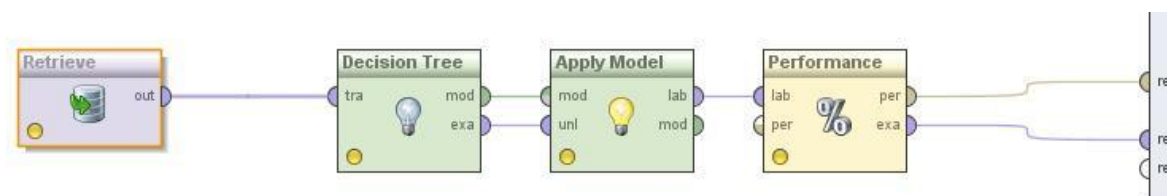
Etapa 02: Processamento

Através do software Rapid Miner 5.2 foi realizada a Mineração de Dados, etapa principal do processo de Descoberta de Conhecimento. No presente trabalho foram utilizados os operadores: Decision Tree, Apply Model, Performance, Weight by Relief.

Nesta fase foram derivadas as regras de classificação e a obtenção dos pesos por relevância dos atributos (questões da avaliação) da base, cujas atividades podem ser visualizadas nas Figuras 10,11e 12 apresentadas no pós-processamento.

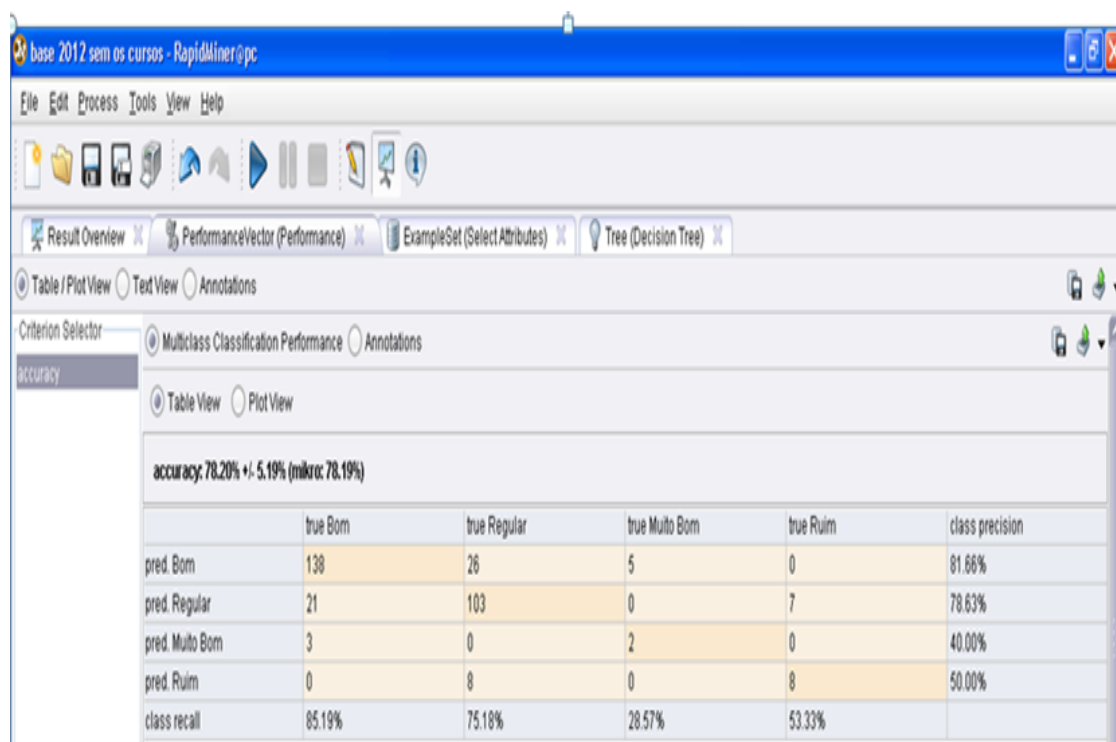
A Figura 9 mostra a acurácia de 78,20% do classificador quando a mineração foi realizada sem os ajustes descritos na Etapa 01(pré-processamento)

Figura 8 - Processo no Rapid Miner para construção da Árvore de Decisão



Fonte: Autor, 2012.

Figura 9 - Acurácia do classificador sem os ajustes



Fonte: Autor, 2012.

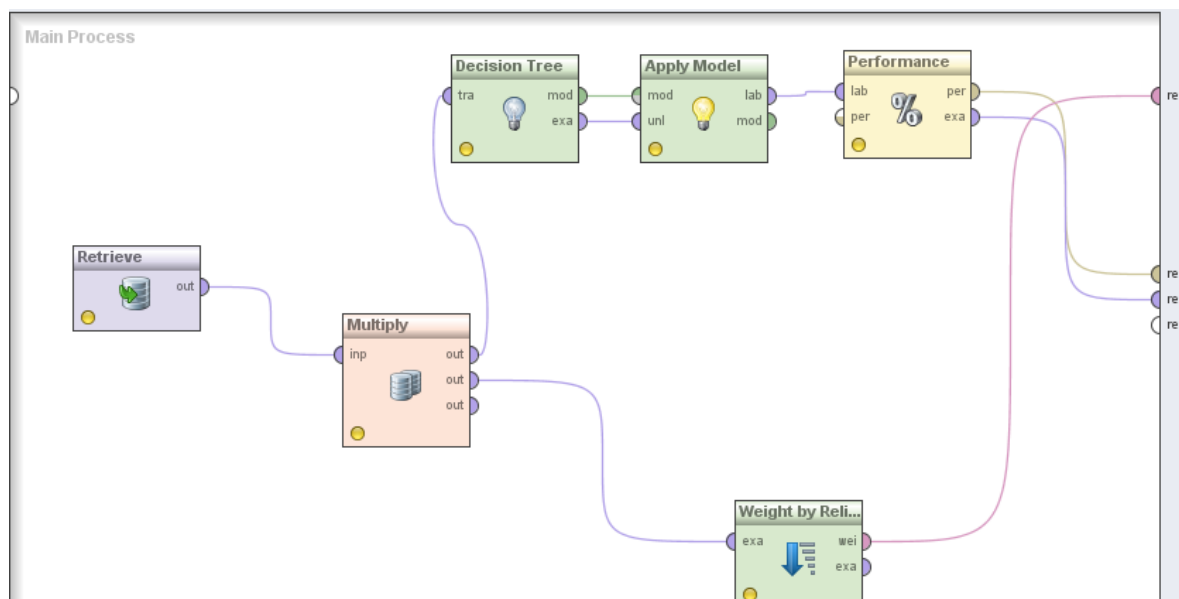
Vários testes foram realizados apresentando resultados similares ao descrito na figura 9. Após estas experiências foi necessário realizar os ajustes mencionados na Etapa 01(pré-processamento) obtendo-se uma acurácia mais significativa conforme descrição na etapa 03 (pós-processamento) a seguir.

Etapa 03: Pós-processamento

Nesta etapa obteve-se a interpretação dos dados da Mineração de Dados, sendo necessário dar semântica aos resultados obtidos para cada atributo, bem como, a identificação dos problemas de relevância (peso) entre as avaliações dos discentes, gestores e a matriz obtida pelo especialista, com base nos resultados de avaliações anteriores realizadas por comissões designadas pelo INEP e através dos documentos oficiais da IES.

Assim, construiu-se um modelo de árvore de decisão aplicando o algoritmo C4.5 ,conforme Figura 10 .Posteriormente verificou-se a sua exatidão quanto à predição da classe (Avaliação Curso) da base de conhecimento, através de um recurso chamado matriz de confusão, visualizado na Figura 11. Também foi calculado o peso relativo aos atributos (questões), e gerada a Figura 12 com pesos por relevância através do operador weight by relief. Este procedimento foi adotado também para os Gestores e Especialista.

Figura 10 - Processo no Rapid Miner para construção da Árvore de Decisão, obtenção dos pesos por relevância e a matriz de confusão



Fonte: Autor, 2012.

Figura 11 - Matriz de confusão para discentes

resultado-08-09 - RapidMiner.ppc

File Edit Process Tools View Help

Result Overview ExampleSet (Retrieve) PerformanceVector (Performance)

Table / Plot View Text View Annotations

Criterion Selector

accuracy

Multiclass Classification Performance Annotations

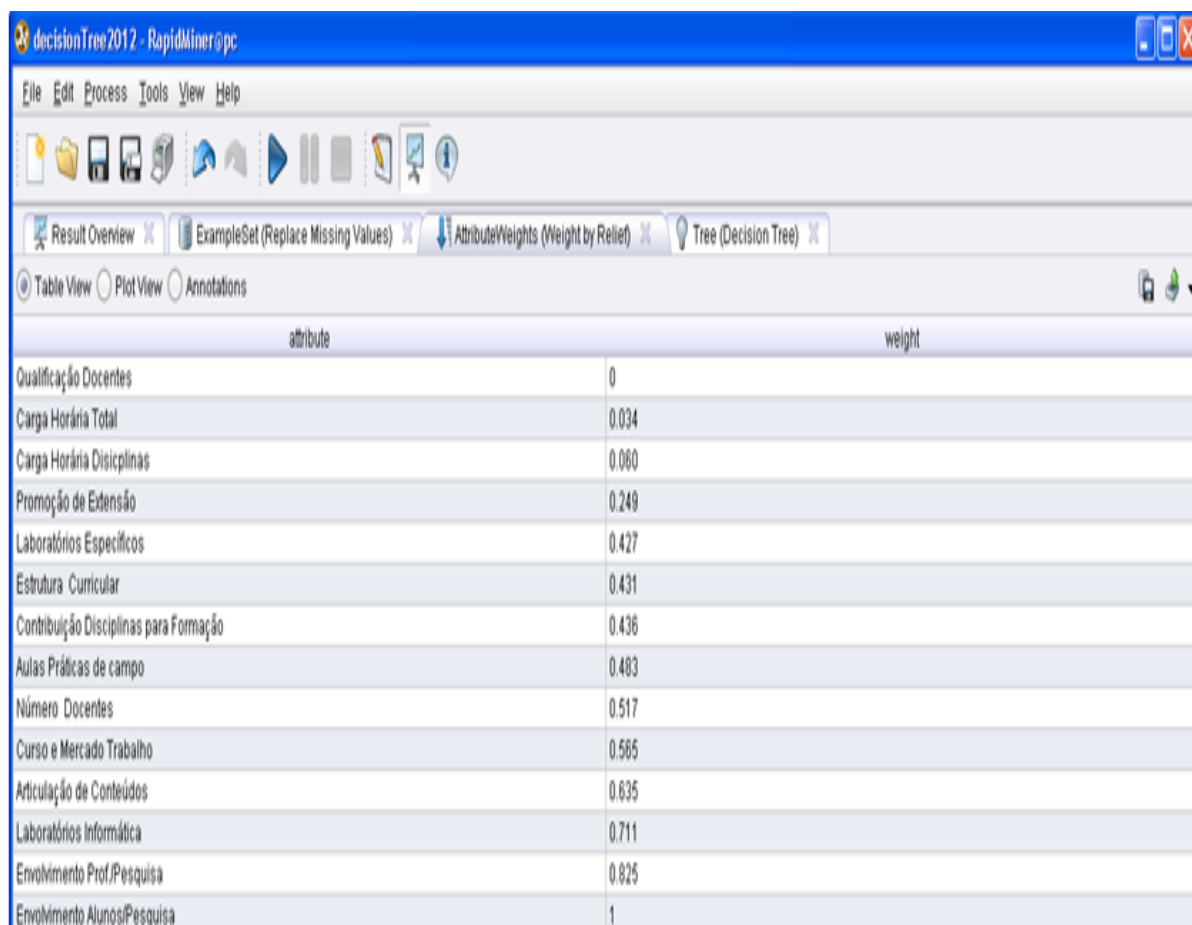
Table View Plot View

accuracy: 94.39%

	true Bom	true Regular	true Muito Bom	true Ruim	class precision
pred. Bom	153	3	2	0	96.84%
pred. Regular	6	131	0	1	94.93%
pred. Muito Bom	1	0	5	0	83.33%
pred. Ruim	2	3	0	14	73.68%
class recall	94.44%	95.62%	71.43%	93.33%	

Fonte: Autor, 2012.

Figura 12 - Peso por relevância dos atributos da base dos discentes



attribute	weight
Qualificação Docentes	0
Carga Horária Total	0.034
Carga Horária Disciplinas	0.060
Promoção de Extensão	0.249
Laboratórios Específicos	0.427
Estrutura Curricular	0.431
Contribuição Disciplinas para Formação	0.436
Aulas Práticas de campo	0.483
Número Docentes	0.517
Curso e Mercado Trabalho	0.565
Articulação de Conteúdos	0.635
Laboratórios Informática	0.711
Envolvimento Prof./Pesquisa	0.825
Envolvimento Alunos/Pesquisa	1

Fonte: Autor, 2012.

De posse dos dados obtidos na Figura 12 foi realizada uma normalização obtendo a Tabela 5 com o objetivo de facilitar a análise destes dados.

A matriz de confusão da Figura 11 revela que a árvore de decisão representa adequadamente os dados obtidos na avaliação institucional, tendo em vista a acurácia de 94,39% do classificador. Ressalta-se que o conceito muito bom teve uma previsão de 71,43% devido à baixa frequência deste conceito na base em estudo. A Tabela 5 mostra os pesos relativos aos atributos da base revelando a importância dos mesmos numa escala de 0 a 1 (Normalização realizada através do operador weight by relief no processo de autoavaliação). Para analisar os dados desta tabela foi necessário normalizar estes conforme descrito na seção 3 (etapa 04) da metodologia desta pesquisa. A 1ª coluna contém os dados obtidos na Mineração e a 2ª coluna através do processo de normalização (Frequência Relativa).

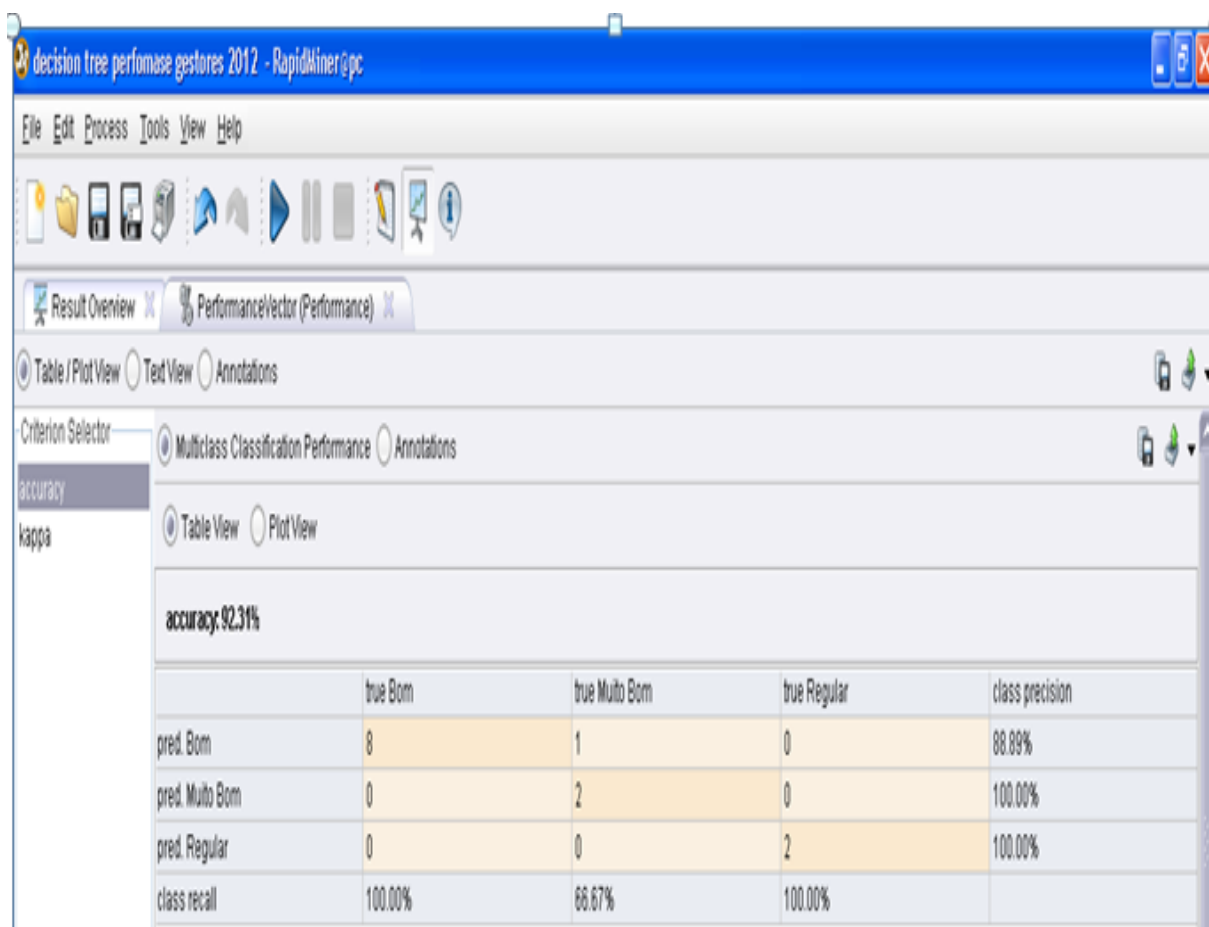
Tabela 5 - Tabela do Rapid Miner. Matriz com pesos e normalização

Atributo	Peso	Normalização%
Qualificação Docente	0	0,00%
Carga Horária Total	0,034	0,534%
Carga Horária Disciplinas	0,06	0,941%
Promoção de Extensão	0,249	3,907%
Laboratórios Específicos	0,427	6,700%
Estrutura Curricular	0,431	6,763%
Contribuição Disciplinas para Formação	0,436	6,841%
Aulas Práticas de Campo	0,483	7,579%
Número Docentes	0,517	8,112%
Curso e Mercado de Trabalho	0,565	8,866%
Articulação de Conteúdos	0,635	9,964%
Laboratórios de Informática	0,711	11,156%
Envolvimento Prof/Pesquisa	0,825	12,945%
Envolvimento Alunos/Pesquisa	1	15,691%

Fonte: Autor, 2012.

O mesmo processo foi realizado com a base de dados obtida pela avaliação dos gestores e especialista frente aos cursos superiores do IFS, obtendo assim as Tabelas 6 e 7 com os pesos relativos aos atributos da base. Para estes dados também foi realizada a normalização com o objetivo de dar um melhor significado a interpretação dos mesmos, e desta forma, facilitar a comparação destes resultados com os relativos aos discentes.

Figura 13 - Matriz de confusão dos gestores



Fonte: Autor, 2012.

A matriz de confusão da Figura 13 revela que a árvore de decisão obtida na avaliação dos gestores também representa adequadamente os dados obtidos na avaliação institucional, pois a mesma apresenta uma acurácia de 92,31% do classificador. As Tabelas 6 e 7 mostram os pesos por relevância dos atributos da base (Gestores e Especialista).

Tabela 6 - Dados dos gestores

Atributo	Peso	Normalização
Qualificação Docente	0	0%
Carga Horária Total	0,02	0.407%
Carga Horária Disciplinas	0,03	0.611%
Promoção de Extensão	0,112	2.280%
Laboratórios Específicos	0,140	2.850%
Estrutura Curricular	0,171	3.481%
Contribuição Disciplinas para Formação	0,234	4.764%
Aulas Práticas de Campo	0,443	9.019%
Número de Docentes	0,458	9.324%
Curso e Mercado de Trabalho	0,513	10.444%
Articulação de Conteúdos	0,560	11.401%
Laboratórios de Informática	0,614	12.500%
Envolvimento Prof/Pesquisa	0,617	12.561%
Envolvimento Alunos/Pesquisa	1	20.358%

Fonte: Autor, 2012.

