

Predykcja kierunku wzrostu twarzy z zastosowaniem metod uczenia maszynowego

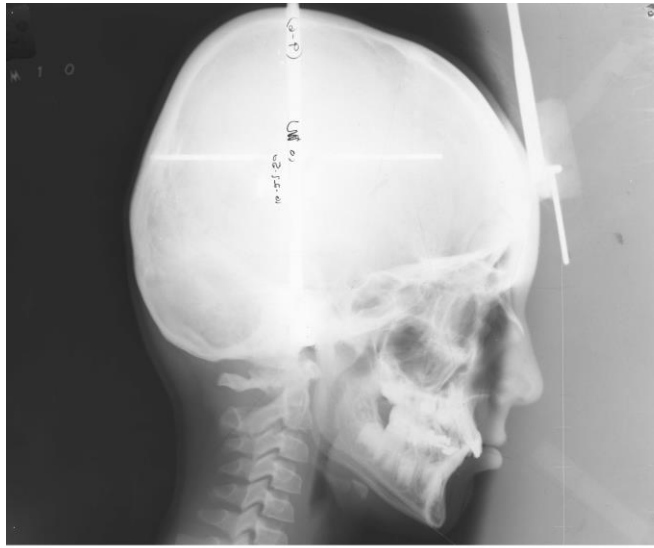
Stanisław Kaźmierczak

Agenda

1. Wprowadzenie
2. Sformułowanie zagadnienia
3. Analiza danych
4. Wyniki predykcji
5. Augmentacja tabelaryczna
6. Inne pomysły

Kierunek wzrostu

Do przodu



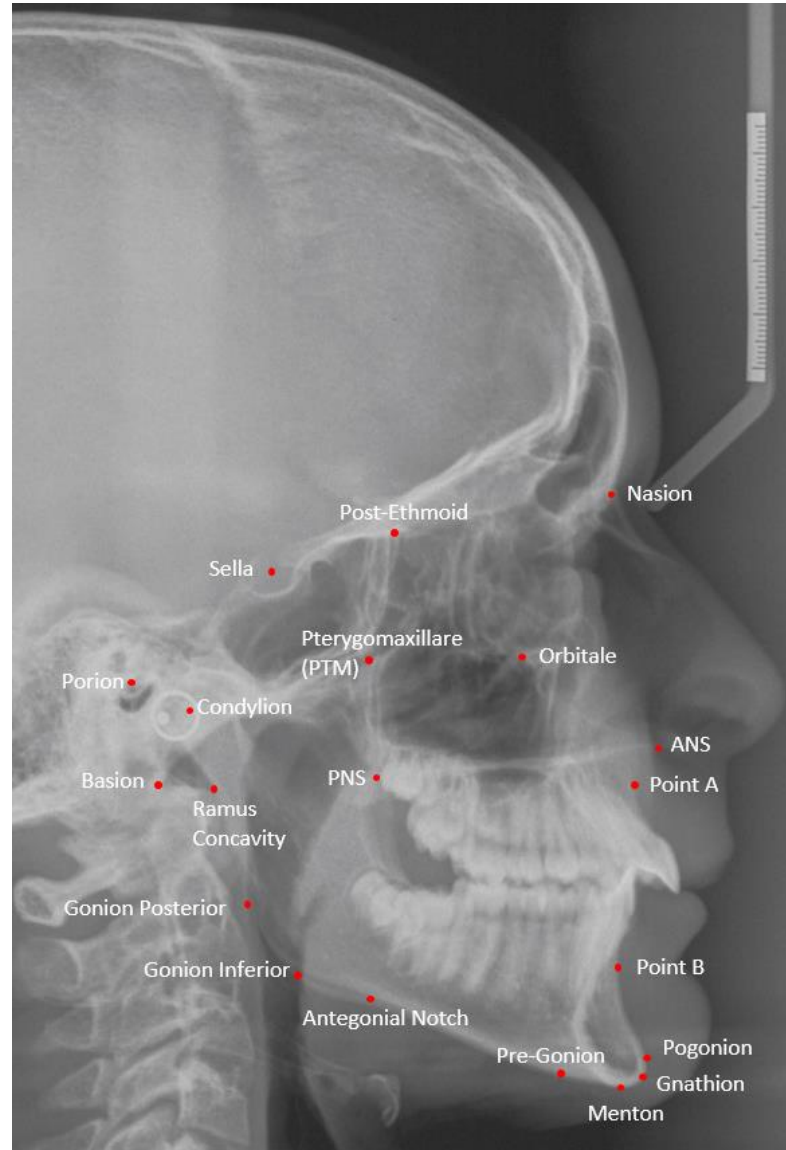
Ku dołowi



Wprowadzenie

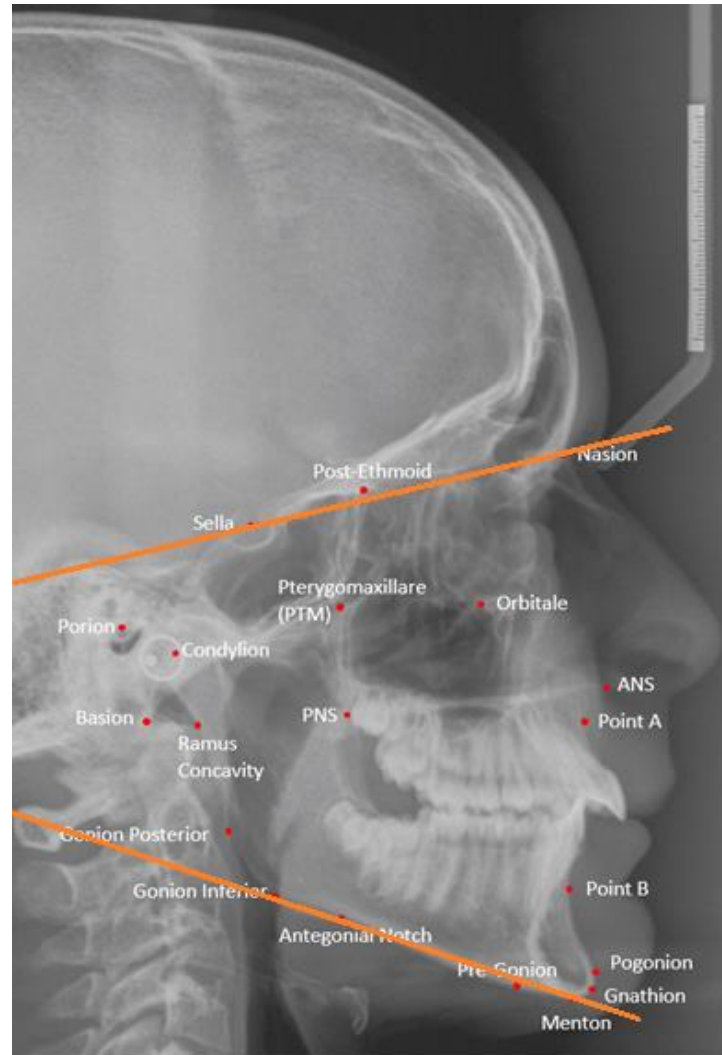
- Pierwsze próby kierunku wzrostu twarzy sięgają lat 70-tych ubiegłego wieku
- Odpowiednio wcześnie podjęte działania mogą w dużym stopniu skorygować wadę
- Dla ekspertów dziedzinowych (ortodontów) zagadnienie to stanowi do tej pory wyzwanie
- Zgodnie z naszą wiedzą, temat ten nie był wcześniej podejmowany w kontekście zastosowania metod uczenia maszynowego
- Wykonywanie zdjęć RTG w celach czysto naukowych jest obecnie uważane za nieetyczne
- Jak zdefiniować kierunek wzrostu twarzy?
- Przewidujemy zmianę, a nie docelową wartość

