

Marcin Pietrzykowski*

ZASTOSOWANIE METODY MINI-MODELI OPARTEJ NA HIPERSZEŚCIANIE W PROCESIE MODELOWANIA DANYCH WIELOWYMIAROWYCH

Streszczenie

W artykule zaprezentowano metodę samo-uczenia mini-modeli (metodę MM) opartą na hiperbryłach w przestrzeni wielowymiarowej. Jest to metoda nowa i rozwojowa, będąca w trakcie intensywnych badań. Bazuje ona na próbkach pobieranych jedynie z lokalnego otoczenia punktu zapytania, a nie z obszarów odległych od tego punktu. Grupa punktów, używana w procesie uczenia mini-modelu jest ograniczona obszarem hiperbryły. Na tak zdefiniowanym lokalnym otoczeniu punktu zapytania metoda MM w procesie uczenia oraz obliczania odpowiedzi można użyć dowolnej metody aproksymacji. W artykule przedstawiono algorytm uczenia i działania metody w przestrzeni wielowymiarowej bazujący na hipersferycznym układzie współrzędnych. Metodę przebadano na zbiorach danych wielowymiarowych, a wyniki porównano z innymi metodami bazującymi na próbkach.

Słowa kluczowe: mini-model, lokalna regresja, algorytm najbliższych sąsiadów, modelowania matematyczne, metody bazujące na próbkach

Wprowadzenie

W artykule przedstawiono zastosowanie metody mini-modeli opartej na hiperbryłach w zadaniach modelowania danych wielowymiarowych. W odróżnieniu od znanych i powszechnie stosowanych metod modelowania, takich jak: sieci neuronowe, sieci neurorozmyte, aproksymacja wielomianowa, które dokonują aproksymacji w całej domenie systemu, metoda MM nie tworzy modelu globalnego, jeżeli nie jest on wymagany. Operuje jedynie na danych znajdujących się w lokalnym otoczeniu punktu zapytania. Metoda MM stara się zidentyfikować

* Marcin Pietrzykowski, mgr inż., Zachodniopomorski Uniwersytet Technologiczny w Szczecinie, Wydział Informatyki, e-mail: mpietrzykowski@wi.zut.edu.pl

funkcję matematyczną opisującą zależność pomiędzy zmiennymi wejściowymi i wyjściowymi. Jednakże w procesie modelowania bardzo często jesteśmy zainteresowani jedynie odpowiedzią na konkretne zapytanie, np. Jak wysoka będzie stopa bezrobocia w sytuacji, gdy podaż pieniądza wynosi 950 mld zł, inflacja jest na poziomie 10%, a liczba ludności wynosi 37,5 mln osób? Odpowiedź na to pytanie wymaga jedynie danych znajdujących się w otoczeniu punktu zapytania, tj. danych, gdzie podaż pieniądza wynosi około 950 mld zł, inflacja jest na poziomie około 10%, a liczba ludności wynosi około 37,5 mln osób. Do odpowiedzi na to pytanie identyfikacja funkcji w całej domenie modelowanego systemu nie jest wymagana. Typowy punkt zapytania jest to zbiór zmiennych niezależnych o znanej wartości. Wartość zmiennej zależnej pozostaje nieznaną. Dla powyższego przykładu punkt zapytania będzie posiadał postać: albo w prostszej formie.

Koncepcja metody została opracowana przez prof. Andrzeja Piegata (Piegat i in., 2010, 2011). Początkowa wersja metody znalazła zastosowanie dla problemów dwu- i trójwymiarowych (Pietrzykowski, 2011a, 2011b, 2012), jednak badania wykazały, że zastosowanie jej dla problemów o większej liczbie wymiarów niesie ze sobą pewne trudności. W artykule przedstawiono modyfikację metody, która jest bardziej efektywna dla zastosowania w problemach wielowymiarowych. Głównym celem artykułu jest przebadanie metody MM bazującej na simpleksie i hipersześcianie oraz porównanie jej dokładności z innymi metodami bazującymi na próbkach: metodą k-najbliższych sąsiadów (k-NN) (Fix, Hodges, 1951; Fukunaga, Narendra, 1975; Bottou, Vapnik, 1992), siecią neuronową typu GRNN (Specht, 1991; Celikoglu, 2006), siecią neuronową typu RBF (Poggio, Girosi, 1990; Park, Wasenberg, 1991), *local linear regression* (Fan, 1992; Ruppert, Wand, 1994).