Tabela 7 - Dados do especialista

Atributo	Peso	Normalização
Qualificação Docente	0.667	9.452%
Carga Horária Total	0.667	9.452%
Carga Horária Disciplinas	0.667	9.452%
Promoção de Extensão	0.111	1.573%
Laboratórios Específicos	0.00	0.00%
Estrutura Curricular	0.333	4.719%
Contribuição Disciplinas para Formação	1	14.170%
Aulas Práticas de Campo	0.167	2.366%
Número de Docentes	0.667	9.452%
Curso e Mercado de Trabalho	0.667	9.452%
Articulação de Conteúdos	1	14.170%
Laboratórios de Informática	0.111	1.573%
Envolvimento Prof/Pesquisa	0.667	9.452%
Envolvimento Alunos/Pesquisa	0.333	4.719%

Fonte: Autor, 2012.

Para obter eficácia no processo de Mineração de dados é necessária a integração entre o especialista do domínio da aplicação, o analista que deve conhecer todo o processo de Mineração e extração do conhecimento e, finalmente, o usuário que utiliza o conhecimento extraído com uma ferramenta de apoio a gestão (REZENDE, 2005). No presente trabalho, os gestores e colégio de dirigentes da IES representam os usuários.

4.3 Avaliação dos Documentos Oficiais do IFS

Após análise dos documentos oficiais da IES e estudo sobre o histórico das avaliações externas realizadas por comissões designadas pelo MEC, quanto aos cursos superiores do IFS. Foi construída a **Matriz do Especialista** apresentada na Tabela 8 onde cada atributo da base recebe uma nota numa escala de 1 a 5, sendo estes dados normalizados para dar um melhor entendimento na comparação das discrepâncias entre as avaliações dos discentes, gestores e especialista, sendo: 1 - Péssimo; 2 - Ruim; 3 - Regular; 4 - Bom; 5 - Muito Bom.

Tabela 8 - Matriz obtida pelo conhecimento do especialista

Atributo	Nota	Normalização %
Qualificação Docente	3	7,317
Carga Horária Total	4	9,756
Carga Horária Disciplinas	4	9,756
Promoção de Extensão	1	2,439
Laboratórios Específicos	2	4,878
Estrutura Curricular	4	9,756
Contribuição Disciplinas para Formação	4	9,756
Aulas Práticas de Campo	1	2,439
Número Docentes	3	7,317
Curso e Mercado de Trabalho	4	9,756
Articulação de Conteúdos	3	7,317
Laboratórios de Informática	2	4,878
Envolvimento Prof/Pesquisa	3	7,317
Envolvimento Alunos/Pesquisa	3	7,317

Fonte: Autor, 2012.

4.4 Comparação da Relevância dos Pesos Obtidos na Avaliação dos Discentes, Gestores e Especialista do IFS

Esta etapa foi realizada após a normalização dos pesos por relevância obtidos para cada atributo da base de dados na visão dos discentes, gestores e Especialista (os pesos por relevância foram gerados na mineração dos dados através do software Rapid - Miner). Pelo exposto, o processo sugere que a comparação das discrepâncias combine as visões dos participantes da autoavaliação (discentes, gestores, especialista). Para realizar estas comparações os dados foram consolidados na Tabela 9 com o objetivo de tornar clara a análise dessas discrepâncias. A primeira coluna da tabela apresenta os atributos considerados no questionário da autoavaliação. Também são apresentados os dados dos pesos por relevância, obtidos via mineração de dados, para cada um dos papéis envolvidos na autoavaliação: aluno (PA), gestor (PG) e especialista (PE). A coluna Nota Esp. (NE) apresenta a nota do especialista na avaliação de cada atributo, que pode variar de 1 (mínimo) a 5 (máximo). A coluna Ganho Potencial apresenta o quanto é possível evoluir em cada atributo, isto é, a diferença entre a nota máxima (cinco) e a nota dada pelo especialista. O Índice de prioridade (IP) é um índice proposto na presente pesquisa que considera tanto o ganho potencial do atributo (GP), quanto a sua importância relativa à opinião do especialista (PE). Sendo assim, a fórmula proposta para a definição do índice de prioridades é: $IP = GP \times PE$. As demais colunas da tabela apresentam os valores normalizados dos dados já comentados.

Tabela 9 - Comparação e relevância dos pesos obtidos

Atributo	Peso por Relevância			Nota Esp, (NE)	Ganho Potencial	Índice De Prioridade	Dados Normalizados(%)					
	Aluno	Gestor	Esp.				Aluno	Gestor	Esp.	Nota (Esp.)	Ganho potencial (GP)	Índice de prioridade PE x GP
QD	0	0	0,667	3	2	1,334	0	0	9,452	7,317	6,897	11,65
CHT	0,034	0,02	0,667	4	1	0,667	0,534	0,407	9,452	9,756	3,448	5,83
CHD	0,06	0,03	0,667	4	1	0,667	0,941	0,611	9,452	9,756	3,448	5,83
PE	0,249	0,112	0,111	1	4	0,444	3,907	2,280	1,573	2,439	13,793	3,88
LE	0,427	0,14	0	2	3	0	6,700	2,850	0	4,878	10,345	0
EC	0,431	0,171	0,333	4	1	0,333	6,763	3,481	4,719	9,756	3,448	2,91
CDF	0,436	0,234	1	4	1	1	6,841	4,764	14,170	9,756	3,448	8,74
APC	0,483	0,443	0,167	1	4	0,668	7,579	9,019	2,366	2,439	13,793	5,84
ND	0,517	0,458	0,667	3	2	1,334	8,112	9,324	9,452	7,317	6,897	11,65
CMT	0,565	0,513	0,667	4	1	0,667	8,866	10,444	9,452	9,756	3,448	5,83
AC	0,635	0,56	1	3	2	2	9,964	11,401	14,170	7,317	6,897	17,47
LI	0,711	0,614	0,111	2	3	0,333	11,156	12,500	1,573	4,878	10,345	2,91
EPP	0,825	0,617	0,667	3	2	1,334	12,945	12,561	9,452	7,317	6,897	11,65
EAP	1	1	0,333	3	2	0,666	15,691	20,358	4,719	7,317	6,897	5,82

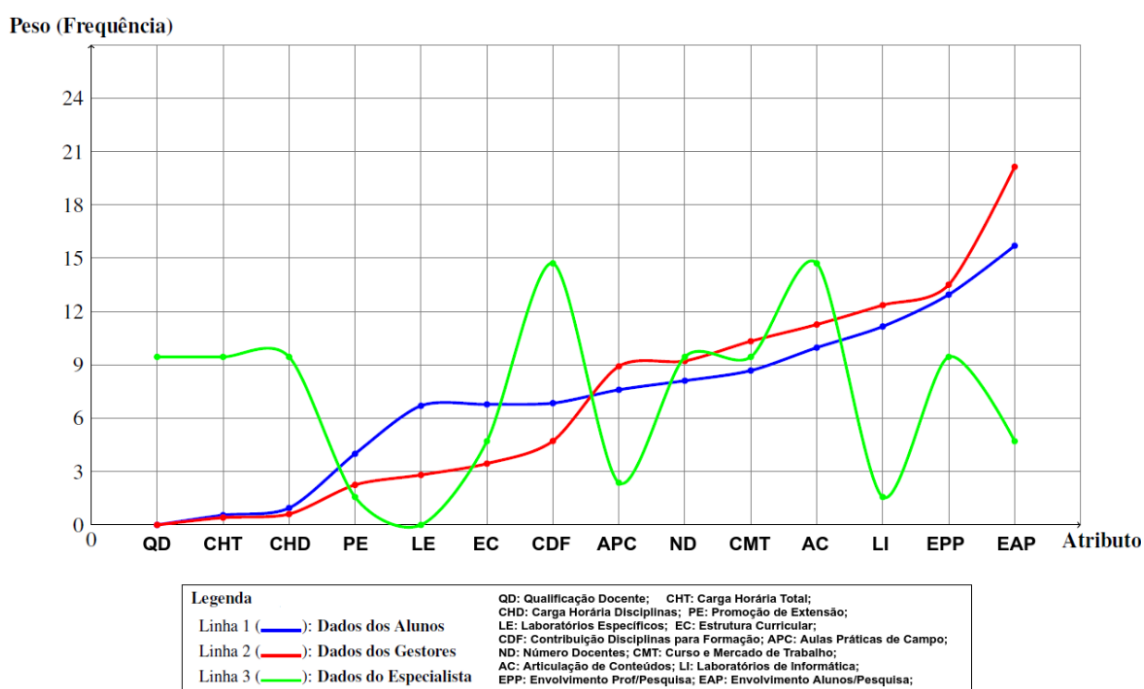
Fonte: Autor, 2012.

No tocante às comparações entre os atributos apresentados na Tabela 9, o modelo proposto apresenta três estratégias complementares envolvendo diferentes campos dos dados da autoavaliação institucional, cada uma com um propósito complementar, sendo todas elas enfatizando a opinião do especialista avaliador: (i) **diagnóstico geral**, que busca identificar inconformidades entre as prioridades buscadas pelos avaliadores, gestores e discentes; (ii) **classificação de fragilidades**, que visa ordenar as fragilidades apontadas pelo especialista de acordo com a sua importância no resultado final da avaliação (peso por relevância); e (iii) **diagnóstico detalhado**, que visa identificar possíveis causas das discrepâncias percebidas e de eventuais fragilidades identificadas no processo avaliativo. A seguir, cada uma dessas comparações serão apresentadas.

4.4.1 Diagnóstico Geral

O diagnóstico geral envolve a identificação de discrepâncias entre o peso por relevância dado por gestores e discentes, em relação ao peso por relevância dado pelo especialista avaliador. A Figura 14 apresenta o gráfico comparativo dos pesos por relevância atribuídos a cada um dos pares envolvidos: gestor, discente e especialista. Nesse gráfico, divergências percebidas em relação aos pesos dados pelo especialista em cada atributo, podem indicar necessidade de ações por parte dos gestores, visando melhores resultados em avaliações futuras. Um exemplo da utilidade dessa estratégia pode ser percebido no atributo “Qualificação Docente”, onde apesar do alto peso dado pelo especialista (9,45%), gestores e discentes consideraram tal atributo como de menor relevância no processo de avaliação do curso. Vale ressaltar que as diretrizes apontadas pelo SINAES não definem pesos entre os critérios da avaliação dos cursos. Sendo assim, para obter tal informação, a solução proposta sugere a utilização de mineração de dados, mais especificamente algoritmos de classificação, com o intuito de identificar pesos por relevância a partir dos dados avaliados pelo especialista.

Figura 14 - Gráfico comparativo Discente X Gestor X Especialista

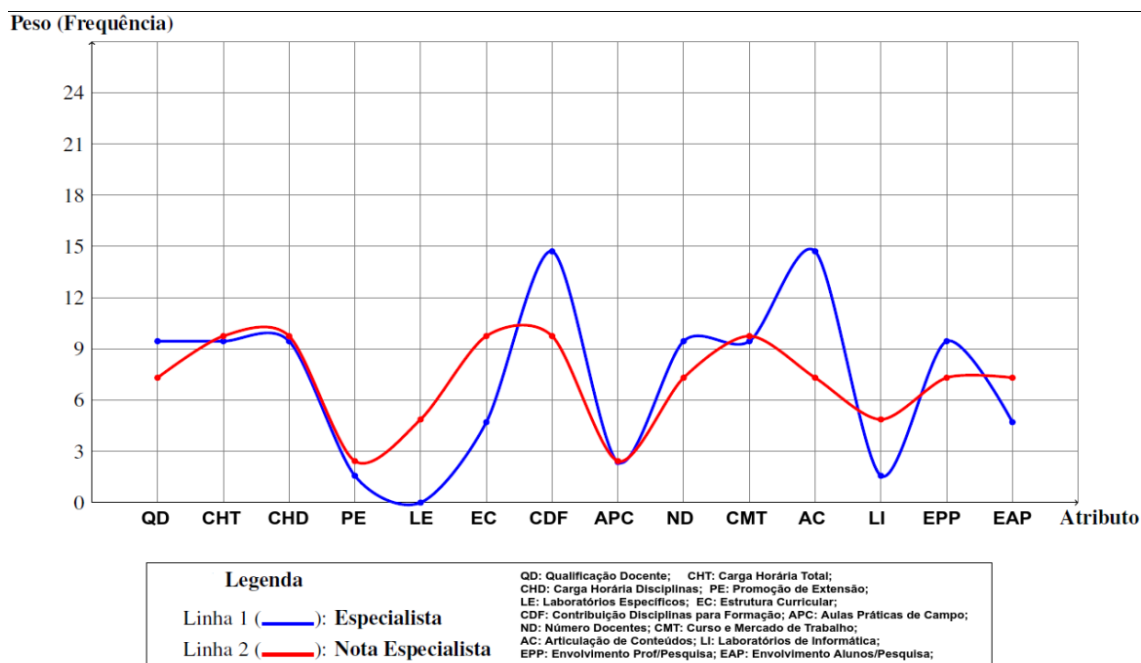


Fonte: Autor, 2012.

4.4.2 Classificação de Fragilidades

O objetivo desta etapa é guiar o gestor nas suas atividades de curto e médio prazo. Para isso, as fragilidades identificadas pelo especialista devem ser priorizadas de acordo com o respectivo impacto na avaliação final do curso. A Figura 15 apresenta um gráfico comparativo envolvendo o peso por relevância dado pelo especialista e a nota atribuída na avaliação de cada atributo pelo Especialista. Cenários que envolvam notas baixas combinadas com pesos altos devem ser priorizados pelos gestores. Por exemplo, o atributo “Articulação de Conteúdos” representa um potencial ponto de melhoria, uma vez que possui grande prioridade, mas a nota atribuída para ele foi considerada mediana. Para classificar os atributos mais prioritários de maneira mais quantitativa, a abordagem proposta se baseia em um índice, denominado índice de prioridade, que é diretamente proporcional ao quanto ainda é possível evoluir no atributo (ganho potencial) e ao peso por relevância dado ao atributo, pelo especialista: *Índice de Prioridade = ganho potencial x peso por relevância do especialista*.

Figura 15 - Gráfico comparativo Especialista X Nota do Especialista



Fonte: Autor, 2012.

Para ilustrar a ordenação dos atributos, de acordo com a sua prioridade, a Tabela 10 apresenta os valores do Índice de Prioridade de todos os atributos considerados.

Tabela 10 - Ranking de Prioridades

Atributo	Índice de Prioridade (PE x GP)
Articulação de Conteúdos – AC	17,47
Qualificação Docente- QD	11,65
Número de Docentes – ND	11,65
Envolvimento de Professor em Pesquisa- EPP	11,65
Contribuição das Disciplinas na Formação - CDF	8,74
Aulas Práticas de Campo – APC	5,84
Carga Horária Total – CHT	5,83
Carga Horária das Disciplinas – CHD	5,83
Curso e Mercado de Trabalho – CMT	5,83
Envolvimento de Aluno em Pesquisa- EAP	5,82
Promoção de Atividades de Extensão – PE	3,88
Estrutura Curricular – EC	2,91
Laboratório de Informática- LI	2,91
Laboratório Específico – LE	0,00

Fonte: Autor, 2012.

4.4.3 Diagnóstico detalhado

O diagnóstico detalhado visa a obter informações complementares que possam justificar as discrepâncias percebidas e as fragilidades identificadas pelo especialista. O detalhamento do diagnóstico deve se basear na comparação dos pesos por relevância atribuídos pelos gestores e discentes com as notas atribuídas pelo especialista durante a sua avaliação. A importância dessa comparação é facilitar a identificação de atributos críticos: (i) que apesar de serem considerados importantes pela instituição, apresentaram resultados insatisfatórios; ou (ii) que

apesar de serem pontos fortes da instituição (bons resultados), não são valorizados na mesma proporção pelos gestores e/ou discentes.

Os dados contidos na Tabela 9 revelam uma convergência em termos de ordenação da relevância dos mesmos atributos entre as avaliações dos discentes e gestores (mesma ordenação de relevância com pesos diferentes, exceção o atributo com peso zero e o atributo com peso um). No entanto, essa convergência não ocorre na avaliação do especialista que possui uma visão mais precisa da IES, pois sua avaliação é pautada em dados técnicos baseados nas avaliações realizadas por comissões externas do INEP e nos documentos oficiais do IFS. Os pesos por relevância obtidos na mineração dos dados dos discentes e gestores mostram que numa escala de zero (0) para menor relevância e um (1) para o atributo com maior relevância obteve-se, respectivamente, os atributos Qualificação Docente e Envolvimento dos Alunos em Pesquisa. Já na visão do especialista, os atributos com menor nota foram Promoção de Atividades de Extensão e Aulas Práticas de Campo, enquanto que os atributos de maior nota foram Carga Horária do Curso e Disciplinas, Estrutura Curricular, Contribuição das Disciplinas para Formação e Curso e Mercado de Trabalho. Esta análise evidencia uma discrepância na visão dos gestores e discentes com relação ao conhecimento do especialista.

A Tabela 11 mostra as medidas de tendência central Média, Mediana e as medidas de dispersão: desvio padrão e coeficiente de variação. Estes dados são utilizados como referência para análise das discrepâncias dos dados obtidos na Tabela 9 e foram calculados de acordo com as relações descritas no apêndice A desta pesquisa.

Tabela 11 - Medidas de centralidade e dispersão (aluno, gestor, nota (especialista))

	Discente	Gestor	Especialista
Média	7,14	7,14	7,13
Mediana	4,62	6,82	7,31
Desvio Padrão	4,62	6,03	2,61
Coeficiente de Variação	64%	84%	36%

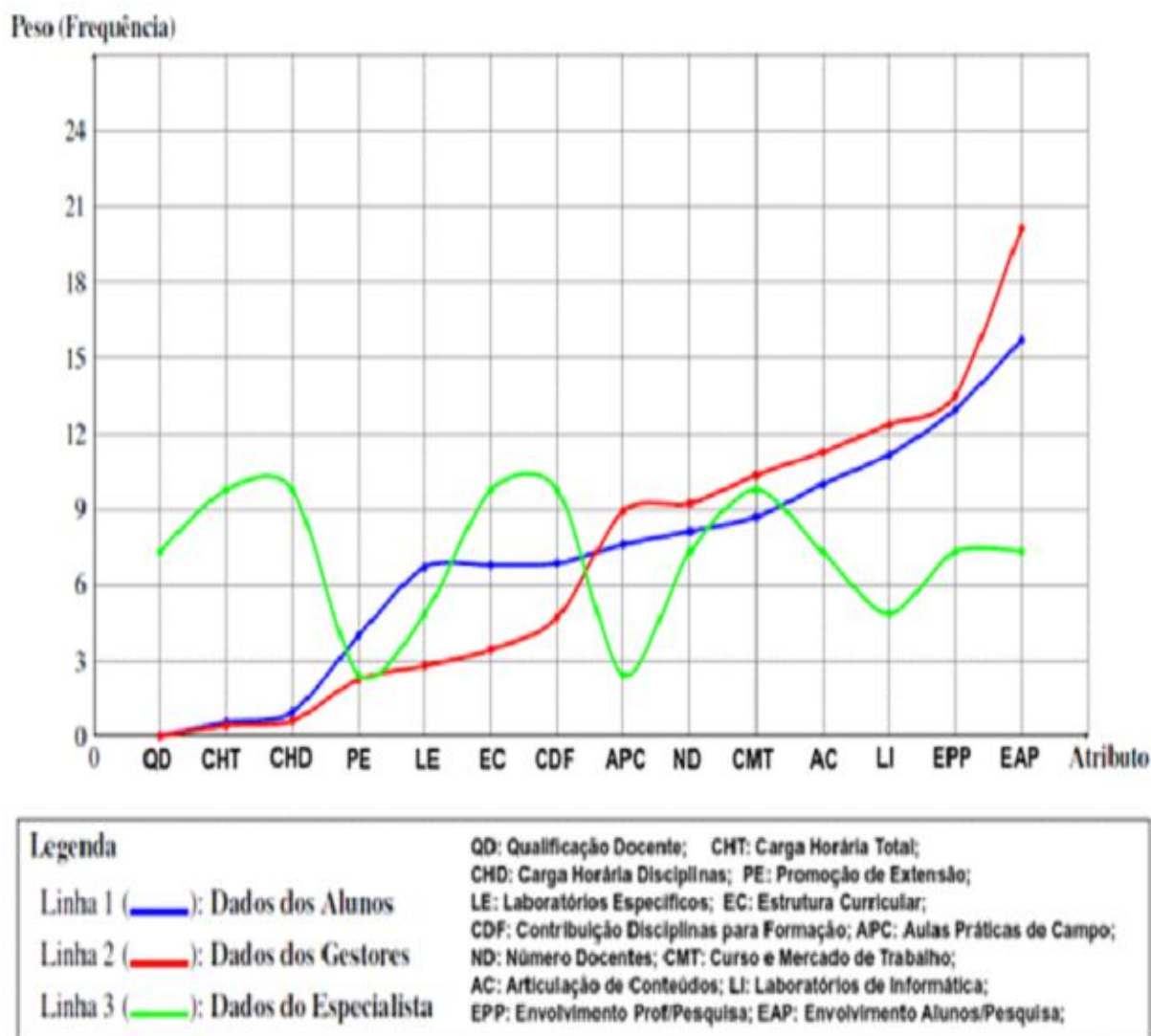
Fonte: Autor, 2012.

