



IBS PAN, Warszawa 9 kwietnia 2008

Obrona rozprawy doktorskiej

Neuro-genetyczny system komputerowy
do prognozowania zmiany indeksu
giełdowego

mgr inż. Marcin Jaruszewicz

promotor: dr hab. inż. Jacek Mańdziuk, prof. PW

Plan prezentacji

- ▶ Cel badań
- ▶ Hipotezy badawcze
- ▶ Sposób realizacji
- ▶ Opis przekształceń danych
- ▶ Opis systemu
- ▶ Gra giełdowa
- ▶ Wyniki
- ▶ Podsumowanie



Cel badań

- ▶ Wykazanie możliwości skutecznej krótkoterminowej prognozy zmiany wartości indeksu giełdowego za pomocą metod inteligencji obliczeniowej
- ▶ Zastosowanie algorytmu genetycznego jako wsparcie procesu doboru architektury sieci neuronowej
- ▶ Zastosowanie przekształceń Analizy Technicznej
- ▶ Zastosowanie informacji z innych giełd

Hipotezy badawcze

- ▶ Hipoteza główna:

Istnieje możliwość względnie skutecznej prognozy wartości indeksu dla horyzontu 1 dnia przy zastosowaniu systemów hybrydowych inteligencji obliczeniowej

- ▶ Hipoteza 1:

Zależności między parametrami opisującymi giełdy są zmienne w czasie

- ▶ Hipoteza 2:

Zależności pomiędzy różnymi giełdami mogą być wykorzystane w prognozie

Założenia sposobu realizacji

- ▶ Mając dany ciąg kolejnych wartości indeksu giełdowego,
- ▶ wykorzystując algorytm genetyczny do doboru zestawu zmiennych,
- ▶ za pomocą sieci neuronowej
- ▶ wykonać prognozę procentowej zmiany indeksu giełdowego od znanej wartości zamknięcia do wartości zamknięcia kolejnego dnia

Plan prezentacji

- ▶ Cel badań
- ▶ Hipotezy badawcze
- ▶ Sposób realizacji
- ▶ Opis przekształceń danych
- ▶ Opis systemu
- ▶ Gra giełdowa
- ▶ Wyniki
- ▶ Podsumowanie



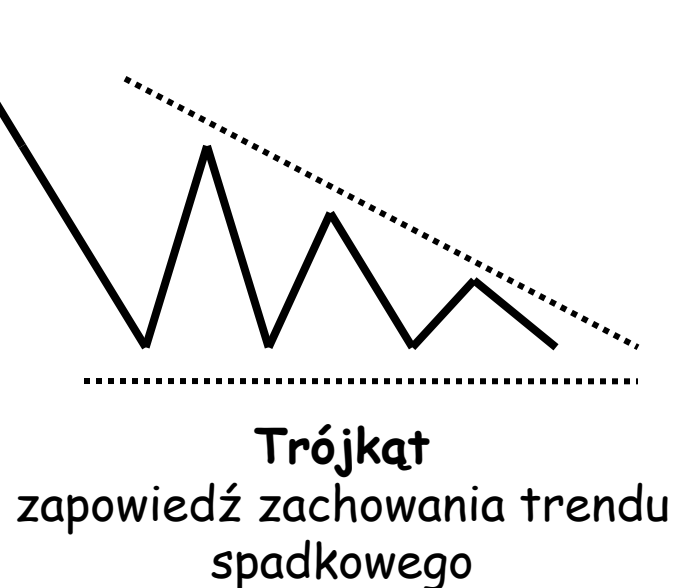
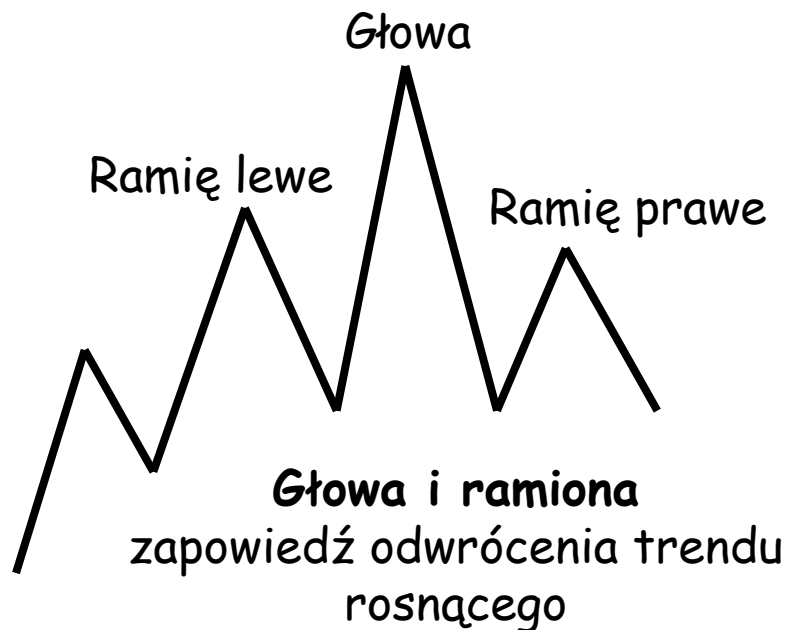
Przekształcenia danych

