

Szkoła Główna Handlowa w Warszawie
Kolegium Analiz Ekonomicznych

**Przewidywanie wydarzeń w szeregach
czasowych za pomocą metod uczenia
maszynowego a Hipoteza Rynków
Efektywnych**

Adam Chudziak

Praca doktorska napisana
pod kierunkiem naukowym
dr. hab. Michała Ramszy, prof. SGH

Warszawa, wrzesień 2020 r.

1 Motywacja i hipotezy badawcze

Hipoteza Rynków Efektywnych (ang. *Efficient Markets Hypothesis*, EMH) jest przedmiotem zainteresowania badaczy od czasu jej sformułowania przez Fama (Fama 1970). Niesłabnące zainteresowanie kwestią, czy możliwe jest przewidywanie zmian cen i wykorzystanie tej wiedzy aby osiągnąć zysk jest podsycane przez potrzeby inwestorów na rynkach finansowych. Dlatego też nie dziwi fakt, że nowe narzędzia prognostyczne są szybko wykorzystywane do prognozowania rynków finansowych. Nie inaczej było w przypadku metod sztucznej inteligencji i eksploracji danych, które stosowane są coraz powszechniej w branży finansowej. Jednakże analiza, której poddawane są nowe metody i algorytmy często skupia się na bardziej technicznych aspektach, takich jak szybkość działania czy wykorzystanie danych.

Celem niniejszej pracy jest sprawdzenie przydatności wybranych metod uczenia maszynowego i eksploracji danych do prognozowania finansowych szeregów czasowych. Główna hipoteza stwierdza, że niektóre rynki nie spełniają Hipotezy Rynków Efektywnych.

H Niektóre rynki finansowe nie spełniają Hipotezy Rynków Efektywnych.

W badaniu tej hipotezy wykorzystujemy dwie hipotezy pomocnicze

HP1 Techniki uczenia się maszynowego pomagają w prognozowaniu zdarzeń w finansowych szeregach czasowych, ujawniając niewykryte nieefektywności rynku.

HP2 Transformacja finansowych szeregów czasowych za pomocą reprezentacji wykorzystujących punkty istotne zwiększa jakość prognoz.

Weryfikacja EMH sprawia trudności z powodu niejednoznacznej definicji hipotezy (Thaler 2009; Guerrien i Gun 2011; Charron 2017). Według Thalera, hipotezę można rozumieć na dwa sposoby. Pierwszym z nich jest „cena jest właściwa”—twierdzenie, że ceny odzwierciedlają *prawdziwą* wartość papieru wartościowego. Warto osobnej dyskusji jest pytanie, czy taka wartość istnieje,

czym ona jest, i czy ceny ją odzwierciedlają, jednak jest to problem, którego nie dotyczy ta praca doktorska. Drugi sposób zrozumienia hipotezy mówi, że na rynku „nie ma darmowych lunchy” (zwany jest też „techniczną efektywnością”). Stwierdza on, że udana spekulacja na rynku za pomocą prognoz cen jest niemożliwa. Analiza przedstawiona w tej pracy koncentruje się na technicznej efektywności w dłuższej perspektywie. Krótkoterminowe naruszenia EMH, w szczególności dotyczące np. high frequency tradingu, są poza zakresem prezentowanej analizy.

Powszechne jest, że analizy nowych algorytmów koncentrują się na innych charakterystykach niż te istotne dla ekonomistów. Często nacisk położony jest na szybkość działania algorytmu, zużycie pamięci i jedynie dość podstawowe miary sprawdzające czy algorytm wykonuje powierzone mu zadanie. Celem pracy jest analiza opłacalności stosowania modeli predykcyjnych łączących reprezentacje punktów istotnych i sztuczne sieci neuronowe dla miesięcznych zmian cen akcji, które wykorzystują jako zmienne prognostyczne tylko przeszłe zmiany cen. Analiza wykorzystuje długie szeregi czasowe, co jest rzadkością w literaturze przedmiotu.

2 Struktura pracy

W skład pracy oprócz rozdziału wprowadzającego i konkluzji wchodzi trzy rozdziały oraz dwa załączniki.

