



UNIVERSIDADE ESTADUAL PAULISTA  
“JÚLIO DE MESQUITA FILHO”  
Campus de São José do Rio Preto

Alciano Gustavo Genovez de Oliveira

Aplicação da Lógica Nebulosa em um Classificador para Identificação de  
Perfis por Aspectos Cognitivos

São José do Rio Preto – SP  
2016

Alciano Gustavo Genovez de Oliveira

Aplicação da Lógica Nebulosa em um Classificador para Identificação de  
Perfis por Aspectos Cognitivos

Dissertação apresentada como parte dos requisitos para obtenção do título de Mestre em Ciência da Computação, junto ao Programa de Pós-Graduação em Ciência da Computação, do Instituto de Biociências, Letras e Ciências Exatas da Universidade Estadual Paulista “Júlio de Mesquita Filho”, Campus de São José do Rio Preto.

Orientador: Prof. Dr. Norian Marranghello

São José do Rio Preto – SP  
2016

Oliveira, Alciano Gustavo Genovez de.

Aplicação da lógica nebulosa em um classificador para identificação de perfis por aspectos cognitivos / Alciano Gustavo Genovez de Oliveira. -- São José do Rio Preto, 2016  
95 f. : il., tabs.

Orientador: Norian Marranghello

Dissertação (mestrado) – Universidade Estadual Paulista “Júlio de Mesquita Filho”, Instituto de Biociências, Letras e Ciências Exatas

1. Computação - Matemática. 2. Mineração de dados (Computação)  
3. Lógica difusa. 4. Computação - Classificação. 5. Inteligências múltiplas. 6. Ensino a distância - Ensino auxiliado por computador.  
I. Universidade Estadual Paulista "Júlio de Mesquita Filho". Instituto de Biociências, Letras e Ciências Exatas. II. Título.

CDU – 518.72

Ficha catalográfica elaborada pela Biblioteca do IBILCE  
UNESP - Câmpus de São José do Rio Preto

Alciano Gustavo Genovez de Oliveira

Aplicação da Lógica Nebulosa em um Classificador para Identificação de  
Perfis por Aspectos Cognitivos

Dissertação apresentada como parte dos requisitos para obtenção do título de Mestre em Ciência da Computação, junto ao Programa de Pós-Graduação em Ciência da Computação, do Instituto de Biociências, Letras e Ciências Exatas da Universidade Estadual Paulista “Júlio de Mesquita Filho”, Campus de São José do Rio Preto.

Comissão Examinadora

Prof. Dr. Aleardo Manacero Junior (Presidente).  
UNESP – São José do Rio Preto, SP.

Prof. Dr. Rodrigo Capobianco Guido.  
UNESP – São José do Rio Preto, SP.

Prof. Dr. Henrique Dezani  
FATEC – São José do Rio Preto, SP.

São José do Rio Preto – SP  
29 de novembro de 2016

*A meus pais, esposa, Tia Vera e amigos.*

## **AGRADECIMENTOS**

Agradeço primeiramente a Deus pelo dom da vida e por minha família, é graças a ele que estou concluindo este mestrado. É Deus quem tem me dado forças não só para concluir este trabalho, mas também para superar todas as adversidades que surgiram em minha vida. É graças a Deus que tenho conseguido conviver com perdas insuperáveis.

A minha mãe Sonia e em memória de meu pai Alceu, por estarem sempre ao meu lado e por terem me ensinado virtudes que me foram fundamentais para a formação de minha pessoa.

A minha esposa pela compreensão e pela ajuda me motivando, cobrando e dando forças para que eu não desistisse durante as dificuldades.

Aos meus amigos que de várias formas também contribuíram para a conclusão de mais esta etapa de minha vida, amigos que não são muitos, mas são valiosos e importantes.

Ao meu orientador Prof. Dr. Norian Marranghello, que foi de fundamental importância para a conclusão deste trabalho, meu muito obrigado pela atenção, disposição, orientação e comprometimento durante o período em que tive a oportunidade de trabalhar com o Sr.

*O maior sentido da vida é lutar, mas não para vencer e sim para não ser  
vencido.  
Bruce Lee.*

## RESUMO

Atualmente, as instituições de ensino, em sua grande maioria, estão disponibilizando cursos na modalidade de ensino a distância com o intuito de possibilitar aos indivíduos que não podem frequentar as aulas regularmente a realização de estudos a distância. Em muitos casos, os indivíduos desistem dessa modalidade de ensino sem concluir os estudos, isso se deve a vários fatores, sendo um deles, a dificuldade no entendimento do conteúdo disponibilizado durante as aulas devido a forma com que o mesmo é apresentado. Esta dissertação apresenta a utilização da lógica nebulosa em um classificador computacional, que tem por objetivo classificar indivíduos por aspectos cognitivos que estão relacionados com a Teoria das Inteligências Múltiplas propostas, originalmente, por Howard Gardner. O resultado dessa classificação possibilita direcionar os indivíduos para ambientes de ensino em que o conteúdo esteja adequado ao seu perfil cognitivo. Os testes foram realizados utilizando uma ferramenta acadêmica de mineração de dados que possibilitou determinar padrões cognitivos em cada indivíduo pela inserção de dados de entrada obtidos por meio da aplicação de um questionário e retornando os aspectos cognitivos mais aflorados de cada indivíduo. Após validados, os resultados mostraram aproximadamente 67% das classificações condizentes com os aspectos cognitivos identificados em aula.

**Palavras-chave:** Lógica Nebulosa. Teoria das Inteligências Múltiplas. Classificador Computacional.



## **ABSTRACT**

*Most educational institutions are offering distance-learning courses, in order to enable individuals who cannot attend classes regularly, to conduct their studies. In many cases, individuals drop out of this mode of teaching without completing their studies. One of the factors for that being the difficulty in understanding the content due to the way content is presented. This dissertation presents the use of fuzzy logic in a computational classifier that aims to classify individuals by cognitive aspects related to Howard Gardner's Theory of Multiple Intelligences. The outcome of such a classifier makes it possible to direct individuals to learning environments in which the content is presented according to his/her main cognitive profile. The tests were performed using a data mining academic tool that allowed determining cognitive patterns for each individual by inputting data obtained from a questionnaire and returning the cognitive aspect most outlined in each individual. After validated, about 67% of the classification outcomes were considered in accordance to the cognitive aspects identified in classroom observations.*

**Keywords:** *Fuzzy Logic. Theory of Multiple Intelligences. Computational Classifier.*

## LISTA DE FIGURAS

Figura 2.1: Exemplo de função de pertinência .....	24
Figura 2.2: Sistema lógico nebuloso .....	26
Figura 2.3: Inferência nebulosa. ....	27
Figura 2.4: Processo de Defuzzificação. ....	28
Figura 2.5: Exemplo de partição nebulosa utilizando o método <i>Simple Fuzzy Grid</i> ..	42
Figura 2.6: Representação da grade utilizada na solução do problema proposto.....	43
Figura 2.7: Estrutura das tabelas de regras nebulosas correspondendo a $L=2$ e $L=3$ .....	47
Figura 3.1: Etapas de processamento realizadas pelo Weka.....	51
Figura 3.2: Representação de um subespaço no espaço bidimensional. ....	58
Figura 3.3: Resultado de um teste de classificação .....	60
Figura 3.4: Matriz de confusão .....	61
Figura 3.5: Valores de acurácia fornecidos pelo classificador .....	62
Figura 3.6: Resultado do agrupamento .....	64
Figura 3.7: Grupos de aspectos cognitivos identificados para um indivíduo .....	65
Figura 4.1: Resultado do processo de classificação com baixo nível de acurácia ....	67

## LISTA DE TABELAS

Tabela 3.1: Respostas correspondentes a cada valor numérico respondido .....	52
Tabela 3.2: Modelo do questionário.....	54
Tabela 3.3: Exemplo de questionário respondido.....	54
Tabela 3.4: Grupo de indivíduos que responderam ao questionário .....	55
Tabela 3.5: Informações referentes ao espaço amostral.....	57
Tabela 3.6: Leitura das linhas da matriz de confusão. ....	61
Tabela 3.7: Leitura das colunas da matriz de confusão. ....	62
Tabela 4.1: Resultado dos testes realizados para a classificação por aspecto cognitivo. ....	66
Tabela 4.2: Resultado da classificação por aspecto cognitivo utilizando o filtro <i>Resample</i> .....	71
Tabela 4.3: Resultado dos testes realizados com e sem o filtro <i>Resample</i> .....	71
Tabela 5.1: Classificação dos grupos de indivíduos por aspecto cognitivo .....	72
Tabela 5.2: Aspecto cognitivo mais aflorado para os indivíduos.....	73
Tabela 5.3: Aspecto cognitivo mais aflorado por área de conhecimento .....	74
Tabela 5.4: Análise do desempenho escolar dos indivíduos .....	76

# SUMÁRIO

1. INTRODUÇÃO .....	13
1.1 MOTIVAÇÃO .....	13
1.2 OBJETIVOS .....	15
1.3 ORGANIZAÇÃO DO TRABALHO.....	16
2. FUNDAMENTAÇÃO TEÓRICA.....	17
2.1. LÓGICA NEBULOSA.....	17
2.1.1 Teoria dos Conjuntos Nebulosos.....	17
2.1.2 Fundamentos Utilizados na Lógica Nebulosa .....	19
2.1.3 Definições e Operações.....	20
2.1.4 Propriedades Algébricas .....	21
2.1.5 Variáveis Linguísticas.....	22
2.1.6 Sistema Lógico Nebuloso .....	26
2.1.7 Aplicações da Lógica Nebulosa .....	28
2.2 TEORIA DAS INTELIGÊNCIAS MÚLTIPLAS .....	29
2.2.1 Tipos de Inteligências Múltiplas.....	29
2.3 SISTEMAS DE RECOMENDAÇÃO.....	34
2.3.1 Classificação dos Sistemas de Recomendação.....	34
2.3.2 Tipos de Descoberta de Conhecimento .....	37
2.4 BALANCEAMENTO DAS CLASSES DE DADOS .....	39
2.5 MÉTODOS DE CLASSIFICAÇÃO DE DADOS .....	40
2.5.1 Método <i>Simple Fuzzy Grid</i> .....	41
2.5.2 Método de Regras Nebulosas “se”...”então” Distribuídas.....	46
2.6 MÓTODO DE AGRUPAMENTO DOS DADOS DE SAÍDA .....	48
3 METODOLOGIA .....	50
3.1 WEKA ( <i>WAIKATO ENVIRONMENT FOR KNOWLEDGE ANALYSIS</i> ).....	50
3.1.1 Processamento realizado pelo Weka .....	51
3.2 QUESTIONÁRIO UTILIZADO .....	52
3.3 GRUPO DE INDIVÍDUOS SELECIONADOS PARA OS TESTES .....	55
3.4 ANALISE ESTATÍSTICA DAS AMOSTRAS .....	56
3.5 REALIZANDO A CLASSIFICAÇÃO DE UM INDIVÍDUO .....	58

3.6	ENTENDENDO O RESULTADO DA CLASSIFICAÇÃO .....	60
3.7	REALIZANDO O AGRUPAMENTO DAS RESPOSTAS EM CLASSES.....	64
3.8	ENTENDENDO OS RESULTADOS DO AGRUPOAMENTO.....	65
4	RESULTADOS OBTIDOS PARA OS PRIMEIROS TESTES.....	66
4.1	ANALISE DOS RESULTADOS OBTIDOS PARA OS PRIMEIROS TESTES .....	66
4.2	UTILIZANDO O FILTRO <i>RESAMPLE</i> .....	69
4.3	REFAZENDO OS TESTES.....	70
5	RESULTADOS OBTIDOS .....	72
5.1	VALIDAÇÃO DOS RESULTADOS .....	75
6	CONCLUSÃO .....	79
	REFERÊNCIAS.....	82
	APÊNDICE A.....	86
	APÊNDICE B.....	88
	APÊNDICE C .....	90
	APÊNDICE D .....	92
	APÊNDICE E.....	94

## **1. INTRODUÇÃO**

O mercado de trabalho, por diversas vezes, necessita de profissionais qualificados para atender suas necessidades, isso faz com que sejam geradas vagas de emprego que, em alguns casos, têm dificuldades para serem preenchidas devido à falta de capacitação dos profissionais.

Hoje em dia não é tão raro encontrar alunos em cursos técnicos ou de nível superior que têm dificuldades para interpretar textos e realizar cálculos simples. Desta forma, para que se possa atingir o mesmo nível de conhecimento em relação há alguns anos, é necessário se capacitar cada vez mais.

Um dos motivos para a falta de profissionais capacitados é a dificuldade de que as pessoas têm em frequentar os cursos de média e longa duração devido à falta de tempo por estarem comprometidas com família, trabalho e outros afazeres.

Dessa forma, uma das alternativas são os cursos a distância que disponibilizam ferramentas de apoio ao ensino para que o aluno possa estudar no horário em que mais lhe convier, permitindo a continuidade de sua formação profissional.

Em alguns casos, essas ferramentas de apoio ao ensino podem ter índice de evasão maior que os cursos presenciais. Um dos prováveis motivos é a ausência da presença física do professor que avalia o aluno e direciona as aulas de forma a explorar as características deste que possam favorecer o ensino. Dessa forma, o aluno deve se adaptar à metodologia de ensino proposta pela ferramenta de ensino, sendo que essa adaptação, na maioria das vezes, pode não acontecer, levando o aluno a deixar os estudos.

### **1.1 MOTIVAÇÃO**

Uma das dificuldades enfrentadas pela educação, nos dias de hoje, é proporcionar ao aluno, que realiza os estudos por meio de ferramentas de apoio ao ensino à distância, um ambiente agradável e de conteúdo de fácil entendimento com o intuito de facilitar o aprendizado e diminuir a evasão.

Levando-se em consideração esse contexto, propõe-se o desenvolvimento de um classificador computacional (RICCI et al, 2011) que possa ser utilizado para

classificar indivíduos de acordo com seus aspectos cognitivos. Esse classificador é fundamentado na utilização da lógica nebulosa (PEDRYCZ, 1989) como método computacional responsável pela classificação dos indivíduos e a Teoria das Inteligências Múltiplas, proposta por Gardner (GARDNER, 2011), é utilizada como base para os aspectos cognitivos utilizados como parâmetros de classificação dos indivíduos.

Freedman realizou estudos sobre os resultados obtidos no processo de ensino e aprendizagem em disciplinas ministradas por professores, que fazem uso das Inteligências Múltiplas individualmente para cada aluno e as prioridades com que eles, efetivamente, empregam estratégias de ensino e técnicas de avaliação por meio da diferenciação (FREEDMAN, 2015), porém, os estudos realizados não citam a utilização de um classificador que possa categorizar os alunos por Inteligências Múltiplas, associá-las a aspectos cognitivos, utilizar essa classificação para determinar o perfil cognitivo dos alunos, comparar com as técnicas aplicadas e analisar se a prática condiz com o aspecto cognitivo mais aflorado do aluno. Isso pode ser utilizado para determinar se a técnica de ensino utilizada para um determinado aluno está relacionada ao aspecto cognitivo mais aflorado do mesmo e se o emprego desse aspecto durante o processo de ensino resultou nos melhores resultados obtidos com relação ao aprendizado do discente.

Kunjai descreveu as inúmeras vantagens da aplicação da teoria das Inteligências Múltiplas no ensino a distância. Flexibilidade, operacionalidade remota, relação custo-eficácia, simplicidade, consistência, entre outras, são algumas vantagens citadas. A utilização de ferramentas e de tecnologias inteligentes, associadas à teoria das Inteligências Múltiplas, possibilitou educação fácil e conveniente de forma efetiva sem barreiras de tempo e lugar. (KUNJAL, 2015).

A utilização de um classificador de indivíduos, a partir de seus aspectos cognitivos associados à Teoria das Inteligências Múltiplas, por meio da aplicação de um questionário aos indivíduos antes de iniciarem as aulas, pode contribuir para que as Inteligências Múltiplas utilizadas no processo de ensino a distância sejam aquelas que fazem parte do perfil dos indivíduos.

Adnan, Chowdury, Taz, Ahmed e Rahman realizaram a implementação de um método que utiliza a lógica nebulosa para encontrar um conjunto de artigos relacionados com outros que podem ser recomendados para um leitor. (ADNAN et

al, 2014). Esta implementação motivou a realização de estudos com o intuito de entender como é possível aplicar a lógica nebulosa em classificadores, bem como, entender que existem vários métodos de classificação que utilizam a lógica nebulosa e que podem ser utilizados em processos de classificação.

Ahmad e Adnan encontraram uma forma de medição do avanço das habilidades cognitivas de um ser humano é uma pesquisa desafiadora devido a essas habilidades terem uma grande relação com as emoções, idade e gênero. Existem pesquisas que têm por objetivo investigar o quanto as emoções, idade, gênero e habilidades cognitivas são dependentes entre si (AHMAD; ADNAN, 2015).

Os estudos realizados contribuíram para entender que os aspectos cognitivos podem sofrer alterações durante a vida dos indivíduos e que os fatores idade e gênero também podem influenciar. Desta forma, foi possível entender que o questionário utilizado para fornecer parâmetros para o classificador utilizado nesta dissertação deve levar em conta a idade, o gênero e o meio em que os indivíduos estão inseridos.

## **1.2 OBJETIVOS**

Os objetivos desta dissertação são:

- Fazer uso da lógica nebulosa para o desenvolvimento do método computacional responsável por classificar indivíduos;
- Classificar indivíduos por aspectos cognitivos, sendo que estes devem estar fundamentados à Teoria das Inteligências Múltiplas;
- Os resultados obtidos no processo de classificação devem oportunizar aos indivíduos que, por ventura, desejarem realizar os estudos por meio de ferramentas de apoio ao ensino, a possibilidade de serem direcionados a ambientes que favoreçam os estudos por meio da utilização dos aspectos cognitivos mais aflorados para o indivíduo.



### 1.3 ORGANIZAÇÃO DO TRABALHO

Esta dissertação está dividida em capítulos que descrevem a fundamentação teórica, a metodologia utilizada para a classificação dos indivíduos, testes realizados, resultados obtidos e propostas para trabalhos futuros.

O primeiro capítulo, é referente à introdução, à motivação, aos objetivos e à organização desta dissertação.

No segundo capítulo são apresentados o estado da arte referente à lógica nebulosa, à Teoria das Inteligências Múltiplas e aos sistemas de recomendação. Também são descritos os métodos utilizados pelo classificador para filtrar, classificar e agrupar os dados de entrada para que possa ser gerado o perfil cognitivo dos indivíduos.

No terceiro capítulo, é descrita a metodologia utilizada para o classificador e é apresentada a ferramenta de acadêmica de mineração de dados Weka versão 3.5.7, como também o questionário utilizado para a coleta dos dados que servem como base de dados de entrada para o classificador. São demonstrados os métodos *Simple Fuzzy Grid* utilizado no processo de classificação e o método de agrupamento das respostas em classes que faz uso do *cluster EM*. São realizados testes iniciais no classificador para determinar a acurácia.

No quarto capítulo, são apresentados os resultados obtidos para os primeiros testes, a análise dos resultados, a aplicação do filtro *Resample* utilizado para a melhora da acurácia do classificador e o procedimento de refazer os testes com a aplicação do filtro *Resample*.

No quinto capítulo, são apresentados os resultados obtidos para os testes realizados. As conclusões e os trabalhos futuros são expostos no sexto e sétimo capítulos, respectivamente.

## **2. FUNDAMENTAÇÃO TEÓRICA**

Neste capítulo são apresentados o estado da arte e os principais conceitos sobre a Lógica Nebulosa, a Teoria das Inteligências Múltiplas e os Sistemas de Recomendação. São descritos os métodos utilizados pelo classificador para balancear as classes de dados, classificar os dados de entrada e agrupar os dados de saída para que possa ser gerado o perfil cognitivo dos indivíduos.

### **2.1 LÓGICA NEBULOSA**

As afirmações para as quais não se pode ter certeza necessitam de uma modelagem diferenciada. Pode-se dizer que não choverá hoje, mas pode-se afirmar isso com um grau de certeza de 0,8, tomando como base de que a certeza completa é de 1,0. Esse tipo de situação é um modelo de lógica nebulosa. A lógica nebulosa opera com propostas que podem ser verdadeiras em um grau de certeza que varia de 0 a 1 (WEBER; KLEIN, 2003).

#### **2.1.1 Teoria dos Conjuntos Nebulosos**

O princípio da lógica nebulosa é baseado na teoria dos conjuntos nebulosos, que por sua vez, possibilita realizar o processamento de dados com certo grau de incerteza. Devido a essa característica, ela pode ser diferenciada dos sistemas lógicos tradicionais. Outra característica que diferencia esses dois tipos de lógica é que na lógica nebulosa o raciocínio exato pode corresponder a afirmações incertas que foram criadas linguisticamente para representar escalas de valores. Com essas afirmações podem ser criadas as chamadas variáveis linguísticas que, diferentemente da lógica tradicional, são de extrema relevância para a lógica nebulosa, pois possibilitam captar o grau de incerteza presente nessas variáveis e traduzir para um modelamento matemático. Esta possibilidade da lógica nebulosa é chamada grau de participação ou também conhecido como função de pertinência ao contrário dos sistemas lógicos binários em que o valor atribuído as variáveis, só pode assumir dois valores que são verdadeiro (1) ou falso (0) (PEDRYCZ, 1989).

Com relação à lógica nebulosa e à composição de relações nebulosas, tem-se que essas relações devem expressar, em nível global, a dificuldade de decidir quais elementos pertencem ou não a um dado conjunto de valores. Por exemplo, os valores considerados como sendo verdade podem ser expressos, linguisticamente, por verdade, muito verdade e não verdade, onde cada termo linguístico é interpretado como um subconjunto nebuloso com grau de pertinência que varia entre 0 e 1 indicando o quanto um determinado valor pertence a um dado conjunto. Quando o grau de pertinência tem o valor igual a zero, indica que o elemento não pertence a um determinado conjunto e o grau de pertinência correspondente a 1 indica que o elemento pertence completamente a um dado conjunto (LEE, 1990).

Considerando que o valor correspondente ao grau de pertinência pode variar de 0 a 1 para cada conjunto de valores, podem existir situações em que um elemento pode pertencer a mais de um conjunto de valores, dependendo do grau de pertinência que ele tem em relação a cada conjunto.