Os dados contidos na Tabela 11 apontam que a avaliação dos gestores possui a maior dispersão relativa (84%) seguido pelos discentes com dispersão relativa (64%). A avaliação do especialista possui a menor dispersão relativa (36%). Os resultados desses valores obtidos indicam que existe uma elevada dispersão nas avaliações dos discentes, gestores e especialista conforme definição de coeficiente de variação descrito no apêndice **A** desta pesquisa. A seguir, são apresentados mais discussões sobre os resultados obtidos na análise detalhada, fruto da comparação dos pesos por relevância atribuídos por discentes e gestores, com as notas dos especialistas.

Para análise detalhada das eventuais causas de fragilidades encontradas, optou-se pela construção de um gráfico contendo os resultados dos pesos por relevância dos atributos na concepção de Alunos, Gestores e nota do Especialista visualizados na Tabela 9, gerando assim, um gráfico Peso x Atributo conforme Figura 16.

A análise foi realizada considerando uma comparação dos resultados entre as três categorias (Gestor X Discente X Especialista) que participaram do processo avaliativo da IES. Esta comparação foi feita considerando todos os atributos da base em estudo com o objetivo de detectar problemas na autoavaliação e sugerir ações paliativas/corretivas. A seguir, são apresentadas discussões relacionadas a cada um dos atributos comparados. Tais discussões foram feitas pelo autor da dissertação, a partir dos gráficos de análise comparativa apresentados anteriormente: diagnóstico geral (Seção 4.4.1), classificação de fragilidades (Seção 4.4.2) e diagnóstico detalhado (Seção 4.4.3).

Figura 16 - Comparação entre os resultados peso discente x peso gestores x nota especialista



Fonte: Autor, 2012.

Qualificação docente (QD)

Através da figura 16, nota-se que este atributo foi considerado o de menor relevância nas avaliações dos discentes e gestores contrariando a visão do especialista durante o diagnóstico geral (Seção 4.4.1). Este conhecimento evidenciado apenas após a interpretação da mineração dos dados, para discentes e gestores, é o chamado conhecimento não trivial, “escondido” na base de dados mostrando um problema evidente na concepção dos gestores e discentes, pois, a qualificação docente deveria ser um dos atributos mais importantes da base, conforme prevê a política do SINAES¹⁷. A divergência do especialista com discentes e gestores neste atributo pode indicar a necessidade de conscientização da gestão da IES em implementar políticas de apoio a qualificação docente. Ressalta-se que as IES que não têm esta política bem definida, normalmente obtêm conceitos insatisfatórios nas avaliações do MEC.

Carga horária total das disciplinas (CHT)

A relevância deste atributo tende ao mesmo valor nas avaliações do gestor e discente, e mostra que os discentes e principalmente os gestores precisam conhecer melhor a carga horária total dos cursos avaliados, pois o seu peso teve pouca relevância com relação à maioria dos atributos da base analisada, sendo extremamente divergente da visão do especialista durante o diagnóstico geral (Seção 4.4.1). Essa constatação mostra que a IES precisa aproximar o colegiado e o núcleo docente estruturante (NDE) dos cursos superiores da comunidade acadêmica para informar melhor as discussões e decisões tomadas em relação à carga horária total dos cursos de acordo com o objetivo dos mesmos.

Carga horária das disciplinas (CHD)

Este atributo teve baixa relevância nas avaliações do discente e gestores, deixando evidente um problema de desconhecimento da comunidade acadêmica quanto à importância desta variável. Este problema deve ser resolvido com a prática

¹⁷ Para entender o SINAES leia seção 2.1

de seminários e fóruns para discutir a carga horária das disciplinas, visando ao conhecimento de todos os atores institucionais quanto à importância deste atributo para o curso. Estas considerações ficam claras quando se observa a avaliação do especialista que conhece a realidade da IES, e considera este atributo como de grande relevância nesta base em estudo divergindo dos outros avaliadores em tela.

Promoção de atividades de extensão (PE)

Apesar do gráfico da Figura 14 no diagnóstico geral (Seção 4.4.1) mostrar uma convergência para o mesmo peso na avaliação do gestor e do especialista, e o discente atribuir mais relevância a esse atributo, vale ressaltar a baixa avaliação obtida na nota do especialista, visto no diagnóstico detalhado (Seção 4.4.3). Este fato evidencia que, apesar de não haver problema detectado na valorização desse atributo, por parte dos gestores e discentes, o especialista tem consciência de que as atividades de extensão devem ser ampliadas ou melhoradas. Assim, com esta ação, espera-se corrigir a fragilidade detectada.

Laboratórios específicos para cada curso (LE)

Os resultados da avaliação deste atributo mostram uma relevância ainda baixa, porém com uma avaliação mais valorada pelo discente em relação aos gestores e especialista, sendo este último o que atribuiu menor relevância. Acredita-se que a baixa valoração por parte dos gestores, deva-se em parte ao fato de que os gestores estão conscientes que para resolver este problema, muito tem a ser feito quanto aos laboratórios. Durante a avaliação detalhada relacionada a este atributo, foi percebido que se trata de uma fragilidade evidente da instituição, que devido ao alto custo para saná-la e ao baixo peso atribuído pelos especialistas, suas ações corretivas não estão classificadas entre as mais prioritárias. Porém, por se tratar de uma fragilidade que pode ter impactos em outros atributos, deve-se ter um tratamento objetivando a redução da mesma.

Estrutura curricular (EC)

Este atributo também possui baixa relevância por gestores e alunos, sendo que a avaliação do discente consegue valorizar um pouco mais o atributo, quando comparado aos gestores. Esse fato evidencia que este atributo deve ser melhor discutido na IES por se tratar de um ponto chave na análise dos cursos, e alunos e gestores ainda desconhecem algumas regulamentações da estrutura dos cursos de acordo com os dados obtidos na autoavaliação. A visão do especialista é otimista, tendo em vista que o mesmo analisa os projetos de cada curso (PPC) no momento da avaliação externa. O fato de um determinado curso ser bem avaliado não implica que todos os cursos da IES estejam no mesmo patamar. Para este atributo, gestores e discentes têm uma visão mais ampla da execução das políticas de todos os cursos da instituição. Podem existir dificuldades na operacionalização do que está proposto no projeto. Como exemplo, podemos citar a dificuldade da IES em ter professores especialistas em todas as disciplinas que compõem a grade curricular.

Contribuição das disciplinas na formação do discente (CDF)

Analisando os dados obtidos quanto a este atributo vê-se uma avaliação positiva do especialista seguido pelo discente. No entanto, o gestor considera uma relevância abaixo dos demais avaliadores. Este conhecimento revela que, uma ampla discussão deve ser realizada entre a comunidade acadêmica para detectar os possíveis problemas no processo de ensino aprendizagem que o gestor apontou na sua avaliação. Ressalta-se que a avaliação do especialista leva em consideração os conteúdos ministrados nas disciplinas podendo ser uma avaliação “fria” do processo e, neste caso, o gestor (principalmente os coordenadores de cursos) tem uma melhor visão deste atributo, pois no dia a dia é quem vive intensamente as dificuldades encontradas pelos professores em evidenciar na prática os conhecimentos adquiridos pelos discentes no curso que está fazendo. Esta dificuldade normalmente está associada à disponibilidade de equipamentos e laboratórios para atender as demandas dos cursos.

Aulas práticas de campo (APC)

Na análise deste atributo vê-se uma tendência para uma avaliação positiva entre gestores e discentes, divergindo de forma significativa da visão do especialista. Este problema ocorre em função do especialista não encontrar no momento da visita in loco a IES a documentação que comprove todas as atividades práticas de campo realizadas pelos cursos. Outra evidência é que o especialista pode considerar necessárias mais aulas práticas para atingir os objetivos que o curso se propõe a fornecer aos discentes de acordo com o Projeto pedagógico do curso. Ações devem ser tomadas para atender as demandas apontadas pelo especialista.

Número de docentes (ND)

O gráfico da Figura 16 mostra que as avaliações tendem ao mesmo ponto com uma avaliação mais otimista do gestor. No entanto, a avaliação do especialista revela que a IES ainda precisa de mais docentes para atender a demanda dos cursos. Outro ponto que vale ressaltar é que apesar de ocorridas várias contratações via concurso público, ainda existem muitos professores contratados, e dessa forma, na visão do especialista, o quadro de docentes deve ser efetivo com tempo integral e com maior titulação possível. Este fato demanda tempo e planejamento da IES.

Curso e mercado de trabalho (CMT)

A avaliação deste atributo revela um cenário positivo, com um otimismo maior dos gestores. Percebe-se pelos discentes que algumas medidas devem ser tomadas para integrar os alunos ao mercado de trabalho. Como exemplo, pode-se citar as políticas da IES voltadas para os estágios e empregabilidade dos discentes no mercado de trabalho.

Articulação de conteúdos (AC)

Este atributo apresenta uma avaliação otimista principalmente por parte dos gestores e discentes. A avaliação do especialista revela que algumas medidas devem ser tomadas para tornar melhor a articulação entre os conteúdos das

disciplinas. Como exemplo, pode-se citar o caso dos discentes do curso de engenharia que utilizam conhecimentos da estatística no 1º período, e estudam esses assuntos no 3º período na disciplina estatística básica. Essas discussões devem ocorrer entre o corpo docente e gestores dos cursos visando a corrigir eventuais distorções na disponibilização dos pré-requisitos e sequência das disciplinas a serem cursadas pelos discentes.

Laboratório de informática (LI)

A análise deste atributo apresenta um cenário otimista pelos discentes e gestores devido aos investimentos que têm sido realizados nos últimos anos. Essa visão divergiu da avaliação do especialista que considera, por exemplo, que o número de equipamentos e laboratórios devem ser aumentados para atender as necessidades dos cursos e desenvolvimento de pesquisas.

Envolvimento do professor em pesquisa (EPP)

O gráfico contido na Figura 16 revela uma avaliação positiva dos discentes gestores quanto a este atributo. Isso se deve ao fato de terem sido desenvolvidas várias ações na IES voltadas para pesquisa. No entanto, a avaliação do especialista diverge mostrando que devem ser implementadas mais ações de incentivo a pesquisa entre o corpo docente, pois a pesquisa é um dos principais pilares na avaliação de uma instituição de ensino superior pelo MEC.

Envolvimento do aluno em pesquisa (EAP)

A análise dos dados relativos a este atributo mostra um cenário extremamente otimista, principalmente pelos gestores. Este fato ocorreu devido ao aumento significativo na produção acadêmica nos últimos dois anos na IES. Esta avaliação positiva diverge consideravelmente da visão do especialista, que enxerga a pesquisa na instituição ainda embrionária, precisando de políticas mais agressivas tendo em vista que este atributo é de suma importância no desenvolvimento de qualquer instituição de ensino superior.

5 INFLUÊNCIAS DA SOLUÇÃO PROPOSTA NO APOIO A TOMADA DE DECISÃO DOS GESTORES. UM ESTUDO DE CASO COM O COLÉGIO DE DIRIGENTES DA IES

Após a descoberta de conhecimento através da Mineração de Dados visualizada e discutida na Seção 4 desta dissertação, foi realizado um seminário com o colégio de dirigentes da IES, formado pelos tomadores de decisão da instituição a exemplo: Reitor, Pró-reitores e diretores de Campi, com o objetivo de avaliar os impactos causados na gestão da IES face ao processo de descoberta de novos conhecimentos através da Mineração de Dados da autoavaliação, Conforme pode ser observado no questionário disponível no Apêndice C. O seminário buscou elucidar aspectos relacionados às divergências entre a visão dos gestores x especialista x discentes quanto às discrepâncias evidenciadas relacionadas a:

- QD: Qualificação docente;
- CHD: Carga Horária das disciplinas;
- CHT: Carga Horária Total;
- EC: Estrutura Curricular;
- AC: Articulação de conteúdos;
- LI: Laboratórios de Informática;
- EPP: Envolvimento do Professor em Pesquisa;
- EAP: Envolvimento do Aluno em Pesquisa.

Todas as oito perguntas relativas aos itens acima mencionados tinham como alternativas:

1. Irrelevante;
2. Pouco importante;
3. Importante;
4. Muito importante.

Também foram avaliadas pelo colégio de Dirigentes 4 perguntas com as alternativas acima descritas contendo subquestões relativas ao impacto na tomada de decisão a:

1. Curto prazo (duas semanas);
2. Médio prazo (seis semanas);
3. Longo prazo (dois anos).

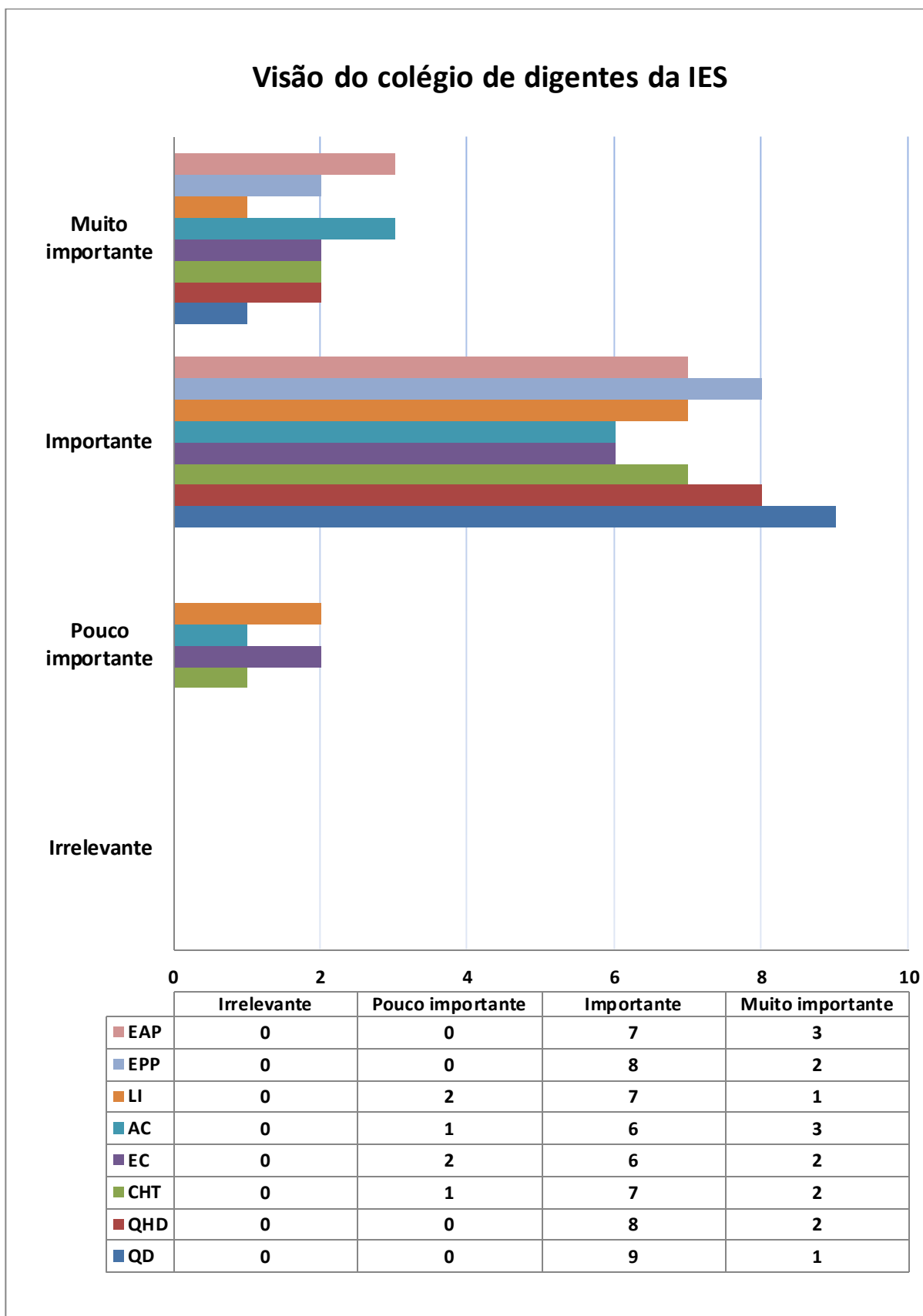
Relativas aos quesitos QD: Qualificação docente, LI: Laboratório de informática, EPA: Envolvimento dos professores em pesquisa, EAP: Envolvimento dos alunos em pesquisa.

Finalmente a Questão 13: solicitava a contribuição com críticas, sugestões e avaliação da possibilidade de uso destes conhecimentos na tomada de decisão no futuro.

Na construção do instrumento de avaliação do colégio de dirigentes da IES também foi levado em consideração às instruções da secção 2.2.5 desta pesquisa Segundo Barbetta (2007) e Richardson (1989).

A Pesquisa contou com uma participação 76% do quadro de representantes do colégio de dirigentes. O gráfico 2 apresenta um resumo das oito primeiras questões mencionadas vislumbrando mostrar a tendência dos resultados das questões avaliadas.

