

1 2 9 0



UNIVERSIDADE D
COIMBRA

Simão Oliveira da Silva

**DEFINIÇÃO E OTIMIZAÇÃO DE ESTRATÉGIAS DE
INVESTIMENTO NO MERCADO ACIONISTA
BASEADAS EM INDICADORES FUNDAMENTAIS**

Trabalho de Projeto no âmbito do Mestrado em Economia, na especialidade de Economia Financeira orientado pelo Professor Doutor Hélder Miguel Correia Virtuoso Sebastião e pelo Professor Doutor Pedro Manuel Cortesão Godinho e apresentado à Faculdade de Economia da Universidade de Coimbra.

Junho de 2019



FACULDADE DE ECONOMIA
UNIVERSIDADE DE
COIMBRA

Definição e Otimização de Estratégias de Investimento no Mercado Acionista baseadas em Indicadores Fundamentais

Simão Oliveira da Silva

Trabalho de Projeto no âmbito do Mestrado em Economia, na especialidade de Economia Financeira, orientado pelo Professor Doutor Hélder Miguel Correia Virtuoso Sebastião e pelo Professor Doutor Pedro Manuel Cortesão Godinho, apresentado à Faculdade de Economia da Universidade de Coimbra para a obtenção do grau de Mestre.

Junho 2019

Resumo

Este trabalho analisa o sucesso de estratégias de investimento baseadas em indicadores globais, nos principais mercados acionistas europeus entre 1992 e 2016, e testa diferentes metodologias assentes em técnicas lineares e não-lineares (técnicas de *machine learning*) com o objetivo de potenciar os resultados das estratégias de investimento.

Tendo como principal referência o artigo de Piotroski (2000), foi adotada a estratégia F_score, capaz de proporcionar avultados retornos no mercado acionista Norte-Americano entre 1976 e 1996, e definiu-se uma nova estratégia, assente num novo indicador, ao qual denominamos de N_score, que incorpora além das variáveis fundamentais do F_score os fatores *book-to-market* (BM) e dimensão. Estas duas estratégias de investimento baseadas no F_score de Piotroski (2000) e no N_score revelam grande capacidade em identificar vencedores e perdedores, proporcionando elevados retornos, ainda que os mesmos dependam em grande parte da exequibilidade das posições curtas sobre as empresas de baixo *score*. Constatamos que a estratégia F_score quer aplicada a empresas de valor quer aplicada a todas as empresas proporciona retornos médios superiores a 20% ao ano, o que demonstra que o sucesso desta estratégia não é restrito ao elevado contexto BM, nem circunstancial ao espaço (noutros mercados que não o Norte-Americano) e ao tempo (diferente período amostral). Por último, os resultados sugerem que a incorporação dos fatores BM e dimensão não contribuem para aumentar a capacidade preditiva do F_score, na medida em que a estratégia N_score possibilita, em média, retornos anuais inferiores à estratégia F_score, quer com a compra quer com a venda a descoberto.

Do processo de otimização das estratégias F_score e N_score mostramos que o método (OLP_rt), em que se ponderam as variáveis *dummy* de modo a maximizar a rentabilidade do portefólio de elevado *score*, possibilita um acréscimo de aproximadamente 1,5% anualmente e um maior rácio de Sharpe. Contudo, concluímos que a metodologia simples e linear, inerente às estratégias de investimento adotadas (F_score e N_score), revela uma elevada eficiência dada a dificuldade em potenciar os retornos conseguidos com a compra de ações com elevado *score*.

Palavras-chave: mercados de capitais; análise fundamental; *book-to-market*.

Abstract

This paper analyses the success of investment strategies based on global indicators in the main European stock markets between 1992 and 2016, and tests different methodologies based on linear and non-linear techniques (machine learning techniques) in order to outperform the investment strategies.

The F_score strategy, from Piotroski (2000), which was able to provide large returns in the North American stock market between 1976 and 1996, was adopted as the main reference, and a new strategy was established, based on a new indicator, the N_score, which incorporates in addition to the fundamental variables of F_score the book-to-market (BM) and size factors. These two investment strategies based on Piotroski's (2000) F_score and N_score show great ability to identify winners and losers on the high returns they guarantee, even if they depend to a large extent on the feasibility of short positions on low score companies. We find that the F_score strategy applied to value companies as well as to all companies provides average returns of more than 20% per year, which demonstrates that the success of this strategy is not restricted to the high BM context, or circumstantial to space (in other markets than North American) and time (different sample period). Finally, the results suggest that the incorporation of BM and size factors do not contribute to increase the predictive capacity of F_score, since the N_score strategy allows, on average, annual returns lower than the F_score strategy, either with the purchase or with the short sale.

From the optimization process of the F_score and N_score strategies we show that the method (OLP_rt), in which the dummy variables are weighted in order to maximize the profitability of the high score portfolio, allows an increase of approximately 1.5% annually and a higher Sharpe ratio. However, we conclude that the simple and linear methodology, inherent to the investment strategies adopted (F_score and N_score), reveals a high efficiency given the difficulty in boosting the returns obtained with the purchase of high score stocks.

Keywords: capital markets; financial statement analysis; book-to-market.

Índice

1. Introdução	1
2. Revisão da Literatura	3
2.1. O Efeito <i>Book-to-Market</i> (BM)	3
2.2 Estratégias baseadas em Análise Fundamental.....	3
2.3 Técnicas de <i>machine learning</i> ao serviço da Análise Fundamental	10
3. Metodologia	13
3.1 Estratégias de Investimento baseadas num Indicador Global.....	13
3.1.1 Estratégia F_score de Piotroski (2000)	13
3.1.2 Estratégia N_score	14
3.1.3 Cálculo dos retornos	14
3.2 Otimização	15
3.2.1 Otimização Linear dos Ponderadores (OLP)	16
3.2.2 Regressão Linear (RL) e Técnicas de <i>Machine Learning</i>	17
4. Dados e Análise Preliminar.....	19
5. Resultados.....	25
5.1 Estratégias de Investimento baseadas em Indicadores Fundamentais	25
5.1.1 Estratégia F_score para empresas de valor.....	25
5.1.2 Estratégia F-score para todas as empresas.....	29
5.1.3 Estratégia N_score	31
5.2 Otimização das Estratégias de Investimento	33
5.2.1 Otimização da Estratégia F_score	34
5.2.2 Otimização da Estratégia N_score	39
6. Conclusão.....	45
7. Bibliografia	47
Apêndice	50
Anexos	51

1. Introdução

Piotroski (2000) demonstra a utilidade da análise fundamental na definição de estratégias de investimento aplicadas às ações de empresas de valor norte americanas entre 1976 e 1996.¹ Este autor desenvolve uma estratégia de investimento baseada num indicador fundamental global, o F_score, que agrega 9 variáveis fundamentais na forma binária com a finalidade de mensurar a força financeira da empresa, a partir da qual identifica potenciais vencedores e perdedores. A simplicidade metodológica aliada aos elevados retornos proporcionados por esta estratégia despertou, ao longo das últimas duas décadas, investigação de estratégias de investimento assentes na análise de informação retirada das demonstrações financeiras. Vários estudos constataam a capacidade da estratégia de investimento criada em Piotroski (2000) aplicada a empresas de valor e mesmo independentemente do seu *book-to-market* (BM), em diferentes mercados acionistas, colocando em causa a hipótese de eficiência dos mercados. Um dos objetivos do presente trabalho de projeto passa por replicar a estratégia F_score de Piotroski (2000) num contexto europeu (1992-2016) quer aplicada às empresas de valor quer a todas as empresas da amostra, a fim de perceber se a mesma é eficaz e se o contexto BM é importante no sucesso da mesma. Pretende-se também estudar o desempenho de uma nova estratégia assente num novo indicador global, o N_score que, para além das variáveis consideradas no F_score, incorpora o fator BM e fator dimensão, tendo por base os resultados de Fama e French (1992), os quais mostram que a rentabilidade das ações está diretamente relacionada com o rácio BM e indiretamente com a capitalização de mercado.

Não obstante o sucesso e eficácia da estratégia baseada no F_score, o indicador fundamental é obtido de forma linear através da soma não ponderada das nove variáveis fundamentais, o que atribui igual peso/importância a cada uma delas na identificação dos vencedores e perdedores. Deste modo, torna-se pertinente a investigação de estratégias de investimento baseadas na análise fundamental que incorporem técnicas de *machine learning*, permitindo modelar uma relação complexa e não-linear entre as diferentes variáveis fundamentais e os lucros obtidos, obtendo dessa forma indicadores fundamentais mais eficientes.² É na simplicidade da metodologia usada por Piotroski (2000) que reside a principal motivação do

¹ As empresas de valor são empresas que apresentam um elevado valor contabilístico face ao valor de mercado, ou seja, um elevado rácio *book-to-market* (BM).

² Shen (2011) crítica o facto de o indicador fundamental ser obtido de forma linear pela soma não ponderada das nove variáveis fundamentais, o que em princípio contrasta com a não-linearidade e extrema volatilidade do mercado acionista.

presente trabalho de projeto, que procura através de métodos lineares e não-lineares (técnicas de *machine learning*) potenciar a capacidade das estratégias baseadas em indicadores fundamentais, F_score e N_score, em identificar vencedores nos principais mercados acionistas europeus.

O trabalho de projeto está organizado da seguinte forma. Após a introdução segue-se a apresentação da revisão da literatura na Secção 2, de seguida explica-se a metodologia empregue na definição das estratégias de investimento e dos métodos de otimização testados. Na Secção 4 descreve-se todo o processo relativo a formação da base de dados e efetua-se uma breve análise preliminar das estatísticas descritivas, enquanto a Secção 5 destina-se à apresentação e discussão dos resultados obtidos pelas estratégias de investimento (F_score e N_score) e pelos métodos de otimização testados. Por último, a Secção 6 apresenta as principais conclusões do presente trabalho.

2. Revisão da Literatura

A presente secção começa por apresentar o efeito BM, de seguida na Subsecção 1.2 são enunciados diversos contributos relativos à construção de estratégias de investimento assentes na análise fundamental, tendo como principal foco a estratégia F_score de Piotroski (2000). A última secção destina-se à apresentação de artigos que incorporaram técnicas de *machine learning* ao serviço de estratégia de investimento fundamental baseada no F_score de Piotroski (2000).

2.1. O Efeito *Book-to-Market* (BM)

Diversos autores, entre os quais Fama e French (1992) e Lakonishok et al. (1994), demonstram que portefólios constituídos por empresas de elevado rácio BM (também designadas de empresas de valor) apresentam retornos superiores aos portefólios de empresas de baixo rácio BM (ou empresas de crescimento), evidenciando uma forte correlação entre o rácio BM e o desempenho futuro das empresas, o designado efeito *book-to-market*. Embora o efeito BM seja consensual na literatura, a sua justificação divide-se em duas teorias distintas. Por um lado, Fama e French (1992) defendem uma explicação baseada no risco, de acordo com a qual os retornos excessivos das empresas de elevado rácio BM face às demais, se devem à compensação do risco enfrentado, uma vez que estas empresas tendem a vivenciar situações de *stress* financeiro (ver também Fama e French, 1995; Chen e Zhang, 1998). Por outro, Lakonishok et al. (1994) argumentam que a diferença dos retornos entre empresas de valor e empresas de crescimento se deve à ineficiência do mercado, *market mispricing*. Isto é, os autores defendem que as dificuldades financeiras que as empresas de valor vivenciam e os fracos resultados que apresentam levam à formação de expectativas demasiado pessimistas sobre o seu futuro desempenho, enquanto para as empresas de crescimento, os investidores formam expectativas demasiado otimistas devido à extrapolação que fazem dos recentes ganhos e do crescimento destas empresas. Deste modo, o desajustamento das expectativas revela-se nos períodos futuros, através de ganhos positivos das empresas de valor, tendencialmente subestimadas, e de retornos negativos nas empresas de crescimento, frequentemente sobrestimadas, (La Porta et al., 1997).

2.2 Estratégias baseadas em Análise Fundamental

Na construção de estratégias de investimento são utilizadas essencialmente duas técnicas de análise para prever a evolução futura dos preços das ações: a análise técnica e a análise

fundamental. A análise técnica procura extrair tendências e regras de investimento a partir das transações passadas do mercado, enquanto a análise fundamental, tema do presente trabalho, tem como objetivo determinar o real valor das ações com base na análise de fatores económicos e financeiros, permitindo a identificação das ações sobre e subvalorizadas mediante a comparação entre o preço de mercado e o valor das ações.

O estudo da análise fundamental é de extrema relevância, sobretudo para os investidores, dada a sua eventual capacidade em prever a evolução dos preços das ações. Por isso, muitos artigos estudam a utilidade das informações retiradas das demonstrações financeiras, nas previsões dos retornos no mercado acionista. Um dos primeiros contributos foi dado por Ou e Penman (1989) que mostram que um conjunto de rácios financeiros pode prever, de forma eficaz, futuras mudanças nos ganhos das empresas, no entanto a metodologia apresentada pelos autores, assente num modelo logit, apenas é aplicável considerando uma grande quantidade de informação histórica dado ao elevado número de rácios financeiros utilizados, o que eleva os custos e reduz a sua praticabilidade. Tendo em consideração essas limitações, Lev e Thiagarajan (1993) utilizam apenas 12 indicadores financeiros reconhecidos como úteis para os analistas financeiros, e provam que esses indicadores estão correlacionados com os retornos contemporâneos. Posteriormente, Abarbanell e Bushee (1998) argumentam ser possível a criação de uma estratégia de investimento baseada nesses indicadores, capaz de proporcionar retornos anormalmente elevados.³ Já Sloan (1996) mostra que um elevado valor de *accruals* nos ganhos das empresas, tende a refletir diminuições da rentabilidade futura.

Ao invés de estudar a relação entre cada um dos indicadores e os retornos, Piotroski (2000) desenvolve uma estratégia de investimento baseada num único indicador global fundamental, que incorpora informações presentes nas demonstrações financeiras, ao qual denomina de *F_score*. A estratégia foi aplicada a empresas de valor, uma vez que o autor defende que as demonstrações financeiras representam a fonte de informação mais confiável e acessível para as empresas de valor. Em primeiro lugar, porque as ações destas empresas tendem a ser negligenciadas e pouco seguidas pelos analistas, o que sugere que as previsões e recomendações dos analistas financeiros sejam escassas ou pouco fiáveis. Outra razão enunciada pelo autor prende-se com o facto de as empresas de elevado rácio BM terem

³ Abarbanell e Bushee (1998) constataam que de entre os 12 sinais analisados, são as relativas mudanças nos inventários, as despesas de capital e as taxas de imposto efetivas, os indicadores com maior capacidade em prever os retornos a 1 ano.

acesso limitado aos grandes canais de informação (os média), e as suas divulgações voluntárias poderem não ser credíveis, visto que estas empresas tendem a enfrentar condições financeiras extremamente difíceis (Fama e French, 1995; Chen e Zhang, 1998). Por último, face a debilidade financeira, Piotroski (2000) argumenta que a avaliação destas empresas deve ter como principal foco aspetos financeiros como a alavancagem financeira, liquidez, tendências de rentabilidade e adequação de *cash flows*, que se obtêm a partir das demonstrações financeiras. A estratégia de investimento em valor de Piotroski (2000) consiste na compra e venda de ações com base na força financeira transmitida pelo indicador global. O F_score compreende três áreas financeiras: rentabilidade, liquidez e eficiência operativa, sendo constituído por nove indicadores fundamentais, obtidos a partir de informações presentes nas demonstrações financeiras, nomeadamente a rentabilidade dos ativos, *cash flows* das atividades operacionais, *accruals*, aumentos de capital, quer variações da rentabilidade dos ativos, da liquidez, da alavancagem financeira, do rácio da margem bruta, da rotação do ativo. Estes *inputs* são depois transformados em variáveis *dummy*, que assumem o valor 1 caso o seu valor seja indicativo de saúde financeira da empresa, e o valor 0 caso contrário. Deste modo, o F_score para cada empresa assume um número inteiro entre 0 e 9. Piotroski (2000) mostra que o retorno médio ajustado ao mercado obtido pelo investimento em ações de elevado rácio BM pode obter um acréscimo de 7,5% anualmente através da seleção de empresas com elevado F_score (8 e 9). Adicionalmente, a estratégia de investimento que consiste simultaneamente na compra das ações de empresas financeiramente fortes, ou seja, com elevado F_score (8 e 9) e na venda a descoberto das ações de empresas financeiramente fracas, isto é, com baixo F_score (0 e 1), possibilitou anualmente, para o mercado acionista Norte-Americano entre 1976 e 1996, um retorno médio ajustado ao mercado de 23%. O autor evidencia a robustez da sua estratégia ao longo do tempo e face a outros fatores, tais como o preço das ações, o volume transacionado, o acompanhamento de analistas e verificou que os maiores benefícios da estratégia estavam concentrados em pequenas e médias empresas, empresas com baixo *turnover*, e empresas sem recomendações de analistas, concluindo que os elevados retornos obtidos da estratégia estavam associados a um ambiente de baixa disseminação de informação.

Baseado no trabalho desenvolvido em Piotroski (2000), Mohanram (2005) prova ser possível a criação de uma estratégia de investimento fundamental baseada igualmente num indicador global (a que denomina de G_score), que incorpora informações das demonstrações financeiras, para as empresas de crescimento (de baixo rácio BM) capaz de proporcionar

elevados retornos. O autor alega que *a priori* não existe uma relação clara quanto à eficácia da análise de informações das demonstrações financeiras neste tipo de empresas, devido ao facto de estas disporem de outras fontes de divulgação de informação (como os média) e por atraírem intervenientes sofisticados, como analistas e investidores institucionais, que face ao bom desempenho e rápido crescimento destas empresas tendem a utilizar outras medidas não financeiras.⁴ No entanto, o autor sugere que embora a análise fundamental tradicional possa ter aplicabilidade limitada para as empresas de crescimento, outras informações presentes nas demonstrações financeiras podem ser úteis para captar a sobrevalorização decorrente da *naive extrapolation* (“extrapolação ingénu”) dos recentes ganhos e do crescimento destas empresas. Em função disso, Mohanram (2005) incorpora no seu *G_score* não só variáveis da análise tradicional financeira, tal como em Piotroski (2000), mas outras informações presentes, segundo o autor, na *conservative accounting* (“contabilidade conservadora”), como despesas em I&D e publicidade, e variáveis relacionadas com a “extrapolação ingénu” (variabilidade dos resultados e dos crescimentos das vendas), de forma a adaptar o indicador global às circunstâncias deste universo de empresas.⁵ O cálculo do indicador global de Mohanram (2005) consiste, à semelhança do *F_score*, na soma dos sinais (0 ou 1) relativos às variáveis que compõem o *G_score*. Todavia, o processo de transformação das 8 variáveis que compõem o indicador global em sinais (ou variáveis binárias), foi realizado através da comparação *cross-section* contemporânea dentro do mesmo sector, usando a mediana como medida de referência. A estratégia baseada no *G_score*, que consiste na compra de ações de empresas com elevado *G_score* (6, 7 e 8) e na venda a descoberto de ações de empresas com baixo *G_score* (0 e 1), proporcionou um retorno médio anual ajustado à dimensão de 20,6%, sendo robusta ao longo do tempo e face a fatores como a dimensão, acompanhamento de analistas, liquidez, *momentum*, *book-to-market* e *accruals*.⁶ O autor constata que os resultados da estratégia são potenciados para os segmentos compostos pelas empresas de maior dimensão e pelas empresas mais seguidas pelos analistas, sugerindo que a análise fundamental é mais eficaz em ambientes de rápida disseminação de informação para as empresas de crescimento. Por último, Mohanram (2005)

⁴ Por exemplo, medidas que incidem sobre a reputação da empresa, influência e valor para o consumidor, competitividade e inovação (alguns exemplos específicos são a preferência pela marca, a quota de mercado, a taxa de resposta “take rate”, etc.).

⁵ Penman e Zhang (2002) demonstram que o mercado tende a ignorar as implicações da “contabilidade conservadora”, como as despesas em I&D e em publicidade, sobre os resultados futuros.

⁶ Mohanram (2005) obtém o retorno ajustado à dimensão pela diferença entre o retorno da empresa e o retorno para o mesmo período do decil de capitalização de mercado da empresa na CRSP.

analisa a importância do contexto BM no sucesso na análise fundamental, através da replicação da estratégia F_score de Piotroski (2000) a empresas de crescimento e da estratégia G_score a empresas de valor. O autor verifica que ambas as estratégias veem os seus retornos significativamente reduzidos nestes novos contextos, e dessa forma, concluiu a importância do contexto BM na análise fundamental baseada em informações retiradas de demonstrações financeiras.