Existem outras características da lógica nebulosa que podem ser descritas, por exemplo, nos sistemas lógicos binários os valores das variáveis que representam os predicados são exatos, tais como, par, ímpar, verdadeiro, falso, entre outros, ao passo que na lógica nebulosa os mesmos valores são nebulosos, tais como, alto, baixo, quente, frio, entre outros.

Nos sistemas lógicos clássicos, existe um modificador muito utilizado que é a negação. Na lógica nebulosa, são utilizados vários modificadores de predicados, por exemplo, muito mais, muito menos, sendo sua utilização importante para a geração dos termos linguísticos, muito alto, mais ou menos perto, entre outros, sendo assim é possível supor que a lógica nebulosa é indicada para solucionar problemas reais nos quais a modelagem matemática é impraticável (LEE, 1990).

A modelagem e o controle nebulosos são técnicas para se manusear informações qualitativas de uma maneira rigorosa. Tais técnicas consideram a falta de exatidão das informações e, fazendo isso, tornam-se suficientemente poderosas para manipular de maneira conveniente o conhecimento. Desta forma, sua utilização em sistemas de controle de processos em tempo real, em computadores ou microcontroladores, é das mais convenientes. A teoria de modelagem e controle nebulosa trata do relacionamento entre entradas e saídas, agregando vários parâmetros de processo e de controle. Isso permite a consideração de processos

complexos, de modo que os sistemas de controle resultantes proporcionam um resultado mais preciso, além de um desempenho estável e robusto. A grande simplicidade de implementação de sistemas de controle nebulosos pode reduzir a complexidade de um projeto a um ponto em que problemas anteriormente intratáveis passam a ser solúveis (LEE, 1990).

### 2.1.2 Fundamentos Utilizados na Lógica Nebulosa

Na teoria clássica de conjuntos, um elemento pertence ou não a um dado conjunto. Dado um universo  $U$  e um elemento particular  $x \in U$ , o grau de pertinência  $\mu_A(x)$  com respeito a um conjunto  $A \subseteq U$  é dado por:

$$\mu_A(x) = \begin{cases} 1 & \text{se } x \in A \\ 0 & \text{se } x \notin A \end{cases}$$

A função  $\mu_A(x): U \rightarrow \{0,1\}$  é chamada de função característica na teoria clássica de conjuntos. Frequentemente, uma generalização desta ideia é utilizada, por exemplo, para manipulação de dados com erros limitados. Todos os números dentro de um erro percentual terão um fator de pertinência 1 (um), tendo todos os demais um fator de pertinência 0 (zero). Para o caso preciso, o fator de pertinência é 1 (um) somente no número exato, sendo 0 (zero) para todos os demais (ZADEH, 1965).

Existe uma caracterização mais ampla, na medida em que se sugere que alguns elementos são mais membros de um conjunto do que outros. O fator de pertinência pode então assumir qualquer valor entre 0 (zero) e 1 (um), sendo que o valor 0 (zero) indica uma completa exclusão e o valor 1 (um) representa completa pertinência (ZADEH, 1965).

Formalmente, seja  $U$  uma coleção de elementos denominados genericamente por  $\{u\}$ .  $U$  é chamado de universo. Um conjunto nebuloso  $A$  em um universo  $U$  é definido por uma função de pertinência  $\mu_A$  que assume valores em um intervalo  $[0,1]$ . Considerando as informações acima pode-se deduzir que o conjunto suporte de um conjunto nebuloso  $A$  é o subconjunto dos valores  $u$  de  $U$  tal que  $\mu_A(u) > 0$  e  $\mu_A(u) \leq 1$ , sendo que se  $\mu_A(u) = 0$ , indica que este elemento não pertence ao conjunto nebuloso  $A$ , pois, na lógica nebulosa o valor zero na função de pertinência indica completa exclusão do elemento em relação ao conjunto em questão.

### 2.1.3 Definições e Operações

Para facilitar o estudo e o entendimento das definições e operações da lógica nebulosa, são considerados dois conjuntos nebulosos  $A$  e  $B$  no universo  $U$ , assim tem-se que:

Conjunto vazio:  $A = \emptyset$  se e somente se  $\forall u \in U, \mu_A(u) = 0$

Complemento  $A'$ :  $\mu_{A'}(u) = 1 - \mu_A(u)$

Conjuntos iguais:  $A = B$  se e somente se  $\forall u \in U, \mu_A(u) = \mu_B(u)$

A subconjunto de  $B$ :  $A \subset B$  se  $\forall u \in U, \mu_A(u) < \mu_B(u)$

Nas operações com conjuntos nebulosos, utilizam-se os conceitos de norma triangular (norma-t) e co-norma triangular (norma-s). Uma norma triangular é uma função:  $J: [0,1] \times [0,1] \rightarrow [0,1]$  tal que,  $\forall x, y, z, w \in [0,1]$ :

$$(i) x J w \leq y J z, \text{ se } x \leq y, w \leq z$$

$$(ii) x J y = y J x$$

$$(iii) (x J y) J z = x J (y J z)$$

$$(iv) x J 0 = 0; x J 1 = x$$

Uma co-norma triangular é tal que  $s: [0,1] \times [0,1] \rightarrow [0,1]$ , satisfazendo as propriedades (i) e (iii) acima e ainda:

$$(iv) x s 0 = x; x s 1 = 1$$

Exemplos de normas-t incluem o mínimo ( $\wedge$ ) e o produto algébrico ( $\cdot$ ) e como exemplo de normas-s podem ser citados o máximo ( $\vee$ ) e a soma limitada ( $\oplus$ ).

Para representar a união e intersecção de conjuntos nebulosos, utilizam-se respectivamente normas triangulares (normas-t) e co-normas triangulares (normas-s). Assim, utilizando-se o operador mínimo como norma-t e máximo como norma-s, partindo desse pressuposto pode-se descrever a união e a intersecção conforme abaixo (BRIN; PAGE, 1998):

$$\text{Uni\~ao de } A \text{ e } B: \mu_{A \cup B}(x) = \mu_A(u) \vee \mu_B(u)$$

$$\text{Intersec\~ao de } A \text{ e } B: \mu_{A \cap B}(x) = \mu_A(u) \wedge \mu_B(u)$$

### 2.1.4 Propriedades Algébricas

Utilizando as definições de união e intersecção implementadas pelos operadores, máximo, mínimo e complemento, é possível verificar que as seguintes propriedades algébricas de conjuntos ordinários também valem para conjuntos nebulosos (LEE, 1990):

Involução:  $(A')' = A$

Idempotência:  $A \cap A = A$  e  $A \cup A = A$

Comutatividade:  $A \cap B = B \cap A$  e  $A \cup B = B \cup A$

Associatividade:  $(A \cap B) \cap C = A \cap (B \cap C)$  e  $(A \cup B) \cup C = A \cup (B \cup C)$

Distributividade:  $A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$  e

$$A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$$

Lei Transitiva: se  $A \subset B$  e  $B \subset C$  então  $A \subset C$

Observando que as funções de pertinência dos conjuntos, vazio e universo são definidas como sendo 0 (zero) e 1 (um), respectivamente,  $\forall u \in U$ , as seguintes propriedades também são válidas:

$$A \cap \emptyset = \emptyset \text{ e } A \cap U = A$$

$$A \cup \emptyset = A \text{ e } A \cup U = U$$

Existem ainda as propriedades de conjuntos clássicos que não se verificam para conjuntos nebulosos, estas estão descritas abaixo:

$$A \cap A' \neq \emptyset \text{ e } A \cup A' \neq U$$

Por exemplo, seja  $A$  um conjunto nebuloso e um elemento  $a \in U$  tal que  $\mu_A(a) = 0.3$ , temos então que  $\mu_{A \cap \bar{A}}(a) = 0.3$  e  $\mu_{A \cup \bar{A}}(a) = 0.7$ , assim,  $A \cap A' \neq \emptyset$  e  $A \cup A' \neq U$ .

### 2.1.5 Variáveis Linguísticas

Conforme citado anteriormente existem algumas questões e problemas em que não é possível ou aplicável respostas determinísticas, assim, é aberto um espaço onde raciocínios aproximados podem levar a um entendimento maior para alguns problemas. Esse modelo de aproximação não é algo raro, esse tipo de problema ocorre corriqueiramente no dia a dia das pessoas, sendo esse o meio termo cabível ao um determinado fato.

A lógica nebulosa traz consigo a ideia de medida, ou aplicação de uma métrica na qual, aparentemente, a questão parece ser vaga. Essas métricas, quando aplicadas, aproximam a questão de algumas suposições possíveis (COX, 1994).

As variáveis linguísticas são consideradas o vocabulário da linguagem nebulosa, pois são expressões presentes na linguagem humana e que, quando traduzidas ou interpretadas podem levar a conclusões importantes. Considerando os aspectos cognitivos como exemplo, a afirmação é algo muito comum (VON ALTROCK, 1996):

***“Durante a realização de trabalhos escolares em grupo, Antônio gosta de ouvir música, João não gosta muito, mas o Carlinhos não consegue realizar os trabalhos ouvindo música e Antônio é o líder e não pode sair do grupo”***

Ao analisar essa frase por meio da lógica nebulosa é possível tomar algumas ações. Colocando-se no lugar do professor da disciplina em questão, ao ter essa frase como matéria prima para uma tomada de decisão e interferir na forma de organização dos grupos de estudo, ele pode levantar a questão:

***“Tenho que retirar um aluno do grupo e trocar por outro aluno, qual dos alunos devo retirar?”***

Sem muitos cálculos ou dados exatos que possam fornecer informações relevantes ao perfil dos alunos, o pensamento humano do professor pode conduzir a reflexão de forma semelhante a esta:

***“O Antônio gosta de ouvir música durante os estudos, João não gosta muito, porém, Carlinhos não gosta de ouvir música durante os estudos e Antônio não pode sair do grupo. O nível de aceitação de Carlinhos para ouvir música em relação a João é bem menor, então vou trocar o Carlinhos de grupo e colocar no lugar de Carlinhos um aluno que tenha uma aceitação maior para estudar ouvindo música que a aceitação de Carlinhos.”***

Desta forma pode-se notar que no exemplo acima não existem regras bem definidas sobre qual aluno deve sair do grupo, porém, por meio de abstrações e analogias é possível concluir algo. Os termos, gosta, não gosta muito e não gosta, são as chamadas variáveis linguísticas, pois possibilitam enumerar matematicamente qual a proximidade desses alunos em gostar de estudar ouvindo música ou não.

Considerando a incerteza presente nestes casos é extremamente válido lembrar que as incertezas somente podem ser levadas em consideração se for possível diante de uma determinada situação efetuar aproximações e cálculos que levem a alguma conclusão válida.

Ao analisar este contexto, considerando a lógica nebulosa, existe algo chamado função de pertinência que vem a ser um mapeamento matemático de cada valor numérico possível para as variáveis linguísticas (VON ALTROCK, 1996).

Continuando a análise do problema de trocar um aluno do grupo de estudos, nota-se a importância em aproximar a léxica do modelo matemático para que assim seja possível gerar conclusões válidas sobre o problema, ou seja, analisar as informações de entrada, que neste caso são as informações sobre a preferência dos alunos em ouvir música durante os estudos e produzir uma sequência de informações que podem ser entendidas e manipuladas pelo professor para que ele possa decidir qual é a melhor opção para trocar um aluno de grupo durante os estudos.

Outro exemplo de aplicação da lógica nebulosa é utilizá-la para análise da temperatura térmica em um determinado ambiente. Detalhando um pouco mais o exemplo da temperatura tem-se que, para exemplificar as funções de pertinência, deve-se considerar a Figura 2.1, onde por meio de alguns gráficos são



representadas três variáveis térmicas: baixa, média e alta. Estas variáveis são relativas a análise sobre a temperatura.

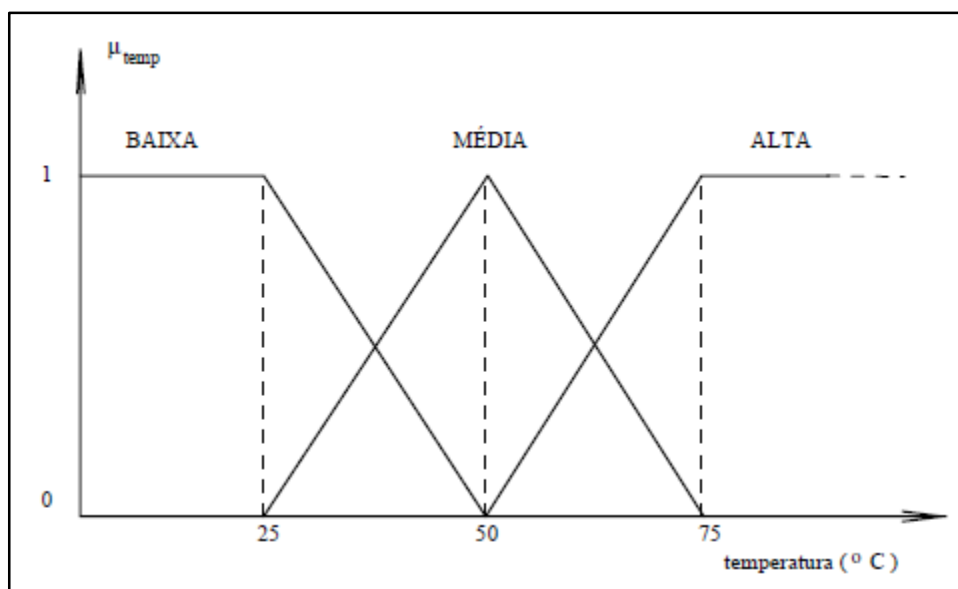


Figura 2.1: Exemplo de função de pertinência.

É comum este tipo de representação ser apresentada em um único gráfico. Para a Figura 2.1 tem-se que, na escala de valores do eixo vertical existem os valores limite 0 e 1 que são os valores mínimo e máximo, respectivamente, para o fator de pertinência da temperatura.

A escala de valores do eixo horizontal corresponde aos valores da temperatura em °C (graus celsius). Conforme a temperatura aumenta dentro do conjunto de valores que representam a baixa temperatura, o fator de pertinência diminui até o momento em que o fator de pertinência atinge o limiar entre dois conjuntos de valores e passa a pertencer ao conjunto de valores referente à temperatura baixa e média. Com o aumento da temperatura, ele atinge o fator de pertinência máximo (fator de pertinência igual a 1) para este conjunto de valores ao atingir 50°C. Ao atingir esta temperatura, ele deixa de pertencer ao conjunto de valores correspondentes à baixa temperatura pertencendo somente ao conjunto de valores para temperatura média.

Com o aumento da temperatura, o fator de pertinência para o conjunto de valores de temperatura média começa a diminuir até que ele atinja o limiar entre os conjuntos de valores entre a temperatura média e alta, passando a pertencer aos dois conjuntos de valores. Com a continuidade no aumento da temperatura, o fator de pertinência para o conjunto de valores para temperatura média diminui, ao tempo

que o fator de pertinência para a temperatura alta sobe até que ele atinja o valor máximo (fator de pertinência igual a 1) ao atingir a temperatura de 75°C, deixando de pertencer ao conjunto de valores para a temperatura média, passando a pertencer somente ao conjunto de valores correspondentes à temperatura alta. Esse processo ocorre de forma inversa para a diminuição da temperatura, deixando de ser alta, passando gradativamente para média até atingir a baixa temperatura.

Ainda considerando o problema da temperatura, pode-se realizar outra abordagem considerando que os valores da variável linguística podem ser sentenças em uma linguagem especificada. Para ilustrar, os valores da variável linguística temperatura podem ser expressos vários valores, tais como, “alta”, “não alta”, “muito alta”, “bastante alta”, “não muito alta”, “alta, mas não muito alta”. Neste caso, os valores da variável linguística são sentenças formadas a partir do nome “alta”, da negação “não”, dos conectivos “e” e “mas”, e dos modificadores “muito” e “bastante”. Para gerar estes valores compostos, utilizam-se diversos termos, que podem ser divididos nas seguintes categorias (VON ALTROCK, 1996):

(a) **Termos Primários:** nomes de conjuntos nebulosos especificados em um determinado universo, por exemplo, alto, baixo, pequeno, médio, grande, zero.

(b) **Conectivos Lógicos:** A negação NÃO, conectivos E e OU, e conectivos mascarados, tais como, mas, porém, entre outros, permitem a inclusão, exclusão ou definir se um determinado valor está em mais de um conjunto.

(c) **Modificadores:** muito, pouco, levemente, extremamente são utilizados para modificar os valores, alterando o grau de pertinência do valor em relação aos conjuntos nebulosos.

(d) **Delimitadores:** parênteses e aspas são utilizados para separar as sentenças.

### 2.1.6 Sistema lógico nebuloso.

Considerando todas as informações apresentadas sobre lógica nebulosa é possível definir um sistema lógico nebuloso e dividi-lo em três operações básicas, são elas: fuzzificação, inferência e defuzzificação. As três operações básicas podem ser visualizadas por meio da Figura 2.2.

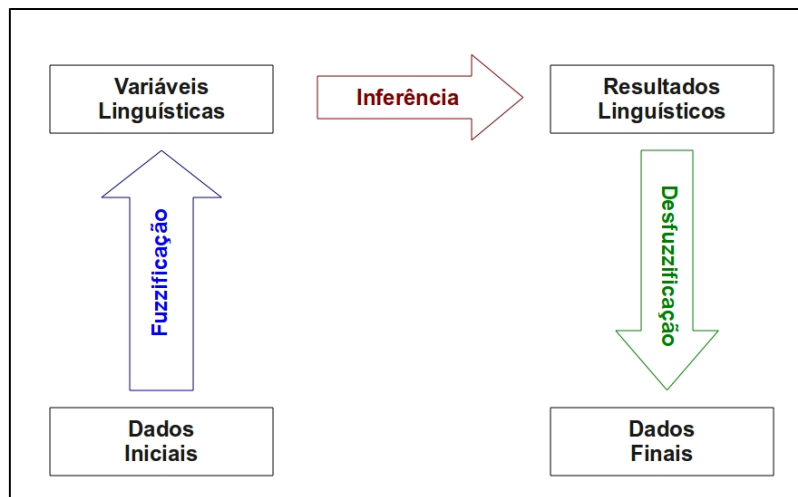


Figura 2.2: Sistema lógico nebuloso.

**Fuzzificação** – Nesta primeira etapa do sistema lógico nebuloso, o problema é analisado e os dados de entrada são transformados em variáveis linguísticas. Neste momento é de extrema importância que todos os dados de imprecisão e incerteza sejam considerados e transformados em variáveis linguísticas. Após esta transformação são determinadas as funções de pertinência (COX, 1994).

Para exemplificar esta etapa e levando em consideração o exemplo citado na Figura 2.1 referente à análise da temperatura, pode-se ter como ponto de partida as seguintes afirmações: a temperatura está alta, muito alta, baixa, muito baixa, razoavelmente alta, razoavelmente baixa.

Considerando as variáveis linguísticas “muito alta”, “muito baixa” e “razoavelmente alta”, entre outras, pode-se determinar que a variável linguística está sendo representada com palavras, mas esta atribuição de graus poderia também ser feita por meio de valores numéricos.

**Inferência** – Considerando que, na etapa anterior, os dados de entrada foram transformados em variáveis linguísticas, nesta segunda etapa é o momento em que serão criadas as regras ou proposições por meio da associação das variáveis já criadas. As proposições são geradas do relacionamento entre as variáveis do modelo e a região nebulosa. Essas regras resultantes das associações podem ser condicionais ou não condicionais. Esta fase do sistema lógico nebuloso pode ser dividida em dois componentes: agregação e composição. Por meio da Figura 2.3 é possível visualizar a inferência nebulosa por meio de agregação e composição. Enquanto a agregação define a validade de uma regra, a composição define o resultado obtido por meio de uma inferência (COX, 1994).

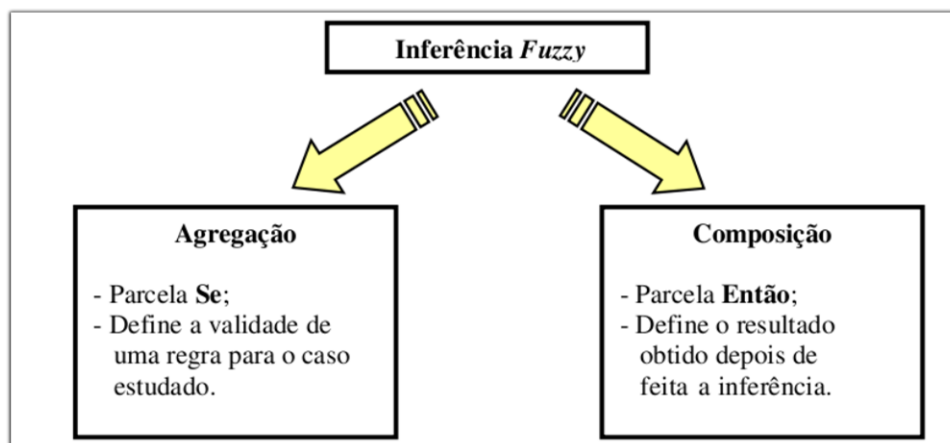


Figura 2.3: Inferência nebulosa.

**Defuzzificação** – A defuzzificação é a etapa em que os valores nebulosos são convertidos em números reais, tendo assim um conjunto de saída matematicamente definido. Para encontrar o valor atual e real correspondente é necessário que se encontre um valor que melhor represente a informação constante no conjunto. Este é o processo chamado de defuzzificação, conforme apresentado na Figura 2.4 (COX, 1994).

