



Politechnika Śląska, Instytut Informatyki



Ewolucyjne uczenie maszynowe

... czyli o moich zainteresowaniach naukowych

Jakub Nalepa

Instytut Informatyki, Politechnika Śląska, Gliwice
Future Processing, Gliwice

`jakub.nalepa@polsl.pl`

`jnalepa@ieee.org`

WWW: <http://sun.aei.polsl.pl/~jnalepa/>

Uroczystość wręczenia Nagrody im. Witolda Lipskiego

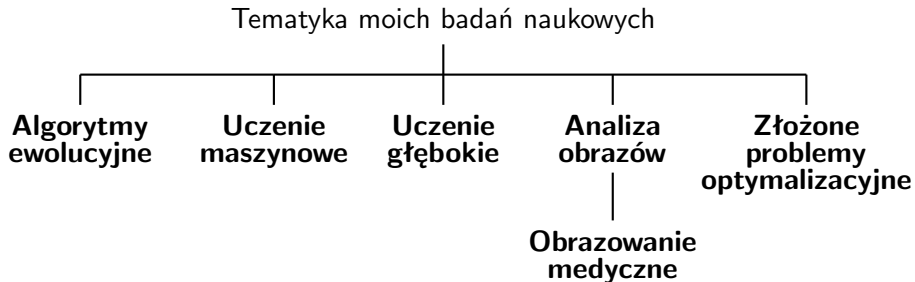
Warszawa, 12 października 2017

Plan prezentacji

- 1 Wprowadzenie
- 2 Dobór zbiorów treningowych dla SVM-a
- 3 Automatyczny dobór parametrów sieci głębokich
- 4 Parę słów o segmentacji obrazów medycznych. . .

Tematyka moich badań naukowych

Tematyka moich badań naukowych



Tematyka moich badań naukowych

• Algorytmy ewolucyjne

- Algorytmy genetyczne i memetyczne
- Schematy adaptacji w algorytmach ewolucyjnych
- Równoległe algorytmy ewolucyjne
- Inne algorytmy ewolucyjne. . .

• Uczenie maszynowe

- Dobór zbiorów treningowych dla maszyny wektorów wspierających (ang. *support vector machine*, SVM)
- Dobór funkcji jądrowej i jej parametrów dla SVM-a
- Jednoczesny dobór zbiorów treningowych i funkcji jądrowych dla SVM-a

• Uczenie głębokie (ang. *deep learning*)

- Automatyczny dobór hiperparametrów sieci głębokich
- Segmentacja obrazów medycznych
- Dogenerowywanie danych (ang. *data augmentation*)

Tematyka moich badań naukowych

• Algorytmy ewolucyjne

- Algorytmy genetyczne i memetyczne
- Schematy adaptacji w algorytmach ewolucyjnych
- Równoległe algorytmy ewolucyjne
- Inne algorytmy ewolucyjne. . .

• Uczenie maszynowe

- Dobór zbiorów treningowych dla maszyny wektorów podpierających (ang. *support vector machine*, SVM)
- Dobór funkcji jądrowej i jej parametrów dla SVM-a
- Jednoczesny dobór zbiorów treningowych i funkcji jądrowych dla SVM-a

• Uczenie głębokie (ang. *deep learning*)

- Automatyczny dobór hiperparametrów sieci głębokich
- Segmentacja obrazów medycznych
- Dogenerowywanie danych (ang. *data augmentation*)

Tematyka moich badań naukowych

• Algorytmy ewolucyjne

- Algorytmy genetyczne i memetyczne
- Schematy adaptacji w algorytmach ewolucyjnych
- Równoległe algorytmy ewolucyjne
- Inne algorytmy ewolucyjne. . .

• Uczenie maszynowe

- Dobór zbiorów treningowych dla maszyny wektorów podpierających (ang. *support vector machine*, SVM)
- Dobór funkcji jądrowej i jej parametrów dla SVM-a
- Jednoczesny dobór zbiorów treningowych i funkcji jądrowych dla SVM-a

• Uczenie głębokie (ang. *deep learning*)

- Automatyczny dobór hiperparametrów sieci głębokich
- Segmentacja obrazów medycznych
- Dogenerowywanie danych (ang. *data augmentation*)

Tematyka moich badań naukowych

● Analiza obrazów

- Wykrywanie i segmentacja obszarów występowania ludzkiej skóry w obrazach barwnych
- Rozpoznawanie gestów

● Obrazowanie medyczne

- Segmentacja obrazów medycznych – różne modalności (CT, MR, PET, PET/CT) i organy
- Analiza teksturalna obrazów medycznych
- Interaktywna analiza i współdzielenie obrazów medycznych

● Rozwiązywanie złożonych problemów optymalizacyjnych

- Problem trasowania pojazdów z oknami czasowymi
- Problem odbioru i dostawy z oknami czasowymi
- Problem szeregowania zadań
 - Algorytmy ewolucyjne (genetyczne i memetyczne)
 - Algorytmy równoległe
 - Uczenie maszynowe w rozwiązywaniu problemów transportowych

Tematyka moich badań naukowych

● Analiza obrazów

- Wykrywanie i segmentacja obszarów występowania ludzkiej skóry w obrazach barwnych
- Rozpoznawanie gestów

● Obrazowanie medyczne

- Segmentacja obrazów medycznych – różne modalności (CT, MR, PET, PET/CT) i organy
- Analiza teksturalna obrazów medycznych
- Interaktywna analiza i współdzielenie obrazów medycznych

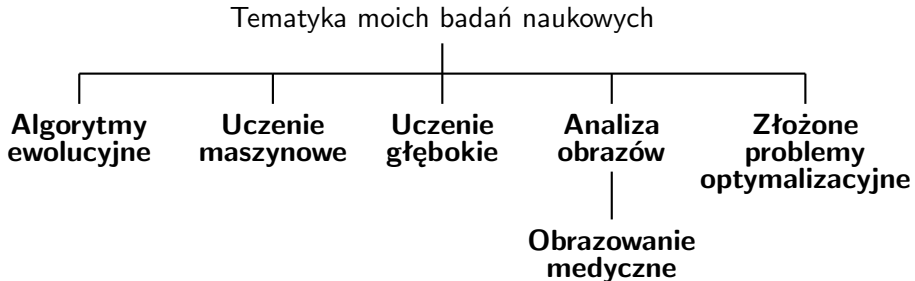
● Rozwiązywanie złożonych problemów optymalizacyjnych

- Problem trasowania pojazdów z oknami czasowymi
- Problem odbioru i dostawy z oknami czasowymi
- Problem szeregowania zadań
 - Algorytmy ewolucyjne (genetyczne i memetyczne)
 - Algorytmy równoległe
 - Uczenie maszynowe w rozwiązywaniu problemów transportowych

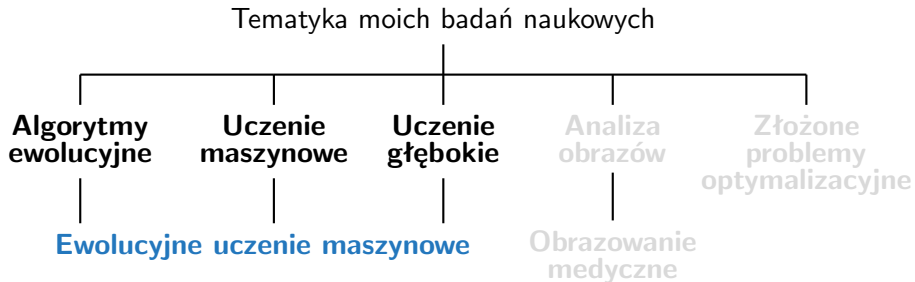
Tematyka moich badań naukowych

- **Analiza obrazów**
 - Wykrywanie i segmentacja obszarów występowania ludzkiej skóry w obrazach barwnych
 - Rozpoznawanie gestów
- **Obrazowanie medyczne**
 - Segmentacja obrazów medycznych – różne modalności (CT, MR, PET, PET/CT) i organy
 - Analiza teksturalna obrazów medycznych
 - Interaktywna analiza i współdzielenie obrazów medycznych
- **Rozwiązywanie złożonych problemów optymalizacyjnych**
 - Problem trasowania pojazdów z oknami czasowymi
 - Problem odbioru i dostawy z oknami czasowymi
 - Problem szeregowania zadań
 - Algorytmy ewolucyjne (genetyczne i memetyczne)
 - Algorytmy równoległe
 - Uczenie maszynowe w rozwiązywaniu problemów transportowych

Tematyka moich badań naukowych



Tematyka moich badań naukowych



Tematyka moich badań naukowych

- **Algorytmy ewolucyjne**

- Algorytmy genetyczne i memetyczne
- Schematy adaptacji w algorytmach ewolucyjnych
- Równoległe algorytmy ewolucyjne
- Inne algorytmy ewolucyjne...

- **Uczenie maszynowe**

- Dobór zbiorów treningowych dla maszyny wektorów podpierających (ang. *support vector machine*, SVM)
- Dobór funkcji jądrowej i jej parametrów dla SVM-a
- Jednoczesny dobór zbiorów treningowych i funkcji jądrowych dla SVM-a

- **Uczenie głębokie (ang. *deep learning*)**

- Automatyczny dobór hiperparametrów sieci głębokich
- Segmentacja obrazów medycznych
- Dogenerowywanie danych (ang. *data augmentation*)

Tematyka moich badań naukowych

- **Algorytmy ewolucyjne**

- Algorytmy genetyczne i memetyczne
- Schematy adaptacji w algorytmach ewolucyjnych
- Równoległe algorytmy ewolucyjne
- Inne algorytmy ewolucyjne...

- **Uczenie maszynowe**

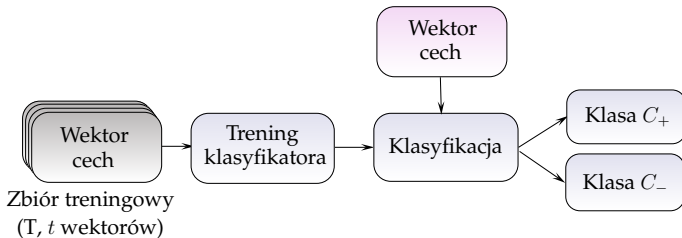
- Dobór zbiorów treningowych dla maszyny wektorów podpierających (ang. *support vector machine*, SVM)
- Dobór funkcji jądrowej i jej parametrów dla SVM-a
- Jednoczesny dobór zbiorów treningowych i funkcji jądrowych dla SVM-a

- **Uczenie głębokie (ang. *deep learning*)**

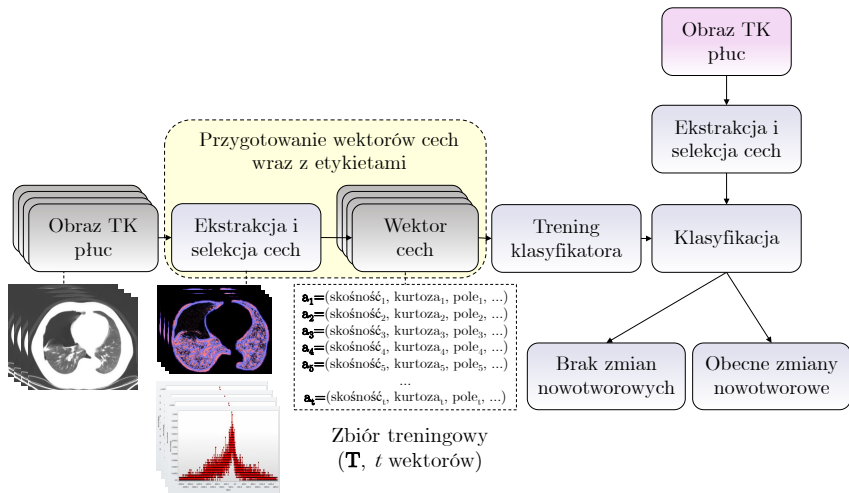
- Automatyczny dobór hiperparametrów sieci głębokich
- Segmentacja obrazów medycznych
- Dogenerowanie danych (ang. *data augmentation*)