Não obstante o artigo de Mohanram (2005) ter demonstrado o sucesso da análise de informações de demonstrações financeiras no segmento de mercado das empresas de crescimento, o G-score está sujeito a algumas críticas, discutidas em Piotroski (2005). Uma das críticas prende-se com a elevada dependência dos retornos provenientes do lado curto da estratégia baseada no G_score, (as vendas a descoberto contribuem com um retorno médio anual ajustado de 17,5%, face aos 3,1% obtidos com a compra de ações com elevado G_score), o que poderá pôr em causa a efetiva existência de retornos avultados caso existam restrições às vendas a descoberto. Outra das críticas apontadas por Piotroski (2005) está relacionada com o objetivo implícito ao G_score consistir na identificação das empresas que não tendem a verificar as fortes expectativas de crescimento. Isto é, a estratégia de Mohanram (2005) tem o fator crescimento como foco principal, negligenciando o papel do desempenho fundamental na determinação do valor da empresa, o que poderá explicar não só a maior capacidade do G_score em diferenciar vencedores de perdedores na subamostra de empresas de rápido crescimento, como o elevado peso das vendas a descoberto nos retornos desta estratégia. Além disso, e ainda que Solimon (2004), tenha demonstrado a utilidade de utilizar *benchmarks* relativos ao setor, Piotroski (2005) defende que esta metodologia pode colocar alguns problemas, pois pode-se avaliar incorretamente metade das empresas de um setor no caso deste estar subvalorizado ou sobrevalorizado num certo momento de tempo.⁷ Apesar das críticas expostas, Piotroski (2005) reconhece a importância do artigo de Mohanram (2005), concluindo que juntos, os dois artigos (Piotroski, 2000; Mohanram, 2005) sugerem que é possível mudar a distribuição dos retornos de um determinado universo de empresas contextualizado pelo BM (empresas de valor para o F_score, e empresas de crescimento para o G_score) com base na análise de demonstrações financeiras. Por último, Piotroski (2005) destaca que em ambos os artigos, os retornos das estratégias parecem ser explicados por

⁷ Imagine-se que um determinado setor está subvalorizado, através desta metodologia discriminada por setores apenas metade das empresas obtêm o valor 1 para cada uma das variáveis fundamentais, logo poderemos estar a excluir empresas que estão subvalorizadas e deveriam fazer parte do portefólio de elevado F_score.

erros de avaliação sistemáticos, ainda que por mecanismos diferentes, o que suporta a explicação do efeito BM com base na teoria *market mispricing*.⁸

Os contributos de Mohanram (2005), e principalmente de Piotroski (2000), promoveram o interesse de muitos investigadores relativamente a estratégias de investimento fundamental assentes num indicador global. Vários autores replicaram a estratégia de Piotroski (2000) para empresas de elevado BM em diferentes mercados, provando o sucesso da metodologia fora do mercado Norte-Americano, como em Lopes e Galdi (2006) e Aggarwal e Gupta, (2009), no Brasil e Índia, respetivamente. Independentemente de Mohanram (2005) e Piotroski (2000) defenderem a importância do contexto BM no sucesso das suas estratégias, muitos investigadores estudaram, recentemente, a capacidade da estratégia F_score, em identificar vencedores e perdedores independentemente do BM em diferentes mercados. Concretamente para o contexto europeu, Amor-Tapia e Tascón (2016) provam o sucesso e robustez da estratégia F_score, entre 1989 e 2011 selecionando empresas de 14 países europeus, assim como Navas et al. (2018) que utilizam empresas pertencentes ao Euronext 100 para o período 2000-2016. Já Dosamantes (2013) demonstra a eficácia da estratégia de Piotroski para a bolsa de valores mexicana entre 1991 e 2011. Conjuntamente, estes três artigos indicam que em diferentes contextos a estratégia F_score (compra e venda das ações com maior e menor F_score, respetivamente) é robusta e garante, em média, uma rentabilidade anual superior em 20% à rentabilidade do mercado, evidenciando dessa forma, o sucesso da estratégia baseada no F_score para todas as empresas, independentemente do contexto BM.⁹

Noma (2010) e Iqbal et al. (2013), por outro lado estudam novas estratégias de investimento fundamental ainda que baseadas na metodologia de Piotroski (2000). Noma (2010) construiu uma estratégia baseada num indicador composto (também designado F_score) por 3 variáveis fundamentais utilizadas no F_score original de Piotroski para empresas da bolsa de Tóquio entre 1986 e 2001.¹⁰ A estratégia consiste na compra de ações com F_score igual

⁸ De acordo com Piotroski (2000) os retornos das empresas de valor devem-se em grande parte à lenta reação ou negligência do mercado, potenciada por ambientes de baixa disseminação de informação, enquanto para Mohanram (2005) os retornos da estratégia para empresas de baixo BM resultam da má interpretação das informações (empresas com elevado G_score apresentam retornos superiores e um nível de risco inferior), uma vez que os retornos da estratégia são superiores em ambientes de elevada disseminação de informação.

⁹ Piotroski e So (2012) verificam o sucesso da estratégia F_score para o mercado acionista Norte-Americano entre 1972 e 2010, apresentando retornos ajustados da dimensão significativos para toda a amostra de empresas (independentemente do rácio BM).

¹⁰ Em Noma (2010) as variáveis fundamentais utilizadas são: variação da rentabilidade dos ativos, *cash flows* das atividades operacionais e variação do rácio da margem bruta.

a 3 e na venda a descoberto das ações com um F_score igual a 0. Noma (2010) constata o sucesso da sua estratégia independentemente do contexto BM, uma vez que quando aplicada a todas as empresas, a empresas de valor e empresas de crescimento permite, em média, um retorno médio ajustado ao mercado de 7,8%, 9,1% e 8,8%, respetivamente. Em conformidade com Piotroski (2000), Noma (2010) mostra que a estratégia F_score utilizada é mais eficaz para o segmento das empresas de elevado rácio BM, e para a subamostras constituídas por pequenas e médias empresas, sendo que os retornos proporcionados pela estratégia são inconsistentes com o ajustamento ao risco, uma vez que a relação entre F_score e o risco não é positiva.¹¹ Por sua vez, Iqbal et al. (2013) desenvolvem uma estratégia baseada num indicador global para empresas cotadas na bolsa de Karachi entre 2000 e 2009. O seu indicador global é composto por cinco indicadores, 4 deles presentes no F_score de Piotroski (rácio de alavancagem, rácio *cash flow*, ROA, rácio dos *accruals*), enquanto o quinto corresponde à variação do rácio BM, sendo atribuído o valor 1 no caso de o valor do rácio BM diminuir de um ano para o outro. Recorrendo a uma estimação através do método dos mínimos quadrados do modelo em que os retornos das ações são explicados pelos retornos do período anterior, pelo F_score, e por duas variáveis *dummy* que controlam os fatores relativos ao sector e ao tempo, os autores constataam que o contributo do seu F_score apesar de positivo não é significativo para identificar vencedores e perdedores no mercado bolsista do Paquistão.

Não obstante, Noma (2010) ter utilizado um indicador global diferente do de Piotroski (2000), o seu principal contributo é demonstrar que a estratégia definida pela compra das ações de valor com elevado F_score (com um valor 3) e venda a descoberto das ações de crescimento com baixo F_score permite aumentar os retornos excedentários. O autor mostra ainda que esta estratégia potencia os resultados da estratégia fundamental baseada no F_score adotado, proporcionando um retorno médio anual ajustado ao mercado de 17,6%, isto é, quase o dobro do obtido pela estratégia aplicada às empresas de valor (compra e venda a descoberto de ações de valor com elevado e baixo F_score, respetivamente). A potencialidade da estratégia desenvolvida por Noma (2010) foi igualmente verificada no artigo de Piotroski e So (2012), o qual investiga se os ganhos da estratégia valor/crescimento (compra de ações de valor e venda das ações de crescimento), no mercado bolsista americano

¹¹ Noma (2010) constata que os retornos obtidos pela estratégia baseada no F_score adaptado aplicada a todas as empresas não é consistente com o risco, uma vez que o portefólio de elevado *score* apresenta um menor risco que o portefólio de baixo *score*. Desta forma é a má avaliação do mercado (*market mispricing*) que explica os elevados retornos obtidos pela estratégia F_score.

entre 1972 e 2010, eram explicados por erros nas expectativas de mercado (teoria *market mispricing*). Neste artigo, a estratégia designada de estratégia valor/crescimento incongruente que consiste na compra de ações de valor com elevado F_score e venda de ações de crescimento com baixo F_score, apresenta, tal como em Noma (2010), retornos excedentários superiores à estratégia baseada no F_score quando aplicada quer para todas as empresas quer para os segmentos de empresas de valor ou de crescimento.¹² Os autores constataam que a estratégia valor/crescimento incongruente gera avultados retornos ajustados pela dimensão (22,64% e 37,66%, valor médio a 1 e 2 anos, respetivamente), enquanto os retornos da estratégia valor/crescimento congruente (compra de ações de valor com baixo F_score e a venda de ações de crescimento com elevado F_score) são economicamente indistinguíveis de zero (0,14% e -2,29%, ambos sem significância estatística). Estes resultados fundamentam a hipótese de que o efeito valor/crescimento ou efeito BM depende fortemente dos erros de avaliação do mercado que podem ser identificados *ex ante* quando a força dos sinais fundamentais é incongruente com as expectativas de mercado implícitas no rácio BM. Isto significa que os altos retornos das ações de valor devem-se a empresas de valor com forte valor fundamental, enquanto os baixos retornos das ações de crescimento estão concentrados nas empresas com reduzido valor fundamental, o que explica os elevados retornos da estratégia valor/crescimento incongruente. As conclusões de Piotroski e So (2012) relativamente à explicação do efeito BM foram estendidas ao mercado europeu, entre 1990 e 2013, por Walkshäusl (2017) o que mais uma vez sustenta a teoria baseada nos erros das expectativas de mercado, *market mispricing*.

2.3 Técnicas de *machine learning* ao serviço da Análise Fundamental

As técnicas de *machine learning* apresentam elevada capacidade de identificar regras e padrões a partir de grandes conjuntos de dados. Fruto do forte desenvolvimento tecnológico nos últimos anos, (que facilitou a recolha e armazenamento de dados), estas técnicas têm sido cada vez mais utilizadas, sendo bastante úteis em diversas áreas desde a deteção de tumores, descoberta de novos medicamentos, reconhecimento facial, previsão de produção

¹² Estratégia valor/crescimento incongruente definida em Piotroski e So (2012) consiste na compra de ações de valor e venda de ações de crescimento cujo valor fundamental contrasta com as expectativas de mercado. Deste modo, a estratégia define-se na compra de ações de valor com elevado F_score (empresas que tendem a estar subvalorizadas, pois apresentam baixas expectativas de mercado, e apresentam um forte valor fundamental) e na venda a descoberto das ações de crescimento com baixo F_score (empresas que contrastam as suas elevadas expectativas de mercado com o reduzido valor fundamental, e por isso tendem a estar sobrevalorizadas). Piotroski e So (2012) adotam o F_score definido em Piotroski (2000).

de energia, até à otimização de estratégias de transação baseadas em análise técnica (Witten e Frank, 2005). Deste modo torna-se relevante investigar a utilização destas técnicas com o objetivo de otimizar estratégias de transação baseadas em indicadores fundamentais. Shen (2011) procura estender a praticabilidade do F_score de Piotroski (2000) através da construção de um modelo ANN (*Artificial Neural Network*) capaz de lidar com a extrema volatilidade e a não-linearidade do mercado acionista. Este autor, selecionou as 400 ações com maior rácio *book-to-market* do mercado bolsista de Taiwan (cerca de 1/3 do total de ações) entre 2008 a 2010. O modelo ANN foi projetado para aprender a relação entre as 9 variáveis fundamentais que compõem o F_score e os retornos obtidos através da compra da ação e venda após 12 e 18 meses.¹³ A estratégia de investimento em valor consistiu na compra e venda a 12 e 18 meses de 30% das ações com previsões de rentabilidade mais elevadas.¹⁴ Os resultados revelaram que a estratégia de valor apresenta retornos médios significativamente superiores assim como maiores rácios de Sharpe face à amostra total das empresas e aos índices de mercado de Taiwan, demonstrando que a utilização de técnicas de *machine learning* (modelo ANN) baseadas em indicadores fundamentais pode ser bastante útil na identificação de ações de valor com elevado potencial.

Ainda que bastante relevante, o artigo de Shen (2011) não permite perceber se a metodologia baseada numa técnica de *machine learning* (modelo ANN) possibilita retornos superiores aos da posição longa (compra de ações de valor) da estratégia de Piotroski (2000). Com essa finalidade, Balachandran et al. (2013) investigam a capacidade de potenciar a estratégia baseada no F_score com recurso a técnicas de *machine learning* no mercado acionista Norte-Americano entre 1982 e 2011. Neste artigo, a estratégia F_score foi usada como *benchmark*, e contemplou somente a compra das ações de valor pertencentes ao decil com maior F_score em cada ano. Foram utilizadas duas abordagens para bater a estratégia F_score base. Na primeira foram incorporadas apenas as 9 variáveis do F_score, enquanto na segunda se incluíram variáveis adicionais para potenciar os resultados obtidos pelas técnicas de aprendizagem avançada.¹⁵ Os modelos preditivos utilizados foram a regressão linear, regressão logística e *support vector machine* (SVM) com janelas de treino de 1 a 4 anos,

¹³ Shen (2011) na construção do F_score substitui a variável fundamental *accruals* pela rentabilidade dos capitais próprios (ROE).

¹⁴ Foram selecionadas aleatoriamente 100 ações compondo a amostra de teste, para as quais foram gerados 100 *outputs* que correspondem aos retornos previstos pelo modelo ANN tendo em consideração as 9 variáveis fundamentais. O retorno da estratégia de valor consistiu na média do retorno efetivo das 30 ações que apresentaram as previsões de rentabilidade mais elevadas.

¹⁵ Como rotação de vendas, passivos, contas a receber, saldo de caixa, entre outras.

sendo utilizados os valores reais das 9 variáveis que compõem o F_score, em vez da sua conversão para variáveis *dummy*, de forma a captar relações mais robustas entre os preditores e os retornos. Todas as estratégias testadas neste artigo apresentaram um retorno médio anual superior ao da estratégia base F_score, no entanto caracterizaram-se por serem mais arriscadas, isto é, com maiores desvios-padrão. Apenas a estratégia do modelo SVM com uma janela de 2 anos de treino obteve um retorno médio superior e um desvio-padrão semelhante ao da estratégia base. Os resultados permitiram ainda verificar que, regra geral, a incorporação de variáveis adicionais permite reduzir significativamente os desvios-padrão, melhorando a qualidade dos investimentos. Embora Balachandran et al. (2013) tenham obtido melhores resultados utilizando um modelo SVM com janelas de treino de 2 anos, concluem que a estratégia F_score de Piotroski (2000), ainda que assente numa metodologia simples e linear, apresenta um elevado desempenho.

3. Metodologia

Nesta secção é descrita a metodologia afeta às estratégias de investimento baseadas no F_score de Piotroski (2000) e no N_score, e aos métodos empregues com o objetivo de otimizar as estratégias de investimento adotadas.

3.1 Estratégias de Investimento baseadas num Indicador Global

3.1.1 Estratégia F_score de Piotroski (2000)

A estratégia designada ao longo do trabalho por F_score consiste na replicação da estratégia F_score de Piotroski (2000), que pretende identificar vencedores e perdedores tendo por base a força financeira mensurada pelo F_score. A lógica subjacente à construção do F_score (Piotroski, 2000) passa por incorporar nove indicadores fundamentais que refletem mudanças na condição económica das empresas, e por isso devem ser úteis na previsão do desempenho futuro das empresas. Esses *inputs* são transformados em variáveis *dummy*, atribuindo-se o valor de 1 no caso de se esperar que o valor desse *input* seja favorável para o desempenho futuro da empresa, e o valor de 0 caso contrário.¹⁶ O F_score é obtido a partir da soma das 9 variáveis *dummy*, podendo dessa forma assumir um valor inteiro entre 0 e 9 para cada empresa. As 9 variáveis fundamentais pretendem avaliar a condição financeira da empresa em três áreas: rentabilidade (ROA, CF, Δ ROA, *Accrual*), liquidez (Δ Lever, Δ Liquid, *Eq_offer*) e eficiência operativa (Δ Margin e Δ Turn), considerando-se que rentabilidade positiva, aumentos de rentabilidade, *cash flows* positivos, *accruals* negativos, aumentos de liquidez, diminuição do peso da dívida, não emissão de ações, aumento da margem de lucro bruta e do *turnover* (rotação do ativo) contribuem para o bom desempenho futuro de uma empresa. No Apêndice está descrita a forma de cálculo do F_score usada no presente trabalho.¹⁷ Após o cálculo anual do indicador F_score para todas as empresas entre 1992 e 2016, formaram-se portefólios com base no valor do indicador global F_score, dando igual peso a todas as empresas de cada portefólio. A Estratégia F_score consiste na compra

¹⁶ Como a estratégia foi criada para as empresas de valor, a transformação dos nove indicadores em variáveis *dummy* é feita por Piotroski (2000) sob um contexto de *stress* financeiro característico deste tipo de empresas (Fama e French, 1995; Chen e Zhang, 1998)

¹⁷ Face a restrições da base de dados algumas das variáveis fundamentais utilizadas são *proxies* das usadas em Piotroski (2000).

do portfólio constituído pelas empresas com F_score de 8 ou 9, e na venda a descoberto do portfólio de empresas com F_score igual ou inferior a 2.¹⁸

3.1.2 Estratégia N_score

Um dos objetivos deste trabalho de projeto consiste na construção de uma nova estratégia de investimento. Por esse motivo, e com o intuito de aumentar a capacidade preditiva do F_score, desenvolveu-se um novo indicador fundamental agregado, o N_score, que adiciona o rácio BM e a capitalização de mercado aos nove indicadores do F_score. A incorporação destes dois indicadores tem como suporte os resultados de Fama e French (1992), que indicam que as empresas com maior rácio BM, assim como as empresas com menor capitalização de mercado, tendem a apresentar retornos mais elevados.¹⁹ Em função disso, o rácio BM foi transformado em 1 caso o valor seja superior à mediana dos valores para o conjunto de todas as empresas no respetivo ano, enquanto o valor de capitalização de mercado foi considerado 1 se fosse inferior à mediana dos valores para o conjunto de todas as empresas no ano considerado. Desta forma, o N_score é composto por 11 indicadores na forma binária (11 variáveis *dummy*) podendo assumir um número inteiro entre 0 e 11. A estratégia N_score consiste, tal como a estratégia F_score, na compra dos portfólios de empresas com elevado *score* e venda a descoberto dos portfólios de empresas com baixo *score*. O portfólio de elevado *score* é composto pelas empresas com *score* de 9, 10 ou 11 enquanto o portfólio de baixo *score* é formado pelo conjunto de empresas que apresentam um *score* igual ou inferior a 3.²⁰

3.1.3 Cálculo dos retornos

Todos os retornos são obtidos anualmente. É utilizado o índice de retorno total, por forma a ter em conta o reinvestimento dos dividendos. Para garantir que toda a informação anual financeira está disponível aquando da formação dos portfólios, os retornos foram calculados a partir do final de abril (momento de formação dos portfólios). Dessa forma, o retorno (bruto) relativo ao F_score do ano t , foi obtido através da diferença entre o logaritmo

¹⁸ Piotroski (2000) definiu a estratégia na compra das ações de empresas com F_score igual ou superior a 8 e na venda das ações com F_score igual ou inferior a 1. Devido relativa escassez de empresas com *score* igual ou inferior a 1, foram incorporadas as empresas com *score* 2 no portfólio de baixo *score*, o que em Piotroski (2000) não altera os resultados qualitativamente.

¹⁹ Fama e French (1992) criaram o modelo geralmente designado de modelo a 3 fatores, no qual os retornos das ações são explicados pelos fatores: risco de mercado, risco de valor e risco de dimensão.

²⁰ De modo a obter-se pesos de empresas com elevado e baixo *score* semelhantes às obtidas em Piotroski (2000).

do índice de retorno total no final de abril do ano $t+2$ e o logaritmo do índice de retorno total no final de abril do ano $t+1$. Os retornos ajustados ao mercado correspondem à diferença entre o retorno (bruto) de uma empresa e o retorno anual do respetivo índice de mercado.