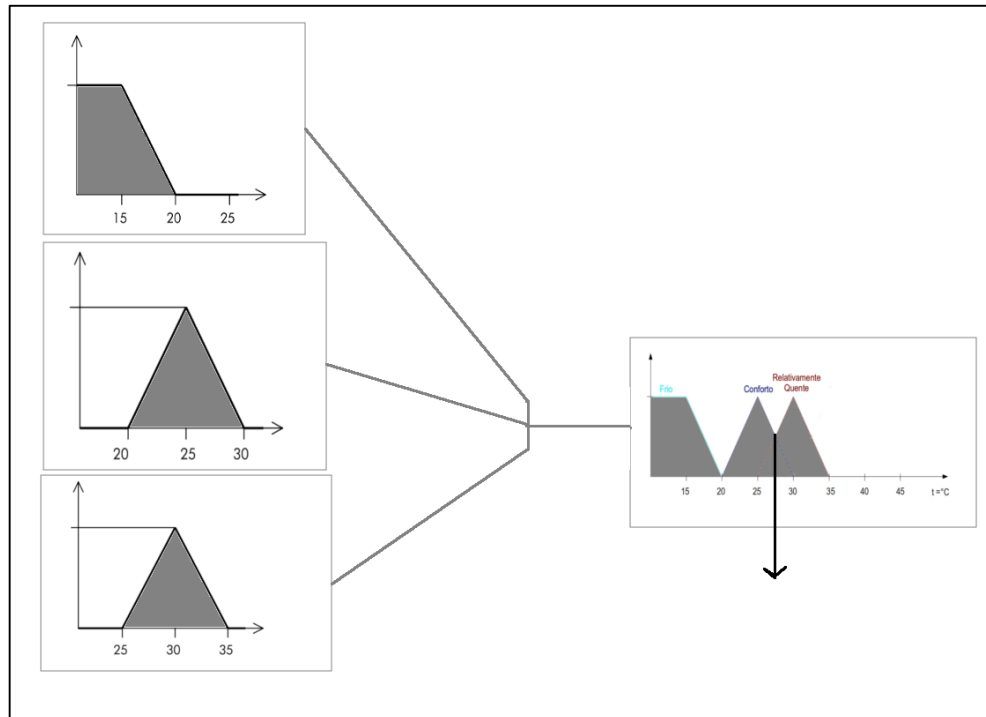


Figura 2.4. Processo de Defuzzificação.

Existem diferentes técnicas de defuzzificação presentes na literatura que fazem uso de diferentes métodos. Alguns dos métodos encontrados são citados por Earl Cox, por exemplo, Centroid é o método no qual a saída precisa a ser considerada é o centro de gravidade do conjunto nebuloso. *Maximum height* é o método em que a saída precisa se obtém tomando a média entre os dois elementos extremos no universo de discurso que correspondem aos maiores valores da função de pertinência do conjunto nebuloso de saída (COX, 1994).

### 2.1.7 Aplicações da lógica nebulosa

Existem diversas áreas em que a lógica nebulosa é aplicada devido sua característica de buscar soluções para problemas reais utilizando um raciocínio próximo do humano. Alguns exemplos de aplicações são (LUGER, 2005):

- Sistemas especialistas e Sistemas multiagentes;
- Reconhecimento de padrões e *Data mining*;
- Robótica e Algoritmos genéticos;
- Sistemas de controle inteligentes;
- Sistemas de apoio à tomada de decisão;

## 2.2 TEORIA DAS INTELIGÊNCIAS MÚLTIPLAS

Este conceito foi desenvolvido por Howard Gardner, psicólogo da Universidade de Harvard na década de 70. O primeiro livro escrito por Gardner intitulado “Estruturas da Mente”, publicado em 1983, gerou grande impacto na área da educação devido a Gardner afirmar que todos os indivíduos dispõem de sete inteligências, sendo elas, linguística, lógico-matemática, espacial, corporal-cinestésica, musical, interpessoal e intrapessoal, e que cada indivíduo tem maior habilidade para desenvolver uma determinada inteligência.

Esta teoria possibilita trabalhar com estímulos específicos e com isso desenvolver, de forma mais adequada, determinada habilidade de cada indivíduo (GARDNER, 2011).

É importante considerar as dificuldades em inserir essa teoria ao ambiente escolar, pois além da falta de esclarecimento por parte dos professores sobre o tema de maneira geral, outros fatores podem dificultar a aplicação dessa metodologia. São eles: a falta de recursos, a pouca disponibilidade de tempo e a não inserção dessa teoria nos Parâmetros Curriculares Nacionais (GARDNER, 2011).

Estudos mais recentes apontam a existência da oitava e nona inteligências denominadas de naturalista e existencial (GARDNER, 2011). Essas por sua vez não fazem parte do universo pertencente aos testes que serão realizados, sendo assim não farão parte desta dissertação.

### 2.2.1 Tipos de Inteligências Múltiplas

Segundo Armstrong (2001) as inteligências múltiplas foram mapeadas, inicialmente, como sendo sete. Gardner ofereceu um meio de mapear a ampla gama de capacidades dos seres humanos nessas sete categorias ou “inteligências”, as quais seriam (ARMSTRONG, 2001):

- **Inteligência Linguística** – Consiste na capacidade de usar as palavras de forma efetiva quer de forma oral, como por exemplo, um orador ou político, ou escrita como um poeta ou jornalista. Esta inteligência inclui ainda a capacidade de manipular a linguagem, ou seus significados, e as dimensões práticas da mesma.

- **Inteligência Lógico-Matemática** – Consiste na capacidade de usar números de forma efetiva como um matemático, assim como a capacidade de raciocínio como um cientista. Esta inteligência inclui sensibilidade a padrões, afirmações, proposições, funções e outras abstrações relacionadas.

- **Inteligência Espacial** – Consiste na capacidade de perceber com precisão o mundo espacial visual e de realizar transformações sobre essas percepções como um decorador de interiores. Esta inteligência envolve sensibilidade à cor, linha, forma, configuração, espaço e às relações existentes entre esses elementos.

- **Inteligência Corporal-Cinestésica** – Consiste na habilidade no uso do corpo todo para expressar ideias e sentimentos, além de facilidade no uso das mãos para manipular, produzir ou transformar coisas. Esta inteligência inclui habilidades físicas específicas, tais como, coordenação, equilíbrio, destreza, força, flexibilidade e velocidade. Tendo como exemplos atletas, malabaristas e médicos cirurgiões.

- **Inteligência Musical** – Consiste na facilidade em perceber, discriminar, transformar e expressar formas musicais. Esta inteligência inclui sensibilidade ao ritmo, tom ou melodia e timbre.

- **Inteligência Interpessoal** – Consiste na capacidade de perceber e fazer distinções no humor, intenções, motivações e sentimentos das outras pessoas. Isso pode incluir sensibilidade a expressões faciais, voz e gestos; a capacidade de discriminar muitos tipos diferentes de sinais interpessoais; e a capacidade de responder efetivamente a esses sinais.

- **Inteligência Intrapessoal** – Consiste na capacidade de autoconhecimento, ou seja, habilidade de agir adaptativamente com base neste conhecimento. Esta inteligência inclui possuir uma imagem precisa de si mesmo, tendo consciência dos estados de humor, intenções, motivações,

temperamento e desejos; e a capacidade de autodisciplina, autoentendimento e autoestima.

Ao observar as sete inteligências citadas, especialmente algumas em específico, como a musical, a espacial e a corporal-cinestésica surge a dúvida de como Gardner conseguiu classificá-las como inteligências, em vez de talentos ou aptidões. Gardner percebeu que as pessoas estão acostumadas a ouvir expressões como: “Ele não é muito inteligente, mas tem uma aptidão maravilhosa para a música”; assim ele usou a palavra inteligência de forma muito consciente e intencional para descrever cada categoria (ARMSTRONG, 2001).

Devido à necessidade de fornecer fundamentos teóricos sólidos para suas afirmações, como forma de comprová-las, Gardner estabeleceu certos “testes”, pelos quais cada inteligência teria de ser provada para ser considerada uma inteligência habilitada e não simplesmente um talento ou aptidão. Os critérios que ele utilizou incluem oito fatores que precisam ser estudados de forma aprofundada para que possam ser utilizados como critério de comprovação de qual é o tipo de inteligência que um indivíduo possui, são eles (GARDNER, 1995):

- Isolamento potencial por lesão cerebral;
- A existência de *Savants*<sup>1</sup>, prodígios e outros indivíduos excepcionais;
- Uma história desenvolvimental distintiva e um conjunto definível de desempenhos peritos;
- Uma história evolutiva e uma plausibilidade evolutiva;
- Apoio de achados psicométricos;
- Apoio de tarefas psicológicas experimentais;
- Uma operação ou um conjunto de operações centrais identificáveis;
- Susceptibilidade à codificação em um sistema simbólico.

---

<sup>1</sup>*Savants* definição: São indivíduos que apresentam uma ou ocasionalmente mais de uma inteligência mais desenvolvida do que as demais.



**Isolamento potencial por lesão cerebral** – Nesses casos, um indivíduo que tenha sofrido algum tipo de lesão ou trauma em uma determinada região do cérebro apresentará algumas funções específicas prejudicadas, como por exemplo, se a região afetada for a da boca, a fala ficará comprometida, enquanto outras funções permanecerão intactas, como a capacidade de realizar cálculos e dançar.

**A existência de Savants, prodígios e outros indivíduos excepcionais** – Existem indivíduos que possuem apenas uma das inteligências em níveis extremamente desenvolvidos, enquanto as demais inteligências apresentam um baixo grau de desenvolvimento e utilização. Esse perfil de inteligências inclui os chamados prodígios e outros indivíduos excepcionais, incluindo os autistas. Já nos casos em que o indivíduo apresenta uma ou, ocasionalmente, mais de uma inteligência mais desenvolvida das demais habilidades, os mesmos são denominados *Savants*.

**Uma história desenvolvimental distintiva e um conjunto definível de desempenhos peritos** – As inteligências são desenvolvidas seguindo um padrão influenciado por atividades culturalmente valorizadas. Cada atividade baseada numa inteligência tem sua própria trajetória desenvolvimental, isto é, cada atividade tem seu momento de surgir na infância, seu momento de pico durante a vida e seu próprio padrão de declínio rápido ou gradual conforme a pessoa envelhece.

**Uma história evolutiva e uma plausibilidade evolutiva** – Todas as sete inteligências passam no teste de ter suas raízes profundamente inseridas na evolução dos seres humanos e mesmo, anteriormente, na evolução de outras espécies. Por exemplo, a inteligência espacial é importante para o ser humano desde os primórdios da humanidade, prova disso são os desenhos feitos pelo homem pré-histórico, encontrados em diversas cavernas espalhadas pelo mundo e na maneira pela qual, certos insetos se orientam no espaço quando procuram flores. Igualmente, a inteligência musical por meio de grande variedade de cantos de pássaros.

**Apoio de achados psicométricos** – As medidas padronizadas de capacidade humana oferecem o “teste” que a maioria das teorias da inteligência utiliza para determinar a validade de um modelo. Embora Gardner não seja nenhum adepto dos testes padronizados, e de fato tenha sido um grande defensor de alternativas de testes formais, ele sugere que muitos testes padronizados existentes apoiam a teoria das inteligências múltiplas.

**Apoio de tarefas psicológicas experimentais** – Ao examinar estudos psicológicos específicos, pode-se perceber as inteligências operando isoladas umas das outras. Da mesma forma, em estudos de capacidades cognitivas com memória, percepção ou atenção, é possível observar evidências de que os indivíduos possuem capacidades seletivas. Certos indivíduos, por exemplo, podem ter uma memória superior para palavras, mas não para rostos; outros podem ter aguda percepção de sons musicais, mas não de sons verbais. Cada uma dessas faculdades cognitivas é específica de uma inteligência, isto é, as pessoas podem demonstrar diferentes níveis de proficiência nas inteligências em cada área cognitiva.

**Uma operação ou um conjunto de operações centrais identificáveis** – Cada inteligência possui um conjunto de operações centrais que servem para acionar as várias atividades inerentes à mesma. Na inteligência musical, por exemplo, esses componentes podem incluir a sensibilidade ao tom ou a capacidade de discriminar uma entre várias estruturas rítmicas.

**Susceptibilidade à codificação em um sistema simbólico** – Grande parte da representação e da comunicação humana de conhecimento ocorre por meio de sistemas simbólicos. Existem linguagens simbólicas que são de grande importância no mundo inteiro.

## 2.3 SISTEMAS DE RECOMENDAÇÃO

O aumento da quantidade e da disponibilidade das informações, principalmente por meio da Internet, faz com que as pessoas tenham um grande número de opções de escolha e, em muitos casos, elas têm dificuldade em realizar essas escolhas devido à falta de informação ou de experiência pessoal para tomar decisões. Diante dessa situação, as pessoas acabam tomando decisões influenciadas por recomendações de outras pessoas ou, no caso de um *software*, do próprio *software*. Quando se está em dúvida, na grande maioria das vezes, segue-se uma recomendação que pode ser direta, por meio de textos de recomendação, opiniões de revisores de filmes e livros, impressos de jornais, entre outros (RESNICK; VARIAN, 1997).

Um dos objetivos dos sistemas de recomendação é auxiliar no aumento da capacidade e da eficácia dos processos de indicação. Em um sistema típico, as pessoas fornecem recomendações como entradas, sendo essas agregadas pelo sistema, que as direcionam para os indivíduos considerados potenciais interessados nesse tipo de recomendação. Um dos maiores desafios de tais sistemas é realizar a combinação adequada entre as expectativas dos usuários e os produtos, serviços, ou pessoas a serem recomendados, sendo o maior problema descobrir esta relação de interesses (RESNICK; VARIAN, 1997).

### 2.3.1 Classificação dos Sistemas de Recomendação

Os sistemas de recomendação podem ser classificados pelo uso de diferentes tecnologias e por aplicarem abordagens distintas, mas possuem como finalidade única a recomendação. Desta forma, pode-se classificá-los em quatro grandes grupos, são eles (MONTANER; LÓPEZ; DE LA ROSA, 2003):

**Recomendação com Base em Conteúdo** – Por muitos anos, os cientistas têm direcionado seus esforços para aliviar o problema ocasionado pela sobrecarga de informações por meio de projetos que integram tecnologias, que, automaticamente, reconhecem e categorizam as informações (REATEGUI; CAZELLA, 2005).

Uma das alternativas é recorrer a utilização de sistemas de recomendação com base em conteúdo, onde o sistema aprende a recomendar itens que são semelhantes àqueles que o usuário gostava no passado. A semelhança dos itens é calculada com base nas características associadas com os itens comparados. Por exemplo, se um usuário tiver classificado positivamente um filme que pertence ao gênero comédia, então o sistema pode aprender a recomendar outros filmes a partir desse gênero.

Alguns sistemas têm como objetivo gerar, de forma automática, descrições dos conteúdos dos itens e comparar essas descrições com os interesses dos usuários, visando verificar se o item é ou não relevante para cada usuário (REATEGUI; CAZELLA, 2005).

Essa técnica é chamada de Filtragem Baseada em Conteúdo por realizar uma seleção baseada na análise de conteúdo dos itens e no perfil do usuário. Muitas ferramentas que utilizam essa abordagem aplicam técnicas como indexação de frequência de termos. Nessa técnica, informações dos documentos e necessidades dos usuários são descritas por vetores com uma dimensão para cada palavra que ocorre na base de dados (REATEGUI; CAZELLA, 2005).

Cada componente do vetor é a frequência que uma respectiva palavra ocorre em um documento ou na consulta do usuário. Claramente, os vetores dos documentos que estão próximos aos vetores da consulta do usuário são considerados os mais relevantes para ele. Outras técnicas aplicadas na Filtragem Baseada em Conteúdo são índices de busca booleana, onde a consulta constitui-se em um conjunto de palavras-chave unido por operadores booleanos e sistemas de filtragem probabilística, onde se aplica o raciocínio probabilístico para se determinar a probabilidade que um documento possui a fim de atender às necessidades de informações (REATEGUI; CAZELLA, 2005).

**Recomendação de Filtragem Colaborativa** – A abordagem da Filtragem Colaborativa é também conhecida por Filtragem Social, sendo esta desenvolvida para atender pontos que estavam em aberto na Filtragem Baseada em Conteúdo (BRESSLER, 2004). A Filtragem Colaborativa se diferencia da Filtragem Baseada em Conteúdo exatamente por não exigir a compreensão ou o reconhecimento do conteúdo dos itens (REATEGUI; CAZELLA, 2005).

A Filtragem Colaborativa permite dois tipos de abordagens. A primeira determina recomendações baseadas em conteúdo consumido por usuários com o mesmo padrão de consumo do usuário atual. Essa abordagem é utilizada principalmente em sistemas de comércio eletrônico, como Amazon e Submarino (KAJIMOTO et al, 2007). A segunda determina recomendações baseadas nas classificações realizadas por outros usuários dentro de um grupo restrito de conteúdo, ordenadas pela soma da relevância de tais classificações, utilizando para isso um sistema de votação com pesos. Essa abordagem é utilizada principalmente em sistemas de notícias, tais como, jornais e revistas *online* disponibilizados na Internet, que de modo geral, oferecem recomendações menos pessoais e mais dirigidas a um grupo determinado de usuários restritos a um tema. Em contrapartida, esse tipo de implementação enfrenta menos problemas de escalabilidade e em geral é mais viável (KAJIMOTO et al, 2007).

**Recomendação Baseada em Conhecimento** – Este método tenta entregar sugestões ao utilizador com base no conhecimento acerca da necessidade de um utilizador para com um determinado item. Por esse motivo, alguns autores chamaram-lhe “Recomendação Utilitária”. As recomendações são efetuadas por meio do estabelecimento de medidas de utilidade derivadas a partir do conhecimento que se possui das relações de um item para com um determinado utilizador. A recomendação baseada em conhecimento pressupõe uma estrutura que guarde essas relações e que permita a sua consulta de modo a determinar a utilidade para o utilizador inferindo novas recomendações. O domínio de conhecimento considerado está associado não só às preferências do utilizador, mas também ao tipo de item a recomendar. Por exemplo, um sistema que recomenda viagens baseado em conhecimento pode tirar partido não só do que se conhece acerca da experiência do utilizador em viagens anteriores, mas também do que se sabe sobre as características dos locais que visitou e dos locais disponíveis para recomendar (GUTTMAN; MOUKAS; MAES, 1998).

**Recomendação Híbrida** – Existe ainda a Recomendação Híbrida, a qual procura combinar os pontos fortes da Filtragem Colaborativa e da Filtragem Baseada em Conteúdo. Por meio da utilização dessa combinação, pode-se criar um sistema que possa melhor atender às necessidades do usuário. Essa abordagem une o melhor das duas técnicas e elimina as fraquezas de cada uma. Os pontos fortes da Filtragem Colaborativa são a descoberta de novos relacionamentos entre usuários e a recomendação de itens diretamente relacionados ao histórico. Os pontos fortes da Filtragem Baseada em Conteúdo são os bons resultados para usuários incomuns e a precisão independente do número de usuários (REATEGUI; CAZELLA, 2005).

### 2.3.2 Tipos de Descoberta de Conhecimento

Existem diversos tipos de descoberta de conhecimento aplicáveis nos sistemas de recomendação, que podem ser mais eficientes do que outras, dependendo do domínio de aplicação. Nesta dissertação são apresentadas três delas (REATEGUI; CAZELLA, 2005):

**Regras de Associação** – As regras de Associação são técnicas que procuram encontrar padrões entre conjuntos de valores ou em gráficos evolutivos de valores. Uma regra de associação é uma indicação de que existe um certo relacionamento entre fatos diferentes. Essa regra constitui-se em um relacionamento  $X \rightarrow Y$ , onde X e Y são conjuntos de itens e a intersecção entre os mesmos constitui-se em um conjunto vazio. Cada regra de associação é associada a um fator de suporte, definido aqui como “FSup”, e a um fator de confiança definido por “FConf” (REATEGUI; CAZELLA, 2005).

O “FSup” é determinado pela razão do número de tuplas que satisfazem X e Y, sobre o número total de tuplas, conforme equação (1) e o “FConf” é determinado pela razão do número de tuplas que satisfazem X e Y, sobre o número de tuplas que satisfazem X conforme equação (2). A função de extração de regras a partir de um banco de dados consiste em encontrar todas as regras possíveis com “FSup” e “FConf” maiores ou iguais a um “FSup” e “FConf” especificados pelo usuário (REATEGUI; CAZELLA, 2005).

$$FSup = \frac{\text{Frequência de } X \text{ e } Y}{\text{Total de amostras}} \quad (1)$$

$$FConf = \frac{\text{Frequência de } X \text{ e } Y}{\text{Frequência de } X} \quad (2)$$

**Classificação** – A utilização do método de classificação para a descoberta de conhecimento consiste em determinar a classe ou categoria de um elemento a partir das características, que podem ser os atributos deste elemento e com base em classes pré-definidas. Uma nova classe pode ser acrescentada ao conjunto, caso seja detectado um elemento com características contraditórias que satisfazem duas classes, mas não podendo ser classificado assim por serem classes disjuntas, ou então quando as características do elemento não se enquadram em nenhuma das classes pré-estabelecidas (REATEGUI; CAZELLA, 2005).

Um exemplo que permite explicar o uso de classificação consiste em supor que um sistema de recomendação guarde algumas informações sobre os seus usuários, por exemplo, o status na Universidade, país e área de interesse. Essas informações podem ser interessantes para prever que tipo de usuário está mais interessado em qual área de pesquisa. Desta forma, o sistema pode então concentrar esforços de recomendação para classificar os usuários por área de pesquisa (REATEGUI; CAZELLA, 2005).

**Agrupamento** – Este método consiste em associar um item a uma ou várias classes categóricas (ou *clusters*), determinando as classes pelos dados, independentemente da classificação pré-definida. Os grupos são definidos por meio do agrupamento de dados baseados em medidas de similaridade ou modelos probabilísticos, visando detectar a existência de diferentes grupos dentro de um determinado conjunto de dados e, em caso de sua existência, determinar quais são eles (SILVA, 2007).

Outra definição de agrupamento, que pode ser realizada de forma simplificada, é o processo de partição de uma população heterogênea em vários subgrupos ou grupos mais homogêneos, sendo esse o processo inverso da classificação. A partir de técnicas de agrupamento, são analisados os atributos de caracterização e os de discriminação para enquadrar os elementos nas suas respectivas classes, que são definidas por essas técnicas. O algoritmo de

agrupamento automaticamente deve criar grupos por meio de partições de um banco de dados em conjunto de tuplas similares. A partição acontece de forma que as tuplas de valores e atributos semelhantes sejam reunidas nos mesmos grupos. Um agrupamento de qualidade surge onde a similaridade intraclasse é alta e a interclasse é baixa (SILVA, 2007).

O agrupamento é um tipo de aprendizado não supervisionado, uma vez que fica a cargo do algoritmo a definição de quais atributos e seus valores serão utilizados para montar agrupamentos de dados similares (SILVA, 2007).

## **2.4 Balanceamento das classes de dados**

Para alguns classificadores, o desbalanceamento de classes é indesejável, pois as classes minoritárias também possuem informações importantes e têm a mesma relevância que as demais. Sendo assim, pode ser utilizado um filtro que utiliza o método de introdução de custos de classificação incorreta ou método de classificação sensível ao custo para realizar o balanceamento das classes (SHENG et al, 2006).