Dobór zbiorów treningowych dla SVM-a

Klasyfikacja binarna (dwuklasowa)

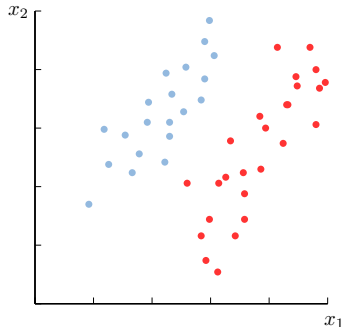


Klasyfikacja binarna (dwuklasowa) – przykład



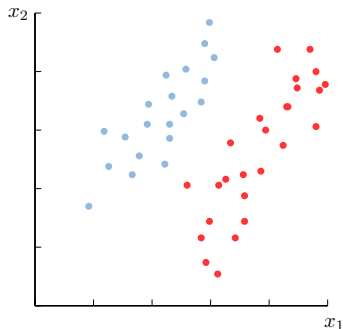
Linowy SVM – dane separowalne liniowo

Zbiór treningowy

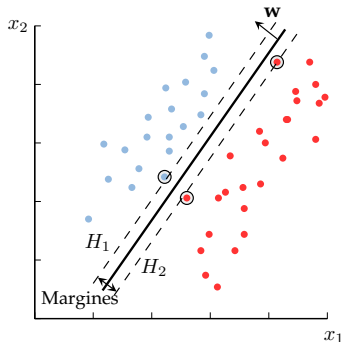


Liniowy SVM – dane separowalne liniowo

Zbiór treningowy



Hiperpłaszczyzna separująca



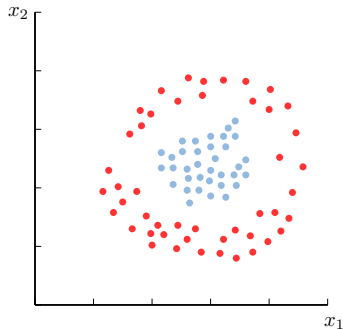
Funkcja decyzyjna: $f(\mathbf{a}) = \text{sgn}(\mathbf{w}^T \mathbf{a} + b)$

\mathbf{a} – wektor cech

b – wyraz wolny

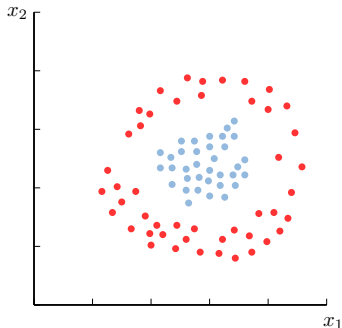
Nieliniowy SVM

Zbiór treningowy

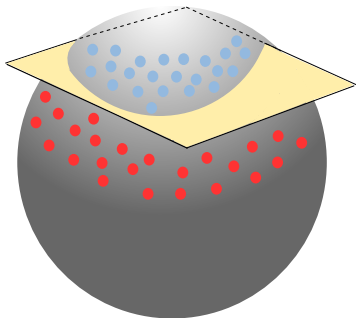


Nieliniowy SVM

Zbiór treningowy



Hiperpłaszczyzna separująca



Wprowadzenie **funkcji jądrowych** (ang. *kernel functions*)

$$\mathcal{K}(\mathbf{a}, \mathbf{a}') = \phi(\mathbf{a})^T \phi(\mathbf{a}')$$

Trening klasyfikatora SVM

- Zagadnienie programowania kwadratowego z ograniczeniami
 - Złożoność czasowa: $O(t^3)$
 - Złożoność pamięciowa: $O(t^2)$

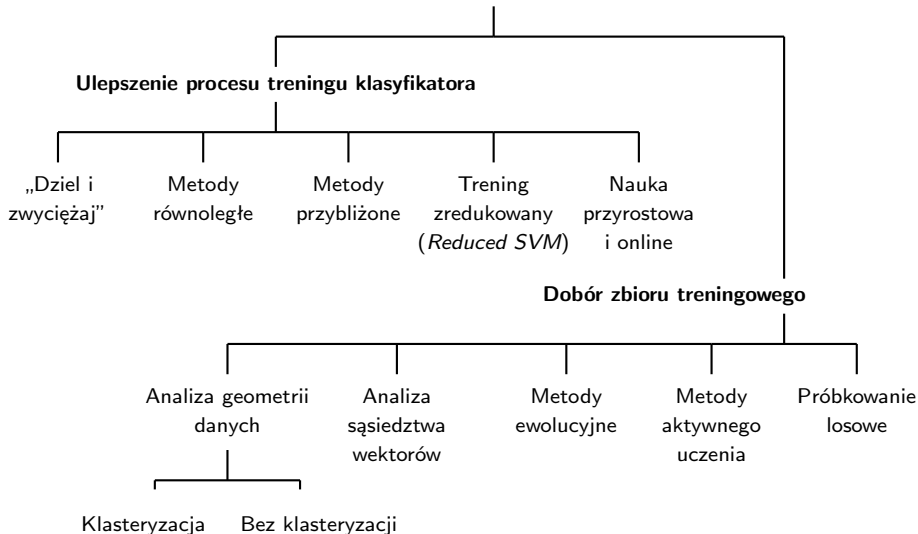
t – liczba wektorów w zbiorze treningowym
- Problem dużych i „trudnych” zbiorów treningowych
 - Długi czas trwania treningu
 - Czasami trening jest niewykonalny
 - Pośrednio – długi czas trwania klasyfikacji
 - Zależny liniowo od liczby wektorów podpierających
 - Wpływ na jakość klasyfikacji

Trening klasyfikatora SVM

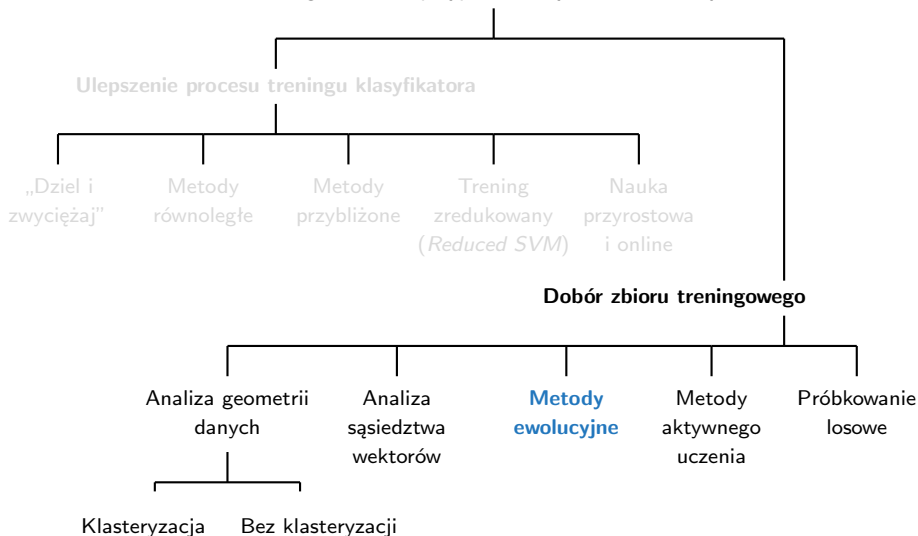
- Zagadnienie programowania kwadratowego z ograniczeniami
 - Złożoność czasowa: $O(t^3)$
 - Złożoność pamięciowa: $O(t^2)$

t – liczba wektorów w zbiorze treningowym
- Problem dużych i „trudnych” zbiorów treningowych
 - Długi czas trwania treningu
 - Czasami trening jest niewykonalny
 - Pośrednio – długi czas trwania klasyfikacji
 - Zależny liniowo od liczby wektorów podpierających
 - Wpływ na jakość klasyfikacji

Trening SVM-a w przypadku dużych zbiorów danych



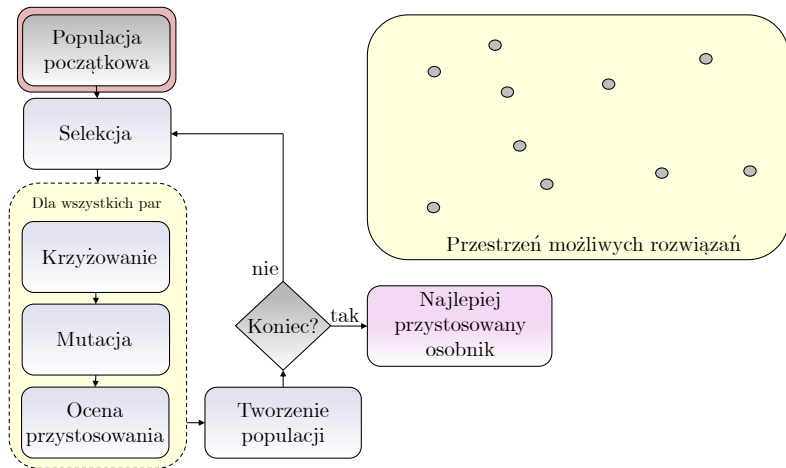
Trening SVM-a w przypadku dużych zbiorów danych



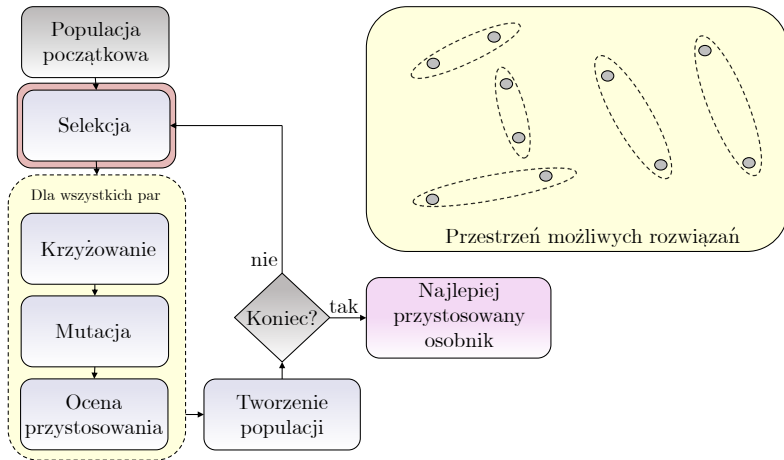
Dobór zbioru treningowego – motywacja

- 1 **Poprawa skuteczności** klasyfikacji klasyfikatora SVM
- 2 **Zmniejszenie** liczby wektorów podpierających
- 3 **Umożliwienie** przeprowadzenia treningu klasyfikatora SVM w przypadku bardzo dużych zbiorów danych

