



Algoritmo evolutivo para predição de dados antropométricos infantis como apoio à tomada de decisão

Evolutionary algorithm for prediction of anthropometric data of children as support to decision making

Algoritmo evolutivo para predicción de los datos antropométricos de niños como apoyo a toma de decisiones

Rúbia Eliza de Oliveira Schultz Ascari¹, Beatriz Terezinha Borsoi¹, Eliane Maria De Bortoli Fávero¹

RESUMO

Descritores: Técnicas de apoio para a decisão; Previsões; Antropometria; Inteligência Artificial

Objetivo: Reportar o emprego de um algoritmo de programação genética no contexto da predição para gerar automaticamente uma equação matemática que represente o mais próximo possível a curva de crescimento infantil (referentes a peso e altura) registrada historicamente. **Método:** O algoritmo de programação genética proposto está focado na indução de equações matemáticas, executando o processo chamado de regressão simbólica sobre dados de quatro crianças do sexo masculino. **Resultado:** Observou-se que a curva gerada torna-se mais próxima da curva de crescimento nos últimos pontos de treinamento. Assim, quanto maior o número de dados para treinamento, maior a precisão da predição. **Conclusão:** A abordagem proposta retrata tendências de evolução com base em informações do histórico da criança analisada e pode ser utilizada para identificar precocemente possíveis problemas de desnutrição, obesidade e outros aspectos relacionados ao desenvolvimento físico, servindo de auxílio na tomada de decisão por parte de especialistas da área da saúde.

ABSTRACT

Keywords: Decision Support Techniques; Forecasting; Anthropometry; Artificial Intelligence

Objective: To report the use of a genetic programming algorithm in the prediction context to automatically generate a mathematical equation that represents the child's growth curve (for weight and height) historically recorded. **Method:** The proposed genetic programming algorithm is focused in the induction of mathematical equations, running the process called symbolic regression on data from four male children. **Results:** It was observed that the generated curve becomes closer of growth curve in the last training points, so with a more number of training data, better the accurate of prediction. **Conclusion:** The proposed approach portrays evolution trends based on historical information of the analyzed child and can be used to identify potential problems of early malnutrition, obesity and other aspects related to physical development, serving as an aid in decision making by specialists health.

RESUMEN

Descriptores: Técnicas de Apoyo para la Decisión; Predicción; Antropometría; Inteligencia Artificial

Objetivo: Describir el uso de un algoritmo de programación genética en contexto de predicción para generar automáticamente una ecuación matemática que representa lo más cercano posible a la curva de crecimiento infantil (peso y altura) registrada históricamente. **Método:** El algoritmo propuesto utiliza la regression sinbolica de los datos de cuatro niños varones para obtener la cuerva y la ecuacion matematica que representa estos datos. **Resultados:** En general, la curva obtenida se aproxima a la curva de crecimiento en los últimos puntos de entrenamiento. Cuanto mayor sea el número de datos de entrenamiento, más precisa será la predicción. **Conclusión:** El enfoque propuesto representa las tendencias de evolución basadas en la información histórica de los niños analizados y puede ser utilizada para identificar los problemas potenciales de desnutrición temprana, obesidad y otros aspectos relacionados con el desarrollo físico, sirviendo como ayuda en la toma de decisiones por parte de especialistas en salud.

¹ Professora do Departamento de Informática da Universidade Tecnológica Federal do Paraná - UTFPR, Pato Branco (PR), Brasil.

INTRODUÇÃO

A evolução do crescimento infantil é acompanhada pelos médicos pediatras desde o nascimento da criança, sendo os dados utilizados como um importante recurso para a identificação e a prevenção de possíveis doenças físicas existentes ou que venham a se desenvolver na criança. Para a Organização Mundial de Saúde (World Health Organization (WHO)), o Ministério da Saúde e a Sociedade Brasileira de Pediatria, o acompanhamento do crescimento deve ser atividade de rotina no atendimento infantil, devido ao reconhecimento da influência que as condições de vida exercem sobre o crescimento⁽¹⁾.

A Antropometria é a ciência que estuda e avalia as medidas de tamanho (altura, diâmetros e comprimentos ósseos, espessuras e dobras cutâneas, circunferência cefálica), peso e alguns índices que avaliam o risco de desenvolver doenças relacionadas ao desenvolvimento físico⁽²⁻³⁾, sendo amplamente aplicada na avaliação do estado nutricional de crianças e adolescentes. As aplicações da antropometria têm se expandido, incluindo a capacidade de selecionar indivíduos que se beneficiarão de determinada intervenção, identificar problemas socioeconômicos e avaliar a resposta às intervenções. A obtenção dos dados antropométricos é realizada por um método barato, não-invasivo, de fácil aplicação, de boa aceitação por parte da população e extremamente útil para rastrear obesidade, subnutrição e outros agravos nutricionais⁽⁴⁾.

Em crianças e adolescentes é comum o uso dos indicadores antropométricos de peso/idade, altura/idade, peso/altura e Índice de Massa Corporal (IMC) segundo idade e sexo⁽⁵⁾. Esses indicadores podem ser comparados com as informações das curvas e tabelas de referência do National Center for Health Statistics (NCHS) e da Organização Mundial de Saúde (OMS)^(6,4,7).