Os custos para o classificador desta dissertação devem ser calculados na fase de construção do classificador, sendo assim, a aprendizagem sensível ao custo é considerada a opção mais ajustada e com melhores resultados. O propósito da aplicação de uma técnica de balanceamento de classes é equalizar os custos de cada classe, isso faz com que a disparidade de custos não seja levada em consideração ou seja minimizada.

Uma possível e comum abordagem na aprendizagem sensível ao custo é a de alterar a distribuição das classes, para que o classificador minimize, num determinado sentido (com aumento da sensibilidade ou da especificidade), os custos com novas instâncias. Procura-se com isto que as classes tenham uma distribuição proporcional à sua importância, ou seja, ao seu custo (ZADROZNY; LANGFORD; ABE, 2003).

Para o classificador utilizado nesta dissertação, de acordo com a literatura, considera-se que a aprendizagem sensível ao custo pode fornecer melhores resultados se comparados a outros métodos de aprendizagem. Desta forma, optou-se por utilizar o filtro *Resample* disponibilizado na ferramenta Weka versão 3.5.7 que



utiliza a sensibilidade ao custo aplicando um metaclassificador que manipula as instâncias de treino de forma a obter classificadores sensíveis ao custo em conjunto com o classificador base, tornando-o sensível a valores de custos que são manipuláveis (MARK et al, 2009).

Nas classificações realizadas pelo Weka, o desbalanceamento de classes está fortemente relacionado com a abordagem de aprendizado sensível ao custo. Desta forma, entende-se, por exemplo, que um caso de falso negativo é mais custoso que um caso de falso positivo (HAN; KAMBER, 2006).

## **2.5 Métodos de classificação de dados**

Existem vários métodos que podem ser utilizados para solucionar problemas de classificação. Mediante a proposta desta dissertação, o método escolhido deve utilizar a lógica nebulosa como método computacional para solucionar o problema de classificação proposto, assim, ficou definido que o método utilizado deve ser fundamentado em regras nebulosas se...então.

A escolha deste método tem sua justificativa na literatura. Os sistemas nebulosos baseados em regras se...então aplicadas a conjuntos nebulosos foram criados com o intuito de representar experiências da vida real, visto que podem ser utilizados para a representação do conhecimento, e, portanto, são muito úteis no desenvolvimento de sistemas inteligentes, em especial os especialistas e os multiagentes. Para muitos problemas reais, a imprecisão dos dados e a incerteza do conhecimento são, por natureza, parte do problema em si, e raciocinar considerando esses aspectos sem uma fundamentação adequada pode gerar inferências imprecisas (MUNAKATA, 2008).

A lógica nebulosa coloca-se como o principal instrumento para uma representação mais adequada do conhecimento e do próprio raciocínio. Isso se deve à sua capacidade de lidar com incertezas, raciocínio aproximado, termos vagos e ambíguos, como os que as pessoas pensam, isso tudo indo além do escopo das lógicas clássicas. Dessa forma, a lógica nebulosa permite aos sistemas computacionais inteligentes “raciocinar” considerando aspectos inerentes à incerteza, possibilitando aos sistemas computacionais, tomar decisões muito próximas as decisões dos “humanos” (MUNAKATA, 2008).

De acordo com o proposto nesta dissertação, o classificador é capaz de determinar qual o aspecto cognitivo predominante em cada indivíduo e para isso é aplicado um questionário com perguntas e respostas fundamentadas nas escolhas humanas. O questionário está disponibilizado no Apêndice B.

O processamento destas informações segue um padrão de computação que se aproxima do raciocínio humano, devido às possibilidades de respostas não exatas para as questões, ou seja, diferentes de nunca e sempre. Além dessas, elas podem ser “quase sempre” ou “quase nunca” e isso diferencia a lógica aplicada ao classificador em relação à lógica convencional. Desta forma, as respostas devem aproximar ou afastar o indivíduo de cada conjunto nebuloso que representa cada um dos aspectos cognitivos utilizados para classificar os indivíduos. Mesmo que as respostas não sejam exatas, essas são convertidas pelo classificador em parâmetros de classificação e não podem ser descartadas. Isso é possível devido ao método computacional utilizado pelo classificador estar fundamentado na lógica nebulosa.

### **2.5.1 Método *Simple Fuzzy Grid***

A Geração de regras nebulosas se...então, a partir de dados numéricos para problemas de classificação padrão, consiste em duas fases: criar subconjuntos nebulosos a partir de um espaço padrão e determinar regras nebulosas se...então para cada subconjunto nebuloso criado.

Na partição nebulosa pelo método *Simple Fuzzy Grid*, o número de subconjuntos nebulosos é definido por  $k^n$ . Sendo que, “ $n$ ” é o número de dimensões do espaço nebuloso e “ $K$ ” é o número de conjuntos nebulosos. Na Figura 2.5, um espaço padrão bidimensional está dividido em 25 subconjuntos nebulosos, este valor é obtido substituindo  $K$  e  $n$  em  $k^n$ , com  $K = 5$  e  $n = 2$ .

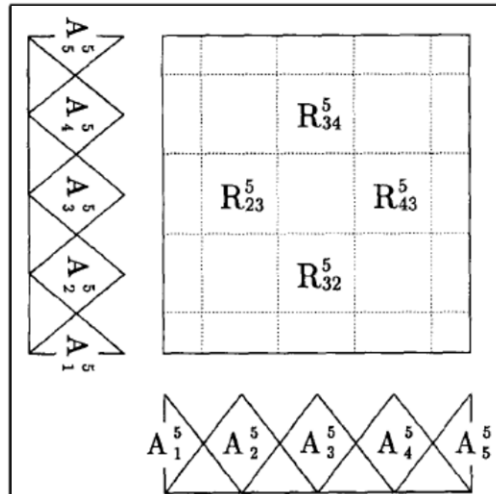


Figura 2.5: Exemplo de partição nebulosa utilizando o método *Simple Fuzzy Grid*.

A aplicação desse método se torna inviável para problemas de partição nebulosas que têm o número de subconjuntos nebulosos aumentado exponencialmente. Isso ocorre quando as dimensões de um espaço padrão aumentam. Por exemplo, um espaço padrão de oito dimensões e cinco conjuntos nebulosos em cada eixo resulta em 390625 subconjuntos nebulosos.

Uma das justificativas da aplicação deste método para este problema em específico é que o número de subconjuntos nebulosos não aumenta exponencialmente. O classificador proposto nesta dissertação utiliza duas dimensões que permitem a comparação de cada conjunto com todos os outros e utiliza sete conjuntos nebulosos para classificar os resultados e isso ocorre em todos os testes. Se, no futuro, a quantidade de conjuntos nebulosos a serem analisados aumentar significativamente, um novo estudo de viabilidade do método poderá ser realizado.

O número de subconjuntos nebulosos utilizados nesse classificador é obtido substituindo “ $k$ ” por 7 e “ $n$ ” por 2 em  $k^n$ . Desta forma, tem-se que o total de subconjuntos nebulosos é de 49, onde  $k$  é igual ao total de conjuntos nebulosos que serão utilizados e  $n$  é igual as dimensões. Por meio da Figura 2.6 é possível visualizar a representação da grade utilizada para o problema proposto com duas dimensões particionadas em sete aspectos cognitivos, sendo que cada um dos aspectos cognitivos é representado por um conjunto nebuloso e os quarenta e nove subconjuntos nebulosos são representados por cada espaço formado da subdivisão da grade.

$S_7^7$	$A_{11}^{49}$	$A_{12}^{49}$	$A_{13}^{49}$	$A_{14}^{49}$	$A_{15}^{49}$	$A_{16}^{49}$	$A_{17}^{49}$
$S_6^7$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
$S_5^7$							
$S_4^7$							
$S_3^7$							
$S_2^7$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
$S_1^7$	$A_{71}^{49}$	$A_{72}^{49}$	$A_{73}^{49}$	$A_{74}^{49}$	$A_{75}^{49}$	$A_{76}^{49}$	$A_{77}^{49}$
	$S_1^7$	$S_2^7$	$S_3^7$	$S_4^7$	$S_5^7$	$S_6^7$	$S_7^7$

Figura 2.6: Representação da grade utilizada na solução do problema proposto.

$S_i^7$  representa cada um dos sete conjuntos nebulosos nas duas dimensões e  $A_j^{49}$  representa um dos quarenta e nove subconjuntos nebulosos criados na subdivisão da grade.

Os valores de pertinência obtidos para cada subconjunto nebuloso são utilizados para classificar os indivíduos e as informações obtidas são cruzadas entre elas. Desta forma é possível utilizar somente duas dimensões para que as comparações possam ser realizadas.

O total de 49 subconjuntos nebulosos pode ser considerado um número baixo, isso possibilita a aplicação desse método em que, serão geradas 49 regras nebulosas se...então, sendo uma para cada subconjunto. Outra característica do método *Simple Fuzzy Grid* é possuir um conjunto de auto geração de regras do tipo se...então. Por meio da equação (3) é possível observar a estrutura da regra utilizada pelo método.

$$\text{Regra } R_{ij}^k: \text{SE } x_1 \text{ é } A_i^k \text{ e } x_2 \text{ é } A_j^k \text{ ENTÃO } (x_1, x_2) \text{ pertencem a } C_{ij}^k \quad (3)$$

Para esta equação tem-se que  $K$  é o número de subconjuntos nebulosos,  $R_{ij}^k$  é o nome da regra,  $(x_1 \text{ e } x_2)$  são os pontos de um padrão a classificar,  $A_i^k$  e  $A_j^k$  são cada um dos subconjuntos nebulosos,  $C_{ij}^k$  é a consequência da regra,  $CF_{ij}^k$  é o grau de certeza da regra e  $i \text{ e } j = 1, \dots, K$  representam cada uma das dimensões do problema. As regras são criadas com o objetivo de testar todas as possíveis combinações dos subespaços (ISHIBUCHI; NOZAKI; TANAKA, 1993).

No método *Simple Fuzzy Grid*, os subconjuntos nebulosos  $A_i^k$  são definidos por uma função triangular simétrica descrita pelas equações (4), (5) e (6), onde o valor central da função  $a_i^k$  e o suporte é o intervalo aberto  $(a_i^k - b^k, a_i^k + b^k)$  (ISHIBUCHI; NOZAKI; TANAKA, 1993).

$$\mu_i^K(x) = \max\{1 - |x - a_i^K|/b^K, 0\}, i = 1, 2, \dots, K \quad (4)$$

$$a_i^K = (i - 1/K - 1), i = 1, 2, \dots, K \quad (5)$$

$$b^K = 1/(K - 1) \quad (6)$$

Existem alguns passos que devem ser seguidos para a geração das regras nebulosas, são eles (ISHIBUCHI; NOZAKI; TANAKA, 1993):

1 – Calcular  $\beta_{Ct}$  para  $t = 1, 2, \dots, M$ , sendo que,  $M$  é o número de classes do problema e  $\beta_{Ct}$  é definido pela equação (7).

$$\beta_{Ct} = \sum_{p \in Ct} \mu_i^K(x_{p1}) \cdot \mu_j^K(x_{p1}) \quad (7)$$

Para a equação (7)  $Ct$  faz referência à classe utilizada na regra em questão e  $p1, p2, \dots, pm$  determinam o número de padrões do problema.

2 – O próximo passo é encontrar as classes  $X$  ( $CX$ ), para isso deve-se utilizar a equação (8).

$$\beta_{CX} = \max\{\beta_{C1}, \beta_{C2}, \dots, \beta_{CM}\} \quad (8)$$

Se duas ou mais classes têm o máximo valor na equação (8), então a regra se...então, que corresponde ao subespaço nebuloso  $A_i^K \times A_j^K$  não é gerada. Caso contrário  $C_{ij}^K$  é determinada como sendo  $CX$  na equação (8) (ISHIBUCHI; NOZAKI; TANAKA, 1993).

3 – Se uma única classe tem o máximo valor na equação (8),  $CF_{ij}^K$  é determinada pela equação (9).

$$CF_{ij}^K = |\beta_{CX} - \beta| / \sum_{t=1}^M \beta_{Ct} \quad (9)$$

A equação (10) determina o valor de  $\beta$  como sendo:

$$\beta = \sum_{Ct \neq CX} \beta_{Ct} / (M - 1) \quad (10)$$

Para este procedimento, a consequência de  $C_{ij}^K$  é determinada a partir da classe que tem a maior soma de  $\mu_i^K(x_{p1}) \cdot \mu_j^K(x_{p2})$ . Neste caso o grau de certeza ( $CF$ ) oscila no intervalo de 0 a 1. Se todos os padrões em  $A_i^K \times A_j^K$  são da classe  $X$  ( $CX$ ), o grau de certeza é  $CF_{ij}^K = 1$ , do contrário  $CF_{ij}^K < 1$ .

Se for necessário classificar um novo padrão, deve ser seguido um novo procedimento que contém duas etapas, são elas (ISHIBUCHI; NOZAKI; TANAKA, 1993):

Assumindo que  $x_p = (x_{p1}, x_{p2})$  é um novo padrão.

1 – Calcular  $\alpha_{Ct}$  para  $t=1, 2, \dots, M$  de acordo com a equação (11).

$$\alpha_{Ct} = \max\{\mu_i^K(x_{p1}) \cdot \mu_j^K(x_{p2}) \cdot CF_{ij}^K \mid C_{ij}^K = Ct; R_{ij}^K \in S_R\} \quad (11)$$

Sendo que  $S_R$  é o conjunto de todas as regras geradas.

2 – Encontrar a classe  $X(CX)$  que satisfaça a equação (12).

$$\alpha_{CX} = \max\{\alpha_{C1}, \alpha_{C2}, \dots, \alpha_{CM}\} \quad (12)$$

Caso existam duas ou mais classes com o máximo valor na equação (12), então  $x_p$  não pode ser classificado, caso contrário, associar  $x_p$  à classe  $X(CX)$  determinada pela equação (12) (ISHIBUCHI; NOZAKI; TANAKA, 1993).

Para esse procedimento, o resultado da inferência nebulosa se...então é a consequência que tem o máximo produto de  $\mu_i^K(x_{p1}) \cdot \mu_j^K(x_{p2}) \cdot CF_{ij}^K$ . Se não existem regras nebulosas se...então tal que,  $\mu_i^K(x_{p1}) \cdot \mu_j^K(x_{p2}) > 0$  e  $CF_{ij}^K > 0$  para  $x_p$ , o novo padrão de  $x_p$  não pode ser classificado (ISHIBUCHI; NOZAKI; TANAKA, 1993).

## 2.5.2 Método de Regras Nebulosas “se”...”então” Distribuídas

Neste método, as regras nebulosas correspondentes a várias partições nebulosas são simultaneamente utilizadas em inferência nebulosa, em outras palavras, várias tabelas de regras nebulosas com diferentes partições nebulosas são sobrepostas. O processo de utilizar um conjunto de regras nebulosas de mais de uma tabela de regras nebulosa é chamado de 'regras nebulosas distribuídas', esse nome é devido ao conhecimento ser derivado de dados numéricos que estão distribuídos em diferentes tabelas de regras nebulosas.

Quando as regras nebulosas derivam a partir de dados numéricos, é importante determinar uma partição nebulosa adequada uma vez que o desempenho das regras nebulosas é afetado pela partição nebulosa.

A determinação da partição nebulosa, no entanto, não é fácil, já que não existe um método geral estabelecido e a sua falta é considerada um dos motivos para a não utilização deste método para o problema proposto nesta dissertação.

Para que se possa obter uma acurácia robusta com a utilização desse método, propõe-se a utilização de soluções simplificadas que consistem em reduzir o número total das regras nebulosas distribuídas por meio da remoção de algumas regras desnecessárias. O procedimento de remoção destas regras será demonstrado mais adiante. Esta remoção pode gerar inconsistência nos resultados em relação ao projeto desenvolvido nesta dissertação, uma vez que tal seleção de regras nebulosas pode exigir um procedimento complicado que pode gerar resultados com precisão inferior ao método *Simple Fuzzy Grid*.

Associar a proposta desta dissertação à possível imprecisão nos resultados da classificação a ser realizada, inviabiliza a utilização desse método, pois um indivíduo pode ser classificado em outro aspecto cognitivo que pode não ser o predominante para ele.

Para exemplificar este método é suposto um espaço padrão definido como sendo um quadrado unitário  $[0, 1] \times [0, 1]$ . É definida também a partição nebulosa  $\{A_1^L, A_2^L, \dots, A_L^L\}$  que é usada para cada um dos dois eixos no espaço padrão, onde  $L$  é o número de subespaços nebulosos em cada eixo. Neste caso, as regras normais nebulosas podem ser representadas da seguinte forma:

$$\text{Se } x_1 \text{ é } A_i^L \text{ e } x_2 \text{ é } A_j^L \text{ então } \dots, \quad i, j = 1, 2, \dots, L \quad (13)$$

O número total de regras nebulosas na equação (13) é  $L^2$ . Por meio da Figura 2.7 é possível visualizar as tabelas de regras nebulosas correspondentes a  $L=2$  e  $L=3$ .

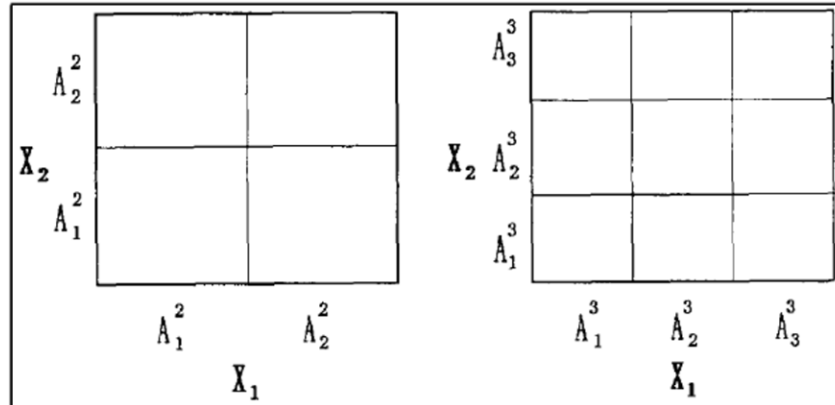


Figura 2.7: Estrutura das tabelas de regras nebulosas correspondentes a  $L=2$  e  $L=3$ .

Por outro lado, no método proposto, podem ser utilizadas simultaneamente  $L - 1$  tabelas de regras nebulosas. Ou seja, no caso de regras nebulosas distribuídas, devem-se utilizar regras nebulosas correspondendo a  $L - 1$  partições nebulosas:  $\{A_1^2, A_2^2\}, \{A_1^3, A_2^3, A_3^3\}, \dots, \{A_1^L, A_2^L, \dots, A_L^L\}$ . No caso de  $L = 3$ , duas tabelas de regras nebulosas são utilizadas simultaneamente. Um conjunto de regras nebulosas distribuídas pode ser representado como:

$$\text{Se } x_1 \text{ é } A_i^K \text{ e } x_2 \text{ é } A_j^K \text{ então } \dots, \quad i, j = 1, 2, \dots, K; K = 2, 3, \dots, L \quad (14)$$

Sendo que  $K$  é o número de subespaços nebulosos em cada eixo em cada partição nebulosa. O número total das regras nebulosas distribuídas na equação (14) é  $2^2 + 3^2 + \dots + L^2$ , ou seja,  $\frac{1}{6}L(L+1)(2L+1) - 1$ . Todas as regras nebulosas correspondentes às partições nebulosas com  $K = 2, 3, \dots, L$  são usadas, simultaneamente, na inferência nebulosa.

O número total das regras nebulosas distribuídas pode ser reduzido por meio da remoção de algumas regras desnecessárias a partir da equação (14), uma vez que tal seleção de regras nebulosas pode exigir um procedimento complicado, porém essa remoção pode gerar inconsistência nos resultados em relação ao projeto desenvolvido nesta dissertação.



Partindo do exposto com relação ao método de classificação utilizando regras nebulosas distribuídas e considerando a aplicação no problema proposto temos  $K=7$ , onde  $K$  é o número de subespaços nebulosos em cada eixo, este valor é correspondente ao número de aspectos cognitivos utilizados como parâmetro de classificação dos indivíduos, sendo este igual a sete.

O número de subespaços nebulosos é utilizado para determinar o número de tabelas de regras nebulosas; cada subespaço nebuloso possui uma tabela em correspondência a ele e o número de tabelas é expresso pela letra “ $L$ ”. Ao realizar o processo de classificação por meio do método que utiliza regras nebulosas distribuídas, tem-se que, para cada subespaço a ser analisado, são utilizadas  $L - 1$  regras, simultaneamente, ou seja, são utilizadas seis regras nebulosas.

Por meio da equação (15) é possível calcular o número total de regras nebulosas que devem ser utilizadas para solucionar o problema proposto.

$$\frac{1}{6}L(L + 1) \cdot (2L + 1) - 1, \text{ onde } L = 7 \quad (15)$$

Se for utilizado o método de regras nebulosas distribuídas no problema proposto, serão necessárias 139 regras nebulosas contra 49 regras nebulosas se for aplicado o método *Simple Fuzzy Grid*.

## 2.6 Método de agrupamento dos dados de saída

Para este processo é aplicado um método de agrupamento de respostas em classes que faz uso do *cluster EM (Expectation Maximization)* que realiza a distribuição de probabilidade para a instância que está sendo analisada, a qual, por sua vez, indica a probabilidade de cada instância pertencer a cada um dos agrupamentos. O *cluster EM* pode ser ajustado para decidir quantos grupos criar por meio de validação cruzada, ou é possível especificar quantos grupos devem ser gerados (DEMPSTER; LAIRD; RUBIN, 1977).

Para os testes realizados é utilizada a validação cruzada, que tem por característica criar somente a quantidade necessária de grupos para classificar as respostas. Para o problema proposto, podem ser criados até sete grupos, cada um referente a um dos sete aspectos cognitivos utilizados para classificar os indivíduos.

O método de agrupamento utilizado faz uso do algoritmo de agrupamento probabilístico que recebe o nome de *Expectation Maximization (EM)*, sendo este, fundamentado em um modelo estatístico que faz uso do modelo de misturas Gaussianas (DEMPSTER; LAIRD; RUBIN, 1977).