Metoda mini-modeli – algorytmy

Na wstępie należy podkreślić zasadniczą różnicę pomiędzy mini-modelem a modelem globalnym. Mini-model operuje tylko i wyłącznie na punktach znajdujących się w granicach jego zasięgu, które są ściśle określone i nie obejmują całej modelowanej dziedziny tak jak w przypadku modeli globalnych. Model globalny operuje i uczy się na wszystkich dostępnych próbkach. Omawianą różnicę obrazuje porównanie pomiędzy mini-modelem liniowym oraz modelem liniowym. Model liniowy (np. regresja liniowa) korzysta z wszystkich dostępnych próbek i obejmuje całą dziedzinę problemu. Mini-model liniowy tworzony jest na podsta-

wie próbek należących do ściśle określonego ciągłego obszaru przestrzeni wejść zwanego dalej otoczeniem punktu zapytania lub domeną mini-modelu. Co prawda w ekstremalnych przypadkach mini-model liniowy może rozszerzyć otoczenie punktu zapytania na całą dostępną dziedzinę, ale jest to sytuacja mogąca mieć miejsce jedynie podczas modelowania danych, w których zachodzi ścisła zależność liniowa, co jest raczej rzadkością.

W skład metody MM wchodzi dwie grupy algorytmów: algorytmy definiowania lokalnego otoczenia punktu zapytania oraz algorytmy modelowania matematycznego. Lokalne otoczenie punktu zapytania można zdefiniować jako obszar bryły geometrycznej określonej w przestrzeni wejść. Dla danych dwuwymiarowych będzie to odcinek, dla trójwymiarowych będzie to figura płaska (np. trójkąt lub czworokąt), dla danych czterowymiarowych będzie to nieregularna bryła wypukła (np. posiadająca cztery (simpleks) lub sześć ścian). Obszar mini-modelu uogólniony do przestrzeni n -wymiarowej przyjmuje postać $n-1$ -wymiarowej wypukłej i nieregularnej hiperbryły. Domena mini-modelu może przyjmować również postać hiperelipsoidalny (Pluciński, 2012a, 2012b, 2014).

Druga grupa algorytmów wchodząca w skład metody MM zawiera algorytmy modelowania matematycznego na próbkach leżących we wnętrzu domeny mini-modelu. Na tak zdefiniowanym obszarze możliwe jest zastosowanie dowolnej metody modelowania matematycznego, np. aproksymacji wielomianowej, wartości średniej, wnioskowania rozmytego.

W początkowych badaniach obejmujących mini-modele w przestrzeni dwu- i trójwymiarowej w procesie definiowania lokalnego otoczenia punktu zapytania operowano wierzchołkami bryły geometrycznej. Podejście to jest obciążone „przekleństwem wymiarowości”, gdyż liczba wierzchołków w bryle opartej na hipersześcianie wzrasta wykładniczo. Ponadto występuje wiele wad, takich jak problem współpłaszczyznowości wierzchołków należących do jednej ściany, problem manipulowania ścianą jako zbiorem wierzchołków, problem rozstrzygania zawierania się danego punktu we wnętrzu mini-modelu. Wyżej wymienione wady doprowadziły do powstania koncepcji manipulowania ścianą, jako całością oraz przeniesienia większości obliczeń z kartezjańskiego do sferycznego układu współrzędnych (Pietrzykowski, 2013, 2014). Centrum układu stanowi punkt zapytania. Ściana zdefiniowana jest poprzez pojedynczy punkt zwany generatorem ściany. Istnieje założenie, że ściana jest ortogonalna do wektora utworzonego przez ww. punkt oraz początek układu współrzędnych. Koncepcja ta posiada następujące zalety: brak problemu współpłaszczyznowości wierzchołków należących do jed-

nej ściany, stosunkowo proste obliczenia rozstrzygające o zawieraniu się punktu we wnętrzu domeny mini-modelu, zmniejszenie liczby parametrów potrzebnych do sterowania mini-modelem, stosunkowo proste manipulowanie położeniem ściany, stosunkowo prosta rozszerzalność mini-modelu do przestrzeni o wyższej wymiarowości. Zastosowanie tego podejścia pozwala uniknąć przekleństwa wymiarowości (mini-model bazujący na obszarze hipersześcianu w przestrzeni wymiarowej będzie posiadał 2^{n-1} wierzchołków, ale tylko $2(n-1)$ ścian). Ma to duże znaczenie, ponieważ pozwala uprościć trudny proces uczenia się mini-modelu.