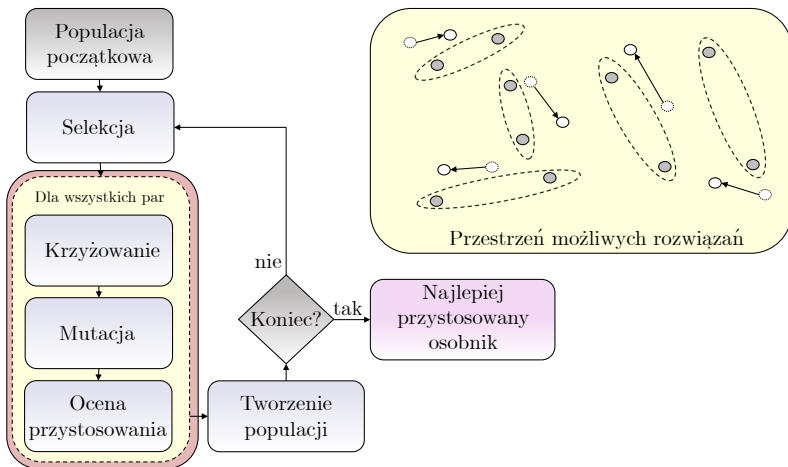
Algorytmy genetyczne



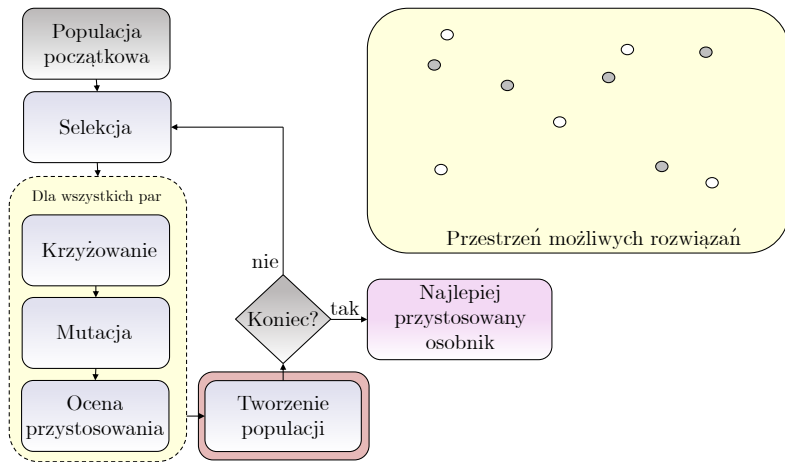
Algorytmy genetyczne



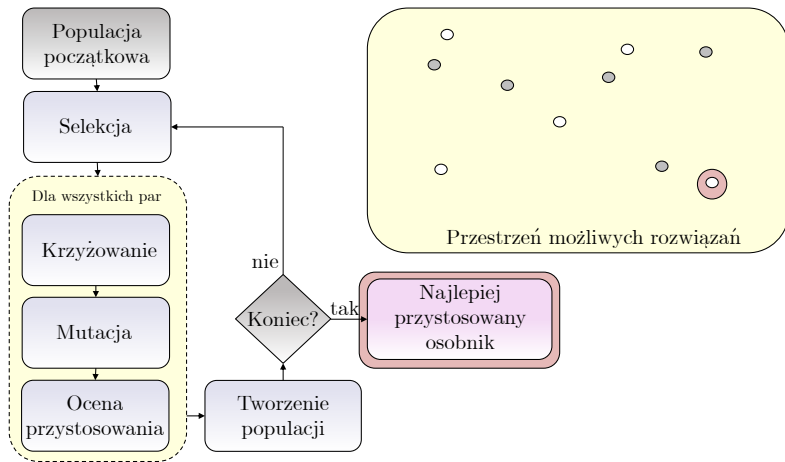
Algorytmy genetyczne



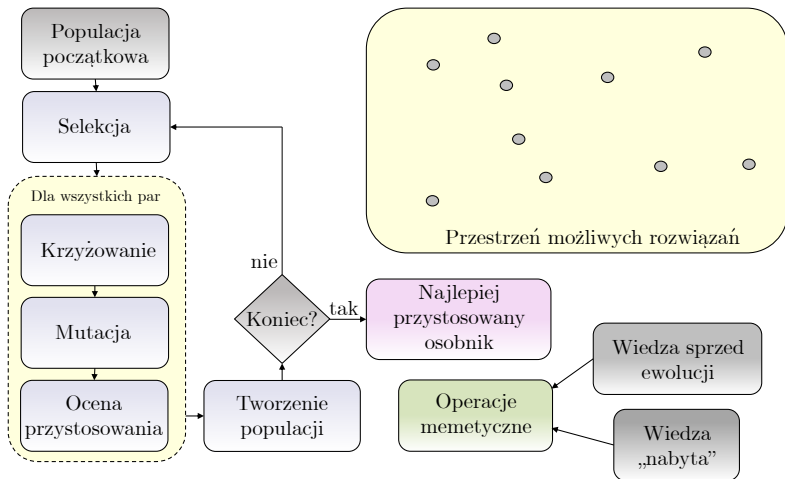
Algorytmy genetyczne



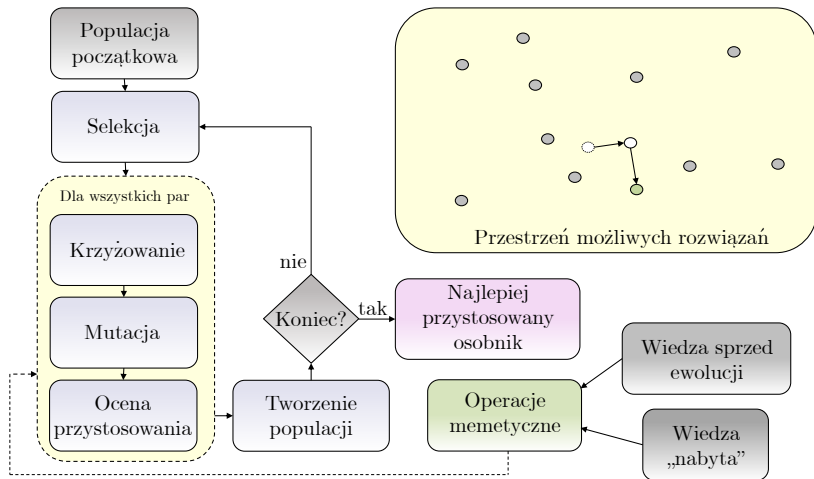
Algorytmy genetyczne



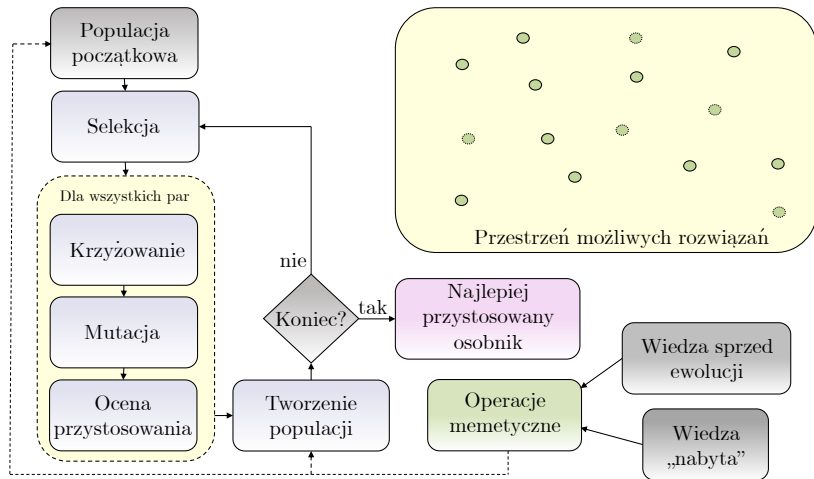
Algorytmy memetyczne (hybrydowe algorytmy ewolucyjne)



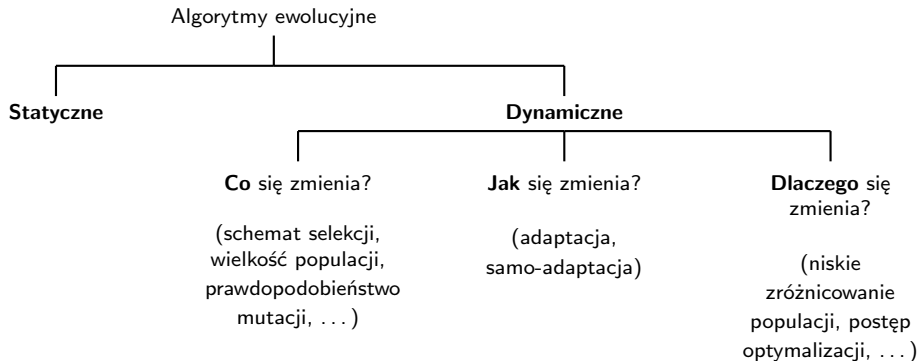
Algorytmy memetyczne (hybrydowe algorytmy ewolucyjne)



Algorytmy memetyczne (hybrydowe algorytmy ewolucyjne)



Adaptacyjność w algorytmach ewolucyjnych



A. Eiben et al. Parameter control in evolutionary algorithms. *Evolutionary Computation*, IEEE Transactions on, 3(2):124–141, 1999.

Ewolucyjny dobór zbiorów treningowych dla SVM-a

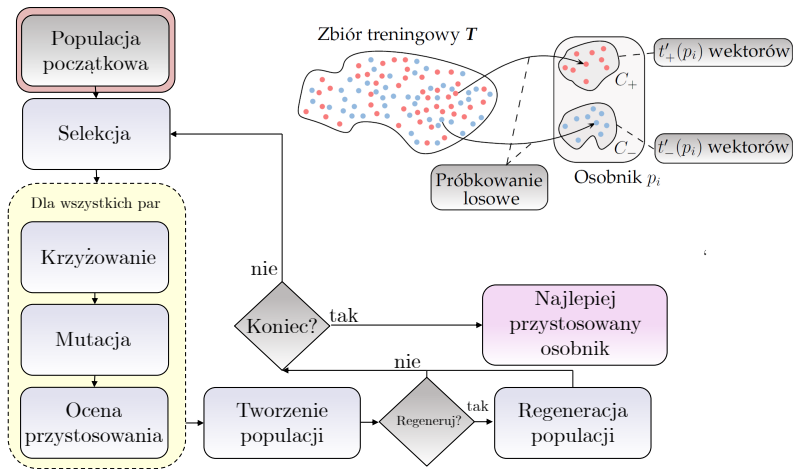
Zaproponowane algorytmy ewolucyjne

- 1 Algorytm genetyczny (ang. *genetic algorithm*, GASVM)
- 2 Adaptacyjny algorytm genetyczny (ang. *adaptive genetic algorithm*, AGA)
- 3 Dynamicznie adaptacyjny algorytm genetyczny (ang. *dynamically adaptive genetic algorithm*, DAGA)
- 4 Algorytm memetyczny (ang. *memetic algorithm*, MASVM)
- 5 Adaptacyjny algorytm memetyczny wykorzystujący dodatkową analizę geometrii danych w zbiorze treningowym (**P**roincipal **C**omponent **A**nalysis **A**daptive **M**emetic **A**lgorithm, PCA²MA)

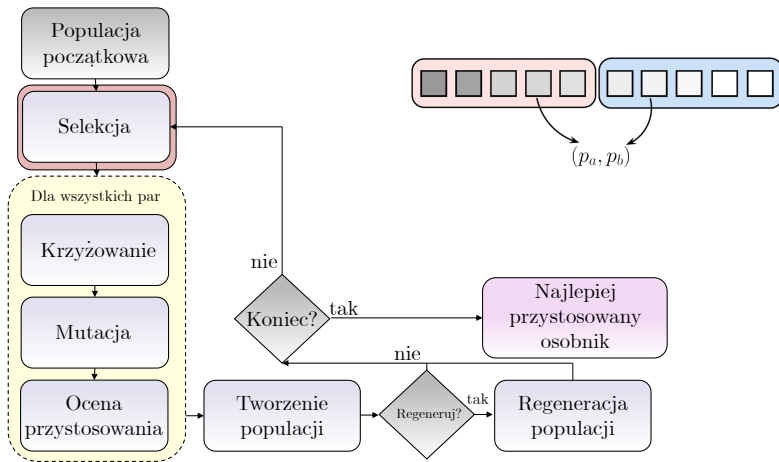
Algorytm genetyczny (GASVM)

Kawulok M., Nalepa J.: Support vector machines training data selection using a genetic algorithm, Structural, Syntactic, and Statistical Pattern Recognition, Joint IAPR International Workshop, SSPR & SPR 2012, Gimel'farb G., Hancock E., Imiya A., Kuijper A., Kudo M., Omachi S., Windeatt T., & Yamada K. (ed.), LNCS, vol. 7626, pp 557-565, DOI: 10.1007/978-3-642-34166-3 61, Springer, 2012.