Rozdział drugi przedstawia Hipotezę Rynków Efektywnych i ukazuje różne podejścia do jej oceny i różne metody jej krytyki. Przedstawiono zagadnienia związane z niejasnym sformułowaniem hipotezy, w tym problem hipotezy łączonej oraz kwestię rozróżnienia „technicznej” i „fundamentalnej” efektywności (Charron 2017). Przedyskutowano różne metody empirycznej weryfikacji hipotezy.

Rozdział trzeci poświęcony jest metodologii pracy. Testowane w rozprawie modele prognostyczne wykorzystują reprezentacje punktów istotnych (LR)

do wstępnego przetwarzania szeregów czasowych, a także sztucznych sieci neuronowych (ANN) do generowania prognoz. Omówiono zarówno algorytmy reprezentacji, jak i sztucznych sieci neuronowych. Konkretnie algorytmy są dobrze ugruntowane w literaturze, jednak zastosowana metodologia jest nowatorska. W związku z tym, oprócz prezentacji metod, omówiono również dokonane meta-wybory.

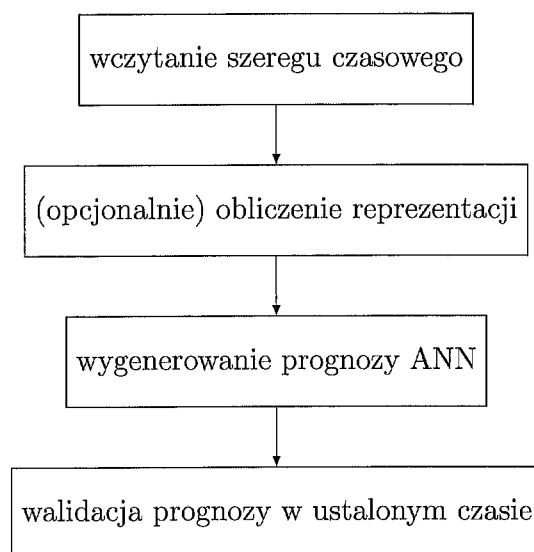
W rozdziale czwartym przedstawiono główne wyniki pracy i przedyskutowano je w kontekście teoretycznym. Rozdział jest podzielony na trzy części. W części pierwszej zaprezentowano ocenę reprezentacji punktów istotnych, omawiając ich błędy, kompresję, i zdolność do zachowania informacji w sygnale, co analizowano metodami spektralnymi. W części drugiej zaprezentowano rezultaty otrzymane z modeli prognostycznych, w tym dokładność kierunkową, test opłacalności strategii handlowych opartych na prognozach, oraz statystyczne testy zdolności predykcyjnych. Odnotowano i przeanalizowano również różnice pomiędzy wynikami w okresach wzrostu gospodarczego i recesji, datowanymi według danych z NBER. Rozdział kończy się dyskusją rezultatów.

3 Metodologia

Jak wspomniano wcześniej, w rozprawie badamy techniczną efektywność cen akcji w długim terminie. Modele przewidują kierunek zmian cen akcji na następny miesiąc, na podstawie kroczącego okna szeregu cen. Dane składają się z cen akcji NYSE w latach 1970-2015.

Istnieje szereg badań analizujących popularne wśród praktyków metody przewidywania, takich jak analiza techniczna wykorzystująca przeszłe ceny instrumentów finansowych (Lo, Mamaysky i Wang 2000; Farias Nazário i in. 2017). Chociaż w niniejszej pracy nie testujemy metod analizy technicznej jako takiej, stosowane modele prognostyczne ANN opierają się na podobnej koncepcji rozpoznawania wzorców w danych. Surowe finansowe szeregi cza-