3.2 Otimização

No processo de otimização das estratégias *F_score* e *N_score* apenas foram consideradas as posições longas da estratégia, coincidentes com a compra das ações com elevado *score*, tal como em Balachandran et al. (2013).²¹ A amostra de treino é constituída por 15 janelas móveis constituídas cada uma por 10 anos, e a amostra de teste é composta por 15 janelas de 1 ano, o que restringiu a amostra de teste ao período 2002-2016. Foram usadas sempre as variáveis *dummy* que compõem os indicadores globais em vez dos valores reais das variáveis fundamentais, tendo por base o estudo de Patel et al. (2015), que no âmbito da análise técnica, conclui que a eficácia das previsões das técnicas de *machine learning* é superior quando se usa a informação através de sinais binários em detrimento dos valores reais das variáveis. O processo de otimização das estratégias assenta em dois procedimentos distintos. Por um lado, pretende-se otimizar o indicador global mediante a determinação de ponderadores das variáveis *dummy*, que o compõem, e que maximizam a rentabilidade média (ajustada ao mercado) ou o equivalente certo do portefólio de elevado *score*. Por outro, são geradas previsões da rentabilidade anual ou do equivalente certo, para cada ação, com base nas variáveis *dummy* que compõe o indicador global.²² A rentabilidade ajustada corresponde à rentabilidade anual da ação menos o retorno anual do respetivo índice de mercado, enquanto o equivalente certo (EC) foi calculado da mesma forma que em DeMiguel et al. (2007), através da seguinte expressão:

$$EC = E(R) - (a/2)Var(R), \quad (1)$$

onde $E(R)$ e $Var(R)$ correspondem respetivamente à esperança e variância do excesso da rentabilidade diária em relação à taxa de juro sem risco, enquanto que a é o parâmetro de aversão ao risco do investidor, cujo valores prováveis estão compreendidos entre 1 e 10. Por hipótese é assumido o valor de 1 para o parâmetro a , tal como em DeMiguel et al. (2007).

²¹ A otimização da estratégia *F_score* foi efetuada num contexto de elevado BM, tendo sido usada a amostra 75% das empresas de valor, enquanto na otimização da estratégia *N_score* foram consideradas todas as observações com dados disponíveis para a capitalização de mercado.

²² Com objetivo de selecionar empresas que superem os respetivos índices foram utilizados os retornos ajustados no processo de otimização. A otimização da estratégia *F_score* através da maximização ou previsão dos retornos brutos gerou piores resultados preditivos, apresentando retornos médios brutos e ajustados inferiores ao processo de otimização via retornos ajustados.

Nas Subsecções 3.2.1 e 3.2.2 são apresentados todos métodos testados com o objetivo de selecionar um portefólio de empresas que supere os portefólios de elevado *score* das estratégias base (F_score ou N_score).

3.2.1 Otimização Linear dos Ponderadores (OLP)

Este procedimento consistiu na determinação de ponderadores para as variáveis *dummy* que compõem o indicador global (F_score ou N_score), de forma a maximizar o valor da rentabilidade média (OLP_rt) ou do equivalente certo (OLP_ec) do portefólio de elevado *score*.²³ Esta otimização foi realizada através do método evolucionário da ferramenta Solver, que é um suplemento do Microsoft Excel. Foram aplicadas as seguintes restrições: i) o valor da soma dos ponderadores teria que ser igual ou inferior ao valor máximo do indicador agregado, (9 no caso do F_score, ou 11 no caso do N_score); ii) os ponderadores teriam que obrigatoriamente selecionar um número mínimo e máximo de empresas com um elevado *score*. Para determinar o valor utilizado para o número mínimo de empresas, começou por calcular-se 75% da percentagem média de empresas que apresentavam elevado *score* em cada um dos 10 anos da janela de treino. O número mínimo de empresas foi definido como a percentagem assim obtida, multiplicada pelo número total de empresas na janela respetiva. O limite máximo foi definido como o dobro do número de empresas imposto como limite mínimo. Estes limites foram definidos com o objetivo de regular a otimização dos ponderadores impedindo que fosse selecionado um número de empresas desajustado (muito baixo ou muito alto). Pois, se o número de empresas fosse muito baixo os ponderadores obtidos poderiam refletir relações espúrias de eventuais *outliers*, por outro lado um número demasiado grande restringe a capacidade do indicador global em identificar vencedores.

Desta forma, através da OLP, é obtido em cada janela de treino um conjunto de ponderadores para as variáveis *dummy*. O *score* otimizado corresponde ao produto escalar entre o conjunto das *dummy* de cada empresa e o conjunto dos ponderadores otimizados. As empresas que apresentam um *score* otimizado elevado (superior ou igual a 8 na otimização do F_score, e superior ou igual a 9 na otimização do N_score) formam o portefólio de elevado *score* do método OLP. Na amostra de teste, o portefólio OLP corresponde às empresas que apresentam um *score* otimizado elevado, sendo que o *score* otimizado é obtido pelo produto

²³ Considerou-se o equivalente certo do portefólio de elevado *score* como a média não ponderada dos equivalentes certos das empresas que constituem esse portefólio. O elevado *score* consiste nos valores 8 e 9, na otimização da estratégia F_score, enquanto na otimização da estratégia N_score o elevado *score* é definido pelos valores 9, 10 e 11.

escalar entre o conjunto das variáveis binárias para uma empresa e os ponderadores otimizados obtidos na respetiva janela de treino.

3.2.2 Regressão Linear (RL) e Técnicas de *Machine Learning*

Recorrendo a uma regressão linear e técnicas de *machine learning* procurou-se prever a rentabilidade (ajustada ao mercado) ou o equivalente certo com base nas variáveis *dummy* que formam os indicadores globais. Todos os portefólios de elevado *score* são formados por uma percentagem de empresas que apresentam as maiores previsões da rentabilidade ou do equivalente certo. A percentagem de empresas seleccionadas correspondeu a 75% da percentagem média de empresas que apresentavam elevado *score* em cada um dos 10 anos da janela respetiva, de acordo com a definição do F_score ou do N_score.

Na regressão linear a rentabilidade (RL_rt) ou o equivalente certo (RL_ec) são explicados pelos valores das variáveis fundamentais na sua forma binária. Os valores das previsões foram obtidos através da soma do termo independente com o produto entre os coeficientes dos regressores obtidos na janela de treino e os valores das respetivas variáveis *dummy* relativas a cada empresa.

As previsões relativas às técnicas de *machine learning* utilizadas foram obtidas com recurso ao software Weka (*Waikato Environment for Knowledge Analysis*). Foram utilizadas 4 técnicas de *machine learning*, nomeadamente, árvores de regressão, *random forest*, *support vector machine* e *bagging*, e mais 2 técnicas designadas de *ensemble 4* e *ensemble 2* que combinam as previsões obtidas pelas 4 técnicas.²⁴ As árvores de regressão permitem lidar com problemas de aprendizagem em que se pretende prever valores pertencentes a intervalos contínuos. Neste trabalho foi usada uma implementação de árvores de regressão baseada num algoritmo designado por M5p, que combina árvores de regressão tradicionais com regressões lineares que são efetuadas nas suas folhas (Wang e Witten, 1996). Quinlan (1992) mostra que estes modelos produzem melhores previsões que as árvores de regressão que se limitam a incorporar valores nas suas folhas. As *random forests* consistem na combinação das previsões provenientes de várias árvores. As árvores são construídas a partir de subamostras dos dados do conjunto de treino, obtidas de forma independente, e cada ramo da árvore é obtido com base apenas num subconjunto das variáveis consideradas. Nos problemas de regressão, as previsões das *random forests* são obtidas através da média das

²⁴ Note-se que, em *machine learning*, os modelos que têm por objetivo obter previsões de valores pertencentes a intervalos contínuos (como no caso presente) são designados por modelos de regressão.

previsões de cada árvore (Breiman 2001). O *support vector machine* (SVM) faz o mapeamento dos pontos num espaço de alta dimensão, recorrendo a uma função Kernel capaz de resolver problemas não-lineares. Deste modo, para problemas de regressão, o SVM procura definir uma função que se aproxime o melhor possível dos pontos de treino através da minimização do erro de previsão (Witten e Frank, 2005). *Bagging* é o acrónimo para *bootstrap aggregating*, consistindo num método que gera várias previsões a partir de vários subconjuntos da amostra de treino, obtendo o valor previsto através da média das previsões, caso se pretendam obter previsões de valores pertencentes a intervalos contínuos (Breiman 1996). Os diversos subconjuntos do conjunto de treino correspondem a diferentes amostras obtidas por um procedimento de *bootstrap*. Por último, as metodologias *ensemble* combinam as previsões das 4 técnicas de *machine learning* de forma a preverem quer a rentabilidade quer o equivalente certo para cada ação. O *ensemble 4* utiliza a média dos valores previstos pelas 4 técnicas, enquanto que *ensemble 2* usa o valor mediano, excluindo assim as duas previsões extremas para cada ação.

4. Dados e Análise Preliminar

O presente trabalho procura replicar e otimizar estratégias de investimento baseadas em indicadores fundamentais no mercado acionista europeu. Tendo em vista esse objetivo, foram selecionadas no total mais de 4 mil empresas cotadas nas principais bolsas de valores de cada um dos seguintes países: Reino Unido, Alemanha, França, Itália, Bélgica e Espanha.²⁵ Para todas as empresas foram recolhidas as seguintes variáveis com periodicidade anual: valor contabilístico por ação, capitalização de mercado, total dos ativos, *cash flows*, margem de lucro bruta, rácio de liquidez, resultado líquido antes de itens extraordinários/dividendos preferenciais, receitas líquidas de venda/emissão de ações ordinárias ou preferenciais, dívida de longo prazo, rotação do ativo, e ainda, com periodicidade diária, o preço não ajustado, e o índice de retorno total (ver o Anexo 1 que descreve as variáveis utilizadas). Foi igualmente recolhido o índice de retorno total para cada um dos seis índices de mercado MSCI, que posteriormente são utilizados para se obter a rentabilidade ajustada ao mercado. Todas estas variáveis, necessárias para a construção das estratégias de investimento, estão expressas na mesma unidade monetária, o euro, tendo sido recolhidas da base de dados Datastream/Eikon para o período 1990-2018 (ver Anexo 1). Por fim, retirou-se do *site* do Banco Central Alemão a taxa de juro *overnight* de empréstimos interbancários na praça financeira de Frankfurt e a EONIA, de modo a obter a *proxy* da taxa de juro sem risco diária entre 1990 e 2018.²⁶

A construção da base de dados utilizada no presente trabalho exigiu vários procedimentos. Face ao elevado volume de dados, foram separadas as variáveis de periodicidade diária das de periodicidade anual de modo a facilitar o processo de recolha e tratamento dos dados. Numa primeira fase, foram eliminadas todas as empresas sem dados para o índice de retorno total (entre 1990 e 2018), restringindo o número de empresas para 3174. De seguida, verificou-se para cada empresa, se existiam quebras nas séries históricas das cotações diárias do índice de retorno total, tendo-se verificado que, para as 3174 empresas, não existem

²⁵ Todas as empresas do presente trabalho estavam cotadas numa das seguintes bolsas de valores: Euronext Bruxelas, Euronext Paris, Londres, Mercado Contínuo Espanhol, Milão, Frankfurt (Xetra). A seleção das empresas foi restringida a um universo de ações de empresas europeias (inglesas, belgas, francesas, italianas, alemãs e espanholas) cotadas nas seis bolsas de valores referidas, tendo sido selecionadas as mais relevantes de acordo com o ranking da base de dados Datastream.

²⁶ A série da taxa de juro consistiu desde 18/10/1990 a 1/01/1999 nos valores da taxa de juro *overnight* da praça financeira de Frankfurt e a posteriormente nos valores da EONIA, até 18/10/2018. Entre 4/01/1999 e 31/05/2012 a EONIA e taxa de juro *overnight* da praça financeira de Frankfurt sobrepõem-se. Nesse período a diferença máxima é de 0,422% ao ano, ou seja, a diferença máxima diária foi 0,001172%, enquanto o valor médio da diferença é de 0,031951% ao ano, isto é 8,87527E-05% ao dia.

quebras na série temporal do índice de retorno total, e a última cotação disponível corresponde ao último dia da série recolhida. Por fim, foram calculados os retornos anuais, tendo sido eliminadas as empresas cuja série temporal do índice de retorno total se iniciava depois do último dia de cotação de abril do ano de 2017 (28/04/2017), o que reduziu o número de empresas para 3009.²⁷ Na segunda fase foram utilizadas as variáveis de periodicidade anual e calculou-se, para cada ano, os 9 indicadores que compõem o F_score de Piotroski (ver Apêndice).²⁸ Como o F_score de um dado ano t contém informação financeira do ano $t-2$ e prevê retornos anuais entre os anos $t+1$ e $t+2$, o período de análise foi restringido para 1992-2016 e contabilizou um total de 28236 observações empresas-ano. As observações de empresas que apresentaram um índice de retorno nulo no último dia de cotação de abril foram eliminadas (por ser impossível o cálculo do retorno), bem como as observações com retorno anual nulo em que o índice de retorno total permanece constante durante todo o ano.²⁹ Posteriormente, obteve-se para cada observação o rácio *book-to-market* (BM), através da divisão entre o valor contabilístico por ação e o preço não ajustado no momento de formação do portefólio (último dia de abril do ano seguinte, de forma a garantir que toda a informação necessária à aplicação da estratégia se encontra disponível), diminuindo a amostra para 28005 observações. Por último, foram eliminadas 71 observações que distorciam as estatísticas descritivas por apresentarem rácios BM muito elevados, fruto do reduzidíssimo valor de mercado, muito próximo de zero. Desta forma, o processo de construção da base de dados resultou numa amostra constituída por um total de 27934 observações empresas-ano ao longo de 25 anos (1992 a 2016).³⁰

Seguindo a metodologia de Piotroski (2000), a estratégia foi aplicada às empresas de valor. Para o devido efeito, em cada ano, as empresas foram ordenadas por rácio BM, selecionando-se as empresas pertencentes ao quintil com maior rácio. O número de observações assim selecionadas foi de 5548 ao longo de 25 anos (1992-2016), o que constitui aproximadamente 2/5 da dimensão utilizada em Piotroski (2000). Tendo por base artigos de investigação mais

²⁷ As empresas cuja série do índice de retorno total se iniciava depois do último dia de cotação de abril de 2017 (28/04/2017) não permitiam o cálculo de pelo menos um retorno anual.

²⁸ Foram eliminadas da base de dados, em cada ano, as observações de empresas que não disponham de informação financeira necessária para o cálculo do indicador global F_score. Desta forma não se incorreu no erro *survival bias*, uma vez que as empresas são selecionadas anualmente, mediante a existência de informação necessária para a aplicação da estratégia F_score (em cada um dos 25 anos).

²⁹ Foram eliminadas 36 observações que apresentavam o índice de retorno total nulo aquando o cálculo do retorno anual, obtido pela diferença de logaritmos, e ainda 73 observações com retornos anuais nulos em que o valor do índice de retorno total permanecia constante ao longo desse ano.

³⁰ Das 27934 observações utilizadas na replicação da estratégia F_score, apenas 27641 apresentam dados para a capitalização de mercado.

recentes, como Amor-Tapia e Tascón (2016), e Navas et al. (2018), foi igualmente aplicada a estratégia para a amostra total, com o intuito de verificar a eficácia da estratégia sem o condicionamento prévio ao BM, e perceber, de forma comparativa, se o contexto BM compromete o sucesso da estratégia baseada no F_score.

No Anexo 2 encontra-se representada, para os dois conjuntos de dados (empresas de valor e todas as empresas), a distribuição das empresas-ano pelas Bolsas de Valores -Painel A-, assim como o número de observações de empresas ao longo dos 25 anos -Painel B-. No Painel A verifica-se que a bolsa de Londres é a mais representada com 43% do total de empresas-ano (49% das observações de empresas de valor), seguindo-se a Euronext Paris, e as bolsas de Frankfurt (Xetra) e de Milão, que contabilizam 25% (21%), 15% (10%) e 9% (13%), respetivamente, enquanto o Mercado Contínuo Espanhol e Euronext Bruxelas contam, na respetiva ordem, com apenas 5% (3%) e 4% (3%) do número de observações. No Painel B constata-se que o número de empresas aumenta consideravelmente ao longo do período de análise, desde as 352 empresas (70 empresas de valor) em 1992, para mais do quántuplo em 2016, com 2070 (414) empresas.

Na Tabela 1 encontram-se estatísticas descritivas de variáveis financeiras para o conjunto das empresas de elevado rácio BM (empresas de valor), no Painel A, e para todas as empresas da amostra, no Painel B. Como indica a Tabela 1, as empresas de valor apresentam em média (mediana) um rácio BM igual a 20 (1,89) e uma capitalização de mercado de 1,13 mil milhões (42 milhões), enquanto a média (mediana) do rácio BM e da capitalização de mercado é respetivamente, para todas as empresas, 3,29 (0,47), e 2,88 mil milhões (158 milhões).³¹ Em conformidade com a literatura, a Tabela 1 evidencia o fraco desempenho das empresas de valor, espelhado principalmente nos reduzidos valores da rentabilidade dos ativos (ROA).³² Estas empresas apresentam em média (mediana) um ROA de -4.4% (0,54%), e apenas 54% das observações contemplam ROA positivos, sendo notório o aumento do valor médio e mediano do ROA (para 0,55% e 3,75%, respetivamente), assim como da percentagem de observações com rentabilidades do ativo positivas (75% face aos 54%) quando se considera o conjunto de todas as empresas. Além da rentabilidade dos ativos reduzida, as empresas de valor apresentam em média (mediana) variações negativas do ROA e da margem bruta, respetivamente -2,34% (-0,31%) e -27,87% (-0,21%), o que também

³¹ As estatísticas descritivas afetas à capitalização de mercado estão restringidas a 27641 observações (para todas as empresas), e a 5514 observações (para o conjunto de empresas de valor).

³² Fama e French (1995) evidenciam que as empresas de valor tendem a enfrentar situações financeiras difíceis associadas aos seus pobres resultados.

evidencia o seu fraco desempenho. Também na Tabela 1, podemos constatar que os dois conjuntos de observações (empresas de valor e todas as empresas) apresentam, em média, *cash flows* positivos, *accruals* negativos, e exibem, em média, face ao ano transato, diminuições da rentabilidade, da alavancagem, da margem de lucro bruta e da rotação do ativo, e em contrapartida verificaram um aumento na liquidez.

Na Tabela 2 constam os retornos brutos e ajustados ao mercado para o conjunto de empresas de valor e para todas as empresas da amostra, ao longo do período de análise (1992-2016). Contrariamente ao usualmente apresentado na literatura, a Tabela 2 mostra que as empresas de valor apresentam um retorno médio ajustado ao mercado negativo (-8,61%) e inferior à média da amostra com todas as empresas (-3,76%).³³ Os resultados da Tabela 2 parecem estar em conformidade com os da Tabela 1, na qual se verifica que as percentagens de observações com sinais financeiros positivos são inferiores para a amostra formada pelo conjunto de empresas com maior rácio BM. Isto é, há relativamente mais empresas que apresentam uma situação financeira mais desfavorável (empresas com baixo *F_score*) na amostra constituída pelas empresas de valor, o que eventualmente pode explicar o retorno médio inferior nesta amostra. Ainda na Tabela 2 constata-se que a maioria das observações exibem retornos ajustados ao mercado negativos, quer para as empresas de valor, quer para todas as empresas (aproximadamente 55% e 51% respetivamente), o que revela a potencialidade da estratégia de investimento em aumentar o retorno médio através da eliminação de parte dessas observações. Por último, há que realçar a grande diferença entre a média dos retornos brutos e ajustados anuais das empresas de valor do presente trabalho (valores negativos) comparativamente com os valores obtidos no estudo de Piotroski (2000).³⁴ Este contraste sustenta um dos objetivos deste trabalho, que consiste em perceber se a estratégia de investimento desenvolvida em Piotroski (2000) continua a ser eficaz num contexto temporal, espacial e economicamente diferente, abrangendo a profunda crise financeira de 2007-2009.

³³ A literatura que argumenta que as empresas de valor apresentam, em média, retornos avultados quer pela explicação baseada no risco (Fama e French, 1992), quer pela explicação baseada no *market mispricing* (Lakonishok et al.,1994).

³⁴ Piotroski (2000) obteve, para as empresas de valor do mercado americano entre 1976 e 1996, um retorno médio bruto e ajustado anual de 23,9% e 5,9%, respetivamente.