- ▶ Dane źródłowe dla giełd GSE, NYSE, TSE: wartości indeksu DAX: otwarcia, zamknięcia, minimalna, maksymalna
- ▶ Średnie i procentowe zmiany wartości (5-,10-,20-dniowe)
- ▶ Oscylatory (8 podstawowych)
 - ▶ Przekształcenia podstawowych parametrów giełdy
 - ▶ Generowane sygnały

Przekształcenia danych

▶ Formacje

- ▶ Charakterystyczne kształty wykresu wartości ceny
- ▶ Prognoza zmiany lub zachowania trendu



Dane wejściowe – podsumowanie

- ▶ Zmienne opisujące giełdę GSE
 - ▶ 30 zmiennych dla bieżącego dnia (wszystkie zmienne)
 - ▶ 40 zmiennych dla wcześniejszych 5 dni (tylko zmienne: wartość zamknięcia, procentowe zmiany, średnie kroczące)
- ▶ Zmienne spoza giełdy GSE
 - ▶ wartości zamknięcia indeksu giełd: NYSE (DJIA), TSE (NIKKEI)
 - ▶ wartości notowań walut: EUR do USD, USD do JPY

Plan prezentacji

- ▶ Cel badań
- ▶ Hipotezy badawcze
- ▶ Sposób realizacji
- ▶ Opis przekształceń danych
- ▶ Opis systemu
- ▶ Gra giełdowa
- ▶ Wyniki
- ▶ Podsumowanie



Elementy systemu

1. Moduł przetwarzania danych
 - ▶ Wyliczenie zmiennych
 - ▶ Wyznaczenie formacji
 - ▶ Skalowanie wartości
2. Algorytm genetyczny jako narzędzie wyboru zmiennych wejściowych dla sieci neuronowej
3. Sieć neuronowa jako narzędzie do prognozy

Sieć Neuronowa

- ▶ Wielowarstwowy perceptron
- ▶ Jednokierunkowa sieć
- ▶ Pojedyncza warstwa ukryta
- ▶ Nauka metodą propagacji wstecznej z momentem
- ▶ Wykorzystanie w algorytmie genetycznym
- ▶ Wykorzystanie do prognozy

Algorytm Genetyczny

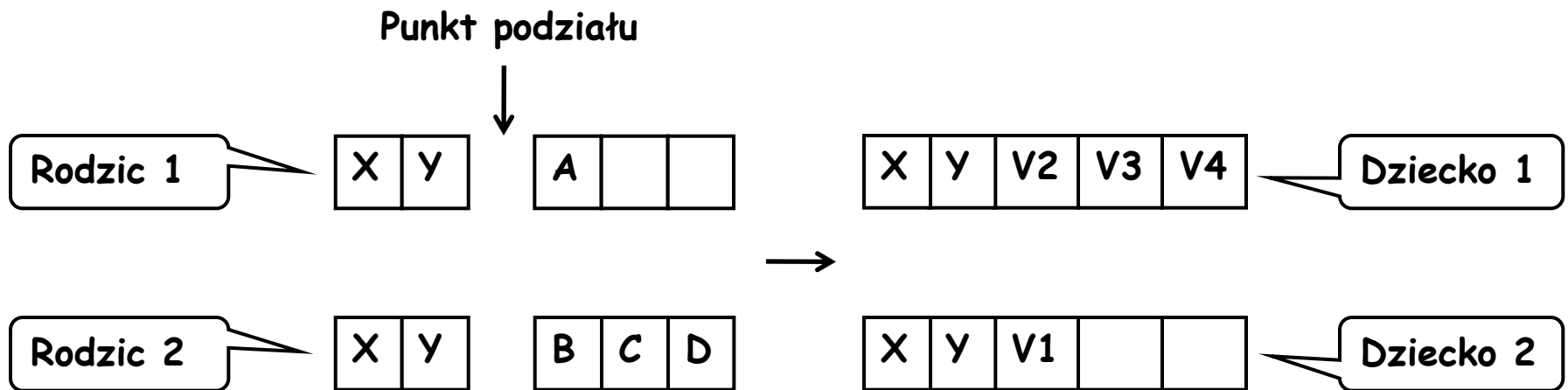
- ▶ Chromosom
 - zbiór zmiennych wejściowych dla sieci neuronowej
- ▶ Przystosowanie chromosomu
 - w oparciu o błąd prognozy sieci neuronowej
- ▶ Martwe chromosomy
 - brak możliwości nauki dla zakodowanych zmiennych