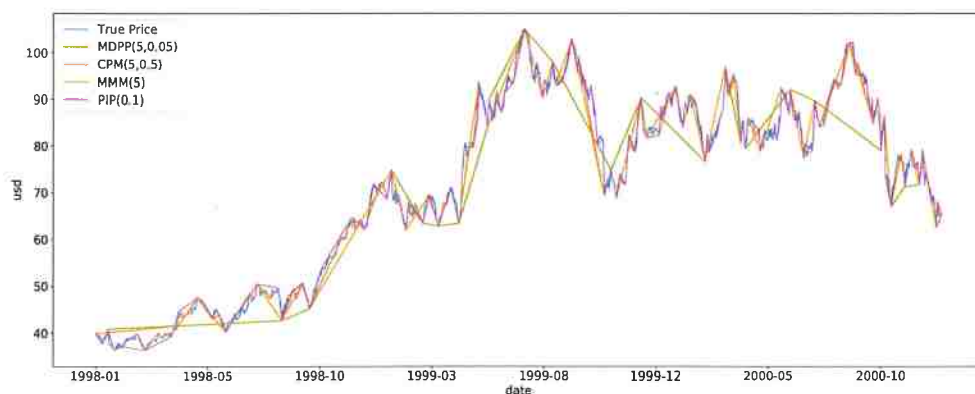
sowe często wykazują wysoką zmienność, która może negatywnie wpływać na jakość prognozy. W związku z tym zastosowano opcjonalny etap wstępnego przetwarzania szeregu czasowego cen polegający na przekształcaniu go za pomocą reprezentacji punktów istotnych (LR) (patrz rys. 2).



Rysunek 1: Schemat procesu predykcyjnego.

Reprezentacje punktów istotnych można zdefiniować jako algorytmy eksploatacji danych szeregów czasowych, które konstruują kawałkami liniową reprezentację szeregu czasowego w oparciu o „istotność” poszczególnych punktów w szeregu. Korzenie tej idei wywodzą się z psychologii, gdyż zaobserwowano, że ludzka pamięć organizuje się wokół zdarzeń o szczególnym znaczeniu (Humphreys, Wiles i Dennis 1994; Glenberg 1997). Algorytmy LR operacjonalizują tę koncepcję na różne sposoby. Albo stosując kryteria decydujące o tym, czy każdy punkt jest istotny, czy też nie, albo tworząc globalny porządek istotności punktów. Rezultatem jest podzbiór pierwotnych obserwacji, pomiędzy którymi wykonywana jest interpolacja liniowa. Rysunek 2 przedstawia przykładowe reprezentacje obliczone dla IBM. W pracy analizujemy cztery algorytmy: Landmark Model (MDPP, Perng i in. 2000), Major Maxima and Minima algorithm (MMM, Fink, Pratt i Gandhi 2003), Perceptually Impor-

tant Points algorithm (PIP, Fu i in. 2008), oraz Critical Point Model (CPM, Bao 2008).



Rysunek 2: Ceny skorygowane IBM przetworzone za pomocą reprezentacji punktów istotnych: Landmark Model (MDPP), Critical Point Model (CPM), Major Maxima and Minima algorithm (MMM) i Perceptually Important Points algorithm (PIP). Parametry reprezentacji znajdują się w nawiasach.

Sztuczne sieci neuronowe odniosły duży sukces w różnych dziedzinach. Są one szczególnie użyteczne w zadaniach wymagających rozpoznawania wzorców, takich jak widzenie komputerowe. Zdolność rozpoznawania wzorców sprawia, że są one dobrym wyborem do testowania przewidywalności w oparciu o ceny akcji. Prognozowanie rynku akcji z wykorzystaniem sieci neuronowych cieszy się dużym zainteresowaniem, np. (White 1988; Chong, Han i Park 2017; Sezer, Ozbayoglu i Dogdu 2017). Wiele badań analizuje jednak krótkie okresy i selektywnie dobiera prezentowane wyniki. W niniejszej pracy przedstawiono analizę przeprowadzoną dla długiego okresu.

W prognozowaniu wykorzystujemy trzy rodzaje topologii sieci neuronowych: wielowarstwowy perceptron (Rosenblatt 1958), proste rekurencyjne sieci neuronowe i sieci LSTM (Hochreiter i Schmidhuber 1997). Możliwości konstrukcji sieci są praktycznie nieograniczone. Nawet w obrębie wybranej topologii należy wziąć pod uwagę kilka meta-parametrów, takich jak liczba warstw i liczba neuronów w każdej warstwie. Topologie wybrane w tej pracy są jednymi z najbardziej popularnych w praktyce ze względu na ich udane zasto-

sowanie w wielu aplikacjach. W pracy testowane są różne kombinacje sztucznych sieci neuronowych i reprezentacji punktów istotnych.

