

## Classificação automática de pacientes para atendimento médico pediátrico multidisciplinar a partir do seu Grau de Vigilância

Juliana Tarossi Pollettini<sup>1</sup> Renato Tinós<sup>1</sup> Sylvia Panico<sup>2</sup>  
Julio César Daneluzzi<sup>2</sup> Alessandra Alaniz Macedo<sup>1</sup>

<sup>1</sup> Departamento de Física e Matemática, FFCLRP – USP

<sup>2</sup> Departamento de Pediatria e Puericultura, FMRP – USP

**Abstract.** *This paper proposes the automatic definition of “Grau de Vigilância – GV” (Alertness Level) towards helping medical decisions. GV is established by associations of factors that puts in risk the patient’s and his family’s health and development.*

**Resumo.** *O presente artigo propõe a definição automática da medida denominada Grau de Vigilância (GV), a qual visa ajudar na identificação de indivíduos que requerem diferentes níveis de atendimento médico.*

### 1. Introdução

No Brasil, os profissionais da área de atenção básica a saúde têm contato direto com pacientes e seus familiares e, muitas vezes, esses profissionais são a primeira fonte de informação sobre desenvolvimento e saúde. O conhecimento sobre o processo de desenvolvimento humano é importante para os profissionais que trabalham na área de saúde, uma vez que as informações coletadas nesse processo podem auxiliar em trabalhos preventivos, aprimorando a qualidade de vida dos pacientes. O papel da equipe de saúde das unidades básicas de saúde (UBSs), principalmente no âmbito da pediatria, é significativo e, portanto, os profissionais que atuam em atenção básica a saúde devem ser capazes de detectar, o mais cedo possível, crianças em situação de risco visando promover encaminhamentos ou intervenções que minimizem ou corrijam os problemas detectados para cada uma delas, assim como identificar crianças que contam com fatores protetores para a promoção da saúde e do desenvolvimento [1]. O uso de suporte tecnológico no contato paciente e profissional da saúde pode aprimorar a manipulação da informação coletada para, por exemplo, identificar e sinalizar crianças cujo processo de desenvolvimento possa ser considerado como em situação de risco.

Com o objetivo de realizar estudos que visam a sistematização de procedimentos para o acompanhamento do processo de desenvolvimento humano e a promoção da saúde de pacientes atendidos em UBSs, pesquisadores de diferentes universidades brasileiras, integrantes de um núcleo de pesquisa (Diretório Lattes CNPq) denominado “Núcleo Interdisciplinar de Pesquisa e Ensino (NIPE)” realizaram, durante as últimas décadas, diversas pesquisas, propondo a utilização de um índice denominado Grau de Vigilância (GV) [2]. O GV consiste em um índice cuja utilização permite a definição de procedimentos pediátricos a serem adotados no atendimento básico a saúde, por meio da associação de fatores de risco e de proteção à saúde e ao processo de desenvolvimento, aos quais estão expostos os pacientes (crianças e adolescentes) e suas famílias. O GV varia em valores gradativos possibilitando ações educativas, terapêuticas ou especializadas requeridas pelos pacientes e respectivas famílias [2].

O presente artigo apresenta mecanismos de definição automática desenvolvidos para a atribuição do GV do paciente, a partir da manipulação e análise de informações constantes em seu cadastro eletrônico (protocolo de atendimento médico pediátrico). Para a classificação automática foram investigadas duas técnicas de Reconhecimento de Padrões: *k*-Vizinhos Mais Próximos e Redes Neurais Artificiais. O resultado da definição automática do GV pode suportar a tomada de decisões por parte dos profissionais da equipe de saúde responsável pelo atendimento médico pediátrico em atenção básica a saúde. Considerando o GV como um indicador de ações para a definição do atendimento requerido, sua definição automática pode contribuir com o trabalho dos profissionais da saúde como uma “segunda opinião informatizada” para o auxílio à tomada de decisões.

## 2. Fundamentos Teóricos

A Inteligência Artificial (IA) pode ser definida como “[Automatização de] atividades que associamos ao pensamento humano, atividades como a tomada de decisões, a resolução de problemas, o aprendizado...” [3]. Russell e Norvig indicam que IA é composta por subcampos de pesquisa, como Sistemas Baseados em Conhecimento, Robótica, Processamento de Linguagem Natural e Aprendizado de Máquina [4].