Algorytm Genetyczny – mutacje

- ▶ Mutacja kodowanych zmiennych
 - ▶ Prawdopodobieństwo zależne od przekroczenia progu liczby iteracji
 - ▶ Prawdopodobieństwo zależne od zróżnicowania przystosowania osobników w populacji
 - ▶ Prawdopodobieństwo zależne od przystosowania chromosomu (*żywy, martwy*)
- ▶ Mutacja liczby kodowanych zmiennych
 - ▶ Występuje w przypadku osiągnięcia progu wartości średniego rozmiaru chromosomów w populacji
 - ▶ Umożliwienie tworzenia osobników o rozmiarze różnym od rodziców

Algorytm Genetyczny – krzyżowanie

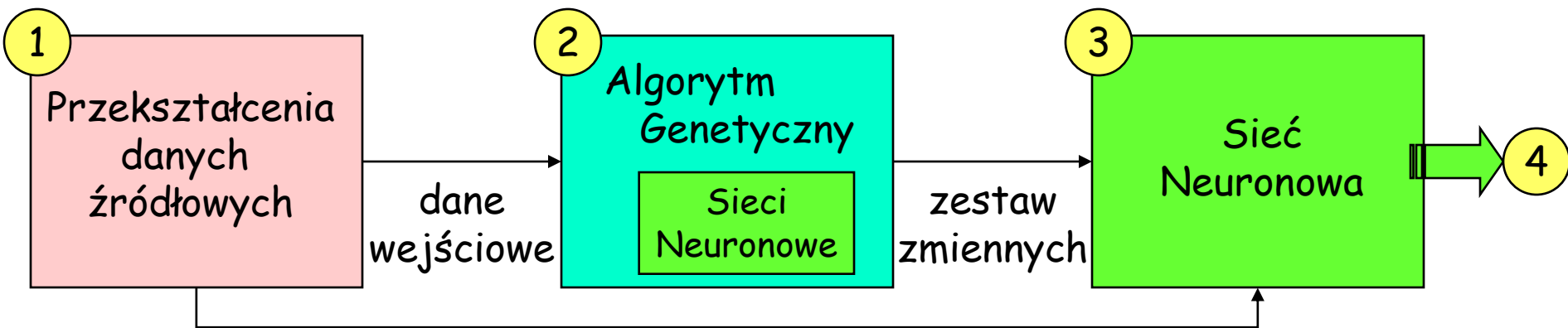
- ▶ Selekcja turniejowa
- ▶ Krzyżowanie „iloczynowe”
- ▶ Zastępowanie rodziców na podstawie przystosowania



X,Y - zmienne wspólne, A-D - pozostałe zmienne V - zmienne wylosowane

Działanie systemu – podsumowanie

1. Wygenerowanie danych wejściowych
2. Znalezienie przez **algorytm genetyczny** zestawu zmiennych dla prognozy
3. Nauka **sieci neuronowej**
4. Prognoza **sieci neuronowej**



Plan prezentacji

- ▶ Cel badań
- ▶ Hipotezy badawcze
- ▶ Sposób realizacji
- ▶ Opis przekształceń danych
- ▶ Opis systemu
- ▶ Gra giełdowa
- ▶ Wyniki
- ▶ Podsumowanie