4 Rezultaty

W pracy analizowano czy model prognostyczny łączący reprezentacje punktów istotnych ze sztucznymi sieciami neuronowymi może być z powodzeniem wykorzystywany do prognozowania ruchów cen na giełdzie. Otrzymane rezultaty są niejednoznaczne. Znaleźliśmy ograniczoną, zmienną w czasie przewidywalność, którą jednak uważamy za trudną do wykorzystania dla osiągnięcia zysku, ze względu na problemy z wyborem modelu.

Są okresy, w których przewidywalność wydaje się być wyższa niż w innych okresach. Sporadyczne wzrosty dokładności kierunkowej prognoz mogą stanowić okna możliwości dla handlujących akcjami. Jednak niejasne jest jaką metodą można identyfikować okresy podwyższonej przewidywalności, a także modele, które pozwalają na ich wykorzystanie. Dalsze badania w tym kierunku są już prowadzone, patrz np. (Harvey i in. 2018).

Istnieją również różnice w jakości prognoz pomiędzy testowanymi szeregami czasowymi. W przypadku niektórych spółek, żaden z modeli prognostycznych nie był w stanie osiągnąć zadowalających wyników. Wynik ten wskazuje na różnice w efektywności na poziomie mikro pomiędzy aktywami z tego samego rynku.

Jednym z zastosowanych przez nas benchmarków był zwrot z inwestycji w ramach strategii inwestycyjnej kup-i-trzymaj. Co do zasady, proponowane podejście nie pozwala na osiągnięcie lepszych wyników. Z drugiej strony, handlowanie z wykorzystaniem testowanych modeli skutkuje zmniejszeniem ryzyka posiadanych portfeli, mierzonego jako ich zmienność. Jest to zgodne z teorią, np. (Merton 1973).

Stwierdziliśmy, że w okresach recesji modele osiągają lepsze wyniki w porównaniu z odpowiednimi benchmarkami, niż w okresach wzrostu. Należy jednak

pamiętać, że ponieważ gospodarka amerykańska przez większość ostatniego półwiecza rozwijała się, próby są nie zrównoważone. Mamy o wiele więcej danych dotyczących ekspansji, a zatem wynik ten należy traktować z ostrożnością.

Reprezentacje punktów istotnych na ogół spełniają swoje zadanie redukcji szumu w szeregu czasowym, przy jednoczesnym zachowaniu ich kształtu. Niestety, uzyskane rezultaty zależą wyraźnie od parametrów, a optymalny wybór tych ostatnich zależy od szeregu. Z analizowanych reprezentacji odrzucono jedną, ponieważ w testach stwierdzono, że nie zachowuje ona częstotliwości, z jaką teoretycznie powinna zachować przy ustalonych parametrach.

Dowody dotyczące wpływu przetwarzania wstępnego z wykorzystaniem reprezentacji punktów istotnych na prognozy są mieszane. W zależności od spółki, prognozy modeli korzystających z przetwarzania wstępnego odbiegały w różnym stopniu od prognoz modelu, który z nich nie korzystał. Różnice pomiędzy prognozami zostały przetestowane przy użyciu testu Diebolda-Mariano i nie wszystkie były statystycznie istotne (Diebold i Mariano 1995). W przypadku niektórych spółek, reprezentacje zwiększyły jakość prognoz. Jednakże żadna reprezentacja nie została jednoznacznie zidentyfikowana jako najlepsza.

Wyniki testów statystycznych zdolności predykcyjnych, zwłaszcza testu Pesarana-Timmermanna (Pesaran i Timmermann 1992), są mieszane. Statystycznie istotna zdolność predykcyjna, połączona z pokonaniem benchmarkowego zwrotu z inwestycji przez symulowany portfel wystąpiła w kilku przypadkach, były to jednak raczej wyjątki niż typowe zachowanie. Nie jest jasne, w jaki sposób można było wcześniej zidentyfikować te modele, ponieważ konfiguracja nie pozwalała na uprzednią walidację modeli.