Tabela 1- Estatísticas descritivas para as empresas de valor e para o conjunto de todas as observações entre 1992 e 2016

Painel A - Empresas de valor -5548 observações empresas-ano entre 1992 e 2016						
Variável	Média	Mediana	Desvio-Padrão	Assimetria	Curtose	Sinais Positivos (%)
BM	20,7791	1,8852	110,8177	12,5150	205,2898	-
CM	1138373	42312	10209389	18,7131	459,7149	-
Ativos	3310411	109906	19934951	10,7993	146,4675	-
ROA	-0,0441	0,0054	0,4391	-30,0684	1820,3107	55,88%
Δ ROA	-0,0234	-0,0031	2,4426	-33,1566	1822,2402	45,89%
CF	0,0179	0,0387	0,5240	1,9685	1985,0406	72,66%
<i>Accruals</i>	-0,0620	-0,0419	0,2502	-38,5848	2126,2077	87,56%
Δ Liquid	0,0204	-0,0200	3,8479	9,1552	428,5060	46,40%
Δ Lever	-0,0018	0,0000	0,1147	-29,7864	1610,9811	61,99%
Δ Margin	-0,2787	-0,0021	21,8073	-28,0532	1327,6590	46,85%
Δ Turn	-0,0158	0,0000	0,4347	-1,1136	505,2283	47,39%
Painel B – Todas as empresas -27934 observações empresas-ano entre 1992 e 2016						
Variável	Média	Mediana	Desvio-Padrão	Assimetria	Curtose	Sinais Positivos (%)
BM	3,2938	0,4685	68,4768	-15,5946	2101,4908	-
CM	2886454	158240	11737376	9,1805	125,9367	-
Ativos	4430256	210106	18509439	8,5753	99,4517	-
ROA	0,0055	0,0375	0,8191	23,9544	6082,3683	75,24%
Δ ROA	-0,2339	0,0002	42,2496	-163,0603	27092,0852	50,29%
CF	0,0308	0,0811	5,3976	-161,4722	26684,6085	84,11%
<i>Accruals</i>	-0,0253	-0,0467	5,3441	165,7817	27629,6594	90,27%
Δ Liquid	0,0639	0,0000	10,2221	142,9896	22675,3389	48,59%
Δ Lever	-0,1139	0,0000	19,9775	-165,7509	27633,9551	62,30%
Δ Margin	-1,2388	0,0012	390,9371	-29,3440	9977,8493	51,96%
Δ Turn	-0,0123	0,0000	0,4521	-28,3166	2617,6379	48,16%

A primeira coluna refere-se aos acrónimos das variáveis, cuja definição é a seguinte: BM é o rácio *book-to-market*, definido pelo quociente entre o valor contabilístico por ação no final do ano fiscal e o preço não ajustado no momento da formação dos portefólios (último dia de abril). CM corresponde à capitalização de mercado (em milhares de euros), que consiste no número de ações ordinárias em circulação multiplicado pelo preço da ação no final do ano fiscal. Apenas 27641 observações apresentam dados para a capitalização de mercado, das quais 5514 correspondem a empresas de valor. Ativos representa o total de ativos no final do ano fiscal (em milhares de euros). ROA é a rentabilidade do ativo. Este indicador resulta da divisão entre o resultado líquido antes de itens extraordinários/dividendos preferenciais no final do ano fiscal e o total dos ativos no início do ano. Δ ROA é igual ao ROA do ano t menos ROA do ano $t-1$. CF é definido pelo quociente entre os *cash flows* no final do ano e o total dos ativos no início do ano. *Accruals* corresponde à diferença entre ROA e CF. Δ Liquid corresponde à diferença entre o rácio de liquidez do ano t e o rácio de liquidez do ano $t-1$. O rácio de liquidez consiste no quociente entre o total do ativo corrente e o total do passivo corrente. Δ Lever é definido

pela variação anual do rácio da dívida de longo prazo sobre o ativo do ano. $\Delta Margin$ corresponde à variação anual da margem de lucro bruta. A margem de lucro bruta é definida pelo quociente entre o lucro bruto e as vendas líquidas ou receitas. $\Delta Turn$ consiste na variação anual do rácio de rotação do ativo. O rácio de rotação do ativo é definido como as vendas líquidas ou receitas sobre o total do ativo. Os sinais positivos correspondem à percentagem de cada variável *dummy* assumir o valor de 1. As variáveis fundamentais, que compõem o *F_score*, foram transformadas em variáveis *dummy*, assumindo o valor de 1 no caso de se esperar que o valor dessa variável tenha um efeito positivo no desempenho futuro da empresa, e de 0 caso contrário.

Tabela 2- Retornos anuais entre 1992-2016

Amostra	Retornos anuais	Média	Mediana	Retornos Positivos (%)	Desvio-Padrão	Assimetria	Curtose
Empresas de valor	Retorno Bruto	-0,0284	0,0116	51,37%	0,5815	-0,6165	5,5647
	Retorno Ajustado	-0,0861	-0,0543	44,54%	0,5279	-0,6253	6,2330
Todas as empresas	Retorno Bruto	0,0187	0,0594	56,53%	0,4908	-0,6914	7,6115
	Retorno Ajustado	-0,0376	-0,0080	48,95%	0,4408	-0,7072	9,2309

Na presente tabela, o retorno bruto corresponde ao lucro resultante da compra de uma ação no final de abril e sua venda no final de abril no ano seguinte, enquanto o retorno ajustado corresponde ao retorno bruto descontado da rentabilidade do mercado (medida através do índice MSCI correspondente). Na coluna designada por “Retornos positivos (%)” são apresentadas as percentagens de retornos (brutos ou ajustados) positivos durante o período de análise (1992 a 2016). O cálculo das estatísticas apresentadas foi efetuado considerando, de forma agregada, todas as observações respeitantes aos 25 anos do período de análise, ou seja, foi dado o mesmo peso a cada observação. A amostra “empresas de valor” contabiliza 5548 observações, e na amostra “todas as empresas” são consideradas 27934 observações.

5. Resultados

A presente secção apresenta os resultados do estudo aplicado. Na Subsecção 5.1 são apresentados os resultados da aplicação das estratégias de investimento baseadas nos indicadores globais F_score e N_score , enquanto a Subsecção 5.2 expõe os resultados relativos à otimização das estratégias de investimento.

5.1 Estratégias de Investimento baseadas em Indicadores Fundamentais

5.1.1 Estratégia F_score para empresas de valor

Na Tabela 3 estão representados os resultados agregados da estratégia F_score aplicada às empresas de valor (conjunto de observações de empresas pertencentes ao quintil com maior rácio BM em cada ano) entre 1992 e 2016. O principal destaque da Tabela 3 é a relação quase monótona entre o indicador global e os retornos anuais (brutos e ajustados ao mercado), espelhando a capacidade do F_score em distinguir vencedores de perdedores. A compra das ações com um F_score elevado (8 ou 9) e venda a descoberto das ações com baixo F_score (igual ou inferior a 2) proporciona um retorno médio ajustado ao mercado de 21,38%, superior a 20% tal como em Piotroski (2000), demonstrando a eficácia da estratégia F_score aplicada a empresas de valor num contexto económico bastante diferente, que inclui a forte crise financeira de 2007-2009. Ainda que o portefólio das empresas de elevado F_score supere não só o portefólio das empresas de baixo F_score em mais de 21%, como o portefólio de todas as empresas de valor em mais de 8% (semelhante aos resultados de Piotroski, 2000), permite apenas um retorno médio anual bruto (ajustado ao mercado) de 5,23% (-0,04%), o que significa que os elevados retornos da estratégia F_score , acima dos 21%, dependem em grande parte da capacidade em identificar os perdedores, e consequentemente da possibilidade de efetuar vendas a descoberto.

A análise da Tabela 3 sugere ainda que a estratégia F_score é eficaz quando aplicada para o período de análise (1992-2016), contudo resta saber qual é a distribuição anual dos resultados da estratégia. Face ao reduzido número de observações de empresas para cada ano, a estratégia consistiu, tal como em Piotroski (2000), na compra do conjunto de ações de empresas que apresentam a maioria das variáveis *dummy* igual a 1 (F_score igual ou superior a 5) e na venda a descoberto das ações de empresas com maioria de variáveis *dummy* igual a 0 (F_score inferior a 5).

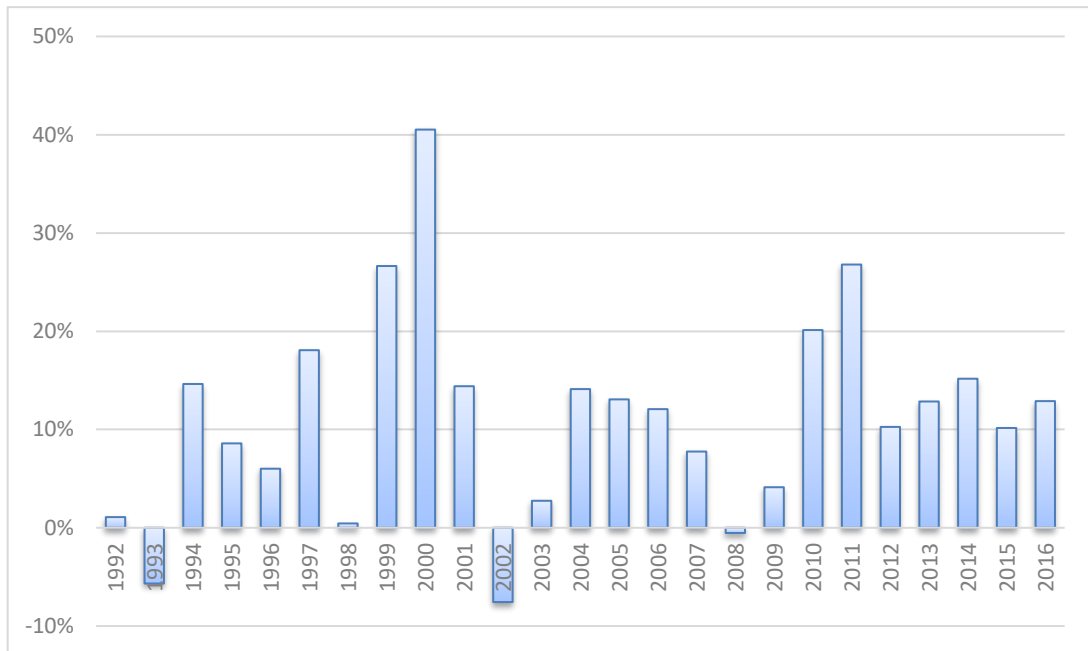
Tabela 3- Estratégia F_score aplicada a empresas de valor (1992-2016)

	N	Retornos Brutos		Retornos Ajustados	
		Média	Mediana	Média	Mediana
Empresas de valor	5548	-0,0284	0,0116	-0,0861	-0,0543
F_score					
0	3	-0,7528	-0,6714	-0,6658	-0,7286
1	51	-0,3827	-0,2279	-0,4168	-0,2966
2	254	-0,1074	-0,0907	-0,1683	-0,1960
3	587	-0,1385	-0,0732	-0,1878	-0,1256
4	1009	-0,0598	-0,0199	-0,1122	-0,0841
5	1210	-0,0026	0,0243	-0,0645	-0,0311
6	1084	-0,0091	0,0417	-0,0692	-0,0391
7	843	0,0234	0,0522	-0,0413	-0,0286
8	398	0,0551	0,0771	0,0077	0,0079
9	109	0,0422	0,0688	-0,0299	-0,0204
Baixo Score (0,1,2)	308	-0,1593	-0,1230	-0,2143	-0,2233
Elevado Score (8,9)	507	0,0523	0,0724	-0,0004	0,0073
Elevado Score - Todas		0,0807	0,0608	0,0857	0,0616
Estratégia F_score		0,2116	0,1955	0,2138	0,2306

Na presente tabela estão representados os retornos anuais brutos e ajustados ao mercado, ao longo do período 1992-2016, de diversos portfólios formados com base no valor do F_score, que pode assumir um valor inteiro entre 0 e 9. A estratégia F_score consiste na compra de ações com elevado *score* e na venda a descoberto das que apresentam um baixo *score*. O portfólio de baixo *score* é constituído por todas as empresas que apresentaram um F_score igual ou inferior a 2, enquanto que o portfólio de elevado *score* é composto pelas empresas que obtiveram um F_score igual a 8 ou 9. “Elevado *Score* – Todas” representa a diferença entre a média dos retornos obtidos pela compra das ações com elevado *score* e a rentabilidade média da amostra considerada. O retorno bruto corresponde ao lucro resultante da compra de uma ação no final de abril e sua venda no final de abril no ano seguinte enquanto o retorno ajustado é obtido através da diferença entre o retorno (bruto) de uma empresa e o retorno do respetivo índice do mercado.

Na Figura 1 e no Anexo 3 encontram-se representados os retornos ajustados da estratégia adotada ao longo dos anos. Observa-se que 22 dos 25 anos apresentam retornos ajustados ao mercado positivos, e que a estratégia garante um retorno médio ajustado de 11,23%, mais 19,84% que a rentabilidade média ajustada ao mercado de todas as empresas de valor (-8,61%). Os resultados evidenciam um bom desempenho ao longo dos anos, sugerindo que a estratégia baseada no F_score é robusta ao longo do tempo.

Figura 1- Retornos ajustados da estratégia F_score aplicada a empresas de valor para cada ano



Fonte: Cálculos do autor

Não obstante a estratégia F_score aplicada às empresas de valor seja eficaz em diferenciar vencedores de perdedores e pareça ser robusta ao longo do tempo, depende em grande parte do lado curto da mesma, o que pode comprometer os seus elevados retornos (acima de 20%) devido à existência de eventuais restrições às vendas a descoberto. Como os retornos obtidos através das vendas a descoberto podem não ser exequíveis, fruto das restrições do mercado, o retorno proporcionado pela compra das ações de empresas com elevado F_score ganha uma maior relevância. Posto isto, e tendo por base as evidências obtidas por Patel et al. (2006) e Dreman (2008) foram eliminadas as empresas pertencentes ao decil e quartil de maior rácio BM em cada ano, assumindo a hipótese de que os maiores valores de rácios BM poderão estar associados a empresas que se encontram numa situação problemática, não captada pelo indicador global, todavia refletida no valor de mercado muito baixo (elevado rácio BM).³⁵ Deste modo, pretende-se eliminar as observações de empresas que, embora apresentem um F_score elevado, enfrentem uma situação difícil (traduzida pelo baixo valor de mercado), de forma a potenciar os retornos anuais obtidos pelo portefólio das empresas

³⁵ Patel et al. (2006) e Dreman (2008) indicam que os portefólios constituídos pelas empresas pertencentes, respetivamente, ao 2º e 3º decil e 2º quintil com maior *dividend yield* são os que proporcionam os maiores retornos, pois algumas empresas que apresentam os maiores valores de *dividend yield* podem ter sofrido uma queda no preço da sua ação. De forma análoga coloca-se a mesma hipótese relativamente ao rácio BM.

de valor fortes financeiramente (com F_score igual a 8 ou 9). A eliminação do decil e quartil de empresas com maior rácio BM originou a formação de duas subamostras de empresas de valor designadas respetivamente por 90% das empresas de valor e 75% das empresas de valor, as quais foi aplicada a estratégia F_score (Anexo 4 e 5, respetivamente).

De modo a facilitar a análise, a Tabela 4 sintetiza os resultados da aplicação da estratégia F_score nas diferentes amostras de valor para o período 1992-2016. Essa tabela mostra que a eliminação do decil e do quartil de empresas com maior rácio BM contribui para o aumento dos retornos médios brutos (ajustados) da amostra das empresas de valor em 2,32% (2,34%) e 5% (5,05%), respetivamente, evidenciando o mau desempenho geral das empresas que apresentam os maiores valores de rácio BM. No entanto, os retornos da estratégia F_score diminuem significativamente com a eliminação destas empresas, sendo reduzidos em quase metade se tivermos em conta a amostra 75% de empresas de valor (sem o quartil de empresas de valor com maior rácio BM). Ainda assim, os resultados mais interessantes e satisfatórios da Tabela 4 prendem-se com o facto dessa diminuição surgir toda do lado curto da estratégia, uma vez que o retorno médio bruto (ajustado) obtido pela compra das ações de empresas fortes financeiramente aumenta de 5,23% (-0,04%) para 6,41% (1,13%) e 7,89% (2,53%) quando são eliminados o decil e quartil de empresas com maior rácio BM, respetivamente. Em suma, os resultados estão em conformidade com a hipótese assumida de que é possível potenciar os retornos obtidos do portefólio constituído de empresas de valor fortes financeiramente através da eliminação das empresas com maior rácio BM, o que apesar de comprometer o retorno total da estratégia pode ser interessante para um investidor num cenário de fortes restrições às vendas a descoberto.³⁶

³⁶ Os resultados da aplicação da Estratégia F_score ao longo dos 25 anos nas amostras resultantes da eliminação das observações pertencentes ao decil e quartil de empresas com maiores rácios BM contidos nos Anexos 6 e 7, respetivamente, sugerem que a estratégia continua a ser robusta ao longo do tempo. A amostra constituída por 90% das empresas de valor apresenta retornos médios ajustados positivos em 21 dos 25 anos e um retorno médio ajustado de 10,04% superior em 16,31% à média do retorno da amostra (Anexo 6), enquanto que a amostra com 75% das empresas de valor apresenta 20 anos com retornos ajustados positivos e um retorno médio ajustado da estratégia de 7,32%, superior em 10,88% ao retorno médio ajustado da amostra (Anexo7).

Tabela 4- Estratégia F_score em diferentes amostras de valor (1992-2016)

				Retornos Brutos		Retornos Ajustados	
Portefólio		N	Média	Mediana	Média	Mediana	
Empresas de valor	Todas as empresas	5548	-0,0284	0,0116	-0,0861	-0,0543	
	Elevado <i>Score</i>	507	0,0523	0,0724	-0,0004	0,0073	
	Baixo <i>Score</i>	308	-0,1593	-0,1230	-0,2143	-0,2233	
	Elevado <i>Score</i> - Todas	-	0,0807	0,0608	0,0857	0,0616	
	Estratégia F_score	-	0,2116	0,1955	0,2138	0,2306	
90% Empresas de valor	Todas as empresas	5006	-0,0052	0,0267	-0,0627	-0,0358	
	Elevado <i>Score</i>	480	0,0641	0,0772	0,0113	0,0079	
	Baixo <i>Score</i>	263	-0,0969	-0,0922	-0,1566	-0,1953	
	Elevado <i>Score</i> - Todas	-	0,0693	0,0504	0,0740	0,0437	
	Estratégia F_score	-	0,1610	0,1694	0,1679	0,2032	
75% Empresas de valor	Todas as empresas	4171	0,0217	0,0451	-0,0356	-0,0211	
	Elevado <i>Score</i>	432	0,0789	0,0781	0,0253	0,0118	
	Baixo <i>Score</i>	198	-0,0349	-0,0674	-0,0972	-0,1431	
	Elevado <i>Score</i> - Todas	-	0,0573	0,0331	0,0609	0,0329	
	Estratégia F_score	-	0,1139	0,1456	0,1226	0,1549	

Todas as definições encontram-se descritas na Tabela 3.

5.1.2 Estratégia F-score para todas as empresas

Com o propósito de analisar o sucesso da estratégia F_score fora do contexto de elevado rácio BM foi aplicada a estratégia à amostra que contempla todas as empresas, constituída por 27934 observações ao longo do período 1992-2016.³⁷ Na Tabela 5 constata-se o sucesso da estratégia F_score aplicada à amostra de todas as empresas, que permite em média, um retorno anual bruto (ajustado) de 21,05% (20,87%). Ainda assim, 12,29% (18,58%) provêm

³⁷ Vários artigos como Piotroski e So (2012), Dosamantes (2013), Amor-Tapia e Tascón (2016) e Navas et al. (2018) mostram a eficácia da estratégia F_score de Piotroski (2000) fora do contexto de elevado rácio BM.

da venda a descoberto e apenas 8,76% (2,29%) da compra do portefólio de elevado F_score. Surpreendentemente, a estratégia F_score apresenta resultados mais interessantes para os investidores quando aplicada à amostra constituída por todas as empresas, uma vez que os retornos médios brutos e ajustados são semelhantes aos obtidos com a amostra das empresas de valor (superiores a 20%), no entanto os investidores conseguem, em média, retornos brutos (ajustados) superiores em 3,53% (2,33%) com a compra de ações com elevado F_score. Isto é, os retornos obtidos pela estratégia F_score dependem menos da venda a descoberto (relativamente aos retornos obtidos com as empresas de valor), o que põe em evidência a sua maior exequibilidade, face a existência de restrições às vendas a descoberto.

