



Instytut Badań Systemowych  
Polskiej Akademii Nauk

---

*Seria:*

**BADANIA SYSTEMOWE**

**TOM 65**

Redaktor Naukowy

Prof. dr hab. inż. Jakub Gutenbaum

## **Rada Redakcyjna:**

1. Prof. dr hab. inż. Olgierd Hryniewicz – *przewodniczący*
2. Prof. dr hab. inż. Jakub Gutenbaum – *redaktor naukowy*
3. Prof. dr hab. inż. Janusz Kacprzyk
4. Prof. dr hab. inż. Tadeusz Kaczorek
5. Prof. dr hab. inż. Roman Kulikowski
6. Doc. dr hab. inż. Marek Libura
7. Prof. dr hab. inż. Krzysztof Malinowski
8. Prof. dr hab. inż. Zbigniew Nahorski
9. Dr. hab. inż. Marek Niezgódka, prof. UW
10. Prof. dr hab. inż. Roman Słowiński
11. Doc. dr hab. inż. Jan Studziński
12. Prof. dr hab. inż. Stanisław Walukiewicz
13. Prof. dr hab. inż. Andrzej Weryński
14. Doc. dr hab. inż. Antoni Żochowski



Instytut Badań Systemowych  
Polskiej Akademii Nauk

**Antoni Wiliński**

**GMDH – metoda grupowania argumentów  
w zadaniach zautomatyzowanej predykcji  
zachowań rynków finansowych**

**Warszawa - Szczecin**

Copyright © by Instytut Badań Systemowych PAN  
Warszawa 2009

*Dr hab. inż. Antoni Wiliński – prof. nadzw.  
Zachodniopomorski Uniwersytet Technologiczny w Szczecinie  
Wydział Informatyki  
Ul. Żołnierska 49, Szczecin 71-210  
Tel. 091- 449 5660 ; fax 091-449 5540*

*Autor jest profesorem Zachodniopomorskiego Uniwersytetu Technologicznego, poprzednio Politechniki Szczecińskiej. Habilitację uzyskał na Politechnice Warszawskiej w zakresie automatyki i robotyki. W obecnej kadencji jest dziekanem Wydziału Informatyki ZUT. Od lat zajmuje się problematyką inteligencji obliczeniowej i rozpoznawania wzorców w szeregach czasowych w celach predykcyjnych lub modelowania rzeczywistości. W Katedrze Systemów Multimedialnych, której jest kierownikiem, prowadzone są badania nad webowym systemem handlu automatycznego. Rozważane są rozmaite strategie analizy technicznej specyficzne dla decyzji podejmowanych z wysoką częstotliwością przez infoboty. Wśród tych strategii istotną rolę odgrywają omawiane tu algorytmy oparte na GMDH (Group Method of Data Handling). Praca przeznaczona jest raczej dla specjalistów będących zwolennikami analizy technicznej.*

### **Recenzenci:**

Prof. dr hab. inż. Zbigniew Banaszak  
Doc. dr hab. inż. Maciej Krawczak

### **Redaktor techniczny:**

Dr inż. Anna Samborska-Owczarek

Powyższej książki w całości lub części nie wolno powielać ani przekazywać w żaden sposób, nawet za pomocą nośników mechanicznych i elektronicznych (np. zapis magnetyczny), w tym też umieszczać ani rozpowszechniać w postaci cyfrowej zarówno w Internecie, jak i w sieciach lokalnych, bez uzyskania pisemnej zgody posiadacza praw autorskich.

ISBN 9788389475237  
**ISSN 0208-8029**



*To, co znane, nie jest jeszcze przez to, że jest znane, czymś poznanym*  
Fryderyk Hegel (1770-1831)

## Wprowadzenie

Zasadniczym celem niniejszej monografii jest potwierdzenie tezy autora, że fuzja matematyki i technologii egzemplifikowana powiązaniem metod predykcji ze strategią inwestycyjną opartą na mechanizmach technologicznych platform brokerskich daje nową synergetyczną wartość. Wartości tej nie daje się uzyskać odrębnie. Przy niczym nieograniczonej fascynacji handlem<sup>1</sup> na platformach internetowych [10, 66, 81, 85, 86, 94], często dziś spotykanej wśród licznych, najczęściej młodych przedstawicieli przyszłego społeczeństwa informacyjnego, nie da się, bez odpowiedniej wiedzy, osiągnąć sukcesu. Podobnie sama wyłącznie teoretyczna wiedza, bez umiejętności jej zastosowania, pozostaje metafizyczną wartością, rzadko satysfakcjonująca twórcę. Trawestując Freddie Mercurę<sup>2</sup> można powiedzieć *The sale must go on* i coś energetycznego jest w tym aforyzmie. Jest w nim imperatyw wdrożenia odkrycia matematycznego, rozwiązania znalezionego w przestrzeni przeszłej. To zdarzenie rzadkie w świecie nauki. Z kolei szaleństwo nieprzygotowanego żywiołowego uczestnictwa w giełdowych<sup>3</sup> i foreksowych<sup>4</sup> transakcjach to zdarzenia częstotliwe, niejednokrotnie z odcieniem chorobliwego hazardu i najczęściej bez happy-endu.

---

<sup>1</sup> Przez handel na platformach internetowych rozumie się najczęściej częstotliwe zawieranie transakcji na odległość polegające na otwieraniu pozycji na rynkach udostępnianych przez daną platformę (najczęściej na rynkach walutowych). Takie zachowanie inwestora nazywane jest często, bez tłumaczenia, tradingiem.

<sup>2</sup> Freddie Mercury – słynny brytyjski piosenkarz ub. wieku, którego jednym z największych przebojów był utwór „The show must go on”.

<sup>3</sup> W pracy rozpatrywane będą predykcje walorów (akcji) giełdowych i kursów walut.

<sup>4</sup> Forex – to międzynarodowy rozproszony rynek walutowy udostępniany właśnie poprzez tysiące internetowych platform brokerskich.

Czego więc syntezą jest ostatecznie sukces na rynku<sup>5</sup> finansowym? Jest zapewne sumą wiedzy (zastosowań matematyki, statystyki, udanych modeli predykcyjnych, znajomości realiów ekonomicznych), umiejętności (znajomości technologii brokerskich, przewidywania reakcji rynków na podstawie symptomów politycznych, ekonomicznych czy społecznych, stosowania oceny percepcyjnej zdarzeń) i możliwości ciągłego uczestnictwa w grze rynkowej (inaczej mówiąc wolnego czasu po podjęciu decyzji inwestycyjnej lub użycia systemu informatycznego do monitoringu rynku i podejmowania automatycznych decyzji wg wcześniej przetestowanej strategii).

W tym aspekcie cytowane dalej motto Kanta o znaczeniu matematyki w obiektywizacji reguł uogólniających opis rzeczywistości, tworzących model wiedzy o niej, jest wyrazem przekonań autora. Matematyki nie należy jednak w tym cytacie zawęzać do jej akademickiej definicji. Dziś Kantowi chodziłoby zapewne także o szersze rozumienie tego konceptu – uwzględniającego np. także dokonania metod sztucznej inteligencji, teorii chaosu czy mechaniki kwantowej. Szerzej, zgadzając się z Włodzisławem Duchem [15], należy raczej posługiwać się tu pojęciem inteligencji obliczeniowej, która jest pojęciem szerszym niż sztuczna inteligencja. Metoda staje się bowiem wartościowa poprzez obliczenia. Przed ich przeprowadzeniem często nie ma podstaw do uznania koncepcji za metodę.

Nie odzegnując się od niczego, autor, odwołując się do klasyfikacji postaw w nauce wg Andrzeja Wierzbickiego [124], skłonny jest zaliczyć prezentowane tu poglądy do skrajnie pragmatycznych, typowych dla przedstawicieli bardziej nauk technicznych niż nauk ścisłych.

Dzisiejsze postawy inwestorów wobec zadania predykcji na rynkach finansowych w najbardziej generalny sposób różnicuje podejście zaproponowane przez Eugena Famę [20], który wprowadzając kilkadziesiąt lat temu pojęcie rynku efektywnego podzielił inwestorów na zwolenników analizy fundamentalnej, analizy technicznej, zwolenników równoczesnego wykorzystywania metod z obu obszarów wiedzy oraz sceptyków przekonanych o nieprzewidywalności absolutnie niestacjonarnych zachowań rynków. Współcześnie, uważa się (np. Hia Jong Teoh [33], Goefrey Friesen[24], Bill Cai [8]), że klasyfikacja przekonań o możliwości przewidywania rynków winna przebiegać jeszcze nieco inaczej. Powinna przebiegać mianowicie wg podziału na trzy kategorie postaw -

---

<sup>5</sup> Rynek to umowna nazwa dziedziny, w której obserwowane są zmiany ceny wybranego instrumentu finansowego np. kursu waluty, ceny akcji na giełdzie, ceny towaru itp.

że nic nie da się przewidzieć, że tylko analiza fundamentalna jest zasadna oraz, że tylko wyrafinowane metody współczesnej analizy technicznej (wykorzystujące sztuczne sieci neuronowe, rozpoznawania wzorców, algorytmy genetyczne itp.) mają sens.

Niewątpliwie do zwolenników tego ostatniego podejścia zaliczyć można większość noblistów w dziedzinie ekonomii (tych, związanych z ekonometrią, ekonomią i inżynierią finansową) ostatnich kilkunastu lat takich jak Myron Scholes, Robert Merton, Robert Engle, Harry Markowitz, Joseph Stiglitz, William Sharpe. Wszyscy oni uważali, że jest sens tworzenia złożonych modeli odwzorowania gospodarczej rzeczywistości i pomimo świadomie popełnianych błędów (na poziomie uproszczeń w założeniach koniecznych do zbudowania modelu) przyczynili się do istotnego poszerzenia wiedzy matematycznej o modelowaniu rynków.

Częste są także opinie, że prawdopodobnie najlepsze wyniki predykcyjne na poziomie zapewniającym ich praktyczną użyteczność inwestycyjną uzyskuje się poprzez kojarzenie elementów analizy technicznej i fundamentalnej. To oczywiście pewna odmiana przytoczonego wyżej poglądu o celowości poszukiwania skutecznego rozwiązania w gąszczu współczesnej inteligencji obliczeniowej. Metametodą jest tu hybrydyzacja metod o różnym charakterze i różnych technik wykonania decyzji rekomendowanych przez te modele. Bardzo dynamicznie rozwijają się ostatnio w rozpatrywanym obszarze badania chińskich analityków np. Sheng-Hsun Hsu [107] i Zhang Yudong [133] łącząc modelowanie za pomocą sztucznych sieci neuronowych z teorią zbiorów przybliżonych [88, 89] uważają, że to najbardziej perspektywiczny kierunek opanowywania niestacjonarnych procesów, a Hia Jong Teoh [33] przedstawia metodę łączącą sieci neuronowe z GMDH (Group Method of Data Handling). S. Dehuri i S. Cho [14] łączą techniki regresyjne z optymalizacją rojem cząstek. Optymalizacja rojem jest najczęściej zabiegiem pomocniczym dla określania najlepszych parametrów w wielowymiarowej przestrzeni predyktorów (indykatorów, zmiennych objaśniających, składowych portfela itp.).

Interesującym aspektem metod stosowanych dla predykcji rynków jest także umowny podział na dwie duże grupy metod (niezależnie od wykorzystywania danych o naturze fundamentalnej lub technicznej). Do pierwszej grupy można byłoby zaliczyć wszelkie metody oparte na idei pozyskiwania pewnych parametrów z przeszłości pozwalających na zbudowanie modelu wykorzystującego te parametry (czasami tylko chwilowo) i umożliwiających niejako kontynuację modelu w przyszłości. Metod tych nie należy jednak zaliczać do parametrycznych wg taksonomii stosowanej w statystyce – nie są to metody oparte na założeniu zna-



jomości rozkładu statystycznego zmiennej losowej. Do tej grupy z pewnością należy rozpatrywana tu metoda GMDH i wszelkie inne metody oparte na regresji, korelacji, średnich kroczących i ich pochodnych, modelach rodziny GARCH, algorytmach genetycznych czy sieciach neuronowych.

Druga duża rodzina to metody oparte na poszukiwaniu wzorców. Podstawowa idea przyświecająca zwolennikom tej grupy metod jest hipoteza, że jeżeli zmiany w szeregu czasowym wielokrotnie obserwowane w przeszłości następowały wg określonego powtarzalnego schematu, to prawdopodobne jest powtórzenie tego schematu także wówczas, gdy zaczyna być obserwowany jego początek. Charakter wzorców może być bardzo zróżnicowany. Wzorcami są powszechnie stosowane w analizie percepcyjnej rynków, wynikające z teorii fal Elliotta formacje graficzne (takie jak np. formacja głowy i ramion, podwójnego szczytu, spodka, diamentu itp.). Wzorcami mogą być określone sekwencje świec japońskich, może to być także zbiór jednocześnie spełnionych wielu warunków związanych z indykatorami czy parametrami rozkładu statystycznego.

Większość uznanych traderów<sup>6</sup> takich jak Larry Williams, Jack Schwager, Joe Krutsinger czy Thomas DeMark, opisani w pracy Joe Krutsingera [66], posługiwała się i posługuje technikami, które dziś można nazwać poszukiwaniem wzorców. Większość z opisywanych technik (spośród tych ujawnianych) to stosunkowo proste zabiegi mające jednak wspólną cechę. Są rezultatem ciężkiej wieloletniej pracy, tysięcy obserwacji i zauważenia trudnych do logicznego objaśnienia, jednak powtarzalnych sekwencji zdarzeń. Przez stosunkowo proste wzorce należy rozumieć takie, które daje się zaobserwować lub obliczyć w czasie rzeczywistym bez budowania złożonych modeli prognostycznych, zawsze jednak, jak twierdzi większość spośród tych wspomnianych zwycięzców rynku, podstawowym narzędziem wspomagającym ich decyzje jest ustawicznie pracujący komputer.

Interesującym podejściem jest kojarzenie analizy technicznej z fundamentalną. Do tej grupy można na przykład zaliczyć techniki prognostyczne oparte na obserwacji zdarzeń gospodarczych np. ogłaszanych o stałych porach dnia czy miesiąca informacjach o nastrojach lub

---

<sup>6</sup> Trader w powyższym kontekście to zawodowo trudniący się prognozowaniem rynków inwestor, często mentor dla innych graczy. Wielcy amerykańscy traderzy to ludzie sukcesu, którzy wbrew rozpowszechnionym poglądom o nieprzewidywalności rynków dowiedli, że stosowane przez nich techniki, rzadko ujawniane, są skuteczne.

wskaźnikach rozwoju gospodarczego, stopach procentowych ustalanych przez banki centralne itp. Posługiwanie się tymi technikami w sposób automatyczny umożliwia serwis [fxengines.com](http://fxengines.com) [142]. Do tej grupy metod można także zaliczyć opisywaną w niniejszej pracy fuzję GMDH z wynikami obserwacji inwestorów uczestniczących w handlu na wybranym rynku.

W Polsce problematyką GMDH bezpośrednio zainteresowało się niewielkie grono naukowców i powstało niewiele publikacji, z których najważniejsze to prace Witolda Pedrycza [2, 90] i Józefa Korbicza [62, 68]. Spośród innych autorów zainteresowanych metodą należy wymienić nazwiska Zbigniewa Banaszaka, Marcina Mrugalskiego, Janusza Kusia. Powstała także jedna (prawdopodobnie) praca doktorska wykorzystująca wprost metodę GMDH napisana przez Justynę Patalas pod kierunkiem Tadeusza Krupy.

Tytułowa metoda grupowania argumentów (atrybutów) to opracowane przez ukraińskiego uczonego Aleksieja Ivakhnenkę [34-51] podejście do interpolacji i predykcji w formie samoorganizującego się opisu rzeczywistości (modelu). W podejściu tym kojarzona jest wspomniana wiedza i umiejętności. Wiedza to zdolność do opracowania najlepszego w danych okolicznościach modelu (np. predykcyjnego), umiejętność to zdolność do zastosowania sugestii wynikających z tego modelu do podjęcia decyzji (np. inwestycyjnej, z wykorzystaniem platformy brokerskiej). Najstarsze z tych prac powstały już w latach siedemdziesiątych ub. wieku [37, 39, 47, 50].

Aleksiej Ivakhnenko [36, 38] i jego następcy Tetsuo Sawaragi [101], Hema Madala [73], Johann Mueller i Frank Lemke [40, 41, 46, 70] rozwinęli kilka algorytmów mieszczących się w obrębie tej metody.

Metoda w części dotyczącej modelu predykcyjnego polega na równoczesnym zastosowaniu dwóch zasad – zbudowania najlepszego regresyjnego modelu według kryterium opartego na informacji nieużywanej do budowy modelu, oraz drugiej zasady – ograniczenia złożoności struktury modelu regresyjnego do najmniejszej akceptowalnej przez badacza (zgodnie z zasadą brzytwy Ockhama). Istotą metody, od której pochodzi jej nazwa – grupowanie argumentów lub atrybutów, jest uwzględnianie w modelu regresyjnym tylko kilku spośród wielu argumentów – zmiennych wejściowych/zmiennych objaśniających. Oznacza to kolejne rozpatrywanie niejednokrotnie tysięcy modeli regresyjnych o różnych konfiguracjach tworzących je danych wejściowych i wyborze tylko najlepszych z nich. W tym sensie metoda jest samoorganizująca się. Na każdym kroku szeregu czasowego najlepszy w danej chwili model regresyjny ma nie tylko różne współczynniki liczbowe przy poszcze-

gólnych wyrazach równania (w ogólnym przypadku nieliniowego), ale także różną strukturę tego równania w sensie wartości stopnia wykładników wyrazów równania – modelu. Wybór kilku (trafniej, wielu) najlepszych, a nie jednego najlepszego modelu, na danym kroku złożoności algorytmu, Ivaknenko uzasadnia koniecznością pozostawienia pewnej swobody w procesie decyzyjnym. Nawiązując do prac Gabora Ivakhnenko [42] powtarza argumentację selekjonerów np. roślin czy zwierząt, którzy pozostawiają zawsze kilka-kilkanaście najlepszych par czy osobników. W ten sposób uzyskuje się niejako szerszą bazę odwzorowania rzeczywistości. Nie wszystkie najlepsze cechy mogą być przekazane przez pojedynczy egzemplarz (przez najlepszy model rzeczywistości). Ta chwilowa dominacja może być przypadkowa i w wielokrotnie powtarzanej predykcji może się nie sprawdzić.

Z punktu widzenia stopnia naszej determinacji w procesie poznania rzeczywistości metoda ta jest zapewne specyficznym konsensem pomiędzy filozoficznym ewolucjonizmem, wyrażonym równaniem regresji, a kreacjonizmem, wyrażonym poprzez odrzucenie przekonania o skuteczności prognozowania na podstawie przeszłości. Autorski kreacjonizm jest tu wyrażony poprzez zwątpienie w możliwość precyzyjnej prognozy. Sceptycyzm ten wyrażany jest wielokrotnie – myślą przewodnią tej pracy jest przecież odrzucenie wiary w gaussowski rozkład ceny walorów względem średniej (np. rozdział 3) i ciągłe stosowanie rozbudowanej przestrzeni zmiennych decyzyjnych związanych z technologiami platform brokerskich. Sceptycyzm ten wyraża się także w rozpatrywaniu wyłącznie prognoz krótko- i bardzo krótkoterminowych, całkowicie wynikających z matematycznej czy statystycznej wiedzy pozyskanej z przeszłości, ignorującej „chiromancje” ekonomiczne czy polityczne. Postawa taka rzutuje zresztą na przyszłe zastosowania metody – jako automatycznie realizowanego scalpingu<sup>7</sup>.