O modelo de misturas Gaussianas assume que todos os atributos são variáveis aleatórias independentes. Uma mistura é um conjunto de  $N$  distribuições probabilísticas na qual, cada distribuição representa um grupo. Uma instância individual é assinalada com uma probabilidade dada por um certo conjunto de valores de atributos em um determinado grupo. Com isto, o procedimento geral do algoritmo *EM* consiste em (DEMPSTER; LAIRD; RUBIN, 1977):

1. Definir os valores iniciais para os parâmetros ( $\sigma$  desvio padrão e  $\mu$  média da distribuição para cada grupo).
2. Assumir o valor de uma instância.
3. Utilizar a função de densidade probabilística de cada grupo para calcular a probabilidade da instância pertencer a cada grupo (sete fórmulas de distribuição probabilística).
4. Determinar o grupo que contém a instância por meio do valor calculado para a probabilidade de pertencer a cada grupo (é classificada no grupo em que mais se aproximar).
5. Retornar ao passo 2.

Para o cálculo da probabilidade de cada instância pertencer a cada grupo é utilizada a função de distribuição probabilística definida pela equação (16).

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}} \quad (16)$$

Para a equação (16) tem-se que  $\sigma$  é o desvio padrão,  $\mu$  é a média da distribuição e  $\sigma^2$  é a variância. Ambos valores são obtidos por meio do algoritmo de agrupamento *EM* e são calculados para cada grupo utilizado na classificação.

### 3 METODOLOGIA

Este capítulo descreve a metodologia utilizada para a implementação do classificador por meio da utilização do software Weka versão 3.5.7. O questionário utilizado para a coleta dos dados serve como base de dados de entrada para o classificador. São demonstrados os métodos *Simple Fuzzy Grid* utilizado no processo de classificação e o método de agrupamento das respostas em classes que faz uso do *cluster EM*. São realizados testes iniciais no classificador para determinar a acurácia.

#### 3.1 WEKA (WAIKATO ENVIRONMENT FOR KNOWLEDGE ANALISYS)

O *software* Weka é uma ferramenta de mineração de dados que é desenvolvida por um grupo de pesquisadores da Universidade de Waikato localizada na Nova Zelândia. É um *software* livre, de código aberto e está disponível para *download* no link [www.cs.waikato.ac.nz/ml/weka/downloading.html](http://www.cs.waikato.ac.nz/ml/weka/downloading.html).

Com o passar dos anos, a ferramenta Weka se consolidou como uma das ferramentas mais utilizadas em ambiente acadêmico para a mineração dos dados, tendo como ponto forte a classificação, mas também é capaz de minerar regras de associação e grupos de dados. Pode ser utilizada no modo console ou por meio de interface gráfica.

Existem várias versões da ferramenta Weka, cada uma delas com características e classificadores diferentes, sendo que a versão utilizada nesta dissertação é a 3.5.7, justamente por fornecer a lógica nebulosa como um dos métodos computacionais de classificação e disponibilizar o classificador *Simple Fuzzy Grid* utilizado para os testes.

A maioria das ferramentas computacionais utiliza arquivos de extensões específicas para realizar o processamento; o Weka utiliza para a entrada de dados os arquivos de extensão *arff*, *csv*, *xrff* entre outros. A extensão utilizada nesta dissertação é a *csv* que pode ser obtida por meio da conversão de planilhas do Microsoft Excel do formato *xls* ou *xlsx* para o formato *csv*. A justificativa para esta utilização é que o questionário utilizado para fornecer os parâmetros de entrada foi desenvolvido utilizando a ferramenta Microsoft Excel e a conversão para o formato

csv permite que os dados inseridos na planilha sejam aceitos e processados pelo Weka.

### 3.1.1 Processamento realizado pelo Weka

As etapas de processamento realizadas pelo Weka para a obtenção do conhecimento podem ser visualizadas por meio da Figura 3.1. Inicialmente, na etapa de seleção, são selecionados, na base de dados, os dados a serem analisados. Na etapa de processamento, os dados são processados e entendidos pelo sistema que, na etapa de transformação, os converte em dados que possam ser minerados e entendidos pelo Weka. Na etapa de mineração é realizado o treinamento que deverá definir os padrões a serem localizados e, logo em seguida, é realizada uma busca por padrões que foram previamente informados como sendo os parâmetros para a definição das respostas, o resultado da mineração dos dados é interpretado e como saída é fornecida uma base de conhecimento. É importante ressaltar que, ao término da última etapa, o Weka pode realimentar uma das etapas anteriores refazendo o processo a partir da etapa em que ele retornou até que um padrão seja obtido e uma resposta seja gerada.

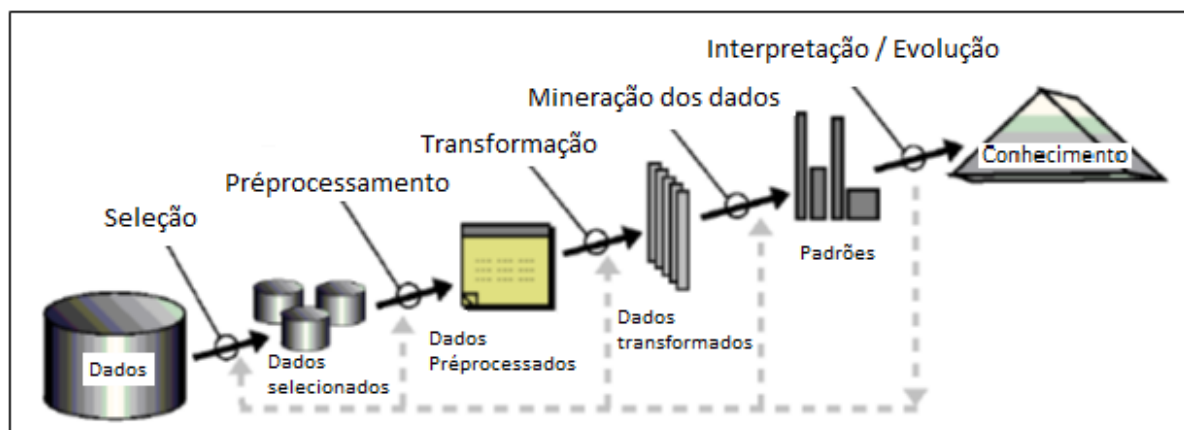


Figura 3.1: Etapas de processamento realizadas pelo Weka.

### 3.2 QUESTIONÁRIO UTILIZADO

Fundamentando-se na Teoria das Inteligências Múltiplas proposta Gardner foi desenvolvido um questionário contendo vinte e uma questões, sendo três questões para cada Inteligência Múltipla que está associada a um aspecto cognitivo, não ficando explícito para os indivíduos que responderam ao questionário, qual é o objetivo deste e quais questões representam cada aspecto cognitivo para que as respostas não fossem influenciadas. Foi comentado somente que o questionário deveria ser respondido com respostas numéricas de um a quatro, conforme Tabela 3.1 e que esse teste tem por finalidade avaliar as preferências de cada indivíduo.

Tabela 3.1: Respostas correspondentes a cada valor numérico respondido.

<b>Valor da resposta</b>	<b>Representação da resposta</b>
1	Nunca
2	Quase nunca
3	Quase sempre
4	Sempre

Fonte: Elaborada pelo próprio autor, 2016.

É importante ressaltar que as questões utilizadas no questionário foram elaboradas pelo próprio autor desta dissertação e que essas foram desenvolvidas após análise de questionário utilizado para teste de descoberta do perfil de indivíduos por Inteligências Múltiplas baseado no modelo proposto por Howard Gardner. O modelo que serve de base para o questionário utilizado está disponibilizado no apêndice A.

O questionário original contém setenta questões, sendo dez questões para cada Inteligência Múltipla, com quatro possibilidades de resposta, dentre elas tem-se agradável, pouco agradável, pouco desagradável e desagradável.

As Inteligências Múltiplas analisadas no questionário original são as mesmas que as utilizadas nesta dissertação. São elas: linguística, lógico-matemática, musical, corporal cinestésica, espacial visual, interpessoal e intrapessoal, sendo estas relacionadas aos aspectos cognitivos que serão utilizados como parâmetros de classificação dos indivíduos.

A diminuição no número de questões de setenta para vinte e uma está relacionada a análise do questionário original. Após análise, foi detectado que a grande maioria das questões não fazem parte do perfil dos indivíduos selecionados para os testes ou estão fora do ambiente em que os indivíduos estão inseridos. Além da exclusão de algumas questões, foi necessário criar outras e alterar algumas das questões originais.

Com relação as possibilidades de resposta para as questões, os indivíduos podem escolher entre quatro níveis de resposta para expressar a opinião deles em relação a cada questão, conforme apresentado na Tabela 3.1.

Outros fatores foram considerados para a elaboração do questionário. Um deles é o público alvo selecionado, todos os indivíduos são estudantes de cursos técnicos com aulas presenciais em diferentes áreas de conhecimento, sendo elas humanas, exatas e biológicas. Outro fator relevante consiste em elaborar questões que exemplifiquem situações inseridas no contexto em que os indivíduos se encontram inseridos, procurando vivenciar o cotidiano destes indivíduos para tentar permitir o fácil entendimento das questões.

Acredita-se também que um questionário extenso é desnecessário devido à demora em respondê-lo por completo, podendo influenciar na qualidade da análise das questões por parte dos indivíduos que responderam ao questionário. Além disso, o desenvolvimento de uma abordagem completamente nova para obter informações relevantes aos testes por parte dos indivíduos não é o objetivo deste trabalho. Assim, ficou decidido a utilização de um questionário composto por vinte e uma questões, sendo três para cada aspecto cognitivo associado a teoria de Gardner.

Por meio da Tabela 3.2, podem ser visualizadas duas questões aplicadas a todos os indivíduos do teste. O questionário contém as colunas código, questão, resposta e aspecto cognitivo. A coluna código foi utilizada como atributo de identificação de cada questão no *software* Weka, a coluna questão possui todas as vinte e uma questões a que os indivíduos responderam numericamente na coluna resposta conforme Tabela 3.1, que tem por finalidade fornecer ao *software* Weka qual o grau de pertinência da resposta do indivíduo para cada pergunta referente a um dos sete aspectos cognitivos que o classifica. Por fim, a coluna aspecto cognitivo

associa a questão e a resposta a cada aspecto cognitivo. O questionário contendo todas as vinte e uma questões está disponibilizado no apêndice B.

Tabela 3.2: Modelo do questionário.

<b>Código</b>	<b>Questão</b>	<b>Resposta</b>	<b>Aspecto cognitivo</b>
1	Faça uma análise sobre as músicas que você costuma ouvir, associe o ritmo musical ao seu estado de espírito. As músicas que você ouve representam o seu estado de espírito no momento em que você as ouve?		Musical
2	Quando você precisa resolver problemas que envolvem cálculos matemáticos, você prefere tentar resolver mentalmente ou procura alguma ferramenta de cálculo que facilite seu trabalho com o objetivo de ganhar tempo? Responda em relação à frequência de calcular mentalmente.		Lógico matemático

Fonte: Elaborada pelo próprio autor, 2016.

Por meio da Tabela 3.3, podem ser observadas duas perguntas de um exemplo de questionário respondido. A coluna aspecto cognitivo foi omitida do formulário para não influenciar as respostas, sendo esta utilizada somente para o processamento. Vale lembrar que o significado das respostas pode ser encontrado na Tabela 3.1. Um questionário respondido está disponível no Apêndice C.

Tabela 3.3: Exemplo de questionário respondido

<b>Código</b>	<b>Questão</b>	<b>Resposta</b>
1	Faça uma análise sobre as músicas que você costuma ouvir, associe o ritmo musical ao seu estado de espírito. As músicas que você ouve representam o seu estado de espírito no momento em que você as ouve?	4
2	Quando você precisa resolver problemas que envolvem cálculos matemáticos, você prefere tentar resolver mentalmente ou procura alguma ferramenta de cálculo que facilite seu trabalho com o objetivo de ganhar tempo? Responda em relação a frequência de calcular mentalmente.	3

Fonte: Elaborada pelo próprio autor, 2016.

### 3.3 GRUPOS DE INDIVÍDUOS SELECIONADOS PARA OS TESTES

O questionário foi aplicado a grupos de indivíduos que pertencem a três diferentes áreas da ciência, sendo elas, exatas, humanas e biológicas. Na área de exatas, os indivíduos são alunos do curso técnico em manutenção e suporte em informática. Na área de humanas, os indivíduos são alunos dos cursos técnicos em administração de empresas e de recursos humanos. Na área de biológicas, os indivíduos são alunos do curso técnico em enfermagem e do curso técnico em zootecnia.

Existe ainda um grupo de indivíduos diferenciado que é constituído por alunos que frequentam o ensino médio regular integrado ao curso técnico em informática. Esse grupo, em particular, pertence a disciplinas que estão classificadas nas três áreas da ciência abordadas nesta dissertação. Os alunos que estão matriculados neste curso, em sua grande maioria, não o escolheram por opção própria e sim por indicação dos pais. Com relação aos demais cursos, os alunos matriculados escolheram-nos de vontade.

Por meio da Tabela 3.4, é possível observar os grupos de indivíduos que responderam ao questionário, a faixa etária de cada grupo bem como a área da ciência a que o grupo pertence.

Tabela 3.4: Grupos de indivíduos que responderam ao questionário.

<b>Área de conhecimento</b>	<b>Faixa etária</b>	<b>Quantidade de indivíduos</b>
Exatas – técnico em manutenção e suporte em informática	De 17 a 20 anos	12
Humanas - administração de empresas	De 16 a 23 anos	16
Humanas - recursos humanos	De 16 a 38 anos	23
Biológicas – técnico em enfermagem	De 18 a 45 anos	58
Biológicas – Técnico em zootecnia	De 16 a 24 anos	21
Exatas, humanas e biológicas - ensino médio integrado com o técnico em informática	De 16 a 18 anos	35
<b>Total de indivíduos que responderam ao questionário</b>		<b>165</b>

Fonte: Elaborada pelo próprio autor, 2016.



Todos os indivíduos são alunos pertencentes à Escola Técnica Estadual Etec de Lins, localizada na cidade de Lins – SP. É importante ressaltar que todos os indivíduos que responderam ao questionário não tiveram dificuldades em interpretar e responder às questões e todos os indivíduos responderam ao teste em no máximo sessenta minutos, sem que uma questão fosse deixada de ser respondida.

### **3.4 ANÁLISE ESTATÍSTICA DAS AMOSTRAS**

Após a definição da instituição e dos grupos de indivíduos para realização dos testes e antes da aplicação questionário utilizado para a coleta dos dados de entrada para o classificador, foi realizada uma pesquisa quantitativa que tem por objetivo determinar a quantidade de amostras mínima para cada grupo de indivíduos. Esta quantidade mínima de amostras foi suficiente para representar todos os indivíduos de cada grupo.

O total de indivíduos matriculados na Escola Técnica Estadual – Etec de Lins era de 176 (cento e setenta e seis) alunos quando a coleta de dados foi realizada, isso levando-se em consideração todos os cursos de nível técnico e o curso de ensino médio, que têm as aulas ministradas nos períodos matutino, vespertino, integral e noturno.

Para que se possam fazer inferências para uma população utilizando-se resultados obtidos a partir de uma amostra, é necessário que essa amostra seja coletada com alguns cuidados, como a atenção ao seu tamanho mínimo e à forma de coleta, que deve ser probabilística, ou seja, deve obedecer ao padrão AAS (amostragem aleatória simples). De acordo com esse padrão, cada elemento coletado tem tanta probabilidade de ser observado quanto qualquer outro elemento da população, e sua observação pode ocorrer uma única vez no processo de amostragem (GIL, 1999). Existem várias formas de se determinar o tamanho de um espaço amostral e a opção utilizada nesta dissertação está descrita pela equação (17):

$$n = \frac{N.Z^2}{((N-1).E^2 + Z^2)} \quad (17)$$

Considerando a equação (16), tem-se “ $Z$ ”=1,645 sendo que “ $Z$ ” é o desvio do valor médio aceitável para alcançar o nível de confiança desejado. Em função do nível de confiança que se pretende alcançar, deve ser utilizado um valor determinado que é dado pela forma da distribuição de Gauss. O nível de confiança adotado é de 90% (noventa por cento), assim, o valor de “ $Z$ ” deve ser igual a 1,645 justificando desta forma o valor adotado e “ $E$ ” é a margem de erro máximo admitido. Considerando que o número total de indivíduos é relativamente baixo e devido à dificuldade de se conseguir entrevistar 100% dos indivíduos devido à ausência de alguns alunos, adotou-se uma margem de erro de 20% (vinte por cento). “ $n$ ” é o total de amostras que devem compor o espaço amostral e “ $N$ ” é o tamanho da população.

Por meio da Tabela 3.5, é possível observar as informações referentes aos grupos de indivíduos que participaram dos testes, bem como, a quantidade de alunos que responderam ao questionário e o total de alunos que pertencem a cada grupo. O número de amostras coletadas para validar os resultados é superior ao mínimo de amostras necessárias para todos os grupos de indivíduos entrevistados, com isso é possível utilizar a quantidade de indivíduos entrevistados para representar todo o grupo de indivíduos.

Tabela 3.5: Informações referentes ao espaço amostral.

Áreas do conhecimento	Cursos	Total de alunos	Nº necessário de alunos	Nº de alunos analisados
Exatas	Técnico em manutenção e suporte em informática	13	11,03	12
Humanas	Técnico em administração	17	13,74	16
Humanas	Técnico em Recursos Humanos	26	18,97	23
Biológicas	Técnico em enfermagem	61	32,29	58
Biológicas	Técnico em Zootecnia	22	16,77	21
Exatas, humanas e biológicas	Ensino médio integrado ao técnico em informática	37	24,13	35

Fonte: Elaborada pelo próprio autor, 2016.

### 3.5 REALIZANDO A CLASSIFICAÇÃO DE UM INDIVÍDUO.

De acordo com o que foi escrito anteriormente, o questionário utilizado nos testes contém vinte e uma questões, sendo três questões para cada tipo de aspecto cognitivo. A classificação dos indivíduos ocorre por meio do cruzamento dos valores de pertinência obtidos para cada um dos subconjuntos nebulosos criados por meio da utilização do método *Simple Fuzzy Grid* (item 2.5.1).

Para o processo de classificação é necessário comparar cada subconjunto nebuloso com todos os demais subconjuntos. Levando-se em conta esta informação, o espaço padrão utilizado para a classificação é bidimensional definido por,  $X_1 \times X_2$  e este espaço padrão é dividido em subespaços de acordo com o método *Simple Fuzzy Grid*. Para o problema proposto, aplicando os conceitos do método, tem-se quarenta e nove subespaços e cada um deles é definido por uma regra. Por meio da Figura 3.2, pode ser visualizada a representação do subespaço nebuloso no espaço padrão bidimensional gerado a partir do produto cartesiano dos conjuntos A e B.

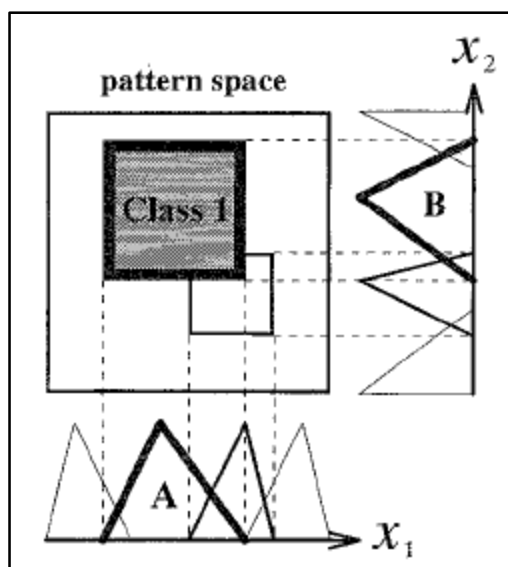


Figura 3.2: Representação de um subespaço no espaço bidimensional.

Este método de classificação emprega, simultaneamente, várias partições nebulosas de tamanhos correspondentes ao padrão de treinamento em um único sistema de classificação baseado em regras nebulosas. Considerando que existem sete conjuntos nebulosos e que cada conjunto é representado por três questões, os conjuntos nebulosos devem ter o mesmo tamanho no plano bidimensional

totalizando uma grade com quarenta e nove subconjuntos nebulosos de mesmo tamanho.

Ao iniciar o processo de classificação, as respostas das questões são analisadas e de acordo com a resposta é atribuído a ela um peso, ou seja, um fator de pertinência desta resposta em relação ao aspecto cognitivo ao qual ela pertence.

O fator de pertinência atribuído a cada resposta do questionário é utilizado para plotar cada resposta na grade, isto ocorre por meio da associação do fator de pertinência a um ponto correspondente às coordenadas cartesianas da grade. Tomemos como exemplo, que a resposta de uma questão teve atribuído a ela o fator de pertinência 0.9 em relação ao aspecto cognitivo a que ela pertence, desta forma, assume-se que  $x_1$  é  $A$ ,  $x_2$  é  $B$ ,  $(x_1, x_2)$  pertencem à classe da resposta,  $A$  e  $B$  são definidos como sendo os subconjuntos em  $X_1$  e  $X_2$ , respectivamente, e que estes representam os eixos de coordenadas da grade e considerando a grade apresentada pelo método *Simple Fuzzy Grid* na Figura 2.6, o ponto correspondente ao fator de pertinência do exemplo é  $(A_{49}^{37}, B_{49}^{37})$ .