Pierwsza część algorytmu polega na konwersji punktów danych z układu współrzędnych kartezjańskich do układu współrzędnych sferycznych (hipersferycznych) (Moon, Spencer, 1988; Bronshtein i in., 2007; Polyenin, Manzhirou, 2010). Transformacja następuje jedynie w przestrzeni wejść, zmienna wyjściowa pozostaje bez zmian. Oznacza to, że trójwymiarowy punkt danych zostanie transformowany do układu współrzędnych biegunowych, a punkt czterowymiarowy do układu współrzędnych sferycznych. W ogólnym przypadku $n + 1$ -wymiarowy punkt zostanie transformowany do układu współrzędnych opartych na n -wymiarowej sferze. Punkt zapytania $Q = \{x_{Q1}, x_{Q2}, \dots, x_{Qn}, y_Q\}$, staje się centrum układu współrzędnych. Wartości zmiennych wejściowych $x_{Q1}, x_{Q2}, \dots, x_{Qn}$ są znane, wartość zmiennej wyjściowej y_Q jest niewiadomą, którą zostanie obliczona w wyniku działania algorytmu. Wszystkie punkty danych p_i są konwertowane do hipersferycznego układu współrzędnych. Punkt jest określony przez promień $r \in [0, \infty)$, (odległość od centrum) i kąty $\varphi_{i1}, \varphi_{i2}, \dots, \varphi_{i(n-2)} \in [0; \pi)$, $\varphi_{i(n-1)} \in [0; 2\pi)$. Zbiór punktów P oznaczamy jako:

$$P = \{p_1, p_2, \dots, p_i, \dots, p_l\} \quad (1)$$

$$p_i = (x_{i1}, \dots, x_{in}, y_i) = (r_i, \varphi_{i1}, \dots, \varphi_{i(n-1)}, y_i)$$

Konwersja z układu współrzędnych opartych na hipersferze do układu współrzędnych kartezjańskich przyjmuje następującą postać:

$$\begin{aligned} x_1 &= r \cos(\varphi_1) \\ x_2 &= r \sin(\varphi_1) \cos(\varphi_2) \\ x_3 &= r \sin(\varphi_1) \sin(\varphi_2) \cos(\varphi_3) \\ &\vdots \\ x_{(n-1)} &= r \sin(\varphi_1) \sin(\varphi_{(n-2)}) \dots \cos(\varphi_{(n-1)}) \\ x_n &= r \sin(\varphi_1) \dots \sin(\varphi_{(n-2)}) \sin(\varphi_{(n-1)}) \end{aligned} \quad (2)$$

Konwersja odwrotna jest następująca:

$$\begin{aligned}
 r &= \sqrt{x_1^2 + x_2^2 + \dots + x_n^2} \\
 \varphi_1 &= \operatorname{arccot} \frac{x_1}{\sqrt{x_2^2 + x_3^2 + \dots + x_n^2}} \\
 \varphi_2 &= \operatorname{arccot} \frac{x_2}{\sqrt{x_3^2 + x_4^2 + \dots + x_n^2}} \\
 &\vdots \\
 \varphi_{n-2} &= \operatorname{arccot} \frac{x_{n-2}}{\sqrt{x_{n-1}^2 + x_n^2}} \\
 \varphi_{n-1} &= 2 \operatorname{arccot} \frac{x_{n-1} + \sqrt{x_{n-1}^2 + x_n^2}}{x_n}
 \end{aligned} \tag{3}$$

Transformacja ta jest jednoznaczna poza następującymi przypadkami:

- jeżeli wszystkie $x_{(k+1)}, \dots, x_n$ są równe zero, a $x_k > 0$ wtedy $\varphi_k = 0$,
- jeżeli wszystkie $x_{(k+1)}, \dots, x_n$ są równe zero, a $x_k < 0$ wtedy $\varphi_k = \pi$,
- jeżeli x_k, \dots, x_n są równe zero to wartość φ_k może wynosi zero.

W ogólnym przypadku domena mini-modelu jest wypukłą nieregularną hiperbryłą składającą się z J ścian. Dla simpleksu $J = n + 1$, dla bryły bazującej na hipersześcianie $J = 2n$, gdzie n jest to wymiarowość przestrzeni. Każda ściana j hiperbryły jest częścią płaszczyzny F_j . Płaszczyznę w sferycznym układzie współrzędnych możemy zdefiniować przy pomocy pojedynczego punktu G_j zwanego generatorem ściany. Istnieje założenie, że płaszczyzna F_j jest ortogonalna do wektora utworzonego przez generator ściany $\overrightarrow{QG_j}$. Każdą płaszczyznę możemy zdefiniować jako:

$$F_j = \left\{ p_i : \varphi_{ij} < \frac{\pi}{2} \wedge r_i = \frac{r_j}{\cos \varphi_{ij}} \right\} \tag{4}$$

gdzie φ_{ij} jest wartością kąta pomiędzy wektorami $\overrightarrow{QG_j}$ i $\overrightarrow{Qp_i}$. Kąt może zostać obliczony przy pomocy iloczynu skalarnego dla współrzędnych kartezjańskich:

$$\varphi_{n-2} = \arccos \frac{x_{i1}x_{j1} + \dots + x_{in}x_{jn}}{\sqrt{x_{i1}^2 + \dots + x_{in}^2} \sqrt{x_{j1}^2 + \dots + x_{jn}^2}} \tag{5}$$