Metoda z pewnością należy do grupy wielu współczesnych metod opartych na myśleniu indukcyjnym takich jak sieci neuronowe, RBF (*radial basis functions*), algorytmy ewolucyjne, teoria zbiorów przybliżonych. W aspekcie rozwiązywanych zadań predykcyjnych i decyzyjnych metoda łączy podejścia charakterystyczne dla ekstrakcji wiedzy i data mining'u (eksploracji danych) z rozpoznawaniem wzorców i klasyfikacją. Ta część metody z pewnością nawiązuje do postawy przeświadczenia o możliwości ekstrakcji wiedzy z bazy danych. Ta postawa

---

<sup>7</sup> Scalping – to specyficzna dla foreksowych platform brokerskich technika inwestowania polegająca na częstotliwym otwieraniu i zamykaniu pozycji, co kilka-kilkanaście minut, a nawet częściej.

występuje wyraźnie u protoplasty metody – prof. Ivaknenki, który pisał [39, 49], cytując Gabora, że świat realny jest bardzo złożony, bardziej, niż to się powszechnie uważa, nie oznacza to jednak, że należy zrezygnować z możliwości rozpoznania tej złożoności. Mało tego, bardzo często złożony model jest niespodziewanie precyzyjny i nie przeuczony<sup>8</sup>.

Zastosowanie indukcji w poszukiwaniu reguły rządzącej rzeczywistością, czyli wyprowadzenie nieznanego do tej pory prawa na podstawie obserwacji, zaczęło być równie ważne jak dedukcyjny sposób myślenia dopiero w epoce informacyjnej i towarzyszącej jej burzliwemu wzrostowi możliwości obliczeniowych maszyn. Pomimo tej technologicznej rewolucji autor konstatuje ciągle duży dyskomfort w zderzeniu potrzeb obliczeniowych rozpatrywanej tytułowej metody GMDH (Group Method of Data Handling)<sup>9</sup> i możliwości aktualnych komputerów. Dodatkową obliczeniową komplikacją podejścia proponowanego przez Aleksieja Ivakhnenkę i naśladowców – Hema Madalę [73], Johanna Mullera [70], jest rozszerzenie regresyjno-neuronowej asocjacji metody GMDH o kryterium zewnętrzne (oparte na informacji spoza rozpatrywanego zbioru danych uczących) związane z uwarunkowaniami technologicznymi. Te uwarunkowania technologiczne to możliwość uwzględnienia ograniczeń brokerskich.

Istotną wnoszoną w niniejszej pracy innowacją jest zastosowanie ograniczeń w strategii inwestycyjnej związanych bezpośrednio z technologiami stosowanymi na platformach internetowych. Innowacja ta, to specyficzna przestrzeń  $SL^{10} \times TP^{11} \times Tdop$  (poziom akceptowalnej straty/poziom satysfakcjonującego zysku/ustalony maksymalny okres

---

<sup>8</sup> To twierdzenie o krok od teorii chaosu, której Ivaknenko nie przywołuje.

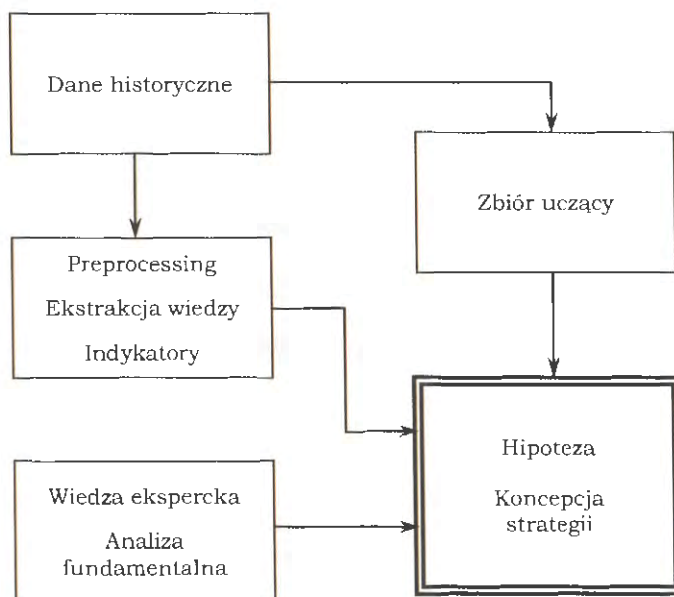
<sup>9</sup> Group Method of Data Handling – angielskie tłumaczenie metody opracowanej przez Ivakhnenkę w oryginale zwane *Metod Gruppovogo Uceta Argumetov* (ros.). W języku polskim nie ma dotąd jednoznacznego przyjętego powszechnie tłumaczenia. Właściwe wydaje się być – Metoda Grupowania Atrybutów czy Metoda Grupowego Uwzględniania Argumentów.

<sup>10</sup> SL – Stop Loss – stosowany na internetowych platformach brokerskich mechanizm półautomatycznego ograniczania starty. Polega on na ręcznym ustawieniu dopuszczalnej przez inwestora ponoszonej przez niego straty, w przypadku zmian rynku w kierunku niezgodnym z oczekiwaniami inwestora i automatycznym wykonaniu tego ograniczenia przez platformę, jeżeli rynek rzeczywiście zmieni się na niekorzyść gracza.

<sup>11</sup> TP – *Take Profit* – inny mechanizm półautomatycznego handlu na platformie brokerskiej. Jeżeli rynek zmieni się zgodnie z oczekiwaniami inwestora, to mechanizm TP dyskontuje zysk. To wyraz obawy inwestora przed utratą chwilowych korzyści, które już osiągnął.

otwarcia pozycji), charakterystyczna dla współczesnych platform internetowych w handlu (tradingu) automatycznym. Tu właśnie technologia miesza się z modelem teoretycznym – modelem wiedzy o obserwowanej rzeczywistości. Rzeczywistością tą dla inwestora jest szereg czasowy danych obserwowanych na rynkach finansowych. W pracy, w dalszej jej części, przedstawione zostaną przykłady zastosowania metody GMDH (dokładniej, niektórych, opracowanych na jej podstawie autorskich algorytmów) na różnych rynkach finansowych. Podstawą do oceny skuteczności algorytmu będzie najczęściej kryterium Sharpe’a (lub Sortino) [106, 109, 140] uwzględniające jednocześnie zysk i ryzyko (średnią stopy zwrotu i jej wariancję).

Ogólna koncepcja kojarzenia dwóch systemów – generowania koncepcji predykcyjnej i systemu podejmowania decyzji w warunkach konkretnej platformy technologicznej przedstawiona jest na dwu schematach.



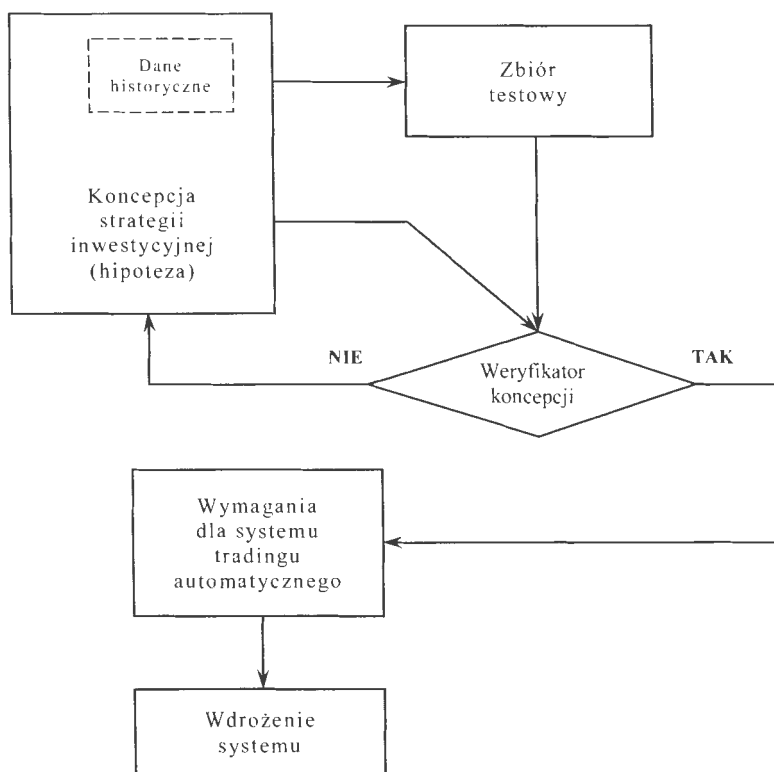
Rys. 1. Schemat blokowy procesu powstawania hipotez badawczych dotyczących koncepcji strategii inwestycyjnych.

Rys. 1 przedstawia schemat poszukiwania koncepcji strategii predykcyjnej. Rodzi się ona w przestrzeni danych historycznych w wyniku ustawicznej penetracji różnymi (z założenia heurystycznymi)

## Wprowadzenie

algorytmami tej przestrzeni, stosowaniu wyrafinowanej filtracji, uznanych i nowych wskaźników wysyłających sygnały do otwarcia lub zamknięcia pozycji. Do tych zabiegów, należących do analizy technicznej, dodać można ekspercką wiedzę o przyszłości rozpatrywanego rynku, wynikającą z przewidywanych skutków zdarzeń politycznych, gospodarczych czy społecznych. Symboliczną sumę tej wiedzy (model wiedzy) przedstawia rys. 1. Jej rezultatem jest pewna strategia, z którą w danej chwili utożsamia się inwestor w swym przekonaniu o jej skuteczności.

Gdy jest już o jej wartości przekonany, to warto podjąć decyzję o jej wdrożeniu. Schemat tego postępowania ilustruje rys. 2.



Rys. 2. Schemat blokowy procesu opracowywania wymagań dla systemu tradingu automatycznego.

Inwestor (gracz giełdowy, użytkownik platformy foreksowej) rozważa decyzje o testowaniu strategii. Racjonalne i logiczne jest, by dokonać tego na nowych, nieuwzględnianych wcześniej danych. To ty-

powy zabieg zwany cross-walidacją lub sprawdzeniem krzyżowym. Może on polegać na sekwencyjnym potraktowaniu przestrzeni danych historycznych (najpierw uczenie, potem testowanie) lub wykonywaniu testów co pewien czas w miarę przemieszczania się wzdłuż szeregu czasowego.

Jeżeli weryfikacja strategii jest pomyślna, to może nastąpić faza wdrożenia systemu z uwzględnieniem wymagań stawianych przez daną platformę inwestycyjną. Są to najczęściej wymagania natury technologicznej, np. uwzględnienie API<sup>12</sup> serwera, jeżeli jest udostępniane, uwzględnienie ograniczeń brokera, np. dopuszczalnej częstotliwości tradingu, dopuszczalnej odległości pomiędzy otwarciem pozycji a zleceniem zamknięcia, wielkości prowizji zależnej od zainwestowanej kwoty (najczęściej nieliniowej) itp.

Ograniczenia te odniesione do konkretnej rozpatrywanej strategii, o obiecujących wynikach uzyskanych w fazie testowania, mogą uczynić ją całkowicie nieprzydatną w praktyce. To właśnie istota rozpatrywanego tu zderzenia teoretycznej sprawności strategii z twardą rzeczywistością.

Twórca metody Group Method of Data Handling – prof. Ivakhnenko, początkowo nie rozpatrywał możliwości jej zastosowania na rynkach finansowych. Rozważał częściej przypadki predykcji ekonomicznej, ekologicznej, czy w systemach technicznych [37, 39, 48, 51]. Zrobił to dopiero pod koniec ubiegłego wieku wspólnie z niemieckimi badaczami Muellerem i Lemke [40, 41]. W tym czasie burzliwie rozwijały się statystyczne metody predykcyjne oparte najczęściej na ułomnych założeniach, np. o normalnym rozkładzie stopy zwrotu lub stacjonarności rynku. Pomimo kilku nagród Nobla dotyczących tego obszaru aktywności człowieka (Markowitz, Sharp, Merton, Scholes, McFadden, Engle) modele budowane przez te znakomitości nie doprowadziły, na szczęście, do ostatecznego zdefiniowania wiarygodnego modelu predykcyjnego. Nie doprowadza do niego, także na szczęście, ta praca. Doprowadza do lepszego zrozumienia aforyzmu Hegla: „To, co znane, nie jest jeszcze przez to, że jest znane, czymś poznanym”.

W odróżnieniu od wielu prac ekonometrycznych praca koncentruje się na pragmatycznym powiązaniu predykcji z możliwościami jej

---

<sup>12</sup> API – Application Programming Interface – interfejs programowania aplikacji, system umożliwiający, w danym rozumieniu, komunikację pomiędzy serwerem, na którym umieszczona jest aplikacja brokerska a klientem, który dzięki API ma możliwość oprogramowania swoich aplikacji w sposób umożliwiający komunikację z serwerem.

internetowej (automatycznej) realizacji. Jak wspomniano, wymaga to uwzględnienia szeregu niespodziewanych ograniczeń natury technologicznej. By udowodnić ten związek w pracy przedstawiono wyniki wielu eksperymentów w przestrzeniach danych historycznych różnych instrumentów finansowych. Wnioski są interesujące – zwłaszcza w wymiarze praktycznym.

Praca składa się z 6 rozdziałów. Pierwsze dwa poświęcone są szczegółowemu badaniu różnych odmian algorytmu kombinatorycznego zbudowanych na koncepcji algorytmu liniowego rozpatrywanego przez Ivakhnenkę [42, 50]. Zmodyfikowane algorytmy GMDH są doprowadzone do takiego stopnia szczegółowości, który umożliwia ich implementację w dowolnie wybranym środowisku programowym. Autor nie publikuje własnych skryptów opracowanych w środowisku Matlab z uwagi na ich rozmiary, być może także ze względu na niedoskonałość kodowania, ale także mając prawo do ochrony tych utworów jako dodatkowej wartości. Skrypty programowe towarzyszące tej pracy są niejednokrotnie bardzo rozbudowane i są często rezultatem wielu miesięcy intensywnych eksperymentów. W dodatku do pracy przedstawiony jest fragment najbardziej złożonego algorytmu GMDH pozwalający Czytelnikowi na opracowanie własnego programu symulacyjnego.

W rozdziale trzecim zbadane zostały wszelkie możliwe fragmenty wielowymiarowej przestrzeni predyktorów, które użyte były w pierwszych rozdziałach do badania skuteczności algorytmów GMDH. Trzy ostatnie rozdziały to prezentacja wyników rozmaitych badań z zastosowaniem tytułowej metody. Badania przeprowadzane są w różnych przestrzeniach danych, także wstępnie przetworzonych i w zaskakująco różnych, odmiennych stylach handlu. Ich wspólnym mianownikiem jest zastosowanie krocząco zmienianej regresji opartej na GMDH – ogólnie podsumowując z wyjątkowo dobrymi rezultatami.

Praca jest przygotowaniem do handlu automatycznego. Wyników takich eksperymentów, z wykorzystaniem systemu informatycznego umożliwiającego handel automatyczny, autor tu nie zamieszcza. Wszystkie symulacje wykonane są na różnych danych historycznych ze starannym odcięciem danych uczących od danych testowych. Celem zasadniczym, jest przygotowanie strategii do zastosowania w kompletnych systemach internetowych – zawierających moduły pozyskiwania bieżących danych, moduły dostępu do zasobów historycznych i moduły realizacji handlu wg przetestowanych i adaptowanych na bieżąco strategii inwestycyjnych.





# 1

## Zmodyfikowany kombinatoryczny algorytm metody grupowania argumentów GMDH w zadaniach predykcji rynku walutowego

*W każdej nauce tyle jest prawdy, ile w niej jest matematyki*  
Emmanuel Kant (1724-1804)

Jednym z najważniejszych i najczęściej opisywanych przez twórcę metody Aleksieja Ivakhnenko [50] i wspomnianych jego uczniów i następców [70,73] jest algorytm kombinatoryczny (combinatorial algorithm). Prof. Ivakhnenko jako współtwórcę tego algorytmu przywołuje V. Stepaszkę [48].

W największym uproszczeniu algorytm ten rozwiązuje zagadnienie aproksymacji lub prognozy poprzez kolejne próby lepszego dopasowania kolejnego modelu do wyników zaobserwowanych, a każdy kolejny model charakteryzuje się wyższym stopniem złożoności.

W rozpatrywanym przykładzie badana będzie skuteczność predykcji modeli liniowych uzyskiwanych poprzez włączanie coraz to większej liczby argumentów (danych pomierzonych, historycznych).

Zgodnie z wieloma podstawowymi badaniami metody GMDH, jeżeli dane wyjściowe w pełni (w miarę) odzwierciedlają istotę procesu, to w wyniku wzrostu złożoności modelu (rzędu modelu) winno uzyskiwać się minimum pewnego kryterium rozbieżności utworzonego poprzez

porównanie wyników modelu z wynikami rzeczywiście uzyskanymi. Jak zwykle w tego rodzaju badaniach kryterium jest uznaniowe, zależy od kreatywności autora, musi jednak logicznie odpowiadać jego oczekiwaniom, tzn. gdy np. będzie mało winno zapewnić aksjologiczną satysfakcję. W poniższym przykładzie wyjaśnione będą wstępne zagadnienia złożoności modelu i kryterium efektywności modelowania (tu prognozowania).

Niech np. będą dane pomierzone wartości pewnych wielkości, co do których istnieje przypuszczenie, że mogą wpływać na przebieg zmiennej wyjściowej (obserwowanej) oznaczanej dalej -  $y$ . W większości rozważań w niniejszej pracy zmienna  $y$  będzie szeregiem czasowym.

Na rynkach finansowych (w szczególności na rozpatrywanym tu rynku walutowym) klasycznym „wyjściem” systemu jest obserwowana w czasie wartość kursu walutowego. Pomimo kilkudziesięciu lat intensywnych badań i praktycznych wysiłków nie udało się ustalić, od jakich to zmiennych „wejściowych” to „wyjście” zależy. Przypuszczeń jest sporo i na tej solidnej wiedzy dziedzinowej bazując można wytypować pewien zbiór argumentów  $X = \{x_1, x_2, \dots, x_i, \dots, x_M\}$ , mogących wpływać na  $y$  – wartości kursu. Każda zmienna  $x_i$  tworzy szereg czasowy:

$$x_{ik} = x_{i,t-k}, x_{i,t-k+1}, \dots, x_{i,t-1}, x_{it} \\ \text{dla } i=1, 2, \dots, M \text{ oraz } k=1, 2, \dots, K, t=K+1, K+2, \dots$$

Podobnie:

$$y_k = y_{t-k}, y_{t-k+1}, \dots, y_{t-1}, y_t \\ \text{dla } k=1, 2, \dots, K, t=K+1, K+2, \dots$$

W szeregach tych  $t$  oznacza indeks chwili bieżącej, a  $(t-k)$  – indeks chwili odległej wstecz od chwili bieżącej o  $k$  okresów (dokonywania pomiaru – tu zwanych świecami). Najczęściej bada się wstecz nie cały nieskończony szereg czasowy, a jedynie pewną skończoną liczbę pomiarów – tu oznaczoną jako  $K$ .

W ten sposób gromadzone są na rynkach walutowych niewyobrażalne ilości informacji np. o kursach walutowych, wartościach akcji, indeksach, derywatach mierzonych (obserwowanych) w sposób niemal ciągły przez wiele ostatnich lat.

Sformułowanie „w sposób niemal ciągły” oznacza, że wiele z tych wartości  $y$  zmienia się częstokroć co kilka sekund i wszystkie te dane są dostępne (może nie powszechnie) do badań i analizy. By zapanować nad tym żywiołem dokonuje się na tych rynkach swoistego próbkowania danych poprzez stosowanie tzw. świec oznaczających zbiór czterech

wartości obserwowanych co wybrany odcinek czasu. Np. świeca OHLC<sup>13</sup> 15min oznacza cztery wartości obserwowanej zmiennej (np. kursu) określone co 15 minut oznaczające kolejno wartość zmiennej na początku 15 minutowego okresu (Open), wartość najwyższą w tym okresie (High), najniższą (Low) i ostatnią/zamknięcia (Close). Ta ostatnia jest najczęściej używana właściwie jako obserwowany wynik (wyjście)  $y$ , trzy pierwsze należą do zbioru  $X$  wartości opisujących (nazywanych także - objaśniających, predyktorów, wejściowych).

