MODEL TUNABILITY

Paweł Pozorski
Katarzyna Rogalska

AGENDA

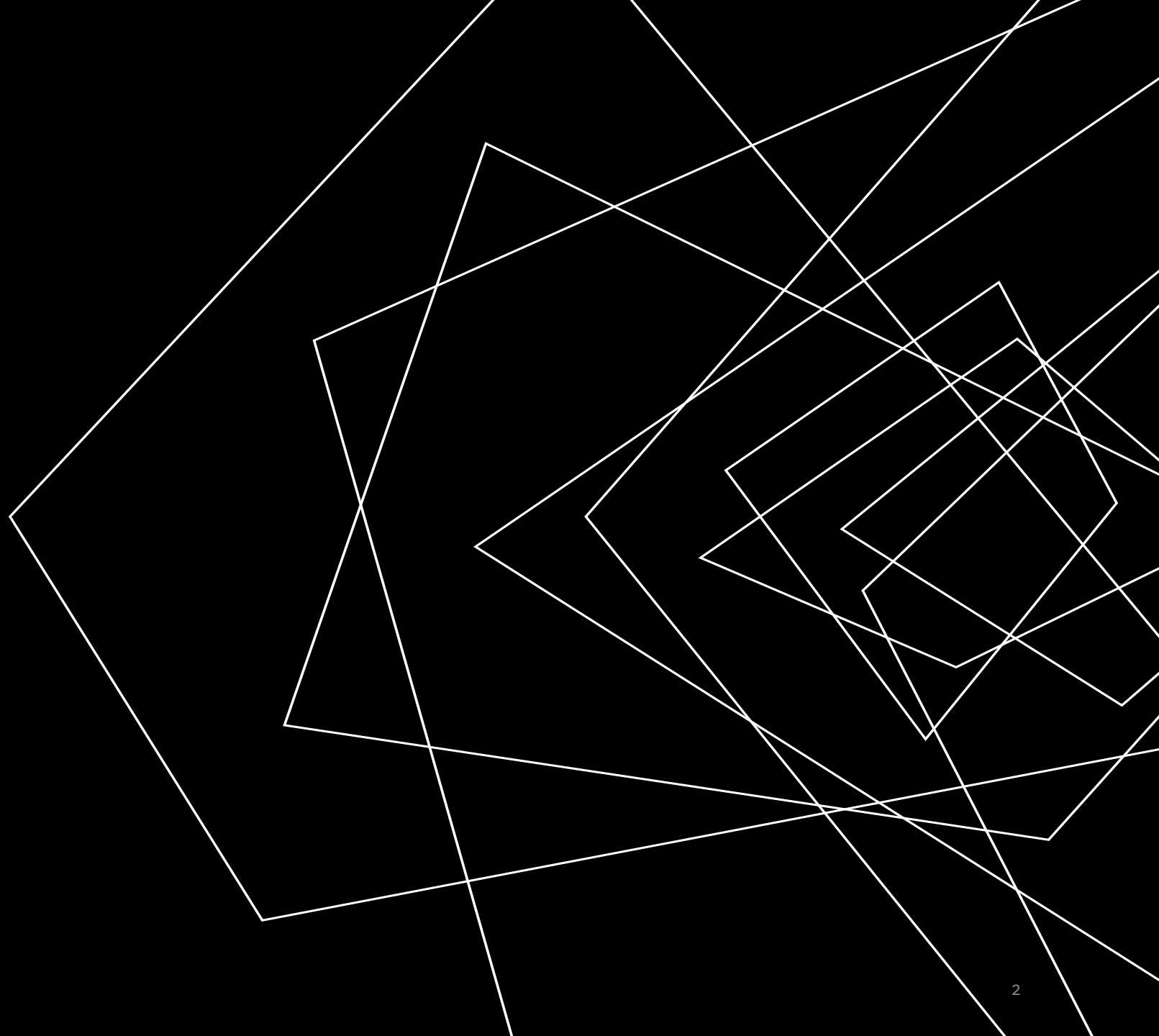
Overview

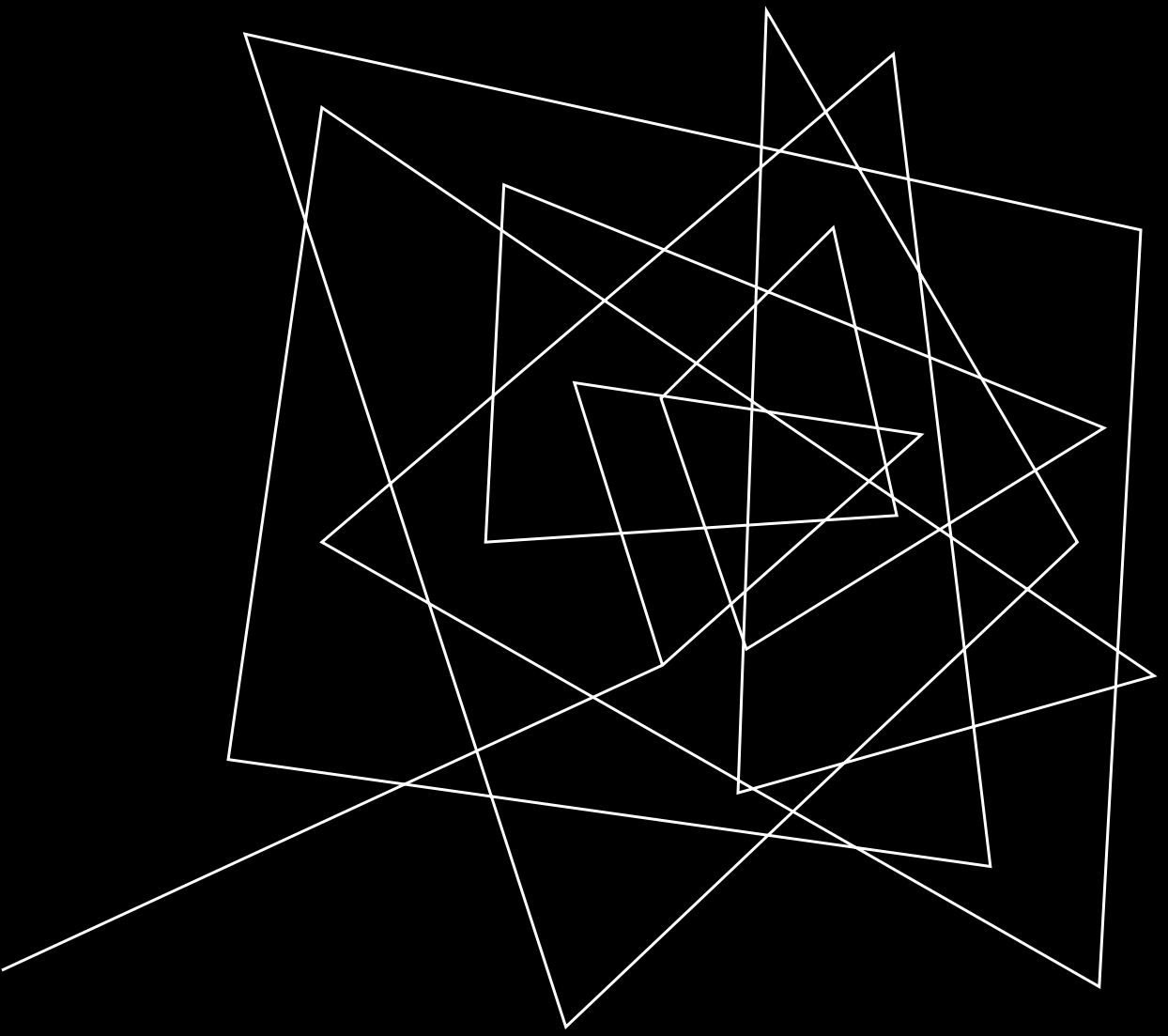
Methods Stability

Methods Time Complexity

Models Tunability

Conclusions





OVERVIEW

OVERVIEW

- 5 avionic-related datasets for binary classification, limited to 10k samples with basic preprocessing
- Optimization methods: GridSearch, BayesSearch
- CV = 3