Dla współrzędnych sferycznych mamy:

$$\begin{aligned} \varphi_{ij} = \arccos & (\cos\varphi_{i1} \cos\varphi_{j1} + \sin\varphi_{i1} \sin\varphi_{j1} (\\ & \cos\varphi_{i2} \cos\varphi_{j2} + \sin\varphi_{i2} \sin\varphi_{j2} (\\ & \dots \\ & \cos\varphi_{i(n-2)} \cos\varphi_{j(n-2)} + \sin\varphi_{i(n-2)} \sin\varphi_{j(n-2)} (\\ & \cos(\varphi_{i(n-1)} - \varphi_{j(n-1)})) \dots)) \end{aligned} \quad (6)$$

Płaszczyzna F_j dzieli całą przestrzeń wejść na dwie półprzestrzenie. Pierwsza półprzestrzeń zawiera punkty danych, które mogą zostać włączone w domenę mini-modelu. Możemy ją zdefiniować następująco:

$$I_j = \left\{ p_i: \varphi_{ij} \geq \frac{\pi}{2} \cup \left(p_{ij} < \frac{\pi}{2} \wedge r_i = \frac{r_j}{\cos \varphi_{ij}} \right) \right\} \quad (7)$$

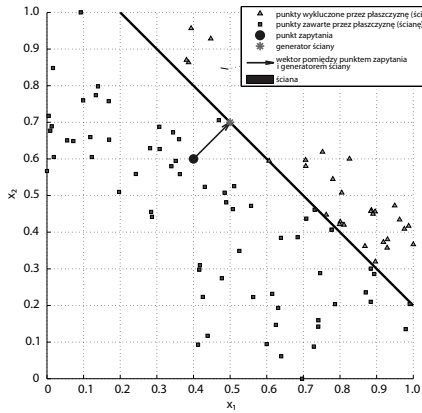
Druga półprzestrzeń zawiera punkty, które nie wejdą w skład domeny mini-modelu, i jest zdefiniowana jako:

$$E_j = \left\{ p_i: \varphi_{ij} < \frac{\pi}{2} \wedge r_i > \frac{r_j}{\cos \varphi_{ij}} \right\} \quad (8)$$

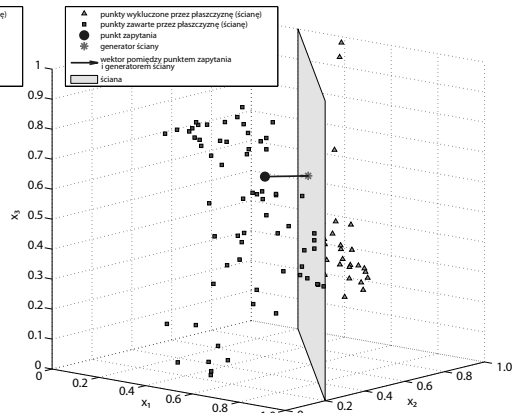
Każda płaszczyzna zawierająca w sobie ścianę bryły definiującej domenę mini-modelu, dzieli przestrzeń wejść w ten sposób bez względu na wymiarowość przestrzeni. Część wspólna wszystkich półprzestrzeni definiuje domenę mini-modelu i zawiera w sobie punkty uczące.

$$Z = I_1 \cap I_2 \cap \dots \cap I_j \quad (9)$$

Sposób, w jaki płaszczyzna dzieli przestrzeń wejść, przedstawiono na rysunku 1. Punkty oznaczone trójkątami na pewno zostaną wyłączone poza domenę mini-modelu, natomiast punkty oznaczone kwadratami mogą zostać w niej zawarte. O tym, czy punkt rzeczywiście wejdzie w skład domeny mini-modelu, decyduje jego położenie względem innych płaszczyzn. Tylko punkty, które zostaną zawarte przez wszystkie płaszczyzny, wejdą w skład MM. Jak ściany tworzą domenę mini-modelu dla problemu trójwymiarowego, pokazano na rysunku 2a. Na rysunku 2b została przedstawiona przykładową domenę dla problemu czterowymiarowego.