A partir de 2005, no Brasil, toda criança recebe ao nascer a caderneta de saúde que contém informações úteis aos pais e os gráficos antropométricos desenvolvidos com base nas referências da NCHS e OMS, para acompanhamento do crescimento da criança, levando em consideração a evolução de seu peso, altura, IMC e perímetro cefálico. Os valores são válidos para crianças nascidas a termo (com 37 semanas ou mais de gestação). A avaliação do estado nutricional, usando os dados antropométricos, tem o objetivo de quantificar e qualificar essas medidas que são comparadas com valores de referência e assim, determinar se os valores encontrados estão ou não dentro dos intervalos de normalidade⁽⁵⁾.

Sendo assim, seria muito importante a possibilidade de predizer se o desenvolvimento da criança será conforme o esperado. Uma das formas de fazer essas predições é utilizando sistemas computacionais. Os algoritmos desses sistemas podem ser baseados, por exemplo, em dados ou séries históricas estatísticas ou em processos naturais. O processo de seleção que ocorre na natureza pode ser simulado no mundo computacional por meio de algoritmos evolutivos. Porém, no lugar de uma população de seres vivos, tem-se um conjunto de possíveis soluções para um determinado problema.

Métodos da Computação Evolutiva (Algoritmos

Genéticos (AG) e Programação Genética (PG) vem sendo utilizados com êxito em uma grande diversidade de aplicações, pois apresentam resultados bastante aceitáveis em relação aos recursos utilizados, sendo de fácil implementação e úteis na resolução de problemas de difícil resolução por outros métodos, especialmente quando se pretende predizer contextos. Conceitualmente a PG pode ser vista como uma extensão dos algoritmos genéticos, representando um método sistemático para tornar os computadores aptos a resolver problemas. A PG é desenvolvida a partir de algoritmos genéticos, mas enquanto a PG representa a solução como uma árvore, em algoritmos genéticos podem ser utilizadas outras formas de representação, como por exemplo, uma *string* binária⁽⁸⁾. Neste trabalho a PG é empregada para gerar automaticamente equações simbólicas que serão utilizadas na predição de dados antropométricos de crianças (peso e altura).

A programação genética se baseia na combinação de conceitos e ideias da teoria da evolução (seleção natural), da genética (reprodução, cruzamento e mutação), da inteligência artificial (busca heurística) e da teoria de compiladores (representação de programas como árvores sintáticas)⁽⁹⁾. O processo evolutivo ocorre na natureza quando quatro condições básicas são satisfeitas⁽⁹⁾: um indivíduo tem a habilidade de se reproduzir; existe uma população desses indivíduos; existe alguma variedade entre esses indivíduos; e a variedade influencia de alguma forma na capacidade de sobrevivência dos indivíduos.

A representação de programas em árvores e a definição de operadores genéticos foram apresentadas inicialmente por Cramer⁽¹⁰⁾. Seguindo essa linha de raciocínio e como base os trabalhos de Holland⁽¹¹⁾, Koza⁽¹²⁻¹³⁾ introduziu o conceito de programação genética utilizando-a como meio de conduzir a pesquisa por programas no espaço de soluções: escalabilidade e reuso de programas⁽¹⁴⁾, resoluções de problemas⁽¹²⁾, soluções comparáveis a soluções humanas⁽¹⁵⁾.

Tanto em algoritmos genéticos quanto em programação genética, através da evolução da população, pode-se chegar ao indivíduo que apresenta a melhor solução ou pelo menos uma boa aproximação da resolução do problema. Para isso, começa-se com uma população inicial aleatória ou pré-determinada. Os indivíduos da população são avaliados por algum critério que indica o grau de aptidão (*fitness*) de cada um deles ao modelo. Os indivíduos com melhor grau de aptidão são selecionados e modificados por meio de operadores genéticos, a fim de gerar novos indivíduos para formação de uma nova população, simulando o processo evolutivo que ocorre na natureza. Em seguida, descarta-se a população original e repete-se o ciclo com essa nova população. Cada ciclo é chamado de geração e se repete até que um determinado critério de término seja satisfeito.

Os algoritmos genéticos permitem definir a solução do problema que pode ser utilizada como apoio ou suporte à tomada de decisão. A decisão cabe ao especialista humano, mas ela pode ser fundamentada por dados obtidos por sistemas computacionais que são implementados utilizando métodos de inteligência artificial, como algoritmos genéticos, lógica *fuzzy* e redes neurais.

Os Sistemas de Apoio à Decisão (SAD) podem proporcionar melhoria no desempenho de profissionais de saúde⁽¹⁶⁾ pelo amparo ao processo decisório,

impactando na qualidade do serviço oferecido. Uma característica importante desse tipo de sistema é a possibilidade de simulação de cenários com base em informações disponíveis sobre determinado assunto, permitindo a predição de algo.

Considerando o objetivo desse trabalho, cabe diferenciar os termos previsão, predição e projeção. A previsão, de uma forma geral, é uma afirmação sobre o futuro e é condicional sobre o que é esperado acontecer se uma série de hipóteses admitidas se torne válida. Assim, a previsão pode ser entendida como uma projeção ou extrapolação de tendências com base no passado⁽¹⁷⁾. Já a predição consiste em ajustar um modelo aos dados de uma série⁽¹⁸⁾. Takashina⁽¹⁹⁾ complementa dizendo que “previsão = projeção + predição”, onde a projeção é baseada nos resultados passados e a predição é baseada na capacidade de obter os recursos para um novo contexto. Sendo assim, considerando essas definições, optou-se no presente trabalho por utilizar o termo predição no sentido de criar uma representação do estado nutricional futuro de crianças, com base nas tendências registradas em seu histórico de desenvolvimento. A predição, quando científica, importa certeza ou alta probabilidade⁽²⁰⁾.

