

DEPARTAMENTO DE ENGENHARIA ELÉTRICA

Business Intelligence Master - Sistemas Inteligentes de Apoio à
Decisão em Negócios

Avaliação de Estratégias de Investimento (no mercado de
ações) com Indicadores Otimizados por Algoritmos Genéticos

Fabio Bruno Pereira

Orientador:

Juan Guillermo Lazo Lazo

Agradecimentos

Agradeço a minha família e todos que me apoiaram e me encorajaram a explorar o mundo de oportunidades que o BI (Business Intelligence) nos abre. Em especial ao meu orientador pelo suporte e dedicação.

Resumo

Esta monografia tem por objetivo estudar e propor um modelo que defina uma estratégia de investimento com base ao estudo do desempenho de alguns indicadores tradicionais de análise técnica do mercado de ações.

Apesar desses indicadores possuírem parâmetros padrões na literatura, entendemos que, de acordo com o mercado de atuação e o Ativo que está sendo avaliado, podem existir outros parâmetros para o indicador analisado que tragam um desempenho superior ao nível de retorno do investimento.

Com o objetivo de estudar o desempenho da estratégia de investimento baseada nesses indicadores com parâmetros otimizados. Para este trabalho foi selecionado um ativo específico (Petr4 – Petrobras PN), negociado na Bolsa de Valores de São Paulo BM&FBOVESPA, foram selecionados dois indicadores para serem analisados: o Moving Average Convergence / Divergence – MACD (Convergência e Divergência de Médias Móveis) e o Índice de Força Relativa - IFR e para encontrar os parâmetros ótimos foi utilizada a técnica de algoritmos genéticos. Além da avaliação individual dos indicadores, foram desenvolvidas algumas estratégias de investimento combinando os indicadores e otimizando os seus parâmetros.

Foi Verificado que o desempenho dos indicadores com parâmetros otimizados foi muito superior ao dos indicadores com parâmetros padrões. Comprova-se também que estratégias combinadas e otimizadas podem ter um desempenho melhor ainda que indicadores otimizados individualmente.

Abstract

This monograph aims to study the performance of investment strategies using traditional indicators as the basis for technical analysis of stock market.

Although traditional indicators have default parameters in the literature, we understand that, according to market performance and the asset being evaluated, there may be other parameters that bring superior return of investment performance.

In order to study the performance of these indicators with optimized parameters, we selected a specific asset (PETR4 - Petrobras PN), traded on the Sao Paulo Stock Exchange BOVESPA, for this work. We selected two indicators to be analyzed: the Moving Average Convergence / Divergence - MACD and Relative Strength Index – IFR. As to find the optimal parameters we used the genetic algorithms techniques. Besides the evaluation of individual indicators, some strategies have been developed by combining indicators.

We concluded that indicators with optimized parameters have higher performance than indicators with default parameters. It also proves that strategies optimized combining indicators can have a better performance than indicators optimized individually.

Sumário

| | |
|--|--------------------------------------|
| 1. INTRODUÇÃO | 6 |
| 1.1 DELIMITAÇÃO | 7 |
| 1.2 OBJETIVOS DO TRABALHO..... | 8 |
| 1.3 DESCRIÇÃO DO TRABALHO..... | 8 |
| 1.4 ORGANIZAÇÃO DA MONOGRAFIA | 9 |
| 2. DESCRIÇÃO DO PROBLEMA | 10 |
| 2.1 PROBLEMA ESTUDADO | 10 |
| 3. METODOLOGIA | ERRO! INDICADOR NÃO DEFINIDO. |
| 3.1 INDICADORES DA ANÁLISE TÉCNICA..... | 12 |
| 3.1.1 MACD - (<i>Convergência e Divergência de Médias Móveis</i>)..... | 13 |
| 3.1.2 IFR - <i>Índice de Força Relativa</i> | 15 |
| 3.2 ALGORITMOS GENÉTICOS..... | 17 |
| 3.2.1 <i>Representação</i> | 20 |
| 3.2.2 <i>Codificação e Decodificação</i> | 20 |
| 3.2.3 <i>Avaliação</i> | 20 |
| 3.2.4 <i>Operadores Genéticos</i> | 21 |
| 3.2.5 <i>Parâmetros da Evolução</i> | 23 |
| 4. ARQUITETURA DO SISTEMA PROPOSTO..... | 25 |
| 4.1 CONSIDERAÇÕES..... | 25 |
| 4.2 MODELO PARA OTIMIZAÇÃO DOS PARÂMETROS DO MACD..... | 26 |
| 4.3 MODELO PARA OTIMIZAÇÃO DOS PARÂMETROS DO IFR | 30 |
| 4.4 ESTRATÉGIAS DE INVESTIMENTO | 32 |
| 5 RESULTADOS..... | 34 |
| 5.1 ESTRATÉGIA COM INDICADOR MACD OTIMIZADO | 34 |
| 5.2 ESTRATÉGIA COM INDICADOR IFR OTIMIZADO..... | 37 |
| 5.3 ESTRATÉGIA MACD E IFR INDEPENDENTES | 42 |
| 5.3 ESTRATÉGIA MACD E IFR DEPENDENTES..... | 43 |
| 6. CONCLUSÃO E TRABALHOS FUTUROS | 44 |
| 7. REFERÊNCIA BIBLIOGRÁFICA..... | 46 |

1. Introdução

O mercado brasileiro de ações movimenta bilhões de reais por dia, onde novas empresas entram no mercado e outras saem, atraindo diferentes tipos de investidores todos os dias. O mercado de capital é extremamente volátil e complexo, o valor de mercado de uma empresa pode oscilar de acordo com uma diversidade muito grande de fatores endógenos e exógenos à empresa.

Atualmente temos duas linhas de estudos importantes para o estudo e análise de investimentos em ações, denominadas elas: análise fundamentalista e análise técnica.

A análise fundamentalista é uma linha de estudo que reúne conceitos e técnicas que buscam vincular o valor real de uma empresa ao preço de sua ação, se utilizando da análise de balanços, interpretação de indicadores macroeconômicos, notícias, pautas de reuniões de Comitê, relatórios de auditoria entre outras fontes de informação. O analista fundamentalista acompanha o dia a dia da empresa buscando a todo momento equiparar a sua capacidade de gerar lucro com o valor comercializado de suas ações.

Enquanto a outra linha acredita que os preços dos ativos se ajustam ao conteúdo das informações, porém não de forma imediata. O ajuste de preços percorre uma série de etapas, dependendo da forma de comportamento dos grupos dos com informações privilegiadas no processo de difusão das informações relevantes. Por isso, este grupo analisa gráficos de preços e indicadores para estudar o comportamento histórico das ações buscando obter informações dos padrões de dados onde seja possível inferir o comportamento futuro das ações analisadas.

Esse projeto tem como objetivo estudar o desempenho de indicadores, utilizados pela corrente de Análise Técnica, otimizados por algoritmos genéticos. Desejamos não só, avaliar o desempenho de indicadores otimizados para um determinado ativo ou indústria, mas também verificar a aplicação e performance de algoritmos genéticos em indicadores financeiros.

Para alcançar tal objetivo, vamos explorar nessa monografia dois indicadores muito conhecidos e utilizados na análise técnica, o Moving Average Convergence / Divergence – MACD (Convergência e Divergência de Médias Móveis) e o Índice de Força Relativa – IFR. Adicionalmente utilizaremos técnicas de otimização por algoritmo genético para tentar melhorar o desempenho desses indicadores.

Nesse projeto desenvolvemos algumas estratégias possíveis de investimento utilizando os indicadores estudados. A fim de manter os resultados comparáveis realizamos todo o estudo em apenas um ativo financeiro (PETR4). Esse ativo foi selecionado pela sua popularidade e estabilidade da sua empresa (Petrobras), o que facilita o entendimento da sua série histórica e análise.

1.1 Delimitação

Muitos investidores utilizam a análise técnica como principal ferramenta de previsão e análise de tendências. Implementações dos indicadores estudados nesse projeto (MACD e IFR) estão disponíveis gratuitamente na internet e oferecido pela maioria das corretoras de valores. No entanto, a maior parte desses indicadores fornecidos aos investidores trabalham com os parâmetros padrões independentemente do Ativo analisado. Acreditamos que para cada mercado, indústria ou até mesmo Ativo, possa existir parâmetros específicos que melhor representem o seu comportamento, isso porque o comportamento de um ativo é muito diferente ao de outro, inclusive quando os ativos pertencem ao mesmo setor.

Existem indicadores parametrizáveis, onde o analista pode escolher os parâmetros que melhor se adequem a sua análise. Por tanto, mesmo para esses indicadores parametrizáveis, a dificuldade está em determinar quais os melhores parâmetros que podemos utilizar. Perguntas frequentes dos que se aventuram por essas alternativas são: será que é vantajoso utilizar parâmetros personalizados? Qual a diferença de desempenho? Qual o esforço necessário para selecionar os parâmetros mais apropriados?

Nesse trabalho buscamos algumas dessas respostas para um conjunto restrito de indicadores, testados sobre um ativo escolhido e com o auxílio da ferramenta (Plug-in para Excel) Evolver que implementa um algoritmo genético.

1.2 Objetivos do trabalho

Este trabalho tem como objetivo:

- Estudar o desempenho dos indicadores MACD e IFR utilizados na análise técnica de Ativos financeiros comercializados nas Bolsas de Valores
- Estudar a performance dos indicadores MACD e IFR com parâmetros otimizados com algoritmos genéticos e comparar os resultados com o desempenho dos mesmos indicadores utilizando os parâmetros padrões.
- Elaborar estratégias de investimentos utilizando a combinação dos indicadores estudados (MACD e IFR) com os parâmetros otimizados para essa estratégia.
- Avaliar as estratégias propostas apresentando os resultados obtidos, pontos positivos e negativos
- Estimular discussões sobre o assunto e incentivar trabalhos futuros

1.3 Descrição do trabalho

Neste trabalho estudamos o comportamento e desempenho de indicadores de análise técnica: o Moving Average Convergence / Divergence – MACD (Convergência e Divergência de Médias Móveis) e o Índice de Força Relativa – IFR e buscamos que otimizar individualmente os seus parâmetros de cada indicador forma que o desempenho em quanto a rentabilidade seja maximizado. Neste trabalho também foram criadas estratégias de investimento usando esses indicadores otimizando os seus parâmetros de forma conjunta. Os resultados deste trabalho foram comparados com os resultados obtidos pelos indicadores usando o valor padrões em seus parâmetros.

1.4 Organização da monografia

Esta monografia está organizada conforme descrito a seguir. No capítulo 2, será apresentado um resumo do problema estudado. O capítulo 3, um resumo sobre a teoria de Algoritmos Genéticos e o detalhamento e explicação dos indicadores da análise técnica utilizados. O capítulo 4 apresenta a descrição da metodologia e estratégias de investimentos desenvolvida, no capítulo 5 apresenta a o estudo de casos com os experimentos e resultados do trabalho e no capítulo 6 são apresentadas as conclusões e trabalhos futuros.