Esta coordenada obtida por meio do fator de pertinência é submetida ao conjunto das quarenta e nove regras nebulosas classificando cada resposta em um dos subconjuntos nebulosos. Esse processo é repetido para cada uma das vinte e uma questões do questionário. A equação 3 representa uma das quarenta e nove regras.

$$\text{Regra } R_{ij}^k: \text{SE } x_1 \text{ é } A_i^k \text{ e } x_2 \text{ é } A_j^k \text{ ENTÃO } (x_1, x_2) \text{ pertencem a } C_{ij}^k \quad (3)$$

A ferramenta Weka versão 3.5.7 é utilizada para a implementação do método e realização dos testes. Os resultados obtidos para a classificação de um indivíduo podem ser visualizados por meio da Figura 3.3. O resultado apresentado contém informações referentes aos dados do indivíduo, os atributos utilizados, a matriz de confusão, além de outras informações que são discutidas no item 3.6 desta dissertação.

```

=== Run information ===

Scheme: weka.classifiers.rules.SimpleFuzzyGrid -X 10 -S 1 -I 3
Relation: Ana Júlia Bergamo Klaus - Feminino -17 anos
Instances: 21
Attributes: 4
    Código
    Questão
    Resposta
    Inteligência
Test mode: evaluate on training data

=== Classifier model (full training set) ===

Time taken to build model: 0.05 seconds

=== Evaluation on training set ===
=== Summary ===

Correctly Classified Instances      9      42.8571 %
Incorrectly Classified Instances    12      57.1429 %
Kappa statistic                    0.3333
Mean absolute error                 0.1833
Root mean squared error             0.4041
Relative absolute error             88.8887 %
Root relative squared error        115.4701 %
Total Number of Instances          21

=== Detailed Accuracy By Class ===

TP Rate  FP Rate  Precision  Recall  F-Measure  ROC Area  Class
0.333    0.056    0.5        0.333   0.4         0.639    Musical
1        0.389    0.3        1        0.462      0.808    Lógico matemática
0        0        0          0        0          0.5      Interpessoal
0        0        0          0        0          0.5      Espacial visual
0.333    0        1          0.333   0.5         0.667    Corporal sinestésica
0.667    0.111    0.5        0.667   0.571      0.778    Intrapessoal
0.667    0.111    0.5        0.667   0.571      0.778    Linguística

=== Confusion Matrix ===

a b c d e f g <-- classified as
1 1 0 0 0 1 0 | a = Musical
0 3 0 0 0 0 0 | b = Lógico matemática
0 2 0 0 0 0 1 | c = Interpessoal
1 1 0 0 0 0 1 | d = Espacial visual
0 1 0 0 1 1 0 | e = Corporal sinestésica
0 1 0 0 0 2 0 | f = Intrapessoal
0 1 0 0 0 0 2 | g = Linguística

```

Figura 3.3: Resultado de um teste de classificação.

### 3.6 ENTENDENDO O RESULTADO DA CLASSIFICAÇÃO

O resultado obtido após a classificação é expresso por meio de uma matriz de confusão que contém informações sobre as classificações atuais e as previsões realizadas pelo algoritmo de classificação. A acurácia do classificador é comumente avaliada utilizando os dados da matriz que tem seus valores gerados a partir da intersecção dos subconjuntos nebulosos gerando uma matriz dos subconjuntos entre si. Por meio da Figura 3.4, pode ser visualizada a matriz de confusão resultante para a classificação realizada.

```

=== Confusion Matrix ===
 a b c d e f g <-- classified as
 1 1 0 0 1 0 | a = Musical
 0 3 0 0 0 0 | b = Lógico matemática
 0 2 0 0 0 1 | c = Interpessoal
 1 1 0 0 0 1 | d = Espacial visual
 0 1 0 0 1 1 | e = Corporal sinestésica
 0 1 0 0 2 0 | f = Intrapessoal
 0 1 0 0 0 2 | g = Linguística

```

Figura 3.4: Matriz de confusão.

A matriz de confusão oferece uma medida efetiva do modelo de classificação ao mostrar o número de classificações corretas *versus* as classificações preditas, ou seja, que podem ocorrer para cada classe. O número de acertos para cada classe localiza-se na diagonal principal da matriz. Os demais elementos para  $i \neq j$  representam erros na classificação, sendo que  $i$  representa as linhas e  $j$  as colunas da matriz. A matriz de confusão de um classificador ideal possui todos os valores para  $i \neq j$  iguais a zero, uma vez que o classificador não comete erros (DAVIS; GOADRICH, 2006).

A matriz de confusão pode ser lida na horizontal e na vertical. Por meio da Tabela 3.6, podem ser visualizados os dados referentes à leitura realizada na horizontal para a matriz de confusão apresentada na Figura 3.4.

Tabela 3.6: Leitura das linhas da matriz de confusão.

<b>Leitura da matriz de confusão na horizontal</b>			
Classe	Total de itens por linha	Itens corretamente classificados	Itens incorretamente classificados
Musical	3	1	2
Lógica matemática	3	3	0
Interpessoal	3	0	3
Espacial visual	3	0	3
Corporal Cinestésica	3	1	2
Intrapessoal	3	2	1
Linguística	3	2	1

Fonte: Elaborada pelo próprio autor, 2016.

Por meio da Tabela 3.7, podem ser visualizados os valores da leitura da matriz de confusão na vertical para a Figura 3.4.

Tabela 3.7: Leitura das colunas da matriz de confusão.

<b>Leitura da matriz de confusão na vertical</b>			
Classe	Total de itens por coluna	Itens corretamente classificados	Itens incorretamente classificados
Musical	2	1	1
Lógica matemática	10	3	7
Interpessoal	0	0	0
Espacial visual	0	0	0
Corporal Cinestésica	1	1	0
Intrapessoal	4	2	2
Linguística	4	2	2

Fonte: Elaborada pelo próprio autor, 2016.

Pode-se constatar que existem alguns erros na classificação que serão analisados, visando melhorar a precisão do classificador.

Além da matriz de confusão, o classificador apresenta os valores detalhados referentes à acurácia. Estes podem ser visualizados por meio da Figura 3.5.

=== Detailed Accuracy By Class ===						
TP Rate	FP Rate	Precision	Recall	F-Measure	ROC Area	Class
0.333	0.056	0.5	0.333	0.4	0.639	Musical
1	0.389	0.3	1	0.462	0.806	Lógico matemática
0	0	0	0	0	0.5	Interpessoal
0	0	0	0	0	0.5	Espacial visual
0.333	0	1	0.333	0.5	0.667	Corporal sinestésica
0.667	0.111	0.5	0.667	0.571	0.778	Intrapessoal
0.667	0.111	0.5	0.667	0.571	0.778	Linguística

Figura 3.5: Valores de acurácia fornecidos pelo classificador.

As informações disponibilizadas na Figura 3.5 são definidas conforme informações fornecidas a seguir (DAVIS; GOADRIC, 2006):

- TP rate (taxa de verdadeiros positivos) é a proporção de casos positivos que foram corretamente identificados.

$$TP Rate = \frac{\text{verdadeiros positivos}}{\text{total de positivos}}$$

- FP rate (taxa de falsos positivos) é a proporção de casos negativos que foram incorretamente classificados como positivos.

$$FP\ Rate = \frac{falsos\ positivos}{total\ de\ negativos}$$

- Precision (precisão) é a porcentagem de amostras positivas classificadas corretamente sobre o total de amostras classificadas como positivas.

$$Precision = \frac{verdadeiros\ positivos}{verdadeiros\ positivos + falsos\ positivos}$$

- Recall (sensibilidade) é a porcentagem de amostras positivas classificadas corretamente sobre o total de amostras positivas.

$$Recall = \frac{verdadeiros\ positivos}{verdadeiros\ positivos + falsos\ negativos}$$

- F-Measure é a média ponderada da *precision* com o *recall*.

$$F_{Measure} = \frac{2x(Precision \times Recall)}{(Precision + Recall)}$$

- ROC Área é a área da curva ROC (**R**eceiver **O**perating **C**haracteristic Curve ou Curva Característica de Operação do Receptor), os gráficos ROC são uma outra maneira, além de matrizes de confusão, para examinar a acurácia dos classificadores. Um gráfico ROC é um gráfico com a taxa de falsos positivos no eixo X e a taxa verdadeiros positivos no eixo Y. Quanto maior a área da curva maior é a exatidão das respostas para o teste.

Levando-se em consideração todas as informações disponibilizadas pelo classificador após cada classificação realizada, a matriz de confusão será a métrica utilizada para determinar a acurácia do classificador a ser considerada para validar a performance do classificador, mediante os testes realizados.



### 3.7 REALIZANDO O AGRUPAMENTO DAS RESPOSTAS EM CLASSES

Após a aplicação do método de classificação, é necessário realizar o agrupamento das respostas em classes para que se possa determinar o perfil cognitivo do indivíduo e, conseqüentemente, o aspecto cognitivo mais afluído para o indivíduo em questão.

Com relação aos testes realizados, utilizou-se a implementação do cluster *EM* disponível na ferramenta WEKA versão 3.5.7.

O resultado obtido pela utilização da técnica de agrupamento pode ser visualizado por meio da Figura 3.9. Na mesma figura, os percentuais localizados em *clustered instances* definem quais são os aspectos cognitivos predominantes para o indivíduo e ao final da Figura 3.6 são definidos os aspectos cognitivos classificados.

```
=== Run information ===
Scheme:   weka.clusterers.EM -I 100 -N -1 -M 1.0E-6 -S 100
Relation: Ana Júlia Bergamo Klaus - Feminino -17 anos
Instances: 21
Attributes: 4

Clustered Instances

0   7 ( 33%)
1   8 ( 38%)
2   6 ( 29%)

Log likelihood: -7.00137

Class attribute: Inteligência
Classes to Clusters:

0 1 2 <-- assigned to cluster
1 2 0 | Musical
1 1 1 | Lógico matemática
0 2 1 | Interpessoal
1 1 1 | Espacial visual
1 0 2 | Corporal sinestésica
1 2 0 | Intrapessoal
2 0 1 | Linguística

Cluster 0 <-- Linguística
Cluster 1 <-- Musical
Cluster 2 <-- Corporal sinestésica
```

Figura 3.6: Resultado do agrupamento.

### 3.8 ENTENDENDO OS RESULTADOS DO AGRUPAMENTO.

Os resultados obtidos após o processo de agrupamento fornecem várias informações que devem ser consideradas, uma delas é fornecer os grupos de aspectos cognitivos nos quais o indivíduo do teste foi classificado. Outra informação relevante é o percentual que define o quanto o indivíduo classificado pertence a cada grupo de aspecto cognitivo em que ele foi agrupado. Essas informações podem ser visualizadas por meio da Figura 3.7.

<b>Cluster 1 --&gt; Musical --&gt; 38%</b>
<b>Cluster 0 --&gt; Linguística --&gt; 33%</b>
<b>Cluster 2 --&gt; Corporal cinestésica --&gt; 29%</b>

Figura 3.7: Grupos de aspectos cognitivos identificados para um indivíduo.

As informações obtidas no procedimento de agrupamento são analisadas no capítulo 4 que trata dos resultados obtidos. Esse procedimento de agrupamento foi realizado para todos os indivíduos que participaram dos testes. No apêndice D estão disponibilizados os resultados obtidos para o grupo de indivíduos referente aos alunos do curso técnico em manutenção e suporte em informática.

## 4 RESULTADOS OBTIDOS PARA OS PRIMEIROS TESTES

Por meio da Tabela 4.1, pode-se observar uma amostra dos resultados obtidos nos testes realizados. Esta amostra fornece os resultados dos testes para dois dos indivíduos. Na coluna indivíduo, foram informados a idade e o gênero de cada indivíduo e na coluna classificação, por aspecto cognitivo, os aspectos cognitivos predominantes correspondentes a cada indivíduo. Estes dois resultados foram selecionados da tabela completa, contendo os resultados para os doze testes que estão disponibilizados no apêndice D.

Tabela 4.1: Resultado dos testes realizados para a classificação por aspecto cognitivo.

<b>Resultado da classificação dos indivíduos individualmente</b>	
<b>Indivíduo</b>	<b>Classificação por aspecto cognitivo</b>
Indivíduo 4 – gênero masculino – idade 18 anos	38% Interpessoal. 29% Lógico matemático. 28% Corporal Cinestésico. 5% Linguístico.
Indivíduo 10 – gênero masculino – idade 17 anos	38% Interpessoal. 38% Musical. 24% Intrapessoal.

Fonte: Elaborada pelo próprio autor, 2016.

### 4.1 ANÁLISE DOS RESULTADOS OBTIDOS PARA OS PRIMEIROS TESTES

Os resultados obtidos não foram satisfatórios, tendo em vista que, ao analisar os resultados obtidos por meio da matriz de confusão no processo de classificação das respostas, é possível observar que o número de respostas classificadas corretamente é muito baixo em relação ao total de respostas. Por meio da Figura 4.1, pode ser observada a matriz de confusão com a quantidade de respostas classificadas corretamente na diagonal com início no lado superior à esquerda até o lado inferior à direita. Esta diagonal possui sete respostas classificadas corretamente

do total de vinte e uma, totalizando aproximadamente 33,33% das respostas classificadas corretamente.

a	b	c	d	e	f	g	<-- classified as
1	1	0	0	0	1	0	a = Musical
0	1	0	0	1	1	0	b = Lógico matemática
0	1	0	0	0	2	0	c = Interpessoal
1	0	0	0	1	0	1	d = Espacial visual
0	1	0	0	1	1	0	e = Corporal Cinestésica
0	0	0	0	0	3	0	f = Intrapessoal
0	0	0	0	0	2	1	g = Linguística

Figura 4.1: Resultado do processo de classificação com baixo nível de acurácia.

Ao analisar o processo de classificação de cada um dos indivíduos, foi possível observar que o máximo de eficiência com relação a classificação das respostas dos indivíduos é de aproximadamente 33,33% dos pontos classificados corretamente e para alguns casos essa eficiência é ainda menor.

Esses dados sugerem a aplicação de algum procedimento que possa melhorar a acurácia do classificador, aumentando o número de respostas classificadas corretamente.

Os resultados dos testes disponibilizados na Tabela 4.1 foram escolhidos dentre os demais para a realização de testes que têm por objetivo a busca de algum procedimento que possa melhorar a acurácia do classificador. A escolha se deve ao fato curioso do empate entre os dois aspectos cognitivos mais predominantes para o indivíduo de número dez, esse empate gera uma dúvida sobre a eficiência do classificador. A melhoria da eficiência do classificador pode fazer com que esse empate não exista mais e um dos aspectos cognitivos possa ser determinado como sendo o mais afluído para o indivíduo em questão.

O indivíduo de número quatro foi escolhido por apresentar um aspecto cognitivo com baixo índice de predominância, que é de 5%. Com a melhoria da eficiência do classificador, provavelmente, esse valor possa aumentar ou talvez, esse aspecto cognitivo possa desaparecer da lista de aspectos cognitivos mais afluídos para o indivíduo em questão.

Com base em pesquisas bibliográficas, foi possível entender que existe o desbalanceamento das classes durante o processo de mineração dos dados, mais especificamente no treinamento do classificador.

Durante o processo de treinamento do classificador é realizado um ajuste de probabilidade de classificação de cada uma das respostas em função das respostas fornecidas, para cada grupo de aspecto cognitivo. Desta forma, os aspectos cognitivos com respostas com fator de pertinência baixo recebem uma probabilidade de ter respostas classificadas a ele menor que os grupos que têm respostas com fator de pertinência mais alto.

O classificador, ao realizar o treinamento desta forma, gera um desbalanceamento de classes, ou seja, respostas que deveriam ser classificadas corretamente são classificadas em outras classes, alterando de forma significativa o resultado do classificador que retorna o perfil do indivíduo por aspecto cognitivo.

Após esta análise, é proposta a aplicação de um filtro que possa permitir que, durante o processo de treinamento, seja mantido o balanceamento das classes referentes aos grupos de aspectos cognitivos, impedindo que o classificador determine probabilidades diferentes para cada grupo, ou seja, após o processo de treinamento, a probabilidade de cada uma das respostas ser classificada em cada um dos grupos de aspectos cognitivos é a mesma.

Quando se trabalha com classes desbalanceadas, é desejável que se possa determinar a acurácia ao invés da precisão do classificador. Considerando que a maioria dos sistemas de aprendizado é projetada para otimizar a precisão, estes classificadores podem apresentar um baixo nível de acurácia se comparados a outros, isso se o conjunto de treinamento se encontrar fortemente desbalanceado, sendo assim, algumas técnicas são utilizadas para lidar com esse problema, tais como a introdução de custos de classificação incorreta, a remoção de amostras redundantes ou prejudiciais, a detecção de exemplos de borda e com ruído, entre outros.

## 4.2 UTILIZANDO O FILTRO *RESAMPLE*

As técnicas de aprendizado de máquinas empregam um princípio de inferência denominado indução, no qual é possível obter conclusões genéricas a partir de um conjunto particular de exemplos. Essas técnicas de aprendizados indutivos podem ser divididas em dois principais tipos: os supervisionados e os não supervisionados (LORENA; CARVALHO, 2007).

No aprendizado supervisionado, é fornecida uma referência do objetivo a ser alcançado, isto é, um treinamento com o conhecimento do ambiente. Este treinamento é constituído por conjuntos de exemplos como sendo as entradas e uma saída esperada. O algoritmo de aprendizado de máquina extrai a representação do conhecimento a partir desses exemplos. O objetivo é que a representação gerada seja capaz de produzir saídas corretas para novas entradas não apresentadas antes. O erro de um aprendizado supervisionado pode ser calculado como a diferença entre a saída desejada e a saída gerada (LORENA; CARVALHO, 2007).

Diferentemente do aprendizado supervisionado, o não supervisionado não utiliza referências, ou seja, não ocorre um treinamento com o conhecimento do ambiente. O algoritmo de aprendizado de máquina não supervisionado aprende a representar (ou agrupar) as entradas submetidas, segundo medidas de similaridade (LORENA; CARVALHO, 2007).

As técnicas de aprendizado não supervisionado são mais utilizadas quando o entendimento dos dados é feito por meio de padrões ou tendências (SOUTO et al, 2003).

O Filtro *Resample* é um filtro não supervisionado, ou seja, utiliza um algoritmo para determinar quais entradas são similares às demais, para que, posteriormente, as respostas possam ser classificadas, sendo esse um dos motivos que me levou a escolhê-lo.

Outro fator relevante para a escolha do filtro *Resample* é o fato de que ele realiza o tratamento do conhecimento obtido na Mineração dos Dados, sendo responsável por balancear os conjuntos de dados por meio de amostragem com reposição. A utilização desta técnica permite variar a proporção de exemplos entre as classes, que pode ser similar à distribuição inicial ou próxima da distribuição balanceada, e cada valor utilizado para o balanceamento pode ser escolhido mais

de uma vez. Isso ocorre devido a uma das características da amostragem de reposição que possibilita que, após cada etapa de seleção de amostras para o balanceamento, o objeto selecionado retorne a classe podendo ser selecionado novamente.

Para que se entenda de forma mais clara como funciona o balanceamento das classes, é dado o seguinte exemplo:

Considerando que cada aspecto cognitivo possui três questões associadas a ele, o mesmo deve ter três respostas válidas durante a fase de treinamento do classificador para que o balanceamento de classes seja mantido.

Sendo assim, se duas respostas associadas ao aspecto cognitivo forem inválidas, o filtro deverá adotar o valor da resposta válida mais duas vezes para que existam três respostas válidas para o aspecto cognitivo em questão.

Outra situação que pode ocorrer é que exista uma resposta inválida e duas válidas. A resposta válida, que vai ser utilizada para substituir a resposta inválida, será aquela que possui o maior fator de pertinência em relação ao aspecto cognitivo associado a este grupo de três questões.

Caso não exista nenhuma resposta válida, o aspecto cognitivo não fará parte dos aspectos cognitivos pertencentes ao indivíduo. Considerando essas situações o filtro pode ser tendencioso para um ou mais aspectos cognitivos, tornando-os mais ou menos significativos para o perfil do indivíduo.

Levando-se em conta o procedimento aplicado para o balanceamento das classes e a possibilidade do filtro aplicado influenciar no resultados do processo de classificação, houve uma maior preocupação em instruir os indivíduos com relação ao correto preenchimento das respostas no questionário.

### **4.3 REFAZENDO OS TESTES**

O filtro *Resample* foi implementado na fase de treinamento do sistema e todo o processo de classificação e agrupamento foi refeito. Os resultados obtidos para os indivíduos quatro e dez com a utilização do filtro *Resample* podem ser visualizados por meio da Tabela 4.2. Todos os demais resultados para os doze indivíduos pertencentes ao grupo de indivíduos que fazem parte dos alunos do curso técnico em manutenção e suporte em informática estão disponibilizados no apêndice E.

Tabela 4.2: Resultado da classificação por aspecto cognitivo utilizando o filtro *Resample*.

Resultado da classificação dos indivíduos individualmente	
Indivíduo	Classificação por aspecto cognitivo
Indivíduo 4	33% Musical. 24% Intrapessoal. 24% Lógico matemático. 14% Corporal Cinestésico. 5% Linguístico.
Indivíduo 10	57% Musical. 43% Linguístico.

Fonte: Elaborada pelo próprio autor, 2016.

Comparando os novos resultados com os resultados anteriores é possível observar que, com a aplicação do filtro *Resample*, os resultados da classificação e consequentemente do agrupamento entre as classes sofreram alterações significativas, sendo alterados até mesmo os aspectos cognitivos predominantes para os indivíduos. As alterações observadas no resultado da classificação por aspectos cognitivos são devido ao balanceamento das classes que ocorre durante o processo de treinamento quando o filtro *Resample* é utilizado. Sem a utilização do filtro, o percentual de pontos classificados é de aproximadamente 33%; quando o filtro é utilizado, o percentual aumenta para aproximadamente 75%, justificando as alterações nos resultados obtidos. Por meio da Tabela 4.3, podem ser visualizados os resultados dos testes para os indivíduos quatro e dez realizados com e sem a utilização do filtro *Resample*.

Tabela 4.3: Resultado dos testes realizados com e sem o filtro *Resample*.

Resultado da classificação dos indivíduos individualmente		
Indivíduo	Classificação por aspecto cognitivo com a utilização do filtro <i>Resample</i>	Classificação por aspecto cognitivo sem a utilização do filtro <i>Resample</i>
Indivíduo 4	38% Interpessoal. 29% Lógico matemático. 28% Corporal Cinestésico. 5% Linguístico.	33% Musical. 24% Intrapessoal. 24% Lógico matemático. 14% Corporal Cinestésico. 5% Linguístico.
Indivíduo 10	38% Interpessoal. 38% Musical. 24% Intrapessoal.	57% Musical. 43% Linguístico.

Fonte: Elaborada pelo próprio autor, 2016.



## 5 RESULTADOS OBTIDOS

A partir dos resultados obtidos, é possível realizar várias análises. Uma delas é determinar o perfil cognitivo de cada grupo de indivíduos analisado. Por meio da Tabela 5.1, é possível visualizar, na primeira coluna, os sete aspectos cognitivos utilizados para classificar os indivíduos e, em cada linha referente ao aspecto cognitivo, são apresentados os percentuais em que ele está presente nos indivíduos para cada curso.

Tabela 5.1: Classificação dos grupos de indivíduos por aspecto cognitivo.