Gráfico 2 - Visão do colégio de dirigentes da IES quanto aos atributos analisados

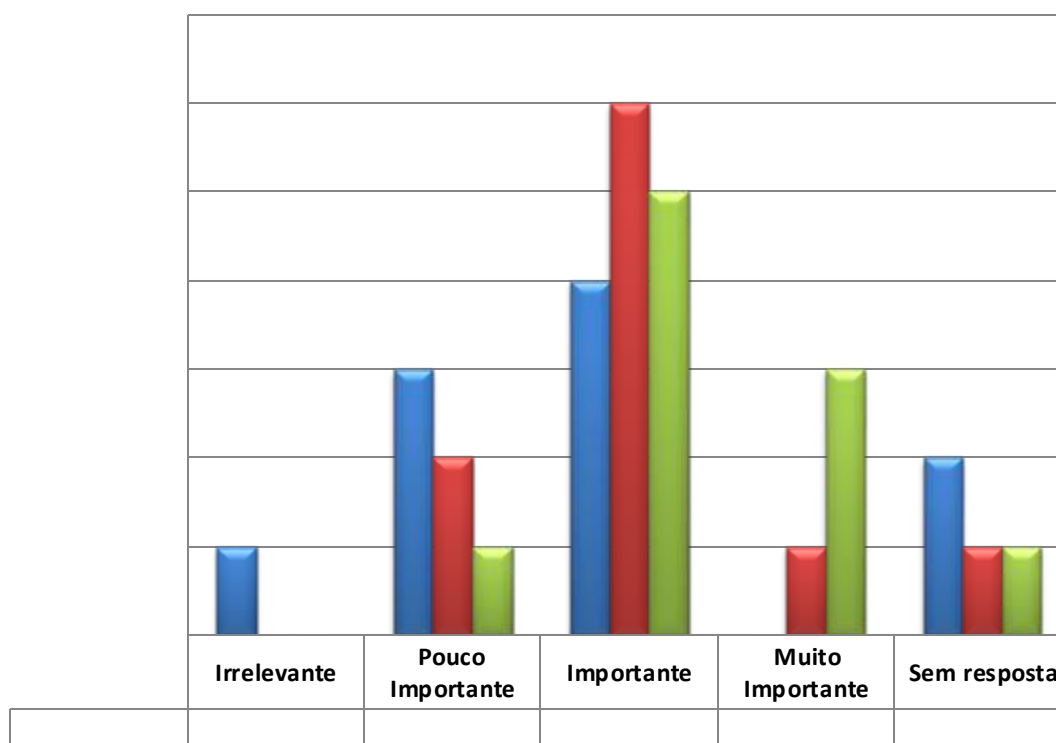


Fonte: Autor, 2012.

Os dados apontam para um resultado na faixa “Importante” tendendo a muito importante o que evidência a importância desta pesquisa como uma ferramenta de apoio a tomada de decisão das IES na visão do colégio de dirigentes.

A seguir é apresentada a discussão das quatro últimas questões do instrumento de avaliação quanto à relevância dos conhecimentos gerados pela mineração de dados em curto prazo, médio prazo, longo prazo, conforme gráficos a seguir:

Gráfico 3 - Questão 9

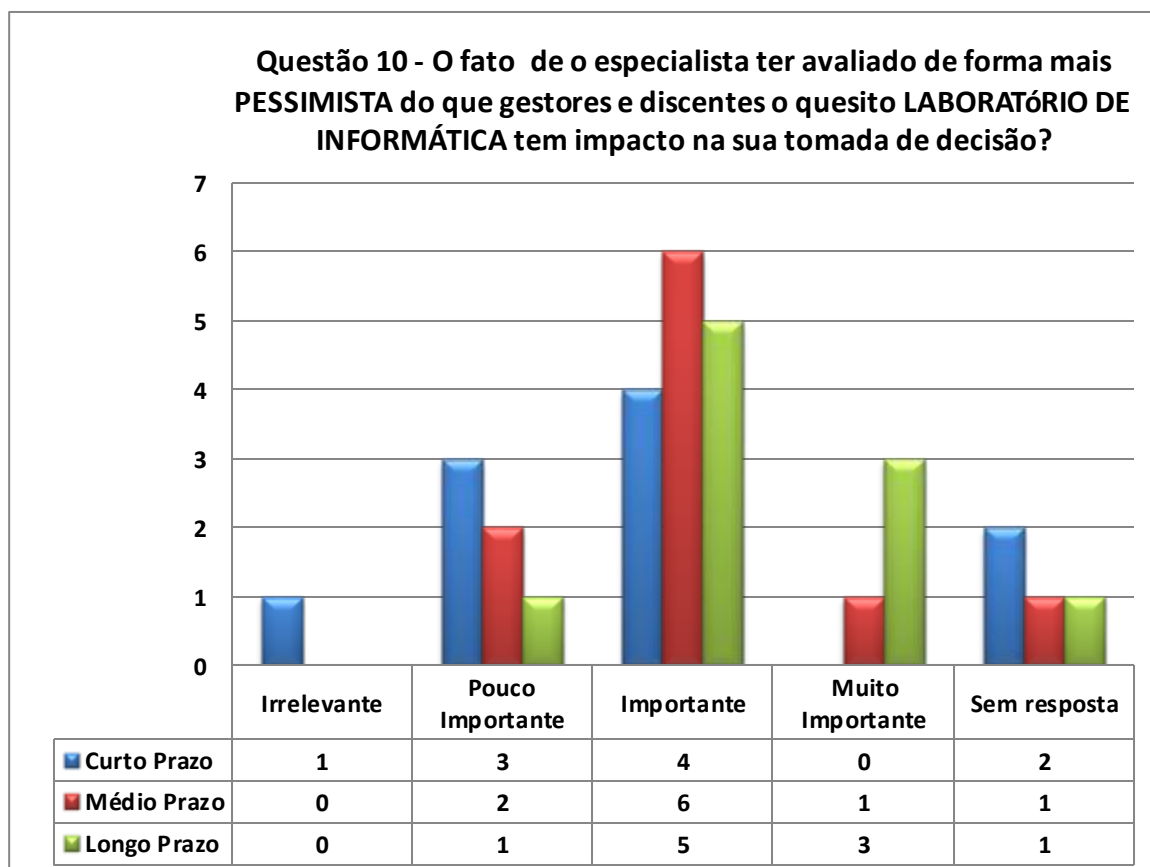


Fonte: Autor, 2012.

Os dados apontam que o quesito QD: Qualificação docente quanto à avaliação do colégio de dirigentes apontam para um resultado na faixa **importante** a curto, médio e longo prazo, e muito importante a médio e longo prazo. Também se percebe que é pouco importante em curto e médio prazo. Esta visão decorre de que

alguns gestores especificam que o processo de qualificação docente não pode ocorrer em curto prazo.

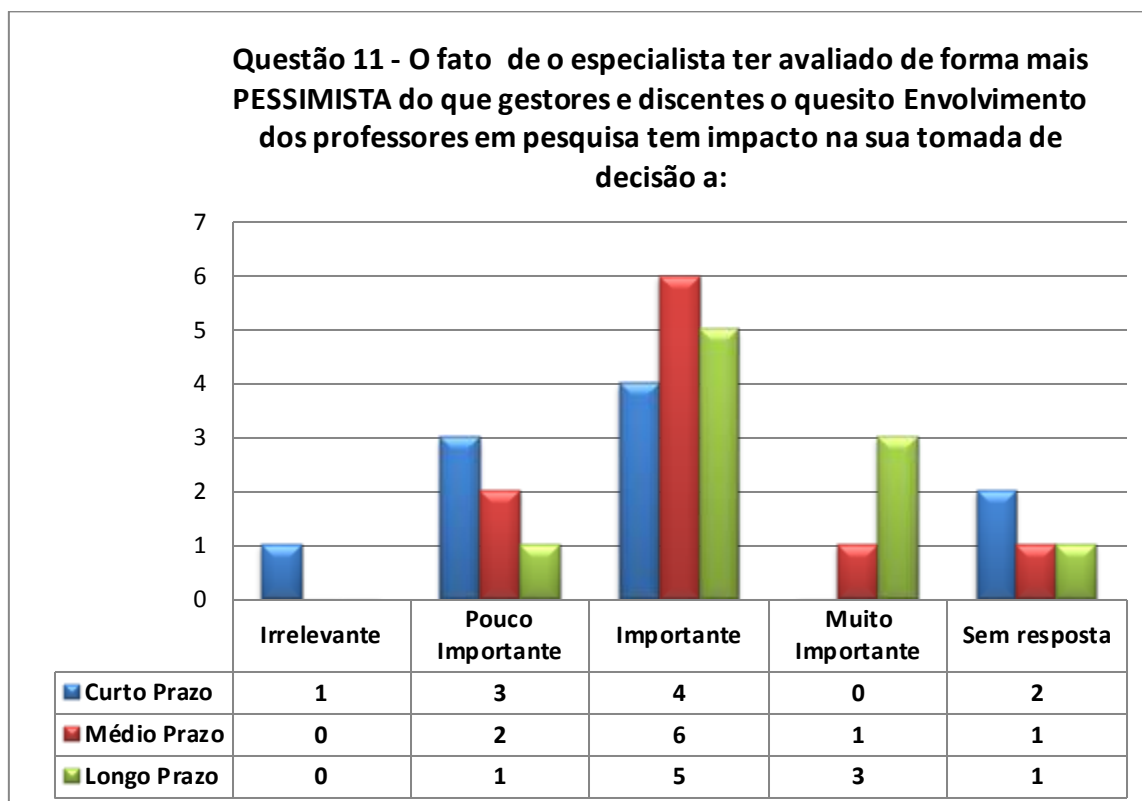
Gráfico 4 - Questão 10



Fonte: Autor, 2012.

Os dados apontam que o quesito LI: Laboratório de informática quanto à avaliação do colégio de dirigentes, apresenta um resultado na faixa “importante” a curto, médio e longo prazo e muito importante a médio e longo prazo. O fato de também ter sido considerado pouco importante a curto e médio prazo ocorre devido alguns gestores avaliarem que medidas corretivas necessitam de um período de tempo maior para serem implementadas.

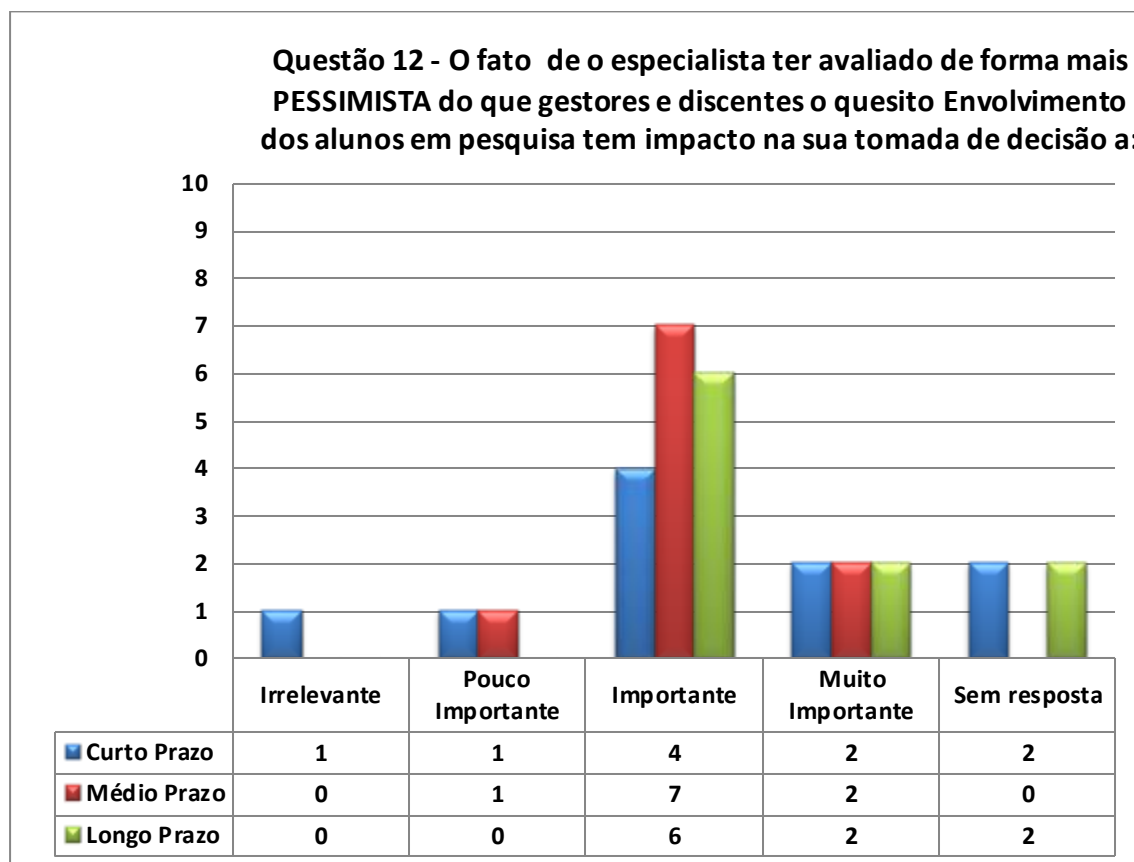
Gráfico 5 - Questão 11



Fonte: Autor, 2012.

O Gráfico 5 revela um resultado na faixa “importante” a curto, médio e longo prazo e muito importante a médio e longo prazo para o quesito Envolvimento dos professores em pesquisa. O fato de também ter sido considerado pouco importante a curto e médio prazo decorre de que políticas de apoio à pesquisa podem demandar um período de tempo maior do que o estipulado na questão em tela.

Gráfico 6 - Questão 12



Fonte: Autor, 2012.

Os dados apontam para um resultado na faixa importante tendendo a muito importante quanto ao quesito Envolvimento dos alunos em pesquisa, o que comprova a importância desses resultados para o apoio a tomada de decisão.

Quanto à questão 13, que solicitou a contribuição com críticas, sugestões e avaliação da possibilidade do uso destes conhecimentos na tomada de decisão no futuro, foi relatado pelo colégio de dirigentes, que os novos conhecimentos serão fundamentais na implementação de políticas para corrigir os problemas evidenciados no processo de descoberta de novos conhecimentos, através da mineração dos dados relativos à avaliação dos gestores, discentes e especialista. O referido colégio sugeriu a realização de um seminário com o corpo docente da IES para discussão de ações que possam ser tomadas para melhoria da IES.

6 CONCLUSÃO E TRABALHOS FUTUROS

Este trabalho apresenta uma solução de apoio para os gestores das IES. Foi apresentada uma abordagem sistemática que contempla obtenção de dados, Descoberta de Conhecimento através da Mineração de Dados e cenários de avaliação dos dados com:

- A comparação das discrepâncias envolvendo a opinião dos discentes, gestores e de um especialista em autoavaliação que teve como referência os resultados das avaliações realizadas por comissões externas designadas pelo MEC;
- Pesquisa e sugestão de melhorias no processo de avaliação;
- Realização de um seminário com o colégio de dirigentes da IES para avaliação da solução apresentada.

Os dados analisados nos resultados obtidos no Capítulo 5 desta pesquisa mostram que a solução apresentada foi considerada “**importante**” pelo colégio de dirigentes da IES sendo esta uma importante ferramenta de apoio a gestão. Portanto, a metodologia adotada nesta pesquisa tornou possível o alcance do objetivo proposto e há indícios qualitativos de que as hipóteses foram validadas.

Este trabalho apresenta como contribuições o ganho de novas informações para auxiliar os gestores das IES na tomada de decisão e a publicação de dois artigos no VI e VII Congresso de Pesquisa e Inovação da Rede Norte e Nordeste de Educação Tecnológica (CONNEPI), relacionados a este trabalho (OLIVEIRA, et al 2011), (OLIVEIRA, et al 2012).

A presente pesquisa aponta a necessidade de aprofundar os estudos em autoavaliação no sentido de construir um sistema especialista, baseado em regras, que recomende ações para minimizar os problemas detectados no processo avaliativo e desta forma contribuir com a evolução da política do SINAES.

Outras sugestões seriam:

- A aplicação dos testes de hipóteses para validar as hipóteses apresentadas no presente trabalho.
- A ampliação do foco desta pesquisa para autoavaliação da infraestrutura, desempenho dos docentes e gestores da IES.

6.1 Limitações da Pesquisa

As limitações para o desenvolvimento desta pesquisa devem-se ao fato da base de dados obtida através da autoavaliação da IES ser pequena, tendo em vista que no momento da coleta dos dados, a referida instituição possuía apenas cinco cursos superiores (Licenciatura em Matemática e Química, Engenharia Civil, Saneamento Ambiental e Turismo) no Campus Aracaju. No entanto, a amostra obtida foi significativa conforme procedimentos estatísticos evidenciados nos Capítulos 2 e 4 desta pesquisa.

Outra limitação desta pesquisa foi a base de dados obtida para os gestores da IES que neste caso é também pequena face ao universo desta categoria na instituição. Também podemos citar como limitação, o fato da presente pesquisa ter foco apenas nos cursos superiores da instituição avaliada. No entanto, a metodologia aplicada pode ser utilizada por qualquer IES, independente do tamanho da sua população, contribuindo para o apoio a tomada de decisão.

REFERÊNCIAS BIBLIOGRÁFICAS

- ADRIAANS, P.; ZANTINGE, D. **Data mining**. Addison Wesley Longman, England, 1996.
- ANDRADE, Maria Margarida de. **Introdução à metodologia do trabalho científico: elaboração de trabalhos na graduação**. 5. ed. São Paulo: Atlas, 2001.
- BATISTA, M. L. S. et al. Uma Abordagem sobre o Sistema de Informação para a Avaliação Institucional (SIAV) da Faculdade Metodista Granbery. **Revista Eletrônica da Faculdade Metodista Granbery**, v. 2, p. sis-004, 2007.
- BARBETTA, Pedro Alberto. **Estatística aplicada às ciências sociais**. 7. Ed. Florianópolis: UFSC, 2007. 315p.
- BARBETTA, Pedro Alberto; REIS, Marcelo Menezes; BORNIA, Antônio Cezar. **Estatística: para cursos de engenharia e informática**. São Paulo: Atlas, 2004.
- BERRY, M. J. A.; LINNOFF, G. **Data Mining techniques – for marketing, sales, and customer support**. 3. Ed. United States: Wiley Computer Publishing, 2011.
- BERTELLI, Eliseu Miguel. **Avaliação institucional: A trajetória da Faculdade Pato Branco-FADEP**. 2004.148 p. Dissertação (mestrado) – Pontífica Universidade Católica do Paraná, Programa de Pós- Graduação em Educação. Defesa: Curitiba, 2004.
- BRASIL. Ministério da Educação e Cultura. CONAES. **Diretrizes para a avaliação das instituições de educação superior**. Brasília: INEP, 2004.
- BRASIL. Ministério da Educação e Cultura. **Sistema Nacional de Avaliação da Educação Superior (SINAES): da concepção à regulamentação**. 2. Ed. Brasília: INEP, 2004.
- CARBONARI, Maria Elisa Ehrhardt. et al. Autoavaliação de IES: base estatística para Índices de Desempenho. **Revista de Ciências Exatas e Tecnologia**, v. 2, n. 2, 2007.
- CARVALHO, D. R. **Data mining através de indução de regras e algoritmos genéticos**. 1999. 126 f. Dissertação (Mestrado em Informática Aplicada) - Pontifícia Universidade Católica do Paraná - PUCPR, Curitiba, 1999.
- CARVALHO, D. R. **Um método híbrido árvore de decisão/algoritmo genético para data mining**, 2002. Tese (doutorado)-Pontifícia Universidade católica do Paraná, 2002.
- CARVALHO, L. A. V. de. **Data mining: a mineração de dados no marketing, medicina, economia, engenharia e Administração**. São Paulo: Érica, 2001.
- DEGENSZAJN, David Mauro; HAZZAN, Samuel; IEZZI, Gelson. **Fundamentos de matemática elementar**. 1 ed. São Paulo: Atual, 2004.
- DIAS, M. A. **Um modelo de formalização do processo de desenvolvimento de sistemas de descoberta de conhecimento em banco de dados**. 2001. 212 f. Tese

(Doutorado do Programa de Pós-Graduação em Engenharia de Produção)-UFSC. Florianópolis, Santa Catarina, 2001.