2. Descrição do problema

2.1 Problema estudado

A Análise Técnica de investimento se baseia fortemente no estudo e na utilização de indicadores que baseiam-se de dados históricos como principal fonte de informação. A maioria dos dados utilizados pelos indicadores tradicionais é de domínio público e podem ser facilmente encontrados na internet. As fórmulas utilizadas pelos indicadores e significado também são públicas e podem ser encontradas facilmente. O problema começa quando precisamos saber se aquele indicador é “confiável” ou não para um determinado ativo ou grupo de ativos. Adicionalmente, gostaríamos de saber quanto preciso é o indicador que estamos analisando e suas fraquezas. Para tentar responder essas perguntas, dividimos o problema em quatro partes e assumimos algumas premissas:

1 – Determinar o ativo ou grupo a ser estudado. Esse passo é importante pois entendemos que o desempenho e confiabilidade de um indicador (ou estratégia de investimento) podem mudar de acordo com o comportamento do ativo. Para esse trabalho foi escolhido o Ativo PETR4, ações referenciais da empresa Petrobras, negociado na Bolsa de valores de São Paulo (Bovespa). Dessa forma os estudos realizados e conclusões obtidas poderão, e possivelmente serão diferentes se aplicados a ativos diferentes ou a épocas muito distantes da atual.

2 – Definir o “arsenal” que será utilizado, isto é, quais os indicadores serão escolhidos para compor a estratégia de investimento e guiar a tomada de decisão. Para esse trabalho foram selecionados o indicador Moving Average Convergence / Divergence – MACD (Convergência e Divergência de Médias Móveis) e o Índice de Força Relativa – IFR. Esses indicadores serão detalhados no capítulo 3.

3 – Elaborar estratégias de investimento. Após ter selecionado o Ativo e os indicadores, está na hora de se definir algumas possíveis estratégias de investimento. Estas estratégias estão relacionadas a como vão-se utilizar os indicadores. Se cada indicador é utilizado separadamente, em conjunto ou

sobrepostos. Definir quais as posições¹ serão operadas. As estratégias adotadas serão detalhadas no capítulo 4 dessa monografia.

4 – Medir e avaliar o desempenho dos indicadores e estratégias definidas. A última parte do problema consistem em medir os resultados apresentados pelos indicadores e estratégias, tentar otimizar os ganhos identificando os melhores parâmetros para cada indicador ou estratégia.

Os ganhos são medidos em percentuais, e como premissa não consideramos valor do aluguel de ações, taxas, impostos e tributos incidentes, custo de oportunidade, disponibilidade financeira, dividendos, juros, inflação, entre outras variáveis econômicas e financeiras.

O Objetivo de simplificar o modelo é permitir a avaliação da performance dos indicadores e estratégias e não de investimento. Este trabalho não busca apresentar ou sugerir investimentos aos leitores, assim como não assegura os resultados financeiros apresentados.

Cada um dos quatro passos acima apresenta novos problemas e desafios ao longo do projeto que serão detalhados ao longo dos capítulos três e quatro. Uma vez realizados os passos acima podemos saber quanto “confiável” é nossa estratégia e ter uma ideia da performance esperada.

¹ Posição Comprado (onde precisamos comprar uma ação primeiro para vender posteriormente) ou Vendido (Aluga-se uma ação e vende a ação para recomprá-la posteriormente por um valor inferior).

3. Metodologia

Dada a problemática descrita no item anterior, o primeiro passo foi determinar o ativo ou grupo a ser estudado. Optou-se por trabalhar com apenas um ativo ao invés de um grupo. Essa escolha se deu pela premissa de que o comportamento de uma única ação é mais facilmente decifrado e explicável do que de um grupo. E tendo um comportamento mais compreensível podemos identificar pontos fracos e fortes dos modelos e algoritmos testados.

Uma vez definido o ativo a ser estudado, iniciamos o estudo dos indicadores de Análise Técnica para selecionar as ferramentas que irão servir de base para a tomada de decisão do investimento.

3.1 Indicadores da análise técnica

A análise técnica ou análise gráfica, de maneira resumida, é uma abordagem que utiliza gráficos como principal ferramenta para determinar o melhor momento e preço para comprar ou vender ativos. A análise técnica tem como premissa que no preço do ativo encontra-se toda a informação necessária para entender o seu comportamento, assim são construídos gráficos como histórico do preço do ativo. Além da utilização de gráficos, a análise técnica também inclui diversas teorias sobre como ocorrem os movimentos do mercado.

A Análise Técnica veio ganhando espaço no século XX conforme os avanços tecnológicos foram ocorrendo. Estes avanços proporcionaram uma visão mais clara dos gráficos e da utilização de indicadores técnicos, ao passo que computadores velozes facilitaram os cálculos matemáticos e reduzem o tempo de resposta para a tomada de decisão.

A comprovação da eficiência das ferramentas técnicas no processo de tomada de decisão no mercado acionário para identificar tendências e pontos de entradas foi o bastante para difundir a análise técnica pelo mundo, principalmente entre os *traders* (comerciantes das Bolsas de Valores).

Com base nas médias móveis e outros conceitos básicos de estatística, operadores de mercado, em especial os norte-americanos, criaram mecanismos para identificar tendências e possíveis situações de entrada e saída.

A partir da década de 60, foram desenvolvidos na Europa e nos EUA diversos estudos para comprovar a eficiência da análise técnica e de seus indicadores. A partir destes experimentos, uma gama de indicadores técnicos foi criada, e muitas dessas ferramentas são utilizadas até hoje.

Como já foi mencionado, foram selecionados dois indicadores (MACD e IFR) para realizar os estudos. Em seguida apresentamos as metodologias utilizadas para cada um desses indicadores.

3.1.1 MACD - (Convergência e Divergência de Médias Móveis)

Dentre os diversos indicadores estatísticos utilizados pelos analistas técnicos, o MACD é um dos mais populares e confiáveis dos indicadores de análise técnica. Criado na década de 60 por Gerald Appel, consiste em um oscilador que pode revelar pontos de reversão de tendência do mercado e também possibilita a identificação visual do ativo "comprado" ou "vendido", facilitando a tomada de decisão para o momento certo de entrar ou sair do mercado.

O MACD (do inglês Moving Average Convergence Divergence) que, como o nome indica, apresenta momentos onde ocorrem com a convergência ou a divergência das médias móveis. Ele é constituído de três médias móveis: Média exponencial de longo prazo, média exponencial de curto prazo e linha de sinal MACD (ou sinalização). Por padrão é utilizado os valores de 26 dias para a média longa, 12 dias para a média curta e 9 dias para a linha de sinalização.

A fórmula clássica da média móvel exponencial (MME) pode ser expressa como:

$$MME = \text{Preço} \times K + MME_{\text{ontem}} \times (1 - K)$$

(1)

Onde Preço representa o valor de fechamento da ação do dia de hoje, MME_{ontem} é o valor anterior da média móvel exponencial, e K é uma variável dependente do Período P como indicado:

$$K = 2 / P + 1$$

(2)

O MACD é formado por duas linhas: a linha de MACD e a linha de Sinal MACD. A linha MACD é calculada subtraindo a média móvel exponencial a 26 dias da média móvel a 12 dias, equação 3.

$$LinhaMACD = MME_F(12) - MME_F(26)$$

(3)

A linha de sinal MACD é calculada pela média móvel exponencial de 9 períodos da série formada pela linha MACD, equação 4

$$LinhaSinalMACD = MME_{LinhaMACD}(9)$$

(4)

O resultado destas linhas é um número que irá oscilar em torno de zero, obtendo-se assim duas linhas como se pode observar na figura 1.

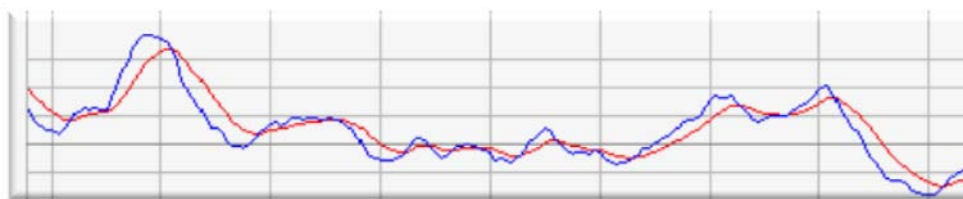


Figura 1 – Indicador MACD, Linha MACD e Linha de sinal MACD

Essas linhas são plotadas em um gráfico ao longo do tempo, e as intercessões, cruzamento ou posição dessas linhas podem sinalizar um momento oportuno do mercado para realizar uma compra ou venda. Os sinais são gerados da seguinte forma:

Sinal de compra: Um sinal de compra é gerado sempre que a linha azul (linha MACD) cruza para cima da linha vermelha (linha de sinal MACD) quando estas linhas estiverem abaixo de 0, zona considerada sobre vendido (oversold).

Sinal de venda: É gerado sempre que a linha azul cruza para baixo da linha vermelha quando estas linhas estão acima de 0, zona considerada sobre comprado (overbought).

3.1.2 IFR - Índice de Força Relativa

O IFR (Índice de Força Relativa ou RSI: Relative Strength Index) é um oscilador que foi criado por Welles Wilder na década de 70. Este oscilador é muito popular entre os analistas técnicos devido aos seus bons resultados.

Em termos de cálculo, a fórmula é bastante simples porém a sua interpretação pode ser um pouco mais complicado, equação 5.

$$\text{IFR} = 100 - (100/(1+S/D)) \quad (5)$$

Onde S é a média das cotações no fechamento dos últimos N dias em que o valor da cotação subiu, D é a média das cotações dos últimos N dias em que o valor da cotação desceu, Welles Wilder recomendou que fosse utilizado um período de 14 dias para o cálculo do IFR, no entanto também é comum a utilização dos períodos de 9 ou 25 dias.

É importante notar que o valor do IFR pode variar entre 0 e 100. Para facilitar o entendimento vamos exemplificar a implementação de um direcionador IFR utilizando o período de 14 dias:

1. Somar o valor das cotações dos dias de alta dos últimos 14 dias e dividir o resultado por 14, determinando a variável S.

2. Somar o valor das cotações dos dias de baixa dos últimos 14 dias e dividir o resultado por 14, determinando a variável D.