Punkty cefalometryczne



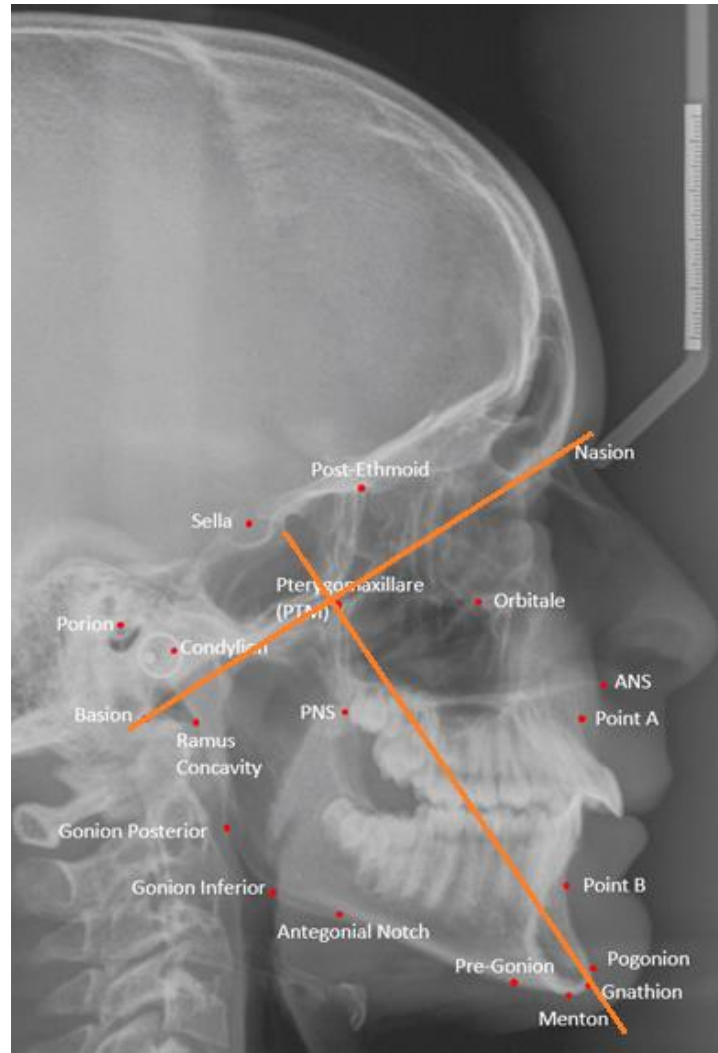
Miary kształtu twarzy: SN/MP

Kąt wyznaczony przez proste Sella-Nasion oraz Menton-Gonion Inferior



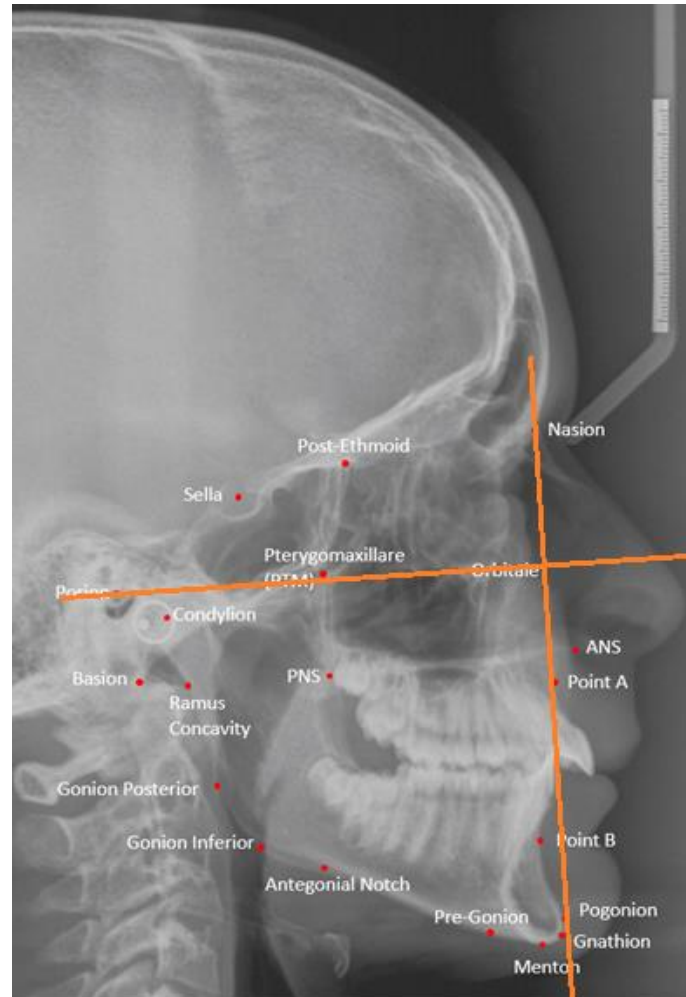
Miary kształtu twarzy: Facial axis (FA)

Kąt wyznaczony przez proste Basion-Nasion oraz Gnathion-PTM



Miary kształtu twarzy: Facial axis (FA)

Różnica odległości Point A oraz Pogonion od prostej przechodzącej przed Nasion i prostopadłej do Porion-Orbitale