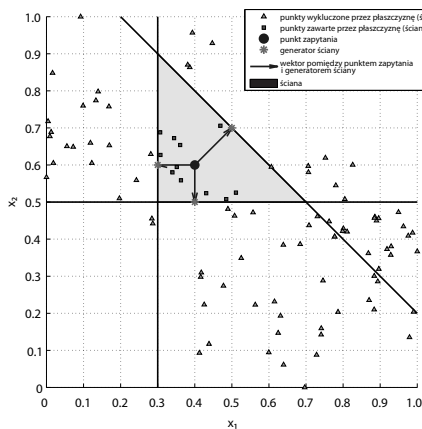
a) widok w przestrzeni dwuwymiarowej



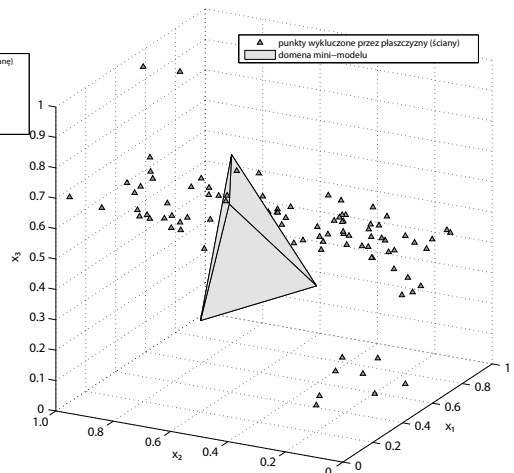
b) widok w przestrzeni trójwymiarowej

Rysunek 1. Przykład podziału przestrzeni przez płaszczyznę (ścianę)

Źródło: opracowanie własne.



a) przykład domeny dla problemu trójwymiarowego



b) przykład domeny dla problemu trójwymiarowego

Rysunek 2. Przykład domeny mini-modelu

Źródło: opracowanie własne.

Punkty danych wyselekcjonowane w powyższy sposób stają się danymi uczącymi dla metody modelowania matematycznego używanej przez mini-model.

Algorytmy uczenia mini-modeli

Uczenie mini-modelu składa się z dwóch części i polega na określeniu optymalnego lokalnego otoczenia punktu zapytania oraz nastrojeniu metody modelowania matematycznego używanego przez MM. Dzięki temu MM posiada zdolność do lokalnej adaptacji w wybranym obszarze dziedziny problemu.

Określenie optymalnego otoczenia punktu zapytania, czyli domeny MM, jest procesem heurystycznym i polega na zmianie położenia ścian bryły otaczającej domenę. Manewrowanie ścianami uczyniło proces uczenia prostszym, w porównaniu do poprzedniej wersji mini-modeli, w której manewrowano wierzchołkami. Pozwoliło to na wyeliminowanie wielu niedogodności, które zostały wymienione w poprzednim podrozdziale. Istnieje wiele potencjalnych lokalnych otoczeń punktu zapytania, które mogą zostać osiągnięte w procesie uczenia. Po zdefiniowaniu domeny mini-modelu, metoda używa punktów wchodzących w jego skład, jako danych uczących dla algorytmu modelowania matematycznego. Następnie mini-model oblicza błąd popełniany przez model matematyczny na próbkach uczących i obliczana jest wartość dla punktu zapytania. W kolejnym kroku mini-model stara się odnaleźć kolejną domenę, następnie oblicza błąd popełniany na próbkach oraz odpowiedź mini-modelu na punkt zapytania. Domena MM, w której został popełniony najmniejszy błąd, zostaje uznana za optymalną. Należy jednak zaznaczyć, że nie każda domena mini-modelu jest poprawna. Punkt zapytania musi znajdować się w domenie, a jej powierzchnia musi być bryłą (hiperbryłą) wypukłą. Poza tymi podstawowymi założeniami domena powinna spełniać następujące warunki:

- minimalna liczba punktów uczących wchodząca w skład mini-modelu,
- maksymalna liczba punktów uczących wchodząca w skład mini-modelu,
- stosunek pomiędzy wektorami $\overrightarrow{QG_j}$ o maksymalnej i minimalnej długości,
- punkt zapytania nie powinien być ekstrapolowany przez próbki uczące (nie zawsze jest to możliwe, a czasami nie jest wymagane).