FAYYAD, U.; PIATETSKY-SHAPIO, G.; SMYTH, P. From data mining to knowledge discovery: an overview. In: **Advances in knowledge Discovery and data mining**, AAAI Press / The MIT Press, MIT, Cambridge, Massachusetts, and London, England, 1996.

FREUND, J.E.; SIMON, G.A. **Estatística aplicada: economia, administração e contabilidade**. 9. Ed. Porto Alegre: Bookman, 2000.

GASPARETTO, A. Avaliação Institucional: processo doloroso de mudança: a experiência da UESC, Ilhéus, Bahia. **Revista da Rede de Avaliação Institucional da Educação Superior**. Campinas: vol. 4, n. 3, set. 1999.

GOLDSCHMIDT, R.; PASSOS, E. **Mineração de Dados: um guia prático**. Rio de Janeiro: Campus, 2005.

KAZMIER, Leonard J. **Estatística aplicada à economia e administração**. São Paulo: Pearson Makron Books, 1982.

MARTINHAGO, Sergio. **Descoberta de conhecimento sobre o processo seletivo da UFPR**. 2005. 125 p. Dissertação (mestrado) - Universidade Federal do Paraná, Setor de Tecnologia e Setor de Ciências Exatas, Programa de Pós- Graduação em Métodos Numéricos em Engenharia. Defesa: Curitiba, 2005.

MARTINS, Gilberto de Andrade. **Estatística geral e aplicada**. 3. Ed. São Paulo: Atlas, 2009.

MATTAR, Fauze Najib. **Pesquisa de marketing**. Edição Compacta. 3. ed. São Paulo: Atlas, 2001.

MITCHELL, Tom M. **Machine learning**. McGraw-Hill, 1997.

NAVEGA, Sergio. Princípios Essenciais do Data mining. In: **Anais Eletrônicos do Info imagem 2002**, Cenadem, Novembro. Intellwise Research and Training, 2002. Disponível em: <http://www.intellwise.com/reports/i2002.pdf> Acessado em 20 out 2012.

OLIVEIRA, L. R. et al. Avaliação Institucional Análise Comparativa do Conhecimento Gerado Através de Árvore de Decisão. In: CONGRESSO DE PESQUISA E INOVAÇÃO DA REDE NORTE E NORDESTE DE EDUCAÇÃO TECNOLÓGICA ,2011,Natal. **Anais...** Natal, Editora do IFRN, 2011.p. 81-90.

OLIVEIRA, L. R. et al. A Contribuição de Mineração de Dados no Processo de Autoavaliação dos Cursos Superiores do Instituto Federal de Sergipe. In: CONGRESSO DE PESQUISA E INOVAÇÃO DA REDE NORTE E NORDESTE DE EDUCAÇÃO TECNOLÓGICA, 2012, Palmas. **Anais...** Palmas, Editora do IFTO, 2012.

OLIVEIRA, Adelize G. de. **Data Warehouse: Conceitos e Soluções**. Florianópolis: Editora Advanced, 1998. 96p.

QUINLAN, J. R. Learning efficient classification procedures and their application to chess end games. In J.G. Carbonell, R. S. Michalski, and T.M. Mitchell, editors: **Machine Learning**, v.1, Tioga, Palo Alto, USA, 1983.

REZENDE, S. O. **Sistemas Inteligentes: fundamentos e aplicações**. Barueri, SP: Manole, 2005.

RICHARDSON, Roberto Jarry. et al. **Pesquisa social: métodos e técnicas**. 3. Ed. São Paulo: Atlas, 2011.

RUMSEY, Deborah. **Estatísticas para Leigos**. Rio de Janeiro. Alta Books, 2010.

SANTOS, Eliane Gomes dos. **Avaliação institucional: Impasses no contexto universitário brasileiro**. 2010. 88 p. Dissertação (mestrado) – Universidade Veiga de Almeida, Mestrado Profissional em Psicanálise, Saúde e Sociedade. Defesa: Rio de Janeiro, 2010.

SANCHES, Raquel C. F. RAPHAEL, Hélia Sonia. Projeto Pedagógico e Avaliação Institucional: articulação e importância. **Revista da Rede de Avaliação Institucional da Educação Superior**. Ano 11. v. 11, n. 1, mar. 2006.p. 103-113.

SANCHES, Raquel Cristina Ferraroni. **Avaliação Institucional e Projeto Pedagógico: articulação imprescindível**. 2007. 185 f. Tese (Doutorado em Educação-Faculdade de Filosofia e Ciências)-Universidade Estadual Paulista, Marília, 2007.

SILVA, Ermes Medeiros da, et al. **Estatística para os cursos de administração, ciências contábeis e economia**. 2. Ed. São Paulo: Atlas, 1996.

SOBRINHO, José Dias; BALZAN, Newton César (org.). **Avaliação institucional: teoria e experiências**. 5. Ed. São Paulo: Cortez, 2011.

SOBRINHO, José Dias. **Avaliação institucional: a experiência da UNICAMP: condições, princípios, processo**. Campinas: Pro-Posições, 1995. (Serie: N. Temático: Avaliação Institucional da Universidade).

_____. **Avaliação: políticas educacionais e reformas da educação superior**. São Paulo: Cortez, 2003.

SPIEGEL, Murray R.; SCHILLER, John J; SRINIVASAN, R. Alu. **Teoria e problemas de probabilidade e estatística**. 2. Ed. Porto Alegre: Bookman, 2004.

SPIEGEL, Murray R. **Probabilidade e Estatística**. São Paulo: Pearson Education do Brasil, 1978.

STEVENSON, William J. **Estatística aplicada à administração**. São Paulo: Harper & Row do Brasil, 1981.

TAN, Pang; STEINBACH, Michael; KUMAR, Vipin. **Introdução ao datamining**. Rio de Janeiro: Ciência Moderna LTDA, 2009.

TOLEDO, Geraldo Luciano; OVALLE, Mo Isidoro. **Estatística básica**. 2. Ed. São Paulo: Atlas, 1985.

APÊNDICE A

A – 1 Distribuição de frequências

No trabalho estatístico em geral lida-se com grandes conjuntos de dados. Para obter uma boa visualização e todas as informações necessárias devem-se agrupar os dados em certos números de classes, intervalos ou categorias. Esse agrupamento é referenciado como Distribuição de Frequência ou simplesmente distribuição (FREUND; SIMON, 2000).

Tabela 12 - Exemplo de uma distribuição de frequência numérica por faixa

Número de assinantes	Número de pessoas
Menos de 1.000	244
1.000 – 3.499	157
3.500 – 9.999	96
10.000 – 19.999	37
20.000 – 49.999	24
50.000 ou mais	6
Total	564

Fonte: FREUND; SIMON, 2000.

Quando os dados apresentam-se agrupados em classes e as mesmas têm uma gama de valores como no exemplo acima, a tabela é referenciada como distribuição numérica ou quantitativa. Existem também distribuições numéricas que cada classe compreende apenas um valor (como no caso da tabela 13). Se os dados estão agrupados em categorias não numéricas, a tabela é referenciada como uma distribuição por categorias ou qualitativa conforme exemplo da tabela 14 (FREUND; SIMON, 2000).

Tabela 13 - Exemplo de uma distribuição de frequências numérica por valor

Número de visitas em Zoô local	Número de pessoas
0	90
1	72
2	26
3	8
4	3
5	0
6	1
Total	200

Fonte: FREUND; SIMON, 2000.

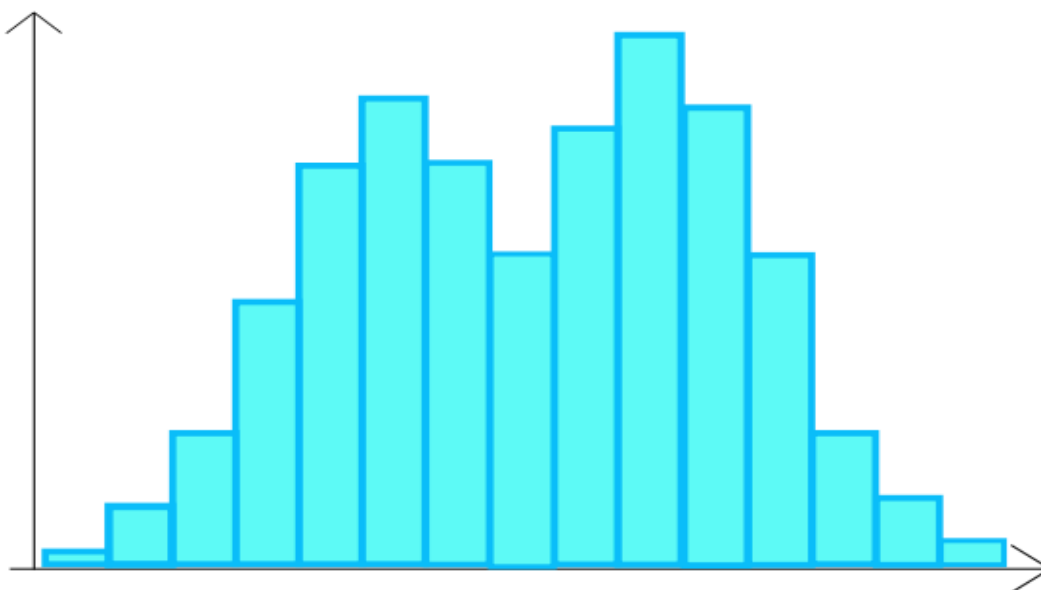
Tabela 14 - Exemplo de distribuição de frequência por categoria

Natureza de reclamações	Número de reclamações
Espaço insuficiente para as pernas	719
Assentos desconfortáveis	914
Corredores estreitos	146
Espaço insuficiente para bagagem de mão	218
Banheiros insuficientes	58
Outras queixas	384
Total	2.439

Fonte: FREUND; SIMON, 2000.

A – 2 Histogramas

Quando o principal objetivo na construção de distribuições de frequências é condensar grandes bases de dados em uma forma de fácil entendimento, é melhor representar esses dados na forma gráfica e a apresentação usual é o histograma (FREUND; SIMON, 2000). Segundo Silva et. al (1996) um Histograma pode ser definido como um conjunto de retângulos justapostos cujas bases coincidem com as amplitudes das classes e as alturas são proporcionais às frequências da classe correspondente.

Figura 17 - Histograma

Fonte: Adaptado de <http://pt.wikipedia.org/wiki/Histograma>.

Pelo exposto, nota-se que um histograma é uma forma prática e útil para mostrar graficamente os dados de uma distribuição de frequência, sendo uma ferramenta muito utilizada em pesquisas com dados quantitativos.

A – 3 Descrição de dados: medidas de tendência central

Uma medida de posição é um valor calculado com base em grupo de dados com o objetivo de descrever todos os dados. Dessa forma, pretende-se que o valor seja representativo de todos os valores do grupo (KAZMIER, 2004). No presente trabalho serão utilizadas as medidas de centralidade Média, Moda e mediana, comumente chamadas de medidas de posição. A seguir, é definida a média Segundo (KAZMIER, 2004).

A – 4 Média aritmética

A média aritmética é definida como a soma de todos os valores do grupo de dados dividido pelo número total de valores. Em geral, uma medida descritiva de uma população é representada por uma letra grega, e uma medida descritiva de uma amostra por uma letra romana. A grande diferença nas fórmulas a seguir é que

o número de elementos da população é representado por N e total de elementos da amostra é n.

A – 5 Média aritmética para dados não agrupados

Quando os dados são agrupados em uma distribuição de frequência com classes o ponto médio de cada classe representa os valores da mesma e, nesse caso, as fórmulas são dadas pelas Expressões 1 e 2.

$$\mu = \frac{\sum x}{N} \text{ Equação 4: Média aritmética para população;}$$

$$\bar{x} = \frac{\sum x}{n} \text{ Equação 5: Média aritmética para uma amostra;}$$

$$\mu = \frac{\sum (fx)}{N} \text{ Equação 6: Média aritmética ponderada para população;}$$

$$\bar{x} = \frac{\sum (fx)}{n} \text{ Equação 7: Média aritmética ponderada para uma amostra;}$$

Segundo Degenszajn et al (1999), a média aritmética pode ser afetada por valores discrepantes em um conjunto de dados, podendo se tornar uma medida pouco representativa do resumo dos dados. Para contornar essa dificuldade é definida a seguir uma medida de tendência central mais robusta a valores discrepantes, chamada mediana.

A – 6 Mediana

Dada uma sequência numérica ordenada $x_1, x_2, x_3, \dots, x_n$ de uma variável x, a mediana - indicada por MD - é definida por:

$$m_d = \begin{cases} x_{\left(\frac{n+1}{2}\right)}, & \text{se } n \text{ é ímpar} \\ \frac{x_{\left(\frac{n}{2}\right)} + x_{\left(\frac{n}{2}+1\right)}}{2}, & \text{se } n \text{ é par} \end{cases}$$

Pelo exposto, a Mediana é definida como o valor que ocupa a posição central quando os dados estão dispostos em ordem crescente e o grupo de itens possui um número N ímpar de termos. Caso N seja par a mediana é definida como a média dos termos que ocupam as posições centrais.

A mediana para dados agrupados em classes pode ser calculada pela relação abaixo de acordo com Silva et al (1996).

$$m_d = I_{m_d} + \frac{\frac{n}{2} - F_{ant}}{f_{m_d}} \cdot h \text{ Equação 8}$$

Onde

I_{m_d} - Limite inferior da classe mediana;

n - Número de elementos da série;

F_{ant} - Frequência acumulada da classe anterior à classe mediana;

f_{m_d} - Frequência simples da classe mediana e h – Amplitude do intervalo de classe.

A – 7 Moda

A moda é o valor que ocorre com mais frequência em um conjunto de dados. Para conjunto de dados com poucos elementos e que não há repetição de elementos não existe a moda. As distribuições de medidas com várias modas são denominadas multimodais (KAZMIER, 2004). De acordo com Silva et al (1996) a Moda para dados agrupados em classes pode ser calculada pela relação de Czuber a seguir:

$$m_o = I_{m_o} + \frac{f_{m_o} - f_{ant}}{2f_{m_o} - (f_{ant} + f_{post})} \cdot h \text{ Equação 8}$$

I_{m_o} – Limite inferior da classe modal;

f_{m_o} - Frequência simples da classe modal;

f_{ant} - Frequência simples da classe anterior à classe modal;

f_{post} - Frequência simples da classe posterior à classe modal;

h – Amplitude do intervalo de classe

A – 8 Resumindo dados: medidas de dispersão

Segundo (Freund; Simon, 2000) em um conjunto de dados distintos (sem repetição) a extensão de sua variabilidade é um problema sério da estatística e as medidas utilizadas para avaliar a dispersão são denominadas medidas de variabilidade ou dispersão. Considere o exemplo adaptado de (Freund; Simon, 2000), um jogador de basquete que faz 23,25 e 24 pontos em seus três primeiros jogos. Um parceiro de equipe faz 40,12 e 20 nesses jogos. Os dois jogadores têm a mesma média 24 de pontos nos três jogos. No entanto, o primeiro jogador é mais consistente, ou seja, os valores estão mais concentrados em torno da média.

A – 9 Desvio padrão

O desvio padrão é sem dúvida, a medida de variação mais utilizada e útil para evidenciar se os dados estão concentrados em torno da média ou se estão afastados da mesma (FREUND; SIMON, 2000).

Segundo (Toledo, ovalle, 1985) dado o conjunto numérico $x_1, x_2, x_3... x_n$, o desvio padrão é definido pela relação:

$$S(x) = \sqrt{\frac{\sum (x_i - \bar{x})^2}{n}}, \text{ Para dados brutos.}$$

Quando o desvio padrão for calculado a partir de uma amostra deve-se substituir n por $n-1$ na fórmula do desvio padrão anteriormente descrita. Se os dados estiverem agrupados por classes a fórmula para o cálculo do desvio padrão é dada por:

$$S(x) = \sqrt{\frac{\sum (x_i - \bar{x})^2 f_i}{n}}, \text{ Para uma população;}$$

$$S(x) = \sqrt{\frac{\sum (x_i - \bar{x})^2 f_i}{n-1}}, \text{ Para uma amostra.}$$

Aplicações do desvio padrão

Teorema de Tchebichev: Para qualquer conjunto de dados (população ou amostra) e qualquer constante $k > 1$, a proporção dos dados que devem estar a menos de k desvios-padrão da média (em qualquer dos dois lados) é ao menos $1 - \frac{1}{k^2}$. De acordo com o teorema acima podemos ter certeza que ao menos

$1 - \frac{1}{2^2} = 75\%$ dos valores em qualquer conjunto de dados estarão ao menos de dois desvios – padrão ($k = 2$) da média, de ambos os lados dela e, por conseguinte, $1 - \frac{1}{3^2} = 88,9\%$ devem estar a menos de três desvios-padrão em ambos os lados da média (FREUND; SIMON, 2000.).

Nota-se que este teorema afirma apenas a porcentagem mínima que deve estar entre certos limites, no entanto, a porcentagem de dados pode ser muito maior do que a especificada no teorema (FREUND; SIMON, 2000).

A – 10 Variância

Segundo (Toledo, ovalle, 1985) a variância é definida como o quadrado do desvio padrão. Pelo exposto, pode-se dizer que a fórmula da variância é igual à relação do desvio padrão sem o radical.

$$S^2(x) = \frac{\sum (x_i - \bar{x})^2 f_i}{n}, \text{ Para uma população.}$$

$$S^2(x) = \frac{\sum (x_i - \bar{x})^2 f_i}{n-1}, \text{ Para uma amostra.}$$

A – 11 Coeficiente de Variação de Pearson

Segundo Martins (2009) o coeficiente de variação é uma medida relativa de dispersão, enquanto a variância e o desvio padrão são medidas de dispersão absoluta e o mesmo pode ser definido pela relação $C.V = \frac{S(x)}{\bar{x}}$, sendo $\bar{x}$ a média amostral e $S(x)$ o desvio padrão amostral. Ainda de acordo com o autor, vale as seguintes regras empíricas:

Se $C.V < 0,15$ há baixa dispersão;

Se $0,15 \leq C.V < 0,30$ há média dispersão;

Se $C.V \geq 0,30$ há elevada dispersão.

De acordo com silva et al (1996) a medida de dispersão relativa (C.V) prevalece sobre a medida de dispersão absoluta por ser mais completa, tendo em vista que seu cálculo leva em consideração a medida do desvio padrão e da média.

APÊNDICE B

B – 1 Ferramentas de mineração de dados com suporte a árvore de decisão

De acordo com (Martinhago, 2006 apud Goebel e Guendwald, 1999) a escolha de uma ferramenta de descoberta de conhecimento deve considerar as seguintes características:

- A habilidade de acesso a uma variedade de fontes de dados, de forma on-line e off-line;
- A capacidade de incluir modelos de dados orientados a objetos ou modelos não padronizados (tal como multimídia, espacial ou temporal);
- A capacidade de processamento com relação ao número máximo de tabelas/tuplas/atributos;
- A capacidade de processamento com relação ao tamanho do banco de dados;
- Variedade de tipos de atributos que a ferramenta pode manipular;
- Tipo de linguagem de consulta.

A tabela 15 mostra algumas ferramentas que permitem a utilização de árvores de decisão com suas aplicações correspondentes.