Aspectos cognitivos	Cursos					
	Manutenção e suporte	Administração	Recursos humanos	Enfermagem	Zootecnia	Ensino médio
Lógico matemático	33,33%	25%	47,82%	43,10%	23,80%	37,14%
Musical	100%	100%	100%	100%	100%	100%
Intrapessoal	58,33%	81,25%	78,26%	63,79%	90,47%	82,85%
Linguístico	83,33%	81,25%	65,21%	46,55%	47,61%	57,14%
Interpessoal	0%	12,5%	13,04%	8,62%	19,04%	8,57%
Espacial visual	0%	18,75%	21,73%	22,41%	23,80%	11,42%
Corporal Cinestésico	25%	31,25%	34,78%	25,86%	28,57%	25,71%

Fonte: Elaborada pelo próprio autor, 2016.

Levando-se em consideração os resultados observados, por meio da Tabela 5.1, é possível determinar que, para todos os cursos técnicos analisados, o aspecto cognitivo musical está presente no perfil cognitivo dos indivíduos. Isso, não quer dizer que ele seja o mais afluído, mas sim, que ele está presente no perfil dos indivíduos analisados e que os perfis cognitivos Interpessoal e Espacial Visual não estão presentes nos indivíduos que pertencem ao grupo Manutenção e Suporte.

Outra análise realizada consiste em determinar qual dos sete aspectos cognitivos é o mais afluído nos indivíduos de cada grupo. Para o curso técnico em manutenção e suporte, tem-se que o aspecto cognitivo musical é o mais afluído em 66,66% dos indivíduos analisados. Para o curso técnico em administração, tem-se que 43,75% dos indivíduos têm o aspecto cognitivo musical mais afluído. Para o

curso técnico em recursos humanos, tem-se que 39,12% dos indivíduos têm os aspectos cognitivos musical e linguístico mais aflorados.

Para o curso técnico em enfermagem, tem-se que 36,21% dos indivíduos têm o aspecto cognitivo intrapessoal mais aflorado. Para o curso técnico em zootecnia, tem-se que 38,1% dos indivíduos têm o aspecto cognitivo intrapessoal mais aflorado e, por fim, para o curso de ensino médio integrado ao curso técnico em informática, tem-se que 39,15% dos indivíduos têm o aspecto cognitivo intrapessoal mais aflorado. Os aspectos cognitivos que estão com percentual igual a zero, não são os mais aflorados em nenhum dos indivíduos analisados. O resultado para esta análise pode ser visualizado por meio da Tabela 5.2.

Tabela 5.2: Aspecto cognitivo mais aflorado para os indivíduos.

Aspectos cognitivos	Cursos					
	Manutenção e suporte	Administração	Recursos humanos	Enfermagem	Zootecnia	Ensino médio
Lógico matemático	0%	6,25%	4,36%	12,07%	19,05%	5,72%
Musical	66,66%	43,75%	38,12%	24,14%	19,05%	30,42%
Intrapessoal	16,67%	12,5%	17,4%	36,21%	38,1%	39,15%
Linguístico	16,67%	31,25%	40,12%	22,42%	19,05%	24,71%
Interpessoal	0%	0%	0%	0%	0%	0%
Espacial visual	0%	0%	0%	0%	0%	0%
Corporal Cinestésico	0%	6,25%	0%	5,16%	4,75%	0%

Fonte: Elaborada pelo próprio autor, 2016.

Uma terceira análise é realizada e consiste em determinar, para as áreas humanas, exatas e biológicas, qual é o aspecto cognitivo mais aflorado. Para essa análise não são considerados os resultados obtidos para os indivíduos classificados no curso de ensino médio integrado ao técnico em informática, uma vez que esse curso em particular está classificado nas três áreas de conhecimento.

Para os resultados observados por meio da Tabela 5.3, é possível determinar que o aspecto cognitivo musical é o mais aflorado para a área de exatas e humanas e o aspecto cognitivo intrapessoal é o mais aflorado para a área de biológicas.

Tabela 5.3: Aspecto cognitivo mais aflorado por área de conhecimento.

Aspectos cognitivos	Áreas de conhecimento		
	Exatas	Humanas	Biológicas
Lógico matemático	0%	5,12%	11,39%
Musical	66,66%	41,05%	22,78%
Intrapessoal	16,67%	15,38%	39,24%
Linguístico	16,67%	35,89%	21,52%
Interpessoal	0%	0%	0%
Espacial visual	0%	0%	0%
Corporal Cinestésico	0%	2,56%	5,07%

Fonte: Elaborada pelo próprio autor, 2016.

Após análise dos resultados obtidos, é possível realizar algumas considerações importantes, tais como: os resultados obtidos fornecem indícios de que é possível utilizar ferramentas computacionais em sistemas de recomendação com o objetivo de classificar indivíduos de acordo com seus aspectos cognitivos e utilizar os resultados obtidos na classificação para diversos fins, como por exemplo, indicar atividades voltadas para o ensino por meio da utilização dos aspectos cognitivos mais aflorados.

Os primeiros testes realizados para o grupo de indivíduos do curso técnico em manutenção e suporte em informática, sem a utilização e com a utilização do filtro *Resample*, indicam a necessidade da utilização de um filtro durante o processo de treinamento do classificador, sendo que a utilização deste filtro elevou a quantidade de respostas classificadas corretamente pelo classificador de aproximadamente 35% para o mínimo de aproximadamente 75%. Porém, a aplicação do filtro está relacionada aos testes realizados nesta dissertação.

## 5.1. VALIDAÇÃO DOS RESULTADOS

Com o objetivo de validar os resultados obtidos no classificador, foi realizada análise de desempenho escolar dos indivíduos pertencentes ao curso Técnico em Manutenção e Suporte em Informática, mais especificamente na disciplina de lógica de programação. Essa análise foi comparada com os resultados correspondentes à turma em questão, obtidos pelo classificador e já apresentados nesta dissertação.

A metodologia utilizada nas aulas da disciplina de Lógica de Programação consistiu em apresentar esta e realizar abordagem inicial do conteúdo por meio de aulas teóricas com a utilização do quadro branco e textos conceituais. Em seguida, foram utilizados fluxogramas para o entendimento de estruturas de decisão e repetição em algoritmos. Por fim, foram realizados exercícios referentes ao desenvolvimento de códigos de linguagem de programação C++, com o objetivo de possibilitar aos indivíduos desenvolver códigos utilizando essa linguagem, de acordo com regras lógicas.

Na Tabela 5.4, pode ser visualizado o comparativo do desempenho escolar dos indivíduos com relação ao perfil cognitivo de cada um deles obtido pelo classificador. Esse desempenho está descrito em função das notas de prova obtidas na disciplina de Lógica de Programação, estando essas associadas ao método de aula utilizado para transmitir o conteúdo e o aspecto cognitivo relacionado ao método utilizado.

Para essas comparações, consideramos que durante aulas teóricas, essencialmente expositivas, o aspecto cognitivo estimulado foi o linguístico; na fase em que usamos fluxogramas para explanação de detalhes dos algoritmos, consideramos que o aspecto espacial-visual era o principal estimulado; e, finalmente, na fase de desenvolvimento de código, por requerer dos alunos o emprego de regras lógicas, consideramos que o principal aspecto cognitivo estimulado foi o lógico-matemático.

Associando-se a terceira e quarta colunas da Tabela 5.4, é possível se ter uma ideia do quanto cada aspecto cognitivo é adequado a cada aluno. Por exemplo, supondo-se um máximo de 30 pontos, é possível dizer que para o aluno 1 aproximadamente 27% ( $8/30$ ) de seu perfil cognitivo é lógico-matemático, 22% ( $6,5/30$ ) é espacial visual e 20% ( $6/30$ ) é linguístico; os 31% restantes seriam

distribuídos entre os outros cinco perfis possíveis. Se, por outro lado, adotarmos a conclusão do classificador, isto é, que só os perfis manifestados estão presentes de forma perceptível, o denominador acima deveria ser 20,5 e as percentagens seriam, respectivamente, 39%, 31,7% e 29,3%.

Contudo, para a preparação das aulas, não foram considerados todos os aspectos cognitivos possíveis, razão pela qual utilizamos nesta comparação a primeira quantificação, qual seja, aquela que admite um percentual remanescente para os perfis eventualmente não considerados na preparação das aulas.

O comparativo dos dados apresentados na coluna “Percentual da metodologia” com os apresentados na coluna “Perfis cognitivos dos indivíduos identificados com o classificador” possibilita determinar se os percentuais dos aspectos cognitivos presentes em cada um dos indivíduos correspondem, aproximadamente, àqueles identificados pelo classificador.

Tabela 5.4: Análise do desempenho escolar dos indivíduos

Indivíduo	Metodologia	Aspecto cognitivo utilizado na metodologia	Nota de prova	Percentual da metodologia	Perfil cognitivo do indivíduo segundo o classificador
1	Aula teórica	Linguístico	6,5	22%	38% Intrapessoal. 33% Musical. 29% Lógico matemático.
	Fluxogramas	Espacial visual	6,0	20%	
	Desenvolvimento de código	Lógico matemático	8,0	27%	
2	Aula teórica	Linguístico	8,0	27%	67% Musical. 33% Linguístico.
	Fluxogramas	Espacial visual	5,0	17%	
	Desenvolvimento de código	Lógico matemático	7,0	23%	
3	Aula teórica	Linguístico	8,5	28%	48% Linguístico. 33% Intrapessoal. 19% Musical.
	Fluxogramas	Espacial visual	6,0	20%	
	Desenvolvimento de código	Lógico matemático	7,0	23%	
4	Aula teórica	Linguístico	6,0	20%	33% Musical. 24% Intrapessoal. 24% Lógico matemático. 14% Corporal Cinestésico. 5% Linguístico.
	Fluxogramas	Espacial visual	7,5	25%	
	Desenvolvimento de código	Lógico matemático	6,0	20%	
5	Aula teórica	Linguístico	7,5	25%	33% Intrapessoal. 29% Linguístico. 28% Musical. 10% Corporal Cinestésico.
	Fluxogramas	Espacial visual	8,0	27%	
	Desenvolvimento de código	Lógico matemático	5,0	17%	
6	Aula teórica	Linguístico	9,0	30%	52% Musical. 48% Linguístico.
	Fluxogramas	Espacial visual	8,0	27%	
	Desenvolvimento de código	Lógico matemático	8,5	28%	
7	Aula teórica	Linguístico	5,0	17%	43% Linguístico. 33% Musical. 24% Intrapessoal.
	Fluxogramas	Espacial visual	5,5	18%	
	Desenvolvimento de código	Lógico matemático	8,0	27%	
8	Aula teórica	Linguístico	6,5	22%	38% Musical. 24% Lógico matemático. 19% Intrapessoal. 19% Linguístico.
	Fluxogramas	Espacial visual	7,0	23%	
	Desenvolvimento de código	Lógico matemático	8,0	27%	
9	Aula teórica	Linguístico	8,0	27%	71% Musical.

	Fluxogramas	Espacial visual	9,0	30%	29% Lógico matemático.
	Desenvolvimento de código	Lógico matemático	5,5	18%	
10	Aula teórica	Linguístico	9,0	30%	57% Musical. 43% Linguístico.
	Fluxogramas	Espacial visual	8,0	27%	
	Desenvolvimento de código	Lógico matemático	7,5	25%	
11	Aula teórica	Linguístico	7,5	25%	57% Musical. 43% Linguístico.
	Fluxogramas	Espacial visual	7,0	23%	
	Desenvolvimento de código	Lógico matemático	7,0	23%	
12	Aula teórica	Linguístico	8,0	27%	33% Musical. 24% Corporal Cinestésico. 24% Intrapessoal. 19% Linguístico.
	Fluxogramas	Espacial visual	9,0	30%	
	Desenvolvimento de código	Lógico matemático	7,5	25%	

Fonte: Elaborada pelo próprio autor, 2016.

Como forma de exemplificar esse comparativo, tomemos como referência o indivíduo 1. Para as aulas teóricas, foi utilizada uma metodologia que faz uso do aspecto cognitivo linguístico. O indivíduo obteve a nota de prova 6.5, que está próxima a nota mediana 5.0, e o perfil cognitivo determinado pelo classificador não contempla o aspecto cognitivo linguístico.

Para as aulas referentes ao entendimento e desenvolvimento de fluxogramas, foi utilizada uma metodologia que faz uso do aspecto cognitivo espacial visual. O indivíduo obteve a nota de prova 6.0 que também está próxima a nota mediana 5.0 e o perfil cognitivo determinado pelo classificador não contempla o aspecto cognitivo espacial visual.

Para as aulas relacionadas ao desenvolvimento de códigos, foi utilizada uma metodologia que faz uso do aspecto cognitivo lógico matemático. O indivíduo obteve a nota de prova 8.0, que está acima da média 5.0, bem como das notas 6.5 e 6.0 obtidas no uso de outras metodologias. A melhora na nota de prova pode ser justificada pelo uso da metodologia que faz parte do aspecto cognitivo lógico matemático que está contido no perfil cognitivo do indivíduo de acordo com o classificador.

Dessa forma, é possível afirmar sobre o indivíduo 1 que seu aspecto cognitivo lógico-matemático é (27%) aproximadamente igual ao previsto pelo classificador (29%). Também, é possível afirmar que parte de seu insucesso nas outras duas provas se deve à exploração das características linguística e espacial-visual, quando seus perfis cognitivos mais afluídos seriam o intrapessoal e o musical, segundo identificado pelo classificador.

Análise semelhante pode ser feita para os indivíduos 2 e 3, que tiveram suas melhores notas nas provas em que o aspecto linguístico foi explorado. Ambos foram classificados, respectivamente, com 33% e 48% de cognição ligada ao aspecto linguístico. Contudo, nenhum foi classificado como apresentando os aspectos cognitivos espacial-visual ou lógico-matemático, razão pela qual suas notas não foram muito boas quando esses aspectos foram explorados em sala de aula.

O indivíduo 4 é um caso interessante. Seu aspecto cognitivo espacial-visual é razoavelmente bem presente, mas o classificador sequer listou esse aspecto cognitivo no seu perfil. Esse comportamento do classificador pode ser devido a uma deficiência no ajuste dos parâmetros ou a certa tendenciosidade nas respostas ao questionário.

É importante ressaltar que existe uma margem de erro no processo de classificação e que dentro dessa margem de erro podem estar os indivíduos 4, 7, 9 e 12, que não obtiveram as melhores notas durante a aplicação das provas, mesmo estando associadas a um dos aspectos cognitivos predominantes para o indivíduo, segundo os resultados do classificador.

A análise dos dados, disponibilizados na Tabela 5.4, possibilita concluir que os indivíduos 1, 2, 3, 5, 6, 8, 10 e 11 (67% do total de alunos considerados para esta validação) obtiveram as melhores notas nas provas referentes aos métodos de ensino que fazem uso de aspectos cognitivos que estão presentes no perfil cognitivo de cada um deles, isso de acordo com perfil determinado pelo classificador.

## 6 CONCLUSÃO

O classificador proposto pode vir a ser utilizado por ferramentas de apoio ao ensino a distância com o objetivo de permitir que tutores responsáveis por gerar conteúdo para as aulas possam utilizar um sistema de recomendação para classificar alunos e direcioná-los para conteúdos que possam explorar seus aspectos cognitivos predominantes, na tentativa de potencializar o ensino, permitindo que o aluno possa realizar os estudos por meio de técnicas que poderão facilitar o aprendizado de forma individualizada.

A integração deste classificador em sistemas de recomendação pode contribuir para validar a proposta desta dissertação, porém, para essa integração, será necessário melhorar os métodos de treinamento do classificador devido ao aumento do número de amostras, bem como ajustar o formulário das questões a serem respondidas pelos indivíduos e realizar mais testes e simulações para que uma análise mais detalhada possa ser realizada e conclusões mais significativas possam ser obtidas.

A aplicação da lógica nebulosa como método computacional empregado no classificador permitiu classificar as respostas com grau de incerteza e de forma semelhante ao raciocínio humano, atribuindo fatores de pertinência para cada resposta de cada uma das vinte e uma questões do questionário. Isso possibilitou avaliar o quanto a resposta fornecida pelo indivíduo o relaciona ao aspecto cognitivo que está sendo avaliado. Um dos objetivos foi justamente trabalhar com respostas que possuem grau de incerteza, possibilitando aos indivíduos mensurarem o quanto eles concordam ou não com cada questão. A aplicação dos conceitos de lógica nebulosa atendeu a essa necessidade.

A utilização de um filtro no processo de treinamento do classificador aumentou a precisão das respostas do classificador e, conseqüentemente, mudou o resultado do agrupamento. Isso foi constatado após os primeiros testes realizados, por meio de análise da matriz de confusão dos resultados do classificador, sendo que este, sem a utilização do filtro *Resample* apresentou em média 35% das respostas classificadas corretamente e, com a utilização do filtro, este índice aumentou para aproximadamente 75% das respostas classificadas corretamente. Desta forma, a utilização do filtro no classificador tornou-se necessária, porém, para



outras aplicações, a necessidade de se utilizar esse filtro, ou qualquer outro, deve ser analisada para que possa ser validada a necessidade de utilização, tendo em vista que existem modelos de classificadores que trabalham com desbalanceamento de classes para que se possa obter resultados satisfatórios.

Os resultados obtidos nos testes realizados remetem à conclusão de que é possível classificar indivíduos por aspectos cognitivos, relacionando estes aos conceitos da Teoria das Inteligências Múltiplas propostas por Gardner por meio da utilização de classificadores que simulem as decisões humanas. Seguindo o modelo de análise proposto por Gardner, pode-se desenvolver um questionário contendo perguntas relacionadas ao contexto em que o indivíduo está inserido e fornecer a ele opções de respostas com certo grau de incerteza tornando possível determinar o aspecto cognitivo mais afluído no indivíduo.

Porém, a aplicação de um questionário para determinar aspectos cognitivos afluídos em indivíduos possui ressalvas. O questionário deve ser elaborado de acordo com vários fatores. A faixa etária dos indivíduos que se pretende analisar, o perfil de cada grupo de indivíduos dentro desta faixa etária, os costumes familiares, crenças e perspectivas de ações futuras, atuação profissional, convívio na sociedade, entre outros devem ser levados em consideração, tornando bem complexo o desenvolvimento de um questionário padrão que possa ser aplicado a todos os indivíduos independente dos fatores apresentados.

O questionário utilizado foi elaborado levando em consideração vários aspectos com relação ao perfil dos indivíduos a serem analisados, e de acordo com conversa informal após alguns dos indivíduos responderem ao questionário, não houve dificuldades em respondê-lo, fornecendo indícios de que o mesmo atendeu ao objetivo de possibilitar a coleta de dados com respostas consistentes.

O resultado do processo de classificação dos indivíduos por aspectos cognitivos realizado nesta dissertação possibilitou várias análises, tais como, determinar o perfil cognitivo dos indivíduos obtendo como resposta os aspectos cognitivos mais afluídos para cada indivíduo, determinar o perfil cognitivo de grupos de indivíduos e determinar qual o perfil cognitivo mais afluído para indivíduos de cada grupo, além de outras análises que podem ser realizadas a partir dos resultados obtidos.

Por fim, são sugeridas algumas propostas para trabalhos futuros. O processamento das informações utilizadas pelo classificador deve seguir um padrão de computação que possa se aproximar do raciocínio humano, assim, acredito que utilizar regras nebulosas “se”...”então” seja uma boa proposta, porém a confirmação pode vir da comparação deste método com outros métodos sendo estes aplicados a mesma ideia abordada nesta dissertação.

Foi utilizado o filtro *Resample* para realizar o balanceamento das classes que representam os aspectos cognitivos propostos como parâmetros de classificação. Esse filtro é não supervisionado, ou seja, utiliza um algoritmo para determinar quais entradas são similares às demais, para que, posteriormente, as respostas possam ser classificadas, sendo este um dos motivos que me levou a escolhê-lo, porém, existem outros filtros com essa característica. Fica como proposta de possíveis trabalhos futuros, a realização da análise do classificador utilizando outros filtros não supervisionados, ou até mesmo supervisionados, para que se possa ter um comparativo da eficiência de cada filtro para este problema em específico.

O conhecimento gerado por este classificador é fundamentado em aspectos cognitivos, porém, outros tipos de conhecimento podem ser gerados. Uma proposta de trabalho futuro consiste em ajustar o classificador para que se possa classificar indivíduos por meio de ontologias, para isso, será necessário ajustar o formulário e os parâmetros utilizados na classificação.

A partir do classificador apresentado nesta dissertação, podem ser desenvolvidos outros classificadores que podem, por exemplo, classificar informações provenientes de redes sociais, com o objetivo de buscar padrões de perfis de indivíduos por meio de mensagens publicadas nestas redes, ao invés de se utilizar questionários como entrada de dados.

## REFERÊNCIAS

ADNAN, Nuruddin M. et al. "Content based news recommendation system based on fuzzy logic". In: **International Conference on Informatics, Electronics & Vision (ICIEV)**, 3., 2014, Dhaka. Bangladesh. ISBN 978-1-4799-5179-6. May, 2014.

AHMAD, Sadique; ADNAN, Awais. "Machine learning based cognitive skills calculations for different emotional conditions". In: **Cognitive Informatics & Cognitive Computing (ICCI\*CC)**, 2015 IEEE 14th International Conference, pp.162-168, 6-8 July, 2015.

ARMSTRONG, Thomas. *Inteligências Múltiplas na Sala de Aula*. ARTMED, Porto Alegre, 2001. 192p.

BRESSLER, Fábio. *Um Protótipo de Aplicação para Recomendação de Produtos Baseado no Interesse e Comportamento de uso do Usuário*. São Leopoldo, 2004. Monografia do Curso de Informática. Universidade do Vale do Rio dos Sinos.

BRIN, Sergey; PAGE, Lawrence. **The anatomy of a large-scale hypertextual websearch engine**. *Computer Network ISDN Systems*, 30(1-7):107-117. 1998.

COX, Earl. *The Fuzzy systems handbook: a practitioner's guide to building, using, and maintaining Fuzzy systems*. New York: AP Professional, 1994.

DAVIS, Jesse; GOADRICH, Mark. - The Relationship between Precision-Recall and ROC Curves. *Proc. 23rd Int. Conf. On Machine Learning*. 2006.

DEMPSTER, Arthur P.; LAIRD, Nan M.; RUBIN, Donald B. Maximum likelihood from incomplete data via the em algorithm. *Journal of the Royal Statistical Society, Series B*, 39(1):1–38. 1997.

FREEDMAN, Rebecca. Enhanced Possibilities for Teaching and Learning: A Whole School Approach to Incorporating Multiple Intelligences and Differentiated Instruction. Dissertação Mestrado – University of Toronto Libraries, May 2015.