XGBOOST

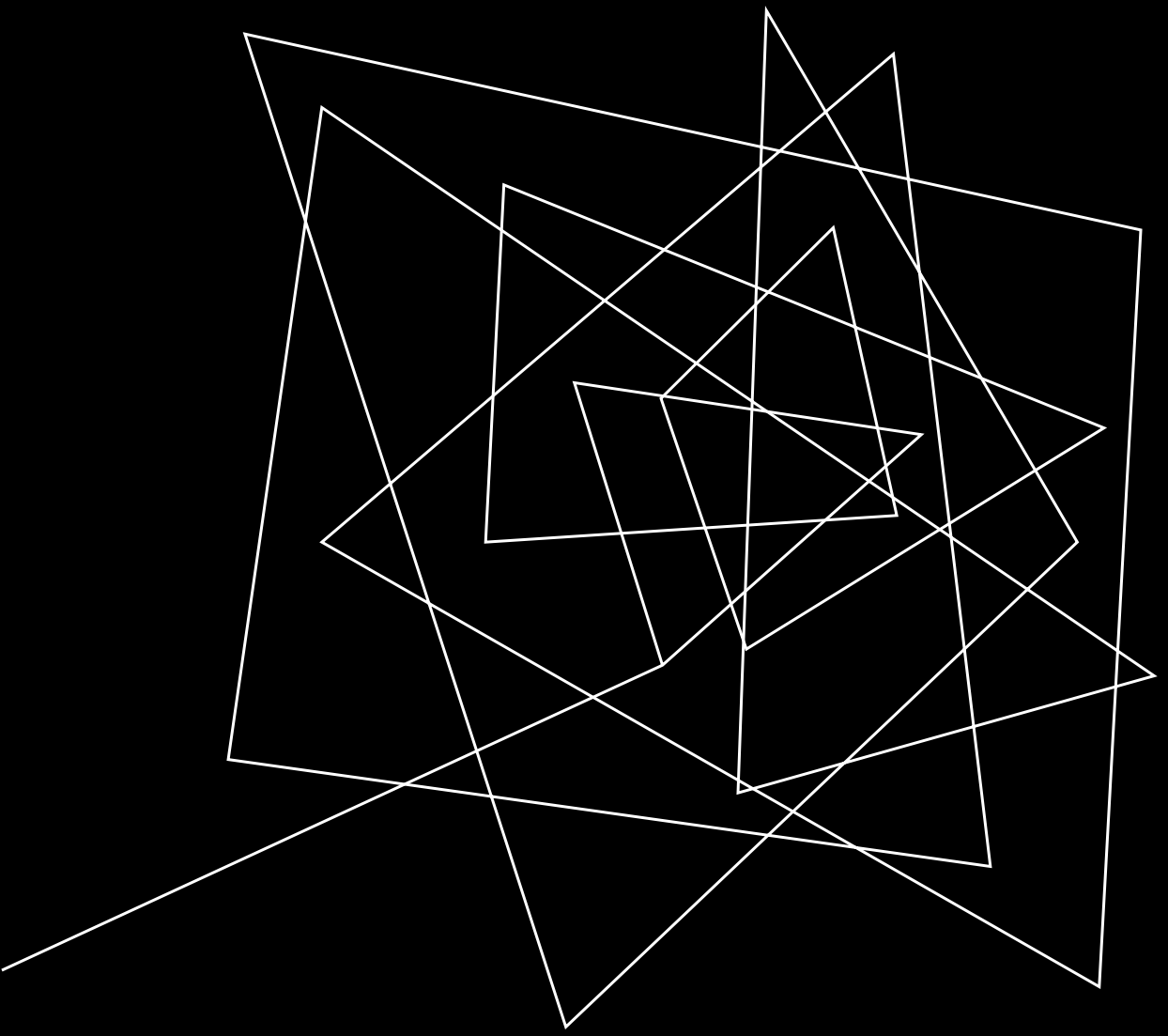
XGBoost Parameter Grid	
n_estimators:	100, 300, 500, 1000
learning_rate:	0.01, 0.05, 0.1
max_depth:	10, 15, 20
min_child_weight:	1, 2, 5, 10
gamma:	0, 0.15, 0.3
eval_metric:	"logloss", "error", "auc"
random_state:	42

LOGISTIC REGRESSION

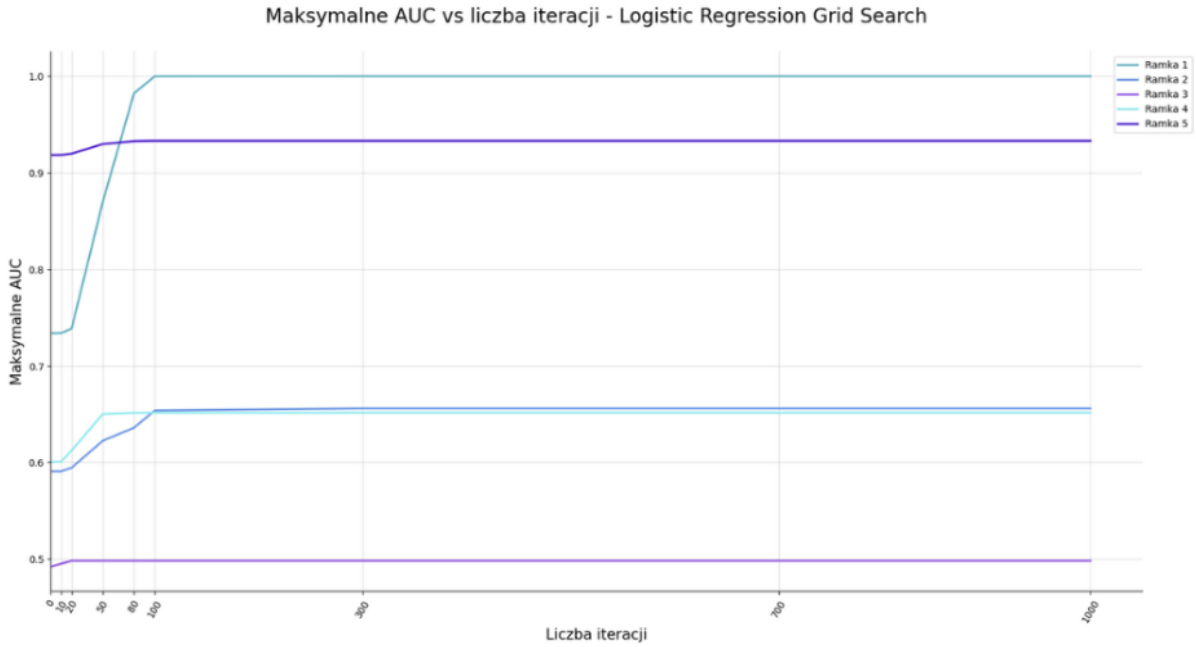
Logistic Regression Parameter Grid	
C:	np.logspace(-4, 4, 10)
max_iter:	50, 100, 200, 300, 500, 1000
tol:	1e-3, 1e-4, 1e-5, 1e-6
fit_intercept:	True, False
class_weight:	None, "balanced"
random_state:	42
max_iter:	500