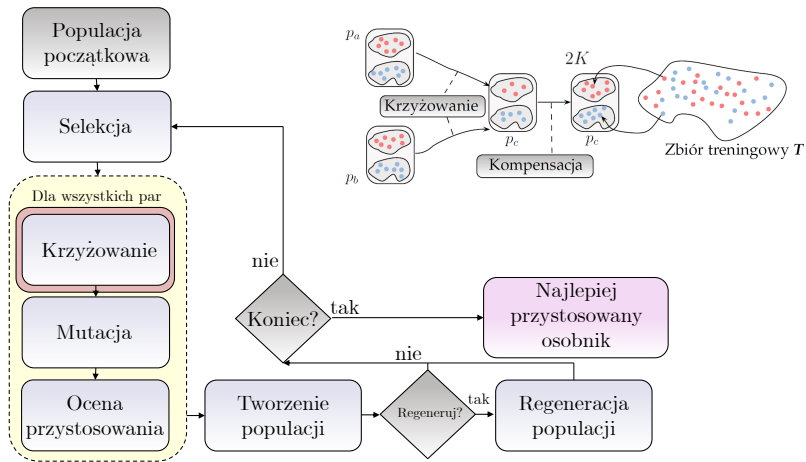
Algorytm genetyczny (GASVM)



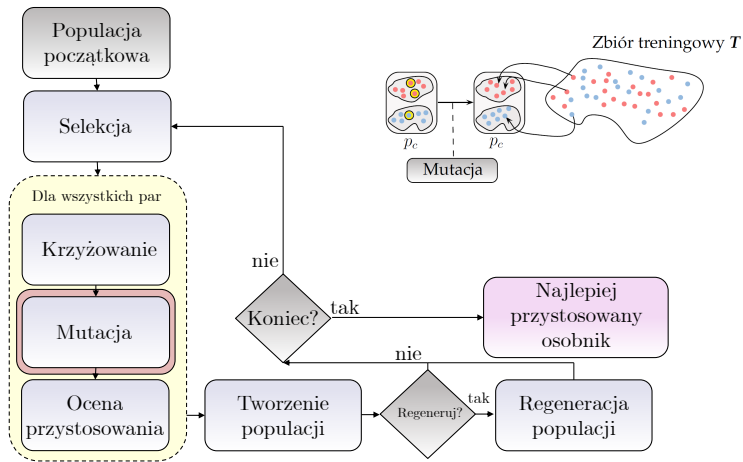
Algorytm genetyczny (GASVM)



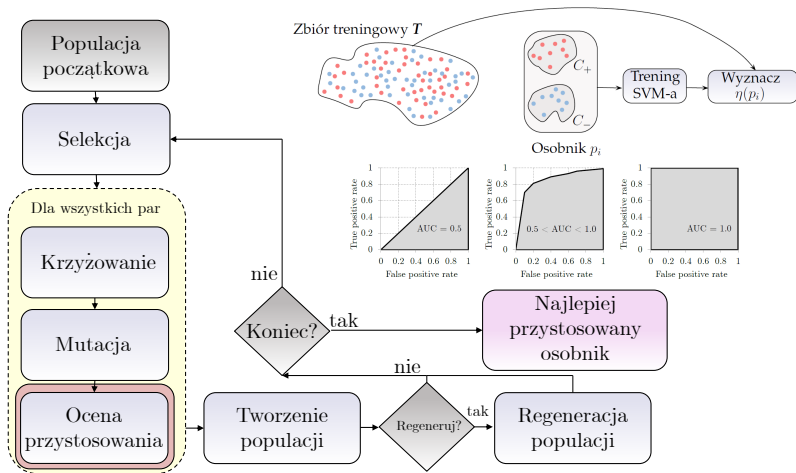
Algorytm genetyczny (GASVM)



Algorytm genetyczny (GASVM)



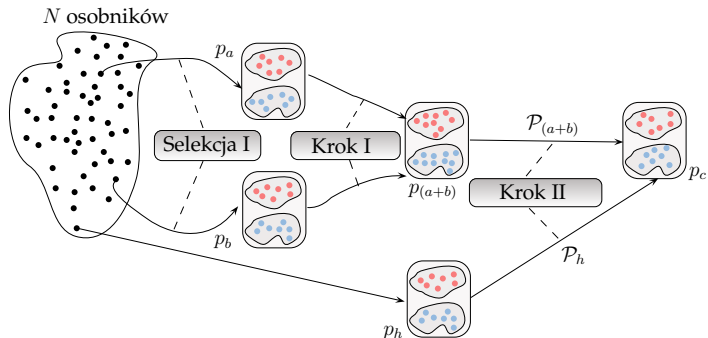
Algorytm genetyczny (GASVM)



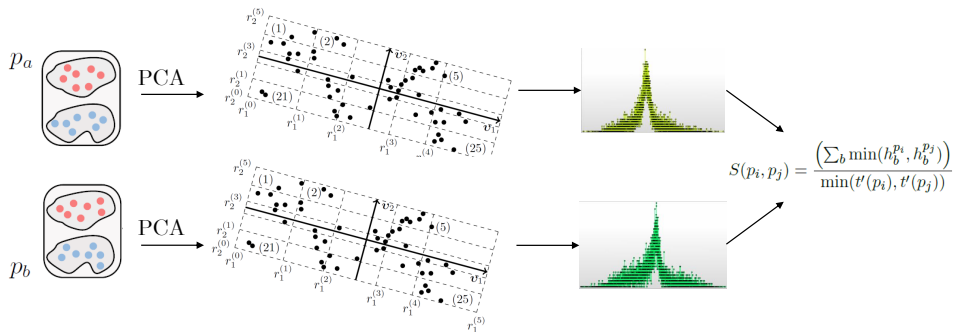
Adaptacyjny algorytm genetyczny (AGA)

Nalepa J., Kawulok M.: Adaptive genetic algorithm to select training data for support vector machines, Applications of Evolutionary Computation, 17th European Conference, EvoApplications 2014, Esparcia-Alcazar A. I. & Mora A. M., (ed.), LNCS, Volume 8602, pp 514-525, DOI: 10.1007/978-3-662-45523-4 42, Springer, 2014.

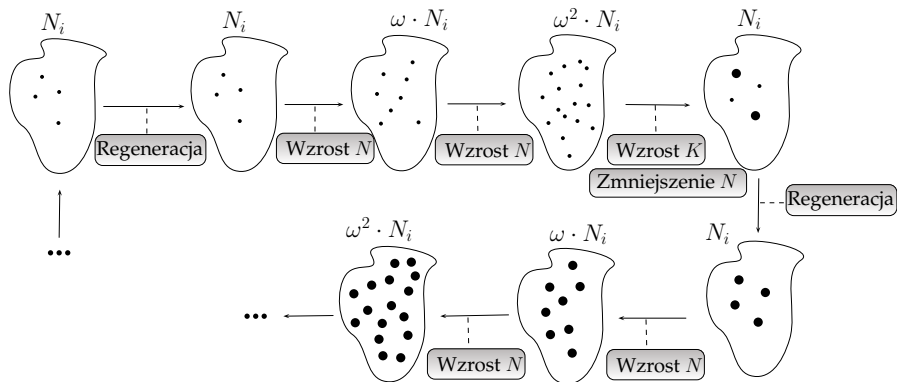
Dwuetaapowe tworzenie osobnika potomnego



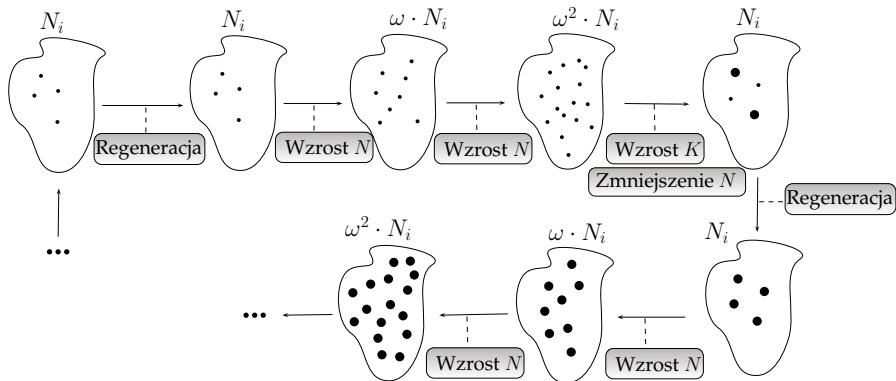
Nowa miara podobieństwa osobników



Adaptacyjna aktualizacja osobników i wielkości populacji



Adaptacyjna aktualizacja osobników i wielkości populacji



Wady: parametry (współczynniki rozrostu, maksymalna liczba regeneracji i generacji bez poprawy η , ...) wymagają strojenia

Dynamicznie adaptacyjny algorytm genetyczny (DAGA)

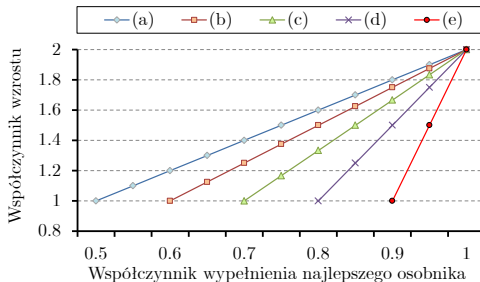
Kawulok M., Nalepa J.: Dynamically adaptive genetic algorithm to select training data for SVMs, Advances in Artificial Intelligence – IBERAMIA 2014, 14th Ibero-American Conference on AI, Bazzan A. L. C. & Pichara K., (ed.), LNCS, Volume 8864, pp 242-254, DOI: 10.1007/978-3-319-12027-0 20, Springer, 2014.

Dynamicznie adaptacyjny współczynnik wzrostu osobników

Współczynnik wypełnienia: $\varrho_{SV}(p^B) = \frac{s(p^B)}{t'(p^B)}$

Wartość pożądana: \mathcal{T}_{SV}

Rozrost osobników zależny od współczynnika wypełnienia najlepszego osobnika



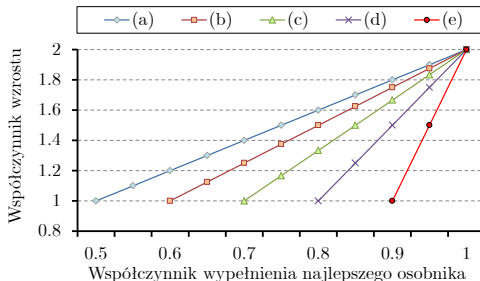
(a): $\mathcal{T}_{SV} = 0.5$, (b): $\mathcal{T}_{SV} = 0.6$, (c): $\mathcal{T}_{SV} = 0.7$, (d): $\mathcal{T}_{SV} = 0.8$, (e): $\mathcal{T}_{SV} = 0.9$

Dynamicznie adaptacyjny współczynnik wzrostu osobników

Współczynnik wypełnienia: $\varrho_{SV}(p^B) = \frac{s(p^B)}{t'(p^B)}$

Wartość pożądana: \mathcal{T}_{SV}

Rozrost osobników zależny od współczynnika wypełnienia najlepszego osobnika



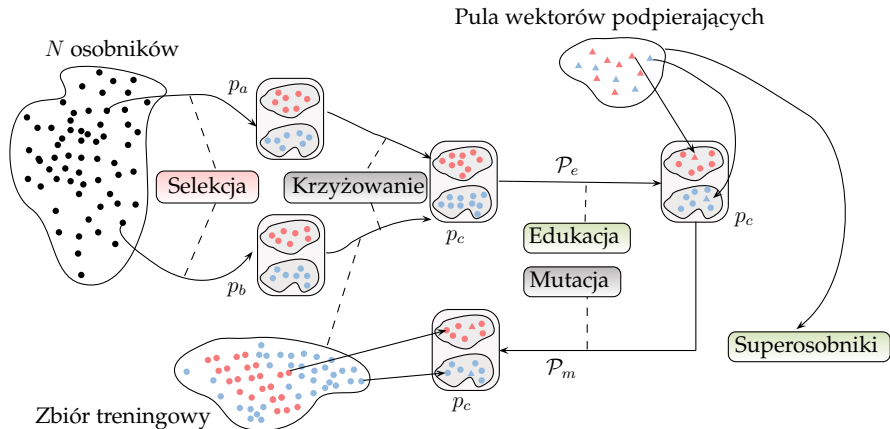
(a): $\mathcal{T}_{SV} = 0.5$, (b): $\mathcal{T}_{SV} = 0.6$, (c): $\mathcal{T}_{SV} = 0.7$, (d): $\mathcal{T}_{SV} = 0.8$, (e): $\mathcal{T}_{SV} = 0.9$

Wady: parametry (pożądana wartość wsp. wypełnienia, minimalna liczba generacji po wzroście K , ...) wymagają strojenia

Algorytm memetyczny (MASVM)

Nalepa J., Kawulok M.: A memetic algorithm to select training data for support vector machines, Proceedings of the 2014 Annual Conference on Genetic and Evolutionary Computation, GECCO 2014, pp 573-580, DOI: 10.1145/2576768.2598370, ACM, 2014.