Tabela 15 - Ferramentas que utilizam árvore de decisão

FERRAMENTA/ EMPRESA FORNECEDORA	APLICAÇÕES
Alice 5.1/ Isoft AS. (1998)	Política de crédito, marketing, saúde, controle de qualidade, recursos humanos.
Clementine/ Integral Solutions Limited (ISL, 1996)	Marketing direto, identificação de oportunidades de venda cruzada, retenção de cliente, previsão de lucro do cliente, detecção de fraude, segmentação e lucro do cliente.
Decision Series/ Neovista Solutions Inc. (1998)	Marketing direcionado, detecção de fraude, retenção de cliente, análise de risco, segmentação de cliente, análise de promoção.
Intelligent Miner / IBM (1997)	Segmentação de cliente, análise de conjunto de itens, detecção de fraude.
Knowledge SEEKER/ Angoss IL (Groth, 1998)	Lucro e segmentação de cliente para detecção de fraude e análise de risco, controle de processo, marketing direto.
MineSet / Silicon Graphics Computer Systems (2000) <i>Rapid-miner</i>	Áreas de saúde, farmacêutica, biotecnologia e química. Uso geral flexível e extensível via plug-in

Fonte (Adaptado DIAS, 2001).

B – 2 Operadores da ferramenta Rapid Miner

Dentre os vários operadores desta ferramenta destacaremos a seguir os que serão utilizados na Mineração dos dados desta pesquisa. A descrição é baseada nas informações contidas na própria ferramenta.

B – 2.1 Decision Tree (árvore de decisão)

Esse operador consegue aprender a partir de dados nominais e numéricos. Árvores de decisão são métodos de classificação poderosos que frequentemente podem ser entendidos com facilidade.

Com o objetivo de classificar um exemplo, a árvore será percorrida no sentido bottom-down cada nodo, em uma árvore de decisão, é rotulado com um atributo. O valor de exemplo para este atributo determina qual das bordas de saída é considerada. Para atributos nominais, nós temos uma borda de saída para cada valor de atributo possível, e para atributos numéricos as bordas de saída são rotuladas com faixas disjuntas de valores.

O aprendizado por árvore de decisão trabalha de forma similar ao C4.5 de Quinlan (ou CART). Explicando sumariamente, o algoritmo de indução de árvore trabalha da seguinte maneira: Quando um novo nodo é criado, em certo estágio, um atributo é selecionado para maximizar o poder discriminativo daquele nodo com relação aos exemplos associados particularmente à subárvore considerada. Esse poder discriminativo é mensurado por um critério que pode ser selecionado pelo usuário (ganho de informação, razão de ganho, índice gini, etc.).

O algoritmo para (é finalizado) em vários casos:

- Nenhum atributo se desenvolve até atingir certo threshold (limiar - ganho mínimo);
- A profundidade máxima da árvore é atingida;
- Existe menos que certo número de exemplos (tamanho mínimo para as divisões) na subárvore atual;
- Finalmente, a árvore é podada, impedindo que todo o poder discriminativo da árvore original seja utilizado.

B – 2.2 Apply Model (aplicação de modelo)

Esse operador aplica um modelo a um conjunto de exemplo. Modelos geralmente contêm informação acerca dos dados sobre os quais eles foram treinados. Essa informação pode ser utilizada para prever o valor de um possível rótulo desconhecido, reproduzindo o comportamento aprendido durante a fase de

treinamento. Todos os parâmetros necessários são armazenados com o objeto que representa o modelo.

B – 2.3 Performance (classificação)

Esse operador de avaliação de performance deveria ser utilizado para tarefas de classificação, isto é, em casos nos quais o rótulo do atributo tem um valor do tipo (poli-nominal). O operador Performance espera um conjunto exemplo teste como entrada, contendo um atributo com o rótulo da regra, e um com a regra de predição. Veja o operador regras de conjunto para mais detalhes. Sobre as bases destes dois atributos um vetor de performances é calculado, contendo os valores do critério performance. Se um vetor de performance foi alimentado com entradas de performance, esses valores serão mantidos se já não existirem valores no referido vetor, caso contrário, os novos valores que devem conter tal vetor serão calculados com base na média dos valores antigos e dos novos valores.

Todos os critérios de performance podem ser considerados/desconsiderados através da configuração de parâmetros booleanos. Seus valores podem ser questionados por um operador "Log" utilizando os mesmos nomes. O critério principal é usado para comparação e precisa ser especificado apenas para processos nos quais os vetores de performance são comparados, como, por exemplo, seleção de atributos ou outros processos organizadores de meta otimização. Se nenhum critério principal foi selecionado, o primeiro critério, no vetor de performance resultante, será considerado o critério principal.

B – 2.4 Weight by Relief (peso por relevância)

A recolocação/comparação mensura a importância das características com base na amostragem de exemplos e na comparação das características (atributos) entre a amostra em análise e a amostra mais próxima (mais parecida) e uma amostra de outra classe (bem diferente). Essa versão também trabalha com múltiplas classes e com conjuntos de regressão de dados. Os pesos resultantes são normalizados para um intervalo que varia entre 0 e 1.

APÊNDICE C

‘Avaliação Sinaes IFS Gestores’

Resultados do seminário com o colégio de dirigentes da IES.

Questionário

Tabela 16 - Dados consolidados

Número de registros nesta consulta:	10
Total de registros no questionário:	10
Percentagem do total:	100.00%

Fonte: Autor, 2012.

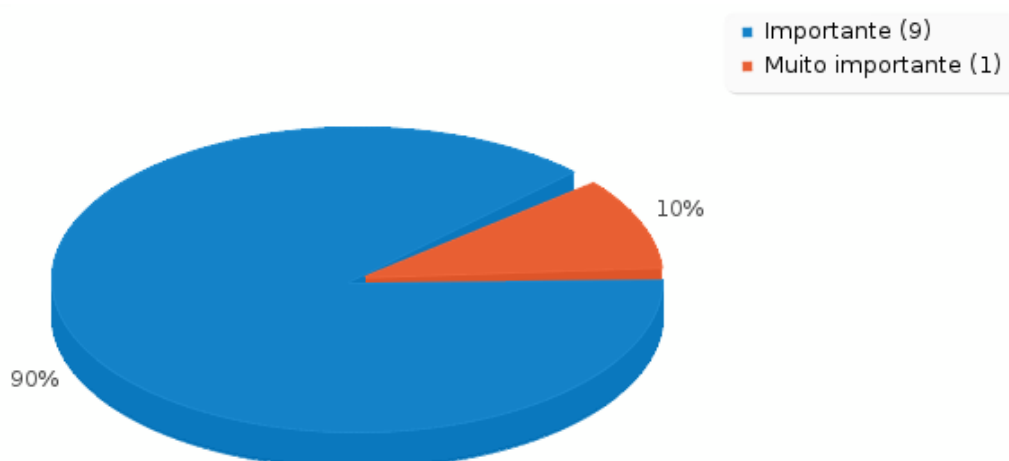
Sumário dos campos para 01

Qual o impacto na gestão do IFS face ao conhecimento gerado através da Mineração de Dados da autoavaliação relacionado às divergências entre a visão dos gestores x especialistas x alunos quanto a QD: Qualificação Docente?

Tabela 17 - Qualificação docente

Resposta	Contagem	Percentagem
Irrelevante (A1)	0	0.00%
Pouco importante (A2)	0	0.00%
Importante (A3)	9	90.00%
Muito importante (A4)	1	10.00%
Sem resposta	0	0.00%
Não completo ou Não à mostra	0	0.00%

Fonte: Autor, 2012.

Gráfico 7 - Qualificação Docente

Fonte: Autor, 2012.

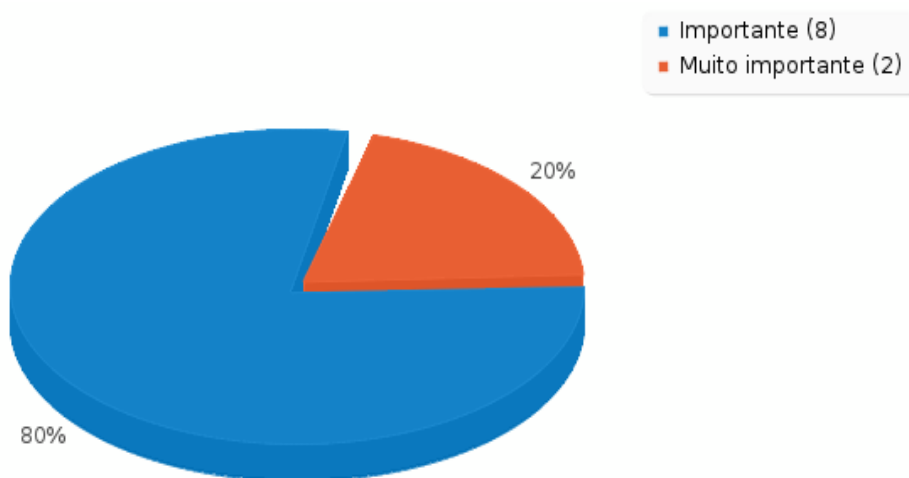
Sumário dos campos para 02

Qual o impacto na gestão do IFS face ao conhecimento gerado através da Mineração de Dados da autoavaliação relacionado às divergências entre a visão dos gestores x especialistas x alunos quanto a CHD: Carga horária das disciplinas?

Tabela 18 - Carga horária das disciplinas

Resposta	Contagem	Porcentagem
Irrelevante (A1)	0	0.00%
Pouco importante (A2)	0	0.00%
Importante (A3)	8	80.00%
Muito importante (A4)	2	20.00%
Comments	0	0.00%
Sem resposta	0	0.00%
Não completo ou Não à mostra	0	0.00%

Fonte: Autor, 2012.

Gráfico 8 - Carga horária das disciplinas

Fonte: Autor, 2012.

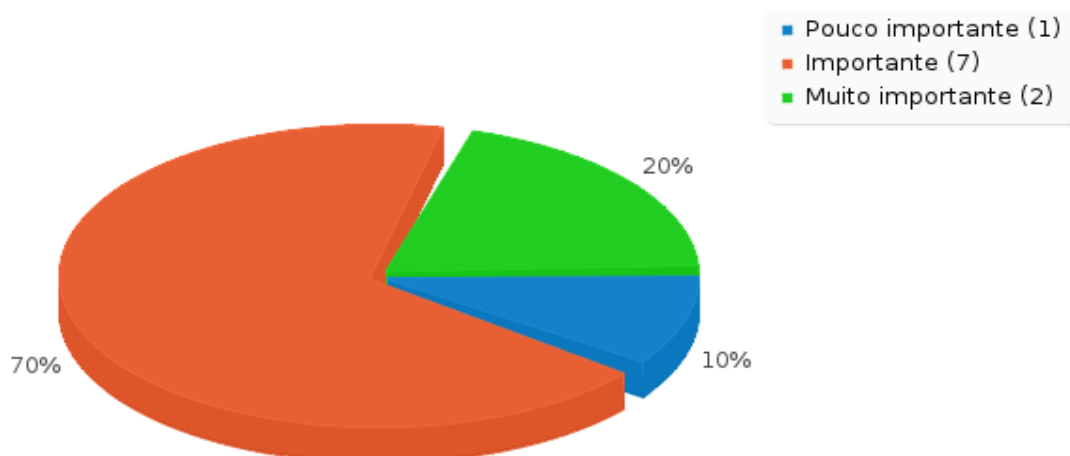
Sumário dos campos para 03

Qual o impacto na gestão do IFS face ao conhecimento gerado através da Mineração de Dados da autoavaliação relacionado às divergências entre a visão dos gestores x especialistas x alunos quanto a CHT: Carga horária total?

Tabela 19 - Carga horária total

Resposta	Contagem	Percentagem
Irrelevante (A1)	0	0.00%
Pouco importante (A2)	1	10.00%
Importante (A3)	7	70.00%
Muito importante (A4)	2	20.00%
Sem resposta	0	0.00%
Não completo ou Não à mostra	0	0.00%

Fonte: Autor, 2012.

Gráfico 9 - Carga horária total

Fonte: Autor, 2012.

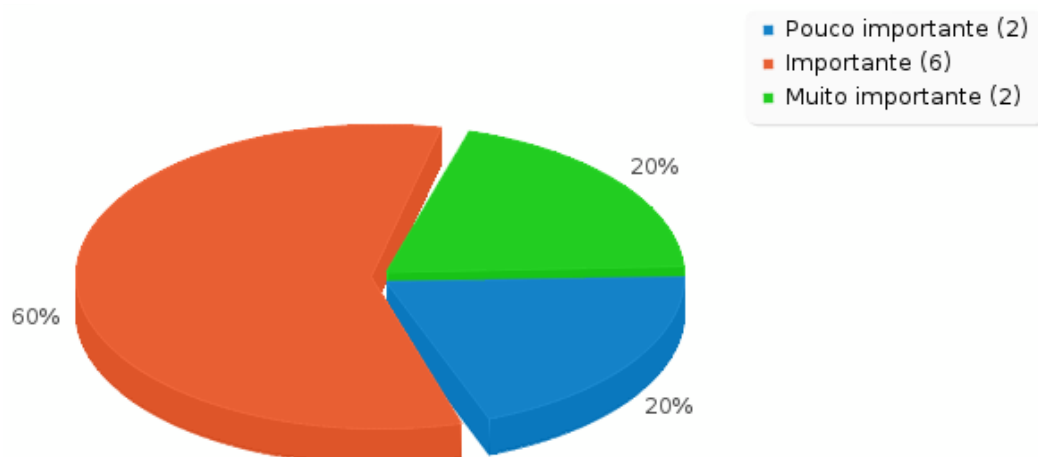
Sumário dos campos para 04

Qual o impacto na gestão do IFS face ao conhecimento gerado através da Mineração de Dados da autoavaliação relacionado às divergências entre a visão dos gestores x especialistas x alunos quanto a EC: Estrutura Curricular?

Tabela 20 - Estrutura Curricular

Resposta	Contagem	Percentagem
Irrelevante (A1)	0	0.00%
Pouco importante (A2)	2	20.00%
Importante (A3)	6	60.00%
Muito importante (A4)	2	20.00%
Sem resposta	0	0.00%
Não completo ou Não à mostra	0	0.00%

Fonte: Autor, 2012.

Gráfico 10 - Estrutura Curricular

Fonte: Autor, 2012.

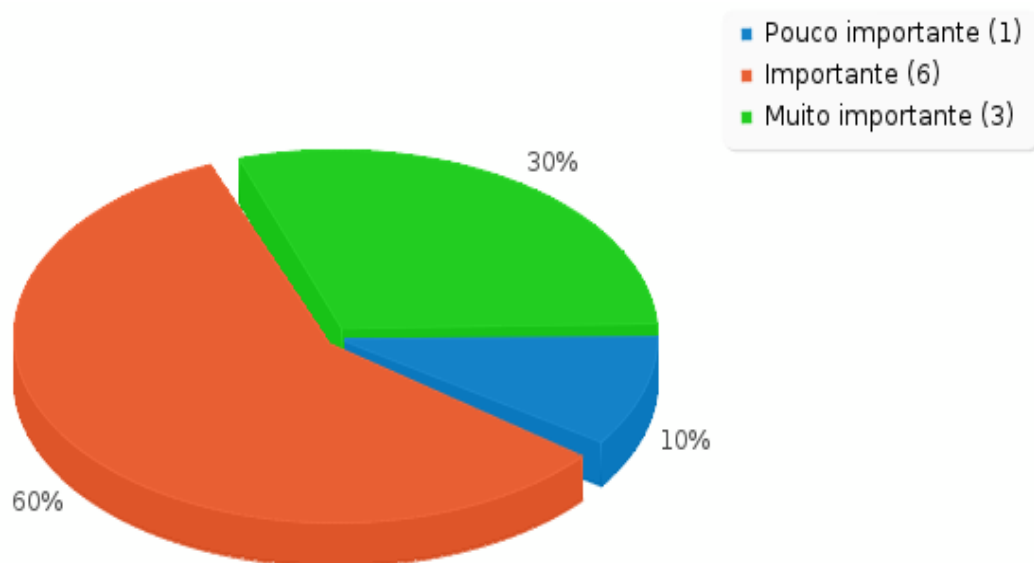
Sumário dos campos para 05

Qual o impacto na gestão do IFS face ao conhecimento gerado através da Mineração de Dados da autoavaliação relacionado às divergências entre a visão dos gestores x especialistas x alunos quanto a AC: Articulação de Conteúdos?

Tabela 21 - Articulação de Conteúdos

Resposta	Contagem	Porcentagem
Irrelevante (A1)	0	0.00%
Pouco importante (A2)	1	10.00%
Importante (A3)	6	60.00%
Muito importante (A4)	3	30.00%
Sem resposta	0	0.00%
Não completo ou Não à mostra	0	0.00%

Fonte: Autor, 2012.

Gráfico 11 - Articulação de Conteúdos

Fonte: Autor, 2012.

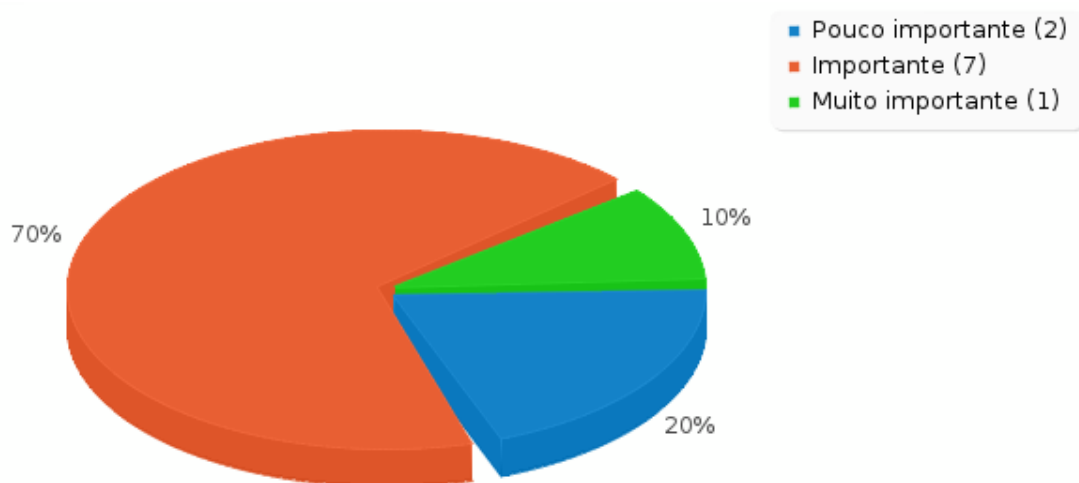
Sumário dos campos para 06

Qual o impacto na gestão do IFS face ao conhecimento gerado através da Mineração de Dados da autoavaliação relacionado às divergências entre a visão dos gestores x especialistas x alunos quanto a LI: Laboratórios de informática?

Tabela 22 - Laboratórios de informática

Resposta	Contagem	Percentagem
Irrelevante (A1)	0	0.00%
Pouco importante (A2)	2	20.00%
Importante (A3)	7	70.00%
Muito importante (A4)	1	10.00%
Sem resposta	0	0.00%
Não completo ou Não à mostra	0	0.00%

Fonte: Autor, 2012.

Gráfico 12 - Laboratórios de informática

Fonte: Autor, 2012.

