

# Preditor de números para lotofácil: uma abordagem usando algoritmos evolutivos

Fernando Barbosa<sup>1</sup>, Lais C. R. S. Lopes<sup>1</sup>, Mikaely Cruciol<sup>1</sup>

<sup>1</sup>Instituto Federal Goiano – Campus Urutaí – Departamento de Informática

Rodovia Geraldo Nascimento, Km 2,5 – Urutaí – GO – Brasil

{nando.cefet, laiscandidolopes, mikaelycruciol42}@gmail.com

**Abstract.** *Genetic algorithms (GA's) are the most widely used and studied evolutionary computing techniques. Based on the theory of evolution proposed by Darwin, has a powerful process of adaptation to certain types of problems, including global and local searches. This article aims, based on genetic algorithms, develop software predictor of numbers for a game of chance very popular in Brazil, Lotofácil. Maintained by the federal government this game reaches all sections of the population of the country.*

**Resumo.** *Os algoritmos genéticos (AG's) são as mais difundidas e estudadas técnicas de computação evolucionária. Baseado na teoria da evolução proposta por Darwin, possui um poderoso processo de adaptação a certos tipos de problema, inclusive buscas locais e globais. O presente artigo pretende, baseado em algoritmos genéticos, desenvolver um software preditor de números para um jogo de azar muito popular no Brasil, Lotofácil. Mantido pelo governo federal este jogo atinge a todas as camadas populacionais do país.*

## 1. Introdução

Os jogos de azar no Brasil são datados desde de 1784, o primeiro jogo que se tem registro teve o intuito de levantar verbas para a construção de obras públicas [Monteiro 2007]. Desde então essa modalidade de jogo caiu no gosto popular, além de representar uma forma de arrecadação para o governo federal.

A Caixa Econômica Federal é o órgão responsável por administrar as loterias federais no Brasil. Ela é mantenedora de nove jogos de azar, são eles, a Mega-Sena, Dupla-Sena, Quina, Lotéca, Lotogol, Lotomania, Timemanía, Loteria federal e a Lotofácil. É com base no último jogo, Lotofácil, que o presente artigo propõe um preditor de números. A cartela da Lotofácil é composta por 25 números, onde o jogador pode marcar de 15 a 18 números, e ganhar caso acerte de 11 a 15 dezenas. O valor do prêmio máximo varia de acordo com o total arrecadado com a venda das cartelas, mais especificamente 46% do valor bruto arrecadado, sendo ainda deduzido deste os valores dos prêmios fixos, R\$2,50 para 11 acertos, R\$5,00 para 12 acertos e R\$12,50 para 13 acertos.

Todo tipo de jogo, seja ele de tabuleiro ou computacional, possui uma problemática, e uma forma de se resolver tais problemas é por meio de heurísticas. As heurísticas são métodos de investigação e solução de problemas baseados na busca algorítmica para

encontrar o ótimo ou uma solução que satisfaça determinados pré-requisitos. Elas procuram usar procedimentos, estratégias, aproximação, sempre na procura da melhor forma de se chegar a um determinado fim, garantindo assim soluções eficientes.

Uma heurística muito conhecida é o Algoritmo Genético (AG), inspirado na teoria de Charles Darwin, o qual afirma que mecanismos de evolução é uma competição que seleciona os indivíduos mais bem adaptados em seu ambiente, transmitindo características aos seus descendentes, permitindo sua sobrevivência. Os AGs são técnicas de busca para encontrar melhores soluções em problemas de otimização através da evolução de populações de soluções de cromossomos artificiais. Um cromossomo ou indivíduo é uma estrutura de dados que contém as possíveis soluções de determinado problema. Os cromossomos são submetidos a um processo evolucionário que envolve avaliação, seleção, cruzamento e mutação. Após vários ciclos de evolução, a população final é composta por indivíduos mais aptos [Pacheco 1999].

O presente artigo propõe um software preditor de números para Lotofácil. Vale ressaltar que é impossível prever com precisão quais números serão sorteados, pois, trata-se de um jogo com problemática de aleatoriedade. Quando a análise combinatória é empregada a esse problema tem-se que, a chance de alguém jogando 15 números acertar todos é de 3.268.760 para uma.