Sformułowanie problemu

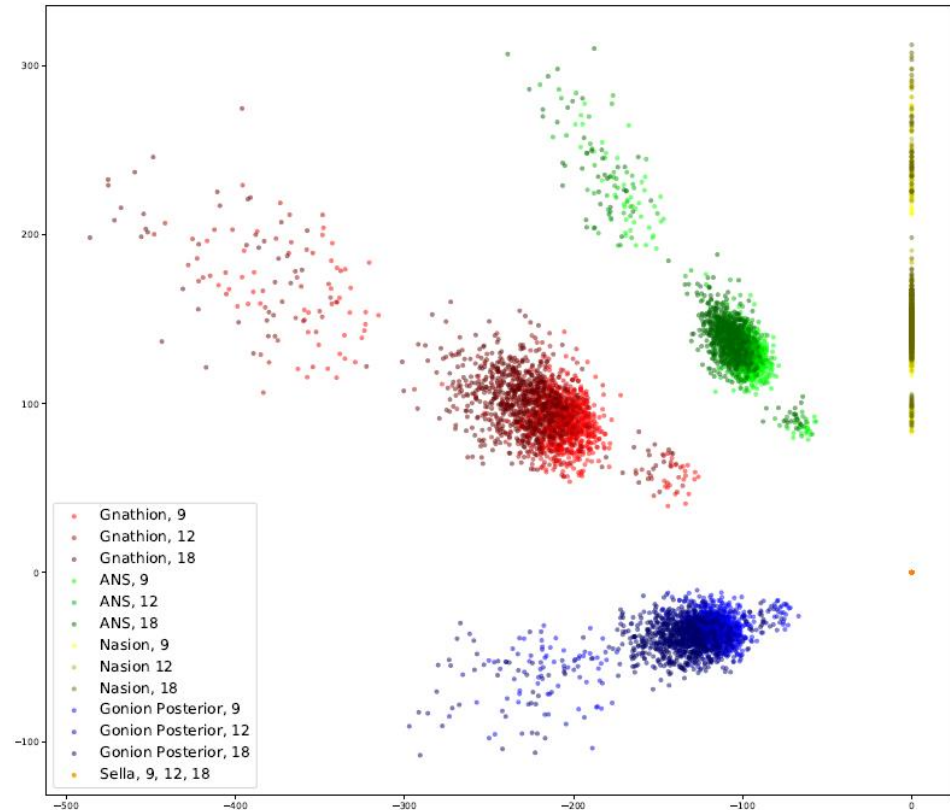
- Mając na wejściu dane 9-cio i/lub 12-latka chcemy przewidywać kierunek wzrostu twarzy, czyli zmianę którejś z trzech metryk
- Dane tabelaryczne – punktowe lub kątowe
- Dla każdego pacjenta obliczono zmianę wspomnianych metryk pomiędzy 9 a 18 rż.
- Na tej podstawie utworzono etykiety w następujący sposób:
 - Klasa 1: wartości mniejsze niż średnia – odchylenie standardowe (wzrost poziomy)
 - Klasa 2: wartości znajdujące się nie dalej niż jedno odchylenie standardowe od średniej
 - Klasa 3: wartości większe niż średnia + odchylenie standardowe (wzrost pionowy)

- Problemy
 - Różne zbiory danych
 - Zdjęcia wykonane na różnych urządzeniach
 - Oznaczanie punktów cefalometrycznych jest obarczone błędem
 - Zdjęcia są w różnej skali, która jest niemożliwa do określenia
 - Jakość zdjęć
- Charakterystyka zbioru
 - 639 instancji

Group	Age	Minimum	Maximum
<i>9-year-olds</i>	9.06 ± 0.45	6.00	10.92
<i>12-year-olds</i>	12.07 ± 0.39	10.00	13.75
<i>18-year-olds</i>	17.41 ± 1.71	15.00	28.42

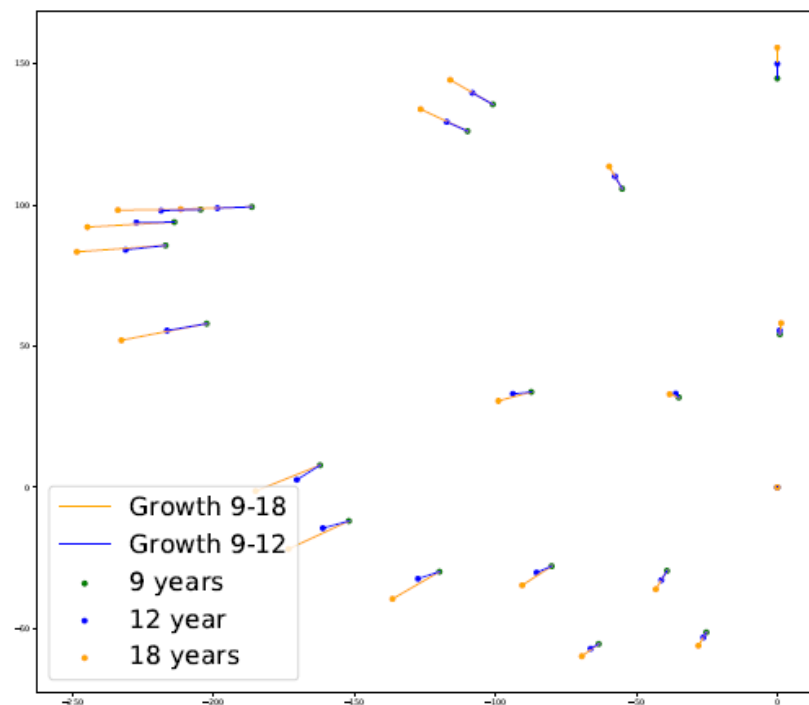
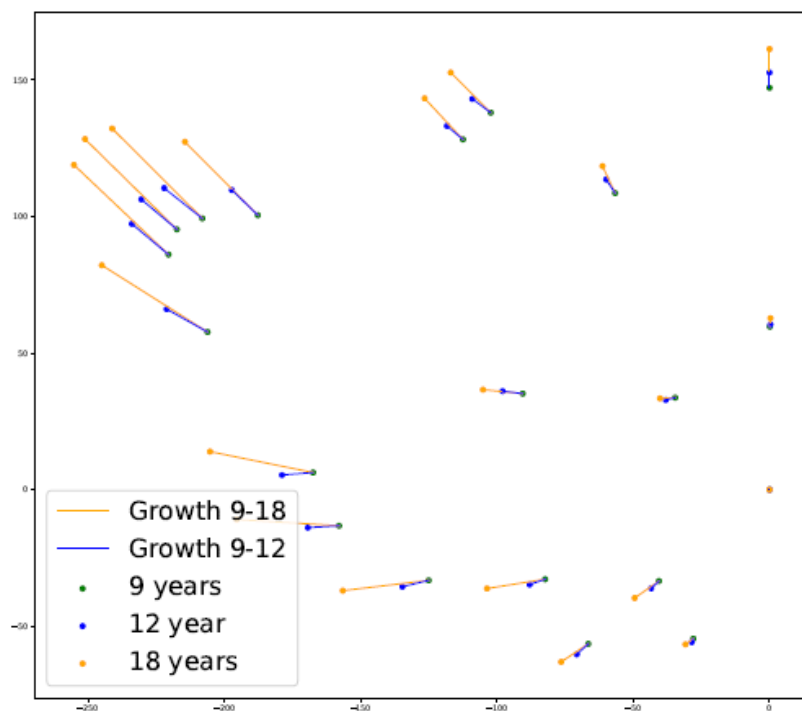
Skala zdjęć

- Co najmniej 3 różne skale zdjęć
- Różnica w skali pomiędzy różnymi pacjentami nie stanowi dużego problemu
- Różnica w skali na poziomie pojedynczych pacjentów jest istotnym problemem
- Nie znamy czynnika skalującego