Desta forma, tomando como base as informações apresentadas nos gráficos antropométricos e a possibilidade de prever situações a partir da aplicação da PG, o objetivo deste trabalho é reportar o emprego de um algoritmo de PG para gerar automaticamente uma equação matemática que represente o mais próximo possível a curva de crescimento infantil (referentes a peso e altura) registrada historicamente. Sendo possível, assim, realizar a predição de como será o seu desenvolvimento nos meses seguintes, servindo como base para um SAD. Este procedimento é chamado de regressão simbólica⁽⁹⁾ e os dados obtidos com a predição podem ser utilizados para identificar precocemente tendências de desnutrição, obesidade e outros aspectos relacionados ao desenvolvimento físico. O algoritmo desenvolvido apresenta como ocorreria a curva de desenvolvimento da criança a partir dos seus dados históricos, possibilitando avaliar se a predição de desenvolvimento da criança está dentro da normalidade ou não, possibilitando ao profissional de saúde indicar os procedimentos de investigação e tratamento adequados, caso necessário. Destaca-se que a precisão no encaminhamento feito pelo profissional de saúde, pode evitar problemas relacionados ao desenvolvimento da criança e até mesmo custos desnecessários com a saúde.

MÉTODOS

O método descreve o algoritmo de PG proposto para prever dados antropométricos de crianças (peso e altura). A solução apresentada determina a classe e os coeficientes da função a ser adotada no modelo de regressão. A saída desse modelo define o peso ou a altura futuros da criança, com base nos respectivos dados de entrada e essa informação pode ser utilizada para identificar tendências para a ocorrência de problemas como desnutrição.

O algoritmo de PG proposto está focado na indução de equações matemáticas baseadas nos dados de estudo,

ou seja, executa o processo chamado de regressão simbólica. O objetivo principal da regressão simbólica é encontrar uma função que se aproxime ao máximo de outra função desconhecida através de um determinado conjunto de dados, ou seja, definir a função representada por esse conjunto de dados e possibilitar a representação da continuidade dessa curva. Assim, o algoritmo de PG desenvolvido busca por equações matemáticas que apresentem uma boa aproximação da curva de evolução do desenvolvimento da criança em estudo. Esse algoritmo considera como entrada o peso ou a altura registrados da criança e a diferença temporal entre esses registros. Então, para um conjunto de medições registradas em “n” instantes de tempo, o peso ou a altura será predito para o instante “tn+1”, com base na melhor solução apresentada pelo algoritmo de PG. A melhor solução neste caso será uma expressão cujos valores de saída resultem no menor erro de aproximação (modificação do Erro Quadrático Médio)(EQM). O método considera todos os pontos registrados pela avaliação destinados para treinamento, assumindo que os pontos finais têm maior influência.

Basicamente, o algoritmo de PG gera uma população de equações matemáticas (indivíduos), avalia cada indivíduo com base na equação que ele representa, usa métodos de seleção para definir quais equações sobreviverão, aplica operadores genéticos para modificar equações existentes e gerar novas equações, a fim de convergir para uma boa solução. A Figura 1 apresenta o algoritmo de PG proposto que é baseado nos passos definidos por Suttasupa et al.⁽⁸⁾, conforme o Quadro 1:

Quadro 1 – Sequência usado pelo algoritmo de PG

1. Criar uma população randômica inicial. O conjunto de funções (como +, -, * e /) torna-se os nós internos da árvore. O conjunto terminal (as variáveis livres) torna-se nós externos (folha).
2. Evoluir a população de acordo com a função de aptidão. A aptidão indica quão bem cada indivíduo resolve o problema.
3. Criar a próxima geração da população pela evolução da sua estrutura da seguinte forma:
 - 3.1 Selecionar os melhores indivíduos da população e copiá-los para a próxima geração usando seleção por torneio (escolha aleatória ou determinística).
 - 3.2 Criar uma nova árvore pelo método *crossover*.
 - 3.3 Criar nova árvore pelo método de mutação.
4. Repetir os passos 2 e 3 até a solução ser encontrada ou a geração máxima seja alcançada.

A seguir são descritas algumas das principais características relacionadas ao algoritmo de PG desenvolvido, representado na Figura 1.

a) Parâmetros de Entrada

- Vetor com todos os instantes temporais (meses) usados no processo de treinamento e suas respectivas medidas antropométricas (peso, altura);
- Profundidade máxima das árvores de sintaxe abstrata que representarão os indivíduos (equações);
- Número total de indivíduos que irão compor a população;
- Condições de parada: número máximo de gerações ou valor de Erro Quadrático Médio modificado (EQMm) aceitável.