Do tego zbioru należą może także wiele zmiennych uzyskiwanych w wyniku przetwarzania danych np. tzw. średnie kroczące, czyli wartości średnie z ostatnich kilku-kilkudziesięciu świec, wariancje zmiennych, inne momenty statystyczne czy specjalnie tworzone wymyślne identyfikatory służące dziełu przewidywania nieprzewidywalnego. Niektóre z tych miar mają fundamentalne znaczenie dla praktyki codziennego tradingu (handlu) na wszystkich giełdach świata, także na walutowych rynkach Foreksu.

Każda z tych zmiennych  $x_i$  może być obserwowana i rejestrowana w tym samym momencie co odpowiadający jej wynik  $y$ , ale interesująca może być nie tylko relacja, co działo się ze zmienną  $x_{it}$ , gdy w chwili bieżącej obserwowano  $y_t$ . Interesująca może być wartość  $x_{ik}$ , gdy obserwowano  $y_t$ , gdzie  $i=1, 2, \dots, M$ , a  $k=t, t-1, t-2, \dots$ . Oznacza to, że interesujące będą przesunięte w czasie wektory zmiennych wejściowych. Przesunięte wstecz, gdyż realnie działające modele predykcyjne będzie można zbudować opierając się wyłącznie na zdarzeniach przeszłych, dokonanych.

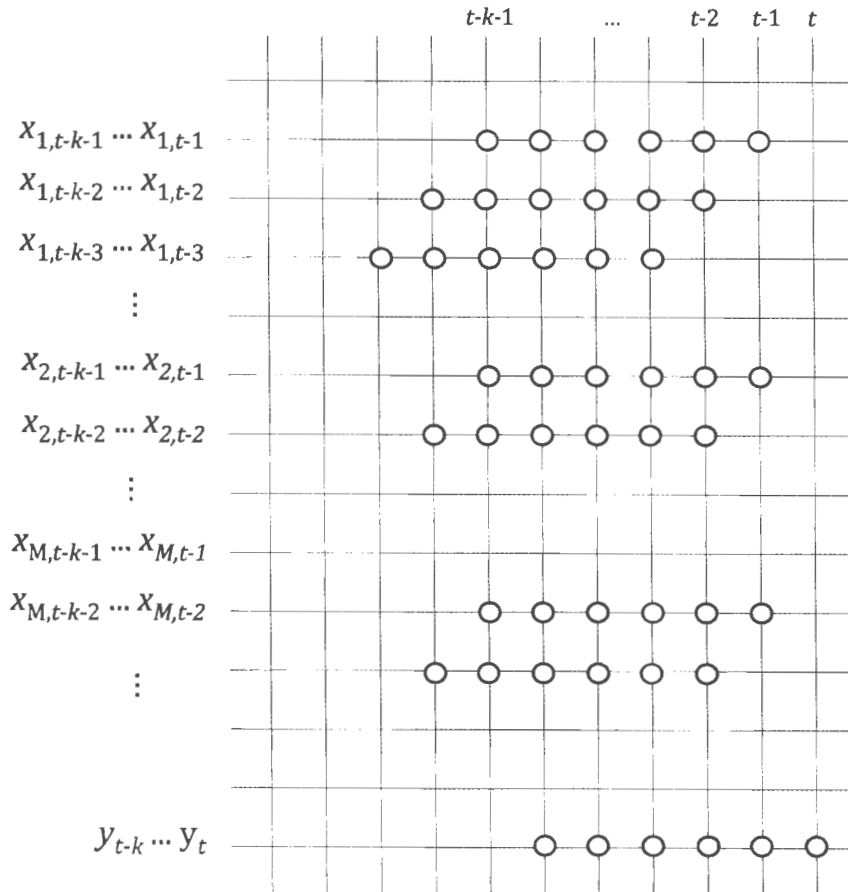
Na dokładność modelowania ma wpływ także wielkość (długość) szeregu czasowego użytego do modelowania, nazywana często oknem, gdyż przypomina okno o stałej szerokości przesuwane po macierzy danych. Niech szerokość tego okna wynosi  $K$ . Jeżeli dane pomierzone przedstawione będą w postaci macierzy o  $M$  kolumnach plus jedna kolumna na wartość  $y$  i  $N$  wierszach kolejnych pomiarów (obserwacji), to  $K$  oznacza liczbę obok siebie umieszczonych wierszy, które brane są pod uwagę przy modelowaniu kolumny wynikowej (rys. 3). Stąd jeżeli „szerokość” okna to liczba położonych obok siebie wierszy, to trafniej byłoby wymiar  $K$  nazwać „wysokością”, byłoby to jednak naruszeniem stosowanej w literaturze konwencji.

---

<sup>13</sup> OHLC – umowne oznaczenie tzw. świecy (*candle*), która jest wektorem czterech kolejnych wartości zmiennej z zadanego przedziału np. 15 min. – kolejno – Open, High, Low, Close. Nazwa pochodzi od sposobu graficznej prezentacji tych czterech danych najczęściej w formie prostokąta z „wąsami”.

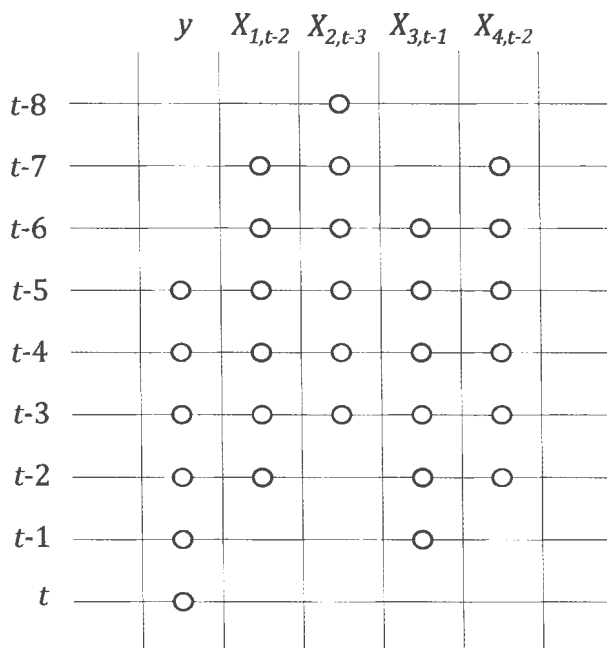


Rozdział 1. Zmodyfikowany kombinatoryczny algorytm metody GMDH...



Rys. 4. Wzajemne ułożenie wektorów zmiennych  $x_i$  przesuniętych o  $k=1,2, \dots, k$  oraz wektora wyników  $y_j, j = t-k, \dots, t$ .

Widać tu wybrane z macierzy obserwacji  $X$  (rys. 4) cztery wektory kończące się co najwyżej w chwili  $t-1$ . W chwili  $t$  obserwowana jest jedynie zmienna wyjściowa  $y_t$  po to, by po nauczaniu systemu, czyli po identyfikacji zależności pomiędzy argumentami  $X_i$  a zmienną wyjściową  $y$ , fakt ten wykorzystać w celu prognozy np. w chwili  $t+1$  (a być może i następnych).

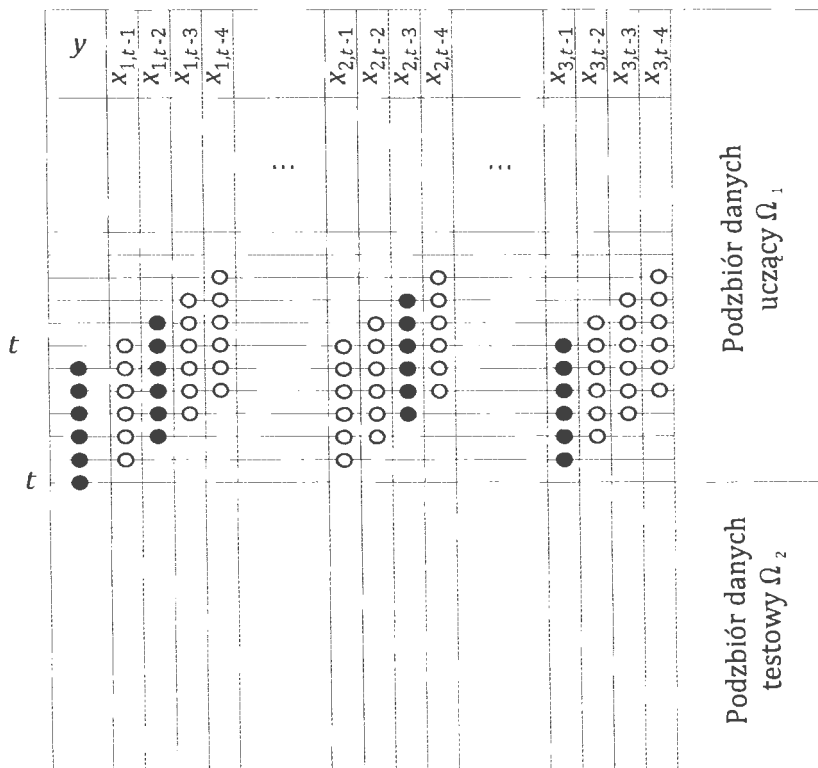


Rys. 5. Przykład wzajemnego rozmieszczenia w macierzy danych nieprzesuniętego wektora zmiennej obserwowanej  $y$  i wybranych wektorów argumentów.

Tak więc na diagramie z rys. 5 widać cztery wektory przesunięte o  $d_{ij}$ , gdzie  $i$  oznacza indeks argumentu  $X_i$  a  $j$  związane jest z wielkością przesunięcia (opóźnienia) tego  $i$ -go szeregu w stosunku do zmiennej  $y$ . Jeżeli np. szereg  $X_1$  jest przesunięty o 2 okresy pomiarowe (świece) to  $d_{1j}=2$  i do wybranego wektora należą wyrazy  $x_{1,t-2}, x_{1,t-3} \dots x_{1,t-k-2}$ . Dla kolejnych wektorów z rys. 5 podobnie obserwuje się przesunięcie  $d_{2j}=3, d_{3j}=1, d_{4j}=2$ .

Wielkość przesunięcia może osiągać wartość  $L$  okresów nie związaną z długością wektora  $K$ . Dopiero poprzez badania konkretnego środowiska danych możliwe będzie określenie racjonalnych wartości ograniczeń  $L$  i  $K$ .

Podsumowując, diagram z rys. 5 można interpretować jako niewielki fragment przestrzeni  $X$  służący prognozie wartości  $y_{t+1}, y_{t+2}$  itd. (rys. 6).



Rys. 6. Wybrane wektory argumentów na tle całej dostępnej przestrzeni danych  $X$ . Możliwy do prognozowania jest tylko jeden okres zmiennej obserwowanej  $y$ .

Przy zastosowaniu powyższej zasady wykorzystania informacji o wartościach argumentów przesuniętych o jeden lub więcej okresów pomiarowych wstecz w stosunku do wektora wynikowego  $y$  możliwe jest wykonanie prognozy o jeden (zaledwie) okres do przodu. Oczywiście możliwe jest „prognozowanie” o więcej niż jeden okres wprzód (przy takiej konfiguracji danych, tzn. z przesunięciem wektora wyników o co najmniej jeden okres), lecz byłoby to działanie pozbawione matematycznego uzasadnienia i logiki. Ewentualny sukces byłby przypadkowy.

W przypadku natomiast wykorzystania takich argumentów, które kończą się więcej niż jeden okres przed  $t$  (np. na diagramie 5 są to zmienne  $X_1$ ,  $X_2$ , i  $X_4$ ) można pokusić się o prognozę na 2 okresy w przód.

Wynika to z możliwości „nauczenia” wektora  $Y$  zależności od pozostałych argumentów i następnie przesuwania go w celu prognozy o tyle okresów wprzód, ile wynosi najmniejsze opóźnienie.



Ostatecznie można w sposób wymuszony określić pożądany okres (liczbę świec) prognozy przemieszczając odpowiednio wektor  $Y$  w stosunku do wektorów  $X_i$ . Prawdopodobnie, gdyż jest to logiczne, im większa będzie liczba prognozowanych okresów wprzód, tym trudniej będzie uzyskać satysfakcjonującą wartość kryterium określającego rozbieżność pomiędzy modelowaną predykcją wartości zmiennej wyjściowej a rzeczywistą wartością zaobserwowaną tej zmiennej. Kryterium takie często jest związane, w sensie niepożądanego skutku, z miarą regularności modelu. Regularnością modelu, a tu raczej algorytmu, Ivakhnenko [47, 50] nazywa taką jego cechę, że spośród wszelkich wymyślonych przez twórcę modelu jego komponentów (czynników, wejść, parametrów itp.) algorytm w wyniku selekcji zbliża się w swoim działaniu do sposobu myślenia człowieka i eliminuje komponenty nadmiarowe.

### 1.1. Kryteria rozbieżności modelu i zmiennej obserwowanej

Dla rozpatrywanych wcześniej przypadków różnych długości horyzontu prognostycznego złożony np. z jednego, dwóch, trzech i więcej okresów (świec) wprzód stosować można różne formy kryteriów rozbieżności.

Proponuje się rozróżnienie prognozy wykonywanej jeden krok wprzód i prognozy wykonywanej na większą liczbę okresów.

Abstrahując od sposobu wyznaczenia modelu  $\hat{y}$  opisującego obserwowaną zmienną wyjściową  $y$  w zależności od wektorów  $X_i$ , na rys. 7 przedstawiona została sytuacja, gdy prognoza wykonywana jest za ledwie jeden okres (krok) wprzód. Ilustrują to dwa stany. Na rys. 7a (pierwszy stan) przedstawiony jest model  $\hat{y}$  wybiegający o jeden krok wprzód w stosunku do chwili bieżącej  $t$  i zarejestrowany do tej chwili rzeczywisty przebieg zmiennej wyjściowej  $y$ .

Na rys. 7b (drugi stan) przedstawiono sytuację, jaka wystąpi (można powiedzieć, wystąpiła, jeżeli badania prowadzone są w zbiorze danych historycznych) po wykonaniu tego jednego kroku wprzód. Ponieważ, z założenia, w eksperymencie używa się danych historycznych, a więc uprzednio zarejestrowanych, to wartość rzeczywista  $y$  w chwili  $t+1$  została także wzięta z przestrzeni danych  $X$ .

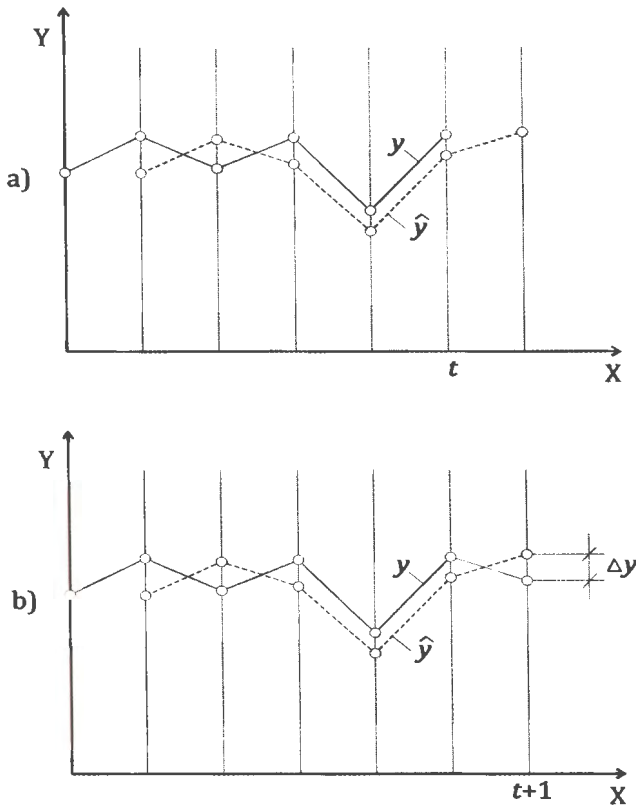
W eksperymencie oczywiście zakładano jednak, że nie była ona znana w chwili  $t$ . Tak więc różnica:

$$\Delta y = \text{abs}(y - \hat{y}) \quad (1)$$

w chwili  $t+1$  może być traktowana jako pewne zewnętrzne kryterium, które o ile zmniejszałoby się, zapewniłoby oczekiwaną satysfakcję.

W tym sensie kryterium  $\Delta y$  lub  $\text{abs}(\Delta y)$  określane niekiedy jako MAPE (maximum absolute percentage error) – gdy wyrażane jest w procentach bezwzględnej wartości  $y$  w chwili obserwacji), jest kryterium użytecznym jakościowo. Jeżeli także satysfakcjonujące jest w sensie ilościowym tzn. błędy są stosunkowo małe – jest to kryterium efektywne – prowadzące do rozwiązania zadania (tu – prognozy).

Jeżeli z kolei, predykcja miałaby dotyczyć więcej niż jednego kroku wprzód, np. dwóch, trzech lub więcej, to powstaje problem jak uogólnić regularność odchylenia na każdym kroku.



Rys. 7. Przebieg rzeczywistych zmian zmiennej rzeczywistej  $y$  oraz jej modelu  $\hat{y}$  dla sytuacji prognozy jeden krok wprzód (a) oraz po weryfikacji rzeczywistego odchylenia  $\Delta y$  (b).

W literaturze [15, 16, 74] często proponuje się kryterium porównujące błędy predykcji do odchyień zmiennej wyjściowej, tak jak poniżej. Podobnie jak poprzednio przy kryterium  $\Delta y$  (1) występuje tu wektor rzeczywistych wartości zmiennej wyjściowej  $y$  i wektor tej samej zmiennej, lecz będący wynikiem modelowania  $\hat{y}$ . Niech przedmiotem predykcji będzie pewien horyzont określony liczbą kroków wprzód  $hp$ . W przypadku prognozy wykonywanej jeden krok wprzód liczba ta wynosi  $hp = 1$ .

Wspomniane kryterium określone jest następująco:

$$C_R = \frac{\sum_i^{hp} (y_i - \hat{y}_i)^2}{\sum_i (y_i - \bar{y})^2} \quad (2)$$

gdzie:

$y_i$  – rzeczywiste wartości zmiennej wyjściowej  $y$  dla chwili bieżącej i od chwili bieżącej do  $i + hp$ ;

$\hat{y}_i$  – modelowane wartości zmiennej wyjściowej;

$\bar{y}$  – średnia wartość rzeczywistej wartości zmiennej wyjściowej dla horyzontu  $hp$ .

Przykład zmienności  $y$  oraz  $\hat{y}$  dla  $hp = 2$  przedstawiono na rys. 8. Na rysunku tym (rys. 8) zaznaczono różnice między modelowaną i rzeczywistą wartością obserwowanej zmiennej wyjściowej. Zmienna wyjściowa  $y$  może być rozpatrywana, jako zmienna normalizowana (względna)  $y_w$  lub z pominięcia tej normalizacji.

Rozróżnienie to (pomiędzy  $y$  i  $y_w$ ) ma niezwykle istotne, często pomijane znaczenie metodyczne i praktyczne. Otóż, w szczególności na rynku walutowym, przy badaniu zmienności kursu par walutowych zmiany obserwowane są w tzw. pipsach, co najczęściej oznacza czwartą po przecinku cyfrę znaczącą, a przy przeciętnych wartościach par walutowych najczęściej pomiędzy 1 – 2 oznacza to (ten jeden pips)<sup>14</sup> wartość względną rzędu  $0.1 \div 0.2\%$ .