Nie ma prostej reguły, jak dobrać wartości powyższych parametrów. Minimalna i maksymalna liczba punktów uczących zależy od zbioru danych. Jednakże dolna granica przedziału punktowego musi być wyższa niż wymiarowość problemu. Czasami, dla wybranego punktu zapytania, nie istnieje żadna domena mini-modelu spełniająca warunki początkowe. W takiej sytuacji mini-model nie jest w stanie zwrócić wiarygodnej odpowiedzi numerycznej.

W eksperymentach numerycznych procedura ucząca polegała na rotacji całej bryły geometrycznej i zmianie promieni generatorów ścian G_j . Parametry zmia-

ny położenia oraz kształtu bryły, takie jak kąt i kierunek obrotu oraz promień r_j , były losowe. Wyniki początkowych eksperymentów numerycznych pokazały, że mini-modele bardzo często były nadmiernie dopasowane do danych uczących w sytuacji, w której położenie płaszczyzn definiujących domenę mini-modelu było zmieniane w zupełnie dowolny sposób. Aby uniknąć przetrenowania, ograniczono możliwość nadmiernego rozciągnięcia figury poprzez ustalenie minimalnego stosunku pomiędzy najkrótszym a najdłuższym wektorem \overline{QG}_j na 0,5. Pojedyncza płaszczyzna nie zmienia swojego położenia, co oznacza, że kąty pomiędzy ścianami figury pozostają bez zmian. Bardzo ważną operacją w procesie uczenia jest operacja rotacji całej figury względem centrum układu współrzędnych, umożliwiająca modyfikację jej położenia w przestrzeni wejść.

Wyniki eksperymentów numerycznych

Metoda mini-modeli została porównana z metodami bazującymi na próbkach: algorytmem k-najbliższych sąsiadów, siecią neuronową typu GRNN, siecią neuronową typu RBF, *local linear regression*. Metoda mini-modeli oraz algorytm k-najbliższych sąsiadów pracują jedynie na danych z najbliższego otoczenia punktu zapytania. Obie metody mogą zostać porównane przy pomocy krzyżowej walidacji typu *leave-one-out*. Obie sieci neuronowe są metodami budującymi model globalny i obie zostały przetestowane przy pomocy 10-krotnej krzyżowej walidacji. Eksperymenty zostały wykonane na sześciu zbiorach danych z repozytorium *UCI Machine Learning* (UCI, 2015):

- Auto MPG – wszystkie dostępne atrybuty poza: x_8 – pochodzenie, x_9 – nazwa samochodu, 391 instancji,
- Boston Housing – x_3 – proporcja handlu niedetalicznego, x_5 – koncentracja tlenków azotu, x_6 – przeciętna liczba pokoi w lokalu, x_7 – procent budynków zajmowanych przez właściciela wybudowanych po roku 1940, x_{10} – wysokość podatku od nieruchomości, x_{11} – liczba uczniów przypadająca na jednego nauczyciela, x_{13} – procent ludności o niskim statusie społecznym, 506 instancji,
- Concrete Compressive Strength – wszystkie dostępne atrybuty, 1029 instancji,
- Concrete Slump Test – wszystkie dostępne atrybuty, 102 instancji,
- Servo – wszystkie dostępne atrybuty, 307 instancji,
- Yacht Hydrodynamics – wszystkie dostępne atrybuty, 307 instancji.

Wyniki eksperymentów numerycznych zostały zaprezentowane w tabeli 1. Błędy zaprezentowane w tabeli są błędami bezwzględnymi. Wszystkie ekspery-

menty zostały przeprowadzone przy optymalnym dobrze parametrów dla wszystkich testowanych metod. Metoda mini-modele została przebadana w dwóch wariantach: bazująca na simpleksie i bazująca na hipersześcianie. Mini-modele używały regresji liniowej jako algorytmu aproksymacji. Metoda odrzucała wyniki w przypadkach, w których nie potrafiła odnaleźć poprawnej domeny dla wybranego punktu zapytania. Wyniki eksperymentów pokazały, że w tego typu przypadkach błąd popełniany przez metodę z reguły był bardzo wysoki.