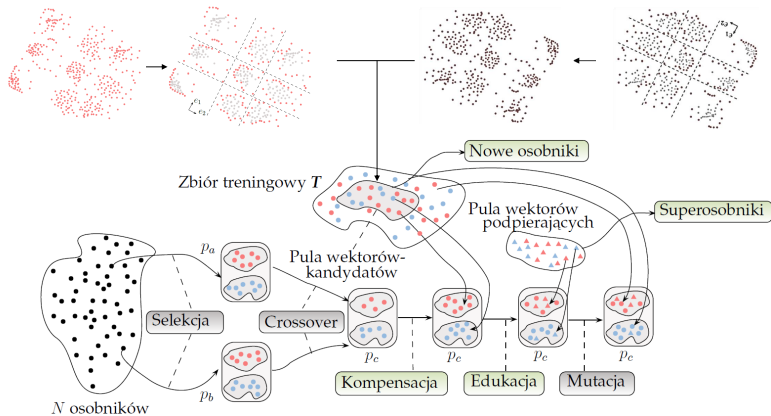
Wiedza „nabyta” – pula wektorów podpierających



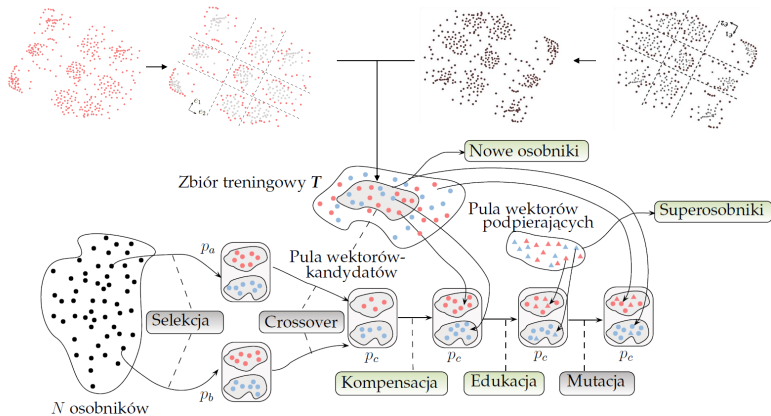
Adaptacyjny algorytm memetyczny wykorzystujący dodatkową analizę geometrii danych w zbiorze treningowym (PCA²MA)

Nalepa J., Kawulok M.: Adaptive memetic algorithm enhanced with data geometry analysis to select training data for SVMs, *Neurocomputing*, Vol. 185, pp 113–132, DOI: 10.1016/j.neucom.2015.12.046, 2016.

Analiza wstępna zbioru (wiedza „sprzed” ewolucji)



Analiza wstępna zbioru (wiedza „sprzed” ewolucji)



Bezparametryczny schemat rozrostu osobników

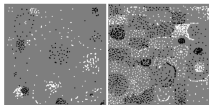
Wyniki badań eksperymentalnych

Szczegóły implementacyjne i zbiory danych

- Rozwiązania algorytmiczne zaimplementowane w języku C++ (przy użyciu biblioteki LibSVM, *sequential minimal optimization*) oraz w środowiskach .NET i R (analiza statystyczna wyników)
- Funkcja jądrowa: radialna funkcja bazowa (ang. *radial basis function*, RBF)
- Intel Xeon, 3.2 GHz, 16 GB RAM

Zbiory danych

Sztucznie wygenerowane (4)
(2D-random, 2D-chessboard)



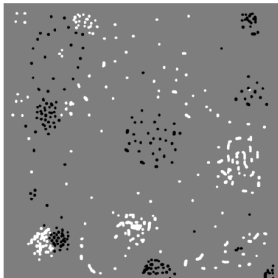
Rzeczywiste (1)
(Skin)



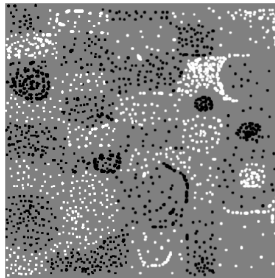
Benchmarkowe (2)
(Mushroom, Adult)

Zbiór	t	v	(t_+/t_-)
<i>Sztuczne</i>			
2D-random ^{α}	6371	6371	1.58
2D-random ^{β}	1507	1507	0.98
2D-chessboard ^{α}	22169	22169	1.21
2D-chessboard ^{β}	3622	3622	1
<i>Rzeczywiste</i>			
Skin	$4 \cdot 10^6$	97560	7.16
<i>Benchmarkowe</i>			
Adult	15082	15082	3.01
Mushroom	2882	2882	0.60

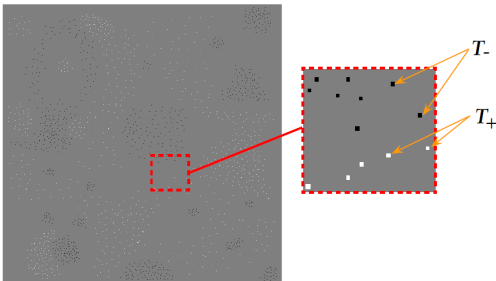
2D-random ^{α}



2D-chessboard ^{α}



2D-random ^{β}



Scenariusze testowe

- 1 Analiza czułościowa
- 2 Porównanie zaproponowanych algorytmów ewolucyjnych
- 3 Porównanie algorytmów ewolucyjnych z innymi algorytmami znanymi z literatury
- 4 Testy statystyczne (Wilcoxona i Friedmana)

Ocena jakości zredukowanych zbiorów treningowych

$$Q(\eta_V, s) = q \cdot \frac{\eta_V}{\eta_V^B} + (1 - q) \cdot \frac{s^B}{s}$$

q – waga pierwszego kryterium (jakość klasyfikacji)

η_V^B – największe AUC spośród analizowanych zbiorów zredukowanych

s^B – najmniejsza liczba wektorów podpierających spośród analizowanych zbiorów zredukowanych

Analiza czułościowa

Analiza czułościowa przeprowadzona dla każdego z zaproponowanych algorytmów ewolucyjnych

- **GASVM** – schemat selekcji, wielkość osobników w populacji
- **AGA** – schemat adaptacji osobników, dwuetapowe krzyżowanie
- **DAGA** – dynamiczny schemat adaptacji, trening zredukowany (RSVM)
- **MASVM** – pula wektorów podpierających (superosobniki, edukacja), schematy regeneracji
- **PCA²MA** – przetwarzanie wstępne T , bezparametryczna adaptacja

Analiza czułościowa

Analiza czułościowa przeprowadzona dla każdego z zaproponowanych algorytmów ewolucyjnych

- **GASVM** – schemat selekcji, wielkość osobników w populacji
- **AGA** – schemat adaptacji osobników, dwuetapowe krzyżowanie
- **DAGA** – dynamiczny schemat adaptacji, trening zredukowany (RSVM)
- **MASVM** – pula wektorów podpierających (superosobniki, edukacja), schematy regeneracji
- **PCA²MA** – przetwarzanie wstępne T , bezparametryczna adaptacja

Wariant	Analiza wstępna T	Adaptacja
PCA²MA	tak	PCA²MA
PCA ² MA(a)	nie	PCA²MA
PCA ² MA(b)	tak	MASVM
PCA ² MA(c)	nie	MASVM

Analiza czułościowa – wnioski

- Zwiększanie wielkości zbiorów zredukowanych **może pogorszyć** jakość klasyfikacji
- Zaproponowane rozwiązania algorytmiczne **istotnie polepszają** jakość zredukowanych zbiorów treningowych
- Wykorzystanie wiedzy dotyczącej zbioru T (wyekstrahowanej *przed i nabytej w czasie ewolucji*) **pozwała na polepszenie jakości zredukowanych zbiorów treningowych** (różnice są statystycznie istotne)

Porównanie zaproponowanych algorytmów ewolucyjnych

Algorytm	Opis
GASVM(4)	Algorytm genetyczny, 4 wektory w zbiorze zredukowanym
GASVM(\bar{t}')	Algorytm genetyczny, \bar{t}' wektorów w zbiorze zredukowanym
AGA	Adaptacyjny algorytm genetyczny
DAGA	Dynamicznie adaptacyjny algorytm genetyczny
MASVM	Algorytm memetyczny
PCA ² MA	Adaptacyjny algorytm memetyczny

Maksymalny czas optymalizacji:

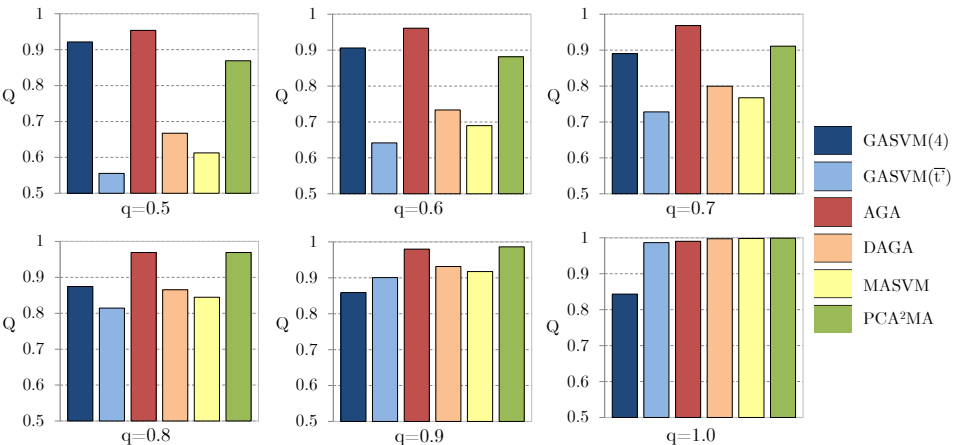
30 sekund (zbiory sztuczne)

300 sekund (zbiory rzeczywiste i benchmarkowe)