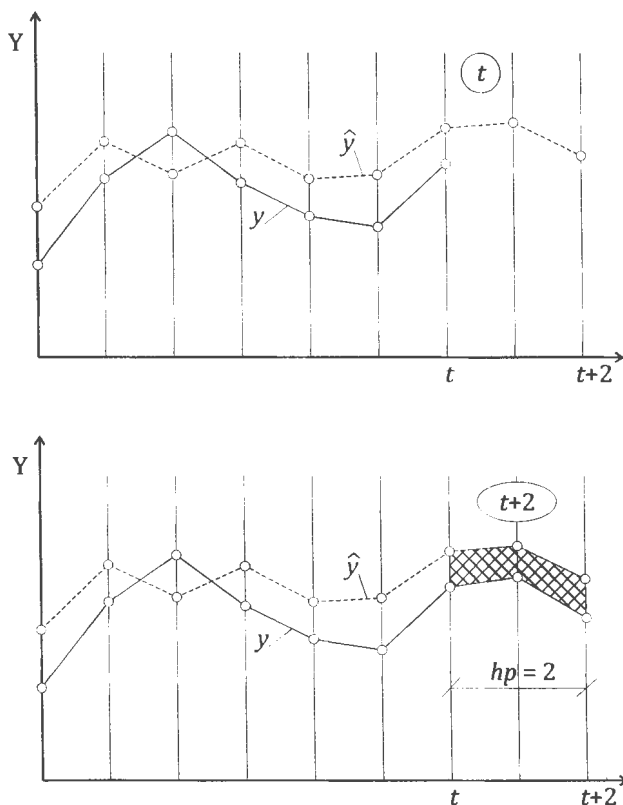
Zmiany (ewentualnie błędy) rzędu 5-10 pipsów mogą mieć istotne praktyczne znaczenie dla strategii inwestycyjnej. Oznacza to, że błąd rzędu  $0.5 \div 1\%$  jest bardzo istotny, gdy tymczasem nawet wśród

---

<sup>14</sup> Takie są przeciętne wartości większości głównych par walutowych np. EUR/USD, GBP/USD, EUR/CHF i innych. Nie dotyczy to par związanych z jennem np. USD/JPY.

entuzjastów tej metody odnajduje się w ich publikacjach zadowolenie z regularności predykcyjnej rzędu  $2 \div 3\%$  (np. [36, 69, 76]).

Tak więc konfrontacja uzyskiwanych wyników modelowania z rzeczywistymi wartościami kursu waluty ma istotne (może nawet decydujące) znaczenie praktyczne. Jednocześnie z powodów metodycznych wskazane jest posługiwanie się wartościami znormalizowanymi (głównie z powodu ogromnej różnorodności argumentów wejściowych) i w tym sensie model obliczeniowy, by był jednorodny i łatwiejsza była percepcja zmian i wpływ argumentów, winien być budowany w przestrzeni zmiennych znormalizowanych.



Rys. 8. Przebieg rzeczywistych  $y$  i modelowanych  $\hat{y}$  wartości zmiennej wyjściowej z zaznaczonym horyzontem predykcyjnym  $hp = 2$ .

Dla sprawdzenia praktycznego znaczenia wyników, winny one być jednak ponownie konwertowane do przestrzeni zmiennych rzeczywistych (bezwzględnych).

Konwersja (normalizacja) przestrzeni zmiennych rzeczywistych (bezwzględnych) na względne przeprowadzona była wg reguły:

$$y_{wi} = \frac{y_i - \min_i y_i}{\max_i y_i - \min_i y_i} \quad (3)$$

gdzie:  $\min y_i$  – minimalne wartość elementu w wektorze  $y_i$ ,  $\max y_i$  – maksymalna wartość elementu w wektorze  $y_i$ .

Formuła ta umożliwi następującą konwersję odwrotną:

$$y_i = y_{wi} (\max_i y_i - \min_i y_i) + \min_i y_i \quad (4)$$

W obu proponowanych kryteriach (1) i (2) niezwykle istotne, fundamentalnie ważne jest to, by przy ich obliczaniu skorzystać ze zbioru wartości zmiennej wyjściowej nieuwzględnianego przy obliczaniu modelu  $\hat{y}$ , a więc wartości niejako zewnętrznych, znajdujących się w testowej (nie uczącej) części danych.

## 1.2. Ogólna koncepcja algorytmu kombinatorycznego

Algorytm kombinatoryczny, jako jeden z wielu możliwych należących do metody GMDH, zaproponowany został w pierwotnej postaci przez jej twórcę prof. Ivakhnenkę [50].

W najogólniejszych zarysach algorytm ten (tu rozpatruje się taką jego odmianę, której zasadniczym celem jest predykcja) ma postać sieci neuronowej (obliczeniowej)<sup>15</sup>, złożonej z wielu warstw z aktywnymi neuronami [40, 45, 51]. Zadaniem każdego neuronu jest selekcja wejść z poprzedniej warstwy i budowanie modelu efektywnej prognozy na podstawie wspomnianych kryteriów zewnętrznych.

Aktywność neuronu (można by powiedzieć także – jego inteligencja) polega na stosowaniu pewnego indywidualnego, charakterystycznego dla danego neuronu kryterium selekcji wejść (z poprzedniej warstwy) i wykorzystywaniu tylko niektórych spośród tych wejść – takich mianowicie, które pozwalają na budowanie dobrego (regularnego) modelu.

---

<sup>15</sup> Niekiedy metodę tę nazywa się PNN (Polynomial Neural Network) np. [www.gmdh.net](http://www.gmdh.net).

Tak więc celem sieci jest sukcesywne wykorzystywanie najlepszego modelu (w sensie zgodności zmiennej obserwowanej  $y$  ze środowiskiem, które generuje sygnały wykorzystywane do prognozy) lub modelu jednego z wielu dostatecznie dokładnych, ale należącego do tych, znalezionych w pierwszej kolejności. Idee te odkryte przez Ivakhnenkę dość dawno stanowią podwaliny dla budowy całej rodziny algorytmów z grupy GMDH. Gabor, Ivakhnenko, Yurashkovskiy przewidywali, a potem dowiedli [49, 50], że budowanie wielu modeli opisujących złożoną rzeczywistość, zwanych niefizycznymi modelami (także Sawaragi [101, 102]), może doprowadzić do określenia modelu o optymalnej złożoności. Złożoność tę można ocenić poprzez liczbę użytych argumentów z przestrzeni  $\mathbf{X}$ , w tym uwzględniając także przesunięcia czasowe (oczywiście wstecz) tych argumentów. Dodatkowo złożoność modelu może być scharakteryzowana poprzez najwyższy stopień wielomianu zastosowanego do opisu.

Zbiór modeli może być określony najogólniej jako podzbiór wielomianu Gabora-Kolmogorova [50]:

$$Y = a_0 + \sum_i^M a_i x_i + \sum_i^M \sum_j^M a_{ij} x_i x_j + \sum_i^M \sum_j^M \sum_k^M a_{ijk} x_i x_j x_k + \dots \quad (5)$$

gdzie:

$a_0$  – wyraz wolny w wielomianie (modelu);

$a_i$  – współczynniki liczbowe przy wyrazach (członach) liniowych wielomianu;

$a_{ij}$  – współczynniki liczbowe przy wyrazach (członach) kwadratowych wielomianu;

$a_{ijk}$  – współczynniki liczbowe przy wyrazach (członach) trzeciego stopnia wielomianu; itd.

W powyższej postaci wielomian nie ma praktycznego zastosowania z uwagi na szybko rosnącą liczbę członów wielomianu wynoszącą  $C_{M+R}^M$ , gdzie  $M$  to liczba zmiennych wejściowych (argumentów, kolumn macierzy danych także z uwzględnieniem przesunięć czasowych) a  $R$  to stopień maksymalny zastosowanego wielomianu (5).

W rozpatrywanych tu badaniach użyto ponad 20 argumentów wejściowych (jeszcze do tej pory nie opisanych). Już przy niewielkim stopniu wielomianu np. kwadratowym lub sześciennym, liczba kombinacji przekroczyłaby możliwości obliczeniowe (w aspekcie pozyskiwania wiedzy o optymalnych modelach dla bieżącego zarządzania zadaniami prognozy) nawet bardzo współczesnych komputerów.

Właśnie rozpatrywanie pewnych wyborów z wielomianu Gabora-Kołmogorowa określa wielość algorytmów należących do rodziny GMDH. Omawiany tu algorytm kombinatoryczny jest stosunkowo prosty, gdyż jest rozwinięta, ale liniową odmianą metody, jednak, jak się okaże niekiedy zaskakująco efektywną.

Odmiana ta, zwana np. w [39, 73] kombinatoryczną (lub COMBI) pozornie jest tylko prostym przypadkiem wyboru z wielomianu (5). Przy dużej liczbie zmiennych wejściowych szybko bowiem rośnie złożoność obliczeniowa i konieczne są ograniczenia natury technicznej, nie matematycznej.

W rozpatrywanym w niniejszej pracy liczba wybranych po wielu próbach argumentów tworzących zbiór  $X_i$  (taki jak np. na rys. 4) wynosiła  $M=26$ , czyli był to zbiór wektorów kolumnowych poszczególnych zmiennych z indeksami kolejno  $i=1, 2, \dots, 26$ . Sposób ich doboru i charakterystyka każdego argumentu będą omówione nieco dalej. Wspomniane jednak ograniczenia techniczne (natury wynikającej z ograniczeń wydolności komputera) spowodowały konieczność tworzenia podzbiorów modeli (zamiast wszystkich możliwych rozwiązań) w poszczególnych warstwach sieci neuronowej.

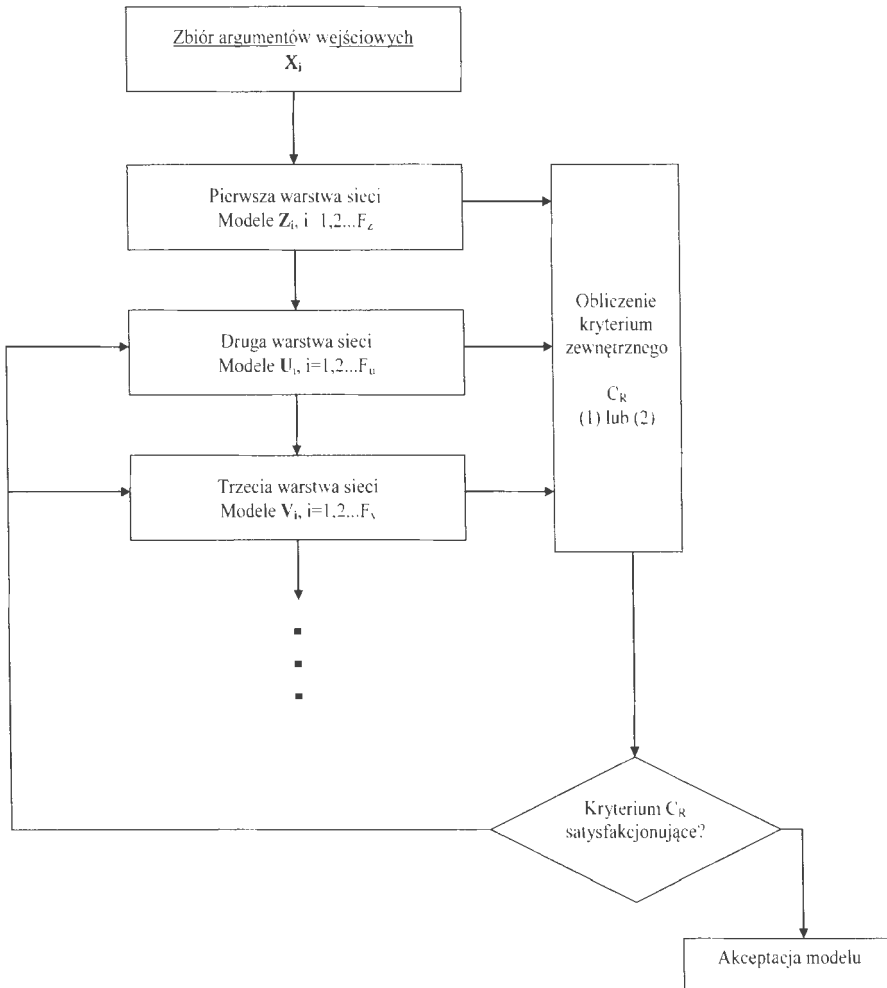
Najbardziej ogólny schemat tej sieci przedstawia rys. 9. Jest to schemat obrazujący przebieg obliczeń dla jednego jedynego kroku dla wybranego odcinka czasu. W przypadku rozpatrywanego tu zadania prognozy przebiegu kursu pary walut, ten odcinek to okres pomiędzy kolejnymi świecami.

Schemat ten objaśnia sposób osiągnięcia wystarczającej złożoności modelu opisującego rzeczywistość (modelu nie-fizycznego). Otóż ze zbioru argumentów wybierane są podzbiory (tu - pary argumentów) i sprawdzany jest model regresyjny wg pewnego kryterium. Kryterium to (np. takie jak (1) lub (2) zbudowane jest na danych nie uwzględnianych przy budowaniu modelu regresyjnego – stąd określenie tego kryterium jako zewnętrznego.

Jeżeli spośród wszystkich rozpatrywanych modeli w warstwie pierwszej (tu par argumentów) w postaci:

$$\begin{aligned} z_1 &= f(x_1, x_2) \\ z_2 &= f(x_1, x_3) \\ &\dots \\ &\dots \\ z_s &= f(x_{M-1}, x_M) \end{aligned} \tag{6}$$

gdzie:  $M$  – liczba argumentów (kolumn w macierzy danych), żaden model nie osiągnie pod względem dokładności (regularności) satysfakcjonującej wartości, to rozpatrywane będą modele drugiej warstwy.



Rys. 9. Ogólny schemat sieci neuronowej generującej kolejne modele o rosnącej złożoności w kolejnych warstwach.



Modele te przyjmą postać:

$$\begin{aligned}
 u_1 &= f(x_1, x_2, z_1, z_2) \\
 u_2 &= f(x_1, x_3, z_1, z_2) \\
 &\dots \\
 &\dots \\
 u_r &= f(x_{M-1}, x_M, z_{S-1}, z_S)
 \end{aligned}
 \tag{7}$$

gdzie: S – liczba modeli warstwy poprzedniej (pierwszej) wynikająca z możliwej do uzyskania liczby kombinacji po 2.

Podobnie jak poprzednio, jeżeli żaden z modeli drugiej warstwy nie zapewni oczekiwanej dokładności, to rozpatrywane będą kolejne warstwy sieci neuronowej z modelami utworzonymi przez dołączenie pary modeli warstwy poprzedniej.

Powstaną modele:

$$\begin{aligned}
 v_1 &= f(x_1, x_2, z_1, z_2, u_1, u_2) \\
 v_2 &= f(x_1, x_3, z_1, z_2, u_1, u_2) \\
 &\dots \\
 &\dots \\
 v_p &= f(x_{M-1}, x_M, z_{S-1}, z_S, u_{R-1}, u_R)
 \end{aligned}
 \tag{8}$$

gdzie: R – liczba modeli warstwy poprzedniej (drugiej) wynikająca z możliwej do uzyskania liczby kombinacji po 4.

Jeżeli także wśród tych modeli nie znajdzie się żaden zapewniający wystarczającą dokładność to rozpatrywana będzie czwarta warstwa itd.

W przedstawionych dalej przykładach obliczeniowych zaprezentowana będzie złożoność sieci sześciowarstwowej. Złożoność uzasadniona zresztą dokładnością modeli regresyjnych. Gdyby okazało się, że w którejkolwiek fazie testowana dokładność modelu byłaby wystarczająca, to zgodnie ze schematem na rys. 9 model ten byłby wykorzystany do predykcji i obliczenia w danym kroku zostałyby zakończone.

Stosując ulubioną notację graficzną Ivakhnenki (np. [50, 80]) schemat sieci o narastającej złożoności przedstawić można byłoby jak na bardziej szczegółowym rys. 10.

Istotą warstwy neuronowej jest generator modelu regresyjnego, kolejno  $z_{ij}$ ,  $u_{ijkl}$ ,  $v_{ijklmn}$  itd. Każda taka warstwa, to szereg równoprawnych neuronów korzystających z różnych wejść poprzedniej warstwy.

Jeżeli zbiór  $X$  będzie traktowany jako zbiór wejść do systemu to pierwszą warstwą sieci będzie generator  $Z_{ij}$ .

$Z_{ij}$  to zbiór modeli regresyjnych  $z_{ij} = a_0 + a_1x_i + a_2x_j$ ,  $i=1, 2, \dots, M$ ,  $j=2,3,\dots,M$ ,  $i \neq j$ , takich, które dla każdej pary argumentów zbioru  $X$  dają

minimum sumy kwadratów różnic pomiędzy danym  $z_{ij}$ , a wektorem zmiennej obserwowanej  $y$ . Współczynniki liczbowe  $a_0, a_1, a_2 \dots$  przedstawione w poszczególnych warstwach to różne liczby – tu dla uproszczenia przedstawione za pomocą tych samych symboli.

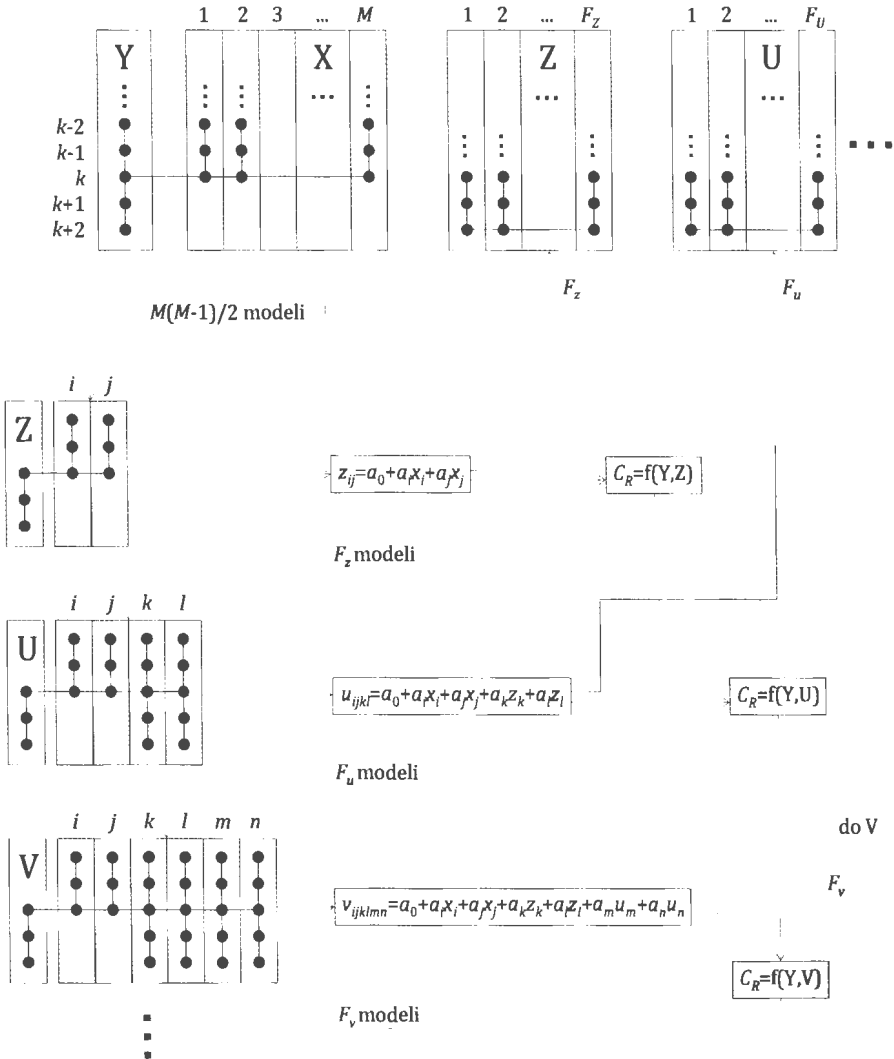
Cały schemat na rys. 10 podobnie jak na rys. 9 dotyczy tylko jednego okresu (jednej świecy walutowej, jednego kroku w szeregu czasowym). Pomimo tego, że krok ten wykonywany jest z uwzględnieniem zbioru wartości argumentu od chwili bieżącej  $k$  wstecz kolejno  $k-1, k-2, \dots$  to predykcja wykonywana jest na jeden lub kilka kroków wprzód (na rys.10 horyzont predykcyjny wynosi  $h_p=2$ ).