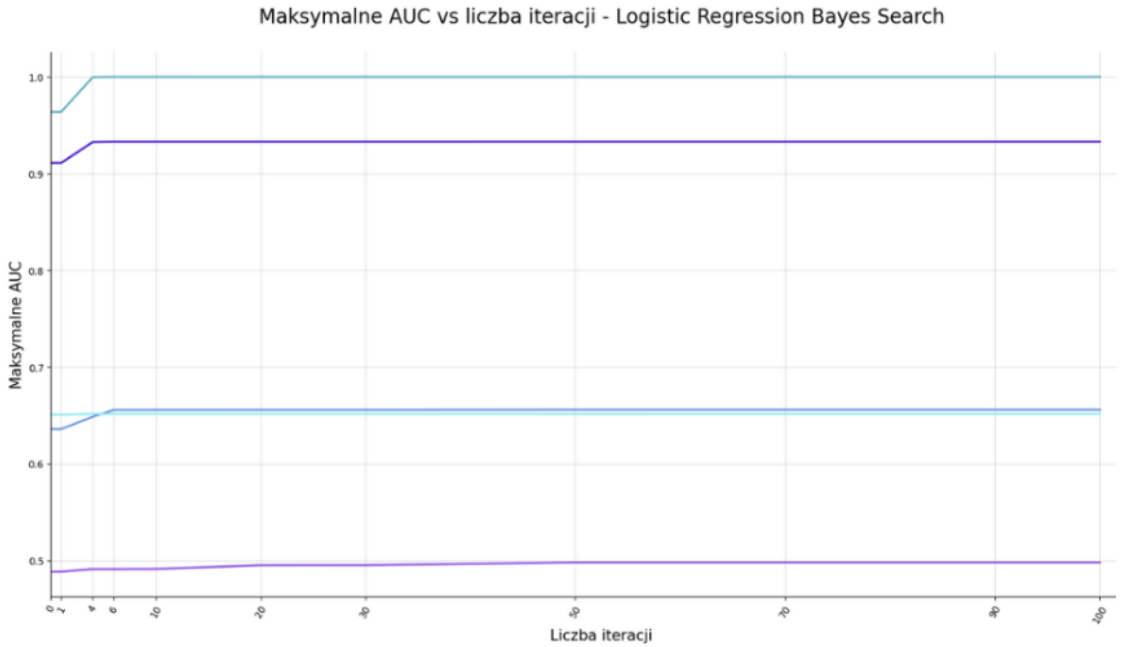
KNN

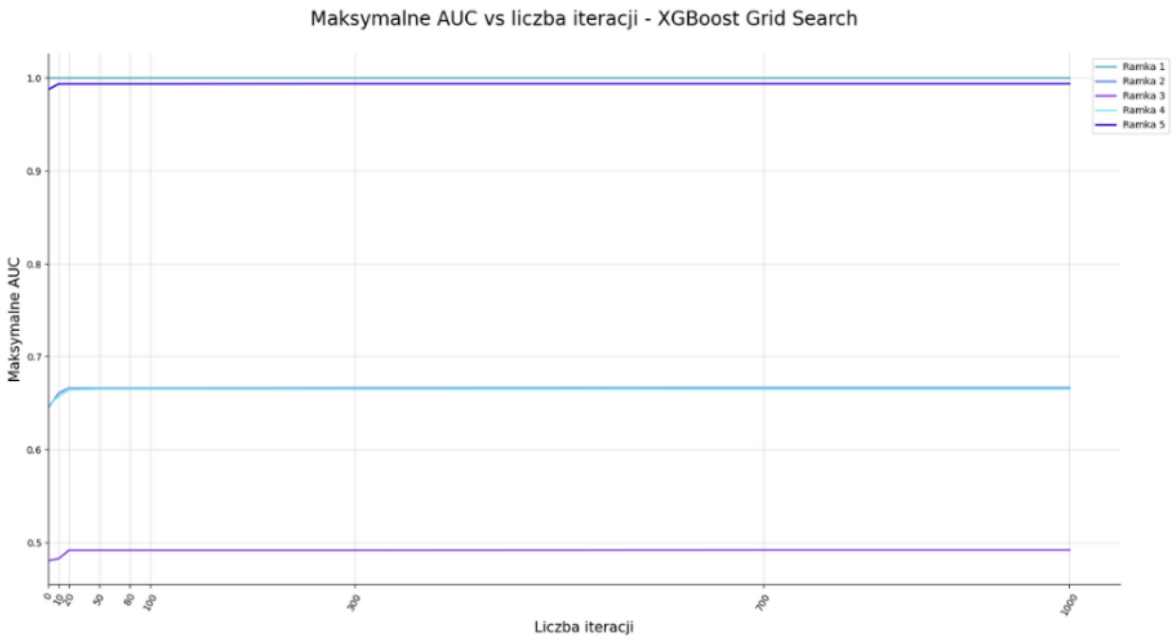
KNN Parameter Grid	
n_neighbors:	3, 5, 7, 9, 11, 15, 30, 50
weights:	"uniform", "distance"
algorithm:	"auto", "ball tree", "kd tree", "brute"
metric:	"euclidean", "manhattan", "minkowski", "chebyshev"
leaf_size:	20, 30, 40, 50