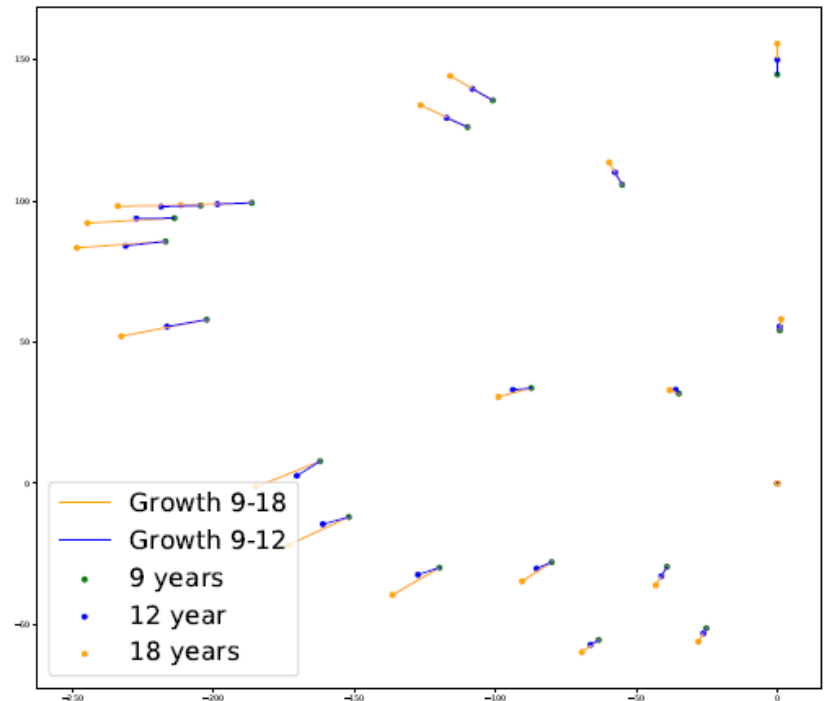
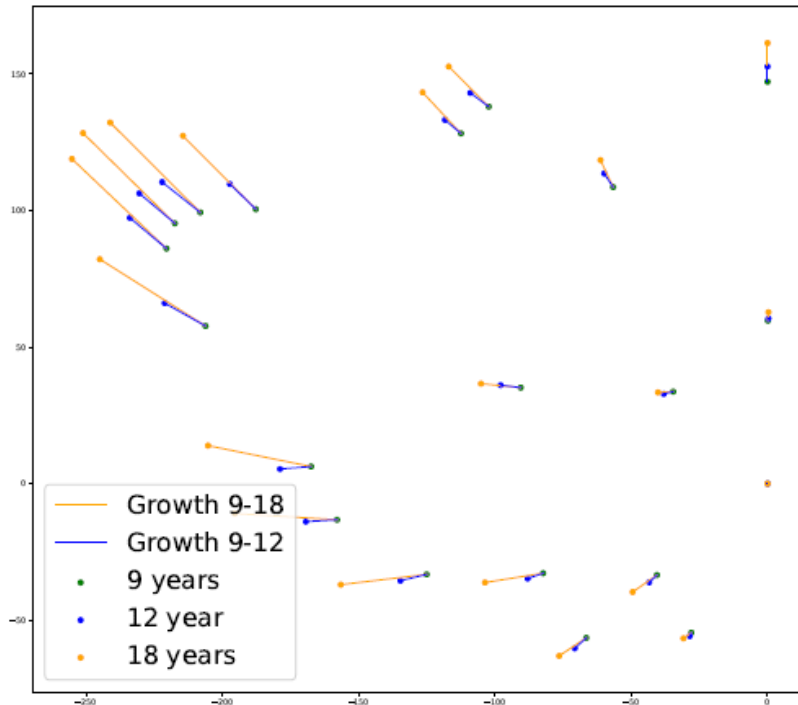
Porównanie grup pacjentów (1)

- Wybrani zostali pacjenci, którzy wg. trzech miar wzrostu oznaczeni są „do przodu” lub „ku dołowi”
- Okazuje się, że tylko 31 z 639 pacjentów spełnia zadane kryteria



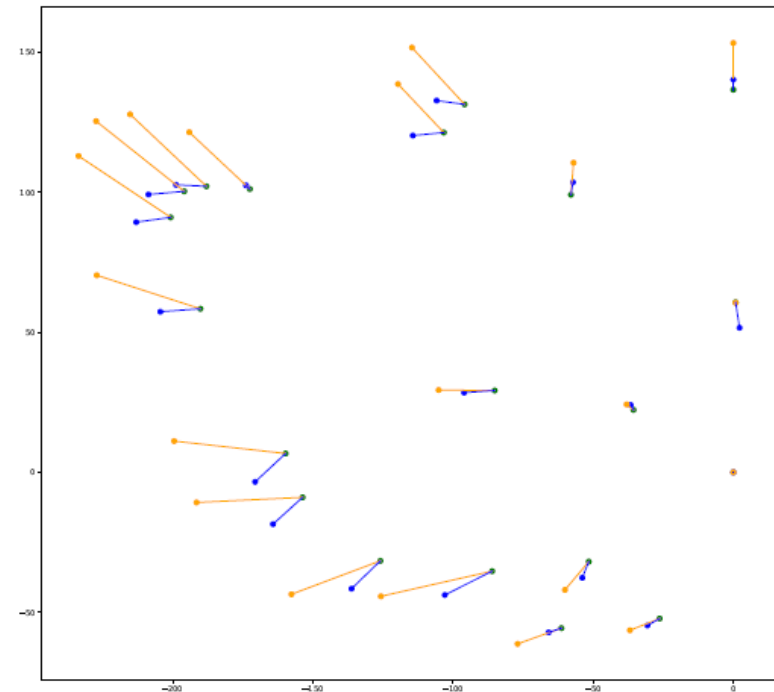
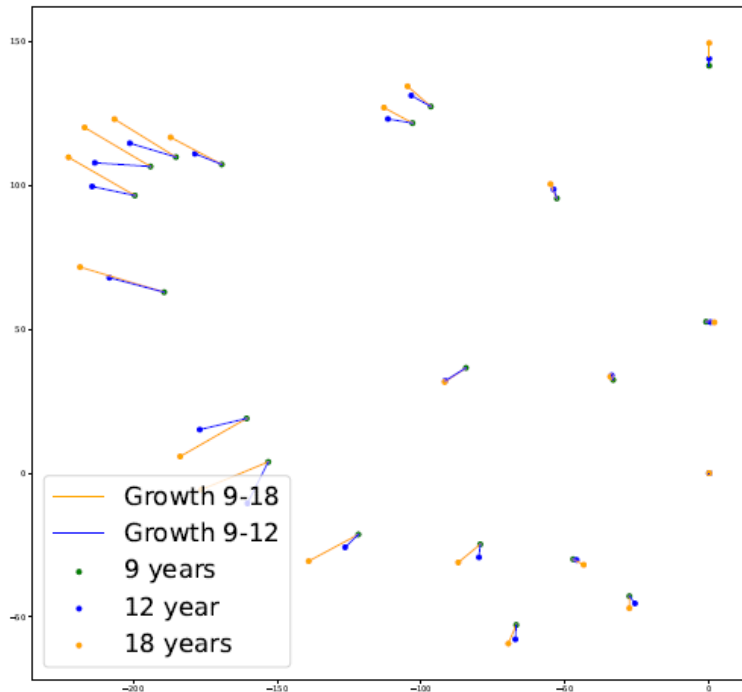
Porównanie grup pacjentów (2)

- Kierunek wzrostu punktów należących do podbródka istotnie różni się dla obu grup
- Kierunek zmiany położenia w okresie 9-12 jest bardzo zbliżony do kierunku zmiany w okresie 9-18



Porównanie grup pacjentów (3)

- Na poziomie pojedynczych pacjentów rozbieżność między kierunkiem 9-12 oraz 9-18 jest niestety duża