Poszukiwana jest wystarczająca złożoność modelu określona liczbą użytych do wygenerowania modelu warstw sieci. Autorzy większości tu cytowanych prac przewidują występowanie wyraźnego minimum funkcji dopasowania modelu w miarę wzrostu złożoności. Podobne rezultaty będą przedstawione także w dalszej części pracy jako efekt implementacji i badań algorytmu.

W modelu obliczania prognozy na schemacie rys. 10 występuje charakterystyczne przesunięcie szeregów czasowych branych pod uwagę przy obliczeniu modeli regresyjnych. Wzajemne przemieszczenie wektorów argumentów takie jak na rys. 4 i rys. 5 zostały tu zaznaczone w zbiorze  $\mathbf{X}$  w górnej części schematu. Tuż obok utworzono nowe zbiory argumentów z warstw przejściowych o licznosci  $F_z, F_u$  itd. Te wektory nowych argumentów nie muszą już być przesunięte wobec zmiennej obserwowanej  $Y_i$  bowiem są wynikiem przetwarzania danych i zawierają także elementy odpowiadające wartościom prognozowanym.

Zachowując więc, dla poprawności obliczeń macierzowych, identyczną długość wszystkich wektorów  $Y_i$  oraz wektorów argumentów wejściowych  $X_i$ , a następnie argumentów pochodnych (przejściowych)  $Z_i, U_i$  oraz  $V_i$  zastosowano, bez zagrożenia dla dostępu do danych, przesunięcie wektorów argumentów o  $h_p$  wprzód w zbiorach  $Z_i, U_i$  oraz  $V_i$ . Przesunięcia tego (wprzód) nie wolno jednak zrobić w zbiorze pierwotnych argumentów  $X$ . Zakłada się tu bowiem, że danych tych jeszcze nie ma w chwili estymacji wartości predykcyjnych  $y_{k+1}, y_{k+2}$ . Z kolei wartości argumentów w zbiorach  $Z_i, U_i \dots$  o indeksach  $k+1, k+2..$  po prostu są wygenerowane na podstawie estymowanych wartości argumentów pierwszego zbioru (pierwszej warstwy sieci neuronowej).

Niezwykle ważną cechą systemu predykcji w formie omawianej sieci neuronowej jest ograniczanie liczby neuronów (wyjść neuronów) wykorzystywanych w kolejnej warstwie. Ograniczenia te mają odzwierciedlenie w licznosciach zbiorów modeli  $F_z, F_u, F_v \dots$



Rys. 10. Schemat algorytmu kombinatorycznego GMDH z trzema warstwami neuronowymi i zbiorami modeli o licznosciach kolejno  $F_z$ ,  $F_u$ ,  $F_v$  jako wejsciami do kolejnych warstw.

Licznosci te są ustalane arbitralnie, aczkolwiek zawsze ze wskazaną dozą rozsądku, z uwzględnieniem określonych przez prowadzącego eksperyment barier natury technicznej, wynikającej ze wspomnianej już ograniczonej wydolności komputera. W dalszym przedstawionym przykładzie, przy licznosci zmiennych wejściowych rzędu  $20 \div 30$  ograniczono licznosci kolejnych zbiorów  $F_z$ ,  $F_u$  ... do  $10 \div 15$ . Były to zbiory wyse-

lekcionowanych najlepszych rozwiązań w poszczególnych warstwach sieci. Bez tego ograniczenia, przy przedstawionej na rys. 10 złożoności wyrażającej się stopniem zagnieżdżenia pętli, niemożliwe byłoby znalezienie optymalnych rozwiązań w rozsądnym czasie.

Określenie *najlepsze rozwiązanie poszczególnych warstw*, oznacza, że do kolejnych zbiorów o licznosciach  $F_z$ ,  $F_u$ ,  $F_v$  itd. przepuszczane były jedynie modele o kryterium (lokalnym) jakości lepszym od wyznaczonego (ponownie arbitralnie) przez analityka prowadzącego poszukiwania. Ta decyzja to kolejny przykład jak wielce heurystyczna jest ta metoda (w szczególności ten algorytm), jak wielkiej wyobraźni i kreatywności wymaga.

Gdyby np. kryterium selekcji modeli danej warstwy było zbyt radykalne, to nie osiągnięto by założonej licznosci  $F_u$ ,  $F_v$ ... Gdyby jednak było zbyt liberalne, to założona licznosc byłaby osiągnięta zbyt szybko poprzez wypełnienie zbioru modelami nie najwyższej jakości, a po prostu pierwszymi w sensie rozpatrywanych kombinacji indeksów neuronów spełniającymi minimum progu selekcji. Rozstrzygnięcie tego problemu – jak ustanowić progi selekcji dla poszczególnych warstw sieci neuronowej może być dokonane wyłącznie eksperymentalnie, gdyż zależy od struktury statystycznej macierzy danych.

Przy okazji tych dywagacji warto wspomnieć o genezie metody GMDH. U jej źródeł leży zasada masowej selekcji, którą Ivakhnenko porównuje do obserwowanych w przyrodzie naturalnych procesów ewolucji (w tym sensie metoda ta klasyfikuje się w szerokiej rodzinie algorytmów ewolucyjnych – przynajmniej w aspekcie inspiracji). Ivakhnenko i jego współpracownicy [73] zauważają bowiem konieczność pozostawienia pewnej swobody wyboru wśród najlepszych rozwiązań generowanych przez warstwę sieci neuronowej. Oznacza to także zachowanie (heurystyczne), że rezygnuje się z najlepszego rozwiązania uzyskiwanego np. po pierwszym kroku (warstwie) i dopuszcza się do poszukiwania lepszych rozwiązań w kolejnych warstwach np. po 10÷20 najlepszych modeli. Podobnie jak w naturalnym procesie ewolucji – wymieszanie najlepszych rozwiązań daje gwarancję wychwycenia najlepszych cech. Poza tym, problemem zasadniczym jak zwykle jest, czy przestrzeń danych wejściowych obejmuje wszystkie możliwe czynniki wpływające na zmienną obserwowaną. Z filozoficznego, ale na szczęście także statystycznego punktu widzenia okazuje się, że szereg czynników może być wyrażany pośrednio poprzez inne.

W matematyce takie relacje nazywane są występowaniem korelacji między zmiennymi. Jej siła może być niejednakowa, lecz jest to co

najmniej szansa, że korzystając z wielu zmiennych, uda się znaleźć model uwzględniający te zmienne najistotniejsze.

Rozpatrywana tu u podstaw metody zasada statystycznego (tu – regresyjnego) wyrażania zależności gwarantuje selekcję tych czynników poprzez bezpośrednie określenie ich siły w postaci wielkości współczynnika liczbowego w równaniu regresji.

Tak więc sieć przedstawioną na rys. 9 i rys. 10 można rzeczywiście przyrównać do pewnej naturalnej populacji poszukującej najlepszych, najlepiej przystosowujących się osobników. Odwieczna zasada „mieszanie krwi” także tutaj zdaje się być niezwykle skutecznym sposobem optymalizacji.

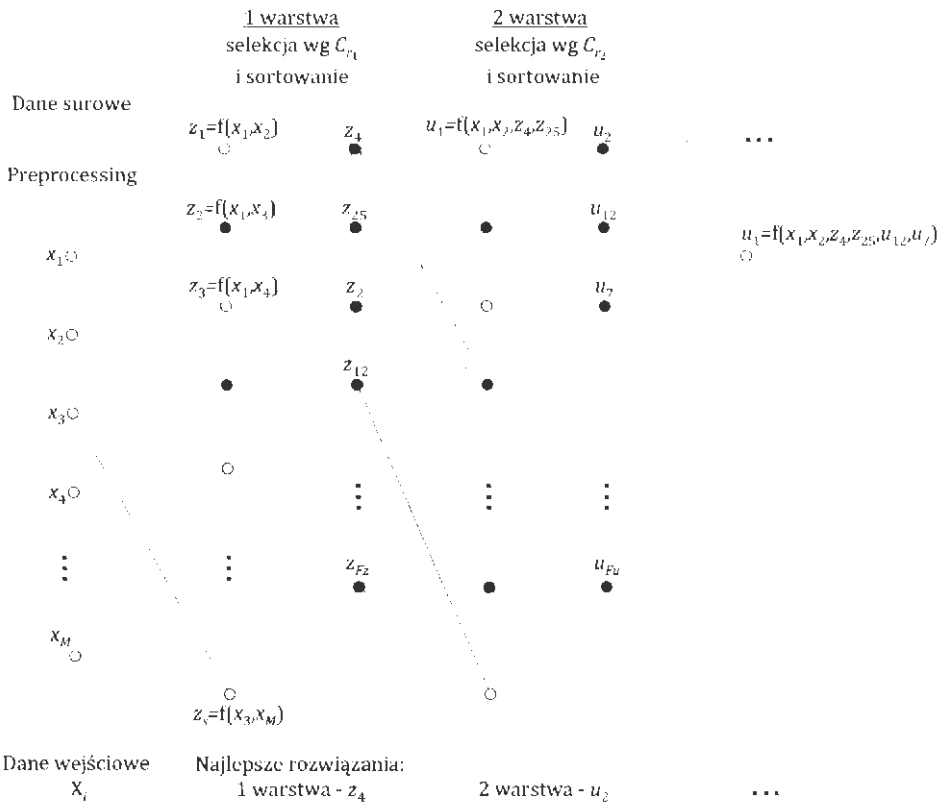
Selekcji najlepszych rozwiązań tworzących po każdej warstwie rodzaj elity nie widać dość wyraźnie na dotychczasowych schematach.

Lemke i Mueller [70] proponują nieco inne ujęcie tego problemu, przedstawione na rys.11, wzbogacone o warstwy sortowania neuronów i odrzucania neuronów słabych, generujących złe rozwiązania.

Ze schematu wynika, że specjalnie wybrany i przygotowany wektor danych wejściowych  $X_i$  stanowi źródło wejść do pierwszej warstwy neuronowej. Każdy  $z_i$  neuron tej warstwy korzysta z pary wejść  $x_j, x_k, k \neq j$ . Liczność neuronów w tej warstwie przed selekcją to  $M(M-1)/2$  (każdy z każdym, lecz nie z sobą). Dla każdej pary argumentów wejściowych obliczane jest równanie regresji w odniesieniu do prognozowanej zmiennej  $y$ . Obliczane są także współczynniki liniowego równania regresji, które minimalizują różnicę pomiędzy wyjściem neuronu a zmienną obserwowaną. Następnie różnica ta (kryterium rozbieżności) porównywana jest z ustalonym dla takiej warstwy kryterium selekcji. Jeżeli dokładność odwzorowania prognozowanego odcinka zmiennej  $y$  jest dobra (wartość kryterium jest mniejsza od ustalonego progu), to neuron kwalifikowany jest jako model należący do elity (zaznaczonej na rys. 11 na czarno) i będzie rozpatrywany jako wejście do następnej warstwy.

W drugiej warstwie w każdym neuronie wykorzystuje się dwa wejścia z pierwszej warstwy i dwa wejścia z sygnałów pierwotnych  $X_i$ . Utworzone w ten sposób równania regresji  $u_n=f(x_i, x_j, z_k, z_l)$  ponownie służą do obliczenia dokładności prognozy badanego odcinka zmiennej  $y$ . Ponownie najlepsze neurony osiągające dokładność lepszą niż założony próg selekcji  $C_{Rz}$  (inny, bardziej rygorystyczny niż w pierwszej warstwie) są kwalifikowane do elity drugiej warstwy (zaznaczone na czarno).

Rozdział 1. Zmodyfikowany kombinatoryczny algorytm metody GMDH...



Rys. 11. Schemat kojarzenie wejść do neuronów kolejnej warstwy oraz selekcji i sortowania modeli w warstwie wg lokalnego kryterium selekcji  $CR_k$  ( $k$  – numer warstwy). Na rysunku widać pozostawione po sortowaniu najlepsze (czarne) neurony jako wejścia do kolejnej warstwy.

W trzeciej warstwie każdy neuron wylicza poprawność prognozy bazując na sześciu wejściach: dwóch argumentach pierwotnych, dwóch z elity pierwszej warstwy i dwóch z elity drugiej warstwy.

Schemat ten powtarza się, aż do uzyskania satysfakcjonującej dokładności prognozy. Warunkiem stopu dla tego algorytmu jest wartość najlepszego neuronu (najlepszego równania regresji wielorakiej) w danej warstwie.

Otóż po zakończeniu obliczenia w wielu pętach wartości wyjść z neuronów w danej warstwie, neurony są sortowane wg kryterium dokładności  $C_R$ . Ułożenie ich w kolejności od najlepszego do ostatniego na pozycji  $F_u$ , potem  $F_u$  itd. symbolizują kolumny czarnych neuronów

w poszczególnych warstwach rys.11. Wszystkie neurony na miejscach gorszych niż  $F_z$ ,  $F_u$ ... po sortowaniu są odrzucane. Ograniczenie liczby najlepszych neuronów do kolejno  $F_z$ ,  $F_u$ ,  $F_v$  itd. ma swoje metodyczne konsekwencje. Jak już wcześniej wspomniano wybór tych ograniczeń, a także sposób ich użycia ma znaczenie natury zarówno technologicznej (wpływ na czas obliczeń) jak i metodycznej (decyduje pośrednio o dokładności predykcji poprzez włączenie lub eliminowanie wartościowych neuronów).

Problemu tego nie byłoby, gdyby zastosowana była zasada sortowania  $F_z$ ,  $F_u$ ... najlepszych neuronów spośród wszystkich modeli danej warstwy. W praktyce okazało się to jednak niemożliwe (czas obliczeń) i zastąpiono tę zasadę heurystyczną regułą znajdowania  $F_z$ ,  $F_u$  itd. neuronów lepszych niż pewien ustalony próg  $C_{Rz}$ ,  $C_{Ru}$  ... po czym następowało przerwanie wszystkich pętli poszukiwań i przejście do sortowania wybranych modeli. Ustalenie wektora progów  $C_{Rk}=[C_{Rz}, C_{Ru}, C_{Rv} \dots]$  wymagało zastosowania najstarszej metody – prób i błędów, gdyż nie ma racjonalnych przesłanek umożliwiających określenie tych wartości na podstawie statystycznych cech macierzy danych. Wektor  $C_{Rk}$  wykazał istotny wpływ na skuteczność algorytmu, gdyż zbyt małe wartości poszczególnych progów powodowały, że z trudem lub w ogóle nie znajdowano w danej warstwie modeli spełniających warunek  $\delta^i_k < C_{Rk}$ , gdzie  $\delta^i_k$  to kryterium rozbieżności i-go modelu w k-tej warstwie. Oznacza to wydłużenie czasu przeszukiwań i obliczanie wszystkich możliwych modeli regresji. Przy liczbie argumentów wejściowych np.  $M=30$  i liczbach  $F_z=F_u=F_v=\dots=10$  w pierwszej warstwie będzie do obliczenia ponad 400 modeli regresji, w drugiej owe  $400 \times 10^2$  a w trzeciej  $400 \times 10^4$  tj.  $4 \cdot 10^6$ . W rozpatrywanym dalej przykładzie rozpatrzono 6 warstw, czyli należałoby obliczyć  $4 \cdot 10^{12}$  modeli.

To liczba obliczeń modeli regresyjnych określona przy założeniu, że w każdej warstwie przybywa równo 10 nowych modeli, gdyby jednak rozpatrywać wszystkie nowe bez ograniczeń, to realizacja obliczeń jest wielokrotnie bardziej niewykonywalna – to klasyczne przekleństwo wielowymiarowości. Nie wspomina się tu nawet o długości wektora argumentów  $X_i$  (liczbie  $K$  rozpatrywanych wierszy okna z  $N$  wierszy macierzy danych), której zwiększenie oznacza nieliniowy wzrost kłopotów z odwracaniem macierzy przy obliczaniu współczynników równania regresji.

Z rozważań tych wynika oczywista konieczność przerywania pętli generowania nowych modeli regresyjnych po znalezieniu pierwszych  $F_z$ ,  $F_u$  ... spełniających warunek dostatecznej dokładności.

Na drugim biegunie tej analizy niech będzie rozpatrzona sytuacja, gdy wektor ograniczeń będzie  $C_{Rk}$  jest bardzo liberalny – łatwy do osiągnięcia w poszczególnych warstwach. Oznacza to niemal natychmiastowe osiągnięcie zakładanych liczebności zbiorów modeli  $F_z$ ,  $F_u$  ... bez gwarancji jednak, że w ogromnym zbiorze pozostałych nie rozpatrzonych znajdzie się wiele wartościowych i niezauważonych rozwiązań.

Tak więc jedynym rozsądnym postępowaniem jest poszukiwanie kompromisu pomiędzy czasem trwania obliczeń a dokładnością metody – zresztą jak zazwyczaj w wielu podobnych problemach optymalizacji w wielowymiarowych przestrzeniach danych.

By uzmysłowić czytelnikowi istotność tego kompromisu należy wspomnieć, że w wielu przedstawianych dalej eksperymentach prowadzonych na współczesnych komputerach dla uzyskania satysfakcjonującego rezultatu konieczne było wykonanie symulacji trwającej niekiedy kilkanaście godzin, a w skrajnych przypadkach nawet kilkadziesiąt.

### **1.3. Planowanie struktury sieci neuronowej**

Opisany algorytm jest dość złożony zarówno z uwagi na zmieniającą się strukturę sieci neuronowej (obliczeniowej) jak i praktycznego wykorzystania lokalnego rezultatu żmudnych wielowarstwowych obliczeń. Przed wyjaśnieniem możliwości owego praktycznego wykorzystania należy jeszcze wyjaśnić bardziej szczegółowo koncepcję samego algorytmu generowania predykcji, lub koncepcję obliczania pewnych zmiennych niezbędnych do prognozowania kierunku i wielkości zmian w rozpatrywanym horyzoncie czasowym. Wartości tych zmiennych będą zapewne wprowadzane do pewnego systemu inwestycyjnego (na rynkach walutowych zwanego systemem tradingowym lub systemem handlu), opisanego i testowanego w kolejnych rozdziałach. Z tego punktu widzenia rozpatrywany tu system wypracowania rekomendacji o predykcji jest systemem wspomagania decyzji dla systemu inwestycyjnego realizującego zadania np. na platformie brokerskiej.

Pierwsze z tych dwu zadań – zadanie wypracowania prognozy wykonuje się za pomocą specjalnie spreparowanej złożonej sieci neuronowej, której pierwszy fragment przedstawiony został na rys. 10. Sieć tę (sieci) można przedstawić jako podwójnie wielowarstwową (niekiedy tak się ją nazywa [50, 73]).

W pierwszej wewnętrznej sieci (sieciach), realizowanej na każdym kolejnym stopniu sieci zewnętrznej, dokonywana jest selekcja najlepszych modeli regresyjnych budowanych na strukturze wybranej, ogra-



nicznej w danym stopniu sieci zewnętrznej. Sieć zewnętrzna z kolei, wykorzystując najlepsze modele generowane przez sieci wewnętrzne wszystkich stopni odpowiada za wybór stopnia złożoności (stopnia lub warstwy sieci zewnętrznej).