Hipotetyczna gra giełdowa

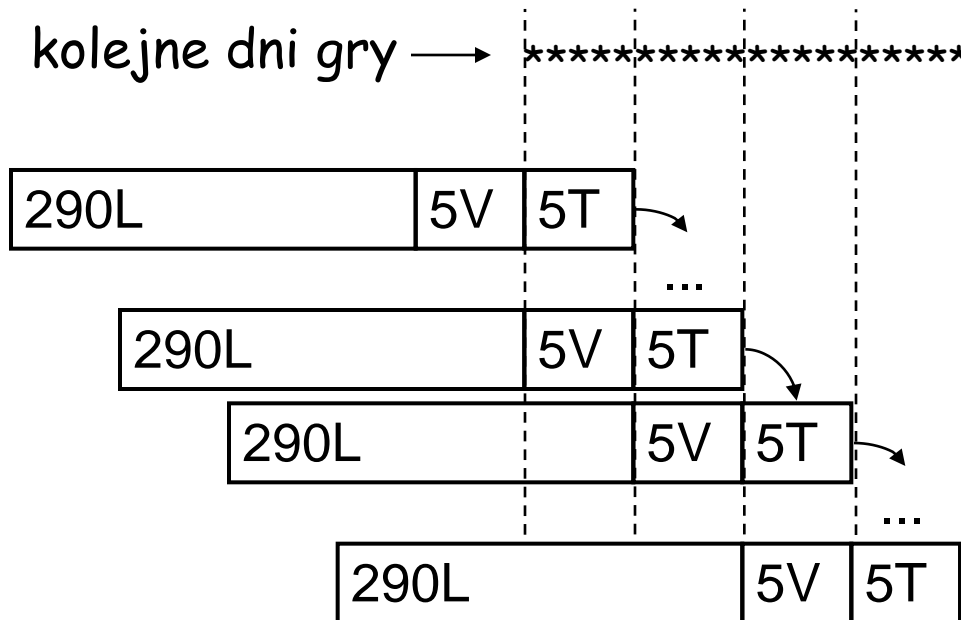
- ▶ Gra rozpoczyna się z budżetem 100 000 jednostek
- ▶ Wszystkie wartości z danego dnia i historii są znane
- ▶ Na podstawie prognozy wartości zamknięcia w następnym dniu podejmowana jest decyzja inwestycyjna: czekaj, kupuj, sprzedaj
- ▶ Decyzja jest realizowana po cenie otwarcia następnego dnia, brak kosztu transakcji
- ▶ Zawsze sprzedawana lub kupowana jest maksymalna liczba indeksów zaokrąglona do liczby całkowitej
- ▶ W ostatnim dniu gry sprzedawane są wszystkie posiadane indeksy

Hipotetyczna gra giełdowa

- ▶ Wynik gry jest zdefiniowany jako kwota pieniędzy pozostała na jej zakończenie
- ▶ Sukces zależy od skuteczności prognozy kierunku zmiany wartości indeksu
- ▶ Możliwy jest zysk w trendzie malejącym, jeśli obecne są chwilowe wzrosty wartości indeksu

Hipotetyczna gra giełdowa – algorytm SNG

- ▶ Sposób działania systemu neuro-genetycznego w czasie gry (100 dni)



wyznaczenie zestawu zmiennych dla sieci neuronowej (algorytm genetyczny), nauka sieci dla dostępnych nowych danych

L – rekordy uczące, V – walidacyjne, T – testowe

Hipotetyczna gra giełdowa – algorytmy

- ▶ Algorytmy testowe
 - ▶ **kupuj i trzymaj (KT)** > w pierwszym dniu zakup maksymalnej ilości indeksów, ostatniego dnia sprzedaż
 - ▶ **utrzymanie ostatniej zmiany (OZ)** > założenie, że kierunek ostatniej zmiany będzie zachowany
 - ▶ **prognoza MACD (MACD)** > decyzja na podstawie sygnałów generowanych przez oscylator MACD
 - ▶ **gra losowa (LS)** > losowa decyzja z równym prawdopodobieństwem dla kupna i sprzedaży
 - ▶ **profetyczny (PRF)** > decyzja oparta o wiedzę najbliższej, przyszłej wartości indeksu