O Aprendizado de Máquina, de acordo com Monard e Baranauskas, é uma área de IA que tem como objetivos o desenvolvimento de técnicas computacionais sobre o aprendizado e a construção de sistemas capazes de adquirir conhecimento de forma automática [5]. Existem diversos tipos de algoritmos de aprendizado de máquina, dentre os quais convém mencionar os de *k*-Vizinhos Mais Próximos e Redes Neurais Artificiais. A fim de preparar o leitor para um entendimento apropriado, as técnicas utilizadas nesta proposta serão brevemente introduzidas.

### 2.1. K-Vizinhos Mais Próximos

De acordo com Monard e Baranauskas, “Uma maneira de classificar um exemplo é lembrar de outro similar cuja classe é conhecida e assumir que o novo exemplo terá a mesma classe”. Essa característica é a base de um modelo de aprendizagem conhecido como aprendizagem baseada em instâncias, ou aprendizagem baseada na memória [5].

O conceito de *k*-Vizinhos Mais Próximos fundamenta-se nessa aprendizagem baseada em instâncias e, segundo Russell e Norvig, sua idéia-chave consiste no fato de que as propriedades de qualquer ponto de entrada específico têm probabilidade de serem semelhantes às propriedades de outros pontos em sua vizinhança [4]. Nesse sentido, o vizinho mais próximo de um padrão de entrada  $x$  é definido como sendo o que apresenta a menor distância entre seu vetor de características e o vetor de características de  $x$ . Portanto, ao classificar um novo padrão de entrada  $x$ , esse modelo se baseia numa função de distância para determinar o quão próximo o novo exemplo se encontra dos padrões armazenados e utiliza a classe associada ao vizinho mais próximo para determinar a classe do novo exemplo.

Haykin afirma que o classificador baseado em *k*-Vizinhos Mais Próximos é constituído de uma variante da regra do vizinho mais próximo e que sua particularidade reside no fato de que, em vez de identificar apenas um vizinho mais próximo, ele identifica *k*-Vizinhos Mais Próximos, sendo *k* um número inteiro [6]. Depois de encontrados os vizinhos, de acordo com Russell e Norvig, a classificação pode ser efetuada de acordo com a classe majoritária entre os *k*-vizinhos, caso o problema em

questão seja considerado discreto. De acordo com Moreno-Seco, Mico e Oncina, entre os diferentes mecanismos de classificação, geralmente o escolhido é o modelo de vizinho mais próximo, devido à sua simplicidade e boas propriedades, entretanto para melhorar as taxas de classificação o critério de *k*-Vizinhos Mais Próximos, em substituição ao anterior, é freqüentemente utilizado, apesar de aumentar o tempo de classificação [7].

Quanto à aplicação do conceito de *k*-Vizinhos Mais Próximos na área médica, segundo Santos Filho et al, a consulta dos *k*-vizinhos se constitui em uma das formas mais comuns de consulta por similaridade baseada em conteúdo [8]. Dentre as aplicações do conceito de consulta por similaridade na área médica, podem ser citadas: aplicação em PACS (Sistemas de Arquivamento e Comunicação de Imagens - *Picture Archiving and Communication Systems*), Cadeias de DNA (utilização da métrica de Levenshtein), Imagens (métrica MiCRoM *Minimum-Cost Region Matching*) e multi-atributos (métricas *Lp* da família Minkowski), como Exame laboratorial e Imagem.

## 2.2. Redes Neurais Artificiais

O conceito de Redes Neurais Artificiais (RNAs) nasceu a partir de tentativas de modelar algumas características simples do processamento de informações do cérebro humano através de redes de neurônios artificiais [4]. Conforme definida por Haykin, “uma rede neural é um processador maciçamente paralelamente distribuído constituído de unidades de processamento simples, que têm a propensão natural para armazenar conhecimento experimental e torná-lo disponível para o uso” [6]. Segundo ele, numa rede neural o procedimento utilizado para realizar o processo de aprendizagem é conhecido como algoritmo de aprendizagem e sua função é modificar as forças de conexão entre neurônios, conhecidas como pesos sinápticos, de forma ordenada, a fim de alcançar um objetivo. Haykin afirma que cada neurônio é uma unidade de processamento em uma rede neural e que possui três elementos básicos: um conjunto de *sinapses* ou elos de conexão entre o neurônio e seus sinais de entrada, um *somador* para somar os sinais de entrada ponderados pelos pesos das respectivas sinapses e uma *função de ativação* para restringir a amplitude da saída do neurônio.