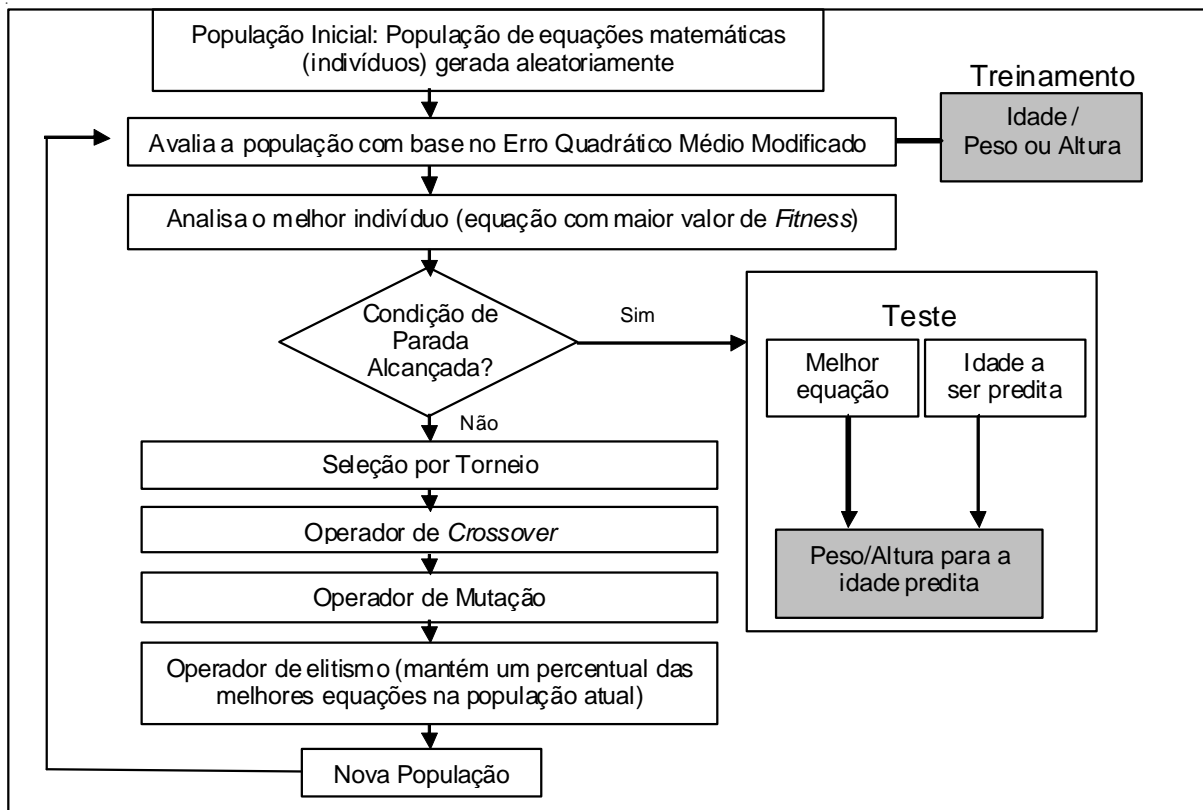


Figura 1 - Esquema do algoritmo de PG desenvolvido

b) Codificação

Os indivíduos são representados por árvores de sintaxe abstrata, compostas pela combinação livre de funções e terminais que são:

- Conjunto de Funções: +, -, /, *, ^, exp, log, sqrt, pow2 (potência de base 2).
- Conjunto de Terminais: t (Variável dependente que representa a unidade temporal), 1, 2, 3, 4, 5, 6, 7, 8, 9.

Para garantir somente soluções factíveis é imposta a condição de que a variável 't' deve estar presente em pelo menos um dos nós que compõem a árvore de sintaxe abstrata.

c) População Inicial

A população é iniciada com base no método *Ramped-half-and-half* (KOZA, 1992). A profundidade máxima das árvores geradas é definida pelo usuário a fim de evitar árvores muito grandes.

d) Avaliação da População (Fitness)

Para cada indivíduo (equação matemática), o valor de t é substituído pelo instante desejado e o resultado obtido corresponde à medida antropométrica (peso ou altura) para este instante específico. Esse processo é repetido com todos os instantes destinados para treinamento. Um EQMm é calculado considerando as saídas geradas pela expressão sugerida pela programação genética e os valores reais, nos quais há maior influência do erro identificado no instante mais recente considerado para treinamento. A Equação 1 foi utilizada para calcular o EQMm neste ponto.

$$EQMm = \frac{\sqrt{\left(\sum_{k=1}^n (Y_k - S_k)^2 \right) + (Y_n - S_n)^2}}{n}$$

<p>Onde: k = instante que está sendo analisado; Y = resultado da equação (indivíduo) para o instante em questão; S = medida antropométrica real para o instante em questão; n = número de instantes analisados;</p>	(1)
---	-----

Para contemplar a propriedade de fechamento (Ex.: divisão por 0), é imposta uma penalidade na função de avaliação como forma de tratar algumas das soluções infactíveis.

e) Processo de Seleção

O método aplicado corresponde ao torneio⁽²¹⁾. Assim, um conjunto composto por 30% dos indivíduos selecionados aleatoriamente da população é criado. O indivíduo deste grupo que possuir o melhor *fitness* (representado pelo menor EQMm) é selecionado para representar um possível pai.

f) Crossover

O *crossover* é baseado na escolha aleatória de um ponto de corte (nó) em cada pai selecionado (árvore). Depois de selecionado o ponto de corte, todos os componentes (funções e terminais) dos nós posteriores a este ponto são trocados.

g) Mutação

Diferente do operador de mutação padrão, a mutação considerada no algoritmo implementado está associada a um indivíduo ao invés de um gene. Portanto, a taxa de mutação adotada na fase de simulações teve que ser aumentada em relação aos valores comumente utilizados nestes algoritmos. Três diferentes tipos de mutação podem ser aplicados:

- Mutação Tipo 1: Cria uma nova sub-árvore para substituir a sub-árvore do nó selecionado.
- Mutação Tipo 2: Altera todas as constantes que

compõem a sub-árvore do nó selecionado.