Dodatkowa trudność wynika z różnic pomiędzy szeregami czasowymi. Zmieniająca się w czasie jakość prognoz sugeruje, że przewidywalność nie jest ich stałą cechą. Ta obserwacja jest zgodna z poglądem przedstawionym w (Timmermann i Granger 2004), gdzie autorzy sugerują, że modele predykcyjne

prawdopodobnie zyskują i tracą moc predykcyjną, nie osiągając stałego, równego poziomu jakości prognoz.

Jakość działania modeli może przynajmniej częściowo być efektem długości horyzontu prognozy. W oknie szeregu cen, na podstawie którego powstaje prognoza, może nie być wystarczającej ilości informacji, aby otrzymać prognozę na miesiąc wprzód o satysfakcjonującej dokładności. Jest to zgodne z poglądem, że okna przewidywalności zmniejszają się (Schulmeister 2009).

Znaleźliśmy ograniczoną przewidywalność, którą jednak trudno wykorzystać tak, by osiągnąć zysk. Mimo, że niektóre modele osiągnęły lepsze rezultaty niż benchmarki, problemem proponowanego podejścia pozostaje wybór modelu. W związku z tym należy stwierdzić, że nie znaleziono wystarczających dowodów na poparcie hipotez badawczych.

Można nakreślić kilka możliwości rozwinięcia badań przedstawionych w rozprawie. Szybkie tempo rozwoju metod uczenia maszynowego sprawia, że udoskonalenie modeli jest jedną z możliwości. Ciekawą perspektywą byłoby na przykład wykorzystanie przetwarzania wstępnego na wzór przetwarzania obrazów, lub wykorzystanie mechanizmu uwagi w uczeniu się statystycznym. Różnice w jakości prognoz pomiędzy spółkami sprawiają, że porównanie wyników z innymi, możliwe, że mniej efektywnymi, rozwijającymi się rynkami byłoby interesujące.

5 Znaczenie uzyskanych wyników

Rozprawa podejmuje aktualny problem, czy postęp technologiczny wpływa na efektywność rynku. Wprowadzono nową metodologię łączącą reprezentacje punktów istotnych i sztuczne sieci neuronowe. Przeanalizowano zachowanie LR i pokazano ich wrażliwość na wybór parametrów. Systematycznie przetestowano możliwości prognozowania ANN. Stwierdzona przewidywalność nie oznacza, zdaniem autora, nieefektywności. Uzyskane wyniki mają istotne konsekwencje dla praktyków handlujących na rynkach finansowych.

Z ich punktu widzenia praca ta pokazuje, że w długiej perspektywie czasowej testowane metody, w oderwaniu od innych narzędzi, nie umożliwiają systematycznego osiągnięcia zysku. Kluczowym problemem staje się dobór właściwej metody predykcji dla właściwego dobra. Ponadto, w świetle tych ustaleń, możliwość generalizacji niektórych wyników publikowanych w literaturze dotyczących wykorzystania sieci neuronowych w finansach staje się wątpliwa.

Bibliografia

- Bao, Depei (2008). „A generalized model for financial time series representation and prediction”. W: *Applied Intelligence* 29.1, s. 1–11. ISSN: 0924-669X, 1573-7497.
- Charron, Jacques-Olivier (2017). „Inefficient Debate. The EMH, the “Remarkable Error” and a Question of Point of View”. W: *Accounting, Economics, and Law: A Convivium* 7.3. ISSN: 2152-2820.
- Chong, Eunsuk, Han, Chulwoo i Park, Frank C. (2017). „Deep learning networks for stock market analysis and prediction: Methodology, data representations, and case studies”. W: *Expert Systems with Applications* 83, s. 187–205. ISSN: 0957-4174.
- Diebold, Francis X. i Mariano, Roberto S. (1995). „Comparing Predictive Accuracy”. W: *Journal of Business & Economic Statistics* 13.3, s. 253–263. ISSN: 0735-0015.
- Fama, Eugene F. (1970). „Efficient Capital Markets: A Review of Theory and Empirical Work”. W: *The Journal of Finance* 25.2, s. 383–417. ISSN: 00221082.
- Farias Nazário, Rodolfo Toríbio i in. (2017). „A literature review of technical analysis on stock markets”. W: *The Quarterly Review of Economics and Finance* 66, s. 115–126. ISSN: 1062-9769.
- Fink, E., Pratt, K.B. i Gandhi, H.S. (2003). „Indexing of time series by major minima and maxima”. W: *SMC'03 Conference Proceedings. 2003 IEEE International Conference on Systems, Man and Cybernetics. Conference Theme - System Security and Assurance (Cat. No.03CH37483)*. SMC '03 2003 IEEE International Conference on Systems, Man and Cybernetics. T. 3. Washington, DC, USA: IEEE, s. 2332–2335. ISBN: 978-0-7803-7952-7.
- Fu, Tak-chung i in. (2008). „Representing financial time series based on data point importance”. W: *Engineering Applications of Artificial Intelligence* 21.2, s. 277–300. ISSN: 09521976.