Segundo Russell e Norvig, as RNAs podem ser usadas tanto para classificação quanto para regressão e, no que se refere à sua estrutura, existem duas categorias principais de redes neurais: redes com alimentação direta e redes recorrentes, sendo que as redes de alimentação direta normalmente estão organizadas em camadas. Essa organização faz com que cada unidade receba apenas a entrada de unidades situadas na camada imediatamente precedente e pode ser formada por apenas uma camada ou por várias camadas. A RNA de camada única, também conhecida como rede de *perceptron*, possui todas as entradas conectadas diretamente às saídas, enquanto as redes de várias camadas são caracterizadas pela presença de uma ou mais camadas de unidades ocultas.

Na literatura científica, podem ser encontradas diversas aplicações de RNAs no auxílio ao diagnóstico médico. Por exemplo, Economou et al desenvolveram um protótipo de sistema de decisão médica baseado em uma subclasse de RNAs que foi aplicado no diagnóstico de doenças pulmonares [9]. Burke, Rosen e Goodman [10] estudaram a aplicação das RNAs em prognóstico relacionado a cânceres e Walczak [11] aplicou RNAs em um instrumento de suporte à decisão médica que pudesse prever a real necessidade de transfusão sanguínea em pacientes do atendimento de emergência.

### **3. Contextualização da Proposta**

Alunos e pesquisadores de Informática Biomédica encontraram no Centro Médico Social Comunitário de Vila Lobato (CMSC Vila Lobato) uma oportunidade para elaborar projetos de pesquisa relacionados ao desenvolvimento de sistemas de computação. Desse modo, esses pesquisadores consolidam os conhecimentos adquiridos e procuram aprimorar a qualidade do cuidado com pacientes.

O CMSC Vila Lobato, uma unidade de atendimento primário à saúde localizada em um bairro periférico da cidade de Ribeirão Preto/SP, exerce atividades clínicas, de ensino e de pesquisa, possui uma tradição de quase 40 anos no acompanhamento em puericultura à saúde da criança e do adolescente e atende aproximadamente 18 mil pessoas. Enquanto usualmente, na grande maioria das unidades de atendimento primário a saúde existentes na realidade brasileira, o atendimento é realizado quando os pacientes chegam ao centro médico apresentando alguma enfermidade, o CMSC Vila Lobato apresenta o diferencial de manter uma proximidade maior entre a equipe de saúde e seus pacientes. Essa proximidade é conquistada pelo fato de, a cada atendimento realizado no centro médico, ser agendada uma nova consulta de rotina para o paciente e também pela existência de programas que incluem atendimentos multidisciplinares voltados ao acompanhamento pediátrico, socioeconômico e comportamental.

Dentre os programas do CMSC Vila Lobato, podem ser citados o PROISE (Programa de Integração Saúde Educação, com a participação da Família e da Comunidade) e o PROESPA (Programa de Educação em Saúde para Pais e Famílias), ambos idealizados pela equipe de pesquisadores e profissionais, que compõem o Grupo de pesquisa NIPE, e atuantes no CMSC Vila Lobato. Estes dois programas permitem a atuação profissional multidisciplinar, possibilitando a adoção de intervenções baseadas na identificação de fatores de risco e de fatores protetores à saúde e ao processo de desenvolvimento humano. O PROISE é um programa voltado para a promoção da saúde e do desenvolvimento de crianças e adolescentes, desenvolvido em parceria entre o CMSC de Vila Lobato e uma escola de ensino fundamental e médio. O PROESPA, por sua vez, é um programa que permite o atendimento a crianças/adolescentes e seus pais/familiares, envolvendo o atendimento médico pediátrico multidisciplinar. Quando em algum atendimento, são identificados determinados fatores de risco, um Grau de Vigilância (GV), baseado nos fatores de risco identificados, é atribuído a ele e um encaminhamento ao PROESPA é efetuado. Diante de cada encaminhamento, a equipe interdisciplinar do CMSC Vila Lobato avalia o caso e determina o tipo de atendimento (individual ou em grupo) que o paciente e seus familiares devem receber. O GV, definido conforme a Tabela 1, possibilita a identificação de grupos que requerem diferenciação nas ações em atenção primária, secundária ou terciária à saúde [2].