Este trabalho está estruturado da seguinte forma: metodologia, sessão em que será tratada a forma como o software foi confeccionado, bem como ferramentas utilizadas, nesta há ainda uma sub-sessão intitulada condução do experimento, que abordará a sistemática usada para realizar os testes. Resultados, sessão que demonstrará os resultados obtidos e por último tem-se a conclusão, sessão encerratória apresentado percepções finais acerca do software.

## 2. Metodologia

Todos os algoritmos e fórmulas, presentes no software, foram implementados por meio da plataforma ®Java<sup>1</sup>, usando-se o paradigma de orientação a objetos [Cardoso 2006].

Dentre a grande variedade de técnicas que anunciam sua eficiência em prever os números da Loteria Lotofácil, foi escolhida uma conhecida como *técnica das catorze dezenas*<sup>2</sup>. Nesta as nove dezenas que mais saíram nos dez últimos jogos são selecionadas e reservadas. Posteriormente cinco dezenas são escolhidas, estas cinco não devem ter sido sorteadas no último jogo e nem estar entre o conjunto das nove dezenas, já separado. Juntando-se as nove mais saídas com as cinco escolhidas totaliza-se catorze dezenas distintas que serão fixas.

As onze dezenas restantes para completar os vinte e cinco números da cartela da Lotofácil, chamadas de dezenas variáveis, são concatenadas com as quatorze fixas. Dessa forma são gerados onze jogos diferentes, todos compostos pelas quatorze dezenas fixas e mais uma dezena das onze variáveis. É mostrado na Figura 1 um exemplo numérico de forma resumida.

Implementou-se a presente técnica para gerar onze jogos, e em seguida um algoritmo genético (AG) capaz de gerar uma população de jogos, cuja a distância fosse a

<sup>1</sup> Sítio oficial da linguagem java: <http://www.oracle.com/technetwork/java/index.html>

<sup>2</sup> Técnica das catorze dezenas, disponível no sítio: <http://www.loteriaspremium.com>

- Nove dezenas mais sorteadas:  
1 - 8 - 12 - 15 - 20 - 22 - 23 - 24 - 25
- Cinco dezenas não sorteadas no último sorteio:  
2 - 5 - 6 - 18 - 19
- Onze dezenas restantes:  
3 - 4 - 7 - 9 - 10 - 11 - 13 - 14 - 16 - 17 - 21
- Jogos gerados:  
1 8 12 15 20 22 23 24 25 2 5 6 18 19 3  
1 8 12 15 20 22 23 24 25 2 5 6 18 19 4  
...  
1 8 12 15 20 22 23 24 25 2 5 6 18 19 21

Figura 1. Exemplo de implementação da técnica das 14 dezenas

menor possível da citada técnica.

Para o algoritmo genético (AG), foi gerada uma população inicial com quinhentos indivíduos, cada indivíduo ou cromossomo representa um jogo com quinze dezenas e mais o valor *fitness*, logo, uma matriz de quinhentas linhas por dezesseis colunas foi gerada para representar a população inicial. Cada cromossomo foi gerado com números aleatórios, respeitando o intervalo de 1 a 25 (números disponíveis na cartela da Lotofácil), tomou-se ainda o cuidado para que nenhum número do indivíduo fosse repetido. Na Figura 2 é demonstrada a representação de um indivíduo.



Figura 2. Representação de um cromossomo

A função de aptidão, considerada como o segundo passo de um AG, consiste em atribuir nota a um determinado cromossomo ou indivíduo. É com base nela que os cromossomos são selecionados e cruzados, a essa nota ou valor dá-se o nome de *fitness*. Como função de aptidão foi escolhido comparar as dezenas encontradas entre as duas técnicas. Para tanto, cada cromossomo foi comparado com os onze jogos gerados pela *técnica das quatorze dezenas*. Jogo a jogo as quinze dezenas entre os dois jogos foram confrontadas, toda vez que uma dezena presente no cromossomo estava também presente em um jogo gerado pela *técnica das quatorze dezenas* o valor do *fitness* aumentava em um ponto gerando um somatório. Quanto maior o *fitness* mais apto o indivíduo estava para 'adivinhar' os números da Lotofácil baseado na *técnica das quatorze dezenas*.