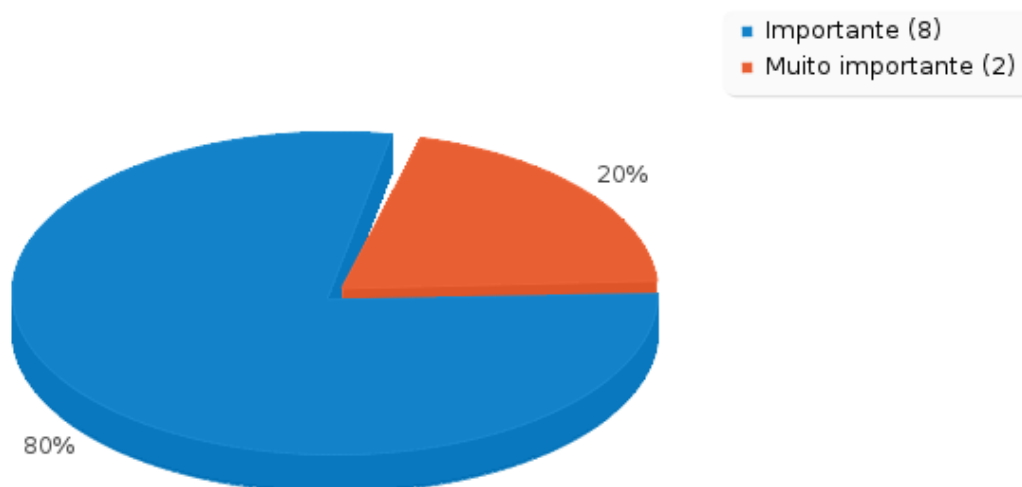
Sumário dos campos para 07

Qual o impacto na gestão do IFS face ao conhecimento gerado através da Mineração de Dados da autoavaliação relacionado às divergências entre a visão dos gestores x especialistas x alunos quanto a EP: Envolvimento do professor em pesquisa?

Tabela 23 - Envolvimento do professor em pesquisa

Resposta	Contagem	Porcentagem
Irrelevante (A1)	0	0.00%
Pouco importante (A2)	0	0.00%
Importante (A3)	8	80.00%
Muito importante (A4)	2	20.00%
Sem resposta	0	0.00%
Não completo ou Não à mostra	0	0.00%

Fonte: Autor, 2012.

Gráfico 13 - Envolvimento do professor em pesquisa

Fonte: Autor, 2012.

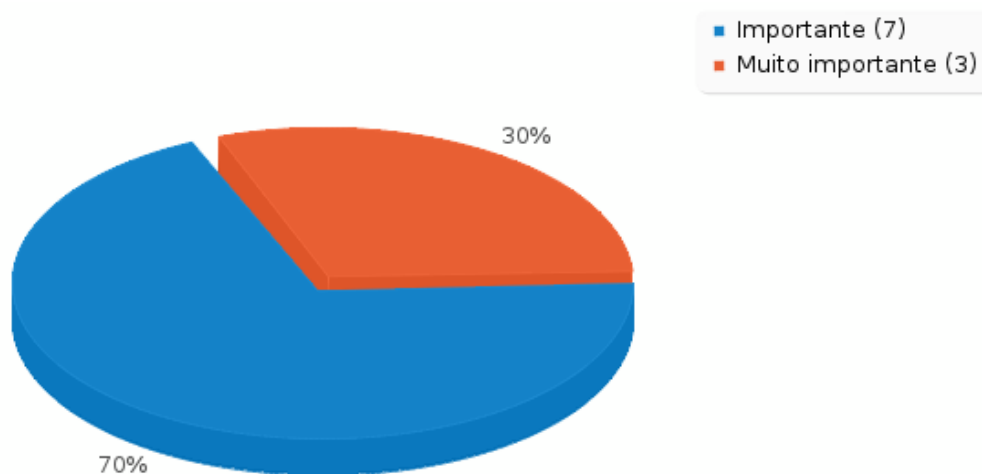
Sumário dos campos para 08

Qual o impacto na gestão do IFS face ao conhecimento gerado através da Mineração de Dados da autoavaliação relacionado às divergências entre a visão dos gestores x especialistas x alunos quanto a EAP: Envolvimento do aluno em pesquisa?

Tabela 24 - Envolvimento do aluno em pesquisa

Resposta	Contagem	Porcentagem
Irrelevante (A1)	0	0.00%
Pouco importante (A2)	0	0.00%
Importante (A3)	7	70.00%
Muito importante (A4)	3	30.00%
Sem resposta	0	0.00%
Não completo ou Não à mostra	0	0.00%

Fonte: Autor, 2012.

Gráfico 14 - Envolvimento do aluno em pesquisa

Fonte: Autor, 2012.

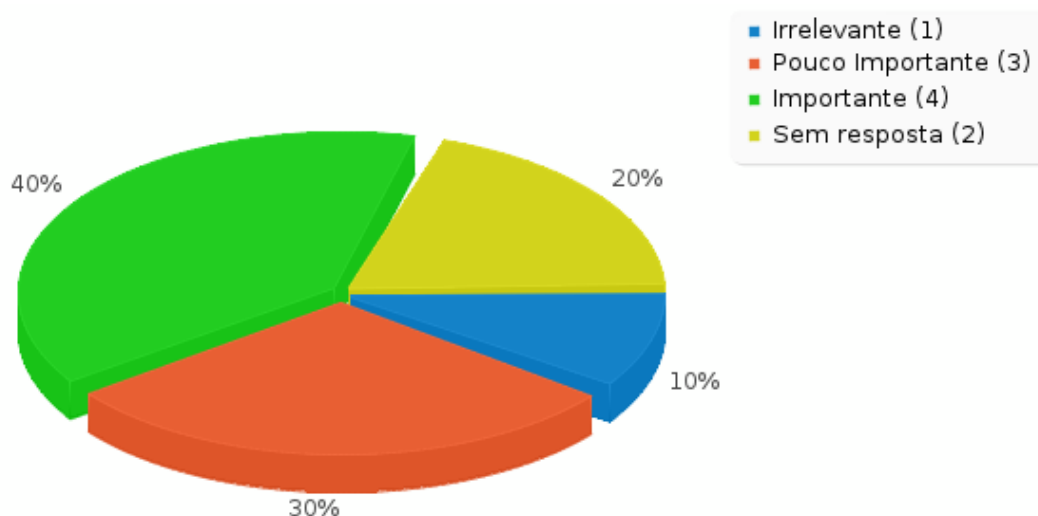
Sumário dos campos para 09(SQ001)

O fato de o especialista ter avaliado de forma mais otimista do que gestores e discentes o quesito Qualificação docente teria qual impacto na sua tomada de decisão? [Curto prazo (duas semanas)]

Tabela 25 - Qualificação docente a curto prazo

Resposta	Contagem	Porcentagem
Irrelevante (A1)	1	10.00%
Pouco importante (A2)	3	30.00%
Importante (A3)	4	40.00%
Muito importante (A4)	0	0.00%
Sem resposta	2	20.00%
Não completo ou Não à mostra	0	0.00%

Fonte: Autor, 2012.

Gráfico 15 - Qualificação docente a curto prazo

Fonte: Autor, 2012.

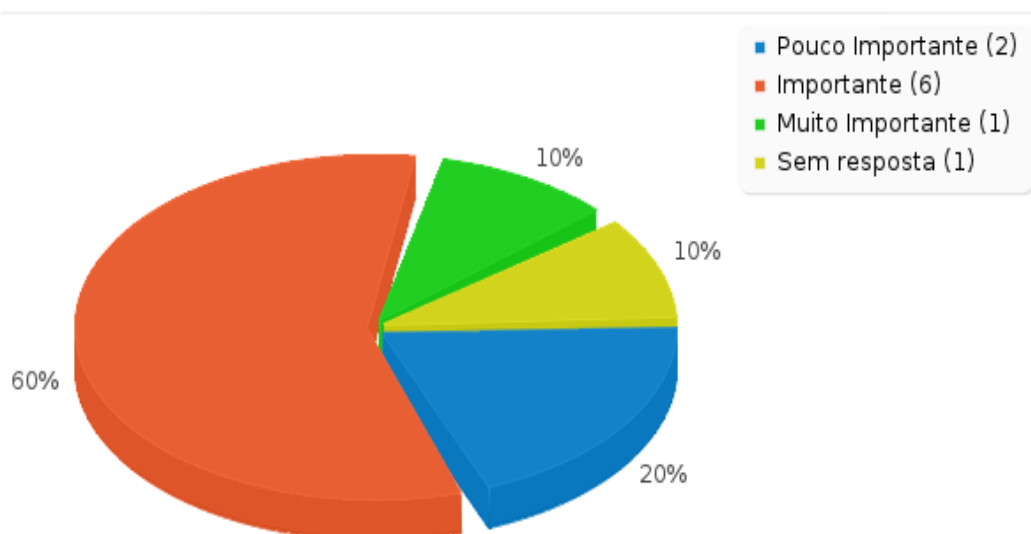
Sumário dos campos para 09(SQ002)

O fato de o especialista ter avaliado de forma mais otimista do que gestores e discentes o quesito Qualificação docente teria qual impacto na sua tomada de decisão? [Médio Prazo (seis meses)]

Tabela 26 - Qualificação docente a médio prazo

Resposta	Contagem	Porcentagem
Irrelevante (A1)	0	0.00%
Pouco importante (A2)	2	20.00%
Importante (A3)	6	60.00%
Muito importante (A4)	1	10.00%
Sem resposta	1	10.00%
Não completo ou Não à mostra	0	0.00%

Fonte: Autor, 2012.

Gráfico 16 - Qualificação docente a médio prazo

Fonte: Autor, 2012.

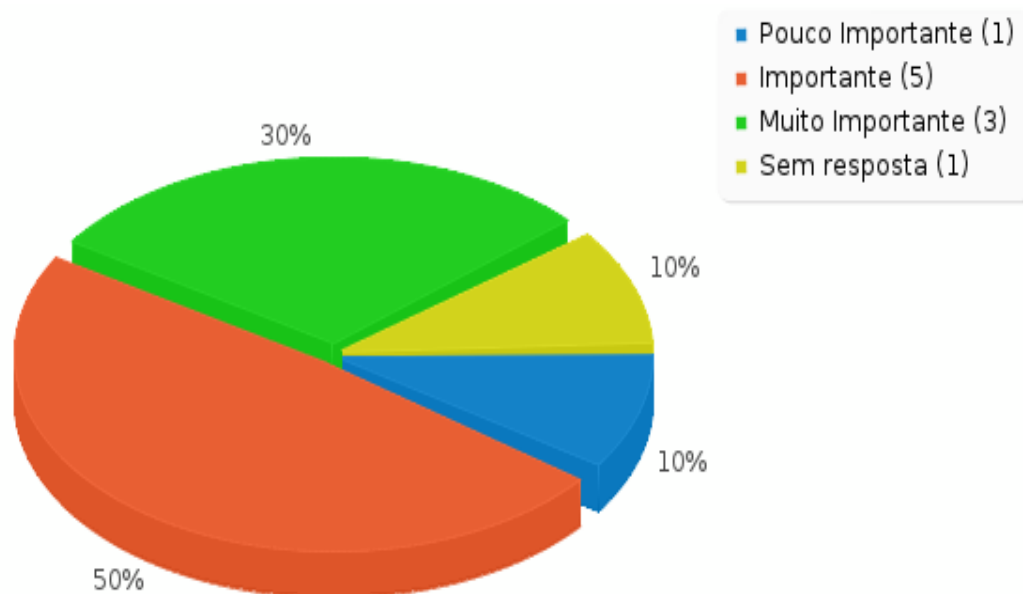
Sumário dos campos para 09(SQ003)

O fato de o especialista ter avaliado de forma mais otimista do que gestores e discentes o quesito Qualificação docente teria qual impacto na sua tomada de decisão? [Longo Prazo (dois anos)]

Tabela 27 - Qualificação docente a longo prazo

Resposta	Contagem	Porcentagem
Irrelevante (A1)	0	0.00%
Pouco importante (A2)	1	10.00%
Importante (A3)	5	50.00%
Muito importante (A4)	3	30.00%
Sem resposta	1	10.00%
Não completo ou Não à mostra	0	0.00%

Fonte: Autor, 2012.

Gráfico 17 - Qualificação docente a longo prazo

Fonte: Autor, 2012.

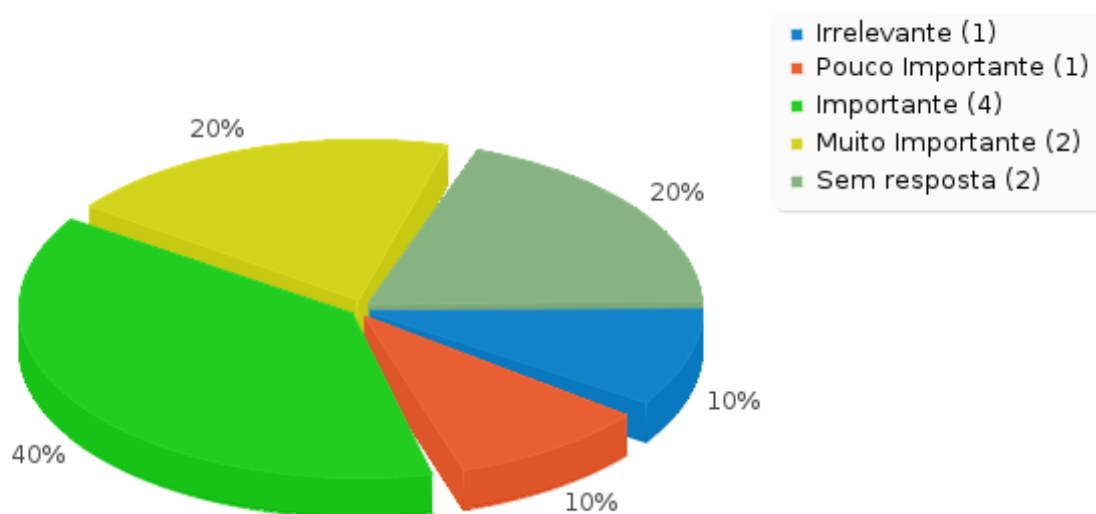
Sumário dos campos para 10(SQ001)

O fato de o especialista ter avaliado de forma mais PESSIMISTA do que gestores e discentes o quesito LABORATÓRIO DE INFORMÁTICA tem impacto na sua tomada de decisão? [Curto prazo (duas semanas)]

Tabela 28 - Laboratório de informática a curto prazo

Resposta	Contagem	Porcentagem
Irrelevante (A1)	1	10.00%
Pouco importante (A2)	1	10.00%
Importante (A3)	4	40.00%
Muito importante (A4)	2	20.00%
Sem resposta	2	20.00%
Não completo ou Não à mostra	0	0.00%

Fonte: Autor, 2012.

Gráfico 18 - Laboratório de informática a curto prazo

Fonte: Autor, 2012.

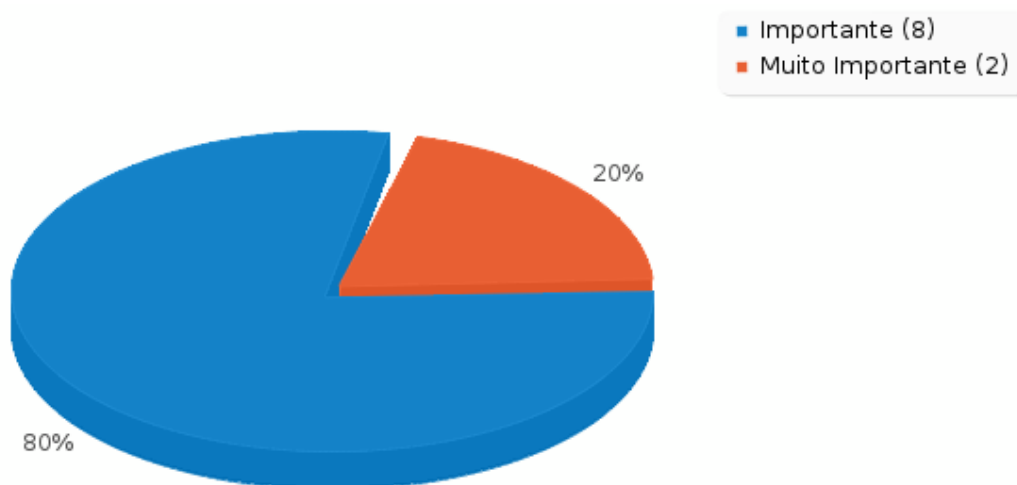
Sumário dos campos para 10(SQ002)

O fato de o especialista ter avaliado de forma mais PESSIMISTA do que gestores e discentes o quesito LABORATÓRIO DE INFORMÁTICA tem impacto na sua tomada de decisão?[Médio Prazo (seis meses)]

Tabela 29 - Laboratório de informática a médio prazo

Resposta	Contagem	Percentagem
Irrelevante (A1)	0	0.00%
Pouco importante (A2)	0	0.00%
Importante (A3)	8	80.00%
Muito importante (A4)	2	20.00%
Sem resposta	0	0.00%
Não completo ou Não à mostra	0	0.00%

Fonte: Autor, 2012.

Gráfico 19 - Laboratório de informática a médio prazo

Fonte: Autor, 2012.

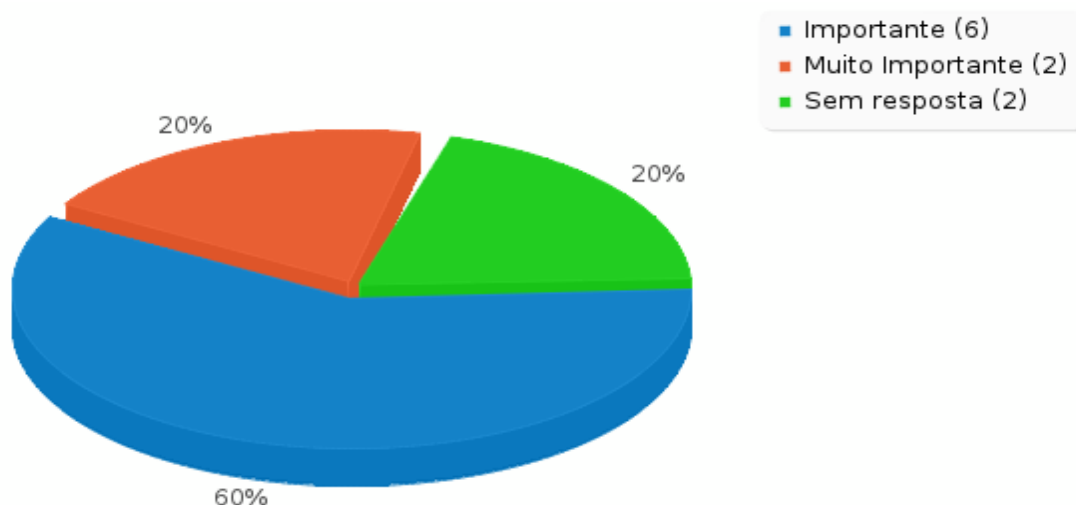
Sumário dos campos para 10(SQ003)

O fato de o especialista ter avaliado de forma mais PESSIMISTA do que gestores e discentes o quesito LABORATÓRIO DE INFORMÁTICA tem impacto na sua tomada de decisão? [Longo Prazo (dois anos)]

Tabela 30 - Laboratório de informática a longo prazo

Resposta	Contagem	Porcentagem
Irrelevante (A1)	0	0.00%
Pouco importante (A2)	0	0.00%
Importante (A3)	6	60.00%
Muito importante (A4)	2	20.00%
Sem resposta	2	20.00%
Não completo ou Não à mostra	0	0.00%

Fonte: Autor, 2012.

Gráfico 20 - Laboratório de informática a longo prazo

Fonte: Autor, 2012.