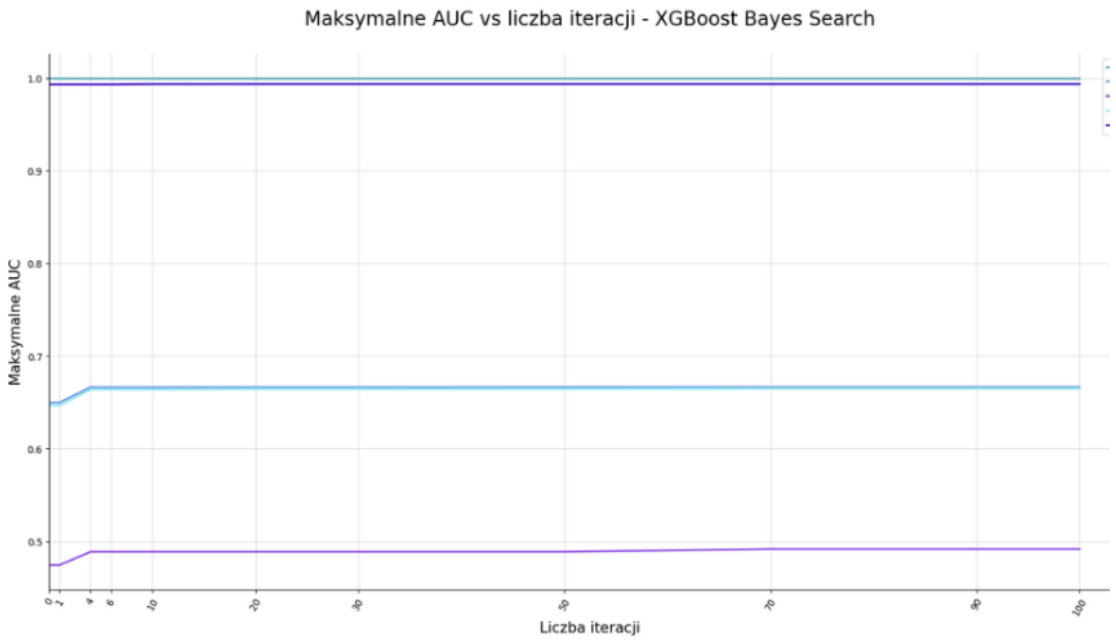


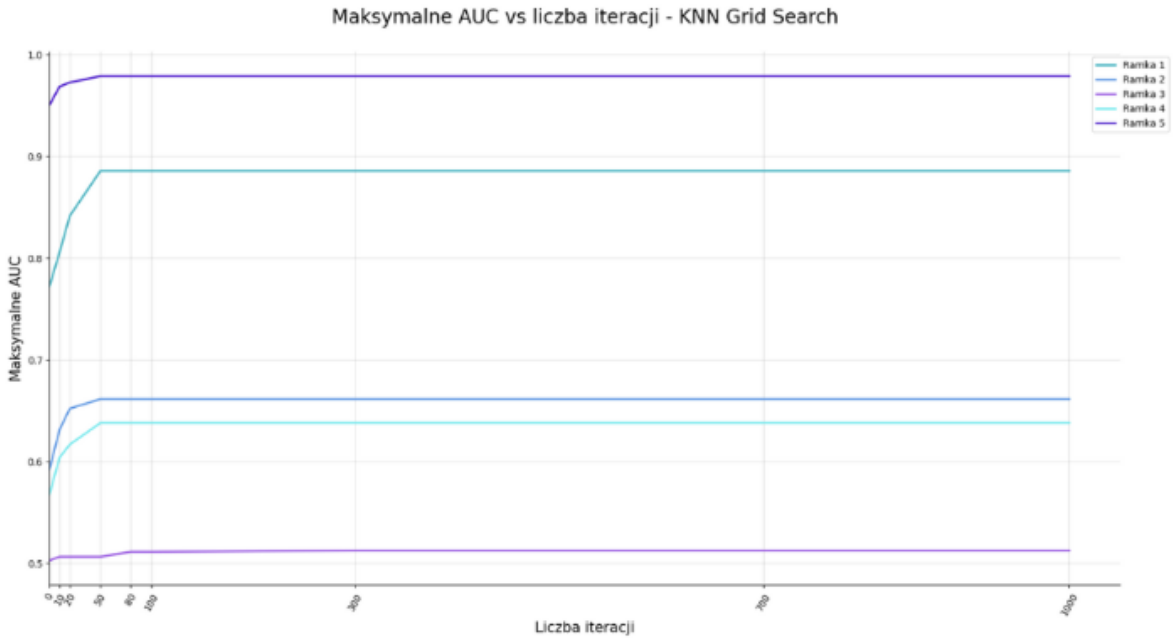
METHODS
STABILITY

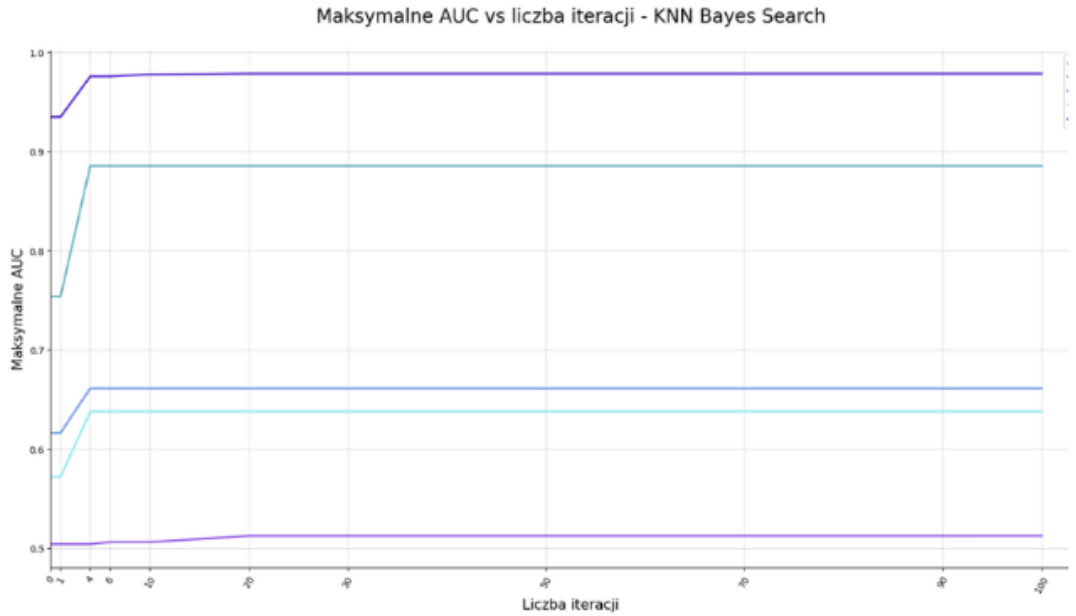

Maksymalne AUC vs liczba iteracji - Logistic Regression Grid Search
Ramka 1
Ramka 2
Ramka 3
Ramka 4
Ramka 5
Maksymalne AUC
Liczba iteracji