Tabela 1

Porównanie efektywności badanych metod

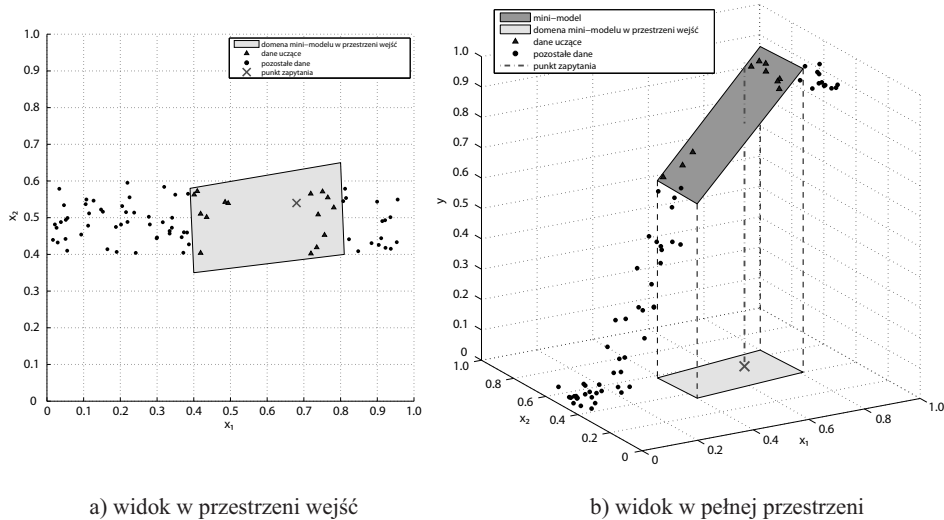
Zbiór danych	k-NN		Metoda mini-modele			GRNN		RBF			Local Linear Regression	
			bazująca na simpleksie	bazująca na hipersześcianie								
	błąd	k	błąd	błąd	liczba próbek	błąd	zasięg	błąd	zasięg	błąd uczący	błąd	zasięg
Auto MPG	0,0531	3	0,0529	0,0534	20–60	0,0502	0,08	0,0512	1,000	0,004	0,0498	0,09
Housing	0,0567	4	0,0551	0,0536	25–40	0,0545	0,1	0,0517	0,800	0,004	0,0534	0,08
Concrete	0,0722	1	0,0483	0,0465	20–40	0,0699	0,03	0,0511	1,000	0,003	0,0705	0,06
Slump Test	0,0600	2	0,0500	0,0590	8–20	0,0621	0,1	0,0114	2,000	0,00008	0,0554	0,14
Servo	0,0437	3	0,0493	0,0529	10–25	0,0381	0,05	0,0528	0,700	0,002	0,0378	0,06
Yacht	0,0371	2	0,0167	0,0127	12–25	0,0385	0,05	0,0220	2,900	0,003	0,0296	0,03

Źródło: opracowanie własne.

Podsumowanie

Wyniki eksperymentów numerycznych udowodniły, że dokładność mini-modele jest wysoka w porównaniu z innymi metodami bazującymi na próbkach. Zaproponowana wersja metody bazująca na hipersześcianie jest w stanie modelować problemy wielowymiarowe, a jej dokładność jest porównywalna z wersją metody opartą na simpleksie. Mini-modele posiadają bardzo zaawansowane właściwości ekstrapolacyjne. Metoda potrafi wziąć pod uwagę tendencję występującą w otoczeniu punktu zapytania, w czym przejawia się jej przewaga nad metodą najbliższych sąsiadów. Ponadto metoda MM potrafi „przeskoczyć” lukę informa-

cyjną i zawrzeć w swojej domenie dane znajdujące się po „drugiej stronie” luki informacyjnej. Przykład takiej sytuacji zilustrowano na rysunku 3.



Rysunek 3. Przykład właściwości ekstrapolacyjnych metody mini-modeli

Źródło: opracowanie własne.

Kolejną zaletą metody MM jest zdolność do wykrycia sytuacji, w których mini-model nie jest w stanie spełnić początkowych kryteriów, a tym samym zwrócić wiarygodnej odpowiedzi na wybrany punkt zapytania. Wyniki eksperymentów numerycznych potwierdziły, że w tego typu sytuacjach błąd popełniany przez metodę z reguły był bardzo wysoki. Pewną wadą metody jest natomiast większe zapotrzebowanie na próbki uczące niż w przypadku metody k-NN. Algorytm uczenia metody MM jest częściowo heurystyczny i niekiedy wyniki mogą się nieznacznie różnić.

Bibliografia

- Bottou L., Vapnik V. (1992), *Local Learning Algorithms*, „Neural Computation”, vol. 4, iss. 6, s. 888–900.
- Bronshtein I., Semendyayev K., Musiol G., Muhlig H. (2007), *Handbook of Mathematics*, Springer.