Plan prezentacji

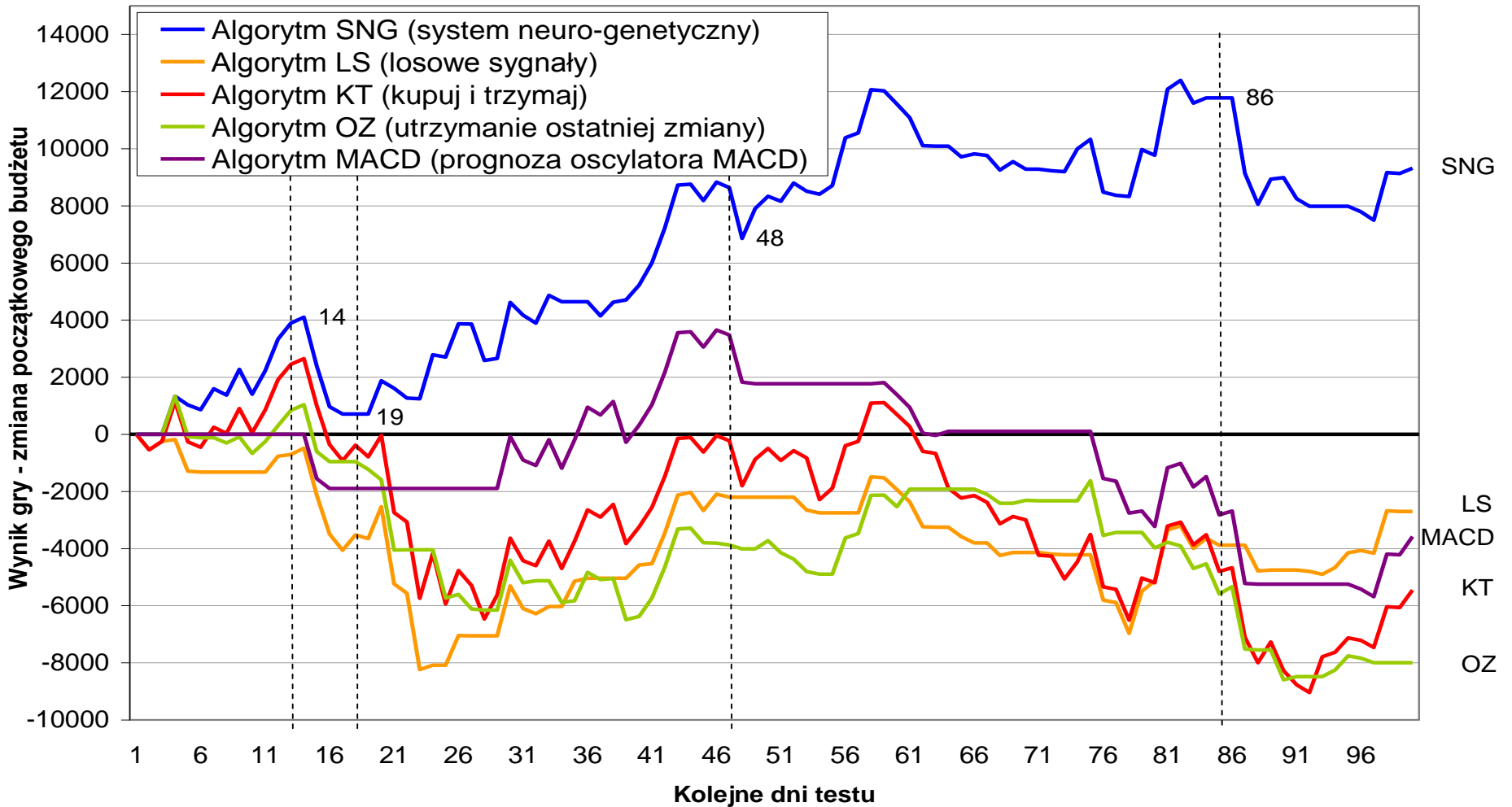
- ▶ Cel badań
- ▶ Hipotezy badawcze
- ▶ Sposób realizacji
- ▶ Opis przekształceń danych
- ▶ Opis systemu
- ▶ Gra giełdowa
- ▶ Wyniki
- ▶ Podsumowanie



Wyniki gry - zestawienie

Algorytm	Wynik [%]
Algorytm SNG (system neuro-genetyczny)	9.32
Algorytm LS (losowe sygnały)	-2.73
Algorytm KT (kupuj i trzymaj)	-5.45
Algorytm OZ (utrzymanie ostatniej zmiany)	-8.00
Algorytm MACD (prognoza oscylatora MACD)	-3.58
Algorytm PRF (profetyczny)	39.56