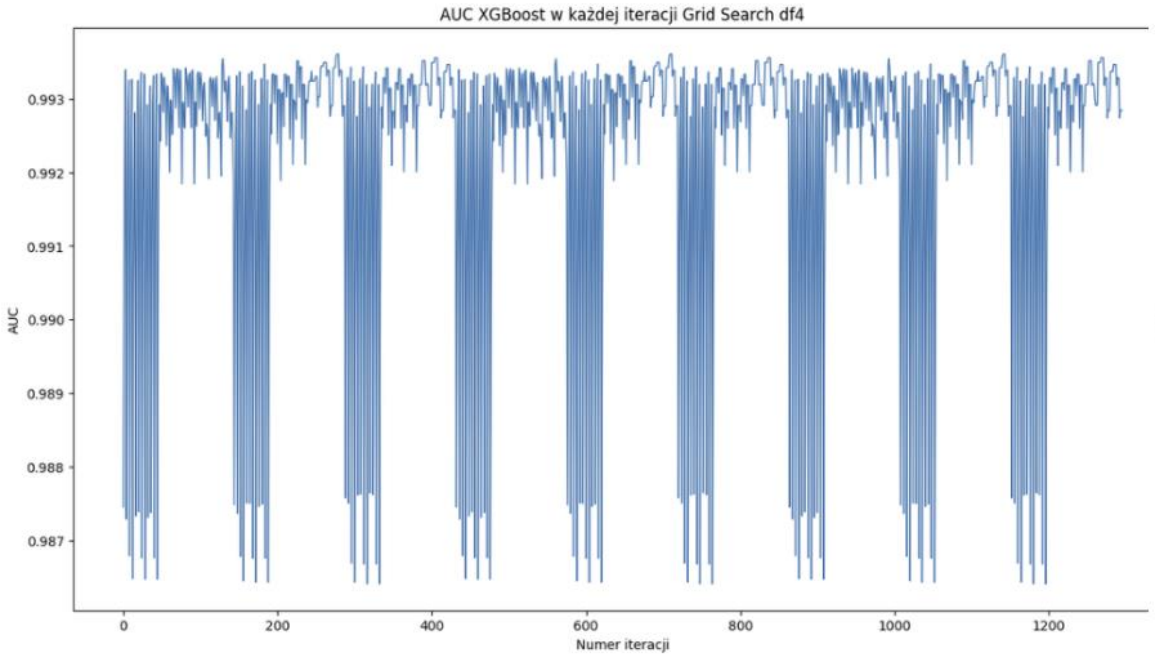

Maksymalne AUC vs liczba iteracji - Logistic Regression Bayes Search
Maksymalne AUC
Liczba iteracji

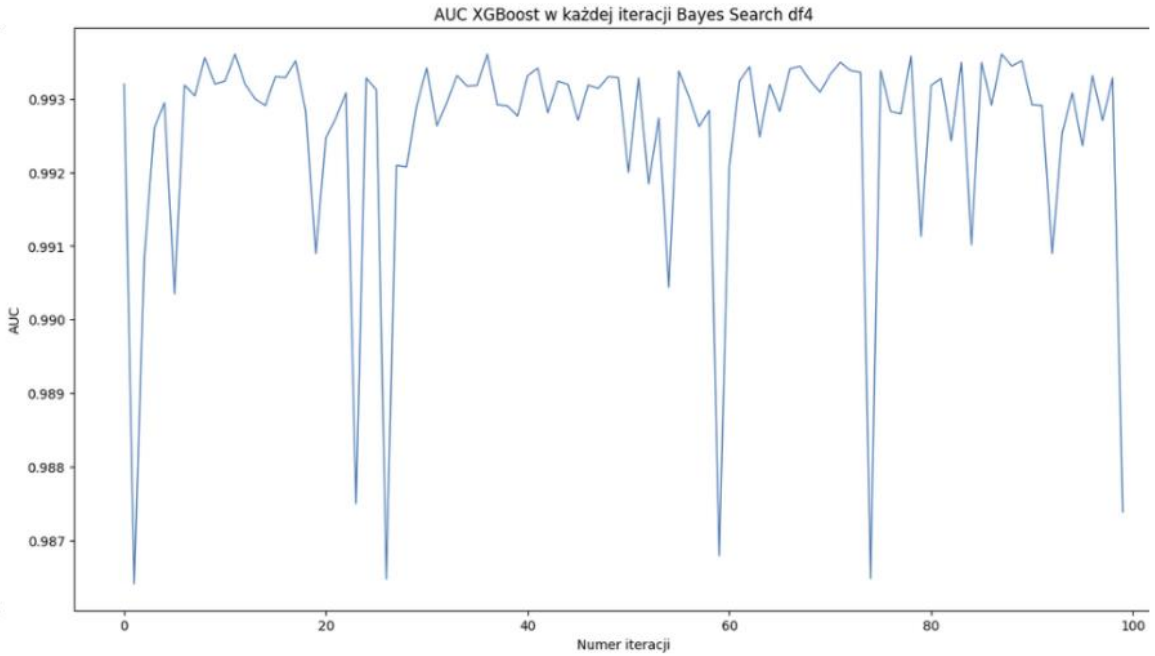

Maksymalne AUC vs liczba iteracji - XGBoost Grid Search
Ramka 1
Ramka 2
Ramka 3
Ramka 4
Ramka 5
Maksymalne AUC
1.0
0.9
0.8
0.7
0.6
0.5
0
10
20
50
80
100
300
700
1000
Liczba iteracji

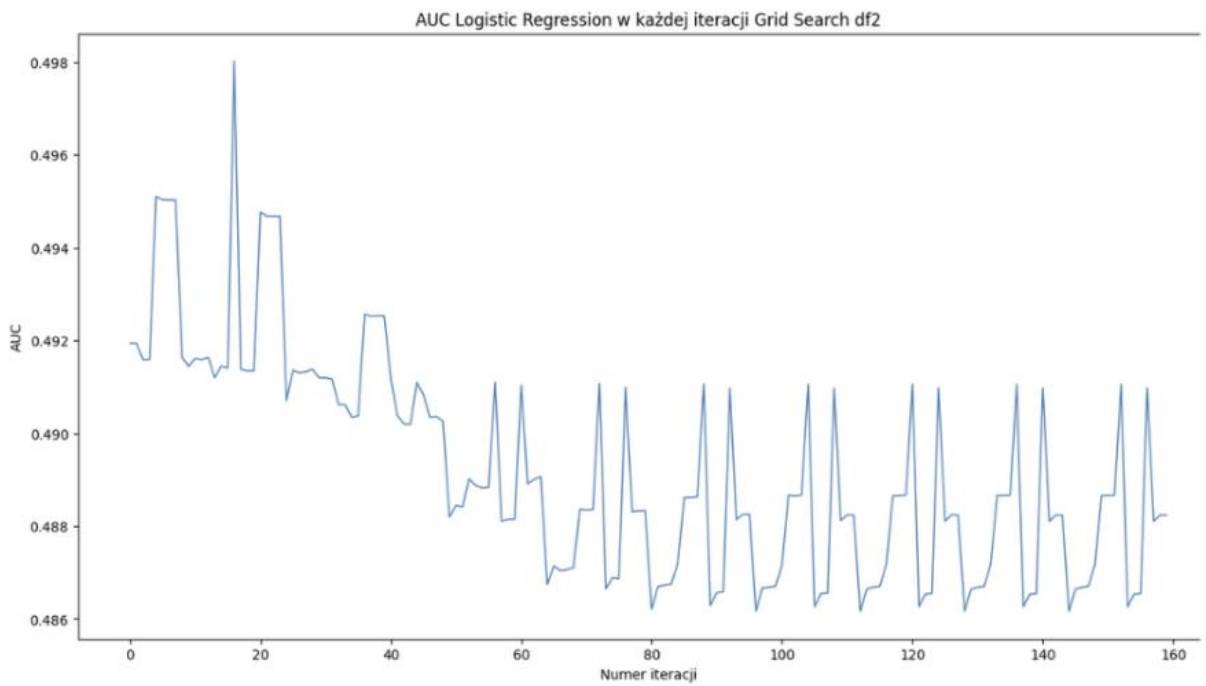

Maksymalne AUC vs liczba iteracji - XGBoost Bayes Search
Maksymalne AUC
1.0
0.9
0.8
0.7
0.6
0.5
0
1
4
6
10
20
30
50
70
90
100
Liczba iteracji