3. Aplicar a fórmula do IFR com os valores de S e D encontrados para determinar o valor do IRF do dia presente. Vale lembrar que assim como o MACD, o IFR também utiliza a média móvel. Dessa forma o cálculo do dia é sempre feito utilizando as últimas cotações iniciando no dia anterior ao cálculo. Os valores diários calculados deverão ser plotados em um gráfico variando de 0 a 100 com a linha intermediária sendo o 50. Abaixo apresentamos um exemplo ilustrativo:

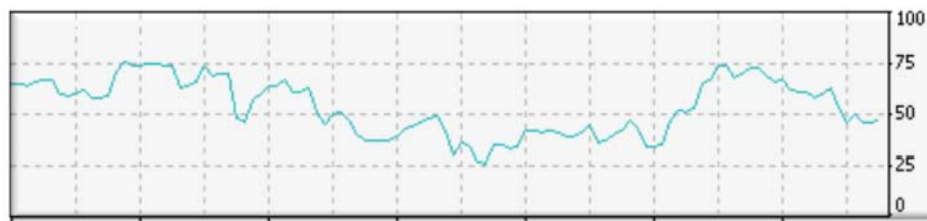


Figura 2 – Indicador IFR

Nesse mesmo gráfico são definidas duas linhas, denominadas grade, que dividem o gráfico em três intervalos: Sobre comprado (overbought), neutro e sobre vendido (oversold). Usualmente a faixa de sobre comprado é definida para valores acima de 70 e a faixa de sobre vendido abaixo de 30. Toda a faixa compreendida entre 30 e 70 é considerada como oscilações imprecisas onde não está bem definido o comportamento do mercado. Essas grades não são fixas, analistas mais conservadores podem utilizar 80 e 20, ou de acordo com o ativo avaliado, o conhecimento do mercado e o tipo de trade, pode se optar por reduzir as grades para 35 e 65, entre outros valores menos tradicionais.

Uma interpretação que pode-se retirar de um gráfico do IFR é sobre a saída das regiões de sobre vendido /sobre comprado. Sempre que o IFR caia abaixo dos 70 pontos (grade superior) depois de ter estado na região de sobre comprado, é gerado um sinal de venda do título. E sempre que o IFR sai de uma região de sobre vendido, isto é, seu valor passa a estar acima dos 30 pontos (grade inferior) é dada uma indicação de compra do título. Outras diversas interpretações podem ser analisadas no gráfico. O gráfico abaixo exemplifica a situação descrita:

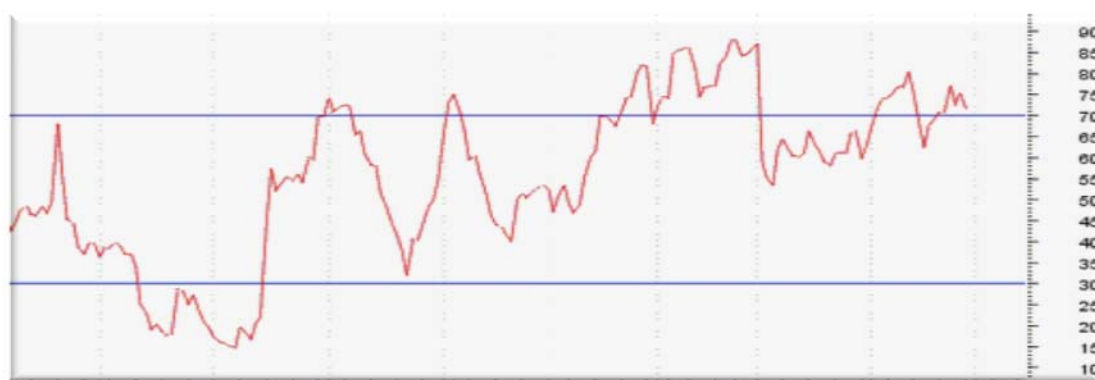


Figura 3 - Indicador IFR com grades a 70/30

3.2 Algoritmos Genéticos

A teoria moderna da evolução combina a genética as ideias de Darwin e Wallace sobre a seleção natural, criando o princípio básico de Genética Populacional, onde a variedade dos indivíduos de uma população é produzida pela recombinação genética e pela mutação.

Este princípio foi desenvolvido durante os anos 40 por biólogos e matemáticos de diversos centros de pesquisa. Nos anos 50 e 60, muitos biólogos e estudiosos começaram a desenvolver simulações computacionais de sistemas genéticos. Nos anos 80 David Goldberg, aluno de John Holland, consegue sucesso na aplicação industrial de Algoritmos Genéticos. Desde então, estes algoritmos vêm sendo aplicados com sucesso nos mais diversos problemas de otimização e aprendizado de máquinas.

Os sistemas desenvolvidos a partir deste princípio são utilizados para procurar soluções de problemas complexos ou com espaço de soluções muito grande (espaço de busca), o que os tornam problemas de difícil modelagem e solução quando se aplicam métodos de otimização convencionais.

Hoje em dia podemos encontrar aplicações de algoritmos genéticos em diversas áreas do conhecimento como na medicina, arquitetura, economia, administração engenharia entre outras. Os algoritmos genéticos também são muito utilizados na implementação de otimizadores, controles, simuladores e para a resolução de problemas matemáticos complexos.

Uma das grandes vantagens dos algoritmos genéticos é sua flexibilidade na modelagem que permite problemas de diversas naturezas serem implementados sem muito esforço de adaptação. Além disso sua estrutura de funcionamento é simples e seu processo evolutivo é o proposto por Charles Darwin, que é de conhecimento de todos.

Independente do problema ou da solução que estejamos trabalhando, os algoritmos genéticos utilizarão sempre uma estrutura básica de funcionamento:

Início:

Cria uma população inicial de indivíduos (aleatória ou com dados fornecidos). Mas a frente veremos como são formados os indivíduos. Esses indivíduos são as possíveis soluções para um problema. Também os chamamos de cromossomas, pois eles carregam o que seria o código genético (solução do problema).

Avaliação:

Verifica se os indivíduos gerados estão dentro dos padrões aceitáveis (universo da solução), elimina os indivíduos que estão fora e classifica o restante em relação à solução desejada. Cada indivíduo recebe uma “nota” em relação ao atendimento a solução, quanto mais próximo maior a nota.

Seleção:

Seleciona os indivíduos que irão se “reproduzir”. Os indivíduos classificados com as maiores notas são considerados os mais aptos a “sobrevivência” e serão selecionados para “dar continuidade a espécie”.

Cruzamento (ou recombinação genética):

Para a realização do cruzamento, os indivíduos são recombinados dois a dois para gerarem novos indivíduos. Essa recombinação pode ser feita de diversas maneiras dependendo da representação do cromossoma, do problema e da metodologia empregada. Para ilustrar um caso simples, imagine que os cromossomos são modelados utilizando letras do alfabeto e o objetivo é encontrar a senha “novidade”.

Exemplo:

Cromossomo pai: “**novicias**”

Cromossomo mãe: “**caridade**”

O cruzamento das palavras é feito, dividindo as palavras ao meio e recombinando as partes. Dessa forma os cromossomos acima gerariam dois filhos:

Cromossomo filho 1: “**novidade**”

Cromossomo filho 2: “**caricias**”

Mutação:

Alguns algoritmos podem optar por implementarem a mutação. A mutação é mais do que a modificação inesperada de um “DNA” do cromossoma. Seguindo exemplo a cima o filho 2 poderia sobre uma mutação tem uma letra alterada aleatoriamente sendo para transformado em ”camicias”. O objetivo da mutação é dar um novo fôlego ao algoritmo em momentos de estagnação. Por exemplo, vamos supor que no exemplo acima que no universo inicial gerado nenhuma palavra inicie pela letra n ou que essa letra tenha sido extinta do universo dos melhores indivíduos. Sem a mutação o algoritmo nunca encontrará a resposta esperada. Por outro lado, um número excessivo de mutações pode impedir a evolução transformando soluções boas em soluções ruins.

O procedimento básico de um algoritmo genético é resumido na Figura 4.

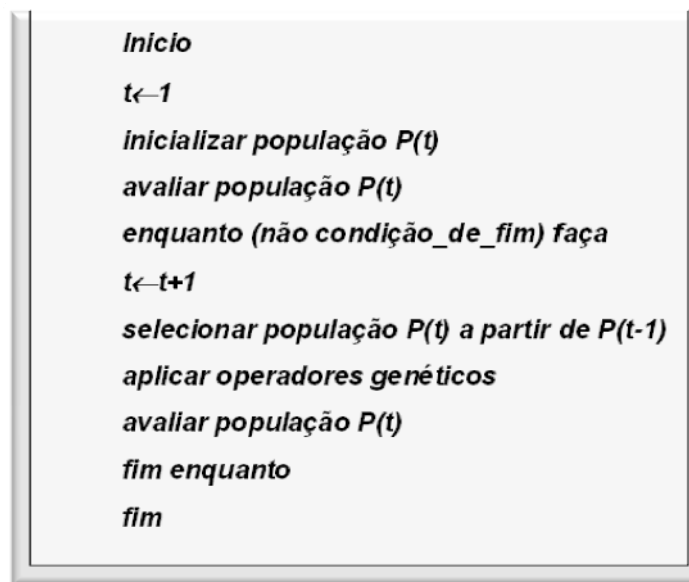


Figura 4 - Procedimento básico do algoritmo genético

Após a mutação, o algoritmo volta para o passo de avaliação e fica nesse processo até encontrar uma solução aceitável ou atingir um limite de tempo ou de gerações. A solução aceitável pode ser a solução exata ou a mais próxima aceitável. Por exemplo, a solução que teve mais de 90% de exatidão é suficiente para o problema estudado e o algoritmo pode parar ao atingir esse percentual.

Outras condições para a parada incluem o tempo de processamento, o grau de similaridade entre os elementos numa população (convergência), número de gerações e o número de indivíduos criados.

Nas seções seguintes apresentamos cada etapa de implementação do algoritmo genético com mais detalhes.

3.2.1 Representação

A representação dos genes é um aspecto fundamental na modelagem de um algoritmo genético. Nessa etapa define-se a estrutura que o cromossomo terá com os genes que o compõem, de forma que o espaço de busca do problema fique definido por toda solução gerada pela recombinação de possíveis genes.

3.2.2 Codificação e Decodificação

A solução de um problema pode ser representada por um conjunto de parâmetros (genes), unidos para formar uma cadeia de valores (cromossomo); a este processo chama-se codificação. As soluções (cromossomos) são codificadas através de uma sequência formada por símbolos de um sistema alfabético. Originalmente, utilizou-se o alfabeto binário (0, 1), porém, novos modelos de Algoritmos Genéticos codificam as soluções com outros alfabetos, como, por exemplo, números reais.

A decodificação do cromossomo consiste basicamente na transcrição do alfabeto do cromossomo na solução a qual representa. Por exemplo, um cromossomo na codificação binária 1011 será decodificado como o número real 11.

3.2.3 Avaliação

A avaliação permite ao algoritmo genético determinar sua proximidade à solução do problema. Ela é feita submetendo o cromossomo decodificado a função de avaliação que irá verificar o grau de aderência da solução proposta pelo cromossomo com a solução ideal esperada para o problema. Em alguns caso a solução ideal não é conhecida, como no caso de maximizações e minimizações, onde deseja-se o máximo ou o mínimo possível. Cabe a função de avaliação comparar a solução apresentada pelo cromossomo utilizando como parâmetro de boa solução, as soluções apresentada pelos outros cromossomos.