Okresy wzrostu

- Korelacja liniowa pomiędzy trzyletnimi okresami jest znikoma
- Dla prognozowanych metryk korelacja pomiędzy 9-12 a 9-18 wynosi 0.50, 0.50 oraz 0.52

SN-MP

	9-12	12-15	15-18	9-15	12-18	9-18
9-12	1.00					
12-15	-0.17	1.00				
15-18	-0.10	0.01	1.00			
9-15	0.62	0.67	-0.07	1.00		
12-18	-0.20	0.81	0.59	0.50	1.00	
9-18	0.50	0.60	0.46	0.86	0.75	1.00

FA

	9-12	12-15	15-18	9-15	12-18	9-18
9-12	1.00					
12-15	-0.24	1.00				
15-18	-0.19	-0.28	1.00			
9-15	0.63	0.60	-0.38	1.00		
12-18	-0.36	0.64	0.55	0.22	1.00	
9-18	0.50	0.40	0.35	0.73	0.63	1.00

PN-AN

	9-12	12-15	15-18	9-15	12-18	9-18
9-12	1.00					
12-15	-0.18	1.00				
15-18	0.05	-0.37	1.00			
9-15	0.52	0.75	-0.28	1.00		
12-18	-0.12	0.62	0.50	0.46	1.00	
9-18	0.52	0.43	0.46	0.72	0.79	1.00

Rozmiar danych i niezbalansowanie

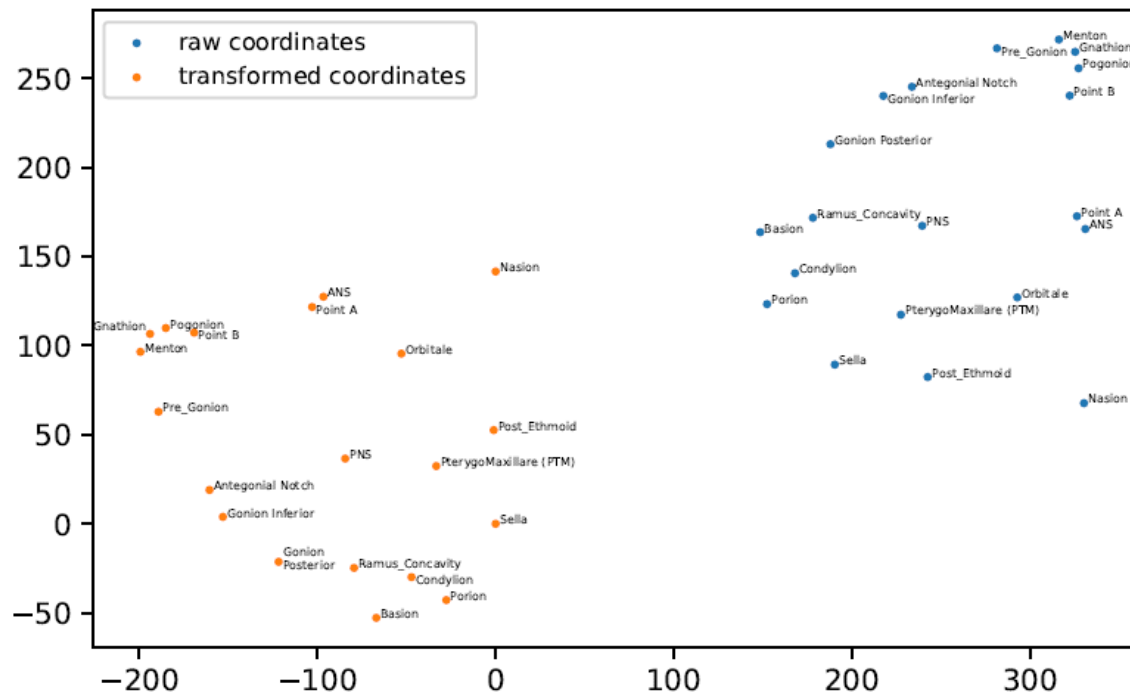
- W zależności od problemu liczba cech wynosi od 16 do 82
- Liczba instancji, które należy podzielić pomiędzy zbiór treningowy, walidacyjny i testowy wynosi 639
- Ze względu na małą liczbę instancji oraz wysoką wartość współczynnika liczba cech/liczba instancji głębokie architektury nie znajdą raczej zastosowania
- Dla metryk SN-MP, FA oraz PN-AN klasa większościowa stanowi odpowiednio 68.23%, 69.95% i 74.80% instancji

Dane wejściowe (1)

- Dane cefalometryczne
 - Kąty wyznaczone przez punkty cefalometryczne
 - 15 cech
- Dane prokrustowe – punkty cefalometryczne znormalizowane w taki sposób, że dla każdej twarzy:
 - Środek ciężkości punktów położony jest w $(0, 0)$
 - Suma kwadratów odległości pomiędzy $(0, 0)$ i wszystkimi przekształconymi punktami wynosi 1
 - Suma kwadratów odległości pomiędzy danym punktem cefalometrycznym oraz jego średnim położeniem jest minimalizowana po wszystkich punktach i pacjentach
 - 40 cech

Dane wejściowe (2)

- Dane przekształcone
 - Punkt Sella znajduje się w początku układu współrzędnych
 - Sella-Nasion wyznacza pion
 - 40 cech
- Dodatkowo w każdym wariancie cechę stanowi wiek



Wyniki (1)