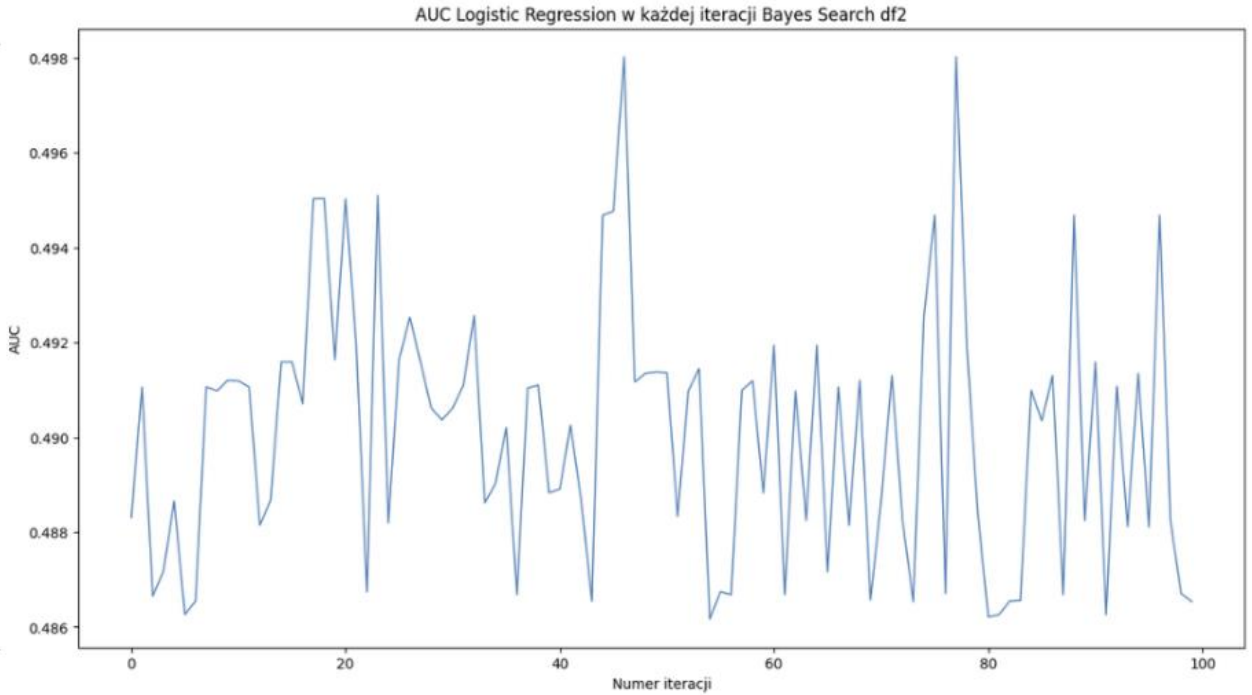

Maksymalne AUC vs liczba iteracji - KNN Grid Search
Ramka 1
Ramka 2
Ramka 3
Ramka 4
Ramka 5
Maksymalne AUC
Liczba iteracji

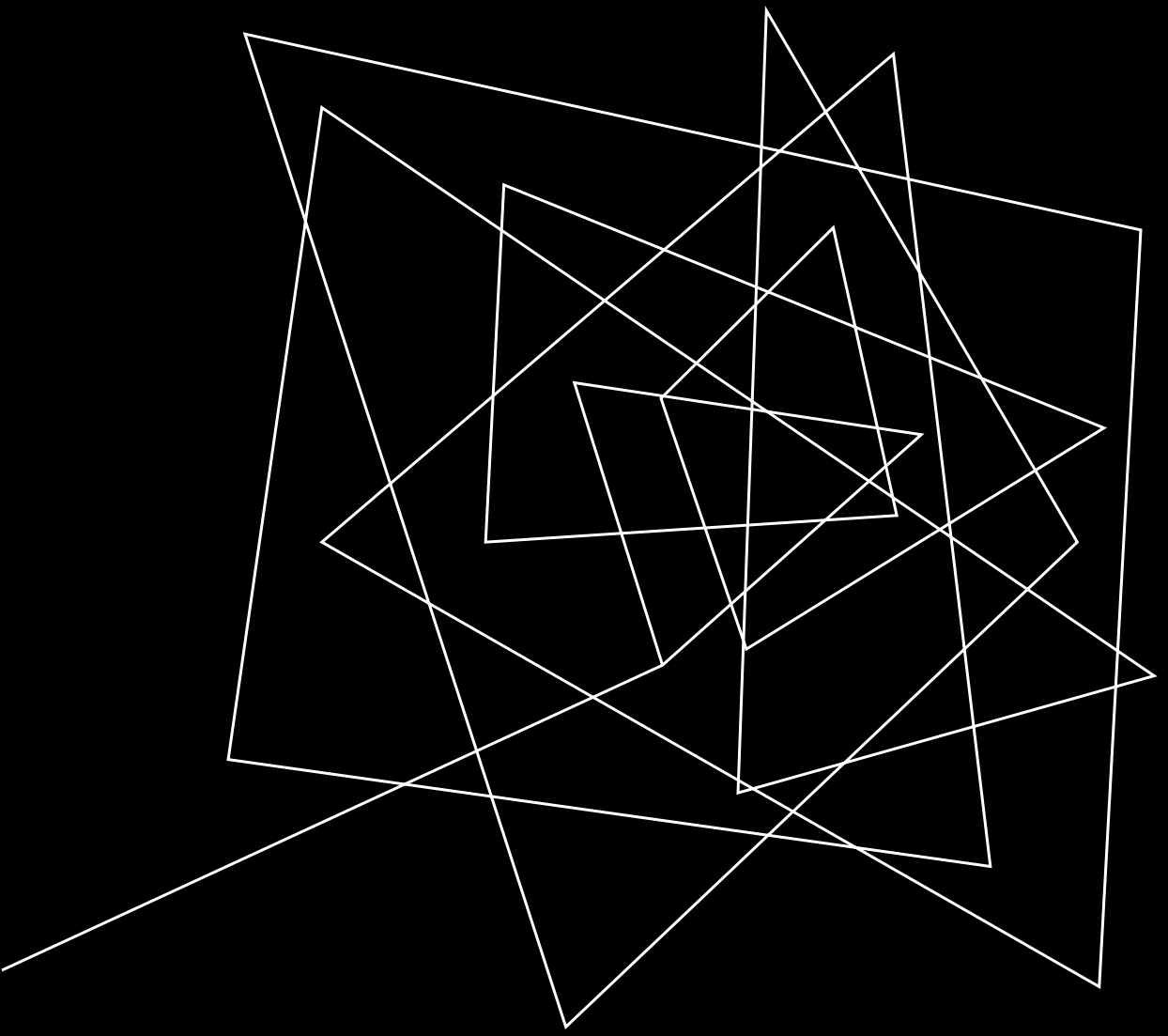

Maksymalne AUC vs liczba iteracji - KNN Bayes Search
Maksymalne AUC
Liczba iteracji