- **Mutação Tipo 3:** Altera o valor do nó selecionado. Se o nó é um terminal, este será substituído por outro terminal. Se o nó é uma função, será substituído por outra função de mesma aridade.

Essas três formas de mutação mostram-se mais interessantes em diferentes estágios da evolução. Isto porque nas gerações iniciais, é importante explorar todo o espaço de busca para encontrar regiões promissoras. Assim, modificações maiores devem ser mais comuns nesta fase. Contudo, nas gerações finais, esperam-se modificações menores a fim de otimizar as possíveis soluções já encontradas. Dessa forma, inicialmente, mutações dos tipos 1, 2 e 3 são aplicadas com probabilidades de 0,5, 0,3 e 0,2, respectivamente. Ao longo da evolução, a taxa de mutação do tipo 1 diminui enquanto as taxas dos tipos 2 e 3 aumentam. Assim, no estágio final do processo evolutivo, a probabilidade do tipo 1 ocorrer é muito pequena, enquanto as probabilidades do tipo 2 e 3 tornam-se maiores (especialmente o tipo 2).

O uso de estratégias elitistas permite manter na população os melhores indivíduos, pois estes poderiam ser perdidos se não fossem selecionados de forma determinística para compor a próxima geração, ou então se fossem modificados por *crossover* ou mutação. Assim, no algoritmo desenvolvido, o operador de elitismo é empregado, a fim de trocar os 10% piores indivíduos da nova população pelos 10% melhores indivíduos da população original, gerando-se uma nova população composta pelos melhores de ambas.

RESULTADOS E DISCUSSÃO

Esta sessão tem como objetivo validar o desempenho

da abordagem proposta, através da realização de predições do desenvolvimento de crianças com diferentes registros históricos de peso e altura. O algoritmo de predição proposto foi aplicado no estudo de quatro crianças do sexo masculino, com até 12 meses de idade. Não foi levado em consideração se o desenvolvimento das crianças selecionadas poderia ser considerado normal ou não. Os dados referentes a peso e altura registrados para cada criança são apresentados nas Tabelas 1 e 2, agrupados em conjuntos de A a D. De cada conjunto de dados foram utilizados para treinamento do algoritmo de programação genética os primeiros registros, sendo que o último registro foi separado para teste. Assim, foram geradas predições para o último período de cada conjunto de dados e os resultados obtidos comparados com o registro existente.

Como se pode notar pelo constante nas tabelas 1 e 2, a distância temporal entre cada registro dos conjuntos de dados pode variar bastante, dificultando a utilização de métodos de predição tradicionais. Neste sentido, a proposta de um método de predição que seja capaz de se adaptar às informações disponíveis para cada criança mostra-se uma alternativa bastante atrativa. O método de predição proposto baseia-se em um mecanismo automático de definição da função de regressão. Este nível maior de autonomia foi alcançado por meio de um algoritmo de PG, descrito na sessão anterior.

Inicialmente foram realizados testes com o algoritmo de PG para identificação do melhor conjunto de parâmetros de entrada. Através da execução do algoritmo durante diversas rodadas variando o número máximo de gerações, número de indivíduos e profundidade das árvores de sintaxe abstrata, o conjunto de parâmetros que apresentou a menor média de EQMm foi identificado como melhor e aplicado para a realização das predições.

Tabela 1 – Conjunto de dados históricos sobre o peso de crianças do sexo masculino

Conjunto A		Conjunto B		Conjunto C		Conjunto D	
Período	Peso	Período	Peso	Período	Peso	Período	Peso
Nascimento	2.390 kg	Nascimento	4.000 kg	Nascimento	3.275 kg	Nascimento	3.735 kg
1º mês	2.540 kg	1º mês	4.100 kg	2º mês	4.740 kg	1º mês	5.050 kg
2º mês	4.570 kg	2º mês	5.380 kg	3º mês	5.540 kg	2º mês	6.780 kg
3º mês	5.180 kg	4º mês	6.640 kg	4º mês	6.120 kg	3º mês	7.500 kg
4º mês	5.820 kg	6º mês	7.980 kg	6º mês	7.420 kg	4º mês	7.870 kg
		9º mês	8.900 kg	9º mês	9.180 kg	5º mês	8.670 kg
		11º mês	9.700 kg	12º mês	10.180 kg	6º mês	9.000 kg
		12º mês	10.100 kg			7º mês	9.390 kg

Tabela 2 – Conjunto de dados históricos sobre a altura de crianças do sexo masculino

Conjunto A		Conjunto B		Conjunto C		Conjunto D	
Período	Altura	Período	Altura	Período	Altura	Período	Altura
Nascimento	47 cm	Nascimento	52 cm	Nascimento	51 cm	Nascimento	55 cm
1º mês	47 cm	1º mês	56 cm	1º mês	54 cm	1º mês	58 cm
2º mês	53 cm	2º mês	60 cm	2º mês	58 cm	2º mês	64 cm
3º mês	55 cm	4º mês	67 cm	3º mês	62 cm	3º mês	65 cm
4º mês	61 cm	6º mês	72 cm	6º mês	68 cm	4º mês	67 cm
		9º mês	77 cm	12º mês	80 cm	5º mês	70 cm
		11º mês	78 cm			6º mês	72 cm
		12º mês	81 cm			7º mês	73 cm
						8º mês	73 cm

Após a definição dos parâmetros de entrada, várias simulações com o algoritmo de PG foram realizadas para cada criança em estudo. Foram executadas 10 rodadas*, cada uma com um número máximo de 100 gerações, 50 indivíduos na população e profundidade máxima das árvores de sintaxe abstrata igual a 4. A taxa de *crossover* aplicada é de 0,6 e a taxa de mutação é de 0,3 para cada indivíduo.