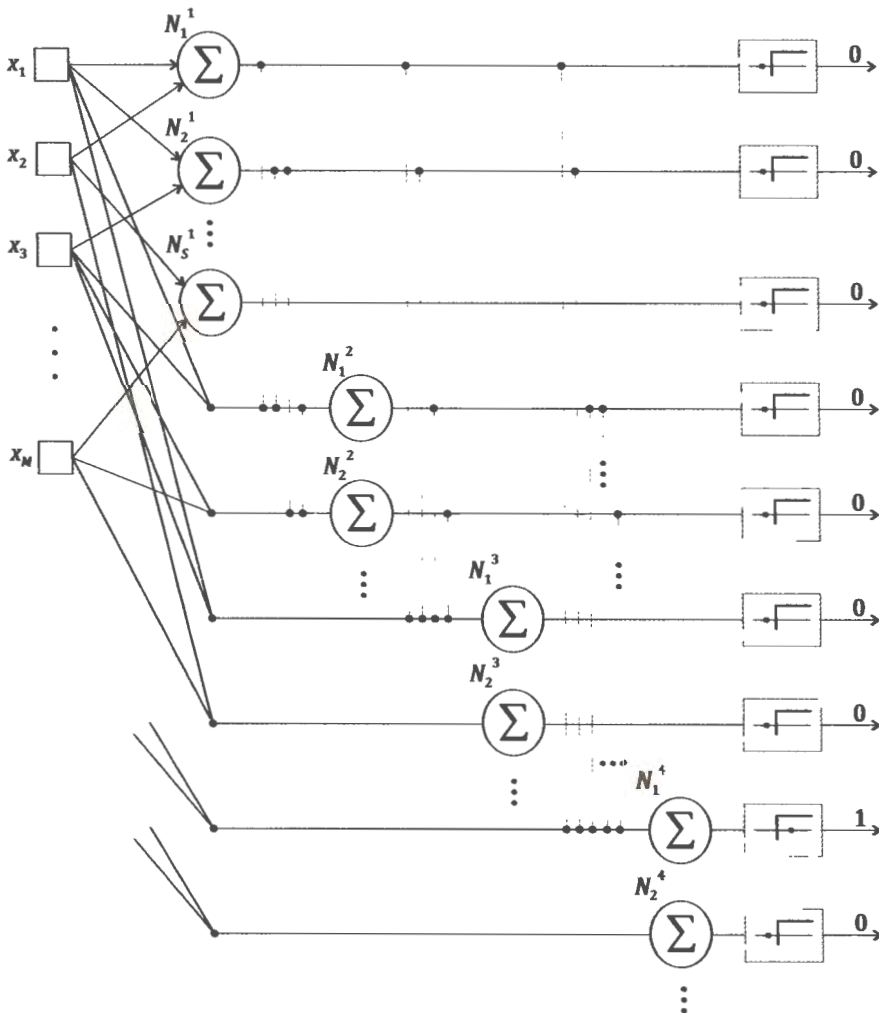
Tak więc oczekiwanym rezultatem szczegółowych obliczeń poprzez modele regresyjne sieci wewnętrznych jest struktura modelu (dobór argumentów równania regresji i ich stopnia). Na wyjściu z każdej sieci wewnętrznej danego stopnia jest więc zbiór najlepszych (w sensie wybranego kryterium) modeli o tej samej złożoności w sensie szeregu Gabora-Kołmogorowa.

Druga zewnętrzna sieć służy do porównywania skuteczności sieci wewnętrznych. Logicznym rozwiązaniem będzie takie porównywanie pomiędzy sobą rezultatów (kryteriów rozbieżności) poszczególnych sieci wewnętrznych, aż od pewnego rosnącego stopnia złożoności nie będzie zauważona poprawa kryterium (lub nastąpi jego pogorszenie).

Tak więc ta zewnętrzna sieć, ten najbardziej ogólny obraz sieci neuronowej, będzie to taki system obliczeniowy, dla którego na wejściu występuje wektor (zbiór) modeli najlepszych, obliczonych w sieciach wewnętrznych, a na wyjściu kilka umownych „neuronów” (tyle, ile przewidywanych stopni złożoności struktury sieci). Funkcje aktywacji tych neuronów mają charakter unipolarny, z zastrzeżeniem, że w danym kroku aktywny może być tylko jeden neuron (rys. 12) – wybierany jest do prognozy jeden najlepszy model o optymalnej w danym kroku złożoności. Ostatnią warstwą tego złożonego systemu obliczeniowego jest więc warstwa perceptronów z wyjściami binarnymi, takimi, których suma równa się 1 (tylko jeden model będzie wykorzystany w rzeczywistym prognozowaniu – ten z najniższą wartością kryterium rozbieżności). Z powyższego opisu wynika, że rozpatrywana „podwójnie wielowarstwowa” sieć ma charakter jednokierunkowy. Jej sprzężenia zwrotne występują w zawaolowanej formie tylko w równaniach regresji, które obejmują  $K$  wierszy macierzy danych wstecz. Modele (warstwy wewnętrznej) są więc syntezą najlepszego dopasowania, czyli ewentualnie takiego, które przy innej strukturze sieci neuronowej byłoby uzyskiwane krok po kroku.

Schemat (rys. 12) dotyczy drugiej, zewnętrznej warstwy sieci i przedstawia naturalną powtarzalność algorytmu - dla istniejącego na podstawie danych historycznych wektora zmiennej wyjściowej obliczone są opisujące go modele, z nich wybierane są najlepsze wg pewnego zewnętrznego kryterium. Następnie budowane są kolejne modele o wyższej złożoności w nadziei na lepsze przybliżenie prognostyczne.

Proces ten jest powtarzany w kolejnych warstwach tak długo, dopóki nie nastąpi pogorszenie jakości prognozy, lub jej poprawa nie będzie miała praktycznego znaczenia.



Rys. 12. Ogólny schemat jednokierunkowej wielowarstwowej sieci neuronowej, której wyjściem jest wskazanie stopnia złożoności modelu predykcyjnego. Na wejściach kolejnych warstw występują wyjścia wszystkich warstw poprzednich tak długo, aż wystąpi akceptowalna zgodność modelu (na wyjściu z ostatniego perceptronu pojawi się 1). Ten warunek na schemacie spełnia przedostatni neuron.

Dla uzmysłowienia różnic w jakości tej prognozy przeprowadzono badanie w różnych miejscach szeregu czasowego (przemieszczanie wzdłuż szeregu czasowego nazywane będzie tu wykonaniem kolejnego kroku – rozważeniem kolejnej świecy). Badania te przeprowadzane będą także w różnych warstwach sieci – o rosnącej złożoności. Złożoność sieci przedstawionej na rys. 12 polega nie tylko na widocznej liczbie warstw. Na diagramie zaznaczono 4 warstwy, poza wejściową – warstwa to kolejna kolumna sumatorów - neuronów. Złożoność ta polega także na potencjalnie dużej liczbie neuronów w każdej warstwie.

W klasyfikacji przyjętej w literaturze sztucznych sieci neuronowych przedstawiony przykład to stosunkowo prosta sieć perceptronowa z unipolarnymi funkcjami aktywacji wyjściu, bez sprzężenia zwrotnego. Ta jednokierunkowość sieci wynika z ogólnego założenia modelu predykcji – postawiona jest hipoteza o pewnej ograniczonej w czasie trwałości struktury modelu predykcyjnego. Sprawdzony w danym krótkim odcinku czasu model jest przez chwile wykorzystywany do wygenerowania prognozy zachowania się funkcji wyjściowej  $\hat{y}$  i następnie po weryfikacji struktury modelu, w kolejnym kroku, zmieniany. Idea ta ujawnia znaczną złożoność przedstawionej sieci w jej wnętrzu.

Każdy z oznaczonych znakiem „ $\Sigma$ ” neuron to złożona inteligentna sieć „niższego rzędu”, ucząca się na podstawie lokalnego modelu regresyjnego i wykorzystująca najlepsze rozwiązania (modelu) poprzedniej warstwy. Na rys. 12 nie są zaznaczone procesy selekcji wejść opisane wcześniej na rys. 11. Schemat (rys. 12) jest jednak syntezą dotychczasowych rozważań i może być najważniejszym uogólnieniem algorytmu umożliwiającym jego implementację.

## 1.4. Koncepcja algorytmów predykcyjnych

Bazując na schemacie sieci neuronowej z rys. 12 dokonano implementacji dwóch podobnych algorytmów należących do rozpatrywanej metody GMDH, o zupełnie różnych celach. Algorytmy te, nazwane CL1 i CL2 przedstawione są w kontekście różnic w korzystaniu z zasobów danych historycznych.

### 1.4.1. Charakterystyka algorytmu estymacji potencjału inwestycyjnego CL1

Pierwszy algorytm CL1 (autorska nazwa od „kombinatoryczny liniowy”), to algorytm, w którym neurony pierwszej warstwy „rozważa-

ją” liniowe modele regresyjne zbudowane z kolejno rozpatrywanych par argumentów ze zbioru  $\mathbf{X}_i$   $i = 1, 2, \dots, M$ . Zbiór  $\mathbf{X}_i$  to zbiór wektorów argumentów  $\mathbf{x}_{ik}$  o jednakowej liczności, ale o różnym przesunięciu  $k = 0, 1, 2, \dots$  w czasie względem chwili bieżącej ( $k=0$  oznacza wektor bez przesunięcia wstecz). Obserwowany wektor zmiennej wyjściowej  $y_k$ ,  $k = 1, 2, \dots, K$  o tej samej liczności był z założenia przesunięty do przodu o  $h_p$  (horyzont predykcyjny) okresów. W teże pierwszej warstwie kolejne neurony badały pary argumentów o długości  $L_w = \lfloor x_i \rfloor$  (czyli w tym przypadku o liczności  $K - \text{szerokości wcześniej rozpatrywanego okna}$ ) poszukując największej zbieżności z wektorem  $y_i$ , poprzez dobór wektora współczynników liczbowych  $\mathbf{A}_i = [a_{i0}, a_{i1}, a_{i2}]$ , takiego, że

$$\mathbf{A}_i = \mathbf{X}_i \setminus \mathbf{Y}_i \quad (9)$$

gdzie  $\mathbf{A}_i$  jest wektorem o wymiarze  $[1 \times 3]$ ,  $\mathbf{X}_i$  macierzą  $[3 \times L_w]$  a  $\mathbf{Y}_i$  wektorem  $[L_w \times 1]$ .

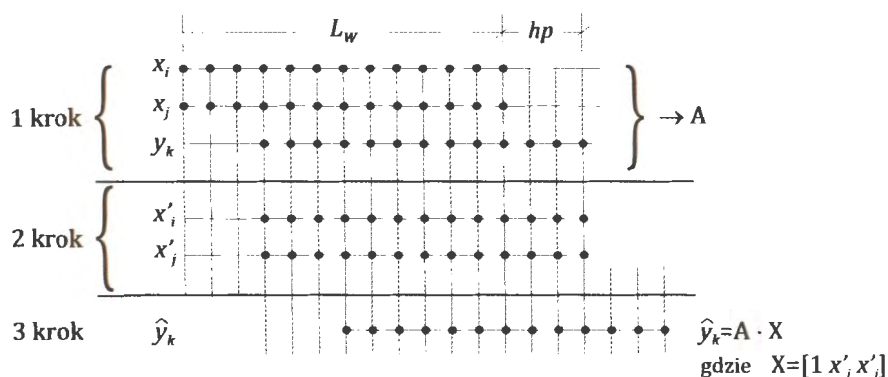
Równanie (9) to inaczej mówiąc rozwiązane równania  $\mathbf{Y}_i = \mathbf{A} \cdot \mathbf{X}_i$  dla  $i=1,2,\dots,s$  (liczba możliwych kombinacji par  $(x_i, x_j)$ ). Ważne jest zrozumienie wymiarowości czynników tego równania. W przypadku rozważania par argumentów  $(x_i, x_j)$  wektor współczynników  $\mathbf{A}_i$  ma wymiar  $[1 \times 3]$ , wektor wyniku  $\mathbf{Y}_i$  ma wymiar  $[1 \times L_w]$  a macierz danych  $\mathbf{X}_i$  ma wymiar  $[3 \times L_w]$ . W równaniu (9) występuje operator dzielenia lewostronnego macierzy „\”.

W algorytmie tym wykorzystuje się informację zgromadzoną w macierzy danych w taki sposób, że ostatni wyraz wektora  $\mathbf{Y}_i$  jest doprowadzony do chwili bieżącej, a najbardziej do przodu przesunięty wektor  $x_i$  ze zbioru  $\mathbf{X}_i$  jest przesunięty wstecz do  $\mathbf{Y}_i$  co najmniej o  $h_p$ . W ten sposób nie są wykorzystane informacje o wartościach argumentów  $x_i$  zgromadzone w  $h_p$  ostatnich chwilach dokonywania obserwacji z wyjątkiem  $\mathbf{Y}_i$ . Dzięki temu jest jednak możliwe przeprowadzenie estymacji predykcji  $\hat{y}_i$  na ostatnich  $h_p$  krokach.

Kolejną czynnością przeprowadzoną w implementacji tego algorytmu jest wykorzystanie wektora współczynników  $\mathbf{A}_i$  (9) dla nowych danych  $\mathbf{X}_i$  uzyskanych poprzez przesunięcie całej macierzy danych o  $h_p$  kroków wprzód. W ten sposób przesunięty wprzód wektor zmiennej obserwowanej  $y_i$  wchodzi w obszar wyników, których w sytuacji symulowanej prognozy jeszcze nie ma. Jeżeli jednak korzysta się z danych historycznych, to można w ten sposób sprawdzić, który z kolejno rozpatrywanych modeli regresyjnych najlepiej sprawdza się na precyzyjnym odcinku.

Postępowanie takie ma znaczenie istotne dla sprawdzenia potencjału jakości predykcyjnej algorytmu. Bazuje się tu bowiem na współ-

czynnikach wielomianów pierwszej warstwy (tu zbudowanych w postaci liniowych równań z dwiema zmiennymi  $x_i, x_j$  oraz wyrazem wolnym) określonych w przeszłości ( $h_p$  kroków wstecz od chwili bieżącej) i wykorzystanych z nowymi wejściami do neuronu. Tymi wejściami są oczywiście wektory  $x_i, x_j$  przesunięte o  $h_p$  wprzód, czyli aż do chwili bieżącej. Sytuację tą można zilustrować dodatkowo schematem na rys. 13.



Rys. 13. Schemat objaśniający kolejność obliczania modelu prognostycznego  $\hat{y}$  na podstawie współczynników  $A$  wielomianu obliczonych w pierwszym kroku i wykorzystanych w równaniu predykcji z nowymi danymi  $x'_i$  i  $x'_j$ .

W ten sposób w pierwszej wartości sieci neuronowej powstanie wiele modeli  $\hat{y}_k$   $k=1,2, \dots, M(M-1)/2$ , które wszystkie, chociaż posiadają wysunięty wprzód o  $h_p$  kroków odcinek predykcyjny, to jednak zbudowane są na posiadanej wiedzy – o wektorach  $x'_i, x'_j$  oraz o współczynnikach z przeszłości  $A$ .

Jeżeli teraz z kolei przyjmie się, że dla celów nie predykcyjnych, lecz dla sprawdzenia potencjału algorytmu (jego możliwości uzyskiwania jak najlepszych rezultatów wg wybranego kryterium), znane są rzeczywiste wartości zmiennej  $y$  na odcinku predykcyjnym (o długości  $h_p$ ), to spośród wszystkich modeli  $\hat{y}_k$  pierwszej warstwy można wybrać te, które wg ustalonego kryterium zewnętrznego są najlepsze.

Na podstawie kombinacji tych najlepszych modeli pierwszej warstwy można z kolei zbudować modele drugiej warstwy i ponownie znaleźć wśród nich najlepsze wg tego samego kryterium w kolejnych warstwach.

Ostatecznie można w ten sposób znaleźć model zapewniający minimum kryterium rozbieżności. Wielce interesujące będzie jak duża będzie złożoność tego modelu ( w sensie liczby wyrazów wielomianu -

lub numeru warstwy w sieci neuronowej) oraz, być może, jeszcze ważniejsza będzie odpowiedź na pytanie, czy zmienia się krok po kroku struktura tego złożonego modelu. Krok po kroku, czyli dla kolejnych chwil bieżących wzdłuż rozpatrywanego szeregu czasowego. Niezmiernie istotną cechą tak opisanego algorytmu (także następnego) jest zdolność do samoorganizacji (*self-organizing*). Algorytm w każdym kroku jednokierunkowej sieci neuronowej sam wybiera zarówno najlepszą strukturę w sensie złożoności (stopnia wielomianu), jak i komponentów składających się na tę strukturę – w sensie zestawu użytych argumentów. Oczywiście ten element sztucznej inteligencji w algorytmie nie jest samorodny, jest wyrazem ludzkiej jednak inteligencji zapisanej w formie nakazu wykonania szeregu określonych operacji w określonej kolejności, w tym także nakazu podejmowania decyzji wg określonych przez człowieka reguł.

„Sztuczne” (czy raczej maszynowe) jest tu zwielokrotnienie możliwości obliczeniowych człowieka. Maszyny, które swoją powolnością obliczeń wywoływały zniecierpliwienie prof. Ivakhnenki już 40 lat temu [50], ciągle są także dziś zbyt wolne dla rozważanego tu algorytmu. Dają jednak przynajmniej przedsmak sukcesu modelowania w pełni rozwiniętego.

#### 1.4.2. Charakterystyka algorytmu estymacji predykcji CL2

Drugi algorytm nazwany przez autora CL2 to typowy algorytm predykcyjny, spełniający zasadnicze oczekiwania badacza. O ile pierwszy algorytm CL1 dawał wyobrażenie jak duża może być precyzja prognozy, nie zapewniając jej jednakże (gdyby okazało się, że struktura modelu jest zmienna), to drugi algorytm CL2 wykorzystując wiedzę z CL1, prognozę wykonuje.

Prognoza na odcinku  $h_p$  okresów wprzód jest realizowana przy najczęściej stosowanym w takich przypadkach założeniu, że wybrany model predykcyjny będzie adekwatny do zmian ewolucyjnego środowiska chociaż przez chwilę. Ten wybrany model może być rezultatem wielowarstwowego przetwarzania w sieci neuronowej i może być co chwilę (co krok) inny. Oznacza to, że może posiadać inną strukturę w sensie liczby wyrazów wielomianu, a także inne współczynniki liczbowe, obliczane co krok. W algorytmie CL2 powstanie też intrygujące zadanie nakładania się kilku prognoz równocześnie na siebie w związku z długością odcinka predykcyjnego  $h_p$  większego niż 1 (choć interesujący jest także przypadek predykcji wyłącznie o jeden krok wprzód np. [66, 83]).