3.2.4. Operadores Genéticos

Os operadores mais conhecidos nos algoritmos genéticos são os de Reprodução, Cruzamento (Crossover) e Mutação.

Reprodução: Refere-se ao processo de selecionar e copiar um determinado cromossomo para a população seguinte de acordo com sua aptidão. Isto significa que os cromossomos mais aptos, têm maior probabilidade de contribuir para a formação de um ou mais indivíduos da população seguinte. Existem basicamente os seguintes métodos: troca de toda população, troca de toda população com elitismo, onde todos os cromossomos são substituídos, sendo o cromossomo mais apto da população corrente copiado para população seguinte, e troca parcial da população (steady state), onde os X melhores indivíduos da população corrente são copiados para população seguinte.

Cruzamento: É um operador baseado na troca de partes dos cromossomos (pais), formando-se duas novas soluções (filhos). Este processo pode ser observado no exemplo a seguir (Figura 5), onde a solução está codificada com alfabeto binário.

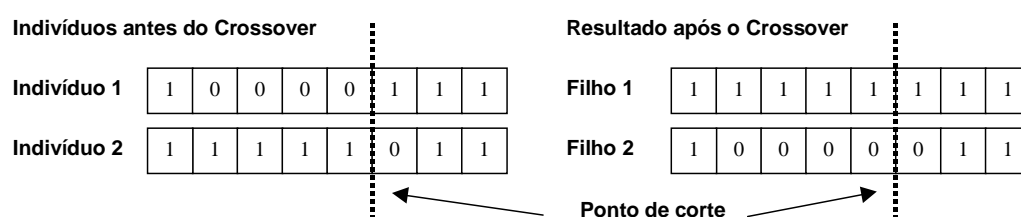


Figura 5 – Cruzamento de um ponto.

O ponto onde ocorre o corte para a realização do cruzamento é escolhido aleatoriamente; no exemplo da Figura 5 utilizou-se um único ponto, mas podem ser realizados cortes em mais de um ponto, caracterizando o multi-point crossover. Para realizar o cruzamento, primeiro é necessária a escolha, por sorteio, dos cromossomos “pais”. Em seguida ocorre a realização ou não do cruzamento segundo um parâmetro, denominado taxa de cruzamento. Deste modo, de acordo com a taxa de cruzamento, pode ocorrer que os cromossomos “pais” sejam repassados sem modificação para a geração seguinte, criando “filhos” idênticos a eles.

A ideia do operador de Cruzamento é tirar vantagem do material genético presente na população.

Mutação: é a troca aleatória do valor contido nos genes de um cromossomo por outro valor válido do alfabeto. No caso de alfabeto binário troca-se de 0 para 1 e vice-versa. Da mesma forma que para o cruzamento, utiliza-se uma taxa de mutação que, para cada bit da sequência de caracteres, sorteia-se se ocorrerá ou não a mutação; no caso de ocorrência, o bit será trocado por outro valor válido pertencente ao alfabeto (Figura 6).

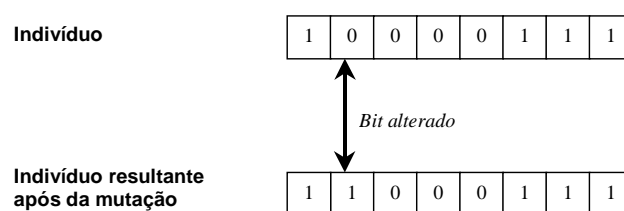


Figura 6 – Mutação

A mutação garante a diversidade das características dos indivíduos da população e permite que sejam introduzidas informações que não estiveram presentes em nenhum dos indivíduos.

3.2.5 Parâmetros da Evolução

Os parâmetros que mais influenciam o desempenho do algoritmo genético são:

Tamanho da População: afeta o desempenho global e a eficiência dos Algoritmos Genéticos. Uma população muito pequena oferece uma pequena cobertura do espaço de busca, causando uma queda no desempenho. Uma população suficientemente grande fornece uma melhor cobertura do domínio do problema e previne a convergência prematura para soluções locais. Entretanto, com uma grande população tornam-se necessários recursos computacionais maiores, ou um tempo maior de processamento do problema. Logo, deve-se buscar um ponto de equilíbrio no que diz respeito ao tamanho escolhido para a população.

Taxa de Cruzamento: probabilidade de um indivíduo ser recombinado com outro. Quanto maior for esta taxa, mais rapidamente novas estruturas serão introduzidas na população. Entretanto, isto pode gerar um efeito indesejável, pois a maior parte da população será substituída, causando assim perda de variedade genética, podendo ocorrer perda de estruturas de alta aptidão e convergência a uma população com indivíduos extremamente parecidos, indivíduos estes de solução boa ou não. Com um valor baixo, o algoritmo pode-se tornar muito lento para oferecer uma resposta aceitável.

Taxa de Mutação: probabilidade do conteúdo de um gene do cromossomo ser alterado. A taxa de mutação previne que uma dada população fique estagnada em um valor, além de possibilitar que se chegue em qualquer ponto do espaço de busca. Porém deve-se evitar uma taxa de mutação muito alta, uma vez que esta pode tornar a busca essencialmente aleatória, prejudicando fortemente a convergência para uma solução ótima.

Intervalo de Geração: controla a porcentagem da população que será substituída durante a próxima geração (substituição total, substituição com elitismo, substituição dos piores indivíduos da população atual, substituição

parcial da população sem duplicatas). Esse número de indivíduos substituídos também é conhecido como GAP.

Número de gerações: representa o número total de ciclos de evolução de um Algoritmo Genético, sendo este um dos critérios de parada do algoritmo genético. Um número de gerações muito pequeno causa uma queda no desempenho: um valor grande faz necessário um tempo maior de processamento, mas fornece uma melhor cobertura do domínio do problema, evitando a convergência para soluções locais.

4. Arquitetura do sistema proposto

4.1 Considerações

Para realizar o estudo dos indicadores escolhemos a ferramenta Excel para a implementação do trabalho. Escolhemos o Excel por ser uma ferramenta simples, de fácil acesso a todas as pessoas e por possuir os instrumentos necessários para a realização do estudo. Utilizamos o plugin Evolver 5.7 da empresa Palisade [1] para Excel para implementar o algoritmo genético, o restante do sistema é composto por fórmulas do Excel e código VBA.

Como mencionado anteriormente, nesse estudo utilizamos o Ativo Petr4, que corresponde à código das ações ordinárias da Petrobras. Obtivemos a série histórica dos valores desse Ativo de 30/07/1998 até 13/05/2011 com os valores de abertura e fechamento diários. Esses dados são disponibilizados ao público e pode ser obtido facilmente na internet do site da BOVESPA ou nos sites de corretoras. A escolha desse ativo se deu por ser um ativo conhecido, ter uma significativa representação no índice da BOVESPA (Ibovespa) e por que a sua série história apresentar movimentos considerados interessantes para o estudo. Abaixo na Figura 7, apresenta-se o gráfico gerado pelo ativo para o período acima mencionado:

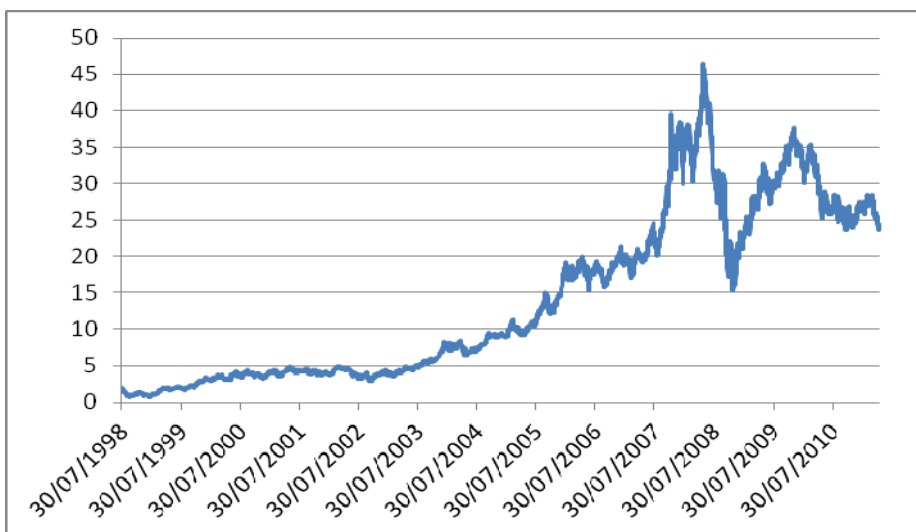


Figura 7 – Evolução do valor da ação Petr4

O valor da ação variou de 1,86 em Julho de 1998 até 46,45 em Maio de 2008, onde apresentou o seu valor máximo nesse período. Para evitar distorções nos resultados devido a eventos econômicos como inflação, juros, variação cambial, etc. Na simulação, optou-se por medir os ganhos e perdas em percentuais.

Para a elaboração do sistema dividimos o estudo em três fases: na primeira implementamos, otimizamos e testamos somente o indicador MACD. Na segunda implementamos, otimizamos e testamos o indicador IFR, e na terceira e ultima fase, elaboramos três estratégias de investimento combinando os dois indicadores e medimos o desempenho dessas estratégias com o desempenho dos indicadores otimizados individualmente.

Para a implementação dos indicadores utilizamos o leiaute a baixo no Excel para estruturar os dados da Petr4:

| Petr4 | | | |
|-------|------|------------|----------|
| id | data | Fechamento | Abertura |

Figura 8 – Cabeçalho dos dados de teste do Ativo Petr4

Onde id é uma chave única que identifica a linha da tabela, iniciando em 1 e continuando sequencialmente até o final da tabela. Data é campo onde é informado a data de cotação do ativo. No campo data, é informado dia, mês e ano da cotação, da mesma forma que o Id inicia em 30/07/1998 e segue sequencialmente até 13/05/2011. Finais de semana, feriados e dias onde não houve atividades na bolsa de valores, não foram considerados.

4.2 Modelo para otimização dos parâmetros do MACD

Para a implementação do indicador MACD criamos cinco colunas no Excel com os valores (calculados por fórmulas) da média longa, média curta, linha MACD, linha de sinalização e linha MACD X linha de sinalização.

A média longa foi implementada considerando a variável “Média móvel longa”, onde o usuário pode entrar o valor (em número de dias) que deseja que a média móvel exponencial seja calculada. A média móvel longa de Y dias é calculada como:

$$\text{MME}(Y) = (\text{Média dos valores de fechamento do Ativo dos } Y-1 \text{ dias anteriores}) \times (K-1) + (\text{o valor do fechamento do Ativo no dia } Y) \times K \quad (6)$$

Onde o K é calculado como $2 / (Y+1)$. O valor padrão para Y é 26 dias.