Os resultados são consistentes com outros artigos da literatura que aplicaram a estratégia F_score de Piotroski (2000) nos mercados acionistas europeus. Por exemplo, Amor-Tapia e Tascón (2016) entre 1989 e 2011, e Navas et al. (2018) entre 2000 e 2016 mostram que a estratégia F_score é capaz de proporcionar no mercado acionista europeu um retorno médio ajustado ao mercado de 28% e 21,74%, respetivamente. Ainda que os retornos brutos dos portefólios de elevado F_score naqueles dois artigos sejam significativamente superiores (29% e 21,16%, respetivamente) aos do presente trabalho (8,76%), quando se tem em consideração a diferença entre o retorno desses portefólios e retorno médio de toda a amostra, verifica-se que os valores dos dois artigos (11% e 8,12%) convergem para os valores aqui apresentados (6,89%), sugerindo que o fraco retorno por nós obtido na compra das ações de empresas com elevado F_score se deve à reduzida rentabilidade geral da amostra selecionada.

Por último, no Anexo 9 é demonstrado o bom desempenho da estratégia ao longo dos anos, em que 22 dos 25 anos apresentam um retorno médio ajustado positivo.³⁸ O retorno médio da estratégia aplicada em cada ano é de 10,09%, superior em 13,85% à média da amostra, sugerindo que a estratégia é robusta ao longo do tempo quando são consideradas todas as empresas.

³⁸ A estratégia considerada na análise do desempenho ao longo dos 25 anos consistiu na compra dos portefólios com F_score igual ao superior a 5 e na venda a descoberto dos portefólios que apresentam um F_score inferior a 5, devido ao reduzido número de observações com F_score elevado (8 e 9) e F_score baixo (0, 1 e 2) em cada ano.

Tabela 5 - Estratégia F_score aplicada a todas as empresas (1992-2016)

	N	Retornos Brutos		Retornos Ajustados	
		Média	Mediana	Média	Mediana
Total de empresas	27934	0,0187	0,0594	-0,0376	-0,0080
F_score					
0	13	-0,1549	-0,0960	-0,1883	-0,1500
1	159	-0,1831	-0,0831	-0,2395	-0,1607
2	750	-0,1096	-0,0905	-0,1743	-0,1651
3	2004	-0,1115	-0,0607	-0,1591	-0,1125
4	4047	-0,0227	0,0298	-0,0769	-0,0379
5	5984	0,0201	0,0622	-0,0328	-0,0026
6	6490	0,0417	0,0800	-0,0175	0,0084
7	5262	0,0526	0,0805	-0,0026	0,0135
8	2602	0,0860	0,1015	0,0205	0,0198
9	623	0,0940	0,0864	0,0326	0,0183
Baixo Score (0,1,2)	922	-0,1229	-0,0866	-0,1858	-0,1627
Elevado Score (8,9)	3225	0,0876	0,0994	0,0229	0,0192
Elevado Score - Todas		0,0689	0,0400	0,0605	0,0272
Estratégia F_score		0,2105	0,1860	0,2087	0,1819

A presente tabela expõe os retornos brutos e ajustados ao mercado de diversos portfólios constituídos por empresas com determinados valores de F_score, ao longo do período 1992-2016. A tabela tem a mesma estrutura que a Tabela 3, onde estão descritas todas as definições.

Resumidamente, os elevados retornos brutos e ajustados juntamente com a menor dependência das vendas a descoberto e o bom desempenho ao longo do tempo, demonstram que a capacidade da estratégia F_score em proporcionar elevados retornos não é restrita a um universo de empresas de valor. Adicionalmente, os resultados da presente subsecção, e da Subsecção 5.1.1 revelam que o sucesso da estratégia F_score de Piotroski (2000) (aplicada a empresas de valor ou a todas as empresas) parece não ser circunstancial ao tempo (diferente período amostral) e ao espaço (outros mercados que não o Norte-Americano).

5.1.3 Estratégia N_score

O indicador N_score incorpora as 9 variáveis do F_score mais o rácio BM e a capitalização de mercado com o objetivo de potenciar a capacidade preditiva do F_score (ver Secção 3). A estratégia N_score definida pela compra de ações com score de 9, 10 e 11 e venda a

descoberto de ações com *score* igual ou inferior a 3, foi aplicada entre 1992 e 2016 à amostra constituída por todas as empresas que disponham de dados relativos à capitalização de mercado.³⁹

Os resultados da estratégia N_score, expostos na Tabela 6, indicam que a estratégia é eficaz em separar vencedores de perdedores, permitindo anualmente um retorno médio bruto e ajustado ao mercado superior a 15%. Essa eficácia está patente na elevada correlação positiva entre os retornos e o N-score, verificando-se que, à exceção dos dois primeiros *scores* (0 e 1), os retornos médios diminuem sempre com a diminuição do N_score. Sobressai, no entanto, o escasso número de observações de empresas com estes dois *scores* (0 e 1), apenas 24 das 27641 observações totais, que juntamente com a presença de eventuais *outliers* poderá explicar os retornos médios mais elevados para estes valores do N_score.

Todavia, os retornos obtidos dependem sobretudo das vendas a descoberto, tendo a compra do portefólio de elevado N_score gerado um retorno médio bruto (ajustado) de apenas 6,75% (0,46%). Além disso, a estratégia N_score parece ser menos eficaz que a estratégia F_score (ver Anexo 10), uma vez que apresenta retornos médios brutos (ajustados) inferiores quer com a compra (de ações com elevado *score*) quer com a venda a descoberto (de ações com baixo *score*), em aproximadamente, 2,15% (1,93%) e 2,47% (3,24%), respetivamente.

No que se refere ao desempenho ao longo dos anos, apresentado no Anexo 11, verifica-se que em 23 dos 25 anos a estratégia N_score apresenta retornos ajustados positivos (empresas fortes com N_score igual ou superior a 6 e empresas fracas com N_score igual ou inferior a 5), proporcionando um retorno médio anual ajustado ao mercado de 6,32%, aproximadamente 10% superior ao retorno médio ajustado de toda a amostra (-3,68%), o que evidencia a eficácia da estratégia ao longo do tempo. Em suma, os resultados sugerem que a estratégia baseada no N_score é eficaz, contudo a incorporação do BM e da capitalização de mercado no indicador global F_score parece não melhorar a capacidade preditiva do mesmo, visto que a estratégia N-score apresenta retornos inferiores aos obtidos pela estratégia F_score.

³⁹ Isto foi feito por forma a que as percentagens de empresas de elevado e baixo *score* fossem semelhantes às obtidas em Piotroski (2000). A amostra de todas as empresas foi restringida a 27641 observações devido a algumas das empresas não disporem de dados relativos à capitalização de mercado.

Tabela 6- Estratégia N_score (1992-2016)

	N	Retornos Brutos		Retornos Ajustados	
		Média	Mediana	Média	Mediana
Total de empresas	27641	0,0195	0,0601	-0,0368	-0,0075
N_score					
0	1	0,0036	0,0036	-0,0813	-0,0813
1	23	0,0416	-0,0269	-0,0714	-0,1482
2	127	-0,1287	-0,0585	-0,1831	-0,1544
3	753	-0,0903	-0,0193	-0,1427	-0,0867
4	2203	-0,0549	-0,0059	-0,1085	-0,0641
5	4332	-0,0128	0,0450	-0,0631	-0,0147
6	5982	0,0205	0,0659	-0,0342	-0,0019
7	5881	0,0355	0,0717	-0,0227	0,0029
8	4577	0,0469	0,0795	-0,0112	0,0114
9	2584	0,0606	0,0763	-0,0049	-0,0076
10	946	0,0789	0,0775	0,0216	0,0149
11	232	0,0983	0,0735	0,0412	0,0176
Baixo Score (0,1,2,3)	904	-0,0922	-0,0242	-0,1465	-0,0948
Elevado Score (9,10,11)	3762	0,0675	0,0756	0,0046	0,0004
Elevado Score – Todas		0,0480	0,0155	0,0414	0,0078
Estratégia N_score		0,1597	0,0998	0,1511	0,0951

Na presente tabela estão representados os retornos anuais brutos e ajustados ao mercado, ao longo do período 1992-2016, de diversos portfólios formados com base no valor do indicador global N_score, que pode assumir um valor inteiro entre 0 e 11. A estratégia N_score consiste na compra de ações com elevado *score* e na venda a descoberto das que apresentam um baixo *score*. O portfólio de baixo *score* é constituído por todas as empresas que apresentaram um N_score igual ou inferior a 3, enquanto que o portfólio de elevado *score* é composto pelas empresas que obtiveram um N_score igual ou superior a 9. “Elevado *Score* – Todas” representa a diferença entre a média dos retornos obtidos pela compra das ações com elevado *score* e a média dos retornos de todas as empresas consideradas. O retorno bruto corresponde ao lucro resultante da compra de uma ação no final de abril e sua venda no final de abril no ano seguinte enquanto o retorno ajustado é obtido através da diferença entre o retorno (bruto) de uma empresa e o retorno do respetivo índice do mercado.

5.2 Otimização das Estratégias de Investimento

Nesta secção são apresentados os resultados da otimização linear e não-linear da estratégia F_score (5.2.1) e da estratégia N_score (5.2.2), considerando somente as posições longas das estratégias, ou seja, a compra das ações com elevado *score*. O processo de otimização consiste quer na ponderação das variáveis *dummy* de modo a maximizar a rentabilidade anual (ajustada ao mercado) ou o equivalente certo, quer na seleção das empresas que apresentam os maiores valores previstos da rentabilidade anual (ajustada ao mercado) ou do equivalente

certo em função das variáveis *dummy* que compõem os indicadores globais. Para esse efeito são utilizados oito métodos, designadamente: otimização linear dos ponderadores (OLP), regressão linear (RL) e 6 métodos que recorrem a técnicas de *machine learning* (árvores de regressão, *random forests*, *support vector machine*, *bagging*, *ensemble 4*, *ensemble 2*), que se encontram descritos na Secção 3. As amostras de treino consistem em janelas móveis de 10 anos enquanto as amostras de teste correspondem ao ano seguinte da respetiva janela móvel, o que restringiu o período de teste desde o ano 2002 até 2016.

5.2.1 Otimização da Estratégia F_score

O processo de otimização da estratégia F_score foi realizado para a amostra 75% das empresas de valor, uma vez que é a amostra constituída por empresas de valor que apresenta um retorno médio superior na compra de ações com elevado F_score (Secção 5.1.1). Os resultados da otimização linear da compra das ações com elevado F_score, através dos métodos OLP e RL, para os anos de teste constam na Tabela 7, onde é perceptível o efeito da grave crise financeira de 2007-2009, refletida nos elevados retornos anuais negativos entre abril de 2007 e abril de 2009, correspondentes à aplicação da estratégia F_score nos anos 2006 e 2007, respetivamente. Estes elevados retornos negativos afetam os retornos médios obtidos nos 15 anos do período de teste, para o qual a estratégia base F_score possibilita um retorno médio anual de 9,8%, potenciando em mais de 5% a rentabilidade média da amostra de empresas de valor considerada (4,56%).

Ainda na Tabela 7 verifica-se que através da maximização da rentabilidade, o método de otimização linear dos ponderadores das variáveis *dummy* que compõem o F_score (OLP_rt) é capaz de potenciar a estratégia base, apresentando não só um retorno bruto médio de 11,27% (ou de 3,36% quando ajustado ao mercado) superior em 1,47% ao da estratégia F_score base, como um maior rácio de Sharpe, o que indica que este método proporciona um retorno médio superior face ao risco enfrentado, e portanto supera a estratégia F_score. Os bons resultados do método OLP_rt contrastam com os do método regressão linear (RL_rt), que ao selecionar as empresas com maior rentabilidade prevista em função das nove variáveis *dummy* que compõem o indicador global, permite um retorno médio de apenas 5,74%, superior em pouco mais de 1% ao retorno médio da amostra.

Tabela 7- Resultados da Otimização linear da Estratégia F_score aplicada a empresas de valor nos anos de teste (2002-2016)

	Otimização linear					
	F_score	Todas as empresas	OLP_rt	RL_rt	OLP_ec	RL_ec
2002	0,5210	0,5098	0,5107	0,4604	0,4418	0,4466
2003	0,0755	0,0955	0,0453	0,0006	-0,0703	0,1097
2004	0,3303	0,2306	0,3369	0,0814	0,3153	0,1819
2005	0,1800	0,1601	0,2069	0,3163	0,1385	0,1791
2006	-0,4663	-0,3835	-0,4055	-0,2990	-0,3779	-0,3370
2007	-0,4471	-0,5922	-0,4625	-0,6643	-0,4872	-0,4575
2008	0,4713	0,3429	0,6871	0,2695	0,3890	0,4413
2009	-0,0657	0,1192	0,0954	0,2180	0,2318	0,2284
2010	-0,1327	-0,2107	-0,1668	-0,2634	-0,0682	-0,2195
2011	0,0129	-0,0253	0,0271	-0,0656	0,0091	-0,0311
2012	0,3304	0,3037	0,3806	0,3476	0,3063	0,4168
2013	0,1568	0,0175	0,1630	0,0736	0,1337	0,1092
2014	0,0947	-0,0810	-0,0221	-0,0287	0,0808	0,0173
2015	0,2642	0,1485	0,2229	0,2610	0,1974	0,2480
2016	0,1448	0,0494	0,0717	0,1534	0,1090	0,1347
Média anual	0,0980	0,0456	0,1127	0,0574	0,0900	0,0979
Desvio-Padrão (diário)	0,0084	0,0063	0,0085	0,0083	0,0080	0,0078
Rácio de Sharpe (diário)	0,0511	0,0428	0,0534	0,0382	0,0484	0,0471
Taxa de sucesso	0,5255	0,4785	0,5169	0,5072	0,5118	0,5156

Na tabela constam os retornos brutos anuais para todas as empresas de valor consideradas, para o portfólio de elevado F_score, e para os portfólios dos métodos de otimização linear (OLP e RL, descritos na Seção 3), entre 2002 e 2016. O retorno bruto corresponde ao lucro resultante da compra de uma ação no final de abril e sua venda no final de abril no ano seguinte. “Média anual” corresponde à média do retorno bruto anual entre 2002 e 2016. O desvio-padrão foi obtido através da média dos desvios-padrão de cada um dos 15 anos (2002-2016). O desvio-padrão em cada ano foi obtido através da série do excesso de rentabilidade diária em relação à taxa de juro sem risco. O rácio de Sharpe corresponde à média dos rácios de Sharpe em cada um dos 15 anos. O rácio de Sharpe de cada ano foi obtido através do quociente entre a média e o desvio-padrão da série do excesso de rentabilidade diária. No cálculo do rácio de Sharpe e do desvio-padrão foi considerado o rebalanceamento diário de modo a manter um portfólio de proporções iguais. A “Taxa de sucesso” corresponde à taxa média de empresas que obtém retornos superiores aos dos respetivos índices de mercado.

A Tabela 7 evidencia também que o método RL é mais eficaz quando a variável dependente é o equivalente certo, uma vez que RL_ec exibe um retorno médio anual de 9,79% (superando em 4,05% o retorno do RL_rt), aproximadamente igual ao da estratégia F_score (9,80%) e apresenta um menor desvio-padrão. Todavia apresenta um rácio de Sharpe inferior, e por esse motivo não é preferível à estratégia base. Já o método OLP não é tão

eficaz quando maximiza o equivalente certo (OLP_{ec}), permitindo um retorno anual de apenas 9%. Ainda assim exibe um rácio de Sharpe superior ao método RL_{ec}, isto é, apresenta um excesso de rentabilidade superior face ao nível de risco, e dessa forma demonstra uma melhor relação entre rentabilidade e risco. Por último, destaca-se o facto de o portefólio da estratégia F_{score} deter em média mais empresas com retornos superiores aos dos respetivos índices de mercado face aos 4 métodos de otimização linear, com aproximadamente 52,55%.

Nas Tabelas 8 e 9 estão expostos os resultados, para os anos de teste, da otimização da estratégia F_{score} com base em técnicas de *machine learning* que preveem quer a rentabilidade quer o equivalente certo, respetivamente. A Tabela 8 mostra que as estratégias baseadas nas técnicas *machine learning* utilizadas não são boas, ao ponto de o maior retorno médio proporcionado ser de apenas 5,30% através da compra do portefólio *ensemble 2*, ainda assim inferior ao método RL_{rt} (5,74%). Somente as árvores de regressão, *support vector machines* (SVM) e a combinação *ensemble 2* apresentam um retorno médio superior à rentabilidade média de todas as empresas da amostra, ainda que tenham rácios de Sharpe inferiores, enquanto as técnicas *random forest* e *bagging* apresentam os retornos mais baixos (1,36% e -0,18%, na respetiva ordem). A Tabela 9 apresenta resultados similares aos expostos na Tabela 8, o que indica que as metodologias baseadas em técnicas *machine learning* utilizadas não são capazes de potenciar a estratégia F_{score} quando selecionam um portefólio constituído pelas empresas que apresentam as maiores previsões do equivalente certo. Todas estas metodologias apresentam rácios de Sharpe inferiores à média de todas as empresas de valor, e só as árvores de regressão permitem um retorno médio de 5,28%, superior à média de toda a amostra (4,56%), ainda assim bastante reduzido tendo em conta os 9,80% da estratégia F_{score}, ilustrando de forma clara a ineficácia destas técnicas em bater a estratégia F_{score}.

Os resultados exibidos nas Tabelas 8 e 9 relativos à otimização com recurso a técnicas de *machine learning* são decepcionantes, pois esperar-se-ia que as técnicas de *machine learning* pudessem otimizar a estratégia F_{score} dada a sua elevada capacidade de aprendizagem a partir dos dados e por apresentarem elevados retornos nas amostras de treino (Anexo 12 e Anexo 13). À exceção das árvores de regressão e SVM, que apresentam retornos médios inferiores aos métodos lineares, os retornos médios para os anos de treino das restantes técnicas de *machine learning* são muito elevados, acima dos 22%, com a técnica *random forest* a apresentar retornos médios acima dos 35%, o que expõe o problema do

sobreajustamento dos dados, na medida em que elevados resultados contrastam de forma clara com os apresentados nos anos de teste.⁴⁰

Tabela 8- Resultados da Otimização da Estratégia F_score baseada em técnicas de *machine learning* para os anos de teste (2002-2016) – Previsão da rentabilidade

Técnicas de <i>machine learning</i> – Previsão da rentabilidade						
	AR	RF	SVM	Bag	Ens.4	Ens.2
2002	0,4466	0,3673	0,3976	0,3994	0,3839	0,3713
2003	0,0200	0,2235	0,0116	0,0158	0,0669	0,0379
2004	0,2252	0,2502	0,1780	0,0968	0,2656	0,3115
2005	0,1906	0,1076	0,2030	0,1477	0,2890	0,2672
2006	-0,3308	-0,4369	-0,2112	-0,4547	-0,2822	-0,3079
2007	-0,5605	-0,7464	-0,6354	-0,8056	-0,7863	-0,7715
2008	0,2406	0,3769	0,2534	0,4271	0,4011	0,4453
2009	0,1698	0,0738	0,1415	0,1827	0,2091	0,3143
2010	-0,2656	-0,2910	-0,2293	-0,3879	-0,4080	-0,2492
2011	-0,0089	-0,0350	-0,0913	-0,1175	-0,1015	-0,1217
2012	0,3732	0,2862	0,4152	0,3318	0,3529	0,3529
2013	0,0582	-0,0486	0,0771	0,0255	0,0241	0,0328
2014	-0,1072	-0,2785	-0,0088	-0,1255	-0,1881	-0,1034
2015	0,1734	0,2365	0,1854	0,2196	0,2685	0,2168
2016	0,1534	0,1187	0,0901	0,0179	0,0455	-0,0005
Média anual	0,0519	0,0136	0,0518	-0,0018	0,0360	0,0530
Desvio-Padrão (diário)	0,0071	0,0093	0,0078	0,0088	0,0085	0,0081
Rácio de Sharpe (diário)	0,0329	0,0162	0,0339	0,0011	0,0274	0,0306
Taxa de sucesso	0,4924	0,4521	0,5174	0,4134	0,4543	0,4690

A presente tabela expõe os principais resultados das metodologias baseadas em técnicas de *machine learning* que geram previsões da rentabilidade com o propósito de otimizar a estratégia F_score aplicada a empresas de valor. Os acrônimos AR, RF SVM, Bag designam, na respetiva ordem, as técnicas de *machine learning* utilizadas: árvores de regressão, *random forest*, *support vector machine* e *bagging*, enquanto que Ens.4 e Ens.2 representam as metodologias em que se combinam as previsões das 4 técnicas de *machine learning*, através da média, (Ens.4), e através da mediana (Ens.2). Todas as outras definições (retornos brutos, média anual, desvio-padrão, rácio de Sharpe e taxa de sucesso) estão descritas na Tabela 7.