- Celikoglu H.B. (2006), *Application of radial basis function and generalized regression neural networks in non-linear utility function specification for travel mode choice modelling*, „Mathematical and Computer Modelling”, vol. 44, iss. 7–8, s. 640–658.
- Fan J.Q. (1992), *Design-adaptive nonparametric regression*, „Journal of the American Statistical Association”, vol. 87, iss. 420, s. 998–1004.
- Fix E., Hodges J.L. (1951), *Discriminatory analysis, nonparametric discrimination: Consistency properties*, Randolph Field, s. 1–21.
- Hollash S.R. (1991), *Four Space Visualization of 4D Objects*, Arizona State University.
- Fukunaga K., Narendra P.M. (1975), *Branch and bound algorithm for computing k-nearest neighbors*, „IEEE Transactions on Computers”, vol. C24, iss. 7, s. 750–753.
- Moon P., Spencer D. (1988), *Field theory handbook: including coordinate systems, differential equations, and their solutions*, Springer.
- Park J., Wasenberg J. (1991), *Universal approximation using radial basis functions network*, „Neural Computation”, vol. 3, s. 246–257.
- Piegat A., Wąsikowska B., Korzeń M. (2010), *Zastosowanie samouczącego się trzypunktowego modelu do modelowania stopy bezrobocia w Polsce*, „Studia Informatyka”, nr 27, s. 59–69.
- Piegat A., Wąsikowska B., Korzeń M. (2011), *Differences between the method of mini-models and the k-nearest neighbors an example of modeling unemployment rate in Poland*, Information Systems in Management IX-Business Intelligence and Knowledge Management, WULS Press, Warszawa, s. 34–43.
- Pietrzykowski M. (2011a), *Comparison of effectiveness of linear mini-models with some methods of modelling*, Młodzi Naukowcy dla Polskiej Nauki. CREATIVETIME, Kraków, s. 113–123.
- Pietrzykowski M. (2011b), *The use of linear and nonlinear mini-models in process of data modeling in a 2D-space*, Nowe trendy w Naukach Inżynierskich. CREATIVETIME, Kraków, s. 100–108.
- Pietrzykowski M. (2012), *Effectiveness of mini-models method when data modelling within a 2D-space in an information deficiency situation*, „Journal of Theoretical and Applied Computer Science”, vol. 6, no. 3, s. 21–27.
- Pietrzykowski M. (2013), *Mini-models working in 3D space based on polar coordinate system*, Nowe trendy w Naukach Inżynierskich 4. Tom II, CREATIVETIME, Kraków, s. 117–125.
- Pietrzykowski M. (2014), *Comparison between mini-models based on multidimensional polytopes and k-nearest neighbor method: case study of 4D and 5D problems*, „Advances in Intelligent Systems and Computing”, vol. 342, s. 107–118.

- Pluciński M. (2012a), *Mini-models – Local Regression Models for the Function Approximation Learning*, w: Proceedings of ICAISC 2012, Part II, LNCS 7268, red. L. Rutkowski, Springer-Verlag, Berlin–Heidelberg, s. 160–167.
- Pluciński M. (2012b), *Nonlinear ellipsoidal mini-models – application for the function approximation task*, „Przegląd Elektrotechniczny”, r. 88, nr 10b, s. 247–251.
- Pluciński M. (2014), *Application of Mini-Models to the Interval Information Granules Processing*, „Advances in Intelligent Systems and Computing”, vol. 342, s. 37–48.
- Poggio T., Girosi F. (1990), *Network for approximation and learning*, „Proceedings of the IEEE”, vol. 78, no. 9, s. 1481–1497.
- Polyanin A., Manzhirov A. (2010), *Handbook of Mathematics for Engineers and Scientists*, Taylor & Francis.
- Ruppert D., Wand M.P. (1994), *Multivariate locally weighted least-squares regression*, „Annals of Statistics”, vol. 22, iss. 3, s. 1346–1370.
- Specht D.F. (1991), *A General Regression Neural Network*, „IEEE Transactions on Neural Networks”, vol. 2, no. 6, s. 568–576.
- Uci machine learning repository*, <http://archive.ics.uci.edu/ml> (25.04.2015).

APPLICATION OF MINI-MODELS METHOD BASED ON HYPERCUBE IN THE MODELING PROCESS OF MULTIDIMENSIONAL DATA

Summary

The article presents self-learning method of mini-models (MM-method) based on polytopes in multidimensional space. The method is new and is an object of intensive research. MM method is the instance based learning method and uses data samples only from the local neighborhood of the query point. Group of points which are used in the model-learning process is constrained by a polytope area. The MM-method can on a defined local area use any approximation algorithm to compute mini-model answer for the query point. The article describes a learning technique based on hyper-spherical coordinate system. The method was used in the modeling task with multidimensional datasets. The results of numerical experiments were compared with other instance based methods.

Translated by Marcin Pietrzykowski

Keywords: mini-model, local regression, k-nearest neighbor, mathematical modeling, instance based learning