A média curta é realizada com o mesmo cálculo porém utilizando a variável “Média móvel curta”, onde o valor padrão 12 dias.

A linha MACD é a diferença entre a média curta e a média longa. A terceira variável utilizada é a linha de sinalização, que é a média dos Z últimos valores da linha MACD, normalmente o seu valor padrão é 9 dias. Para auxiliar a visualização do comportamento das linhas, criamos o campo linha MACD X linha de sinalização, que tem o valor TRUE toda vez que a linha de sinalização está acima da linha MACD e FALSE quando a linha de sinalização está abaixo da linha MACD. Dessa forma pode ser facilmente visualizado o cruzamento entre as linhas.

Além das variáveis: “Média móvel longa”, “Média móvel curta” e “linha de sinalização”, utilizamos as variáveis “Intervalo de Estimação” e “Intervalo de Validação”. Com a variável Intervalo de Estimação, é possível parametrizar o número de dias que serão utilizados para realizar os testes de desempenho e otimização do indicador, enquanto na variável Intervalo de validação é possível informar o período em que será utilizado para apurar o lucro (ou prejuízo) aferido pela aplicação do indicador otimizado.

O modelo opera de duas forma, como Comprado, onde realiza uma ordem de compra é sinalizada antes de uma ordem de venda. Então a operação comprada

é iniciada e só é finalizada ao receber um sinal de venda. O ganho da operação é o retorno calculado como:

$$\text{Retorno} = \frac{\text{valor de compra} - \text{valor de venda}}{\text{valor de compra}} \quad 7)$$

Dessa forma obtêm-se um percentual do retorno do investimento. Ressaltando, mais uma vez, que não foi considerado inflação, taxas, impostos ou outros fatores econômicos fala fim de avaliação da performance do indicador.

A segunda forma de operação possível é a forma Vendida, onde uma ordem de venda é sinalizada antes de uma ordem de compra. Dessa forma consideramos que uma ação é vendida e a operação é finalizada quando um sinal de compra é emitido. Esta última operação é especialmente vantajosa em momentos de queda do ativo, onde no momento seguinte o valor é menor do que o valor atual. Adotamos como premissa do modelo que as operações na forma Comprado e Vendido não são aplicadas paralelamente. Assim o sinal ativa uma das duas formas, a outra não pode ser iniciada até que a anterior seja encerrada.

Uma vez definida a arquitetura e modus operandi do modelo utilizamos o plugin do Excel Evolver 5.7 da Palisade para realizar a otimização do indicador. Assumimos algumas premissas para limitar o espaço de busca do algoritmo como: A média longa pode variar entre 10 e 150 dias. A média Curta pode variar entre 7 e 149 dias desde que seja menor do que a média longa. E por fim a linha de sinalização pode variar entre 1 e 148 dias desde que seja menor do que a média curta. As restrições utilizadas para a parametrização podem ser visualizadas na tela de parametrização do Evolver, Figura 9.

$$10 \leq \text{Média Longa} \leq 150$$

$$7 \leq \text{Média Curta} \leq 149$$

$$\text{Média Curta} \leq \text{Média Longa}$$

$$1 \leq \text{Linha Sinalização} \leq 149$$

$$\text{Linha Sinalização} \leq \text{Média Curta}$$

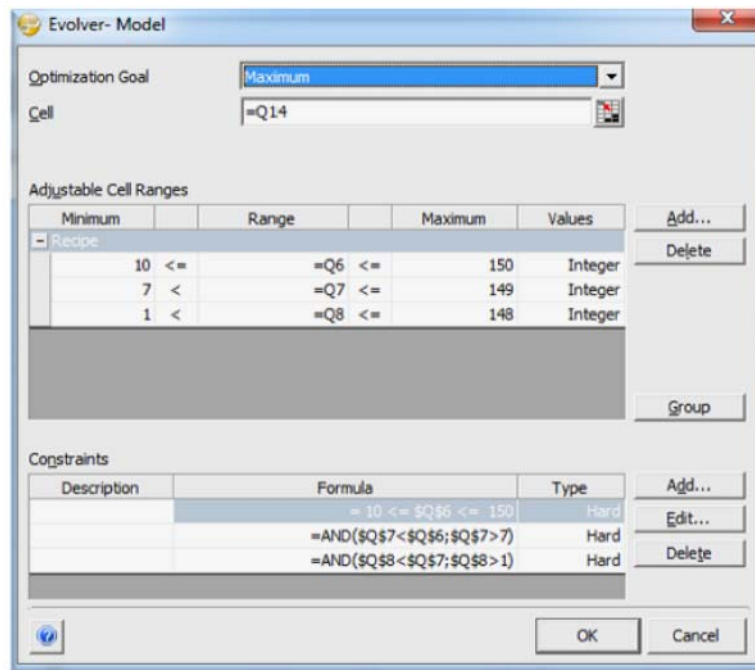


Figura 9 – Parametrização do Evolver para o MACD

Messa implementação de algoritmo genético, o cromossomo é formado por três genes: a média longa, média curta e sinalização.

A função de avaliação utilizada é a função de maximização onde o resultado gerado pelo modelo nas formas Comprado e Vendido são somados e avaliados contra melhor resultado obtido até o momento, equação 7.

$$F_{\text{avaliação}} = \text{Max} \left[\sum \text{Retornos}_{\text{Pass Comprado}} + \sum \text{Retornos}_{\text{Pass Vend}} \right] \quad 8)$$

Os melhores indivíduos (resultados) têm seus genes reproduzidos e levados para a próxima geração.

4.3 Modelo para otimização dos parâmetros do IFR

Para a implementação do indicador IFR utilizamos as mesmas datas, períodos e valores de abertura e fechamento do Ativo PETR4 utilizado pela implementação do MACD. Em seguida identificamos para cada dia se ele foi um dia de alta ou de baixa. Esse cálculo foi feito subtraindo o valor do fechamento do Ativo do dia anterior, do valor de fechamento do Ativo no dia calculado. Se o resultado foi positivo, então o dia foi de alta e se o resultado der negativo, então o dia foi de baixa. Com isso, dividimos o restante do cálculo em duas partes, a parte das baixas e a parte das altas. Na parte das baixas, calculamos com duas colunas. Na primeira coluna, para cada dia, somamos os valores dos dias de baixas dos últimos N dias incluindo o dia atual. Onde N é a variável Período de Cálculo do IFR. Na segunda coluna calculamos a média dos dias de baixa, a primeira coluna por N. A segunda parte do cálculo (altas) é realizada de forma análoga a parte das baixas descrita anteriormente. Por fim, realizamos o cálculo do IFR para cada dia utilizando a equação 8:

$$IFR(N) = 100 - \left[\frac{100}{1 + \frac{\text{Média dos N dias de alta}}{\text{Média dos N dias de baixa}}} \right] \quad (9)$$

Abaixo na Figura 10, apresenta-se uma imagem ilustrativo da representação gráfica o IFR entre duas grades (grade inferior a 20 e superior a 80).

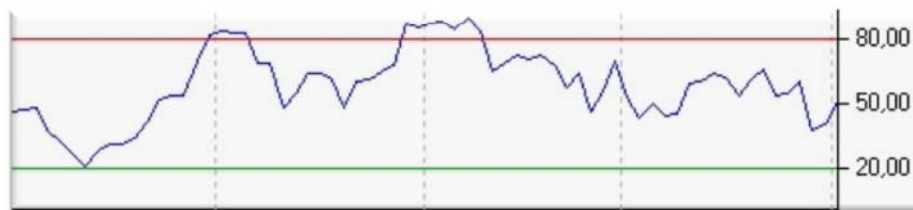


Figura 10 – Gráfico gerado pelo IFR

Como vimos a cima a fórmula do IFR varia de acordo com as médias dos dias, que por sua vez variam de acordo com N, o período de cálculo. Por tanto, o

período de cálculo é a principal variável desse modelo. Adicionalmente a essa variável, fizemos alguns teste de desempenho considerando as grades variáveis, deixando que o algoritmo genético busque o espaçamento das grades e o período para o IFR que traga o maior retorno possível para o investimento. A tabela a baixo ilustra o quadro das variáveis implementadas e o quadro de avaliação dos resultados:

| Parâmetros ajustáveis (dias) | |
|------------------------------|-------------|
| Período de cálculo do IFR | 14 |
| Grade superior | 60 |
| Grade Inferior | 40 |
| Intervalo de Estimação | 2484 |
| Intervalo de validação | 680 |

| | |
|--|---------------|
| <i>Lucro comprado estimação</i> | 197% |
| <i>Lucro comprado real</i> | 15% |
| <i>Lucro vendido estimação</i> | -244% |
| <i>Lucro vendido real</i> | 20% |
| <i>Lucro total estimação</i> | -48% |
| <i>Lucro total real</i> | 35% |
| <i>Lucro médio total Est. p/ 30 dias</i> | -0,57% |
| <i>Lucro médio total real p/ 30 dias</i> | 1,55% |

Tabela 1 – Variáveis utilizadas na implementação do IFR

Nesse exemplo o cálculo foi feito utilizando um período de 14 dias, para as médias dos dias de alta e baixa. As grades fixas a 40 e 60.

Foram utilizados 2484 dias, desde Julho de 1998 até Agosto de 2008, para estimação do resultado e um período de 680 dias, de Agosto de 2008 até Maio de 2011, para a avaliação da performance do algoritmo.

Na tabela de resultados avaliamos em separado cada uma das estratégias Comprado e Vendido, tanto para o universo de estimação como para o universo de validação. O Lucro total estimado é o somatório dos lucros estimados da forma Comprado e Vendido, da mesma forma que o Lucro total real é o somatório dos

lucros gerados no período de validação nas formas Comprado e Vendido. Os dois últimos campos (Lucro médio total Est. p/ 30 dias e Lucro médio total real p/ 30 dias) apresentam as médias mensais dos lucros obtidos tanto para o período de estimação como para o período de validação.

Para os testes onde consideramos as grades fixas a única variável otimizada pelo algoritmo genético com o período de cálculo do IFR. Já nos testes onde foram considerados as grades flexíveis, o cromossomo foi composto do período de cálculo mais as grades superior e inferior. A função de avaliação é a maximização, mesma utilizada no MACD, onde o objetivo é encontrar o maior lucro possível. Utilizando a mesma premissa do MACD, uma vez que uma modalidade de investimento é iniciada (Comprado ou Vendido), não é permitido iniciar outra simultaneamente até que a anterior finalize.

4.4 Estratégias de Investimento

Além de testar cada um dos dois indicadores individualmente nos modos Vendido e Comprado, elaboramos duas estratégias combinando os indicadores.