	9	12	12-9	9, 12	9, 12-9
No.	Features, model Accuracy [%]	Features, model Accuracy [%]	Features, model Accuracy [%]	Features, model Accuracy [%]	Features, model Accuracy [%]
Prediction of SN-MP(18-9)					
1	ceph, SVM 68.23 ± 0.33	ceph, SVM 68.23 ± 0.33	ceph, LR 71.25 ± 2.31	ceph, LR 69.12 ± 2.94	ceph, RF(300) 70.38 ± 2.25
2	proc, SVM 68.23 ± 0.33	proc, SVM 68.23 ± 0.33	ceph, MLP 70.07 ± 2.84	ceph, SVM 68.23 ± 0.33	ceph, RF(100) 70.19 ± 2.45
3	trans, SVM 68.23 ± 0.33	trans, SVM 68.23 ± 0.33	ceph, RF(300) 70.06 ± 2.21	proc, SVM 68.23 ± 0.33	ceph, XGB(100) 69.90 ± 2.84
4	proc, MLP 68.22 ± 0.35	proc, LR 68.23 ± 0.33	ceph, RF(100) 69.77 ± 2.33	trans, SVM 68.23 ± 0.33	ceph, XGB(300) 69.59 ± 2.87
5	proc, MLP(50, 50) 68.21 ± 0.39	proc, MLP 68.23 ± 0.41	proc, XGB(100) 69.06 ± 2.50	proc, RF(300) 68.14 ± 0.79	ceph, LR 69.06 ± 2.92
Prediction of FA(18-9)					
1	proc, MLP(50, 50) 69.98 ± 0.42	proc, MLP 69.97 ± 0.37	proc, MLP(50, 20) 71.12 ± 2.41	proc, MLP(20) 70.05 ± 1.18	proc, MLP(50, 10) 70.92 ± 2.26
2	ceph, SVM 69.95 ± 0.35	ceph, SVM 69.95 ± 0.35	proc, MLP(50, 50) 71.01 ± 2.45	ceph, SVM 69.95 ± 0.35	proc, MLP(50, 20) 70.74 ± 2.28
3	proc, SVM 69.95 ± 0.35	proc, SVM 69.95 ± 0.35	proc, MLP(20) 70.99 ± 2.09	proc, SVM 69.95 ± 0.35	proc, MLP(50) 70.69 ± 2.24
4	trans, SVM 69.95 ± 0.35	trans, SVM 69.95 ± 0.35	proc, MLP(50) 70.99 ± 2.35	trans, SVM 69.95 ± 0.35	proc, MLP(20) 70.66 ± 1.88
5	proc, LR 69.95 ± 0.35	proc, LR 69.95 ± 0.35	ceph, LR 70.95 ± 2.86	proc, LR 69.95 ± 0.35	proc, MLP(100) 70.64 ± 2.11
Prediction of PN-AN(18-9)					
1	ceph, LR 75.05 ± 0.94	ceph, LR 75.25 ± 1.14	ceph, LR 75.14 ± 2.06	trans, RF(300) 74.99 ± 1.58	ceph, SVM 74.80 ± 0.30
2	trans, RF(300) 74.95 ± 1.46	trans, RF(300) 74.95 ± 1.50	ceph, MLP 74.97 ± 2.23	trans, RF(100) 74.88 ± 1.64	proc, SVM 74.80 ± 0.30
3	trans, RF(100) 74.91 ± 1.49	ceph, SVM 74.80 ± 0.30	proc, SVM 74.80 ± 0.30	ceph, LR 74.82 ± 2.05	proc, LR 74.80 ± 0.30
4	ceph, SVM 74.80 ± 0.30	proc, SVM 74.80 ± 0.30	proc, LR 74.80 ± 0.30	ceph, SVM 74.80 ± 0.30	trans, RF(300) 74.80 ± 1.70
5	proc, SVM 74.80 ± 0.30	proc, LR 74.80 ± 0.30	trans, SVM 74.75 ± 0.37	proc, LR 74.80 ± 0.30	ceph, LR 74.78 ± 2.17

Wyniki (2)

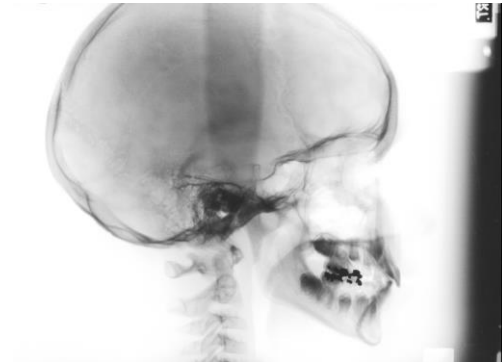
- Udział (procent), jakim jest klasa większościowa stanowi punkt odniesienia (MFC)
- Modele bazujące na danych z jednego punktu czasu tylko w dwóch przypadkach przekroczyły MFC
- Aby móc jakkolwiek przewidywać wzrost potrzebne są dane zarówno z 9-ciu jak i 12-stu lat
- Zmiana wartości pomiędzy 9 a 12 lat jest istotniejsza niż wartości w poszczególnych punktach czasowych
- Wszystkie istotnie statystycznie wyniki zostały osiągnięte na danych cefalometrycznych lub prokrustowych
- PN-AN(18-9) jest najtrudniejsze do predykcji
- Wszystkie istotne statystycznie wyniki zostały osiągnięte przez LR, którąś z architektur MLP lub zespół modeli – RF lub XGB

Wyniki (3)

- Wykonano eksperyment polegający na sztucznym zbalansowaniu klas tak, aby każda stanowiła 1/3
- Osiągnięta dokładność wahała się od 51% do 55% w zależności od modelu
- Im bardziej problem jest niezbalansowany, tym ciężiej przekroczyć MFC (model może uznać, że mając na uwadze minimalizację funkcji celu, bardziej opłaca się predykować klasę większościową)

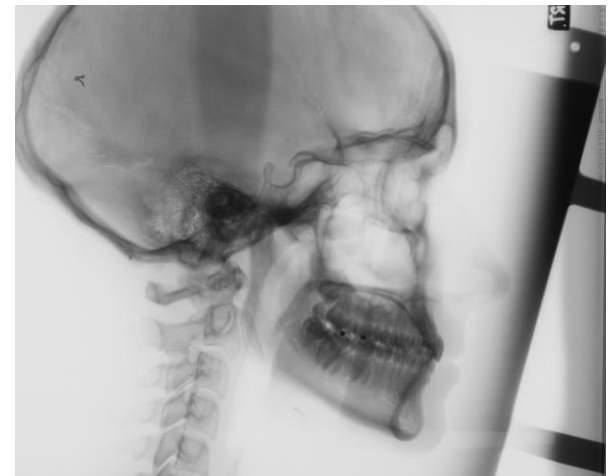
Surowe dane obrazowe (1)

- Problemy z jakością zdjęć:
 - różny poziom naświetlenia
 - białe paski
 - widoczne kawałki fartucha
 - negatywy
 - fragmenty ręcznych notatek
 - zdjęcia robione z lewego profilu
 - różne proporcje obrazów
 - Inne szумы
 - różne miejsce usytuowania twarzy na zdjęciu



Surowe dane obrazowe (2)

- W jaki sposób wycinać istotne fragmenty zdjęcia?
- Wydaje się, że wobec tak zniuansowanego/detalicznego problemu predykcyjnego, działanie na obrazach o tak słabej jakości nie przyniesie wymiernych rezultatów
- Aby stosować augmentację, model powinien mieć pewną bazową siłę predykcyjną
- Augmentacja może stać się źródłem dodatkowego zaszumienia