O presente projeto está inserido no contexto de um projeto maior, realizado entre equipe de pesquisa, ensino e extensão da área de informática biomédica e o CMSC de Vila Lobato que tem como objetivo geral a informatização de procedimentos adotados de atendimento médico pediátrico multiprofissional, o SI-VilaLobato (Sistema de Informação Vila Lobato). O SI-VilaLobato é um sistema de informação que visa a informatização dos procedimentos de atendimento e acompanhamento de saúde e do processo de desenvolvimento das crianças e adolescentes atendidos pelo CMSC Vila Lobato, assim como suas respectivas famílias [12]. O SI-VilaLobato inclui o subsistema de apoio ao PROISE [13], um subsistema em desenvolvimento de apoio ao PROESPA e o subsistema proposto neste artigo.

Tabela 1 – Definições e valores gradativos para Grau de Vigilância (GV)

Grau de Vigilância	Definição
GV Rotina	Identificação de necessidade de ações de rotina, usualmente prestadas nos serviços de atenção primária à saúde.
GV1	Identificação de fatores de risco que requerem ações educativas.
GV2	Identificação de fatores de risco associados que requerem, além de ações educativas, ações terapêuticas.
GV3	Identificação de diversos fatores de risco associados que requerem, além de ações educativas e terapêuticas, ações terapêuticas especializadas.
GV Urgência	A atribuição de qualquer um dos graus de vigilância pode vir associada à atribuição de indicação de urgência no atendimento.

FONTE: Indicadores NIPE (1997).

#### 4. Proposta

Conforme mencionado anteriormente, GVs auxiliam na identificação de pacientes que requerem diferentes níveis de atendimento em atenção básica a saúde. Os autores deste artigo propõem a definição automática do GV do paciente a partir da manipulação e análise de informações provenientes dos prontuários de atendimento do CMSC Vila Lobato e abastecidas na base de dados do sistema de informação.

Pretende-se explorar k-Vizinhos Mais Próximos e Redes Neurais Artificiais para o cálculo automático ou semi-automático de GV e, de acordo com os resultados, estudar o melhor a ser adotado no contexto de manipulação de informações médicas pediátricas e multidisciplinares no CMSC Vila Lobato. No momento, a definição manual ou automática do GV está sendo realizada no CMSC de Vila Lobato, enquanto uma unidade de pesquisa, ensino e extensão, mas com potencial para ser adotada em outras UBSs, desde que considerem as informações necessárias para sua definição. A seguir são apresentadas as etapas de definição dos módulos de cálculo automático de GVs.

##### 4.2. Levantamento de Variáveis e Modelagem

A partir de estudo da proposta do índice de GV e de observação do profissional da saúde durante a determinação do GV, verificou-se que dentre as diversas informações contidas no prontuário do paciente algumas possuíam maior impacto na definição da medida, por exemplo, as impressões diagnósticas. Entretanto, somente as impressões diagnósticas não são suficientes para uma adequada classificação. Outras informações como idade do paciente, tipo de agendamento do atendimento e condutas associadas podem também influenciar na atribuição do índice GV, visto que podem fornecer informações extras a respeito do estado de saúde do paciente. Portanto, foram consideradas necessárias para definir o GV para cada atendimento de cada paciente as seguintes variáveis: a *idade* do paciente, o *sexo* do mesmo, as *impressões diagnósticas* e *condutas* associadas ao atendimento e o *tipo de agendamento* do mesmo. Outras variáveis, que podem influenciar na definição do GV, deverão ser analisadas e modeladas posteriormente.

O estudo do comportamento da classificação dos GVs, perante o maior número possível de variações em seus atributos de interesse, permitiu a criação de diversas tabelas e gráficos que suportaram a investigação da interferência exercida pela

distribuição desses diferentes fatores, assim como seus relacionamentos e dependências, na definição final do GV. Mecanismos foram criados para combinar as variáveis selecionadas para definir GV's, resumir e modelar o processo de estudo das variáveis e, assim, determinar o GV para atendimentos de acordo com os atributos de entrada.

#### **4.3. Definição da Coleção de Treinamento**

Com o objetivo de organizar as informações presentes no banco de dados do sistema SI-VilaLobato, foi implementada e incorporada ao sistema uma classe de software responsável por gerar automaticamente uma planilha que contivesse as meta-informações necessárias para a definição do GV e de seus conteúdos. A planilha gerada foi utilizada pelos profissionais da equipe de saúde para definir manualmente o GV para cada atendimento, de forma que servissem como exemplo para o aprendizado dos módulos de classificação k-Vizinhos Mais Próximos (VPGV) e Redes Neurais (RNGV).