Estratégia Conjunta 1: Nesta primeira estratégia os dois indicadores são iniciados ao mesmo tempo e assim que um dos indicadores sinaliza o início de uma operação (compra ou venda) o outro não pode iniciar uma operação (mesmo emitindo um sinal) até que a operação iniciada pelo primeiro indicador termine. A otimização, por algoritmo genético, avaliou o resultado do conjunto e considerou os parâmetros do MACD e IFR no mesmo cromossomo para a otimização do conjunto. Os testes realizados para essa estratégia, foi similar aos testes realizados para cada indicador individual. A apresentação dos resultados e discussões serão feitas no capítulo seguinte. As parametrizações do evolver podem ser visualizadas na imagem abaixo:

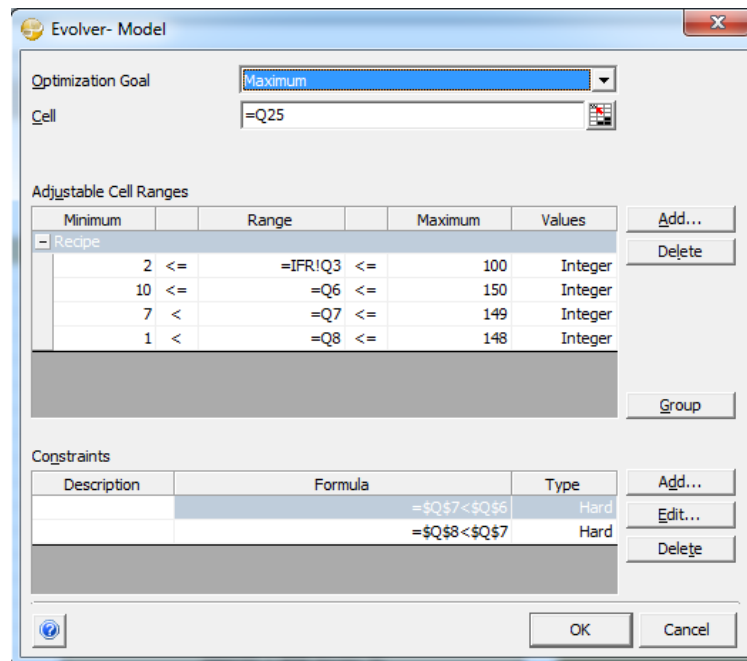


Figura 11 – Parametrização do Evolver para os indicadores MACD e IFR combinados

A célula Q6 tem o valor da média longa, Q7 o valor da média curta e Q8 a linha de sinalização. IFR!Q3 tem o período de cálculo para o IFR. As condições “Hard” garantem que a média curta nunca será maior que a média longa e que a sinalização nunca será maior que a média curta.

Estratégia Conjunta 2: Nesta segunda estratégia os dois indicadores são iniciados ao mesmo tempo e podem operar simultaneamente independente do outro indicador. No entanto a otimização é feita avaliando o resultado do conjunto, os parâmetros do MACD e IFR fazer parte do mesmo cromossomo para a otimização. Assim como na primeira estratégia e nos testes individuais, são consideradas as modalidades Comprado e Vendido para cada indicador. As parametrizações do Evolver são as mesmas apresentadas na imagem da primeira estratégia.

5 Resultados

Durante o processo de otimização dos parâmetros dos indicadores a fim de identificar que valores dos parâmetros são capazes de gerar o melhor resultado (maior lucro), foram realizados diversos testes assumindo algumas premissas. A seguir são apresentados os resultados, premissas e análise de cada um dos testes realizados para os indicadores individuais e das estratégias combinando indicadores.

5.1 Estratégia com Indicador MACD Otimizado

Neste experimento foram realizados seis testes buscando medir o ganho gerado pelo indicador MACD para a ação Petr4 no histórico de cotações do dia 30/07/1998 até 12/08/2008, intervalo o qual consideramos de estimação. O teste inicial considera o indicador MACD com os valores padrões de média longa (26 dias), média curta (12 dias) e sinalização (9 dias), os seguintes testes (denominados teste1, teste 2, teste3, teste 4 e teste 5) consideram diferentes otimizações com o algoritmo genético. Os testes realizados, em uma mesma seção, foram feitos a partir do reprocessamento do algoritmo na tentativa de encontrar o parâmetro ótimo. O que chamamos de teste aqui são apenas reprocessamentos do algoritmo genético. Em cada um dos testes, o algoritmo pode variar os parâmetros: média curta, média longa e sinalização para determinar a combinação mais rentável entre eles (melhores soluções).

A tabela 2 apresenta os valores dos parâmetros do MACD e os resultados de todos os testes realizados para os dados de estimação (lucro estimação) e os dados de validação (lucro real).

Observa-se que para o teste padrão (valores padrões para os parâmetros do MACD), tabela 2, que o valor da aplicação foi de 182% em um período de aproximadamente 10 anos. O que seria equivalente a um rendimento anual aproximado de 18,2%. No entanto quando testamos o indicador no período de

13/08/2008 a 13/05/2011 (dados de validação), obtivemos apenas 1% de lucro no período.

| Resultados Preliminares MACD (melhores testes) | | | | | | |
|---|---------------|---------------|---------------|---------------|---------------|---------------|
| | <i>Padrão</i> | <i>Teste1</i> | <i>Teste2</i> | <i>Teste3</i> | <i>Teste4</i> | <i>Teste5</i> |
| Média longa | 26 | 67 | 100 | 12 | 25 | 10 |
| Média Curta | 12 | 34 | 12 | 8 | 23 | 8 |
| Sinalização | 9 | 27 | 9 | 2 | 11 | 2 |
| Lucro Est. | 182% | 255% | 310% | 327% | 320% | 322% |
| Lucro Real | 1% | 15% | 27% | 4% | 12% | 12% |

Tabela 2 – Resultado dos testes do MACD

Para entender melhor os resultados a tabela 3 apresenta os rendimentos do teste padrão por tipo de operação:

| Operação | Resultado |
|--|------------------|
| <i>Lucro comprado estimacão</i> | 250% |
| <i>Lucro comprado real</i> | -21% |
| <i>Lucro vendido estimacão</i> | -68% |
| <i>Lucro vendido real</i> | 22% |
| <i>Lucro total estimacão</i> | 182% |
| <i>Lucro total real</i> | 1% |
| <i>Lucro médio total Est. p/ 30 dias</i> | 2,20% |
| <i>Lucro médio total real p/ 30 dias</i> | 0,04% |

Tabela 3 – Resultado do MACD para parâmetros padrão

Na tabela 3, observa-se que o indicador apresentou um retorno de -21% em operações Comprado e de 22% para operações Vendido. Observando o gráfico do comportamento do preço da ação no período de validação (figura 12) nota-se que houve mais períodos de baixa que de alta, o que justifica um melhor desempenho das operações Vendidas. Em contra partida observando o gráfico do preço da ação no período de estimacão (figura 13), observa-se uma tendência de alta constante, o que favorece a atuação Comprado.

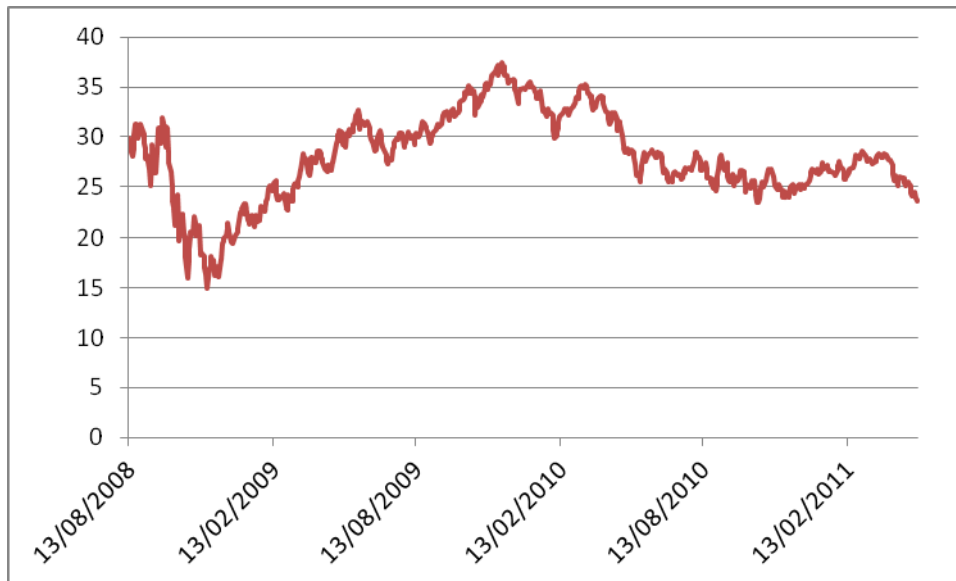


Figura 12 – Gráfico do valor das ações da Petr4 para o período de teste

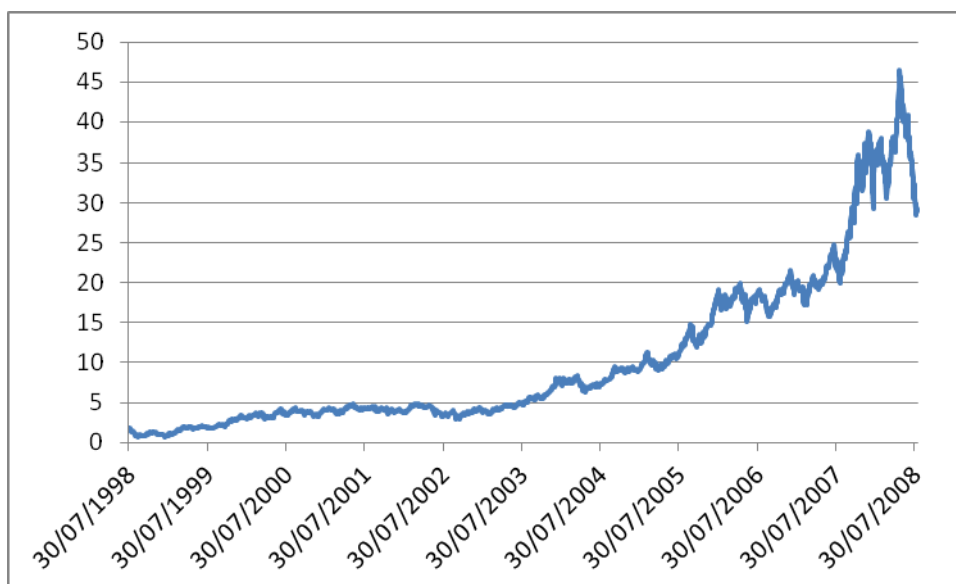


Figura 13 – Gráfico do valor das ações da Petr4 para o período de estimação

Nos testes 1 e 2, o algoritmo optou por médias longas altas e maior diferente entre as médias longa e curta. Tal combinação apresentou um comportamento um pouco inferior para o período de estimação, no entanto um resultado melhor em ambiente mais volátil. A tabela 4 abaixo apresenta o desempenho detalhado por tipo de operação para os testes 1 e 2. r

| Operação | Resultado teste | |
|--|-----------------|-------------|
| | 1 | 2 |
| <i>Lucro comprado estimação</i> | 305% | 329% |
| <i>Lucro comprado real</i> | 15% | 3% |
| <i>Lucro vendido estimação</i> | -50% | -20% |
| <i>Lucro vendido real</i> | 0% | 24% |
| <i>Lucro total estimação</i> | 255% | 310% |
| <i>Lucro total real</i> | 15% | 27% |
| <i>Lucro médio total Est. p/ 30 dias</i> | 3,08% | 3,74% |
| <i>Lucro médio total real p/ 30 dias</i> | 0,67% | 1,18% |