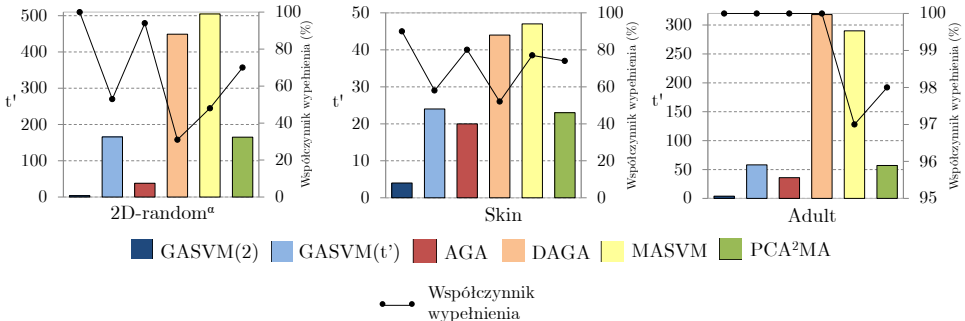
Każdy z wariantów uruchomiony (co najmniej) 30 razy

Porównanie algorytmów ewolucyjnych – wartości Q

Wartości średnie dla trzech (2D-random $^{\alpha}$, Skin, Adult) zbiorów danych



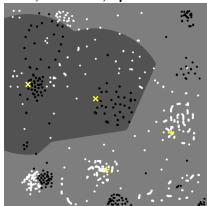
Porównanie algorytmów ewolucyjnych – t'



Porównanie algorytmów ewolucyjnych – 2D-random $^{\alpha}$

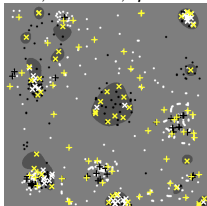
GASVM(4)

$t' = 4, s = 4, \eta_V = 0.6049$



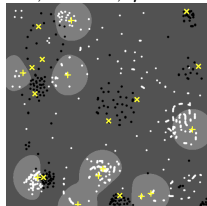
GASVM(166)

$t' = 166, s = 89, \eta_V = 0.9926$



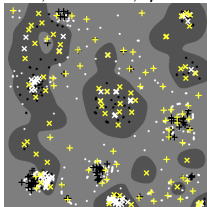
AGA

$t' = 20, s = 20, \eta_V = 0.9512$



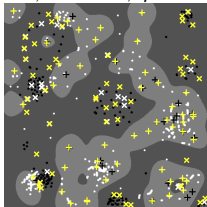
DAGA

$t' = 262, s = 108, \eta_V = 0.996$



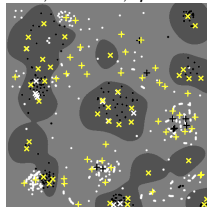
MASVM

$t' = 200, s = 120, \eta_V = 0.9971$



PCA²MA

$t' = 133, s = 96, \eta_V = 0.9983$



Porównanie algorytmów ewolucyjnych – wnioski

- Zastosowanie technik adaptacji pozwala na **uzyskanie wyraźnie lepszych zbiorów zredukowanych**
- Zastosowanie technik adaptacji umożliwia **polepszenie i znaczne ułatwienie** procesu doboru zbioru treningowego
- Zastosowanie adaptacyjnego algorytmu memetycznego (PCA²MA) pozwala na **uzyskanie najlepszych zredukowanych zbiorów treningowych**
- Różnice są **statystycznie istotne** (dla 55 z 60 par porównywanych wariantów)

Porównanie z innymi algorytmami z literatury

Algorytm	Opis
SVM ¹	SVM (trening przy użyciu pełnego zbioru T)
RS ²	Próbkowanie losowe
RSVM ³	Trening zredukowany (<i>reduced SVM</i>)
SR-DSA ⁴	Analiza geometrii danych w zbiorze T

¹ C. Cortes and V. Vapnik. Support-Vector Networks. *Machine Learning*, 20(3):273–297, 1995.

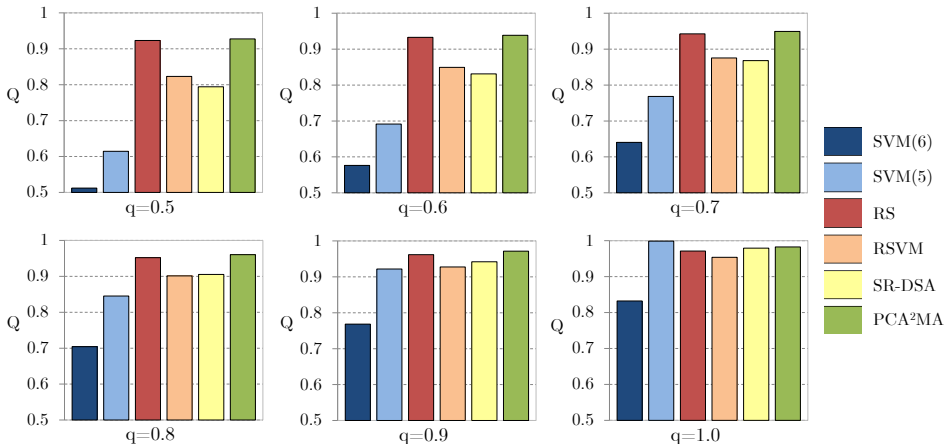
² J. L. Balcázar et al. A random sampling technique for training support vector machines. In *Proc. ALT*, 119–134. Springer, 2001.

³ Y.-J. Lee and S.-Y. Huang. Reduced support vector machines: A statistical theory. *Neural Networks, IEEE Transactions on*, 18(1):1–13, 2007.

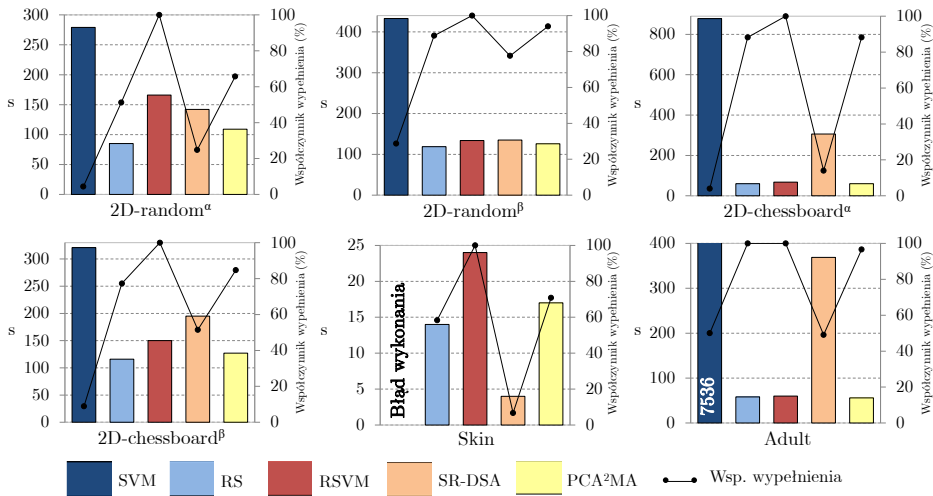
⁴ D. Wang and L. Shi. Selecting valuable training samples for SVMs via data structure analysis. *Neurocomputing*, 71:2772–2781, 2008.

Porównanie z innymi algorytmami z literatury – wartości Q

Wartości średnie dla sześciu ($2D\text{-random}^\alpha$, $2D\text{-random}^\beta$, $2D\text{-chessboard}^\alpha$, $2D\text{-chessboard}^\beta$, Skin, Adult) zbiorów danych



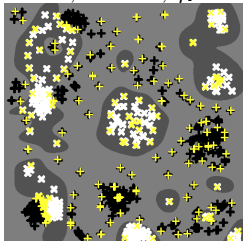
Porównanie z innymi algorytmami z literatury – s



Porównanie z innymi algorytmami – 2D-random $^{\alpha}$

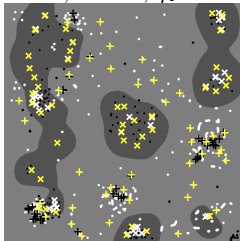
SVM

$t' = 6371, s = 279, \eta_V = 0.99$



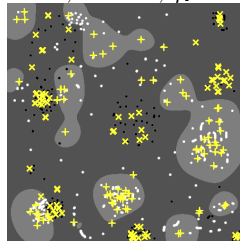
RS

$t' = 166, s = 82, \eta_V = 0.971$



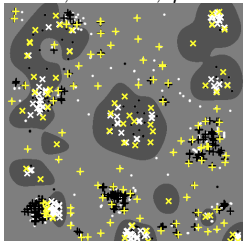
RSVM

$t' = 166, s = 166, \eta_V = 0.97$



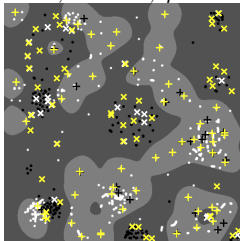
SR-DSA

$t' = 573, s = 142, \eta_V = 0.99$



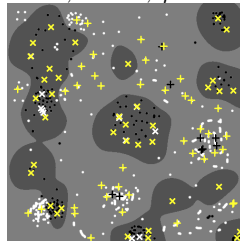
MASVM

$t' = 200, s = 120, \eta_V = 0.997$



PCA²MA

$t' = 133, s = 96, \eta_V = 0.998$



Porównanie z innymi algorytmami z literatury – wnioski

- Zastosowanie adaptacyjnego algorytmu memetycznego (PCA²MA) **pozwala na uzyskanie zredukowanych zbiorów treningowych o jakości lepszej** w porównaniu z innymi przeanalizowanymi metodami
- Czas wykonania zaproponowanych algorytmów ewolucyjnych **jest kontrolowany**
- Zbiory wyekstrahowane przy użyciu metod ewolucyjnych **mogą zawierać te wektory, które zostałyby usunięte ze zbioru zredukowanego przez inne metody**
- Wykorzystanie pełnego zbioru T do treningu klasyfikatora SVM **skutkuje uzyskaniem znacznie większej liczby wektorów podpierających**

Wnioski i kierunki dalszych badań

Wnioski

- Zastosowanie algorytmów ewolucyjnych **pozwała na usunięcie tych wektorów ze zbioru T , które negatywnie wpływają na jakość klasyfikacji**
- Zastosowanie algorytmów ewolucyjnych do doboru zbiorów zredukowanych **umożliwia przeprowadzenie treningu klasyfikatora SVM dla bardzo dużych zbiorów danych**
- Zastosowanie algorytmów ewolucyjnych – zwłaszcza memetycznych – **pozwała na zmniejszenie liczby wektorów podpierających**

Automatyczny dobór parametrów sieci głębokich

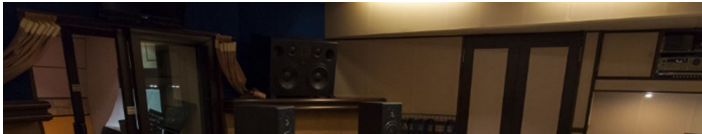
Ribalta P., Nalepa J. et al.: Particle Swarm Optimization for Hyper-Parameter Selection in Deep Neural Networks, Proceedings of the 2017 Annual Conference on Genetic and Evolutionary Computation, GECCO 2017, pp 481-488, DOI: 10.1145/3071178.3071208, ACM, 2017.

Ribalta P., Nalepa J. et al.: Hyper-parameter Selection in Deep Neural Networks Using Parallel Particle Swarm Optimization, Proceedings of the 2017 Annual Conference on Genetic and Evolutionary Computation, GECCO 2017, pp 1864-1871, DOI: 10.1145/3067695.3084211, ACM, 2017.