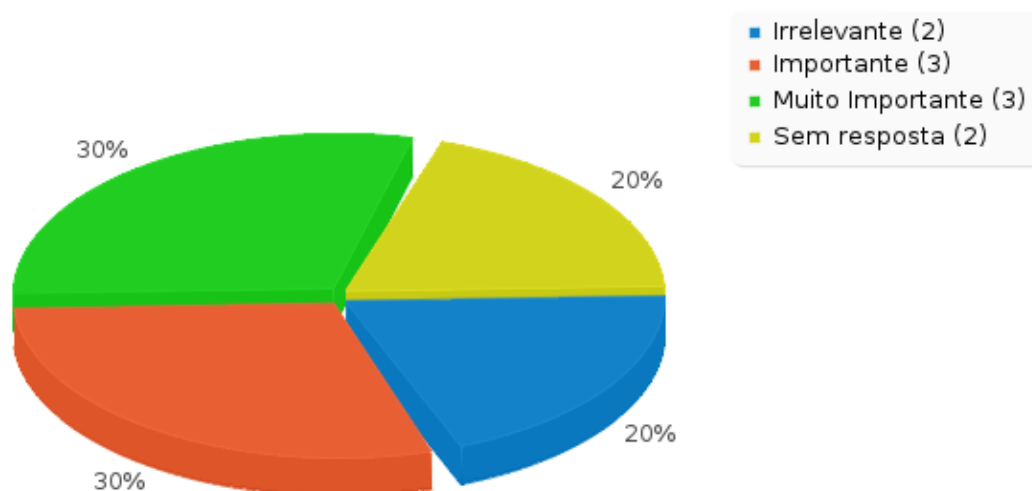
Sumário dos campos para 11(SQ001)

O fato de o especialista ter avaliado de forma mais PESSIMISTA do que gestores e discentes o quesito Envolvimento dos professores em pesquisa tem impacto na sua tomada de decisão a: [Curto prazo (duas semanas)]

Tabela 31 - Envolvimento dos professores em pesquisa a curto prazo

Resposta	Contagem	Porcentagem
Irrelevante (A1)	2	20.00%
Pouco importante (A2)	0	0.00%
Importante (A3)	3	30.00%
Muito importante (A4)	3	30.00%
Sem resposta	2	20.00%
Não completo ou Não à mostra	0	0.00%

Fonte: Autor, 2012.

Gráfico 21 - Envolvimento dos professores em pesquisa a curto prazo

Fonte: Autor, 2012.

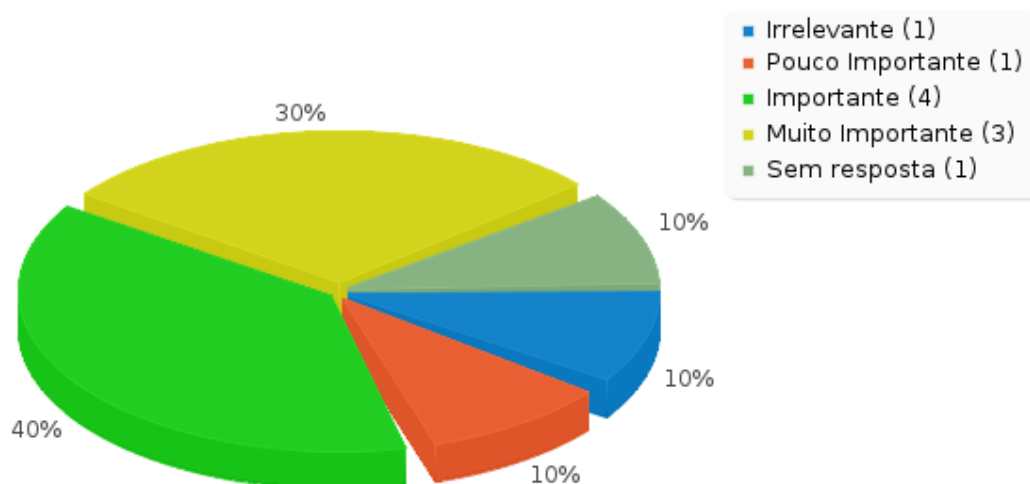
Sumário dos campos para 11(SQ002)

O fato de o especialista ter avaliado de forma mais PESSIMISTA do que gestores e discentes o quesito Envolvimento dos professores em pesquisa tem impacto na sua tomada de decisão a: [Médio Prazo (seis meses)]

Tabela 32 - Envolvimento dos professores em pesquisa a médio prazo

Resposta	Contagem	Porcentagem
Irrelevante (A1)	1	10.00%
Pouco importante (A2)	1	10.00%
Importante (A3)	4	40.00%
Muito importante (A4)	3	30.00%
Sem resposta	1	10.00%
Não completo ou Não à mostra	0	0.00%

Fonte: Autor, 2012.

Gráfico 22 - Envolvimento dos professores em pesquisa a médio prazo

Fonte: Autor, 2012.

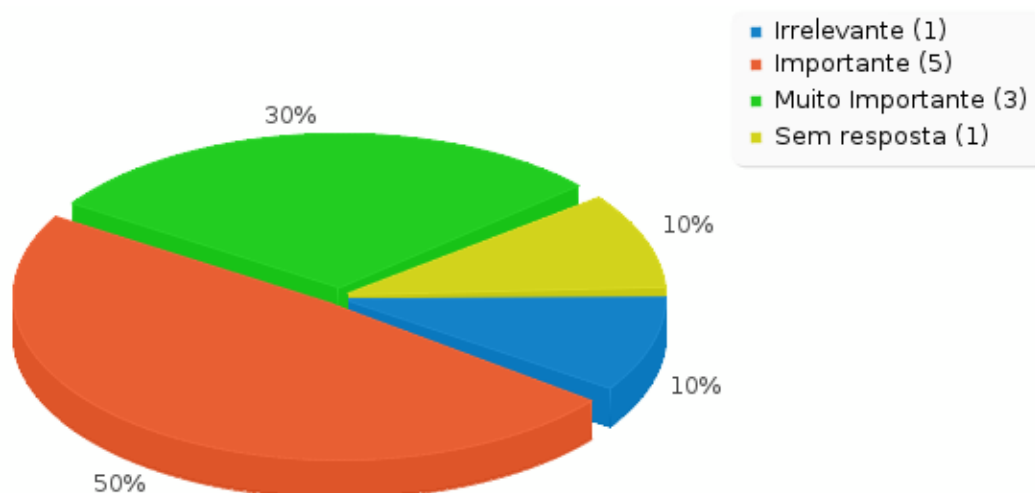
Sumário dos campos para 11(SQ003)

O fato de o especialista ter avaliado de forma mais PESSIMISTA do que gestores e discentes o quesito Envolvimento dos professores em pesquisa tem impacto na sua tomada de decisão a: [Longo Prazo (dois anos)]

Tabela 33 - Envolvimento dos professores em pesquisa a longo prazo

Resposta	Contagem	Porcentagem
Irrelevante (A1)	1	10.00%
Pouco importante (A2)	0	0.00%
Importante (A3)	5	50.00%
Muito importante (A4)	3	30.00%
Sem resposta	1	10.00%
Não completo ou Não à mostra	0	0.00%

Fonte: Autor, 2012.

Gráfico 23 - Envolvimento dos professores em pesquisa a longo prazo

Fonte: Autor, 2012.

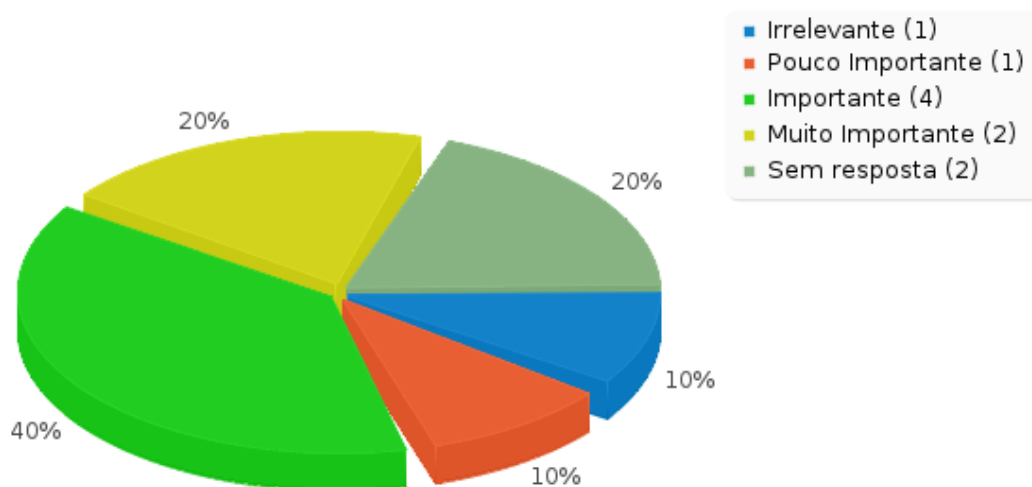
Sumário dos campos para 12(SQ001)

O fato de o especialista ter avaliado de forma mais PESSIMISTA do que gestores e discentes o quesito Envolvimento dos alunos em pesquisa tem impacto na sua tomada de decisão a: [Curto prazo (duas semanas)]

Tabela 34 - Envolvimento dos alunos em pesquisa a curto prazo

Resposta	Contagem	Porcentagem
Irrelevante (A1)	1	10.00%
Pouco importante (A2)	1	10.00%
Importante (A3)	4	40.00%
Muito importante (A4)	2	20.00%
Sem resposta	2	20.00%
Não completo ou Não à mostra	0	0.00%

Fonte: Autor, 2012.

Gráfico 24 - Envolvimento dos alunos em pesquisa a curto prazo

Fonte: Autor, 2012.

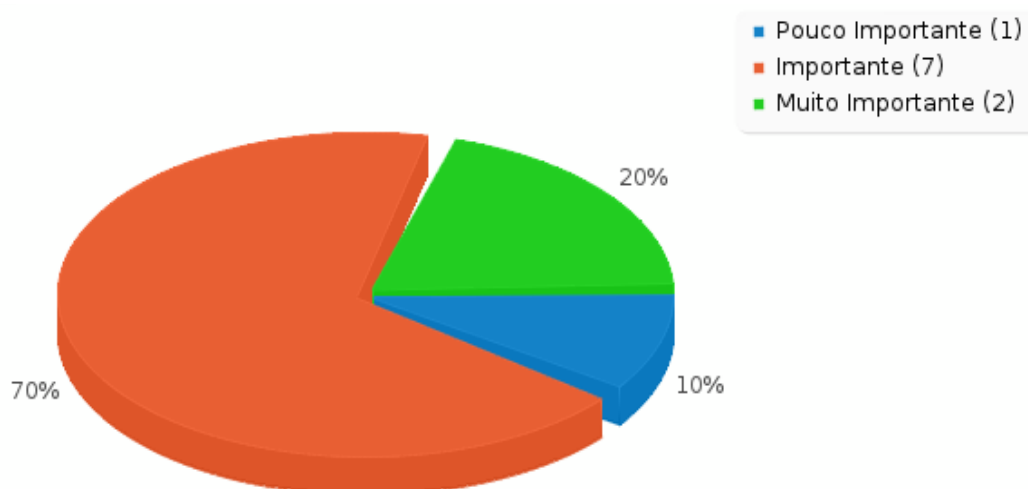
Sumário dos campos para 12(SQ002)

O fato de o especialista ter avaliado de forma mais PESSIMISTA do que gestores e discentes o quesito Envolvimento dos alunos em pesquisa tem impacto na sua tomada de decisão a: [Médio Prazo (seis meses)]

Tabela 35 - Envolvimento dos alunos em pesquisa a médio prazo

Resposta	Contagem	Porcentagem
Irrelevante (A1)	0	0.00%
Pouco importante (A2)	1	10.00%
Importante (A3)	7	70.00%
Muito importante (A4)	2	20.00%
Sem resposta	0	0.00%
Não completo ou Não à mostra	0	0.00%

Fonte: Autor, 2012.

Gráfico 25 - Envolvimento dos alunos em pesquisa a médio prazo

Fonte: Autor, 2012.

Sumário dos campos para 12(SQ003)

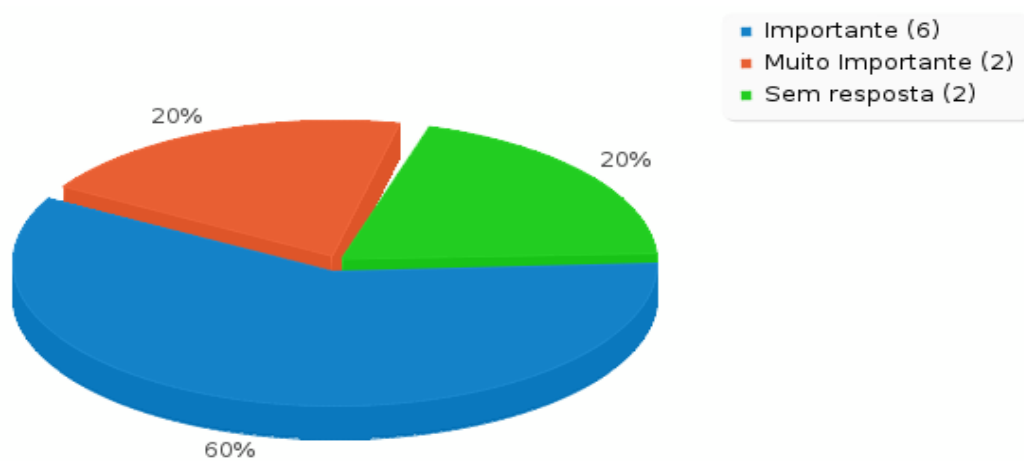
O fato de o especialista ter avaliado de forma mais PESSIMISTA do que gestores e discentes o quesito Envolvimento dos alunos em pesquisa tem impacto na sua tomada de decisão a: [Longo Prazo (dois anos)]

Tabela 36 - Envolvimento dos alunos em pesquisa a longo prazo

Resposta	Contagem	Porcentagem
Irrelevante (A1)	0	0.00%
Pouco importante (A2)	0	0.00%
Importante (A3)	6	60.00%
Muito importante (A4)	2	20.00%
Sem resposta	2	20.00%
Não completo ou Não à mostra	0	0.00%

Fonte: Autor, 2012.

Gráfico 26 - Envolvimento dos alunos em pesquisa a longo prazo



Fonte: Autor, 2012.

APÊNDICE D



INSTRUMENTO DE AVALIAÇÃO DOS GESTORES

INSTITUTO FEDERAL DE SERGIPE COMISSÃO PRÓPRIA DE AVALIAÇÃO-CPA

AVALIAÇÃO INSTITUCIONAL DOS GESTORES EM RELAÇÃO AOS CURSOS SUPERIORES DO IFS

Caro (a) gestor (a),

No sentido de melhorarmos o processo ensino/aprendizagem e dimensionarmos as potencialidades e fragilidades do Instituto Federal de Sergipe, bem como atender as exigências do MEC, através da Lei Federal de Nº 10.861, que criou o Sistema Nacional de Avaliação da Educação Superior (SINAES), solicitamos sua colaboração no sentido de responder esta avaliação institucional.

Não há necessidade de se identificar e seus dados serão mantidos em absoluto sigilo. Para responder as questões, observe a legenda no rodapé da página. Agradecemos desde já sua contribuição para a melhoria da qualidade no IFS.

Participe! O IFS necessita de sua avaliação!

Comissão de Avaliação



AVALIAÇÃO DOS CURSOS PELOS GESTORES (Marcar com “X”)

Como você avalia a qualidade dos cursos superiores com relação a:

	D	F	R	B	MB
1. Estrutura Curricular					
2. Carga horária das disciplinas					
3. Carga horária total do curso					
4. Laboratórios de Informática					
5. Número de Docentes					
6. Qualificação dos docentes					
7. Promoção de atividades de Extensão (seminários, palestras, etc.).					
8. Envolvimento de professores em projetos de pesquisa					
9. Envolvimento de alunos em projetos de pesquisa					
10. Articulação de conteúdos entre as disciplinas do curso					
11. Laboratório de atividades específicas do curso					
12. Aulas práticas de campo (visitas técnicas e microestágios)					
13. Atendimento do curso às necessidades do mercado de trabalho					
14. Contribuição das disciplinas para a formação técnica do aluno					

D = Desconheço; F = Fraco; R = Ruim; B = Bom; MB = Muito Bom;



INSTRUMENTO DE AVALIAÇÃO DOS DISCENTES

INSTITUTO FEDERAL DE SERGIPE COMISSÃO PRÓPRIA DE AVALIAÇÃO-CPA

AVALIAÇÃO INSTITUCIONAL DOS DISCENTES EM RELAÇÃO AOS CURSOS SUPERIORES DO IFS

Caro (a) aluno (a),

No sentido de melhorarmos o processo ensino/aprendizagem e dimensionarmos as potencialidades e fragilidades do Instituto Federal de Sergipe, bem como atender as exigências do MEC, através da Lei Federal de Nº 10.861, que criou o Sistema Nacional de Avaliação da Educação Superior (SINAES), solicitamos sua colaboração no sentido de responder esta avaliação institucional.

Não há necessidade de se identificar e seus dados serão mantidos em absoluto sigilo. Para responder as questões, observe a legenda no rodapé da página. Agradecemos desde já sua contribuição para a melhoria da qualidade no IFS.

Participe! O IFS necessita de sua avaliação!

Comissão de Avaliação



AVALIAÇÃO DOS CURSOS PELOS DISCENTES (Marcar com “X”)

Como você avalia a qualidade dos cursos superiores com relação a:

	D	F	R	B	MB
1. Estrutura Curricular					
2. Carga horária das disciplinas					
3. Carga horária total do curso					
4. Laboratórios de Informática					
5. Número de Docentes					
6. Qualificação dos docentes					
7. Promoção de atividades de Extensão (seminários, palestras, etc.).					
8. Envolvimento de professores em projetos de pesquisa					
9. Envolvimento de alunos em projetos de pesquisa					
10. Articulação de conteúdos entre as disciplinas do curso					
11. Laboratório de atividades específicas do curso					
12. Aulas práticas de campo (visitas técnicas e microestágios)					
13. Atendimento do curso às necessidades do mercado de trabalho					
14. Contribuição das disciplinas para a formação técnica do aluno					

D = Desconheço; F = Fraco; R = Ruim; B = Bom; MB = Muito Bom;

AVALIAÇÃO INSTITUCIONAL DISCENTE ANTES DOS AJUSTES

Caro(a) aluno(a),

No sentido de melhorarmos o processo ensino/aprendizagem e dimensionarmos as potencialidades e fragilidades do Instituto Federal de Sergipe, bem como atender as exigências do MEC, através da Lei Federal de Nº 10.861, que criou o Sistema Nacional de Avaliação da Educação Superior (SINAES), **solicitamos sua colaboração no sentido de responder esta avaliação institucional.**

Não há necessidade de colocar nome ou matrícula e seus dados serão mantidos em absoluto sigilo. Para responder as questões a partir do item 3 observe a legenda no rodapé da página. Agradecemos desde já sua contribuição para a melhoria da qualidade no IFS.

Participe! O IFS necessita da sua avaliação!

Comissão de Avaliação

Como você avalia a qualidade do seu curso com relação a:	D	F	R	B	MB
1 - Estrutura curricular					
2 - Carga horária das disciplinas					
3 - Carga horária total do curso					
4 - Turno de funcionamento					
5 - Orientação aos alunos na matrícula					
6 - Coordenação de estágio					
7 - Laboratórios de informática					
8 - Número de docentes					
9 - Qualificação dos docentes					
10 - Promoção de atividades de extensão (seminários, palestras, etc.)					
11 - Envolvimento de professores em projetos de pesquisa					
12 - Envolvimento de alunos em projetos de pesquisa					
13 - Articulação de conteúdos entre as disciplinas do curso					
14 - Laboratório de atividades específicas do curso					
15 - Aulas práticas de campo (visitas técnicas e microestágios)					
16 - Didática dos professores					
17 - Critérios de avaliação nas disciplinas do curso					
18 - Atendimento do curso às necessidades do mercado de trabalho					
19 - Contribuição das disciplinas para a formação técnica do aluno					
20 - Contribuição das disciplinas para a formação cidadã do aluno					

D = Desconheço; F = Fraco; R = Ruim; B = Bom; MB = Muito Bom;