Tabela 4 – Resultado do MACD para os testes 1 e 2

Nos testes 3, 4 e 5 observa-se que o algoritmo genético optou pela linha MACD tendo poucos dias de diferença entre a média longa e média curta, o que trouxe bons resultados para o período de estimação, onde o comportamento das ações segue um comportamento mais constante e crescente. No entanto para o período de teste apresentou um desempenho aproximado entre 1,5% e 4% ao ano. A tabela 5 abaixo apresenta o desempenho detalhado por tipo de operação para os testes 3, 4 e 5.

| Operação | Resultado | Resultado teste | Resultado teste |
|--|-------------|-----------------|-----------------|
| | teste 3 | 4 | 5 |
| <i>Lucro comprado estimação</i> | 278% | 318% | 273% |
| <i>Lucro comprado real</i> | 13% | 11% | 1% |
| <i>Lucro vendido estimação</i> | 49% | 2% | 49% |
| <i>Lucro vendido real</i> | -10% | 1% | 11% |
| <i>Lucro total estimação</i> | 327% | 320% | 322% |
| <i>Lucro total real</i> | 4% | 12% | 12% |
| <i>Lucro médio total Est. p/ 30 dias</i> | 3,95% | 3,87% | 3,89% |
| <i>Lucro médio total real p/ 30 dias</i> | 0,16% | 0,54% | 0,53% |

Tabela 5 – Resultado do MACD para os testes 3,4 e 5

5.2 Estratégia com Indicador IFR Otimizado

O indicador IFR é composto por basicamente uma variável, o período de cálculo. O que poderia ser otimizado através de uma busca sequencial pelo melhor período. No entanto para tentar encontrar melhores resultado consideramos diversos testes para IFR, desde teste otimizando apenas o período de cálculo com otimização das grades superiores e inferiores. Em seguida são apresentados os quatro testes realizados para diferentes considerações das grades:

A. Grades Fixas

Nesse teste consideramos as grades fixas, figura 14, com diferentes valores (além do padrão 70/30). Foram realizados seis testes sendo que o primeiro considera os valores padrões para o período e para as grades (14 dias com grade a 70/30), nos outros testes foi otimizado o período do indicador para grades predefinidas (80/20, 75/25, 70/30, 65/34 e 60/40).

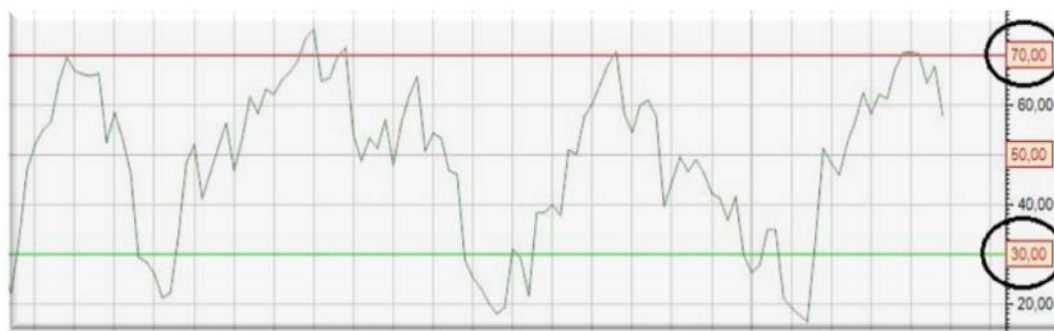


Figura 14 – Grades do IFR a 70 (vermelha) e 30 (verde)

A tabela 6 apresenta os resultados obtidos através da otimização do período considerando os intervalos de grade mencionados acima. Como pode-se observar na tabela, que o teste com os parâmetros padrões tiveram os piores resultados, registrando 4% na estimativa e um prejuízo de -69% no período de teste. Independente da grade utilizada, podemos concluir que o melhor valor para o parâmetro período foi dois dias (o menor admissível). Um período de cálculo curto constitui um indicador muito volátil, onde qualquer oscilação nos preços das ações é captada. Na média dos testes realizados obtivemos um retorno de 98,6% no período de teste, o que corresponde a aproximadamente 35% ao ano, contra um retorno negativo dos parâmetros padrões. Vale ressaltar que um período muito

curto pode exigir um grande número de emissões de ordem de compra e venda que em casos reais gere um custo superior a períodos mais conservadores.

| Resultados IFR (melhores testes - grades fixa) | | | | | | |
|--|--------|-------|-------|-------|-------|-------|
| | Padrão | 80/20 | 75/25 | 70/30 | 65/35 | 60/40 |
| Período de cálculo do IFR | 14 | 2 | 2 | 2 | 2 | 2 |
| Grade superior | 70 | 80 | 75 | 70 | 65 | 60 |
| Grade Inferior | 30 | 20 | 25 | 30 | 35 | 40 |
| Lucro Est. | 4% | 538% | 554% | 588% | 659% | 516% |
| Lucro Real | -69% | 132% | 108% | 78% | 80% | 95% |

Tabela 6 – Resultado do IFR p/ grades fixas

A. Grades equidistantes variáveis e limitadas a 80/20

Como uma tentativa de obter melhores resultados foram realizados cinco testes sendo que o primeiro considera os valores padrões para as grades (70/30) e otimiza só o período. Nos outros testes foi otimizado pelo algoritmo genético o período e as grades, deixando o algoritmo genético livre para escolher a melhor combinação de período e altura das grades. Para esse teste criamos duas restrições para o algoritmo: as grades podem ser assumir qualquer valor desde que equidistantes e a grade superior limitada entre 51 e 80 e a grade inferior entre 49 e 20. A tabela 7 apresenta os resultados obtidos nos testes:

$$51 \leq \text{Grade Superior} \leq 80$$

$$20 \leq \text{Grade Inferior} \leq 49$$

$$\text{Grade Inferior} = 100 - \text{grade superior}$$

| Resultados IFR (melhores testes - grades equidistantes e limitadas a 80 / 20) | | | | | |
|---|--------|--------|--------|--------|--------|
| | Padrão | Teste1 | Teste2 | Teste3 | Teste4 |
| Período de cálculo do IFR | 25 | 2 | 2 | 2 | 2 |
| Grade superior | 70 | 79 | 62 | 64 | 77 |
| Grade Inferior | 30 | 21 | 38 | 36 | 23 |
| Lucro Est. | -114% | 404% | 395% | 346% | 293% |
| Lucro Real | -8% | 36% | 11% | 6% | 41% |

Tabela 7 – Resultado do IFR p/ grades equidistantes limitado a 80/20

Observa-se que os testes realizados com esse método não melhoraram os resultados obtidos anteriormente.

B. Grades equidistantes variáveis

A fim de obter melhores resultados que no teste anterior, foram colocadas menos restrições ao algoritmo genético e repetimos o teste retirando a limitação de 80/20. Os resultados estão apresentados na tabela 8:

| Resultados IFR (melhores testes - equidistantes) | | | | | |
|--|--------------|-------------|--------------|--------------|-------------|
| | Padrão | Teste1 | Teste2 | Teste3 | Teste4 |
| Período de cálculo do IFR | 9 | 2 | 27 | 28 | 2 |
| Grade superior | 70 | 80 | 87 | 85 | 66 |
| Grade Inferior | 30 | 20 | 13 | 15 | 34 |
| Lucro Est. | -135% | 538% | 1061% | 1018% | 641% |
| Lucro Real | 38% | 132% | 0% | 0% | 57% |

Tabela 8 – Resultado do IFR p/ grades equidistantes e ilimitadas

Observa-se que os testes 2 e 3 tiveram resultados bem superiores os resultados encontrados nos teste apresentados na seção A e B. Os testes 2 e 3, dessa seção, apresentaram um retorno superior a 1000% no período de estimativa enquanto, quando testados no período real tiveram resultado nulo.

Analisando cuidadosamente os dados concluí-se que uma leitura inteligente do algoritmo. Se observa-se a figura 13 apresentada no início desse capítulo, podemos ver que durante esse período a ação sofreu uma valorização constante desde o valor inicial de 1,86 reais até 46,45 reais. Dessa forma o algoritmo configurou os parâmetros do IFR de forma que durante o período de estimação poucas operações fossem realizadas, no entanto as operações realizadas apresentaram um alto retorno por adquirirem o ativo a baixo custo e vende-lo após anos por um valor muito superior. Infelizmente essa não é a realizada visualizada no intervalo de teste da figura 12. Nesse cenário o algoritmo não conseguiu finalizar nenhuma operação. Adquirindo o Ativo no início e não identificando uma variação significativa para finalizar a operação.

Observamos com isso que o algoritmo genético optou por um investimento mais conservador de longo prazo, porém com isso perdeu algumas boas oportunidades de investimento no curto prazo.

B. Grades variáveis

Neste experimento foi decidido retirar todas as restrições e deixar o algoritmo genético otimizar livremente o período de cálculo e as grades. A tabela 9 apresenta os melhores parâmetros descobertos pelo algoritmo genético e seus respectivos (duas últimas linhas):

| Resultados IFR (testes livres) | | | | | |
|--------------------------------|--------|--------|--------|--------|--------|
| | Padrão | Teste1 | Teste2 | Teste3 | Teste4 |
| Período de cálculo do IFR | 14 | 11 | 12 | 34 | 44 |
| Grade superior | 70 | 98 | 97 | 88 | 84 |
| Grade Inferior | 30 | 30 | 49 | 40 | 49 |
| Lucro Est. | 4% | 1182% | 1753% | 1937% | 1898% |
| Lucro Real | -69% | 0% | 0% | 0% | 0% |

Tabela 9 – Resultado do IFR p/ otimização livre

Observamos que dessa forma o algoritmo segue o mesmo caminho dos testes 2 e 3 da seção anterior para todos os testes. Porém com os parâmetros variáveis a sua disposição ele consegue um lucro maior que na seção C para o período de estimação e não realiza nenhuma operação no período real.

Após esses testes concluímos que o melhor resultado obtido no período real foi de 132%, encontrado na primeira seção, onde o período foi 2 e grades de 80/20. Vale ressaltar que com um período curto o indicador é volátil e foram necessárias diversas operações de compra e venda, algumas com lucro muito pequeno que em um ambiente real talvez não valesse a pena a sua realização.

Já no período de estimação o melhor resultado foi de 1937% para o período de 34 dias e grade 88/40. Onde o indicador atuou de forma mais conservadora realizando poucas operações.