AUC XGBoost w każdej iteracji Grid Search df4
AUC
0.993
0.992
0.991
0.990
0.989
0.988
0.987
0
200
400
600
800
1000
1200
Numer iteracji


AUC XGBoost w każdej iteracji Bayes Search df4
AUC
0.993
0.992
0.991
0.990
0.989
0.988
0.987
0
20
40
60
80
100
Numer iteracji

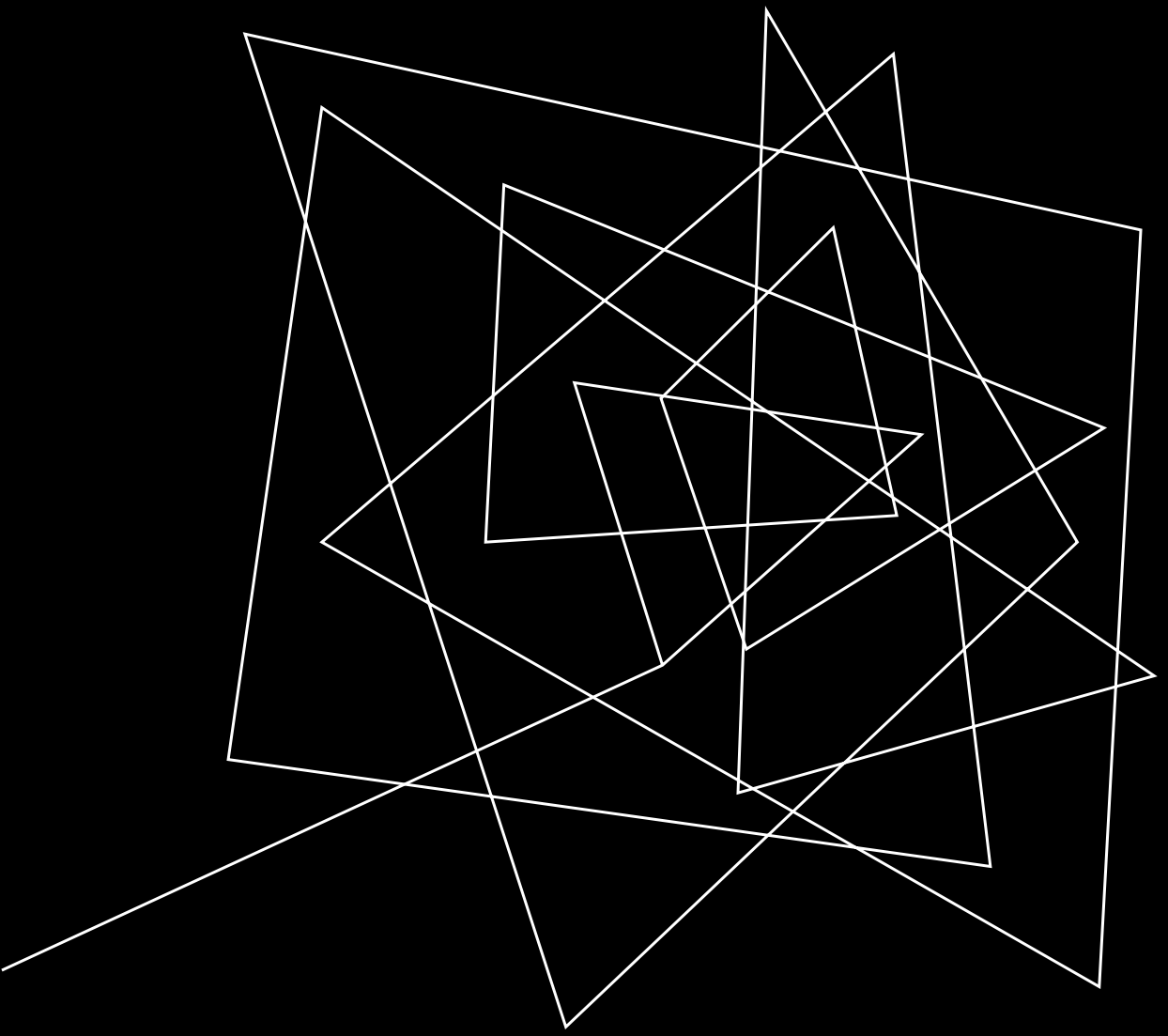

AUC Logistic Regression w każdej iteracji Grid Search df2
AUC
Numer iteracji
0.498
0.496
0.494
0.492
0.490
0.488
0.486
0
20
40
60
80
100
120
140
160


AUC Logistic Regression w każdej iteracji Bayes Search df2
AUC
Numer iteracji
0.498
0.496
0.494
0.492
0.490
0.488
0.486
0
20
40
60
80
100