Dobór wartości hiperparametrów sieci głębokich

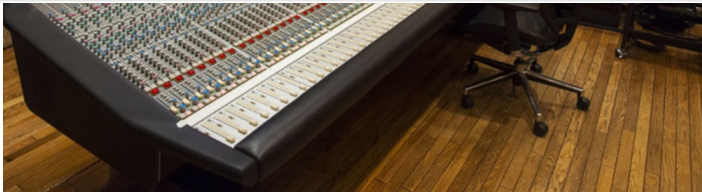


Dobór wartości hiperparametrów sieci głębokich



Główne problemy

- Liczba hiperparametrów rośnie wraz ze *złożonością* sieci
- Wartości parametrów w praktyce dobierane są *ręcznie*

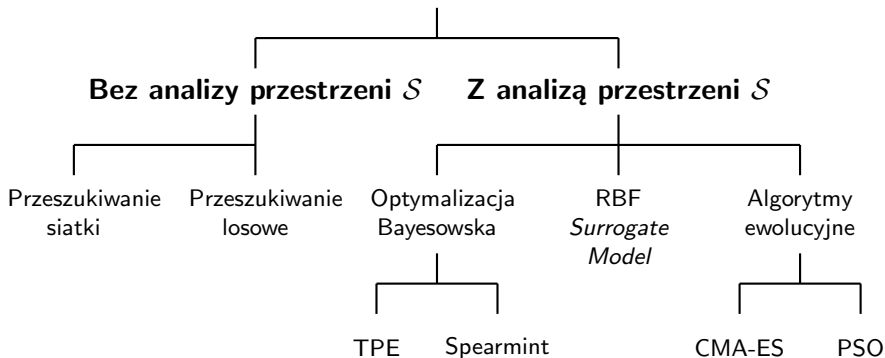


Dobór parametrów jako problem optymalizacyjny

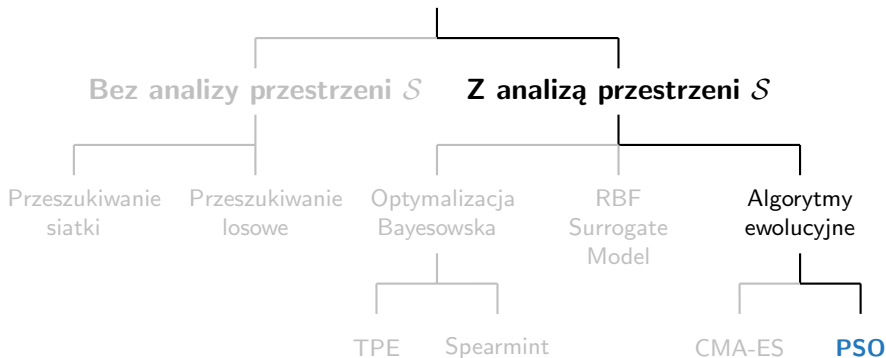
$$\lambda^* = \underset{\lambda}{\operatorname{arg\,min}} \mathcal{L}(\mathbf{T}; \mathcal{M}) = \underset{\lambda}{\operatorname{arg\,min}} f(\lambda; \mathcal{A}, \mathbf{T}, \mathbf{V}, \mathcal{L})$$

- f – funkcja celu
- λ – zbiór wartości hiperparametrów
- $\mathcal{L}(\mathbf{T}; \mathcal{M})$ – funkcja kosztu dla modelu \mathcal{M} i zbioru treningowego \mathbf{T}
- \mathcal{M} – model sieci wyznaczony przy użyciu algorytmu \mathcal{A} na zbiorze \mathbf{T} i zwalidowany na zbiorze \mathbf{V}

Dobór hiperparametrów sieci głębokich

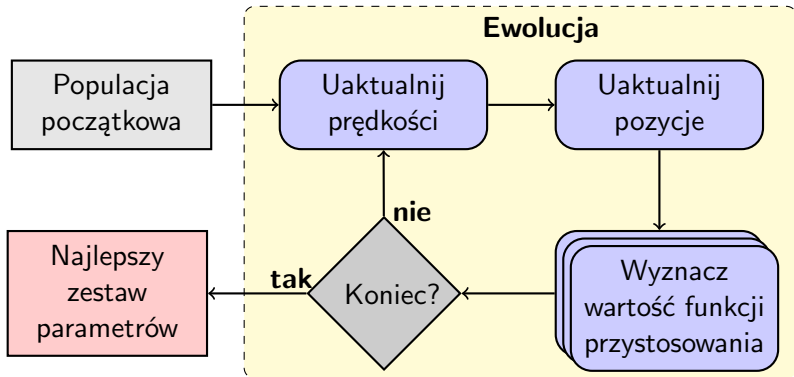


Dobór hiperparametrów sieci głębokich

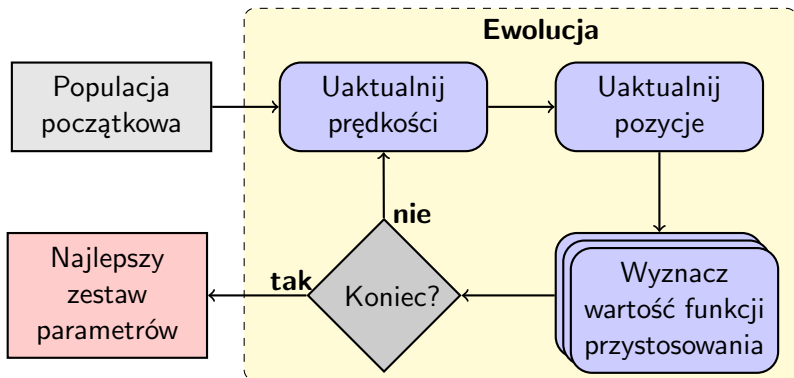


Optymalizacja sieci głębokich rojem cząstek (ang. *particle swarm optimization*)

Optymalizacja sieci głębokich rojem cząstek

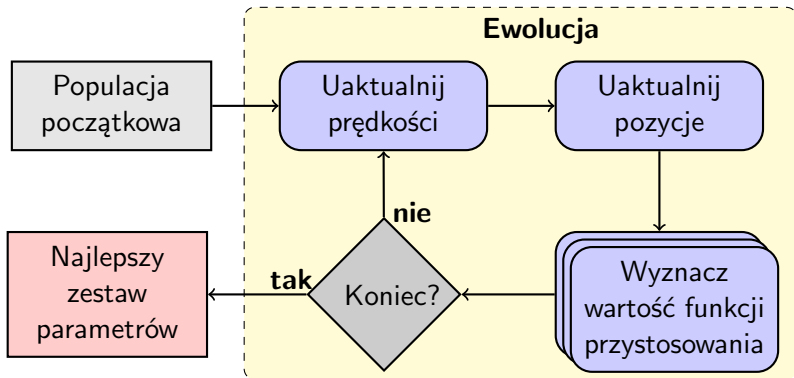


Optymalizacja sieci głębokich rojem cząstek



Populacja początkowa: $\lambda \in \mathbb{R}^k$, $\mathcal{U}(b_l, b_u)$

Optymalizacja sieci głębokich rojem cząstek



Populacja początkowa: $\lambda \in \mathbb{R}^k$, $\mathcal{U}(b_l, b_u)$

Aktualizacja prędkości i pozycji cząstek:

$$v_i \leftarrow \omega v_i + \phi_p r_p (\lambda_i^* - \lambda_i) + \phi_g r_g (\lambda^S - \lambda_i)$$

$$\lambda_i \leftarrow \lambda_i + v_i$$

Wyniki badań eksperymentalnych

Szczegóły implementacyjne

Konfiguracja sprzętowa:

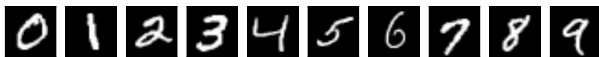
- Intel Xeon E5-2698 v3 (40M Cache, 2.30 GHz) with 128GB of RAM and **NVIDIA Tesla K80** GPU 24GB DDR5
- Intel i7-6850K (15M Cache, 3.80 GHz) with 32 GB RAM and **NVIDIA Titan X** Ultimate Pascal GPU 12GB GDDR5X

Szczegóły implementacyjne:

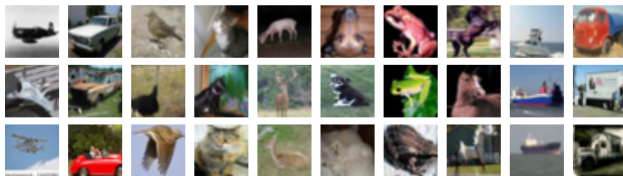
- Optymalizacja rojem cząstek zaimplementowana w języku Python
- Trening sieci głębokich: Keras (Tensorflow)
- Wykorzystanie **archiwum odwiedzonych pozycji** w przestrzeni S

Zbiory danych

- MNIST: 70,000 obrazów w skali szarości ($28 \times 28 \times 1$), 10 klas ($\sim 7,000$ przykładów w klasie)



- CIFAR-10: 60,000 obrazów barwnych ($32 \times 32 \times 3$), 10 klas (6,000 przykładów w klasie)

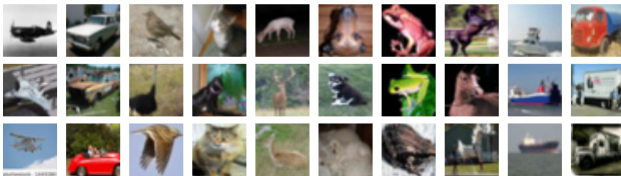


Zbiory danych

- MNIST: 70,000 obrazów w skali szarości ($28 \times 28 \times 1$), 10 klas ($\sim 7,000$ przykładów w klasie)



- CIFAR-10: 60,000 obrazów barwnych ($32 \times 32 \times 3$), 10 klas (6,000 przykładów w klasie)

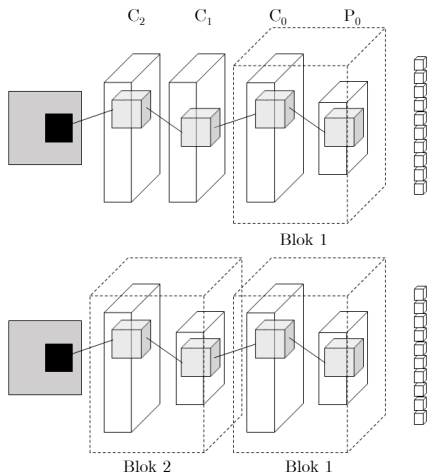


10-krotna walidacja krzyżowa

Scenariusze testowe

- 1 Wpływ wielkości populacji na jakość rozwiązań
- 2 Inkrementacyjne zwiększanie „*minimalnej*” architektury
- 3 Optymalizacja istniejącej architektury (LeNet-4)

Eksperymentalna architektura sieci głębokiej – SimpleNet



P – Pooling, C – Convolutional

Parametry eksperymentalnej sieci głębokiej

SimpleNet- N_k

- N : Liczba **bloków**
- k : Liczba **warstw konwulucyjnych**

Parametry warstw:

Rodzaj warstwy	Parametry	Wartości
Konwulucyjna (C)	Wielkość filtra ($s_F \times s_F$)	$s_F \geq 2$
	Liczba filtrów (n)	$n \geq 1$
Pooling (grupowanie cech) (P)	Krok ($stride$) (ℓ)	$\ell \geq 2$
	Wielkość filtra (s_P)	$s_P \geq 2$

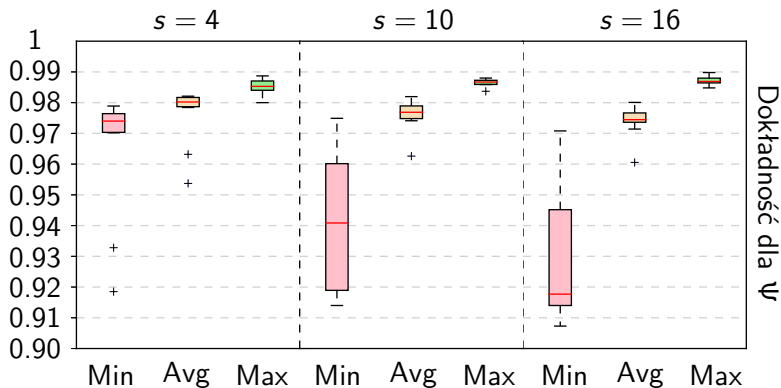
Wartości graniczne:

Warstwa	b_l	b_u
C_n	$\{n = 1, s_F = 2\}$	$\{n = 16, s_F = 8\}$
P_n	$\{s_P = 2, \ell = 2\}$	$\{s_P = 4, \ell = 4\}$

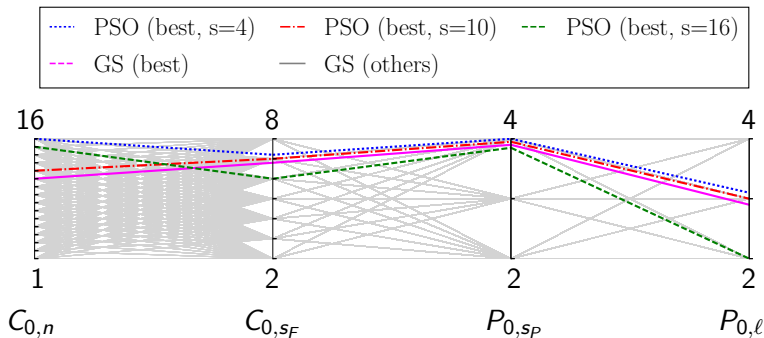
Wpływ wielkości populacji na otrzymywane wyniki

Algorytm	s	Czas (sek.)	Liczba pozycji	g_s	Dokł. dla Ψ
Przesz. siatki	—	87,356	1,008	—	0.9897
Przesz. losowe	—	39,906	400	—	0.9897
PSO	4	934	14	14	0.9852
PSO	10	2,091	29	20	0.9864
PSO	16	13,892	49	23	0.9871

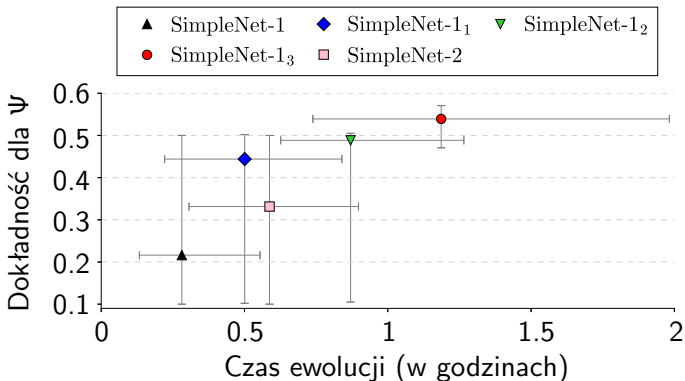
Wpływ wielkości populacji na otrzymywane wyniki



Wpływ wielkości populacji na otrzymywane wyniki



Inkrementacyjne zwiększanie architektury (CIFAR-10)



Optymalizacja istniejącej architektury – LeNet-4 (MNIST)

Klasyfikator	Błąd (%)
Pairwise linear classifier	7.6
Convolutional Clustering	1.4
SimpleNet-1, s=4	1.13
SimpleNet-1, s=10	1.12
LeNet-4	1.1
SimpleNet-1, s=16	1.08
Product of stumps on Haar features	0.87
Boosted LeNet-4	0.7
LeNet-4 (PSO w czasie ok. 2h)	0.66
K-NN with non-linear deformation	0.52
NiN	0.47
Maxout Networks	0.45
DSN	0.39
R-CNN	0.31
MultiColumn DNN	0.23

Wnioski i kierunki dalszych badań

Wnioski

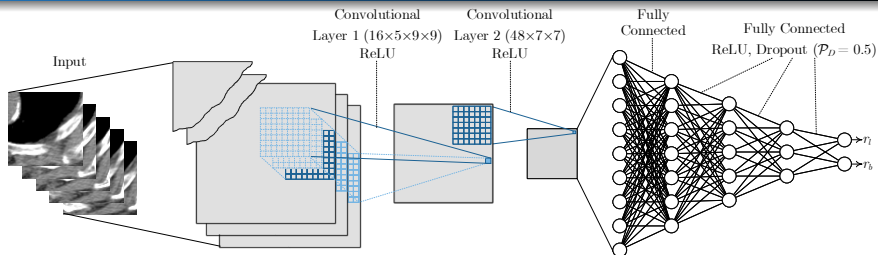
- Wartości hiperparametrów sieci głębokich (o dowolnej architekturze) mogą być ewolucyjnie optymalizowane
- Inkrementacyjne rozszerzanie minimalnych sieci głębokich pozwala na uzyskanie wyższej skuteczności klasyfikacji w przypadku *trudnych* zbiorów danych
- Algorytmy ewolucyjne mogą być łatwo zrównoleglane (przyspieszenie obliczeń oraz uzyskanie wyników lepszej jakości)

Kierunki dalszych badań

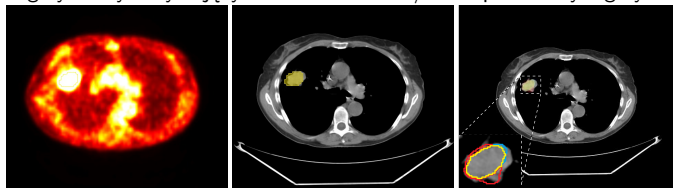
- Ewolucja architektury sieci głębokiej
- Jednoczesna ewolucja architektury i jej hiperparametrów
- Nowa funkcja przystosowania osobników w czasie ewolucji
- Segmentacja obrazów medycznych za pomocą sieci głębokich. . .

Parę słów o segmentacji obrazów medycznych. . .

Segmentacja zmian o wysokim wychwycie w TK



Algorytm wykorzystujący obrazowanie PET/CT Opracowany algorytm

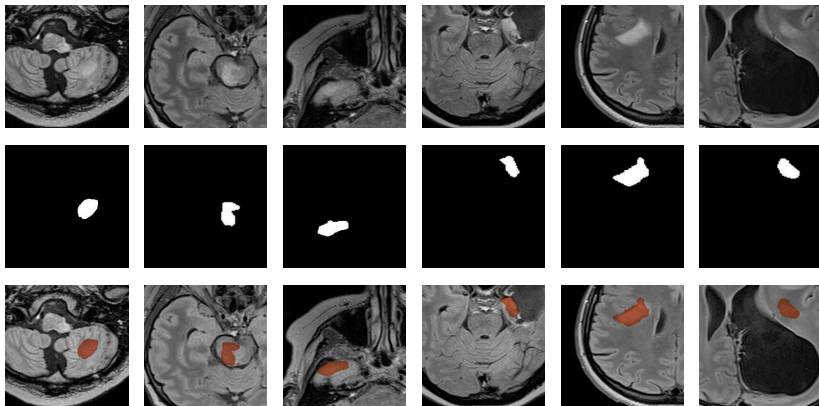


Żółty – TP, Czerwony – FP, Niebieski – FN

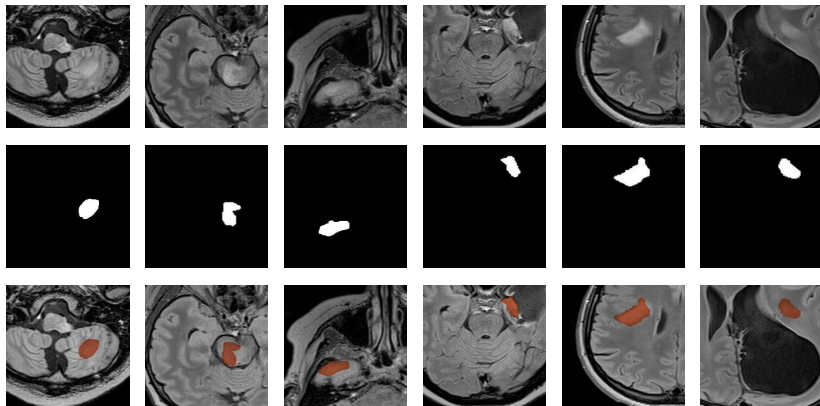
Nalepa J., Hayball M.P.: Seeing beyond the visible: Segmentation of medical images using deep neural networks, CMIH seminar, University of Cambridge (wykład zaproszony), 2017.

Pawełczyk K., Kawulok M., Nalepa J. et al.: Towards Detecting High-Uptake Lesions from Lung CT Scans Using Deep Learning, ICIAP 2017.

Segmentacja zmian w obrazowaniu głowy (FLAIR)



Segmentacja zmian w obrazowaniu głowy (FLAIR)



11 pacjentów, WHO I–IV, DICE > 0.8

Podziękowania

Dzięki!

Paulina Nalepa

Mirek Błoch

Przemysław Szmaj

Krzysztof Simiński

Balaji Ganeshan

Wojtek Dudzik

Krzysztof Pawełczyk

Łukasz Żak

Tomek Grzeszczak

Sarah McQuaid

Urszula Boryczka

Michał Kawulok

Stephen Brown

Zbigniew J. Czech

Marcin Ćwiąk

Kenneth Miles

Maksym Walczak

Luciano Sánchez Ramos

Jola Kawulok

Marcin Dublański

Vineet Prakash

Alex Menys

Mike Hayball

Stanisław Kozielski

Janusz Szymanek

Pablo Ribalta Lorenzo

Bogdan Smołka

Iza Burda

José Ranilla

Maciek Papież

Kamil Dworak

Karolina Nurzyńska

UCL Team

Dzięki!

Paulina Nalepa

Mirek Błoch

Przemysław Szmaj

Krzysztof Simiński

Balaji Ganeshan

Wojtek Dudzik

Krzysztof Pawełczyk

Łukasz Żak

Tomek Grzeszczak

Sarah McQuaid

Urszula Boryczka

Michał Kawulok

Stephen Brown

Zbigniew J. Czech

Marcin Ćwiąk

Kenneth Miles

Maksym Walczak

Luciano Sánchez Ramos

Jola Kawulok

Marcin Dublański

Vineet Prakash

Alex Menys

Mike Hayball

Stanisław Kozielski

Janusz Szymanek

Pablo Ribalta Lorenzo

Bogdan Smołka

Iza Burda

José Ranilla

Maciek Papież

Kamil Dworak

Karolina Nurzyńska

UCL Team

Dzięki!

Paulina Nalepa
Mirek Błoch
Przemysław Szmaj
Krzysztof Simiński
Balaji Ganeshan
Wojtek Dudzik
Krzysztof Pawełczyk
Łukasz Żak
Tomek Grzejszczak
Sarah McQuaid
Urszula Boryczka

Michał Kawulok
Stephen Brown
Zbigniew J. Czech
Marcin Ćwięk
Kenneth Miles
Maksym Walczak
Luciano Sánchez Ramos
Jola Kawulok
Marcin Dublański
Vineet Prakash
Alex Menys

Mike Hayball
Stanisław Kozielski
Janusz Szymanek
Pablo Ribalta Lorenzo
Bogdan Smołka
Iza Burda
José Ranilla
Maciek Papież
Kamil Dworak
Karolina Nurzyńska
UCL Team

Dzięki!



NARODOWE CENTRUM NAUKI

Fundusze Europejskie
Inteligentny RozwójUnia Europejska
Europejski Fundusz
Rozwoju RegionalnegoNarodowe Centrum
Badań i Rozwoju

Future Processing
Ministerstwo Nauki
i Szkolnictwa Wyższego



Politechnika Śląska, Instytut Informatyki



Ewolucyjne uczenie maszynowe

... czyli o moich zainteresowaniach naukowych

Jakub Nalepa

`jakub.nalepa@polsl.pl`

`jnalepa@ieee.org`

WWW: `http://sun.aei.polsl.pl/~jnalepa/`

Dziękuję za uwagę!

Uroczystość wręczenia Nagrody im. Witolda Lipskiego

Warszawa, 12 października 2017