GARDNER, Howard. Frames of mind: The theory of multiple intelligences. Chicago: Basic Books, 3ª Edition, March 29, 2011.

GARDNER, Howard. Inteligências Múltiplas: A Teoria na Prática. Artes Médicas. Porto Alegre, 1995.

GIL, Antonio C. Métodos e técnicas de pesquisa social. 5. ed. São Paulo: Atlas, 1999.

GUTTMAN, Robert H.; MOUKAS, Alexandros G.; MAES, Pattie. “Agent-mediated electronic commerce: a survey.” **Journal: The Knowledge Engineering Review**. Volume13 Issue 2, p147–149. July, 1948.

HAN, Jiawei; KAMBER, Micheline. **Data Mining. Concepts and Techniques**. Second edition. The Morgan Kaufmann Series in Data Management Systems. Elsevier Inc., 2006.

ISHIBUCHI, Hisao; NOZAKI, Ken; TANAKA, Hideo. Efficient Fuzzy partition of pattern space for classification problems, vol. 59, pp. 295-304, Nov. 1993.

KAJIMOTO, A. P. et al. Sistemas de Recomendação de Notícias na Internet Baseados em Filtragem Colaborativa. São Paulo, 2007. Trabalho de Formatura Supervisionado. Universidade de São Paulo.

KUNJAL, Mankad Bharatkumar. Article The Role of Multiple Intelligence in E-Learning. **IJSRD - International Journal for Scientific Research & Development**. Vol. 3, Issue 05, 2015. ISSN (online): 2321-0613.

LEE, C.C. Fuzzy Logic in Control Systems: Fuzzy Logic Controller, part I and II. *IEEE Trans. On Systems, Man and Cybernetics*, vol. 20, pp 404-435. 1990.

LORENA, A. C.; CARVALHO, A. C. P. L. F. Uma Introdução às Support Vector Machines. *Revista de Informática Teórica e Aplicada*, vol.14, no2, pp 43-67, 2007.

LUGER, George F. Artificial Intelligence: Structures and strategies for complex problem solving – 5th ed. Pearson Education/Addison Wesley. 2005.

MARK, Hall, et al. The WEKA Data Mining Software: An Update; *SIGKDD Explorations*, Volume 11, Issue 1. 2009.

MONTANER, M.; LÓPEZ, B.; DE LA ROSA, J. L. A Taxonomy of Recommender Agents on the Internet. *Artificial Intelligence Review*. Netherlands : Kluwer Academic Publishers, pp. 285-330. Aug, 2003.

MUNAKATA, Toshinori. Fundamentals of the new Artificial Intelligence: neural, evolutionary, fuzzy and more – 2nd ed. Springer-Verlag London. 2008.

PEDRYCZ, W. Fuzzy Control and Fuzzy Systems. - John Wiley and Sons Inc, NY. 1989.

REATEGUI, E. B.; CAZELLA, S. C. Sistemas de Recomendação. **XXV Congresso da Sociedade Brasileira de Computação. Universidade do Vale do Rio dos Sinos (UNISINOS)**. São Leopoldo, 2005.

RESNICK, P.; VARIAN, H. R. Recommender Systems. *Communications of the ACM*, New York, v.40, n.3, pp. 55-58, Mar. 1997.

RICCI, Francesco, et al. Recommender Systems Handbook. Springer Science Business Media, New York, 2011.

SHENG, V. et al. “Cost-Sensitive Test Strategies”. In: **Proceedings of the 21st National Conference on Artificial Intelligence**, 2006.

SILVA, I. A. F. Descoberta de Conhecimento em Base de Dados de Monitoramento Ambiental para Avaliação da Qualidade da Água. Programa de Pós-Graduação em Física e Meio Ambiente. Universidade Federal de Mato Grosso, Cuiabá, 2007.

SOUTO, M. C. P. et al. Técnicas de Aprendizado de Máquina para problemas de Biologia Molecular. P.103–152. Minicursos de Inteligência Artificial, **Jornada de Atualização Científica em Inteligência Artificial, XXIII Congresso da Sociedade Brasileira de Computação**, 2003.

VON ALTROCK, Constantin. Fuzzy logic and neuroFuzzy applications in business and finance. New Jersey: Prentice Hall PTR, 1996.

WEBER, L.; KLEIN, P. A. T.; Aplicação da Lógica Fuzzy em Software e Hardware. Editora ULBRA. 1ª Edição, 2003.

ZADEH, L. **Fuzzy Sets - Information and Control**, vol. 8, pp 338-353. 1965.

ZADROZNY, B.; LANGFORD, J.; ABE N. “Cost-Sensitive Learning by Cost-Proportionate Example Weighting”. In: **Proceedings of the 3rd IEEE International Conference on Data Mining (ICDM'03)**, 2003.

# APÊNDICE A

Modelo de questionário utilizado como referência para a elaboração do questionário utilizado nos testes.

Multiple Intelligences Test - based on Howard Gardner's MI Model		<a href="http://businessballs.com">more info at businessballs.com</a>				
(manual version - see businessballs.com for self-calculating version)						
Score the statements: 1 = Mostly Disagree, 2 = Slightly Disagree, 3 = Slightly Agree, 4 = Mostly Agree						
Alternatively for speed, and if easier for young people - tick the box if the statement is more true for you than not.						
Adults over 16 complete all questions. Young people between 8-16 answer red questions only. This is page 1 of 2.						
A short version featuring the young people's questions only is available free from the businessballs website.						
Score or tick the statements in the white-out boxes only	Score					
I like to learn more about myself	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	1
I can play a musical instrument	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	2
I find it easiest to solve problems when I am doing something physical	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	3
I often have a song or piece of music in my head	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	4
I find budgeting and managing my money easy	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	5
I find it easy to make up stories	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	6
I have always been physically well co-ordinated	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	7
When talking to someone, I tend to listen to the words they use not just what they mean	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	8
I enjoy crosswords, word searches or other word puzzles	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	9
I don't like ambiguity, I like things to be clear	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	10
I enjoy logic puzzles such as 'sudoku'	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	11
I like to meditate	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	12
Music is very important to me	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	13
I am a convincing liar	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	14
I play a sport or dance	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	15
I am very interested in psychometrics (personality testing) and IQ tests	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	16
People behaving irrationally annoy me	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	17
I find that the music that appeals to me is often based on how I feel emotionally	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	18
I am a very social person and like being with other people	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	19
I like to be systematic and thorough	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	20
I find graphs and charts easy to understand	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	21
I can throw things well - darts, skimming pebbles, frisbees, etc	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	22
I find it easy to remember quotes or phrases	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	23
I can always recognise places that I have been before, even when I was very young	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	24
I enjoy a wide variety of musical styles	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	25
When I am concentrating I tend to doodle	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	26
I could manipulate people if I choose to	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	27
I can predict my feelings and behaviours in certain situations fairly accurately	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	28
I find mental arithmetic easy	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	29
I can identify most sounds without seeing what causes them	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	30
At school one of my favourite subjects is / was English	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	31
I like to think through a problem carefully, considering all the consequences	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	32
I enjoy debates and discussions	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	33
I love adrenaline sports and scary rides	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	34
I enjoy individual sports best	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	35
I care about how those around me feel	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	36
My house is full of pictures and photographs	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	37
I enjoy and am good at making things - I'm good with my hands	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	38
I like having music on in the background	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	39
I find it easy to remember telephone numbers	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	40

I set myself goals and plans for the future																	41
I am a very tactile person																	42
I can tell easily whether someone likes me or dislikes me																	43
I can easily imagine how an object would look from another perspective																	44
I never use instructions for flat-pack furniture																	45
I find it easy to talk to new people																	46
To learn something new, I need to just get on and try it																	47
I often see clear images when I close my eyes																	48
I don't use my fingers when I count																	49
I often talk to myself – out loud or in my head																	50
At school I loved / love music lessons																	51
When I am abroad, I find it easy to pick up the basics of another language																	52
I find ball games easy and enjoyable																	53
My favourite subject at school is / was maths																	54
I always know how I am feeling																	55
I am realistic about my strengths and weaknesses																	56
I keep a diary																	57
I am very aware of other people's body language																	58
My favourite subject at school was / is art																	59
I find pleasure in reading																	60
I can read a map easily																	61
It upsets me to see someone cry and not be able to help																	62
I am good at solving disputes between others																	63
I have always dreamed of being a musician or singer																	64
I prefer team sports																	65
Singing makes me feel happy																	66
I never get lost when I am on my own in a new place																	67
If I am learning how to do something, I like to see drawings and diagrams of how it works																	68
I am happy spending time alone																	69
My friends always come to me for emotional support and advice																	70

<p><b>Add the scores or ticks in each column and write the total for each column in the boxes on the right.</b></p> <p><b>Your highest scores indicate your natural strengths and potential - your natural intelligences.</b></p> <p><b>There are no right or wrong answers.</b></p> <p><b>My strongest intelligences are (write them here):</b></p>	<b>Intelligence type</b>	<b>your totals</b>														
	Linguistic															
	Logical-Mathematical															
	Musical															
	Bodily-Kinesthetic															
	Spatial-Visual															
	Interpersonal															
	Intrapersonal															

We are **happiest** and **most successful** when we **learn, develop, and work** in ways that make **best use** of our **natural intelligences** (our strengths and style and brain-type in other words).

As such this indicator helps you to focus on the sorts of learning and work that will be most fulfilling and rewarding for you.

See the multiple intelligence types definitions on sheet 2 of this file - or if you have only hard-copy, go to [www.businessballs.com](http://www.businessballs.com) for details about multiple intelligences and/or a free file version of this tool.

© V Chislett MSc and A Chapman 2005-06, based on Gardner's Multiple Intelligences Model. Available free from [www.businessballs.com](http://www.businessballs.com). Not to be sold or published. The authors accept no liability.

[more info at businessballs.com](http://www.businessballs.com)



## APÊNDICE B

Modelo do questionário utilizado nesta dissertação

Código	Questão	Resposta	Aspecto cognitivo
1	Faça uma análise sobre as músicas que você costuma ouvir, associe o ritmo musical ao seu estado de espírito. As músicas que você ouve representam o seu estado de espírito no momento em que você as ouve?		Musical
2	Quando você precisa resolver problemas que envolvem cálculos matemáticos, você prefere tentar resolver mentalmente ou procura alguma ferramenta de cálculo que facilite seu trabalho com o objetivo de ganhar tempo? Responda em relação à frequência de calcular mentalmente.		Lógico matemático
3	Quando você percebe que alguma pessoa da família, ou um amigo mais próximo ou até mesmo um desconhecido não está bem, estando ele com aparente tristeza, com que frequência você procura saber o que está acontecendo e se dispõe a ajudar?		Interpessoal
4	Quando você vai ler revistas, livros ou até mesmo visitar sites na Internet, com que frequência você olha primeiro as imagens e depois lê o texto?		Espacial visual
5	Ao estudar, ler ou realizar alguma atividade em casa ou no trabalho, com que frequência você faz esta atividade ouvindo alguma música de fundo ou mesmo cantando alguma música em voz baixa ou mentalmente não importando o ritmo da música?		Musical
6	Ao consertar ou montar equipamentos e objetos, com que frequência você prefere tentar montar de acordo com a forma que considera ser a melhor deixando de lado o manual de instruções?		Corporal Cinestésico
7	Se você tivesse que anotar tudo que se passa na sua vida em um diário para que um dia possa se lembrar de tudo que já fez e deixar anotações para outras pessoas, com que frequência você anotaria as coisas?		Intrapessoal
8	Se você fosse um atleta profissional e tivesse que escolher um esporte para praticar, com que frequência você iria escolher esportes coletivos pois, teria mais chances de ser reconhecido, podendo ser beneficiado da ajuda dos outros atletas?		Interpessoal
9	Quando você está na sala de aula e o professor esta explicando a matéria utilizando gestos e palavras, se o professor gesticular demais, com que frequência isso acaba distraindo você?		Espacial visual
10	Quando você não está bem emocionalmente e alguma coisa está incomodando você, com que frequência você prefere estar junto a várias pessoas independente do que estejam conversando, mesmo que seja sobre você?		Intrapessoal
11	Quando o professor esta explicando um conteúdo novo e utiliza a fundamentação teórica, com que frequência isso é o suficiente para que você possa aprender e resolver os exercícios propostos ao final da aula?		Corporal Cinestésico
12	Você costuma guardar seu dinheiro e ir gastando ele aos poucos comprando somente o necessário e o que sobra você guarda para uma emergência?		Lógico matemático
13	Quando você vai realizar alguma tarefa sozinho(a) você costuma ser detalhista mesmo que isso possa fazer você perder muito tempo em detalhes pequenos?		Lógico matemático
14	Para se manter atualizado com relação as notícias do cotidiano, com que frequência você busca notícias escritas em jornais, revistas e na Internet?		Linguístico
15	Quando você está ouvindo música e percebe os sons gerados por instrumentos musicais, com que frequência você sente vontade de tocar algum instrumento musical, mesmo que não seja o instrumento utilizado para tocar a música que você está ouvindo?		Musical
16	Quando você não está se sentindo bem emocionalmente, com que frequência você comenta seus problemas para algum amigo(a) e permite que ele dê conselhos a você?		Interpessoal

<b>Código</b>	<b>Questão</b>	<b>Resposta</b>	<b>Aspecto cognitivo</b>
17	Quando você está em ambientes com várias pessoas e os assuntos vão surgindo, você gosta de debater mesmo que não tenham perguntado sua opinião, considerando assim, que sua opinião é importante?		Linguístico
18	Quando você viaja para lugares diferentes que você ainda não conhece, com que frequência você repara em detalhes nos formatos de árvores, cores e formas geométricas das casas, modelos dos carros, entre outras coisas?		Espacial visual
19	Com que frequência você se oferece como voluntário(a) para realizar tarefas manuais, tais como utilizar chaves de fenda, cortar arames com alicate, pregar pregos, cortar papéis, lavar a louça, passar roupa, entre outras?		Corporal Cinestésico
20	Com que frequência você prefere realizar tarefas que deixam você feliz independente da tarefa que você vai realizar, mesmo sabendo que existem outras tarefas que você pode realizar que não façam mal a você, mas também são necessárias?		Intrapessoal
21	Suponha que você esteja participando de um concurso musical e você tem facilidade em cantar músicas nacionais e em outro idioma independentemente do idioma. Com que frequência você faria a opção por cantar músicas internacionais?		Linguístico

## APÊNDICE C

### Exemplo de questionário respondido

Código	Questão	Resposta
1	Faça uma análise sobre as músicas que você costuma ouvir, associe o ritmo musical ao seu estado de espírito. As músicas que você ouve representam o seu estado de espírito no momento em que você as ouve?	4
2	Quando você precisa resolver problemas que envolvem cálculos matemáticos, você prefere tentar resolver mentalmente ou procura alguma ferramenta de cálculo que facilite seu trabalho com o objetivo de ganhar tempo? Responda em relação à frequência de calcular mentalmente.	3
3	Quando você percebe que alguma pessoa da família, ou um amigo mais próximo ou até mesmo um desconhecido não está bem, estando ele com aparente tristeza, com que frequência você procura saber o que está acontecendo e se dispõe a ajudar?	3
4	Quando você vai ler revistas, livros ou até mesmo visitar sites na Internet, com que frequência você olha primeiro as imagens e depois lê o texto?	4
5	Ao estudar, ler ou realizar alguma atividade em casa ou no trabalho, com que frequência você faz esta atividade ouvindo alguma música de fundo ou mesmo cantando alguma música em voz baixa ou mentalmente não importando o ritmo da música?	3
6	Ao consertar ou montar equipamentos e objetos, com que frequência você prefere tentar montar de acordo com a forma que considera ser a melhor deixando de lado o manual de instruções?	1
7	Se você tivesse que anotar tudo que se passa na sua vida em um diário para que um dia possa se lembrar de tudo que já fez e deixar anotações para outras pessoas, com que frequência você anotaria as coisas?	4
8	Se você fosse um atleta profissional e tivesse que escolher um esporte para praticar, com que frequência você iria escolher esportes coletivos pois, teria mais chances de ser reconhecido, podendo ser beneficiado da ajuda dos outros atletas?	3
9	Quando você está na sala de aula e o professor esta explicando a matéria utilizando gestos e palavras, se o professor gesticular demais, com que frequência isso acaba distraindo você?	2
10	Quando você não está bem emocionalmente e alguma coisa está incomodando você, com que frequência você prefere estar junto a várias pessoas independente do que estejam conversando, mesmo que seja sobre você?	3
11	Quando o professor esta explicando um conteúdo novo e utiliza a fundamentação teórica, com que frequência isso é o suficiente para que você possa aprender e resolver os exercícios propostos ao final da aula?	2
12	Você costuma guardar seu dinheiro e ir gastando ele aos poucos comprando somente o necessário e o que sobra você guarda para uma emergência?	2
13	Quando você vai realizar alguma tarefa sozinho(a) você costuma ser detalhista mesmo que isso possa fazer você perder muito tempo em detalhes pequenos?	3
14	Para se manter atualizado com relação as notícias do cotidiano, com que frequência você busca notícias escritas em jornais, revistas e na Internet?	3
15	Quando você está ouvindo música e percebe os sons gerados por instrumentos musicais, com que frequência você sente vontade de tocar algum instrumento musical, mesmo que não seja o instrumento utilizado para tocar a música que você está ouvindo?	4

<b>Código</b>	<b>Questão</b>	<b>Resposta</b>
16	Quando você não está se sentindo bem emocionalmente, com que frequência você comenta seus problemas para algum amigo(a) e permite que ele dê conselhos a você?	1
17	Quando você está em ambientes com várias pessoas e os assuntos vão surgindo, você gosta de debater mesmo que não tenham perguntado sua opinião, considerando assim, que sua opinião é importante?	1
18	Quando você viaja para lugares diferentes que você ainda não conhece, com que frequência você repara em detalhes nos formatos de árvores, cores e formas geométricas das casas, modelos dos carros, entre outras coisas?	4
19	Com que frequência você se oferece como voluntário(a) para realizar tarefas manuais, tais como utilizar chaves de fenda, cortar arames com alicate, pregar pregos, cortar papéis, lavar a louça, passar roupa, entre outras?	3
20	Com que frequência você prefere realizar tarefas que deixam você feliz independente da tarefa que você vai realizar, mesmo sabendo que existem outras tarefas que você pode realizar que não façam mal a você, mas também são necessárias?	3
21	Suponha que você esteja participando de um concurso musical e você tem facilidade em cantar músicas nacionais e em outro idioma independentemente do idioma. Com que frequência você faria a opção por cantar músicas internacionais?	4

## APÊNDICE D

Resultado dos testes realizados para a classificação por inteligência sem a utilização do filtro *Resample*.

Resultado da classificação dos indivíduos individualmente	
Indivíduo	Classificação por aspecto cognitivo
Indivíduo 1 – gênero feminino – idade 17 anos	43% Lógico matemático. 19% Linguístico. 19% Musical. 14% Espacial visual. 5% Interpessoal.
Indivíduo 2 – gênero feminino – idade 19 anos	38% Lógico Matemático. 38% Linguístico. 24% Corporal Cinestésico.
Indivíduo 3 – gênero masculino – idade 17 anos	43% Linguístico. 29% Espacial visual. 28% Intrapessoal.
Indivíduo 4 – gênero masculino – idade 18 anos	38% Interpessoal. 29% Lógico matemático. 28% Corporal Cinestésico. 5% Linguístico.
Indivíduo 5 – gênero masculino – idade 17 anos	52% Lógico matemático. 29% Espacial visual. 19% Musical.
Indivíduo 6 – gênero feminino – idade 20 anos	43% Musical. 33% Linguístico. 24% Intrapessoal.
Indivíduo 7 – gênero masculino – idade 18 anos	24% Lógico matemático. 24% Linguístico. 19% Musical. 19% Interpessoal. 14% Espacial visual.
Indivíduo 8 – gênero masculino – idade 20 anos	38% Lógico matemático. 33% Linguístico. 29% Musical.
Indivíduo 9 – gênero masculino – idade 17 anos	34% Lógico matemático. 34% Musical. 19% Intrapessoal. 13% Interpessoal.

Indivíduo 10 – gênero masculino – idade 17 anos	38% Interpessoal. 38% Musical. 24% Intrapessoal.
Indivíduo 11 – gênero masculino – idade 18 anos	29% Corporal Cinestésico. 24% Musical. 19% Intrapessoal. 14% Interpessoal. 14% Lógico Matemático.
Indivíduo 12 – gênero masculino – idade 17 anos	43% Musical. 29% Lógico matemático. 28% Interpessoal.

## APÊNDICE E

Resultado dos testes realizados para a classificação por inteligência com a utilização do filtro *Resample*.

Resultado da classificação dos indivíduos individualmente	
Indivíduo	Classificação por aspecto cognitivo
Indivíduo 1 – gênero feminino – idade 17 anos	38% Intrapessoal. 33% Musical. 29% Lógico matemático.
Indivíduo 2 – gênero feminino – idade 19 anos	67% Musical. 33% Linguístico.
Indivíduo 3 – gênero masculino – idade 17 anos	48% Linguístico. 33% Intrapessoal. 19% Musical.
Indivíduo 4 – gênero masculino – idade 18 anos	33% Musical. 24% Intrapessoal. 24% Lógico matemático. 14% Corporal Cinestésico. 5% linguístico.
Indivíduo 5 – gênero masculino – idade 17 anos	33% Intrapessoal. 29% Linguístico. 28% Musical. 10% Corporal Cinestésico.
Indivíduo 6 – gênero feminino – idade 20 anos	52% Musical. 48% Linguístico.
Indivíduo 7 – gênero masculino – idade 18 anos	43% Linguístico. 33% Musical. 24% Intrapessoal.
Indivíduo 8 – gênero masculino – idade 20 anos	38% Musical. 24% Lógico matemático. 19% Intrapessoal. 19% Linguístico.
Indivíduo 9 – gênero masculino – idade 17 anos	71% Musical. 29% Lógico matemático.
Indivíduo 10 – gênero masculino – idade 17 anos	57% Musical. 43% Linguístico.
Indivíduo 11 – gênero masculino – idade 18 anos	57% Musical. 43% Linguístico.

Indivíduo 12 – gênero masculino – idade 17 anos	33% Musical. 24% Corporal Cinestésico. 24% Intrapessoal. 19% Linguístico.
---	--