Durante análise da planilha, percebeu-se que as impressões diagnósticas e, principalmente, as condutas médicas não possuíam vocabulário padronizado, pois usuários diferentes definiam um mesmo conceito utilizando termos diferentes. Após um estudo detalhado de algumas possibilidades de se estabelecer atributos e valores para pré-processamento dos dados, optou-se por utilizar uma abordagem que, quando possível, relacionasse impressões diagnósticas envolvidas com conceitos semelhantes e estabelecesse cada uma delas como um valor diferente para um mesmo atributo e, de forma análoga, relacionasse condutas semelhantes. Após a elaboração da relação dos atributos e respectivos valores que podem ser assumidos, foram codificados métodos, em linguagem Java, que são responsáveis por gerar um arquivo contendo os exemplos de atendimentos com os atributos organizados. Para treinamento dos classificadores foi utilizada uma amostra de 103 atendimentos com o GV classificado pelos profissionais da equipe de atendimento pediátrico multidisciplinar do CMSC Vila Lobato.

#### **4.4. Criação dos Módulos Classificadores**

Para a criação dos módulos classificadores, inicialmente foram levantados requisitos gerais para um sistema que definisse automaticamente o GV, de forma que abrangessem os dois módulos classificadores propostos: módulos de classificação usando k-Vizinhos Mais Próximos (VPGV) e Redes Neurais Artificiais (RNGV).

O primeiro requisito observado foi que seria imprescindível que os módulos VPGV e RNGV pudessem ser integrados ao sistema de informação do CMSC Vila Lobato e, para isso, deveriam compor um subsistema denominado GV-Automático. Além disso, outros requisitos funcionais foram levantados: (i) o sistema deveria permitir que o GV atribuído em cada atendimento por equipe médica pediátrica e multidisciplinar fosse inserido no cadastro eletrônico do paciente no momento em que um oficial administrativo cadastrasse o atendimento no sistema; (ii) permitir que profissionais mais familiarizados com o contexto do GV e que, por isso, possuem maior experiência na atribuição desse índice visualizassem um resumo dos atendimentos, incluindo o GV atribuído no momento do atendimento; e (iii) disponibilizar o GV atribuído automaticamente e permitir que os profissionais da equipe multidisciplinar corrigissem o GV atribuído no momento do atendimento, de forma que o GV automático funcionasse para eles como uma “segunda opinião informatizada”. Quando futuramente os médicos e alunos tiverem contato direto com o subsistema, este deverá disponibilizar a eles uma interface gráfica com as informações de um atendimento e ter acesso ao GV atribuído automaticamente, de forma que o GV automático funcione

como base para a atribuição do GV manual. Pelo fato do treinamento dos classificadores ter sido realizado com uma amostra, tornou-se necessário permitir que novos classificadores sejam configurados pelo administrador do sistema, para que, conforme o número de exemplos confiáveis aumente, os classificadores possam ser retreinados para melhorar o aprendizado dos algoritmos e, conseqüentemente, melhorar também suas taxas de acerto. Além desse novo requisito, percebeu-se que seria necessário que o GV-Automático tivesse uma funcionalidade de comparação de resultados advindos dos diferentes classificadores, cujos resultados estão na Seção 5.

Os módulos VPGV e RNGV, mais especificamente, são responsáveis pela classificação de novos exemplos de atendimentos entre os diferentes valores do GV e permitem também que os classificadores sejam atualizados à medida que o número de exemplos classificados na base de dados cresce. É interessante ressaltar que para retreinar os classificadores apenas exemplos com o GV confirmado pelos profissionais da equipe médica pediátrica e multidisciplinar mais familiarizados com o índice (GV) devem ser utilizados. Os módulos possuem uma interface com classes do pacote Weka[14], fazendo chamadas aos algoritmos IBk e MultilayerPerceptron, conforme mostrado na Figura 1. No diagrama dessa figura podem ser visualizadas as classes envolvidas no processo de classificação e os métodos usados pelos módulos propostos.