5.3 Estratégia MACD E IFR independentes

Após testar cada um dos indicadores individualmente, se elaborou uma estratégia de investimento onde operamos com os dois indicadores simultaneamente, onde as ações de um indicador não interferem no outro. O Algoritmo genético otimizou o ganho dos operadores em conjunto pois considerou os parâmetros do MACD e IFR (período e grade) em um cromossomo só. Os testes foram realizados com os mesmos dado da Ação e os intervalos de dados de estimação e de validação foram os mesmos mencionados no início do capítulo. Foram realizados seis testes, sendo que o primeiro teste utiliza os parâmetros padrão tanto para MACD como para o IFR, nos outros testes o algoritmo genético selecionou a melhor combinação de parâmetros após executar mil gerações. A tabela 10 apresenta os resultados.

| Resultados MACD + IFR independentes | | | | | | | |
|-------------------------------------|---------------------------|------------|-------------|-------------|-------------|-------------|-------------|
| | Parâmetros | Padrão | Teste1 | Teste2 | Teste3 | Teste4 | Teste5 |
| MA CD | Média longa | 26 | 33 | 63 | 66 | 12 | 10 |
| | Média Curta | 12 | 14 | 36 | 35 | 8 | 8 |
| | Sinalização | 9 | 13 | 34 | 34 | 2 | 2 |
| IFR | Período de cálculo do IFR | 9 | 2 | 2 | 2 | 2 | 2 |
| | Grade superior | 70 | 70 | 70 | 70 | 70 | 70 |
| | Grade Inferior | 30 | 30 | 30 | 30 | 30 | 30 |
| | Lucro Est. | 47% | 907% | 838% | 839% | 915% | 910% |
| Lucro Real | 39% | 64% | 106% | 104% | 81% | 90% | |

Tabela 10 – Resultado do MACD + IFR independentes

Os resultados da estratégia combinando os dois operadores foi bem vantajosa de forma geral. Para o IFR não houve alteração em relação ao seu desempenho sozinho. Assim como nos testes individuais o melhor período ficou em dois dias.

Já o MACD parece que não houve uma convergência total para um melhor parâmetro. Novamente ele apresenta resultados bons para média curta e média longa. No entanto os melhores resultados durante o período de teste foram utilizando-se médias longas de 63 e 66 dias e linha de sinalização e média curta bem próximas.

5.3 Estratégia MACD E IFR dependentes

Neste experimento, considera-se a estratégia de atuar com os dois indicadores, no entanto não operando em paralelo. Uma vez que um indicador sinaliza o início de uma operação, o outro operador não pode iniciar até que o primeiro tenha finalizado a operação que iniciou. Assim como na estratégia anterior, o algoritmo genético considerou os parâmetros do MACD e IFR em um cromossomo só. A tabela 11 apresenta os parâmetros do algoritmo genético e seus respectivos resultados obtidos durante os testes (duas últimas linhas da tabela):

| Resultados Preliminares MACD + IFR (Teste 4) | | | | | | | |
|--|---------------------------|-------------|-------------|-------------|-------------|-------------|-------------|
| | Parâmetros | Padrão | Teste1 | Teste2 | Teste3 | Teste4 | Teste5 |
| MA | Média longa | 26 | 12 | 15 | 12 | 12 | 27 |
| | Média Curta | 12 | 9 | 9 | 8 | 7 | 21 |
| | Sinalização | 9 | 2 | 2 | 2 | 2 | 11 |
| IFR | Período de cálculo do IFR | 9 | 2 | 2 | 2 | 2 | 2 |
| | Grade superior | 70 | 70 | 70 | 70 | 70 | 70 |
| | Grade Inferior | 30 | 30 | 30 | 30 | 30 | 30 |
| | Lucro Est. | -59% | 636% | 630% | 631% | 622% | 583% |
| | Lucro Real | 39% | 85% | 67% | 90% | 102% | 70% |

Tabela 11 – Resultado do MACD + IFR dependentes

Os resultados obtidos foram bons de forma geral, tendo um resultado médio de 82,8% o que corresponde a aproximadamente 30% ao ano. Podemos observar que os parâmetros, tanto do MACD quando do IFR convergiram para determinados valores.

No período de estimação os resultados foram bastante próximos tendo um valor médio de 620,4% no período, o que corresponde a aproximadamente a 62% ao ano.

Com isso conseguimos concluir que o desempenho do algoritmo em cenários mais estáveis é superior a cenários muito voláteis. Porém mesmo em cenário volátil o desempenho do algoritmo otimizado foi bem superior ao algoritmo com parâmetros padrões.

6. Conclusão e trabalhos futuros

Ao final desse trabalho podemos concluir que a otimização dos parâmetros dos indicadores para um determinado tipo de ação é vantajosa e melhora consideravelmente o seu desempenho. Da mesma forma os algoritmos genéticos se demonstraram bem eficientes e rápidos na otimização dos parâmetros. A tabela 12 abaixo apresenta os resultados médios para os dados do período de validação de cada uma das estratégias de investimento apresentadas nesse trabalho.

| Indicador | Parâmetros Padrão | Parâmetros otimizado |
|----------------------|-------------------|----------------------|
| MACD | 1% | 14% |
| IFR | -13% | 54% |
| MACD + IFR Independ. | 39% | 89% |
| MACD + IFR depend. | 39% | 83% |

Tabela 12 – Resultado médios das estratégias analisadas

Nesse trabalho seleciona-se propositalmente intervalos de dados de estimação e validação um pouco distintos, conforme mostra as figuras 12 e 13 no capítulo Resultados. Selecionamos para o período de estimativa um momento mais estável e de alta crescente e para teste, escolhemos um período de altas e baixas intercalados. Essa escolha foi proposital para testar o desempenho do algoritmo em cenário um pouco diferente daquele que ele foi treinado. Observamos que apesar do desempenho ter reduzido, os resultados continuaram satisfatórios e superiores a utilização de parâmetros padrões.

Esse trabalho apresentou um indicador do benefício de utilizar indicadores com parâmetros otimizados no investimento de ações. No entanto ainda existem muitos outros estudos que podem ser feitos para comprovar esse benefício ou aprimorar os ganhos obtidos. Os testes realizados nesse trabalho poderiam ser

estendidos a uma gama maior de ativos para verificar se o benefício acontece de fato para qualquer ativo e qual seria o benefício médio obtido pela otimização.

Outras estratégias de investimentos poderiam complementar esse trabalho em busca de melhores resultados, assim como a utilização de outras técnicas de otimização além de algoritmos genéticos poderiam ser utilizado para tentar obter os parâmetros ótimos.

7. Referência bibliográfica

1. Palisade. Empresa desenvolvedora do software de otimização genética Evolver. Disponível em: <<http://www.palisade.com/>>. Acesso em: 02 set. 2011.
2. Goldberg, David E. Genetic Algorithms in Search, Optimization, and machine Learning, Addison-Wesley Publishing Company, Inc. 1989.
3. Koza, John R.: Genetic Programming: On the Programming of Computers by Means of Natural Selection, MIT Press, USA, 1992.
4. Michalewicz. Zbigniew, Genetic Algorithms + Data Structures = Evolution programs, Springe-Verlag, USA, 1996.
5. Holland, J. H.: Adaptation in Natural and Artificial Systems. MIT Press, Boston, MA, 1992.
6. Mitchell, Melanie. An Introduction to Genetic Algorithms, The MIT Press, USA, 1994.
7. Back, T. Evolutionary Algorithms in Theory and Practice, Oxford University Press, 1996.
8. Fogel, L. J., Owens, A.J, Walsh, M.J.: Artificial Intelligence Through Simulated Evolution. John Wiley and Sons, NY, 1966.
9. Davis, Lawrence, Handbook of Genetic Algorithms, Van Nostrand Reinhold, USA, 1991.
10. VIDOTTO, Rodrigo Silva. O Moving Average Convergence-Divergence como ferramenta para a decisão de investimento no mercado de ações. RAC, Curitiba, v. 13, n. 2, art. 7, p. 291-309, abr/jun. 2009.
11. LOGDEMSN, Como usar o indicador IFR (Índice de Força Relativa) na análise técnica. Disponível em: <<http://www.logdemsn.com/2008/03/24/como-usar-o-indicador-ifr-indice-de-forca-relativa-na-analise-tecnica/>>. Acesso em: 14 nov. 2011.
12. Abe Marcos, Análise Técnica: Divergência entre Preços e Indicadores. Disponível em:

<http://wiki.advfn.com/pt/An%C3%A1lise_T%C3%A9cnica:_Diverg%C3%Aancia_entre_Pre%C3%A7os_e_Indicadores>. Acesso em 14 nov. 2011.

13. ADVFN Brasil, Índice de Força Relativa. Disponível em: <http://wiki.advfn.com/pt/%C3%8Dndice_de_For%C3%A7a_Relativa>. Acesso em 18 nov. 2011.

14. Vieira, Dalton. IFR (Índice de Força Relativa) – Identifique momentos interessantes de compra ou venda. Disponível em: <<http://daltonvieira.com/ifr-indice-de-forca-relativa-identifique-momentos-interessantes-de-compra-ou-venda>>. Acesso em: 20 out. 2011.

15. Investmax. IFR - Índice de Força Relativa RSI - Relative Strength Index. Disponível em: <<http://www.investmax.com.br/iM/content.asp?contentid=649>>. Acesso em: 10 out. 2011.

16. Investmax. MACD - Média Móvel Divergente/Convergente. Disponível em: <<http://www.investmax.com.br/iM/content.asp?contentid=659>> Acesso em: 07 nov. 2011

17. Ribeiro Rafael. Teste empírico revela a eficiência do indicador MACD com as ações brasileiras. Disponível em: <<http://www.infomoney.com.br/mercados/noticia/1756796>>. Acesso em: 22 set. 2011

18. Martins Leandro. Análise Gráfica – Indicadores MACD e IFR. Disponível em: <<http://dinheirama.com/blog/2009/10/20/analise-grafica-indicadores-macd-e-ifr/>>. Acesso em: 12 nov. 2011.

19. Manual de Análise Técnica. Disponível em: <<http://pt.scribd.com/doc/6569504/Acoes-Manual-de-Analise-Tecnica>>. Acesso em: 20 set. 2011.

20. StockCharts.com. Moving Average Convergence-Divergence (MACD). Disponível em: <http://stockcharts.com/school/doku.php?id=chart_school:technical_indicators:moving_average_conve>. Acesso em :19 out. 2011.

21. Nonel, All. O Indicador MACD: Review. Disponível em: 18 out. 2011. <<http://aespeculadora.blogspot.com/2011/03/macd-indicador.html>>. Acesso em:

22. Investopedia. Moving Average Convergence Divergence – MACD. Disponível em: <<http://www.investopedia.com/terms/m/macd.asp#axzz1dmpxQIZt>>. Acesso em: 24 out. 2011.