⁴⁰ Face aos resultados muito fracos das técnicas de *machine learning* foram previstos os valores da rentabilidade tendo por base os valores reais dos 9 indicadores que compõe o F_score. Verificou-se que os resultados ainda são piores, o que está em harmonia com as evidências de Patel et al. (2015) que demonstram, no âmbito da análise técnica, que as técnicas *machine learning* são mais eficazes quando se usa variáveis binárias em vez dos valores reais das variáveis.

Tabela 9- Resultados da Otimização da Estratégia F_score baseada em Técnicas de *machine learning* para os anos de teste (2002-2016) – Previsão do equivalente certo

Técnicas de <i>machine learning</i>- Previsão do equivalente certo						
	AR	RF	SVM	Bag	Ens.4	Ens.2
2002	0,4928	0,5229	0,5127	0,5084	0,4553	0,4607
2003	0,0787	0,1872	-0,0263	0,0646	0,0008	0,0008
2004	0,2327	0,2497	0,1960	0,2260	0,2409	0,3013
2005	0,1801	0,1171	0,2244	0,1700	0,1473	0,1473
2006	-0,3186	-0,5265	-0,2721	-0,2455	-0,3448	-0,3194
2007	-0,5847	-0,7987	-0,5766	-0,7175	-0,7808	-0,7144
2008	0,2346	0,4414	0,1721	0,3874	0,4618	0,3656
2009	0,1420	0,0213	0,0931	0,1499	0,1240	0,1524
2010	-0,1947	-0,1896	-0,2913	-0,2565	-0,2554	-0,2722
2011	-0,0548	-0,0564	-0,0638	-0,2012	-0,1083	-0,1053
2012	0,3782	0,3330	0,2931	0,3693	0,3013	0,3969
2013	0,0582	0,0288	-0,0057	-0,0197	-0,0133	-0,0139
2014	-0,0956	-0,1680	-0,0369	-0,0877	-0,2521	-0,2343
2015	0,1480	0,1761	0,1054	0,1606	0,1811	0,2125
2016	0,0957	0,0261	0,1742	-0,0171	0,0210	0,0158
Média anual	0,0528	0,0243	0,0332	0,0327	0,0119	0,0262
Desvio-Padrão (diário)	0,0063	0,0091	0,0088	0,0085	0,0086	0,0084
Rácio de Sharpe (diário)	0,0387	0,0209	0,0182	0,0242	0,0171	0,0244
Taxa de sucesso	0,4681	0,4729	0,4637	0,4648	0,4487	0,4785

A presente tabela expõe os principais resultados das metodologias baseadas em técnicas de *machine learning* que geram previsões do equivalente certo com o propósito de otimizar a estratégia F_score aplicada a empresas de valor. Os acrônimos AR, RF SVM, Bag designam, na respetiva ordem, as técnicas de *machine learning* utilizadas: árvores de regressão, *random forest*, *support vector machine* e *bagging*, enquanto que Ens.4 e Ens.2 representam as metodologias em que se combinam as previsões das 4 técnicas de *machine learning*, através da média, (Ens.4), e através da mediana (Ens.2). Todas as outras definições (retornos brutos, média anual, desvio-padrão, rácio de Sharpe e taxa de sucesso) estão descritas na Tabela 7.

Em suma, de todos os métodos utilizados o único que supera a estratégia F_score é o OLP_rt, isto é, o método que pondera as variáveis *dummy* que compõe o F_score de forma a maximizar a rentabilidade do portefólio constituído pelas empresas com F_score igual ou superior a 8. Os resultados indicam ainda que as técnicas de otimização linear (OLP e RL) são mais eficazes que as técnicas de *machine learning*, apresentando maiores retornos médios assim como rácios de Sharpe superiores. Não obstante os elevados retornos nos anos de treino dispostos pela generalidade das técnicas de *machine learning*, os resultados nos anos de teste são muito fracos, o que pode ser explicado pelo problema do sobreajustamento

dos dados, sobretudo na técnica *random forest*, mas também devido à necessidade de um maior volume de observações. Posto isto, conclui-se que apesar da simplicidade metodológica, a estratégia F_score aplicada a empresas de valor funciona muito bem, pois apenas um dos métodos testados consegue um melhor desempenho.

5.2.2 Otimização da Estratégia N_score

Esta subsecção tem como objetivo otimizar a posição longa da estratégia N_score (aplicada a todas as empresas), que consiste na compra das ações com um N_score igual ou superior a 9. Na Tabela 10 estão representados os resultados da otimização linear, pelas abordagens em que se maximiza (OLP) ou prevê (RL) quer a rentabilidade quer o equivalente certo, da estratégia N_score, que permite um retorno médio anual de 8,78% entre 2002 e 2016. Tal como na subsecção anterior, os elevados retornos anuais negativos entre abril de 2007 e abril de 2009 afetos aos anos 2006 e 2007, em que se aplica a estratégia de investimento, evidenciam o impacto da grave crise financeira de 2007-2009 nos principais mercados acionistas europeus.

A Tabela 10 mostra ainda que os resultados da otimização linear dos métodos OLP e RL apresentam, em média, retornos superiores à estratégia que se pretende otimizar, quer pela abordagem da rentabilidade quer pela do equivalente certo. À semelhança da otimização linear do F_score na subsecção anterior, o método linear OLP_rt apresenta o maior retorno médio, aproximadamente 10,28% (ou 2,71% quando ajustado ao mercado), superando em 1,5% a estratégia N_score, enquanto os restantes métodos, OLP_ec, RL_ec e RL_rt permitem obter retornos médios anuais de 9,39%, 9,30% e 8,89%, respetivamente. Contudo, quando se tem em consideração a relação rentabilidade risco, mensurada pelo rácio de Sharpe, verifica-se que apenas os métodos OLP apresentam rácios de Sharpe superiores face à estratégia base. Por outras palavras, os métodos RL apesar de apresentarem retornos médios superiores à estratégia N_score, apresentam uma relação rentabilidade-risco menos favorável.

Independentemente do sucesso dos métodos lineares (OLP) que permitem potenciar o retorno obtido pela estratégia N_score, a estratégia F_score continua a conferir um retorno médio anual superior, de 10,41%. No entanto, na Tabela 10 verifica-se que quer a estratégia N_score quer os métodos OLP apresentam desvios-padrão inferiores, e desta forma os resultados obtidos por estas estratégias aproximam-se dos obtidos pela estratégia F_score quando se tem em consideração a relação rentabilidade-risco, dada pelo rácio de Sharpe.

Essa aproximação dos valores dos rácios de Sharpe face aos valores médios da rentabilidade anual é de tal forma notória, que o método OLP_rt apresenta um retorno médio anual inferior em 0,13%, mas exibe um rácio de Sharpe superior ao da estratégia F_score, e desse modo é preferível num ponto de vista de rentabilidade e risco.

Tabela 10- Resultados da Otimização linear da Estratégia N_score para os anos de teste (2002-2016)

	Otimização linear						
	F_score	Todas as empresas	N_score	OLP_rt	RL_rt	OLP_ec	RL_ec
2002	0,4497	0,4099	0,4859	0,4570	0,3764	0,4471	0,4227
2003	0,1940	0,1064	0,1819	0,1923	0,1924	0,1438	0,1835
2004	0,3100	0,2681	0,2443	0,3056	0,3186	0,3003	0,3216
2005	0,1535	0,1333	0,1418	0,1521	0,1448	0,1384	0,1811
2006	-0,2715	-0,3543	-0,3369	-0,2956	-0,2765	-0,2772	-0,2302
2007	-0,4611	-0,5454	-0,5273	-0,4898	-0,5123	-0,4828	-0,5205
2008	0,3607	0,3001	0,3691	0,3224	0,2986	0,3129	0,3089
2009	0,1731	0,1191	0,1598	0,1930	0,2038	0,1580	0,1565
2010	-0,1158	-0,1487	-0,1173	-0,0552	-0,0767	-0,1058	-0,0684
2011	0,0907	0,0181	0,0475	0,0835	0,1163	0,0979	0,1421
2012	0,2892	0,2476	0,3355	0,2665	0,2159	0,3143	0,2119
2013	0,0862	0,0265	0,0176	0,0690	0,1247	0,0914	0,1138
2014	0,0240	-0,1116	0,0278	0,0265	-0,0022	0,0091	-0,0050
2015	0,1987	0,1298	0,1769	0,1619	0,1903	0,1892	0,1745
2016	0,0808	0,0012	0,1102	0,1526	0,0194	0,0719	0,0021
Média anual	0,1041	0,0400	0,0878	0,1028	0,0889	0,0939	0,0930
Desvio-Padrão (diário)	0,0065	0,0066	0,0055	0,0057	0,0073	0,0062	0,0078
Rácio de Sharpe (diário)	0,0946	0,0522	0,0899	0,1018	0,0833	0,0932	0,0820
Taxa de sucesso	0,5369	0,4943	0,5091	0,5276	0,5323	0,5224	0,5383

A presente tabela apresenta a mesma estrutura que as anteriores. Estão representados para cada ano os retornos médios brutos de todas as empresas consideradas, dos portfólios de elevado *score* das estratégias F_score e N_score, e dos portfólios dos métodos que realizam otimizações lineares (OLP e RL, ver Secção 3 para mais detalhes). Todas as outras definições (média anual, desvio-padrão, rácio de Sharpe e taxa de sucesso) estão descritas na Tabela 7.

Uma das razões para os fracos resultados da otimização da estratégia F_score, aplicada a empresas de valor, com recurso a técnicas de *machine learning* pode estar relacionada com o facto de estas técnicas precisarem de um maior volume de dados. Por esse motivo, espera-

se que os resultados demonstrem uma maior eficácia destas técnicas na otimização da estratégia N_score já que o número de observações é muito superior face a amostra 75% de empresas de valor.⁴¹

Na Tabela 11 são apresentados os resultados, para os anos de teste, da otimização da estratégia N_score por meio de técnicas de *machine learning* que preveem a rentabilidade anual de cada empresa em função das variáveis *dummy* que compõe o N_score. Observa-se na Tabela 11 que todas as técnicas, à exceção da *random forest*, apresentam retornos médios anuais e rácios de Sharpe superiores à média da amostra, o que revela desde logo uma maior eficácia destas técnicas relativamente à otimização da estratégia F_score para as empresas de valor. Tal como nos resultados da secção anterior, as técnicas *machine learning* mais eficientes na previsão da rentabilidade são as árvores de regressão e SVM que apresentam respetivamente um retorno médio anual de 8,72% e 8,52%, valores esses que estão muito próximos do retorno médio assegurado pela estratégia base (8,78%). Todavia, a rentabilidade proporcionada por essas técnicas está associada a um maior risco refletido nos desvios-padrão superiores, o que explica o facto de os rácios de Sharpe serem inferiores aos da estratégia N_score. Face ao exposto nenhuma das técnicas exibidas na Tabela 11 é capaz de bater a estratégia N_score, no entanto há que salientar que o método árvores de regressão utilizado para prever a rentabilidade apresenta um rácio de Sharpe relativamente elevado (0,0883), de tal forma que garante um excesso de rentabilidade face ao nível de risco superior ao método de otimização linear RL_rt.

A otimização mediante técnicas de *machine learning* que procuram prever o equivalente certo revela ser mais eficaz, dado que os resultados apresentados na Tabela 12 para os anos de teste são, na generalidade, melhores comparativamente com os da Tabela 11, quer ao nível do retorno médio quer ao nível do rácio de Sharpe. A Tabela 12 mostra que com este procedimento todas as técnicas permitem em média retornos e rácios de Sharpe superiores aos do portefólio composto por todas as empresas da amostra. A técnica *random forest* exhibe de forma clara o pior desempenho, com um retorno médio anual de 4,65% e um rácio de Sharpe de apenas 0,0528, enquanto as árvores de regressão e SVM são as mais eficazes, apresentando retornos médios anuais de 9,36% e 8,96% respetivamente, superiores aos 8,78% da estratégia N_score. Porém, esta rentabilidade superior à da estratégia N_score

⁴¹ A amostra utilizada na otimização da estratégia N_score é composta por 27641 observações de empresas entre 1992 e 2016, consideravelmente superior à amostra utilizada na otimização do F_score que era constituída por 4171 observações.

deve-se a uma maior volatilidade visto que as duas técnicas (AR e SVM) apresentam rácios de Sharpe inferiores. Desse modo, conclui-se que não é possível bater a estratégia N_score com recurso às técnicas de *machine learning* utilizadas que preveem o equivalente certo, contudo observa-se que quer as árvores de regressão quer SVM apresentam uma relação entre rentabilidade e risco mais favorável que o método RL_ec.

Tabela 11- Resultados da Otimização da Estratégia N_score com base em Técnicas de *machine learning* para os anos de teste (2002-2016) – Previsão da rentabilidade

Técnicas de <i>machine learning</i>- Previsão da rentabilidade						
	AR	RF	SVM	Bag	Ens.4	Ens.2
2002	0,3879	0,3774	0,3732	0,3493	0,3798	0,3700
2003	0,1654	0,0679	0,1877	0,1252	0,1055	0,1916
2004	0,3222	0,2820	0,3158	0,2964	0,3189	0,3299
2005	0,1600	0,0663	0,1496	0,0902	0,0815	0,0912
2006	-0,2820	-0,3803	-0,3051	-0,3667	-0,3299	-0,3263
2007	-0,5156	-0,6069	-0,5287	-0,5385	-0,5215	-0,5078
2008	0,3038	0,2032	0,3121	0,2474	0,2663	0,2818
2009	0,2055	0,1175	0,1770	0,2255	0,2041	0,2103
2010	-0,1375	-0,1039	-0,0542	-0,1153	-0,1167	-0,1094
2011	0,0925	0,0325	0,1028	0,0664	0,0780	0,0996
2012	0,2465	0,2099	0,1869	0,2399	0,2179	0,2200
2013	0,1207	0,0326	0,1512	0,0332	0,0813	0,0664
2014	0,0226	-0,1601	-0,0046	-0,0778	-0,0587	-0,0492
2015	0,1730	0,1597	0,1746	0,1294	0,1606	0,1694
2016	0,0436	0,0535	0,0394	0,0636	0,0625	0,0296
Média anual	0,0872	0,0234	0,0852	0,0512	0,0620	0,0712
Desvio-Padrão (diário)	0,0065	0,0065	0,0075	0,0062	0,0065	0,0066
Rácio de Sharpe (diário)	0,0883	0,0362	0,0807	0,0596	0,0665	0,0747
Taxa de sucesso	0,5277	0,4728	0,5378	0,4985	0,5114	0,5160

Nesta tabela encontram-se apresentados os retornos médios para cada um dos anos de teste da aplicação das metodologias baseadas em técnicas de *machine learning* (que preveem a rentabilidade) com o propósito de potenciar a rentabilidade obtida com compra de ações com elevado N_score. Os acrónimos AR, RF SVM, Bag designam, na respetiva ordem, as técnicas de *machine learning* utilizadas: árvores de regressão, *random forest*, *support vector machine* e *bagging*, enquanto que Ens.4 e Ens.2 representam as metodologias em que se combinam as previsões das 4 técnicas de *machine learning*, através da média, (Ens.4), e através da mediana (Ens.2). Todas as outras definições (média anual, desvio-padrão, rácio de Sharpe e taxa de sucesso) estão descritas na Tabela 7.

Tabela 12- Resultados da Otimização da Estratégia F_score mediante Técnicas de *machine learning* para os anos de teste (2002-2016) – Previsão do equivalente certo

Técnicas de <i>machine learning</i>- Previsão do equivalente certo						
	AR	RF	SVM	Bag	Ens.4	Ens.2
2002	0,4174	0,3664	0,4615	0,3701	0,3695	0,3859
2003	0,1836	0,0983	0,2059	0,1246	0,1630	0,1475
2004	0,3380	0,3461	0,3380	0,3324	0,3422	0,3338
2005	0,1823	0,1185	0,1387	0,1554	0,1334	0,1446
2006	-0,2470	-0,3397	-0,2485	-0,2886	-0,2739	-0,2547
2007	-0,5158	-0,5614	-0,5089	-0,5265	-0,4751	-0,4790
2008	0,3235	0,3171	0,3297	0,3147	0,3215	0,3158
2009	0,1710	0,1497	0,1468	0,2329	0,1730	0,2137
2010	-0,0546	-0,1628	-0,0666	-0,1453	-0,1246	-0,1250
2011	0,1426	0,0087	0,1144	0,0508	0,0989	0,1038
2012	0,2069	0,2204	0,2083	0,2769	0,2110	0,2233
2013	0,1128	0,0511	0,1068	0,0669	0,1251	0,1182
2014	-0,0360	-0,0905	-0,0177	-0,0453	-0,0319	-0,0267
2015	0,1880	0,1508	0,1393	0,1555	0,1622	0,1534
2016	-0,0090	0,0246	-0,0042	0,0490	0,0302	0,0427
Média anual	0,0936	0,0465	0,0896	0,0749	0,0816	0,0865
Desvio-Padrão (diário)	0,0077	0,0069	0,0079	0,0072	0,0075	0,0076
Rácio de Sharpe (diário)	0,0857	0,0528	0,0832	0,0724	0,0769	0,0786
Taxa de sucesso	0,5445	0,4910	0,5346	0,5136	0,5237	0,5254

Nesta tabela encontram-se apresentados os retornos médios para cada um dos anos de teste da aplicação das metodologias baseadas em técnicas de *machine learning* (que preveem o equivalente certo) com o propósito de potenciar a rentabilidade obtida com compra de ações com elevado N_score. Os acrónimos AR, RF SVM, Bag designam, na respetiva ordem, as técnicas de *machine learning* utilizadas: árvores de regressão, *random forest*, *support vector machine* e *bagging*, enquanto que Ens.4 e Ens.2 representam as metodologias em que se combinam as previsões das 4 técnicas de *machine learning*, através da média, (Ens.4), e através da mediana (Ens.2). Todas as outras definições (média anual, desvio-padrão, rácio de Sharpe e taxa de sucesso) estão descritas na Tabela 7.

Sumariamente, as duas Tabelas, 11 e 12, evidenciam um desempenho muito superior das técnicas de *machine learning* utilizadas relativamente à otimização da estratégia F_score (Subseção 5.2.1), sobretudo quando se prevê o equivalente certo. Para tal poderá ter contribuído o facto de o volume de observações ser superior. Todavia, este desempenho não é satisfatório, na medida em que nenhuma das técnicas é capaz de proporcionar retornos superiores aos da estratégia N_score sem que o nível de risco aumente. Por último, há que

destacar o facto de as técnicas de *machine learning random forest* e *bagging* que contêm os piores resultados nos anos de teste, apresentarem quer pela abordagem da previsão da rentabilidade (Anexo 14) quer pela do equivalente certo (Anexo 15), os retornos mais elevados nos anos de treino, muito superiores aos das técnicas lineares (OLP e RL). Em contrapartida, as técnicas de *machine learning* árvores de regressão e SVM, que dispõem dos retornos mais baixos nos anos de treino, (inferiores aos respetivos métodos lineares OLP e RL), são as que evidenciam os melhores desempenhos nos anos de teste.

6. Conclusão

O presente trabalho visa vários objetivos. Primeiro, pretende verificar a eficácia da estratégia de investimento baseada no F_score de Piotroski (2000), num contexto temporal e espacial diferente, nomeadamente analisando os principais mercados acionistas europeus durante o período de 1992 a 2016, o qual engloba a grave crise financeira de 2007-2009. Segundo, desenvolver uma análise comparativa entre o F-score e outro indicador fundamental, mais amplo, que denominamos de N_score, que para além das variáveis incluídas no F_score considera outras variáveis fundamentais tais como o *book-to-market* e a dimensão. E terceiro, analisar o efeito de várias metodologias, que no contexto daqueles indicadores globais fundamentais, possam potenciar o desempenho das estratégias de investimento, nomeadamente recorrendo a técnicas de otimização linear e não-linear.

Os resultados mostram claramente que quer o F_score quer o N_score apresentam capacidade para diferenciar vencedores de perdedores, proporcionando retornos médios (brutos ou ajustados ao mercado) acima dos 20% e 15%, respetivamente. No entanto, os elevados retornos são obtidos em grande parte pelas posições curtas sobre as empresas de baixo *score*. O facto de o F_score proporcionar retornos superiores quer com compra quer com a venda a descoberto sugere que a incorporação do BM e da capitalização de mercado não permite potenciar a capacidade preditiva do F_score. Verifica-se ainda que num cenário de fortes restrições às vendas a descoberto poderá ser útil, para um investidor em empresas de valor, deixar de fora as empresas que apresentam os maiores rácios BM, dado que tais valores poderão sinalizar dificuldades nessas empresas, refletidas num valor de mercado muito baixo. Por último, constata-se que o sucesso da estratégia F_score não é restrito ao contexto de elevado rácio BM, pois a aplicação da mesma para todas as empresas apresenta retornos médios similares aos obtidos para as empresas de valor (acima dos 20%). Deste modo, os resultados indicam que a elevada eficácia da estratégia fundamental baseada no F_score de Piotroski (2000) parece não ser circunstancial ao tempo (diferente período amostral) e ao espaço (outros mercados acionistas que não o Norte-Americano).