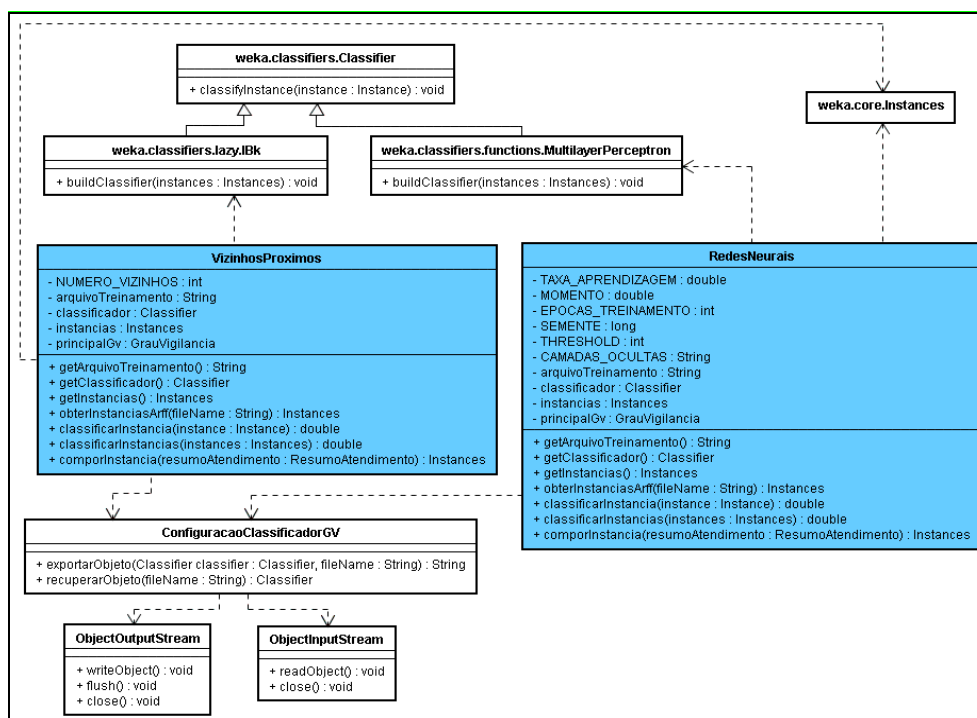


Figura 1: Diagrama de classes ressaltando a ligação entre classes

Para a criação do classificador do módulo VPGV, a classe **VizinhosProximos** gera um arquivo contendo os atributos das variáveis selecionadas na fase de modelagem e os exemplos de atendimentos que serão utilizados para o treinamento. Em seguida, a mesma classe faz uma chamada ao construtor **weka.classifiers.lazy.IBk**, passando como parâmetro o número  $k$  de vizinhos mais próximos. Para representar os exemplos de treinamento, o módulo instancia a classe **weka.core.instances** considerando o arquivo com os atributos e exemplos, e estabelece como classe o atributo que indica o

GV dos mesmos. Finalmente o algoritmo faz uma chamada ao método *buildClassifier* da classe **weka.classifiers.lazy.IBk** e exporta o classificador instanciado em uma chamada ao método *exportarObjeto* da classe **ConfiguraçãoClassificadorGV**, a qual acessa a classe **ObjectOutputStream** para salvar de forma persistente o estado do objeto naquele instante. Em novos atendimentos, o módulo recupera o classificador cujo estado foi salvo e classifica os novos exemplos com os métodos *classificarInstancia* e *classificarInstancias*.

O módulo RNGV é similar ao VPGV no que se refere a acessos a classes externas, mas, ao invés de utilizar a classe **weka.classifiers.lazy.IBk**, são chamados métodos da classe **weka.classifiers.functions.MultilayerPerceptron**.

#### 4.4. O Sistema de Classificação Automática de GV

Seguindo as especificações e o projeto, o sistema GV-Automático possui quatro funcionalidades bem definidas: *verificação e correção do grau de vigilância*, *consulta do mesmo índice*, *configuração de classificadores* e *comparação e seleção dos classificadores*. Essas funcionalidades são apresentadas na interface inicial do sistema.

A funcionalidade de *verificação e correção do GV* apresenta um resumo dos atendimentos de um determinado período de interesse, permitindo que os profissionais mais familiarizados com o referido índice analisem cada atendimento e corrijam o GV atribuído no momento do atendimento, caso julguem necessário. Disponibiliza o GV atribuído pelo GV-Automático. É interessante ressaltar que, para cada atendimento, esses profissionais avaliam o GV atribuído e marcam o grau de vigilância como “confirmado”, caso concordem com a atribuição realizada no momento do atendimento e, caso não concordem, podem alterar o valor desse índice.

Quanto à funcionalidade de *consulta do GV*, o sistema disponibiliza aos profissionais de saúde uma segunda opinião a respeito da atribuição do índice, de acordo com informações fornecidas pelo próprio profissional de um dado atendimento.