A equação obtida como resultado da execução do algoritmo de PG, é composta por operadores diversos, como pode ser verificado na equação 2, originando a curva apresentada na Figura 2, Conjunto A. Essa equação é aplicada para 12 instantes de tempo (meses), tendo como resultado o peso médio da criança em cada mês. Assim, tem-se uma ideia de como poderá ocorrer o ganho de peso da criança em meses futuros, visto que para a criança em questão, existe registro de peso apenas até o seu quarto mês de vida. Essa mesma lógica é aplicada para gerar a predição do ganho de altura.

$$Y = \left(\sqrt{\left(2^{\left(\sqrt{\left(\sqrt{\left((x)^{(3)} \right)} \right)} \right)} \right)} \right) * \left(\sqrt{\left(2^{\left(\sqrt{\left(\sqrt{\left(2^x \right)} \right)} \right)} \right)} \right)$$

Onde:
 x = instante de tempo (mês) que está sendo analisado;
 Y = resultado da equação para o instante em questão (peso da criança);

(2)

As Figuras 2 e 3 apresentam os resultados obtidos pela regressão simbólica baseada em programação genética. Nessas figuras, os pontos utilizados para treinamento são representados pelo símbolo ‘o’ interligados por um traço, indicando os dados históricos

(Tabela 1 para Figura 1 e Tabela 2 para Figura 2) da criança em questão. O ponto utilizado para teste é representado pelo símbolo ‘*’ em cor vermelha e a curva gerada com base na função resultante do algoritmo de PG é formada por pontos representados pelo símbolo ‘*’ na cor azul.

Com os resultados apresentados nas Figuras 2 e 3 é possível verificar uma proximidade da curva gerada pela equação sugerida pelo algoritmo de PG com a curva de desenvolvimento registrada para as crianças analisadas, tanto no caso de predição do peso corporal quanto da altura das crianças. Percebe-se também que, em geral, a curva gerada torna-se mais próxima da curva de desenvolvimento nos últimos pontos de treinamento, o que indica que quanto maior o número de dados para treinamento, melhor a chance da curva se aproximar do ponto de teste, aumentando a precisão da predição.

O algoritmo de programação genética reduz a dependência de um especialista para definição da classe e coeficientes da função que prediz o desenvolvimento da criança. Além disso, a abordagem proposta trabalha com a não-linearidade existente nos dados de uma maneira mais eficiente, permitindo expressões simbólicas obtidas de forma flexível, e cuja previsão seja otimista (aumento dos valores registrados) ou pessimista (diminuição dos valores registrados).

CONCLUSÃO

O objetivo deste trabalho foi apresentar uma abordagem alternativa para realizar a predição de dados antropométricos de crianças. O método proposto baseia-se na comparação dos dados históricos de peso e altura