Relativamente ao processo de otimização das estratégias, no qual foram excluídas as vendas a descoberto, conclui-se que os métodos mais simples, que executam otimizações lineares, são os mais eficazes e ainda que o aumento da quantidade de observações permite um aumento da eficácia geral dos métodos, sobretudo das técnicas de *machine learning*.

Da otimização da estratégia F_score aplicada a empresas de valor sobressaem os pobres resultados obtidos pelas técnicas de *machine learning* e constata-se que somente o método OLP_rt consegue bater a estratégia F_score apresentando um retorno médio superior, em 1,47%, e um maior rácio de Sharpe. Quando se otimiza a estratégia N_score, os resultados dos métodos de otimização são mais eficazes, sobretudo das técnicas de *machine learning*, o que sugere que a eficácia destes métodos pode depender do volume de observações utilizado. Contudo, apenas o método OLP quer através da maximização do retorno médio quer do equivalente certo, consegue obter rentabilidades superiores, (em 1,5% e 0,61%, respetivamente) e maiores rácios de Sharpe que a estratégia N_score.

De todos os métodos testados, o método OLP_rt, em que se otimizam os ponderadores das variáveis fundamentais na forma binária de forma a maximizar a rentabilidade do portefólio de elevado *score*, é o mais eficaz, conseguindo bater as estratégias base F_score e N_score. Esta metodologia permite, entre 2002 e 2016, em média, um retorno anual bruto (ajustado ao mercado) de 11,27% (3,36%) e um rácio de Sharpe de 0,0534 com a compra de ações de valor mediante o F_score otimizado. E, quando aplicada a todas as empresas, tendo por base a informação do indicador global N_score proporciona, em média, para o mesmo período, um retorno anual bruto (ajustado ao mercado) de 10,28% (2,71%) e um rácio de Sharpe de 0,1018. Finalmente, observa-se, de forma clara, que os retornos médios referentes ao período de otimização das estratégias (2002-2016) são afetados pela grave crise financeira (2007-2009), durante a qual se registaram perdas avultadas. Ainda que a metodologia OLP_rt disponha de um melhor desempenho nesse período conturbado, comparativamente com a estratégia base a otimizar, contabiliza, em média, um retorno bruto (ajustado ao mercado) anual de -43,40% (-14,69%), com a compra de ações de valor através do F_score otimizado, e de -39,27% (-9,22%) quando aplicada a todas as empresas tendo por base o indicador N_score otimizado.

Em suma, o presente trabalho permite concluir que a metodologia simples e linear assente nestas estratégias de investimento fundamentais é bastante eficiente, não só pelo facto de o único método a conseguir bater as estratégias consistir numa otimização linear, como pela elevada consistência e robustez da estratégia F_score em diversos contextos económicos.

7. Bibliografia

- Abarbanell, J. S., & Bushee, B. J. (1998). Abnormal Returns to a Fundamental Analysis Strategy. *The Accounting Review*, 73(1), 19-45.
- Aggarwal, N., & Gupta, M. (2009). Do High Book-to-Market Stocks Offer Returns to Fundamental Analysis in India?. *Decision (0304-0941)*, 36(2), 155-175.
- Amor-Tapia, B., & Tascón, M. (2016). Separating Winners from Losers: Composite Indicators Based on Fundamentals in the European Context. *Finance a Uver - Czech Journal of Economics and Finance*, 66(1), 70-94.
- Balachandran, A., Saraph, D., & Ang, E. (2013, dezembro 11). Learning to Identify Winning Stocks. Obtido em junho 18, 2019, de <http://cs229.stanford.edu/proj2013/AngBalachandranSaraph-LearningToIdentifyWinningStocks.pdf>
- Breiman, L. (1996). Bagging predictors. *Machine Learning*, 24(2), 123-140. <https://doi.org/10.1007/BF00058655>
- Breiman, L. (2001). Random Forests. *Machine Learning*, 45(1), 5-32. <https://doi.org/10.1023/A:1010933404324>
- Chen, N. F., & Zhang, F. (1998). Risk and Return of Value Stocks. *The Journal of Business*, 71(4), 501-535.
- DeMiguel, V., Garlappi, L., & Uppal, R. (2007). Optimal Versus Naive Diversification: How Inefficient is the 1/N Portfolio Strategy?. *The Review of Financial Studies*, 22(5), 1915-1953. <https://doi.org/10.1093/rfs/hhm075>
- Dosamantes C. (2013). The Relevance of Using Accounting Fundamentals in the Mexican Stock Market. *Journal of Economics. Finance and Administrative Science*, 18, 2-10.
- Dreman, D. (2008). *Contrarian Investment Strategies: The Next Generation*. New York: Simon and Schuster.
- Fama, E. F., & French, K. R. (1992). The Cross-Section of Expected Stock Returns. *The Journal of Finance*, 47(2), 427-465.
- Fama, E. F., & French, K. R. (1995). Size and Book-to-Market Factors in Earnings and Returns. *The Journal of Finance*, 50(1), 131-155.
- Iqbal, N., Khattak, S. R., & Khattak, M. A. (2013). Does Fundamental Analysis Predict Stock Returns? Evidence from Non-Financial Companies Listed on KSE. *Knowledge Horizons. Economics*, 5(4), 182-190.
- Lakonishok, J., Shleifer, A., & Vishny, R. W. (1994). Contrarian Investment, Extrapolation, and Risk. *The Journal of Finance*, 49(5), 1541-1578.
- La Porta, R., Lakonishok, J., Shleifer, A., & Vishny, R. (1997). Good News for Value Stocks: Further Evidence on Market Efficiency. *The Journal of Finance*, 52(2), 859-874.
- Lev, B., & Thiagarajan, S. R. (1993). Fundamental Information Analysis. *Journal of Accounting Research*, 31(2), 190-215. <https://doi.org/10.2307/2491270>
- Lopes, A. B., & Galdi, F. C. (2006). *Financial Statement Analysis also Separate Winners from Losers in Brazil* (Working Paper). University of Sao Paulo.

- Mohanram, P. S. (2005). Separating Winners from Losers among Low Book-to-Market Stocks using Financial Statement Analysis. *Review of Accounting Studies*, 10(2-3), 133-170. <https://doi.org/10.1007/s11142-005-1527-3>
- Navas, R., Gama, A. P. M., & Bentes, S. R. (2018). Can Fundamental Analysis Provide Relevant Information for Understanding the Underlying Value of a Company?. In V. Bobek (Ed.), *Trade and Global Market* (pp.155-170). Local: IntechOpen. <https://doi.org/10.5772/32009>
- Noma, M. (2010). Value Investing and Financial Statement Analysis. *Hitotsubashi journal of commerce and management*, 44(1), 29-46. <http://doi.org/10.15057/18701>
- Ou, J. A., & Penman, S. H. (1989). Accounting Measurement, Price-Earnings Ratio, and the Information Content of Security Prices. *Journal of Accounting Research*, 27, 111-144. <https://doi.org/10.2307/2491068>
- Patel, J., Shah, S., Thakkar, P., & Kotecha, K. (2015). Predicting stock and stock price index movement using Trend Deterministic Data Preparation and machine learning techniques. *Expert Systems with Applications*, 42(1), 259-268. <https://doi.org/10.1016/j.eswa.2014.07.040>
- Patel, P. N., Yao, S., & Barefoot, H. (2006, agosto 15). High Yield, Low Payout. Obtido em junho 18, 2019, de https://papers.ssrn.com/sol3/papers.cfm?abstract_id=946448
- Penman, S. H., & Zhang, X. J. (2002). Accounting conservatism, the quality of earnings, and stock returns. *The accounting review*, 77(2), 237-264.
- Piotroski, J. D. (2000). Value Investing: The Use of Historical Financial Statement Information to Separate Winners from Losers. *Journal of Accounting Research*, 38, 1-41.
- Piotroski, J. D. (2005). Discussion of “Separating Winners from Losers among Low Book-to-Market Stocks using Financial Statement Analysis”. *Review of Accounting Studies*, 10(2-3), 171-184. <https://doi.org/10.1007/s11142-005-1527-3>
- Piotroski, J. D., & So, E. C. (2012). Identifying Expectation Errors in Value/Glamour Strategies: A Fundamental Analysis Approach. *The Review of Financial Studies*, 25(9), 2841-2875. <https://doi.org/doi:10.1093/rfs/hhs061>
- Quinlan, J. R. (1992). Learning with continuous classes. In A. Adams & L. Sterling (Eds.), *Proceedings of the 5th Australian Joint Conference on Artificial Intelligence* (pp. 343-348). Singapura: World Scientific.
- Shen, K. Y. (2011). Implementing Value Investing Strategy by Artificial Neural Network. *International Journal of Business and Information Technology*, 1(1), 12-22.
- Sloan, R. G. (1996). Do Stock Prices Fully Reflect Information in Accruals and Cash Flows About Future Earnings?. *The Accounting review*, 71(3), 289-315.
- Soliman, M. (2004). Using Industry-Adjusted Dupont Analysis to Predict Future Profitability (Working Paper). Stanford University. Obtido de https://papers.ssrn.com/sol3/papers.cfm?abstract_id=456700
- Walkshäusl, C. (2017). Expectation Errors in European Value-Growth Strategies. *Review of Finance*, 21(2), 845-870. <https://doi.org/10.1093/rof/rfw012>
- Wang, Y., & Witten, I. H. (1996). *Induction of Model Trees for Predicting Continuous Classes* (Working Paper No. 96/23). Hamilton, New Zealand: University of Waikato, Department of Computer Science. Obtido de

<https://researchcommons.waikato.ac.nz/bitstream/handle/10289/1183/uow-cs-wp-1996-23.pdf?sequence=1&isAllowed=y>

Witten, I. H., & Frank, E. (2005). *Data Mining: Practical Machine Learning Tools and Techniques*. (2.^a ed.). San Francisco: Morgan Kaufmann.

Apêndice

Construção do indicador Fundamental F_score

O F_score é obtido através da soma das nove variáveis *dummy*, $F = \sum_{i=1}^9 Fi$, podendo assumir um valor inteiro entre 0 e 9.

Variável Fundamental	Definição da variável	Obtenção da <i>dummy</i>
Rentabilidade (ROA)	Resultado líquido antes de itens extraordinários e dividendos preferenciais no final do ano fiscal dividido pelo total dos ativos no início do ano	F1 = 1 se ROA > 0, caso contrário F1 = 0
Variação da rentabilidade (ΔROA)	Rentabilidade do ano corrente menos a rentabilidade do ano anterior	F2=1 se $\Delta ROA > 0$, caso contrário F2 = 0
Rácio <i>Cash Flows</i> (CF)	Soma do lucro líquido e de todos os encargos ou créditos não monetários (fluxo de caixa da empresa) sobre o total do ativo no início do ano	F3 = 1 se CF > 0, caso contrário F3 = 0
<i>Accruals</i>	ROA - CF	F4 = 1 se <i>Accruals</i> < 0, caso contrário F4 = 0
Variação da liquidez ($\Delta Liquid$)	Variação anual do rácio de liquidez [rácio de liquidez = ativo corrente / passivo corrente]	F5 = 1 se $\Delta Liquid > 0$, caso contrário F5 = 0
Variação da dívida ($\Delta Lever$)	Variação anual do rácio de dívida [rácio de dívida = dívida de longo prazo / total dos ativos]	F6 = 1 se $\Delta Lever < 0$, caso contrário F6 = 0
Alterações do capital (<i>Eq_offer</i>)	Receitas líquidas de venda/emissão de ações ordinárias ou preferenciais	F7 = 1 se <i>Eq_offer</i> = 0, caso contrário F7 = 0
Alterações na margem de lucro bruta ($\Delta Margin$)	Variação anual da margem de lucro bruta [margem de lucro bruta = lucro bruto / (vendas líquidas ou receitas)]	F8 = 1 se $\Delta Margin > 0$, caso contrário F8 = 0
Alterações na rotação do ativo ($\Delta Turn$)	Variação anual da rotação do ativo [rotação do ativo = (vendas líquidas ou receitas) / total do ativo]	F9 = 1 se $\Delta Turn > 0$, caso contrário F9 = 0

Anexos

Anexo 1 – Variáveis recolhidas da base de dados Datastream/Eikon

- **Capitalização de Mercado** - corresponde ao número de ações ordinárias em circulação multiplicado pelo preço da ação no final do ano fiscal.
- **Total de Ativos**- representam a soma do total do ativo corrente, contas a receber de longo prazo, investimentos em subsidiárias não consolidadas, outros investimentos, ativos tangíveis e outros ativos, no final do ano fiscal.
- **Valor contabilístico por ação**- representa o valor contabilístico sobre o número de ações em circulação no final do ano fiscal
- **Resultado líquido antes de itens extraordinários/dividendos preferenciais** - representa o lucro antes dos itens extraordinários e dos dividendos preferenciais e ordinários, mas após as receitas e despesas operacionais e não operacionais, as reservas, os impostos sobre o lucro, as participações minoritárias e os lucros.
- **Margem de lucro bruta** – corresponde à percentagem do lucro bruto/margem bruta sobre as vendas líquidas ou receitas.
- **Cash Flows**- representam a soma do lucro líquido e de todos os encargos ou créditos não monetários. É o fluxo de caixa da empresa. (Se uma Demonstração de Alterações na Posição Financeira não tiver sido fornecida, mas a empresa divulgar um fluxo de caixa agregado, esse valor foi utilizado. Quando o fluxo de caixa não foi divulgado de qualquer maneira, é estimado com base no lucro líquido antes dos dividendos preferenciais mais depreciação, encargos de reservas, provisão para perdas com empréstimos para bancos e provisão para benefícios futuros para seguradoras.)
- **Rácio de liquidez** - corresponde ao rácio entre o total dos ativos correntes sobre o total dos passivos correntes.
- **Dívida de longo prazo** - representa todas as obrigações financeiras que geram juros, excluindo as quantias devidas no prazo de um ano. É apresentado o valor líquido de prémio ou desconto.
- **Rotação do ativo total** – representa as vendas líquidas ou receitas sobre o total dos ativos
- **Receitas líquidas de venda/emissão de ações ordinárias ou preferenciais**- representam o valor que uma empresa recebeu da venda de ações ordinárias e/ou preferenciais. Inclui os valores recebidos da conversão de debêntures ou ações preferenciais em ações ordinárias, permuta de ações ordinárias por debêntures, venda de ações em tesouraria, ações emitidas para aquisições e saída de opções de ações.
- **Índice de retorno total**- o índice de retorno está disponível para ações individuais e fundos de investimento. Este indicador mostra um crescimento teórico no valor de uma participação detida ao longo de um período especificado, assumindo que os dividendos são reinvestidos para comprar unidades adicionais de um capital ou *unit trust* ao preço de fecho aplicável na data ex-dividendo.
- **Preço não ajustado** - Este é o preço de fecho que não foi historicamente ajustado para bónus e *rights issues*.

Anexo 2- Descrição da amostra

Painel A. Observações empresas-ano por Bolsa de Valores (1992-2016)

		Frankfurt (Xetra)	Euronext Bruxelas	Mercado Contínuo Espanhol	Euronext Paris	Milão	Londres	Total
Empresas de valor	Nº de Observações	554	180	187	1176	719	2732	5548
	%	10%	3%	3%	21%	13%	49%	100%
Todas as empresas	Nº de Observações	4165	1035	1292	6854	2499	12089	27934
	%	15%	4%	5%	25%	9%	43%	100%

Painel B. Número de empresas por ano

Ano	Empresas de valor	%	Todas as empresas	%
1992	70	1,26%	352	1,26%
1993	73	1,32%	368	1,32%
1994	78	1,41%	387	1,39%
1995	83	1,50%	417	1,49%
1996	88	1,59%	440	1,58%
1997	96	1,73%	483	1,73%
1998	109	1,96%	545	1,95%
1999	126	2,27%	630	2,26%
2000	135	2,43%	682	2,44%
2001	166	2,99%	848	3,04%
2002	184	3,32%	936	3,35%
2003	201	3,62%	1012	3,62%
2004	215	3,88%	1090	3,90%
2005	237	4,27%	1211	4,34%
2006	267	4,81%	1354	4,85%
2007	281	5,06%	1425	5,10%
2008	287	5,17%	1450	5,19%
2009	304	5,48%	1539	5,51%
2010	329	5,93%	1658	5,94%
2011	334	6,02%	1676	6,00%
2012	334	6,02%	1678	6,01%
2013	349	6,29%	1747	6,25%
2014	383	6,90%	1914	6,85%
2015	405	7,30%	2022	7,24%
2016	414	7,46%	2070	7,41%
Total	5548	100,00%	27934	100,00%

Anexo 3 – Desempenho da Estratégia F_score ao longo do tempo para todas as empresas de valor

ano	Retornos Ajustados de F_score igual ou superior a 5	N	Retorno Ajustado de F_score inferior a 5	N	Retorno Ajustado Estratégia F_score
1992	0,2029	46	0,1921	24	0,0109
1993	-0,1826	48	-0,1259	25	-0,0567
1994	-0,0652	62	-0,2115	16	0,1463
1995	-0,1335	67	-0,2193	16	0,0859
1996	-0,1427	71	-0,2028	17	0,0601
1997	-0,3121	75	-0,4928	21	0,1807
1998	0,0821	77	0,0777	32	0,0044
1999	0,0678	93	-0,1986	33	0,2664
2000	0,0578	87	-0,3475	48	0,4052
2001	-0,0713	89	-0,2153	77	0,1441
2002	0,2233	105	0,2989	79	-0,0756
2003	-0,0064	128	-0,0339	73	0,0275
2004	-0,0988	151	-0,2399	64	0,1411
2005	-0,0021	155	-0,1327	82	0,1307
2006	-0,2750	178	-0,3957	89	0,1207
2007	-0,1924	177	-0,2699	104	0,0776
2008	-0,0300	161	-0,0246	126	-0,0053
2009	-0,0723	165	-0,1135	139	0,0413
2010	-0,1531	228	-0,3543	101	0,2012
2011	-0,1939	240	-0,4617	94	0,2679
2012	0,1353	213	0,0327	121	0,1026
2013	-0,1564	227	-0,2849	122	0,1285
2014	0,0357	259	-0,1160	124	0,1517
2015	0,0493	246	-0,0522	159	0,1016
2016	0,0191	296	-0,1098	118	0,1289
Média	-0,0486	-	-0,1601	-	0,1115

O retorno ajustado é obtido através da diferença entre o retorno (bruto) de uma empresa e o retorno do respetivo índice do mercado. O retorno bruto corresponde ao lucro resultante da compra de uma ação no final de abril e sua venda no final de abril no ano seguinte. A estratégia F_score adotada na presente tabela consiste na compra de ações com F_score igual ou superior a 5 e venda a descoberto das ações com um indicador global inferior a 5.

Anexo 4– Estratégia F_score aplicada à amostra 90% das empresas de valor no período 1992-2016

	N	Retornos Brutos		Retornos Ajustados	
		Média	Mediana	Média	Mediana
Total de empresas	5006	-0,0052	0,0267	-0,0627	-0,0358
F_score					
0	3	-0,7528	-0,6714	-0,6658	-0,7286
1	41	-0,2371	-0,1507	-0,2598	-0,2283
2	219	-0,0617	-0,0714	-0,1303	-0,1809
3	485	-0,1082	-0,0321	-0,1585	-0,0895
4	895	-0,0383	0,0000	-0,0871	-0,0712
5	1108	0,0110	0,0313	-0,0494	-0,0202
6	986	0,0146	0,0555	-0,0470	-0,0251
7	789	0,0364	0,0567	-0,0281	-0,0243
8	380	0,0624	0,0773	0,0146	0,0109
9	100	0,0705	0,0737	-0,0010	-0,0036
Baixo Score (0,1,2)	263	-0,0969	-0,0922	-0,1566	-0,1953
Elevado Score (8,9)	480	0,0641	0,0772	0,0113	0,0079
Elevado Score – Todas		0,0693	0,0504	0,0740	0,0437
Estratégia F_score		0,1610	0,1694	0,1679	0,2032

A presente tabela apresenta os resultados da aplicação da estratégia baseada no F_score à subamostra composta por 90% das empresas de valor entre 1992 e 2016. Todas as definições da presente tabela encontram-se definidas na Tabela 3.