A funcionalidade de *configuração de classificadores* foi criada para tornar o sistema dinâmico, permitindo que os classificadores a serem usados pelas demais funcionalidades do subsistema sejam reconfigurados e tenham condições de aprimorar o aprendizado. Para que os classificadores sejam retreinados apenas exemplos confiáveis devem ser considerados. A interface para configuração de classificadores apresenta duas opções, uma para cada tipo de classificador proposto: VPGV e RNGV. De acordo com testes de treinamentos realizados no decorrer da pesquisa, foram estabelecidas combinações de valores padrão para os parâmetros necessários à configuração dos classificadores, entretanto, o sistema permite que sejam alteradas pelo administrador para que o mesmo possa testar o treinamento para diferentes combinações em busca dos melhores resultados de aprendizado. Tornou-se necessário também disponibilizar um mecanismo para que o administrador pudesse comparar os resultados desses classificadores no momento da atualização. Quando um módulo para determinado conjunto de variáveis apresentar uma maior taxa de acerto, pode ser necessário que o administrador realize, de forma prática, a alteração do classificador a ser usado pelo sistema. Assim, elaborou-se um módulo composto pelos dois classificadores o qual permite que o próprio administrador selecione o classificador mais adequado para ser utilizado pelas demais funcionalidades do GV-Automático.

#### **4.5. Resultados Experimentais**

Para realizar experimentos com o GV-Automático, foram estudadas metodologias de avaliação de algoritmos de Aprendizado de Máquina, pois os módulos de categorização implementados foram desenvolvidos de acordo com essa teoria. Após o estudo, optou-se por usar *Stratified Cross-Validation* [15] com divisão dos exemplos em 10 partições.

Os classificadores VPGV e RNGV foram avaliados para o conjunto original de 103 exemplos fornecidos pelos profissionais do CMSC Vila Lobato e foram obtidas as seguintes taxas de acerto: 37,25 % de acerto para o classificador VPGV e 61,76% para o classificador RNGV. A princípio, os resultados obtidos foram bem melhores para o RNGV com os exemplos fornecidos. Por considerar que as taxas de acertos de cada classificador podem variar dependendo da coleção, decidiu-se permitir o classificador a ser utilizado pelo GV-Automático seja selecionado e configurado.

#### **5. Conclusão**

Para apoiar a identificação de pacientes que requerem diferentes níveis de atendimento, um índice de classificação para atendimento denominado Grau de Vigilância (GV) pode ser utilizado. O GV é estabelecido segundo a existência e associação a fatores de risco e fatores protetores à saúde e ao processo de desenvolvimento humano, aos quais estão expostos os pacientes e suas famílias. Este artigo propõe a classificação automática do GV do paciente a partir da manipulação e análise de informações constantes em seu cadastro eletrônico. Essa proposta tem sido suportada por Redes Neurais Artificiais (RNGV) e K-Vizinhos Mais Próximos (VPGV).

Apesar da taxa de acertos para o classificador RNGV ser de 61,76%, acredita-se que possa ser melhorada, pois esses resultados podem estar relacionados à forma como os atributos foram estabelecidos no início da pesquisa. Pretende-se ampliar o estudo de classificadores e aprimorá-lo. Uma análise a respeito da taxa de acertos da atribuição manual do GV por parte de profissionais de saúde depende do início da utilização do subsistema GV-Automático, principalmente no que se refere à funcionalidade de Verificação/Correção do GV. Por esse motivo ainda não é possível avaliar se a taxa obtida para o RNGV é superior ou inferior à porcentagem de acertos manuais.

Durante a modelagem das informações do cadastro eletrônico do paciente a serem consideradas pelos classificadores, percebeu-se que as impressões diagnósticas e, principalmente, as condutas não possuíam um vocabulário padronizado. Usuários diferentes definiam um mesmo conceito com termos diferentes. Pretende-se dar continuidade a esta pesquisa por meio do uso e extensão de thesaurus e ontologias.

#### **6. Agradecimentos**

À FAPESP pelo financiamento concedido ao projeto (2006/06715-8 e 2004/12477-7). Ao Prof. Dr. Evandro Ruiz e Thiago M. da Costa, pelas etapas iniciais de pesquisa e definição da temática central do presente trabalho. Ao Prof. Dr. Joaquim Felipe e ao Prof. Dr. Domingos Alves pela colaboração técnica a domínios de conhecimento relacionados à pesquisa. À equipe de técnicos, profissionais e pesquisadores do CMSC Vila Lobato, pela colaboração e participação durante o processo de desenvolvimento e implantação do SI-VilaLobato.