O cruzamento, terceiro passo do AG, é utilizado para recombinar material genético de dois pais previamente selecionados. Para selecioná-los foi utilizado o método de seleção por torneio [Lacerda and P. 2003], que consiste em atribuir um duelo entre dois ou mais pais prevalecendo o mais forte. O torneio foi aplicado entre dois pais escolhidos de forma aleatória, foram realizados duzentos e cinquenta duelos, metade da população inicial. Como de cada torneio um pai prevalecia, formou-se uma população de duzentos e cinquenta pais para participarem do cruzamento.

Para o cruzamento ou *crossover* os pares de pais selecionados foram escolhidos de forma aleatória. O operador de *crossover* aplicado foi o crossover de 1 ponto

[Lacerda and P. 2003], onde um ponto aleatório é fixado no filho sendo preenchido com genes de um pai até o ponto especificado e genes do outro pai após o ponto. De cada cruzamento foram gerados dois filhos. O processo de mutação, quarto passo do AG, serve para assegurar alguma diversidade entre os indivíduos. A mutação foi feita selecionando de forma aleatória dois genes em um determinado cromossomo e trocando seu material genético de posição, dessa forma testificava-se que não haveria inconsistência nos cromossomos, como números repetidos por exemplo. A fim de fazer um sistema mais dinâmico e flexível as taxas de mutação e quantidade de gerações ficaram a cargo do usuário, o que possibilita maior gama e formas de realização de testes, além de interação destes com o sistema.

## 2.1. Condução do experimento

Para gerar resultados legítimos, foi feito a descarga de todos os resultados existentes da Lotofácil, serviço oferecido pelo órgão mantenedor das loterias no Brasil, presente no sítio oficial do banco da Caixa Econômica <sup>3</sup>.

Com esses resultados em mãos, foram selecionados, de forma aleatória, onze jogos sequenciais, dez serviram de entrada para o software preditor que gerou uma população ordenada de forma decrescente pelo seu *fitness*. O jogo que sobrou dos onze selecionados foi usado para verificar a quantidade de dezenas acertadas.

O indivíduo mais apto gerado pelo preditor foi comparado com o jogo reservado para realizar a verificação. A quantificação das dezenas acertadas foi feita de forma manual, onde cada dezena do indivíduo presente no jogo reservado para a verificação correspondia a um acerto. Todo o processo foi fielmente repetido quinze vezes, mudando apenas os dez jogos inseridos no sistema e o seu respectivo jogo usado como resposta. Na primeira execução dos testes, os jogos numerados de um a dez serviram de entrada para o sistema e o jogo de número onze foi usado com o intuito de realizar a verificação, dando sequência na segunda execução foram selecionados os jogos de número doze ao vinte e um, sendo o de número vinte e dois usado para a verificação e assim por diante.

Para a condução dos experimentos os parâmetros variáveis do AG foram setados de forma igual para todas as execuções, seus valores podem ser vistos na Tabela 1:

Tabela 1. Valores dos parâmetros utilizados para o preditor

Parâmetro	Valor
<i>pop_inicial</i>	500
<i>qtd_geracao</i>	200
<i>qtd_torneio</i>	250
<i>tx_mutacao</i>	5%

## 3. Resultados

Apesar de o foco do presente trabalho não tenha sido o de fazer um comparativo entre o preditor utilizando a *técnica das quatorze dezenas* e o preditor utilizando um AG, essa comparação foi feita, mas apenas com o intuito de enriquecer os resultados apresentados. Os resultados obtidos pela condução dos experimentos são mostrados na Tabela 2. Na

<sup>3</sup>Sítio oficial do banco Caixa: <http://www.caixa.gov.br/>

primeira coluna é exibido o número da execução, na segunda o número do concurso da Lotofácil utilizado para validar o desempenho, na terceira coluna é demonstrado a quantidade de dezenas acertadas pelo preditor utilizando um AG e na quarta e última coluna é exposto o desempenho do preditor aplicando a técnica das catorze dezenas somente, sem a utilização do AG. Nota-se que para gerar os dois resultados um mesmo jogo foi utilizado como validador.