Wyniki gry – zmiana w czasie



Eksperymenty weryfikujące

- ▶ Powtórzenie eksperymentu przy identycznych ustawieniach systemu

Wynik [%] w poszczególnych przebiegach algorytmu SNG						
1	2	3	4	5	6	wartość średnia
9.32	4.68	3.51	3.41	5.96	7.67	5.44

- ▶ Rozpiętość wyników wynosi 6 p.p.
- ▶ Porównanie wyborów zmiennych przez algorytm genetyczny (dla trzech pierwszych przebiegów)

43% identyczne zmienne we wszystkich przebiegach

26% identyczne zmienne w dwóch przebiegach

30% zmienne nie występujące w innych przebiegach

Eksperymenty weryfikujące

- ▶ Przeprowadzenie eksperymentu dla innego roku
 - ▶ Różnica SNG a KT równa 15 p.p.
 - ▶ Różnica SNG a MACD i LS równa 7 p.p.

Algorytm	Wynik [%]
Algorytm SNG (system neuro-genetyczny)	0.19
Algorytm LS (losowe sygnały)	-7.14
Algorytm KT (kupuj i trzymaj)	-15.21
Algorytm OZ (utrzymanie ostatniej zmiany)	-26.45
Algorytm MACD (prognoza oscylatora MACD)	-7.39
Algorytm PRF (profetyczny)	222.03

Eksperymenty weryfikujące

- ▶ Stabilność wyników – kolejne 5-dniowe przesunięcia okresu 100 dni testów

Przesunięcie okresu testów algorytmu SNG	Wynik	Różnica w stosunku do poprzedniego wyniku
Brak przesunięcia – pierwotny eksperyment	a=109 315.36	n.d.
Przesunięcie o 5 dni	b=110 076.75	0.70%
Przesunięcie o 10 dni	c=108 214.21	-1.69%
Przesunięcie o 15 dni	d=110 467.53	2.08%
Przesunięcie o 20 dni	e=109 297.39	-1.06%

Uwzględnienie kosztów transakcji

- ▶ Koszt [%] każdej transakcji od 0% do 0.4%

algorytm/koszt	Wynik [%]				
	0%	0.1%	0.2%	0.3%	0.4%
Algorytm SNG	9.32	3.96	-0.83	-5.63	-10.00
Algorytm LS	-2.73	-7.36	-11.73	-15.89	-19.88
Algorytm KT	-5.45	-5.64	-5.83	-6.01	-6.20
Algorytm OZ	-8.00	-12.48	-16.98	-21.25	-25.17
Algorytm MACD	-3.58	-4.67	-5.87	-6.89	-8.03
Algorytm PRF	39.56	32.12	24.98	18.21	12.03

Plan prezentacji

- ▶ Cel badań
- ▶ Hipotezy badawcze
- ▶ Sposób realizacji
- ▶ Opis przekształceń danych
- ▶ Opis systemu
- ▶ Gra giełdowa
- ▶ Wyniki
- ▶ Podsumowanie



Wyróżniające cechy rozwiązania

- ▶ Zastosowanie algorytmu genetycznego do adaptacyjnego doboru zmiennych wejściowych dla sieci neuronowej
- ▶ Obecność martwych chromosomów
- ▶ Krzyżowanie iloczynowe
- ▶ Mutacja rozmiaru chromosomu
- ▶ Kontekstowe prawdopodobieństwo mutacji
- ▶ Szerokie zastosowanie Analizy Technicznej (średnie, oscylatory, formacje).

Podsumowanie

- ▶ Potwierdzone zostały wszystkie postawione hipotezy
- ▶ System autonomicznie przystosowuje się do zmian warunków na giełdzie
- ▶ Sensownie wykorzystane są popularne zmienne Analizy Technicznej
- ▶ Uzyskane zostały wyniki lepsze o 15 p.p. od algorytmu „kupuj i trzymaj”
- ▶ Osiągnięto zadowalające wyniki przy uwzględnieniu kosztu transakcji
- ▶ Wykazana została skuteczność gry zarówno dla trendu wzrostowego jak i spadkowego.

Dziękuję za uwagę

Materiały są dostępne są na stronie: jaruszewicz.com