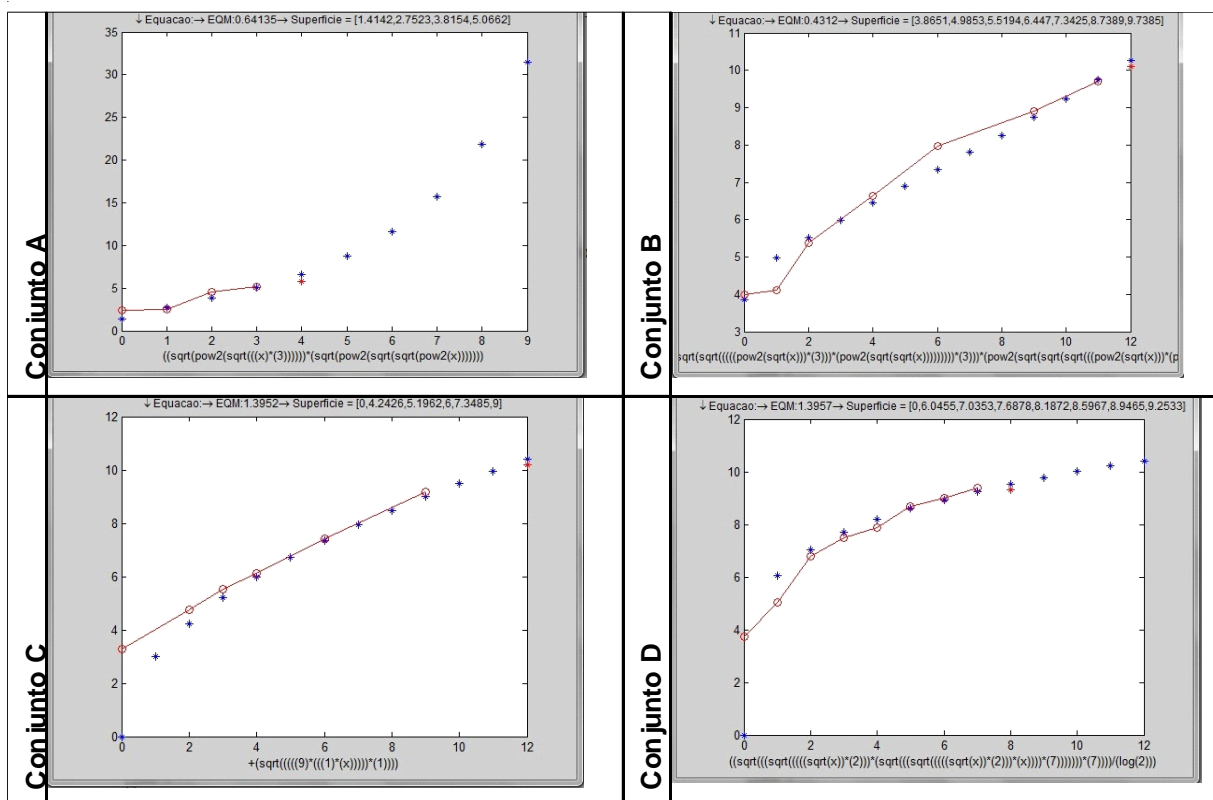


Figura 2 – Predição do peso corporal de crianças com base em sua idade.

* Cada rodada considera uma semente diferente na geração da população inicial.

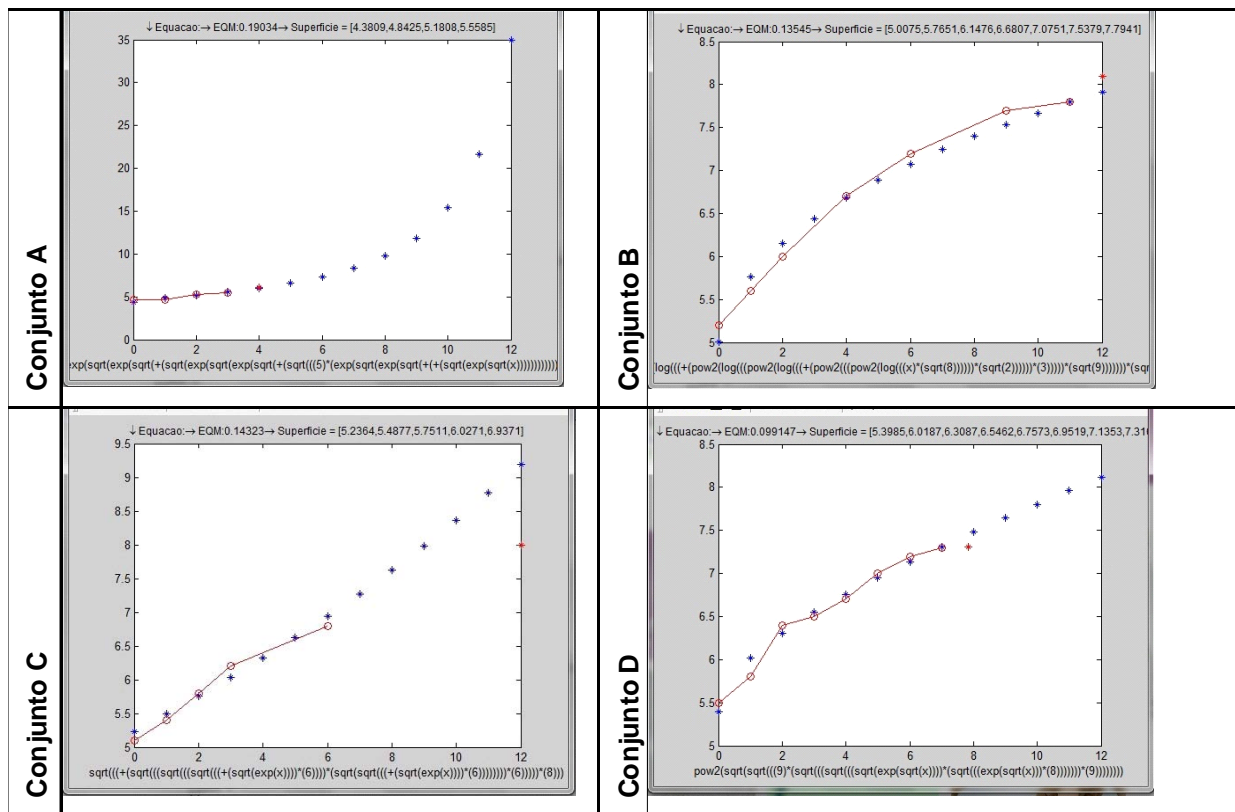


Figura 3 – Predição da altura de crianças com base em sua idade

de crianças em diferentes instantes de tempo (idades). Um algoritmo de programação genética foi aplicado sobre estes dados para analisar e prever o estado nutricional futuro dessas crianças.

Os resultados obtidos mostraram que o uso do algoritmo de programação genética é uma alternativa interessante, pois de um modo geral permitiu gerar previsões com resultados próximos ao valor real de peso ou de altura registrados para cada criança nos instantes de tempo desejados.

Contudo, considera-se reduzido o número de conjuntos de dados utilizados para as simulações, além do fato de ter sido trabalhado apenas com dados históricos de crianças do sexo masculino. Desta forma, a possibilidade de realizar simulações com dados de crianças de ambos os sexos e em maior número, incentiva a realização de trabalhos futuros.