Oversampling

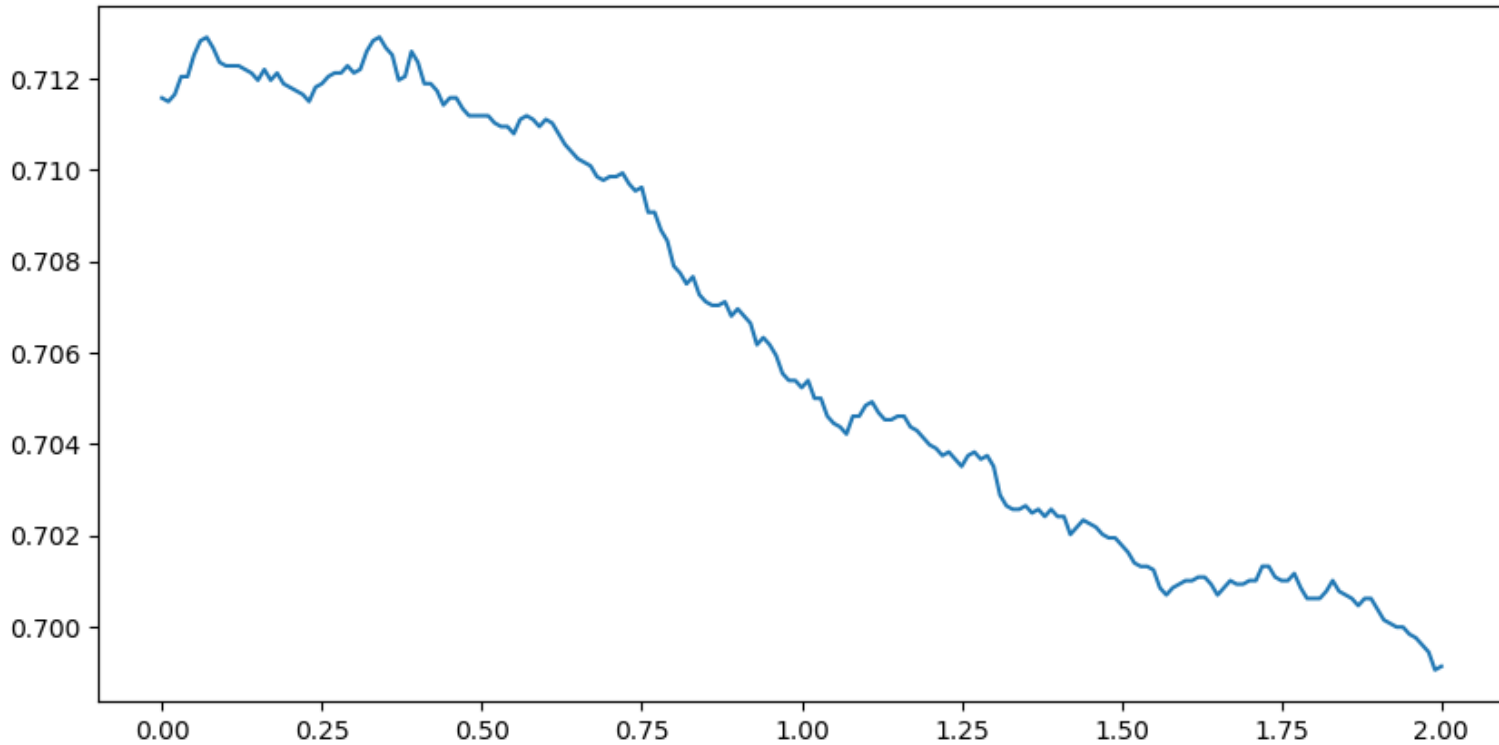
Klasa mniejszościowa

0	71.25 +/- 2.31	4	59.67 +/- 10.54
[[2.27 14. 0.32]		[[11.22 3.37 2.]	
[1.61 64.27 2.35]		[27.52 22.72 18.]	
[0. 10.47 4.71]]		[1.49 3.18 10.52]]	
1	69.86 +/- 3.46	5	55.97 +/- 12.76
[[5.57 10.36 0.66]		[[12.14 2.27 2.18]	
[6.71 55.29 6.23]		[32.63 14.16 21.44]	
[0.16 7.41 7.61]]		[1.89 2.14 11.16]]	
2	67.05 +/- 5.39	6	52.65 +/- 14.39
[[8.15 7.17 1.28]		[[12.83 1.54 2.21]	
[13.81 44.38 10.04]		[36.39 8.32 23.52]	
[0.31 5.99 8.89]]		[2.19 1.41 11.58]]	
3	63.47 +/- 7.99	7	49.81 +/- 15.44
[[9.95 4.99 1.64]		[[13.23 1.06 2.3]	
[21.17 32.9 14.17]		[38.48 4.88 24.87]	
[0.84 4.46 9.88]]		[2.48 0.88 11.82]]	

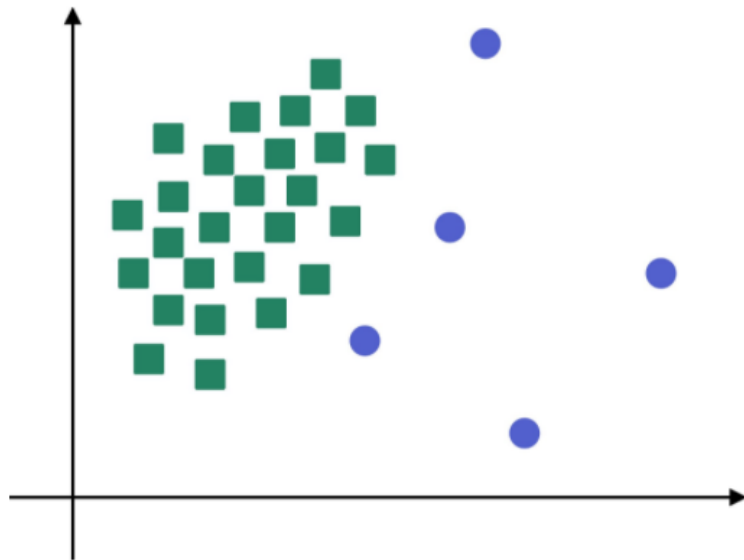
Klasa większościowa

0	71.25 +/- 2.31
[[2.27 14. 0.32]	
[1.61 64.27 2.35]	
[0. 10.47 4.71]]	
1	70.89 +/- 2.01
[[0.61 15.81 0.17]	
[0.41 66.91 0.92]	
[0. 12.18 3.]]	
2	70.63 +/- 1.88
[[0.32 16.11 0.16]	
[0.1 67.52 0.61]	
[0. 12.9 2.28]]	