Os resultados obtidos demonstraram que a técnica das catorze dezenas por si só obteve acertos inferiores se comparados aos acertos encontrados pelo AG. Embora o AG não tenha conseguido acertar os quinze números em nenhuma das quinze execuções, ele demonstrou resultados significativos sendo seu menor índice de acertos de cinquenta e três por cento e seu maior de noventa e três por cento. A média de dezenas acertadas pelo AG nas quinze execuções foi de 10,4 acertos por jogo o que equivale a 69,3%, enquanto que a média da técnica das 14 dezenas foi de 8,6 acertos por jogo, logo 57,3% .

Percebe-se que, o preditor quando usando a técnica das 14 dezenas não conseguiu em nenhuma das execuções acertar mais dezenas do que quando o preditor implementava o AG.

**Tabela 2. Resultados obtidos nas execuções do preditor**

execução	concurso	acertos - AG	acertos - s/ AG
1	11	11	10
2	22	10	08
3	33	09	09
4	44	10	08
5	55	09	09
6	66	10	10
7	77	10	06
8	88	14	08
9	99	10	10
10	110	08	08
11	121	13	12
12	132	11	10
13	143	10	10
14	154	11	07
15	165	10	06

#### 4. Conclusão

O algoritmo genético se mostrou um preditor no mínimo interessante. Embora não tenha conseguido, durante a execução dos testes, predizer as quinze dezenas de algum jogo obteve uma média de acertos consideravelmente boa, em torno de 10 dezenas por jogo.

É provável também que obtivesse resultados mais promissores se a quantidade de testes fosse aumentada, pois das quinze execuções realizadas, em uma o preditor utilizando o AG conseguiu prever 14 das 15 dezenas sorteadas. Vale lembrar que a probabilidade de se atingir 14 dezenas em um jogo da Lotofácil é de 21.791 para um, empregando o preditor esse número caiu para 8 tentativas. Seguindo esses resultados é de se esperar que,

ao realizar uma maior quantidade de teste, o preditor consiga acertar as quinze dezenas com uma probabilidade menor do que a existente, 3.268.760 para uma

Outro fator que pode ser levado em conta para melhorar ou piorar os resultados são as configurações dos parâmetros livres do algoritmo genético. Uma mudança na quantidade de gerações ou de indivíduos por população por exemplo poderia gerar resultados diferentes dos apresentados.

Como trabalhos futuros pretende-se desenvolver preditores aplicando outras técnicas bio inspiradas, ou mesmo, outro AG com técnicas de seleção, como o elitismo [Ricardo 2008] e o método da roleta [Ricardo 2008], para ao final fazer um comparativo e assim eleger o que melhor se adapta ao problema de prever os quinze números sorteados pela Lotofácil. Pretende-se ainda realizar estudos sobre análise combinatória e probabilidade para encontrar outras técnicas matemáticas válidas e assim gerar outras funções de aptidão.

## Referências

- Cardoso, C. (2006). *Orientação a objetos na prática: aprendendo orientação a objetos com Java*. Ciência Moderna, 1th edition.
- Lacerda, E. G. M. and P., C. A. C. (2003). Introdução aos algoritmos genéticos.
- Monteiro, C. (2007). Howstuffworks - como funcionam as loterias. Disponível em <http://lazer.hsw.oul.com.br/loterias-brasil2.htm>. Acessado em 25 de janeiro de 2013.
- Pacheco, M. A. C. (1999). Algoritmos genéticos: Princípios e aplicações. *ICA*.
- Ricardo, L. (2008). *Algoritmos genéticos: uma importante ferramenta da inteligência computacional*. Brassport, 2th edition.