## 7. Referências

- [1] SALLES R. F. **Análise de um programa de intervenção com bebês e famílias atendidas em unidades básicas de saúde – SUS**. 2001. 0 f. Dissertação (Mestrado em Educação Especial) - Universidade Federal de São Carlos. *Orientador*: Sylvania Panico.
- [2] PANICO, S. R. G.; CANZIANI, M. L.; GUERCHON, N. Políticas Públicas Municipais. In: Sylvania Rosalina Grasseschi Panico. (Org.). **Indicadores Nipe**: Subsídios para Políticas Municipais de Saúde. 1 ed. São Carlos: NIPE, 1997, v.1, n1.
- [3] BELLMAN, R.E. **An Introduction to Artificial Intelligence: Can Computers Think?** Boyd & Fraser Publishing Company, San Francisco. 1978 *apud* RUSSELL, S. J.; NORVIG, P. **Inteligência Artificial**. 2. ed., 2004.
- [4] RUSSELL, S. J.; NORVIG, P. Métodos estatísticos de aprendizagem. In: \_\_\_\_\_. **Inteligência Artificial**. Tradução: Vandenberg D. de Souza. 2.ed., p.690-737/2004.
- [5] MONARD, M. C.; BARANAUSKAS, J. A. Conceitos sobre Aprendizado de Máquina. In: REZENDE, S.O (org). **Sistemas Inteligentes**: fundamentos e aplicações. 1.ed. Barueri, SP: Manole, 2003. cap.4, p.89-114.
- [6] HAYKIN, S. Introdução. In: \_\_\_\_\_. **Redes neurais**: princípios e prática. 2.ed. Porto Alegre: Bookman, 2001. cap.1, p.27-74.  
\_\_\_\_\_. Processos de Aprendizagem. In: \_\_\_\_\_. **Redes neurais**: princípios e prática. 2.ed. Porto Alegre: Bookman, 2001. cap.2, p.75-142.
- [7] MORENO-SECO, F.; MICO, L.; ONCINA, J. A new classification rule based on nearest neighbour search. In: Proceedings of the 17th International Conference on Pattern Recognition, 2004. **ICPR 2004**. v.4, p.408-411, 2004.
- [8] SANTOS FILHO, R.F.; SOUSA, E.; TRAINA, A.; TRAINA JR, C. Desmistificando o Conceito de Consultas por Similaridade: A Busca de Novas Aplicações na Medicina. In **Anais do 2º Workshop de Informática Médica**, p.1-4, 2002.
- [9] ECONOMOU,G.; HALLAS,J.; MARIATOS,E.; GOUTIS,C. Artificial neural networks in medical decision making systems: an application to pulmonary diseases' diagnosis through VHDL synthesis. In: European Design and Test Conference, p.590, 1995.
- [10] BURKE, H.B.; ROSEN, D.B.; GOODMAN, P.H. Comparing artificial neural networks to other statistical methods for medical outcome prediction. In: **IEEE World Congress on Computational Intelligence**, USA, v. 4, p. 2213 – 2216, 1994.
- [11] WALCZAK, S. Artificial neural network medical decision support tool: predicting transfusion requirements of ER patients. In: **IEEE Transactions on Information Technology in Biomedicine**. v.9, n.3, p.468-474, 2005.
- [12] COSTA, T. M.; et al. **SiVilaLobato: desenvolvimento e implantação incremental de um registro eletrônico em saúde**. Artigo em preparação a ser submetido.
- [13] COSTA, T. M.; RUIZ, E. E. S.; PANICO, S. **Projeto Interdisciplinar para a criação de uma ferramenta computacional que facilite a pesquisa acadêmica e o acompanhamento da saúde e do desenvolvimento de adolescentes em atenção básica à saúde**. Projeto financiado pela Fundação de Amparo ao Ensino e Pesquisa (FAEPA), 2006.
- [14] WITTEN, I. H.; FRANK, E. **Data Mining**: Practical machine learning tools and techniques, 2nd Edition, Morgan Kaufmann, San Francisco, 2005.
- [15] KOHAVI, R. A Study of Cross-Validation and Bootstrap for Accuracy Estimation and Model Selection. In: **International Joint Conference on Artificial Intelligence (IJCAI)**, 1995.