Vale ressaltar que a abordagem em questão não é e não tem pretensão de ser uma modelagem exata. Ela retrata tendências de evoluções com base em informações do histórico da criança analisada e pode servir como auxílio na tomada de decisão por parte de especialistas da área da saúde.

Embora não seja uma modelagem exata, o algoritmo desenvolvido pode ser utilizado por profissionais da área de saúde no auxílio ao acompanhamento e avaliação do desenvolvimento de crianças, servindo como base para um sistema de apoio a decisão. Esse algoritmo poderia ser utilizado, por exemplo, em locais em que o acompanhamento de um médico pediatra não é constante, sendo o atendimento feito por clínicos gerais, médicos de outras especialidades ou mesmo outros profissionais da área de saúde, seja por motivos econômicos ou pela dificuldade de acesso desses profissionais ao local de atendimento, o que é comum em muitas regiões do Brasil. Assim, outros profissionais da área de saúde podem fazer o acompanhamento e encaminhar a criança para uma avaliação mais aprofundada, caso identifiquem que a curva projetada do desenvolvimento da mesma não se apresenta adequada.

O diferencial do algoritmo apresentado nesse trabalho é a possibilidade de prever dados sobre o processo de crescimento infantil, de forma a prever e prevenir possíveis doenças (ex. obesidade infantil e retardo no desenvolvimento) e não somente visualizar graficamente os dados já registrados para a mesma.

REFERÊNCIAS

1. Zeferino AMB. Acompanhamento do crescimento. J Pediat (Rio J). 2003;79(Sup1):S23-S32.
2. Fernandes JFA. Prática da avaliação física. Rio de Janeiro: Shape; 1999.
3. Vasconcelos FAG. Avaliação nutricional de coletividades. 3a. ed. Florianópolis: Editora UFSC; 2000.
4. WHO – World Health Organization. How much physical activity needed to improve and maintain health [Internet]. 2002 [citado 2012 ago 13]. Disponível em: www.who.int/hpt/physactiv/p.a.how.much.html
5. Araújo ACT, Campos JADB. Subsídios para a avaliação do estado nutricional de crianças e adolescentes por meio de indicadores antropométricos. Alim. Nutr, Araraquara. 2008;19(2):219-25.
6. CDC Growth Charts. Advance date from vital and health statistics. United States: National Center for Health Statistics [Internet]. 2012. Disponível em: www.cdc.gov
7. WHO Multicentre. WHO Multicentre Growth Reference Study Group. WHO child growth standards based on length/height, weight and age. Acta Paediatr Suppl. 2006; 450:76-85.

8. Suttasupa Y, Rungrazungsilp S, Pinyopan S, Wungchusunti P, Chongstitvatana PA. Comparative study of linear encoding in genetic programming. Proceedings of the Ninth International Conference on ICT and Knowledge; 2011 Nov 24-25; Bangkok, Thailand. p.13-7.
9. Koza JR. Genetic programming: on the programming of computers by means of natural selection. Massachusetts: MIT Press; 1992.
10. Cramer N. A representation for the adaptive generation of simple sequential programs. Proceedings of the International Conference on Genetic Algorithms and their Applications; 1985 Jul 24-26; Carnegie-Mellon University, Pittsburgh, (USA). p.183-7.
11. Holland JH. Adaptation in natural and artificial systems. Ann Arbor, MI: The University of Michigan Press; 1975.
12. Koza JR, Bennett FH, Andre D, Keane MA. Genetic Programming III: darwinian invention and problem solving. San Francisco: Morgan Kaufmann Publishers; 1999.
13. Koza JR. Hierarchical genetic algorithms operating on populations of computer programs. Proceedings of the Eleventh International Joint Conference on Artificial Intelligence; 1989 Aug 24-26; Detroit (USA). p. 768-74.
14. Koza JR. Genetic Programming II: automatic discovery of reusable programs. Massachusetts (USA):MIT Press; 1994.
15. Koza JR, Keane MA, Streeter MJ, Mydlowec W, Yu J, Lanza G. Genetic Programming IV: routine human-competitive machine intelligence. Massachusetts (USA): Kluwer Academic Publishers; 2003.
16. Garg AX, Adhikari NK, McDonald H, Rosas-Arellano MP, Devereaux PJ, Beyene J, Sam J, Haynes RB. Effects of computerized clinical decision support systems on practitioner performance and patient outcomes, a systematic review. JAMA. 2005; 293(10):1223-38.
17. Boland JJ. Forecasting water use: A tutorial. In: Torno HC. Computer applications in water resources. Nova York: The society; 1985. p.907-16.
18. Savi MA. Dinâmica não-linear e caos. Rio de Janeiro: E-papers Serviços Editoriais Ltda; 2006.
19. Takashina NT. Era da gestão da sabedoria: modismo ou tendência da qualidade para o próximo milênio? [Internet]. 2002 [citado 2012 jul 10]. Disponível em: www.kmpress.com.br/portal/artigos/preview.asp?id=184
20. Teixeira AS. Democracia e educação. Educação e o mundo moderno. São Paulo: Companhia Editora Nacional; 1969.
21. Goldberg DE, Deb K. A comparative analysis of selection schemes used in genetic algorithms. Foundations of Genetic Algorithms (FOGA). San Francisco: Morgan Kaufmann Publishers; 1991.