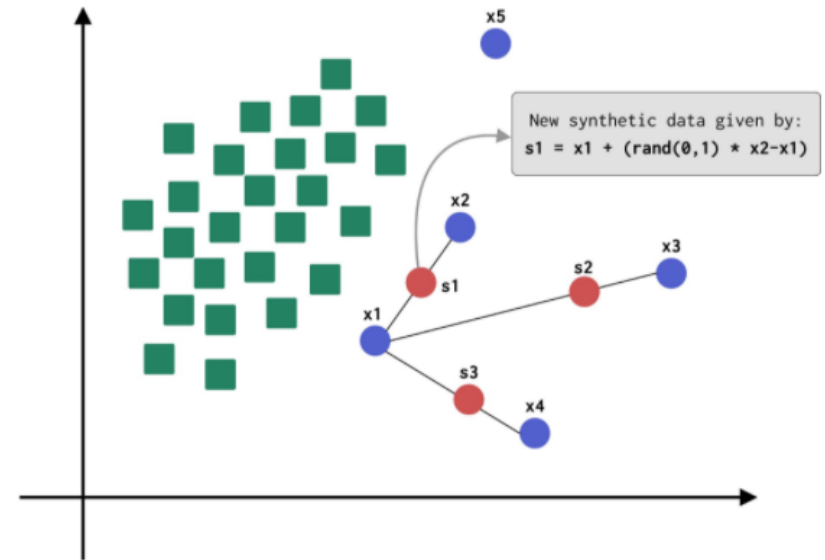
Augmentacja szumem gaussowskim



Augmentacja SMOTE



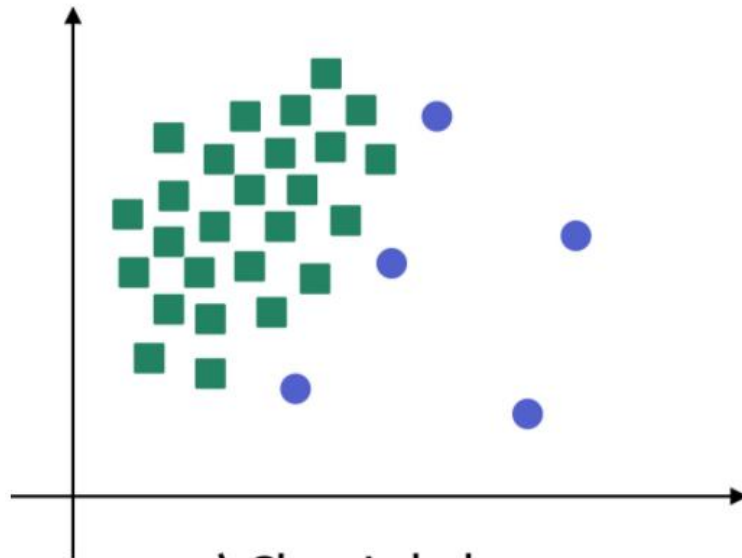
a) Class Imbalance



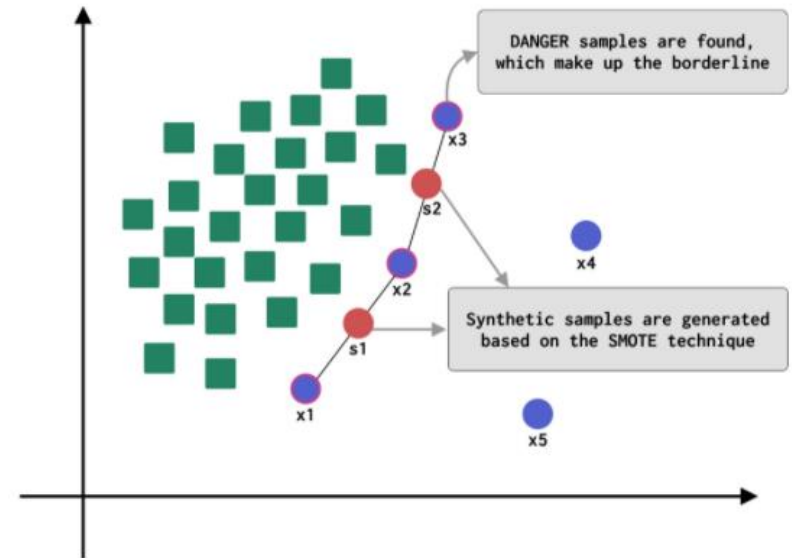
b) SMOTE

[1, 2]

Augmentacja Borderline-SMOTE



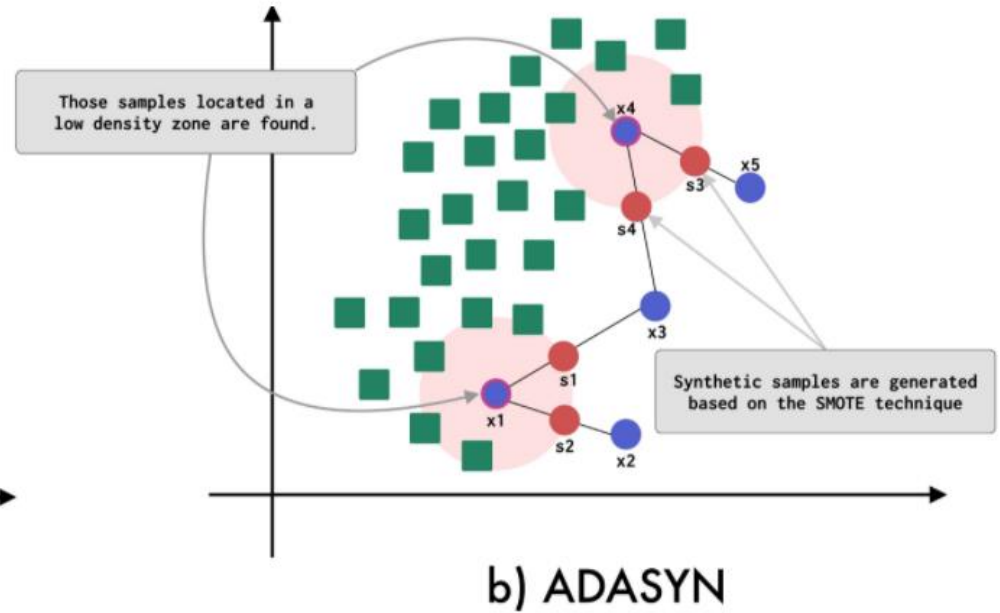
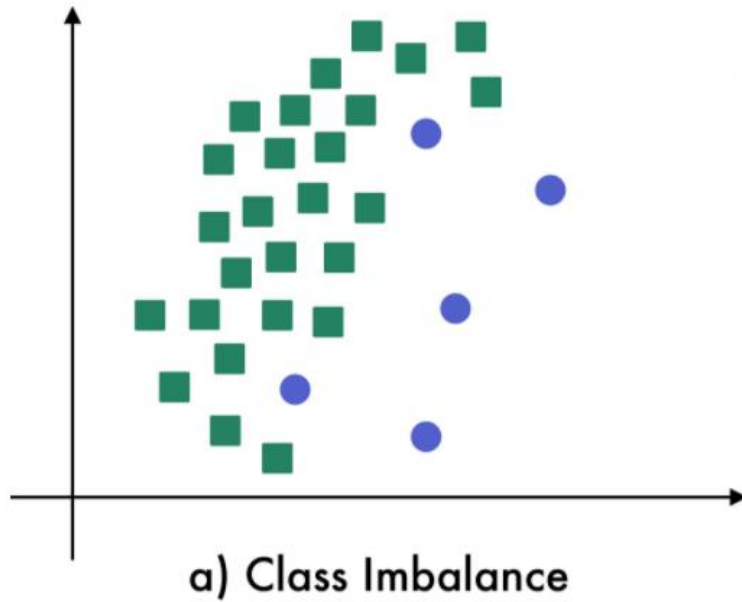
a) Class Imbalance



b) Borderline-SMOTE

[1, 3]

Augmentacja ADASYN



[1, 4]

Maksymalizacja jakości predykcji SN/MP

- MFC: 68.23%
- Najsilniejszy model bazujący na [9], [12], [12-9] [9, 12], [9, 12-9]: 71.25%
- Najsilniejszy model jednocechowy (SN/MP_dif): 72.48%
- Najsilniejszy model dwucechowy (SN/MP_dif, Y Antegonial Notch_dif): 73.75%

Czy informacje niezbędne do wysokiej jakości predykcji znajdują się w danych?

1. <https://towardsdatascience.com/smote-synthetic-data-augmentation-for-tabular-data-1ce28090debc>
2. Chawla, N. V., Bowyer, K. W., Hall, L. O., & Kegelmeyer, W. P. (2002). SMOTE: synthetic minority over-sampling technique. *Journal of Artificial Intelligence Research*, 16, 321-357.
3. Han, H., Wang, W. Y., & Mao, B. H. (2005, August). Borderline-SMOTE: a new over-sampling method in imbalanced data sets learning. In *International Conference on Intelligent Computing* (pp. 878-887). Springer, Berlin, Heidelberg.
4. He, H., Bai, Y., Garcia, E. A., & Li, S. (2008, June). ADASYN: Adaptive synthetic sampling approach for imbalanced learning. In *2008 IEEE IJCNN(IEEE WCCI)* (pp. 1322-1328). IEEE.