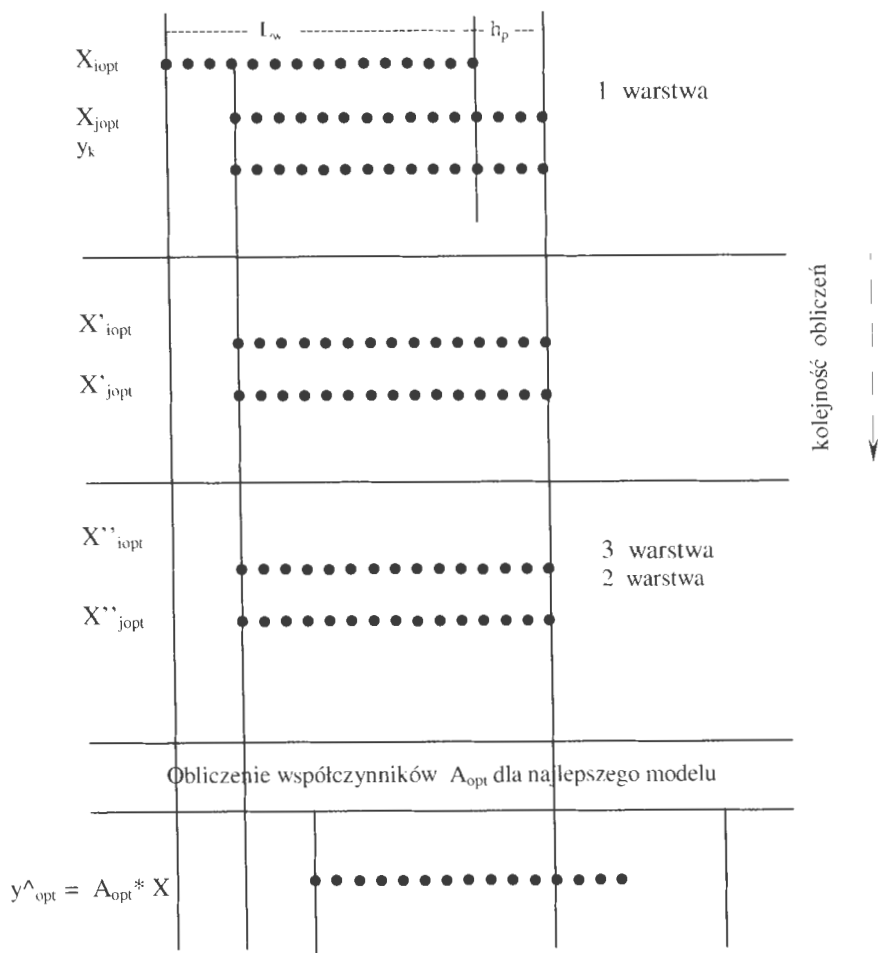
Zasadnicza różnica między CL1 i CL2 polega więc na tym, że w CL1 poszukiwana jest struktura najlepszego modelu predykcyjnego i możliwa do osiągnięcia dokładność predykcji przy założeniu, że znana (z danych historycznych) jest informacja o zmianach w przyszłości („przeszłej” przyszłości) na odcinku  $h_p$ . W modelu tym można także przeprowadzić badanie możliwości osiągnięcia dużej regularności modelu cofając się w tablicy danych o  $h_p$  wierszy i badając kryterium  $C_R$  wyłącznie na danych dostępnych w chwili bieżącej (wg zasady: „co by było, gdyby”). Badania takie pozwalają wówczas przewidzieć strukturę systemu predykcyjnego i jego potencjalne możliwości, nie pozwalają jednak na przeprowadzenie bezpośredniej prognozy.

W modelu CL2 realizowana jest prognoza bez wiedzy o przyszłości, na podstawie ocenionej w algorytmie złożoności modelu (liczby warstw sieci neuronowej) i na podstawie danych wejściowych zbioru  $X$  wykorzystanych aż do chwili bieżącej. W ostatecznym modelu prognostycznym wykorzystuje się też najlepsze wektory w poszczególnych warstwach  $x_{iopt}$ ,  $x_{jopt}$  (w warstwie pierwszej),  $x_{iopt}$ ,  $x_{jopt}$  (w warstwie drugiej),  $x_{iopt}$ ,  $x_{jopt}$  (w warstwie trzeciej) itd.

Do porównania obu algorytmów CL1 i CL2 schemat obliczenia prognozy wg tego drugiego przedstawiono na rys.14.

Na tym schemacie przedstawiającym sytuację dla  $h_p=3$  widać wektor prognozy  $\hat{y}_{opt}$  wyprzedzający chwilę bieżącą o trzy okresy.

Podsumowując – algorytm CL1 określa w każdym kolejnym kroku w szeregu czasowym ile można zyskać przy najlepszej możliwej do przeprowadzenia prognozie, gdyby zastosowano najlepszy z rozpatrywanych modeli, algorytm CL2 przeprowadza tę prognozę wg logicznego (niekoniecznie trafnego) założenia, że ostatnio najlepszy model, będzie poprawny jeszcze przez chwilę, np. jeszcze przez jeden-dwa okresy predykcyjne.



Rys. 14. Schemat objaśniający sposób obliczania prognozy wg algorytmu CL2. Współczynniki  $A_{opt}$  wielomianu prognostycznego obliczone są na podstawie znanych danych pierwszej warstwy oraz danych przetworzonych kolejnych warstw.





## Podsumowanie

Metoda GMDH sprzężona z mechanizmami półautomatycznego obrotu na platformach brokerskich tworzy bardzo interesujący i obiecujący system transakcyjny. Przykłady użycia w różnych obszarach rynków finansowych wskazują na jej uniwersalność w aspekcie miejsca użycia, ale też na ograniczone możliwości wykorzystania. Podstawowym ograniczeniem metody jest niska moc obliczeniowa współczesnych maszyn. W metodzie konieczne jest bowiem transformowanie przestrzeni predyktorów do nowych wielowymiarowych przestrzeni pochodnych. To z kolei wymusza czasochłonne przeszukiwanie przestrzeni zmiennych objaśniających w celu określenia najbardziej obiecujących modeli predykcyjnych. Dodatkowo, eksperymentalnie sprawdzono, że budowane modele tylko przez chwilę są zdolne do opisanie przyszłości i wymagają ustawicznej adaptacji, zarówno poprzez zmianę parametrów modelu (tu wielomianowego) jak też poprzez zmianę struktury tego modelu. Stąd, trudne wyzwanie obliczeniowe – sedno tzw. inteligencji obliczeniowej, które sprowadzić można do poszukiwania kompromisu pomiędzy dokładnością i pracochłonnością algorytmu. Można tu także skonstatować, że przydatność metody daje się stwierdzić wyłącznie poprzez obliczenia, poprzez eksperymenty na fizycznej rzeczywistości reprezentowanej przez macierze predyktorów. Przyszłość tej metodologii – chyba w wykorzystaniu mocy rozproszonych systemów obliczeniowych zgodnie z ideą *cloud computing*, ale także w poszukiwaniu nowych specyficznych dla metody algorytmów heurystycznego przeszukiwania przestrzeni zmiennych objaśniających.

Zapotrzebowanie na moc obliczeniową i lepsze algorytmu przeszukiwania wynika z przekonania, że warto budować modele o dużej złożoności. W rozdziale 1 widać było związek pomiędzy dokładnością predykcji a złożonością modelu.

Zaimplementowana przez autora do sześciu warstw sieć obliczeniowa z trudem daje się wyjaśnić do poziomu czwartej warstwy tak, by można było odtworzyć pierwotne współczynniki równań regresji (te, związane z pierwotnymi, nieprzetworzonymi danymi) dla najlepszych modeli. Z uwagi na złożoność tego opisu przeniesiony on został do dodatku, gdzie zajął ok. 5 stron. Opis warstwy piątej byłby już 4-5 krotnie obszerniejszy, a warstwy szóstej zająłby już kilkadziesiąt stron. W praktycznej implementacji tych algorytmów autor nie naśladował Aleksieja Ivakhnenki, który zazwyczaj przedstawiał modele wielomianowe w formie

uwzględniającej pierwotne predyktory z wyższymi stopniami złożoności (np. jako człony kwadratowe, sześciennie itd.). Ponieważ eksperymentalnie w niniejszej pracy dowiedziono, że model jest poprawny tylko przez przysłowiową chwilę, to dociekanie, które z danych wejściowych mają wpływ na ciągle zmieniającą się strukturę modelu nie było w istocie ważne. Ważna była zdolność systemu do samoorganizacji, do ciągłej zmiany struktury modelu i parametrów. Adaptacyjność tę wielokrotnie uzyskiwano i potwierdzono empirycznie jej zasadność i skuteczność.

Intencją było umożliwienie implementacji tej metody w dowolnym środowisku programowym. Środowiskiem wybranym przez autora jest Matlab. Zainteresowany czytelnik może wybrać dowolne inne o podobnych lub lepszych funkcjonalnościach.

Warto zauważyć, że obliczenia te dają się zrównoleglić z uwagi na wielokrotnie zagnieżdżaną pętlę. To istotny aspekt rozpatrywanego ograniczenia. Współczesne metody zrównoleglenia obliczeń uwzględniają możliwość tworzenia farm komputerów, czy zastosowania szybkich procesorów graficznych w celach naukowych. To jeden z kierunków rozwoju rozpatrywanej metody GMDH. Drugi, to spokojne oczekiwanie na spełnienie się prawa Moora „gwarantującego” wykładniczy rozwój mocy obliczeniowej.

Zupełnie inne podejście, to użycie intelektu zamiast maszyn. To zapewne, na razie, dopóty górujemy nad maszynami, najlepsze przyspieszenie wzrostu skuteczności metody. To poszukiwanie usprawnień metody GMDH w obszarze poszukiwania lepszych baz dla reprezentacji modeli i odpowiedniejszych kryteriów, poszukiwanie lepszych metod przeszukiwania przestrzeni predyktorów.

W pracy stosowano kilka kryteriów jakości prognoz charakterystycznych dla praktycznego handlu. Był więc stosowany w analizie portfelowej wskaźnik Sharpe’a czy Sortino, był także wskaźnik uwzględniający awersję do maksymalnego obsunięcia kapitału. W fazach testowania stosowane były w przypadku poszukiwania modeli regresyjnych nie tylko tradycyjne kryteria metody najmniejszych kwadratów, ale także minimaksowe. Zwłaszcza kryteria związane z awersją do maksymalnego obsunięcia są, zdaniem autora, bardziej naturalne i związane z osobowością inwestora niż kryteria stosowane w modelach Markowitza czy Sharpe’a. Oczywiście zawsze lepiej jest znaleźć model, który przy danym poziomie ryzyka pozwoli na osiągnięciu wyższego niż niższego zysku. Zawsze lepiej jest osiągnąć założony zysk przy niższym niż wyższym poziomie ryzyka. Jednakże jak zmierzyć tę szczególną cechę inwestora, który wybiera poziom ryzyka „nierozsądnie”? Dlaczego istnieją na rynku tysiące alternatyw wobec bezpiecznych obligacji?

Praca jest w jakiejś mierze wyrazem fascynacji autora dorobkiem zmarłego niedawno (w 2008 roku) ukraińskiego wybitnego uczonego prof. Aleksieja Ivakhnenki. Ivakhnenko nigdy nie koncentrował się na zadaniach predykcji na rynkach finansowych. Z przeglądu licznych źródeł tego Autora wynika, że do kilku publikacji dotyczących tej problematyki namówili Go raczej niemieccy partnerzy – Johann-Adolf Mueller i Frank Lemke. Kontakt ten nie przyniósł spektakularnych praktycznych rezultatów, ale nie zniechęcił też (np. Franka Lemkego, z którym autor niniejszej pracy ma bezpośredni kontakt) do kontynuacji usprawnień metody. Nigdy jednak nie był, w tej grupie uczonych, rozpatrywany aspekt handlu automatycznego, częstotliwego, o zupełnie nieoczekiwanych cechach. Wprowadzenie infobotów do zarządzania kontami klientów to, w przekonaniu autora, niedaleka przyszłość. To nieodłączny atrybut członka przyszłego społeczeństwa informacyjnego, który będzie poszukiwał specyficznego, outsourcingowego wsparcia osobistych finansów. To, z dużą dozą prawdopodobieństwa, będzie typowe zachowanie drobnego inwestora zniecierpliwionego natarczywością i nieodpowiedzialnością banków, zwłaszcza po fali bankructw tych ostatnich w okresie trwającego kryzysu finansowego.

Handel tradycyjny i w pełni automatyczny to dwa różne światy – w aspekcie teorii podejmowania decyzji, strategii matematycznych, psychologii inwestora, szybkości działania, wpływu technologii komunikacyjnych na teorie racjonalnego inwestowania (i zarządzania ryzykiem). Zastosowanie tradycyjnych koncepcji dla handlu automatycznego może być zupełnie bezzasadne. To tak, jakby rozpatrywać dwa środki komunikacji – wóz konny i samolot. Przyczepiwszy skrzydła z boków wozu konnego, konstatować ze zdziwieniem, że nie unosi się.

Rozwiązaniem dla wozu jest zmiana koni. Na bardzo szybkie.



## Bibliografia

1. Banaszak Z., Józefczyk J. *Towards dedicated decision support tools*. Applied Computer Science vo. 1 no 1 2005, ATH, Bielsko-Biała 2005, pp. 161-181.
2. Bargiela A., Pedrycz W. *Recursive Information Granulation: aggregation and interpretation issues*. IEEE Transactions on Systems, Man and Cybernetics Part B – Cybernetics 33 (1), 96-112.
3. Borawski M. *Analysis of Properties of Mean Value - Pseudo -Variance Representation*. Polish Journal of Environmental Studies vol.17 No 3b, 2008, p. 39-43.
4. Breiman L., Friedman J.H. , Olshen R.A., Stone C.J., *Classification and Regression Trees*, Wadsworth International Group, Monterey, 1984.
5. Brock W., Lakonishok j., LeBaron B. *Simple technical trading rules and stochastic properties of stock returns*. Journal of Finance 47(1992), 1731-1764.
6. Brown G., Wyatt J., Harris R., Yao X. *Diversity creation methods: A Survey and Categorisation*. University of Birmingham, 2004.
7. Burda Z. *Fizyka i zarządzanie ryzykiem finansowym*. Postępy fizyki, tom 57, zes.3, 2006
8. Cai B.M., Cai C.X., Keasey K. *Market Efficiency and Returns to Simple Technical Trading Rules: Further Evidence form US, UK, Asian and Chinese Stock Markets*. Asia-Pacific Financial Markets, Springer 2005.
9. Chan Kim W., Mauborgne R. *Strategia błękitnego oceanu*. MT Biznes, Warszawa 2005
10. Chua S. *Sammy Chua's Day Trade*. Wiley Trading, 2007.
11. Cizek P., Härdle W, Weron R. *Statistical Tools for Finance and Insurance*, Springer-Verlag, Berlin 2005.
12. Czapiewski P., Samborska-Owczarek A. *Applicability Verification of Pattern – Based Decision Making Method in Real-life Conditions*. Polish Journal of Environmental Studies vol.16 No 4a, 2007, p. 57-62.
13. DeBontd W., Thaler R. *Does the stock market overreact?* Finance of Journal 40(3) 793-805
14. Dehuri S., Cho S.-B. *Multicriterion Pareto based particle swarm optimized polynomial neural network for classification: A review and state-of-the-art*. Computer Science Review 3(2009) 19-40.

15. Duch W. *Computational Intelligence and Where Is It Going*. In Challenges for Computational Intelligence (Ed. J. Kacprzyk). Springer. vol. 63/2007 s. 1-13.
16. Duda R.O., Hart P.E., Stork D.G., *Pattern Classification*, John Wiley, New York (2000).
17. Elder A. *Come Into My Trading Room*. Wiley Trading, 2002.
18. Elder J.F. *A review of Machine Learning, Neural and Statistical Classification* (eds. Michie, Spiegelhalter & Taylor; Ellis Horwood, 1994), *Journal of the American Statistical Association* 91, no. 433: 436-437.
19. Elton E.J, Gruber M.J. *Nowoczesna teoria portfelowa i analiza papierów wartościowych*. WIG-PRESS, Warszawa 1998.
20. Fama E. Market efficiency, long-term returns, and behavioral finance. *Journal of Financial Economics* 49 (1998) 283-306.
21. Fiszeder P. *Dynamiczna alokacja aktywów – model Markowitza, Rynki finansowe – prognozy a decyzje*, Acta Universitatis Lodziensis, Folia Oeconomica, 177, Wydawnictwo Uniwersytetu Łódzkiego, Łódź 2004.
22. Fiszeder P. *Forecasting the Volatility of Polish Stock Index – WIG20*, w: Forecasting Financial Markets. Theory and Applications, Wydawnictwo Uniwersytetu Łódzkiego, Łódź 2005.
23. Fiszeder P. *Modelling Financial Processes with Long Memory in Mean and Variance*, Dynamic Econometric Models 7, UMK, Toruń, 2006.
24. Friesen G.C.F, Weller P, Dunham L.M. *Price trends and patterns in technical analysis: A theoretical and empirical examination*. *Journal of Banking & Finance* 33(2009) 1089-1100.
25. Fujimoto K., Nakabayashi S. *Applying GMDH Algorithm to Extract Rules from Examples*. Systems Analysis Modelling Simulation, Volume 43, Number 10, 2003/10 pp. 1311-1319.
26. Gatnar E. *What is Data Mining*. *Statistics in Transition* vol.5 no. 5 p. 837-842
27. Gencar R. *Linear, non-linear and essential foreign exchange rate prediction with simple technical trading rules*. *Journal of International Economics* 47(1999) 91-107.
28. Grabczewski K. *Zastosowanie kryterium separowalności do generowania reguł klasyfikacji na podstawie baz danych*. pr. dokt. IBS, Warszawa, 2003.
29. Greshilov A.A., Stakun V.A., Stakun A.A. *Matematicheskiye metody postroenija prognozov*. Radio i sviaz. Moskva 1997.

## Bibliografia

30. Hashem S. , Schmeiser B. *Improving Model Accuracy using Optimal Linear Combinations of Trained Neural Networks*. IEEE Transactions on Neural Networks, 1995.
31. Hashem S. , Schmeiser B. *Improving Model Accuracy using Optimal Linear Combinations of Trained Neural Networks*. IEEE Transactions on Neural Networks, 1995.
32. Hia Jong Teoh, Tai Liang Chen, Ching Hsue Cheng, Hsing Hui Chu. *A hybrid multi-order fuzzy time series for forecasting stock markets*. Expert Systems with Applications 36 (2009) 7888-7897.
33. Hia Jong Teoh, Tai-Liang Chen, Ching-Hsue Cheng. *Frequency-Weighted Fuzzy Time-Series Based on Fibonacci Sequence for TAIEX Forecasting*. Emerging Technologies in Knowledge Discovery and Data Mining. Springer 4819/2007.
34. Ivakhnenko A., Ivakhnenko G. *Problems of Further Development of the Group Method of Data Handling Algorithms*. Part I. Pattern Recognition and Image Analysis vol.10 No.2, 2000, pp. 187-194.
35. Ivakhnenko A., Ivakhnenko G. *Problems of Further Development of the Group Method of Data Handling Algorithms, Part I*. Pattern Recognition and Image Analysis vol.10 No.2, 2000, pp. 187-194.
36. Ivakhnenko A., Ivakhnenko G., Mueller J. *Self- organization of Neural Network with Active Neurons*. Pattern Recognition and Image Analysis 1999 v.4 no. 2 pp. 185-196.
37. Ivakhnenko A., Stepashko V. *Pomekhoustoicivost modelirovanija*, Icjier 1986 , Naukovaja Dumka.
38. Ivakhnenko A.G., *An Inductive Sorting Method for the Forecasting of Multi-dimensional Random Processes and Events with the Help of Analogs Forecast Complexing*, Pattern Recognition and Image Analysis, 1991, vol. 1, no.1, pp.99-108.
39. Ivakhnenko A.G., *Dolgosrochnoe Prognozirovanie i Upravlenie Slozhnymi Sistemami* (Long-Term Forecasting and Control of Complex Systems), Kiev: Tekhnika, 1975.
40. Ivakhnenko A.G., Ivakhnenko G.A., Müller J.A., *Self-Organization of Neural Nets with Active Neurons*, Pattern Recognition and Image Analysis, 1994, vol. 4, no. 2. pp. 185-196.
41. Ivakhnenko A.G., Ivakhnenko G.A., Müller J.A., *Self-Organization of Optimum Physical Clustering of the Data Sample for Weakened Description and Forecasting of Fuzzy Objects*, Pattern Recognition and Image Analysis, 1993, vol. 3, no. 4, pp. 415-422.