- Glenberg, Arthur M. (1997). „What memory is for”. W: *Behavioral and Brain Sciences* 20.1, s. 1–19. ISSN: 0140-525X, 1469-1825.
- Guerrien, Bernard i Gun, Ozgur (2011). „Efficient Market Hypothesis: What are we talking about?” W: *Real-World Economics Review* 56, s. 12.
- Harvey, David i in. (2018). *Detecting Regimes of Predictability in the U.S. Equity Premium*. Essex Finance Centre Working Paper. University of Essex, Essex Business School.
- Hochreiter, Sepp i Schmidhuber, Jürgen (1997). „Long Short-Term Memory”. W: *Neural Computation* 9.8, s. 1735–1780. ISSN: 0899-7667, 1530-888X.
- Humphreys, Michael S., Wiles, Janet i Dennis, Simon (1994). „Toward a theory of human memory: Data structures and access processes”. W: *Behavioral and Brain Sciences* 17.4, s. 655–667. ISSN: 0140-525X, 1469-1825.
- Lo, Andrew W., Mamaysky, Harry i Wang, Jiang (2000). „Foundations of Technical Analysis: Computational Algorithms, Statistical Inference, and Empirical Implementation”. W: *The Journal of Finance* 55.4, s. 1705–1765. ISSN: 1540-6261.
- Merton, Robert C. (1973). „An Intertemporal Capital Asset Pricing Model”. W: *Econometrica* 41.5, s. 867–887. ISSN: 0012-9682.
- Perng, C.-S. i in. (2000). „Landmarks: a new model for similarity-based pattern querying in time series databases”. W: *Data Engineering, 2000. Proceedings. 16th International Conference on*. IEEE, s. 33–42.
- Pesaran, M. Hashem i Timmermann, Allan (1992). „A Simple Nonparametric Test of Predictive Performance”. W: *Journal of Business & Economic Statistics* 10.4, s. 461–465. ISSN: 0735-0015, 1537-2707.
- Rosenblatt, F. (1958). „The perceptron: A probabilistic model for information storage and organization in the brain.” W: *Psychological Review* 65.6, s. 386–408. ISSN: 1939-1471, 0033-295X.
- Schulmeister, Stephan (2009). „Profitability of technical stock trading: Has it moved from daily to intraday data?” W: *Review of Financial Economics* 18.4, s. 190–201. ISSN: 1058-3300.

- Sezer, Omer Berat, Ozbayoglu, Murat i Dogdu, Erdogan (2017). „A Deep Neural-Network Based Stock Trading System Based on Evolutionary Optimized Technical Analysis Parameters”. W: *Procedia Computer Science* 114, s. 473–480. ISSN: 18770509.
- Thaler, Richard (2009). „Markets can be wrong and the price is not always right”. W: *Financial Times*, s. 4.
- Timmermann, Allan i Granger, Clive W. J. (2004). „Efficient market hypothesis and forecasting”. W: *International Journal of Forecasting* 20.1, s. 15–27. ISSN: 0169-2070.
- White, Halbert (1988). „Economic prediction using neural networks: the case of IBM daily stock returns”. W: *IEEE International Conference on Neural Networks*. Proceedings of 1993 IEEE International Conference on Neural Networks (ICNN '93). San Diego, CA, USA: IEEE, 451–458 vol.2. ISBN: 978-0-7803-0999-9.