Anexo 5 - Estratégia F_score aplicada à amostra 75% das empresas de valor no período 1992-2016

	N	Retornos Brutos		Retornos Ajustados	
		Média	Mediana	Média	Mediana
Total de empresas	4171	0,0217	0,0451	-0,0356	-0,0211
F_score					
0	3	-0,7528	-0,6714	-0,6658	-0,7286
1	36	-0,1697	-0,1134	-0,1777	-0,1865
2	159	0,0091	-0,0399	-0,0683	-0,1372
3	355	-0,0547	0,0115	-0,1083	-0,0400
4	712	0,0021	0,0146	-0,0480	-0,0427
5	944	0,0148	0,0312	-0,0420	-0,0170
6	847	0,0336	0,0795	-0,0276	-0,0050
7	683	0,0565	0,0644	-0,0066	-0,0058
8	342	0,0690	0,0773	0,0193	0,0109
9	90	0,1168	0,0877	0,0482	0,0121
Baixo Score (0,1,2)	198	-0,0349	-0,0674	-0,0972	-0,1431
Elevado Score (8,9)	432	0,0789	0,0781	0,0253	0,0118
Elevado Score – Todas		0,0573	0,0331	0,0609	0,0329
Estratégia F_score		0,1139	0,1456	0,1226	0,1549

A presente tabela apresenta os resultados da aplicação da estratégia baseada no F_score à subamostra composta por 75% das empresas de valor entre 1992 e 2016. Todas as definições da presente tabela encontram-se definidas na Tabela 3.

Anexo 6 – Desempenho da Estratégia F_score ao longo do tempo para a amostra 90% das empresas de valor

ano	Retornos Ajustados de F_score igual ou superior a 5	N	Retorno Ajustados de F_score inferior a 5	N	Estratégia F_score
1992	0,1921	42	0,2581	21	-0,0660
1993	-0,1563	44	-0,0957	22	-0,0606
1994	-0,0574	58	-0,1974	13	0,1400
1995	-0,1441	61	-0,2177	14	0,0736
1996	-0,1137	65	-0,1589	15	0,0451
1997	-0,2877	68	-0,5121	19	0,2244
1998	0,1001	72	0,0265	27	0,0736
1999	0,0959	82	-0,2204	32	0,3163
2000	0,0713	81	-0,3001	41	0,3714
2001	-0,0332	82	-0,1756	68	0,1424
2002	0,2229	100	0,3105	66	-0,0876
2003	-0,0026	120	-0,0371	61	0,0345
2004	-0,0818	141	-0,1841	53	0,1023
2005	0,0061	147	-0,0857	67	0,0918
2006	-0,2495	165	-0,3386	76	0,0891
2007	-0,1563	163	-0,2325	90	0,0763
2008	-0,0103	151	0,0284	108	-0,0386
2009	-0,0340	149	-0,0935	125	0,0595
2010	-0,1282	214	-0,2912	83	0,1630
2011	-0,1760	221	-0,4323	80	0,2563
2012	0,1599	192	0,0516	109	0,1083
2013	-0,1403	206	-0,2796	109	0,1393
2014	0,0424	242	-0,0196	103	0,0620
2015	0,0472	227	-0,0213	138	0,0685
2016	0,0277	270	-0,0966	103	0,1243
Média	-0,0322	-	-0,1326	-	0,1004

O retorno ajustado é obtido através da diferença entre o retorno (bruto) de uma empresa e o retorno do respetivo índice do mercado. O retorno bruto corresponde ao lucro resultante da compra de uma ação no final de abril e sua venda no final de abril no ano seguinte. A estratégia F_score adotada na presente tabela consiste na compra de ações com F_score igual ou superior a 5 e venda a descoberto das ações com um indicador global inferior a 5.

Anexo 7 - Desempenho da Estratégia F_score ao longo do tempo para a amostra 75% das empresas de valor

ano	Retornos Ajustados de F_score igual ou superior a 5	N	Retorno Ajustados de F_score inferior a 5	N	Estratégia F_score
1992	0,1921	38	0,2681	15	-0,0759
1993	-0,1411	37	-0,0937	18	-0,0473
1994	-0,0415	50	-0,1185	9	0,0770
1995	-0,1179	53	-0,3362	10	0,2183
1996	-0,1244	53	-0,1359	13	0,0115
1997	-0,3010	56	-0,5049	16	0,2039
1998	0,0769	59	0,0237	23	0,0532
1999	0,1467	72	-0,1076	23	0,2542
2000	0,0856	70	-0,2250	32	0,3106
2001	-0,0421	74	-0,1039	51	0,0618
2002	0,2345	88	0,3522	50	-0,1177
2003	-0,0180	106	0,0318	45	-0,0498
2004	-0,0225	119	-0,1246	43	0,1021
2005	0,0100	128	-0,0170	50	0,0271
2006	-0,2246	147	-0,2471	54	0,0225
2007	-0,1365	140	-0,1650	71	0,0285
2008	0,0079	129	0,1379	87	-0,1300
2009	0,0233	129	-0,0184	99	0,0417
2010	-0,0921	190	-0,2095	57	0,1174
2011	-0,1537	192	-0,3254	59	0,1717
2012	0,1540	163	0,0546	88	0,0994
2013	-0,1174	179	-0,2091	83	0,0917
2014	0,0722	206	-0,0162	82	0,0883
2015	0,0308	199	-0,0712	105	0,1020
2016	0,0340	229	-0,1336	82	0,1676
Média	-0,0186	-	-0,0918	-	0,0732

O retorno ajustado é obtido através da diferença entre o retorno (bruto) de uma empresa e o retorno do respetivo índice do mercado. O retorno bruto corresponde ao lucro resultante da compra de uma ação no final de abril e sua venda no final de abril no ano seguinte. A estratégia F_score adotada na presente tabela consiste na compra de ações com F_score igual ou superior a 5 e venda a descoberto das ações com um indicador global inferior a 5.

Anexo 8 - Estratégia F_score aplicada à amostra composta por todas as empresas no período 1992-2016

	N	Retornos Brutos		Retornos Ajustados	
		Média	Mediana	Média	Mediana
Total de empresas	27934	0,0187	0,0594	-0,0376	-0,0080
F_score					
0	13	-0,1549	-0,0960	-0,1883	-0,1500
1	159	-0,1831	-0,0831	-0,2395	-0,1607
2	750	-0,1096	-0,0905	-0,1743	-0,1651
3	2004	-0,1115	-0,0607	-0,1591	-0,1125
4	4047	-0,0227	0,0298	-0,0769	-0,0379
5	5984	0,0201	0,0622	-0,0328	-0,0026
6	6490	0,0417	0,0800	-0,0175	0,0084
7	5262	0,0526	0,0805	-0,0026	0,0135
8	2602	0,0860	0,1015	0,0205	0,0198
9	623	0,0940	0,0864	0,0326	0,0183
Baixo Score (0,1,2)	922	-0,1229	-0,0866	-0,1858	-0,1627
Elevado Score (8,9)	3225	0,0876	0,0994	0,0229	0,0192
Elevado Score – Todas		0,0689	0,0400	0,0605	0,0272
Estratégia F_score		0,2105	0,1860	0,2087	0,1819

A presente tabela apresenta os resultados da aplicação da estratégia baseada no F_score para a amostra total de empresas, composta por 27934 observações ao longo do período 1992-2016. Todas as definições da presente tabela encontram-se definidas na Tabela 3.

Anexo 9 – Desempenho da Estratégia F_score ao longo do tempo para a amostra composta por todas as empresas

ano	Retornos Ajustados de F_score igual ou superior a 5	N	Retorno Ajustados de F_score inferior a 5	N	Estratégia F_score
1992	0,1102	260	0,1251	92	-0,0148
1993	-0,0891	299	-0,1393	69	0,0502
1994	0,0388	338	-0,1352	49	0,1740
1995	-0,1011	346	-0,1538	71	0,0528
1996	-0,1420	362	-0,2165	78	0,0746
1997	-0,2179	412	-0,2834	71	0,0656
1998	-0,0148	419	0,0258	126	-0,0407
1999	0,0843	481	-0,1174	149	0,2017
2000	0,0978	530	-0,1771	152	0,2749
2001	0,0368	595	-0,1383	253	0,1751
2002	0,1596	705	0,2195	231	-0,0599
2003	0,0274	789	-0,0796	223	0,1070
2004	0,0027	899	-0,0743	191	0,0770
2005	-0,0044	946	-0,1000	265	0,0956
2006	-0,1818	1022	-0,3006	332	0,1188
2007	-0,0622	1060	-0,1922	365	0,1300
2008	0,0351	960	-0,0017	490	0,0368
2009	0,0049	1031	-0,0214	508	0,0263
2010	-0,0329	1344	-0,2022	314	0,1693
2011	-0,1067	1290	-0,2852	386	0,1785
2012	0,0869	1260	0,0417	418	0,0452
2013	-0,1027	1273	-0,2477	474	0,1449
2014	0,0650	1403	-0,1130	511	0,1780
2015	0,0119	1431	-0,1084	591	0,1204
2016	-0,0140	1506	-0,1558	564	0,1417
Média	-0,0123	-	-0,1132	-	0,1009

O retorno ajustado é obtido através da diferença entre o retorno (bruto) de uma empresa e o retorno do respetivo índice do mercado. O retorno bruto corresponde ao lucro resultante da compra de uma ação no final de abril e sua venda no final de abril no ano seguinte. A estratégia F_score adotada na presente tabela consiste na compra de ações com F_score igual ou superior a 5 e venda a descoberto das ações com um indicador global inferior a 5.

Anexo 10 – Estratégia F_score aplicada a amostra composta por 27641 observações

	N	Retornos Brutos		Retornos Ajustados	
		Média	Mediana	Média	Mediana
Total de empresas	27641	0,0195	0,0601	-0,0368	-0,0075
F_score					
0	12	-0,1577	-0,0784	-0,2083	-0,1511
1	156	-0,1718	-0,0569	-0,2287	-0,1527
2	734	-0,1046	-0,0827	-0,1678	-0,1573
3	1977	-0,1109	-0,0604	-0,1580	-0,1123
4	4019	-0,0239	0,0293	-0,0779	-0,0384
5	5934	0,0215	0,0636	-0,0314	-0,0017
6	6428	0,0422	0,0808	-0,0172	0,0087
7	5209	0,0536	0,0806	-0,0017	0,0138
8	2555	0,0875	0,1021	0,0219	0,0213
9	617	0,0954	0,0908	0,0319	0,0182
Baixo Score (0,1,2)	902	-0,1169	-0,0770	-0,1789	-0,1527
Elevado Score (8,9)	3172	0,0890	0,1004	0,0239	0,0201
	Elevado Score – Todas	0,0695	0,0403	0,0607	0,0276
	Estratégia F_score	0,2059	0,1774	0,2027	0,1729

A presente tabela apresenta os resultados da aplicação da estratégia baseada no F_score à amostra composta pelas observações de empresas com dados relativos à capitalização de mercado entre 1992 e 2016. Todas as definições da presente tabela encontram-se definidas na Tabela 3.

Anexo 11 – Desempenho ao longo do tempo da Estratégia N_score

Ano	Retornos Ajustados de N_score igual ou superior a 6	N	Retorno Ajustados de N_score inferior a 6	N	Estratégia N_score
1992	0,1263	244	0,0928	106	0,0335
1993	-0,0957	288	-0,1037	78	0,0080
1994	0,0288	314	-0,0297	70	0,0584
1995	-0,0984	334	-0,1719	80	0,0735
1996	-0,1566	351	-0,1626	86	0,0060
1997	-0,2300	393	-0,2089	87	-0,0211
1998	-0,0165	412	0,0290	133	-0,0455
1999	0,0871	458	-0,1001	169	0,1872
2000	0,0619	509	-0,0385	172	0,1004
2001	0,0313	563	-0,1087	284	0,1400
2002	0,1780	689	0,1655	245	0,0125
2003	0,0266	773	-0,0718	235	0,0984
2004	-0,0062	878	-0,0213	195	0,0151
2005	-0,0115	910	-0,0795	273	0,0680
2006	-0,1979	993	-0,2355	336	0,0376
2007	-0,0810	1009	-0,1331	406	0,0521
2008	0,0330	914	0,0107	532	0,0223
2009	0,0010	975	-0,0076	545	0,0086
2010	-0,0372	1318	-0,1715	329	0,1343
2011	-0,1191	1248	-0,2296	421	0,1105
2012	0,0946	1242	0,0224	425	0,0722
2013	-0,1253	1200	-0,1725	509	0,0472
2014	0,0565	1353	-0,0763	516	0,1328
2015	0,0108	1395	-0,1144	600	0,1251
2016	-0,0208	1439	-0,1232	607	0,1024
Média	-0,0184	-	-0,0816	-	0,0632

O retorno ajustado é obtido através da diferença entre o retorno (bruto) de uma empresa e o retorno do respetivo índice do mercado. O retorno bruto corresponde ao lucro resultante da compra de uma ação no final de abril e sua venda no final de abril no ano seguinte. A estratégia N_score adotada na presente tabela consiste na compra de ações com N_score igual ou superior a 6 e venda a descoberto das ações com um indicador global inferior a 6.

Anexo 12 – Resultados nos anos de treino da otimização da Estratégia F_score através da maximização ou da previsão da rentabilidade (rt)

Otimização linear (rt)

Amostras de Treino	Estratégia F_score	Todas	OLP_rt	RL_rt
Retorno Bruto médio	0,0454	0,0014	0,1186	0,1327
Desvio-Padrão	0,0650	0,0424	0,0469	0,0503
Percentagem de amostras de treino com Retorno positivo	66,67%	33,33%	100%	100%
Percentagem de amostras de treino que apresentam retornos superiores à Estratégia F_score		13,33%	100%	100%

A Estratégia F_score contempla apenas a compra das ações de empresas com um *score* igual a 8 e 9. O Desvio-Padrão mede a volatilidade dos retornos anuais médios obtidos nos anos treino.

Otimização com base em técnicas de *machine learning* (rt)

Amostras de Treino	AR	RF	SVM	Bag	Ens.4	Ens.2
Retorno Bruto médio	0,0835	0,3635	0,0825	0,2643	0,2974	0,2448
Desvio-Padrão	0,0513	0,0658	0,0477	0,0490	0,0456	0,0519
Percentagem de amostras de treino com Retorno positivo	100%	100%	93,33%	100%	100%	100%
Percentagem de amostras de treino que apresentam retornos superiores à Estratégia F_score	86,67%	100%	66,67%	100%	100%	100%

Os acrónimos AR, RF SVM, Bag designam, na respetiva ordem, as técnicas de *machine learning* utilizadas: árvores de regressão, *random forest*, *support vector machine* e *bagging*, enquanto que Ens.4 e Ens.2 representam as metodologias em que se combinam as previsões das 4 técnicas de *machine learning*, através da média, (Ens.4), e através da mediana (Ens.2).

Anexo 13 - Resultados nos anos de treino da otimização da Estratégia F_score através da maximização ou da previsão do equivalente certo (ec)

Otimização linear (ec)				
Amostras de Treino	Estratégia F_score	Todas	OLP_ec	RL_ec
Retorno Bruto médio	0,0454	0,0014	0,1241	0,1181
Desvio-Padrão	0,0650	0,0424	0,0472	0,0373
Percentagem de amostras de treino com Retorno positivo	66,67%	33,33%	100%	100%
Percentagem de amostras de treino que apresentam retornos superiores à Estratégia F_score		13,33%	100%	93,33%

A Estratégia F_score contempla apenas a compra das ações de empresas com um *score* igual a 8 e 9. O Desvio-Padrão mede a volatilidade dos retornos anuais médios obtidos nos anos treino.

Otimização com base em técnicas de <i>machine learning</i> (ec)						
Amostras de Treino	AR	RF	SVM	Bag	Ens.4	Ens.2
Retorno Bruto médio	0,0857	0,3933	0,1124	0,2709	0,2661	0,2361
Desvio-Padrão	0,0545	0,0656	0,0468	0,0404	0,0447	0,0409
Percentagem de amostras de treino com médio Retorno Bruto positivo	100%	100%	100%	100%	100%	100%
Percentagem de amostras de treino que apresentam retornos superiores à Estratégia F_score	80%	100%	93,33%	100%	100%	100%

Os acrónimos AR, RF SVM, Bag designam, na respetiva ordem, as técnicas de *machine learning* utilizadas: árvores de regressão, *random forest*, *support vector machine* e *bagging*, enquanto que Ens.4 e Ens.2 representam as metodologias em que se combinam as previsões das 4 técnicas de *machine learning*, através da média, (Ens.4), e através da mediana (Ens.2).

Anexo 14 - Resultados nos anos de treino da otimização da Estratégia N_score através da maximização ou da previsão da rentabilidade (rt)

Otimização linear (rt)

Amostras de Treino	Todas as empresas	Estratégia F_score	Estratégia N_score	OLP_rt	RL_rt
Retorno Bruto médio	0,0065	0,0843	0,0543	0,0916	0,0742
Desvio-Padrão	0,0408	0,0505	0,0482	0,0500	0,0356
Percentagem de amostras de treino com Retorno médio Bruto positivo	53,33%	100%	100%	100%	100%
Percentagem de amostras de treino que apresentam retornos superiores à Estratégia N_score				93,33%	73,33%

A Estratégia F_score contempla apenas a compra das ações de empresas com um *score* igual a 8 e 9. A Estratégia N_score consiste na compra das ações de empresas com um *score* igual ou superior a 9. O Desvio-Padrão mede a volatilidade dos retornos anuais médios obtidos nos anos treino.

Otimização com base em técnicas de *machine learning* (rt)

Amostras de Treino	AR	RF	SVM	Bag	Ens.4	Ens.2
Retorno Bruto médio	0,0728	0,2388	0,0636	0,1953	0,2028	0,1767
Desvio-Padrão	0,0338	0,0489	0,0349	0,0465	0,0530	0,0480
Percentagem de amostras de treino com Retorno médio Bruto positivo	100%	100%	100%	100%	100%	100%
Percentagem de amostras de treino que apresentam retornos superiores à Estratégia N_score	66,67%	100%	53,33%	100%	100%	100%

Os acrónimos AR, RF SVM, Bag designam, na respetiva ordem, as técnicas de *machine learning* utilizadas: árvores de regressão, *random forest*, *support vector machine* e *bagging*, enquanto que Ens.4 e Ens.2 representam as metodologias em que se combinam as previsões das 4 técnicas de *machine learning*, através da média, (Ens.4), e através da mediana (Ens.2).

Anexo 15 - Resultados nos anos de treino da otimização da Estratégia N_score através da maximização ou da previsão do equivalente certo (ec)

Otimização linear (ec)					
Amostras de Treino	Todas as empresas	Estratégia F_score	Estratégia N_score	OLP_ec	RL_ec
Retorno Bruto médio	0,0065	0,0843	0,0543	0,1022	0,0839
Desvio-Padrão	0,0408	0,0505	0,0482	0,0491	0,0487
Percentagem de amostras de treino com Retorno médio Bruto positivo	53,33%	100%	100%	100%	100%
Percentagem de amostras de treino que apresentam retornos superiores à Estratégia N_score				100%	100%

A Estratégia F_score contempla apenas a compra das ações de empresas com um *score* igual a 8 e 9. A Estratégia N_score consiste na compra das ações de empresas com um *score* igual ou superior a 9. O Desvio-Padrão mede a volatilidade dos retornos anuais médios obtidos nos anos treino.

Otimização com base em técnicas de <i>machine learning</i> (ec)						
Amostras de Treino	AR	RF	SVM	Bag	Ens.4	Ens.2
Retorno Bruto médio	0,0860	0,2597	0,0832	0,1972	0,1855	0,1743
Desvio-Padrão	0,0509	0,0493	0,0488	0,0425	0,0510	0,0439
Percentagem de amostras de treino com Retorno médio Bruto positivo	100%	100%	100%	100%	100%	100%
Percentagem de amostras de treino que apresentam retornos superiores à Estratégia N_score	86,67%	100%	86,67%	100%	100%	100%

Os acrónimos AR, RF SVM, Bag designam, na respetiva ordem, as técnicas de *machine learning* utilizadas: árvores de regressão, *random forest*, *support vector machine* e *bagging*, enquanto que Ens.4 e Ens.2 representam as metodologias em que se combinam as previsões das 4 técnicas de *machine learning*, através da média, (Ens.4), e através da mediana (Ens.2).