42. Ivakhnenko A.G., Ivakhnenko G.A., *Obzor zadac po algoritmam Metoda Gruppovogo Uceta Argumentov* . Pattern Recognition and Image Analysis, 1995 vol. 5 no. 4 pp. 527-535.
43. Ivakhnenko A.G., Ivakhnenko G.A., *Perceptron Synthesis according to Clustering-Balance Criterion*. Pattern Recognition and Image Analysis, 1995, vol. 5, no. 3, pp. 337-341.
44. Ivakhnenko A.G., Ivakhnenko G.A., *Simplified Linear Programming Algorithm as Basic Tool for Open-Loop Control*, Systems Analysis, Modelling, and Simulation, (SAMS), 1995, vol.18-19, pp.315-319.
45. Ivakhnenko A.G., Müller J.A., *Parametric and Non-parametric Selection Procedures in Experimental Systems Analysis*, Systems Analysis, Modeling, and Simulation, 1992, vol. 9, pp. 157-175.
46. Ivakhnenko A.G., Müller J.A., *Problems of Computer Clustering of the Data Sampling of Objects under Study*, Sov. J. Automation Inf. Sci., 1991, vol. 24, no. 1, pp. 58-67.
47. Ivakhnenko A.G., *Samoobujascijesja systemy raspoznavanija i avtomatycznego upravlenija*. Tekhnika, Kijev 1969.
48. Ivakhnenko A.G., Stepashko V.S., *Pomekhoustoichivost' Modelirovaniya* (Noise Immunity of Modeling), Kiev: Naukova Dumka, 1985.
49. Ivakhnenko A.G., Yurachkovskij Yu.P., *Modelirovanie Slozhnykh Sistem po Eksperimental'nym Dannym* (Modeling of Complex Systems from Experimental Data), Moscow: Radio i Svyaz', 1987.
50. Ivakhnenko A.G., Zaichenko J.P., Dimitrov V.D. *Priniatije resenij na osnove samoorganizacii*. Sovetskoe Radio, Moskwa 1976.
51. Ivakhnenko G.A., *Self-Organization of Neuronet with Active Neurons for Effects of Nuclear Test Explosions Forecastings*. System Analysis Modeling Simulation (SAMS), 1995, vol.20, pp.107-116.
52. Jajuga K. *Metody statystyczne w finansach*. Statsoft Polska, 2003.
53. Jajuga K., Jajuga T. *Jak inwestować w papiery wartościowe*. PWN, Warszawa 1994.
54. Jajuga K., Kuziak K., Markowski P. *Inwestycje finansowe*, Wydawnictwo Akademii Ekonomicznej im. Oskara Langego we Wrocławiu, Wrocław 1998.
55. Jing Dang, Brabazon A., Edelman D., O'Neill M. *An Introduction to Natural Computing in Finance*. LNCS, PublisherSpringer Berlin / Heidelberg Volume 5484/2009.

56. Jinhong You, Xian Zhou, Li-Xing Zhu. *Inference on a regression model with noised variables and serially correlated errors*. Journal of Multivariate Analysis 100(2009) 1182-1197.
57. Kacprzyk J., Wilbik A., Zadrozny S. *A linguistic quantifier based aggregation for a human consistent summarization of time series*. In: J. Lawry, E. Miranda, A. Bugarin, S. Li, M. A. Gil, P. Grzegorzewski, O. Hryniewicz (Eds.): *Soft Methods for Integrated Uncertainty Modeling*. Springer-Verlag, Berlin Heidelberg 2006, ss. 183-190, Seria: *Advances in Soft Computing*.
58. Kacprzyk J., Wilbik A., Zadrozny S. *Using a genetic algorithm to derive a linguistic summary of trends in numerical time series*. In: *Proceedings of 2006 International Symposium on Evolving Fuzzy Systems*, Ambleside, UK, September 7-9, 2006, IEEE Press, ss. 137-142.
59. Kacprzyk J., Zadrozny S. *Fuzzy linguistic data summaries as a human consistent, user adaptable solution to data mining*. In: Gabrys B., Leiviskä K., Strackeljhan J. (Eds.): *Do smart adaptive systems exist?* Springer Verlag, Heidelberg New York 2005, ss. 321-339, Seria: *Studies in fuzziness and soft computing*.
60. Kohavi R. *A Study of Cross-Validation and Bootstrap for Accuracy Estimation and Model Selection*. Computer Science Department, Stanford University, 1995.
61. Kompa K., Matuszewska Janica A. *Examination of Warsaw Stock Exchange Indexes Behaviour: Applications of Rolling Windows Variance Ratio Test*. Polish Journal of Environmental Studies vol.17 No 3b, 2008, p. 150-154.
62. Korbicz J., Marcin Mrugalski M. *Confidence estimation of GMDH neural networks and its application in fault detection systems*. International Journal of Systems Science, 2008, Vol. 39, no 8, pp. 783-800.
63. Krawczak M. *Theoretical foundations of neural networks prediction*. W: Studziński J., Drelichowski L., Hryniewicz O., Kacprzyk J. (red.) *Technologie informatyczne w zarządzaniu, systemy wspomaganie decyzji*. IBS PAN, Warszawa 2000, pp. 245-254.
64. Krawczak M., Jakubowski A., Konieczny P., Kulikowski R., Miklewski A., Szkatuła G. *Aktywne zarządzanie inwestycjami finansowymi – wybrane zagadnienia*, Akademicka Oficyna Wydawnicza EXIT, Warszawa 2003.
65. Krogh A., Vedelsby J., *Neural Network Ensembles, Cross Validation and Active Learning*. NIPS, 7 of 1995.
66. Kruttsinger J. *Systemy transakcyjne. Sekrety mistrzów*. Wig-Press, Warszawa 1999.
67. Kulikowski R. *Wspomaganie inwestycji obarczonych ryzykiem*. W: Krawczak M., Jakubowski A., Konieczny P., Kulikowski R., Miklewski A., Szkatuła

- G.: *Aktywne zarządzanie inwestycjami finansowymi – wybrane zagadnienia*, Akademicka Oficyna Wydawnicza EXIT, Warszawa 2003, ss. 253-268.
68. Kuś J., Korbicz J. *Statyczne i dynamiczne sieci GMDH w Sieci neuronowe* (eds. Duch, Korbicz, Rutkowski, Tadeusiewicz) w serii pod red. M. Nałęcza *Biocybernetyka i inżynieria medyczna*, PAN, Warszawa 2000.
  69. Kutner R. *Niegaussowskie procesy stochastyczne i niedebye'owska relaksacja oraz zdarzenia rzadkie w świecie realnym*. Uniwersytet Warszawski, 2006.
  70. Lemke F., Mueller J.A. *Self-Organizing Data Mining A Portfolio Trading System*. *Journal of Computational Intelligence in Finance*, 1997/05 pp 12-26.
  71. LeSage J., Sudipto Banerjee, Fischer M.M., Congdon P. *Spatial Statistics: Methods, models & computation*. Computational Statistics and Data Analysis 53(2009) 2781-2785.
  72. Łuniewska M. *The Construction and Evaluation of Sectorial Portfolios*. Polish Journal of Environmental Studies vol.16 No 4a, 2007, p. 161-165.
  73. Madala H.R., Ivakhnenko A. G. *Inductive Learning Algorithms for Complex Systems Modelling*. CRC Press 1994.
  74. Makridakis S., Wheelwright S., Hyndman R. *Forecasting: Methods and Applications*. John Wiley and Sons Inc. 1998.
  75. Marjak H. *The Architecture Selection for Neural Network – a Financial Criterion or a Prediction Criterion*. Polish Journal of Environmental Studies vol.17 No 3b, 2008, p. 255-261.
  76. McLachlan G.J. *Discriminant Analysis and Statistical Pattern Recognition*, Wiley Interscience, New York, 1992.
  77. Merz Ch. J. *Classification and Regression by Combining Models*. Ph.D. thesis, University of California, Irvine, 1998.
  78. Mohsen Mehrara, Ali Moeini, Mehdi Ahrari, Ali Erfanifard. *Investing the efficiency in oil futures market based on GMDH approach*. *Expert Systems with Applications* 36(2009) 7479-7483.
  79. Mrugalski, M., Arinton, E., Korbicz, J. *Sieci neuronowe typu GMDH w diagnostyce systemów przemysłowych*. *Pomiary, Automatyka, Kontrola*. 2002, R. 48, nr 2, s. 10-14.
  80. Mueller J., Lemke F. *Self-organizing Data Mining. An Intelligent Approach to Extract Knowledge from Data*. Berlin, 1999.
  81. Muriel A., *Short-term predictions in forex trading*, *Physica A* 344 (2004) 190-193

## Bibliografia

82. Niederliński A. *Systemy komputerowe automatyki przemysłowej – zastosowania*. WNT, Warszawa 1985.
83. Nowak R. *Statystyka dla fizyków*. WN PWN, Warszawa 2002.
84. Onwubolu G. *Hybrid Self-Organizing Modeling Systems*. Springer 2009.
85. Owens S., Lizotte O. *When to trade*. FX Engines, 2004.
86. Palit A., Popovic D. *Computational Intelligence in Time Series Forecasting*. Springer 2005.
87. Park B.J., Oh S.K., Ahn T.C., Pedrycz W. *A study on multilayer fuzzy polynomial interference system based on an extended GMDH algorithm*. Proc. 8<sup>th</sup> IEEE Inf. Conference on Fuzzy Systems, Seoul 1999.
88. Pawlak Z. *Rough sets and intelligent data analysis*. Information Sciences, 147, 2002, 1-12.
89. Pawlak Z., Polkowski L., Skowron A. *Rough Set Theory*. In: B. Wan, (ed.), *Encyclopedia of Computer Science and Engineering*, Wiley, New York 2004.
90. Pedrycz W. *Computational Intelligence: An Introduction*. CRC Press 1997.
91. Penrose R. *Droga do rzeczywistości*. Prószyński i Ska. Warszawa 2006.
92. Piotrowski E. *Dwoistości wartości kapitału*. Politechnika Białostocka 2002.
93. Piotrowski E., Śładkowski J. *Quantum Diffusion of Prices and Profits*. Physica A, 345(2005) s. 185-195.
94. Piper J. *The Way to Trade*. Harriman Modern Classic, 2006.
95. Raghuraj R.K., Lakshminarayanan S. *Variable Predictive Models – A new multivariate classification approach for pattern recognition application*. Elsevier, Pattern Recognition vol. 42/1, 2009 p.7-17 [sciencedirect.com].
96. Raghuraj R.K., Lakshminarayanan S. *VPMCD: Variable Interaction Modeling Approach for Class Discrimination in Biological Systems*. FEBS Letters, 581, 826-830, 2007.
97. Rejer I. *Fuzzy Expert Model of Index WIG20*. Polish Journal of Environmental Studies vol.17 No 3b, 2008, p. 365-370.
98. Rua A., Nunes C. *International comovement of stock market: A wavelet analysis*. Journal of Empirical Finance, 2009.
99. Rutkowski L. *Metody i techniki sztucznej inteligencji*, WNT, Warszawa 2006.
100. Satchwell Ch. *Pattern Recognition and Trading Decisions*. McGraw Hill, New York 2005.

101. Sawaragi, Y., Soeda, T., Tamura, H. et al., *Statistical Prediction of Air Pollution Levels Using Non-Physical Models*, Automatica (IFAC), 1979, vol. 15, no. 4.
102. Sawaragi, Y., Soeda, T., Tamura, H. et al., *Statistical Prediction of Air Pollution Levels Using Non-Physical Models*, Automatica (IFAC), 1979, vol. 15, no. 4, pp. 441-452.
103. Schwager J.D. *Analiza techniczna rynków terminowych*. WIG-Press, Warszawa 2002.
104. *Self-Organizing Methods in Modeling*, Statistics: Textbooks and Monographs, Farlow, S.J., Ed., New York: Marcel Dekker Inc., 1984, vol. 54.
105. Shahwan T. , Lemke F. *Forecasting Commodity Prices for Predictive Decision Support Systems*. EFITA/WCCA 2005, Via Real, Portugal.
106. Sharpe W.F. *The Sharpe Ratio*. The Journal of Portfolio Management Fall, 1994 s.49-58
107. Sheng-Hsun Hsu, Po-An Hsieh, Ting-Chih Chih, Kuei-Chu Hsu. *A two-stage architecture for stock price forecasting by integrating self-organizing map and support vector regression*. Expert Systems with Applications 36(2009) 7947-7951.
108. Shleifer A. *An inefficient market: An introduction to behavioral finance*. Clarendon Lectures. Oxford University Press 2000.
109. Sortino F., Van der Meer R. *Downside Risk*. The Journal of Portfolio Management Fall., 1991 s.27-31.
110. Sung-Kwun Oh, Pedrycz W. *The design of self-organizing Polynomial Neural Networks*. Information Sciences, 141 (2002) pp.237-258.
111. Szymański S. *Integracja metod klasyfikacji odkrywania wiedzy w systemie ekspertowym*. pr. dokt. Wydział Informatyki Politechniki Szczecińskiej, 2006.
112. Tadeusiewicz R. *Sieci neuronowe*. Akademicka Oficyna Wydawnicza RM, Warszawa 1993.
113. Tak-chung Fu, Fu-lai Chung, Luk R., Chak-man Ng. *Representing Financial Time Series based on data point importance*. In: Engineering Applications of Artificial Inteligence 21(2008) p.277-300.
114. Tarczyński W, Łuniewska M. *Dywersyfikacja ryzyka na polskim rynku kapitałowym*. Placet, Warszawa 2004.
115. Tarczyński W, Łuniewska M. *Metody wielowymiarowej analizy porównawczej na rynku kapitałowym*. PWN, Warszawa 2006.

## Bibliografia

116. Triseev, Y. P., *Approaches to the Solution of Mathematical Programming Problems on the Basis of Heuristic Self-Organization*, Sov. J. Automation Inf. Sci., 1987, vol. 20, no. 2, pp. 30-37.
117. Tsoumakas G., Katakis J., Vlahavas I. *Effective Voting of Heterogeneous Classifiers*. Department of Informatics, Aristotele University of Thessaloniki, 2004.
118. Tsymbal A., Pechenizkiy M., Cunningham P. *Diversity in search strategies for Ensemble Feature Selection*. Department of Computer Science, Trinity College Dublin, 2003.
119. Vapnik V. *Statistical Learning Theory*, Wiley-Interscience, New York, 1998.
120. Wagner N., Michalewicz Z., Khouja M., Mc Gregor R.R. *Forecasting with a Dynamic Window of Time: The DyFor Genetic Program Model*. Springer Berlin/Heidelberg vol.3490/2005, 205-215.
121. Wagner N., Michalewicz Z., Khouja M., Mc Gregor R.R. *Time Series Forecasting for Dynamic Environment: The DyFor Genetic Program Model*. IEEE Transaction on Evolutionary Computation, vol. 11 n. 4 /2007, s.433-453.
122. Weiss S., Kulikowski C., *Computer Systems that Learn: Classification and Prediction Methods from Statistics, Neural Networks, Machine Learning and Expert Systems*, San Francisco, 1991.
123. Weron A., Weron R. *Inżynieria finansowa: Wycena instrumentów pochodnych, Symulacje komputerowe, Statystyka rynku*. WNT, Warszawa 2005.
124. Wierzbicki A. *Badania operacyjne i systemowe w epoce cywilizacji wiedzy*. Instytut Badań Systemowych PAN seria Badania systemowe tom 64, Warszawa 2008, s. 13-35.
125. Wiliński A. *Analiza szeregów czasowych w ujęciu bayesowskim dla funkcji dyskryminacyjnej opartej na rozkładzie gaussowskim*. XI Sesja Informatyki Wydział Informatyki Politechniki Szczecińskiej, Wydawnictwo Informa, Szczecin 2006.
126. Wilinski A. *GMDH Based Computational Intelligence Method for Prediction of Financial Markets*. "Intellectual Systems of Decision Making and The Problems of Computational Intelligence (ISDMCI'2008)", Eupatoria, Tom 1 cz.1, 2008.
127. Wiliński A. *Risk Estimation of Consecutive Losses in Automated Currency Trading*. Polish Journal of Environmental Studies vol.16 No 4a, 2007, p. 380-385.
128. Wiliński A., Czapiewski P. *GMDH-based Multiple Prediction of Financial Markets*. Polish Journal of Environmental Studies vol.17 No 3b, 2008 p.73-78.

129. Wilinski A., Czapiewski P., Sienkiewicz S. *Model interfejsu graficznego w systemie wspomagania decyzji w obrocie instrumentami finansowymi*. X Sesja Informatyki Wydział Informatyki Politechniki Szczecińskiej, Wydawnictwo Informa, Szczecin 2005.
130. Witkowska D. *Artificial Neural Networks Application to Forecast the Participation Units of the Open-End Investment Fund*. Polish Journal of Environmental Studies vol.17 No 3b, 2008, p. 523-529.
131. Witkowska D., Matuszewska A., Kompa K. *Wprowadzenie do ekonometrii dynamicznej i finansowej*. Wydawnictwo SGGW, Warszawa 2008.
132. Yungho Leu, Chien-Pang Lee, Yie-Zu Jou. *A distance-based fuzzy time series for exchange rates forecasting*. Expert Systems with Applications 36(2009) 8107-8114.
133. Zhang Yudong, Wu Lenan. *Stock market prediction of S&P 500 via combination of improved BCO approach and BP neural network*. Expert Systems with Applications 36(2009) 8849-8854.
134. Zhohnarskij, A. A., *Agglomerative Cluster Analysis Procedures for Multidimensional Objects: A Test for Convergence*. Pattern Recognition and Image Analysis, 1992, vol. 2, no. 4, pp. 388-390.
135. Weisstein, Eric W. *Moore-Penrose Matrix Inverse*. Wolfram MathWorld [online] <http://mathworld.wolfram.com/Moore-PenroseMatrixInverse.html> [dostęp: 02/2008]
136. Stokalski B. Cloud computing – informatyka bez kłopotów [dostęp 05/2008] <http://www.infovidematrix.pl/inspiracje/?p=61#more-61>
137. [www.e-forex.ro](http://www.e-forex.ro) [dostęp 08/2008]
138. [www.oanda.com](http://www.oanda.com) [dostęp 05/2007 – 09/2008]
139. [www.investopedia.com](http://www.investopedia.com) [dostęp 02/2005-10/2008]
140. [www.wanadoo.com](http://www.wanadoo.com) [dostęp 02/2006-11/2008]
141. [www.gmdh.com](http://www.gmdh.com) [dostęp 06/2007-11/2008]
142. [www.fxengines.com](http://www.fxengines.com) [dostęp 12/2007 – 02/2009]







Zasadniczym celem niniejszej monografii jest potwierdzenie tezy autora, że fuzja matematyki i technologii egzemplifikowana powiązaniem metod predykcji ze strategią inwestycyjną opartą na mechanizmach technologicznych platform brokerskich daje nową synergetyczną wartość.

Praca koncentruje się na pragmatycznym powiązaniu predykcji z możliwościami jej internetowej (automatycznej) realizacji. Wymagało to uwzględnienia szeregu niespodziewanych ograniczeń natury technologicznej. W pracy przedstawiono wyniki wielu eksperymentów w przestrzeniach historycznych różnych instrumentów finansowych.

**ISSN 0208-8029**

**ISBN 9788389475237**

---

---

**Instytut Badań Systemowych PAN**

**Tel. Centrala 022-38 10 100 / fax 022-38 10 105 e-mail: [ibs@ibspan.waw.pl](mailto:ibs@ibspan.waw.pl)**