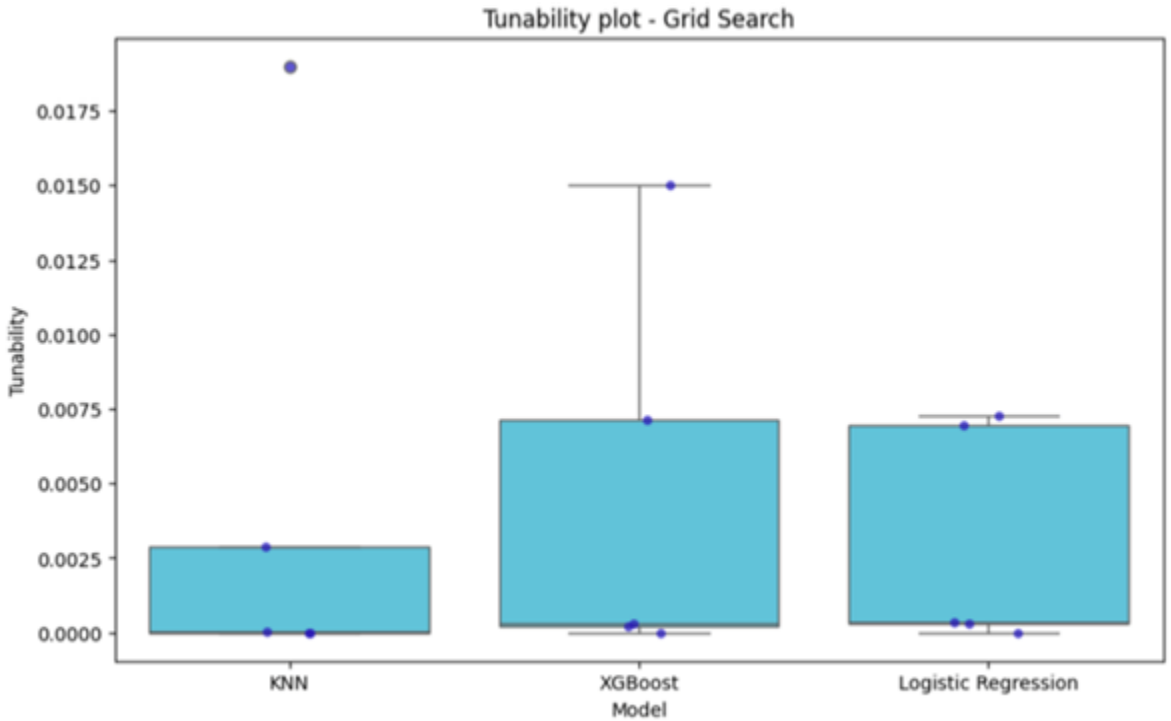


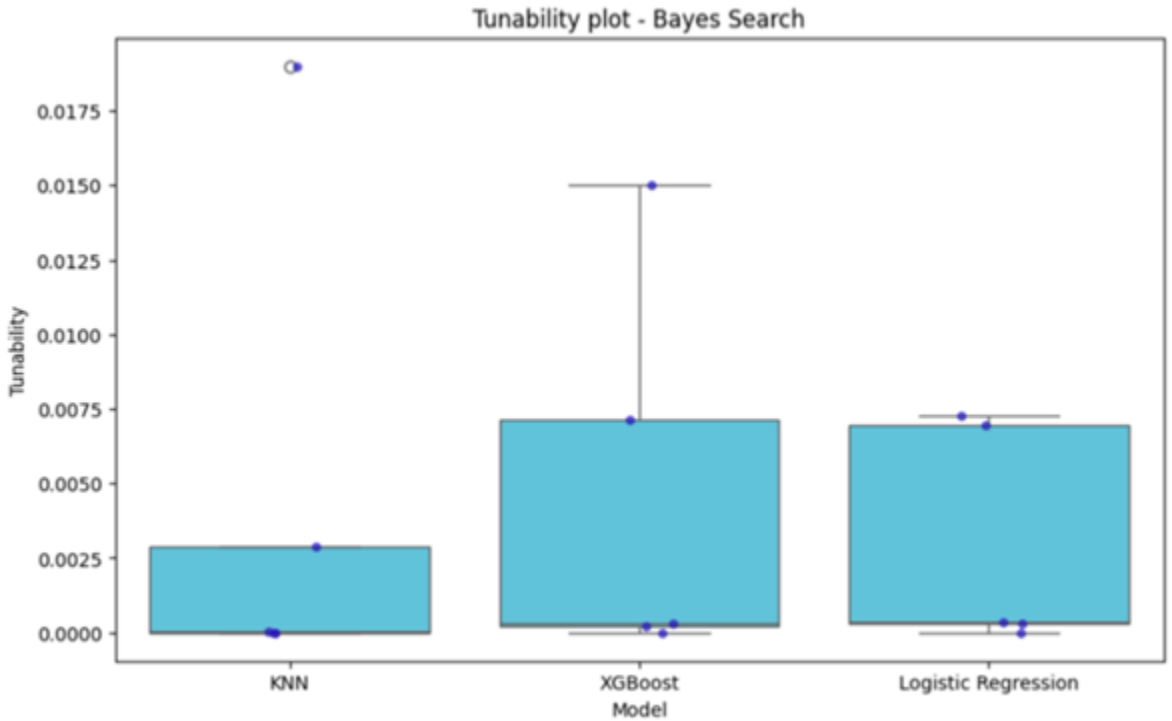
METHODS TIME COMPLEXITY

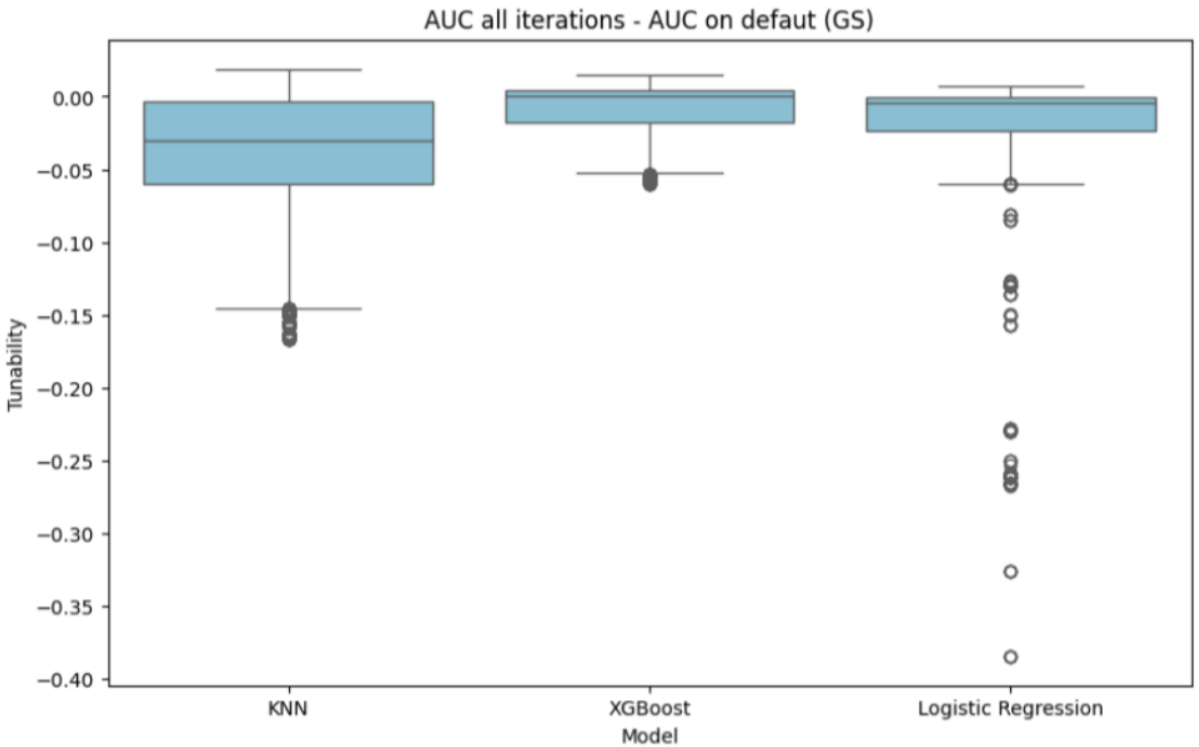
Model	Search Method	Dataset	Iterations	Time	Avg It Time
KNN	Bayes Search	1	100	207.470	2.070
KNN	Grid Search	1	1024	119.660	0.120
KNN	Bayes Search	2	100	240.060	2.400
KNN	Grid Search	2	1024	175.210	0.170
KNN	Bayes Search	3	100	232.420	2.320
KNN	Grid Search	3	1024	66.700	0.070
KNN	Bayes Search	4	100	274.810	2.750
KNN	Grid Search	4	1024	480.140	0.470
KNN	Bayes Search	5	100	228.410	2.280
KNN	Grid Search	5	1024	203.230	0.200
LR	Bayes Search	1	100	210.940	2.110
LR	Grid Search	1	160	3.200	0.020
LR	Bayes Search	2	100	233.240	2.330
LR	Grid Search	2	160	1.530	0.010
LR	Bayes Search	3	100	206.840	2.070
LR	Grid Search	3	160	0.860	0.010
LR	Bayes Search	4	100	233.010	2.330
LR	Grid Search	4	160	9.510	0.060
LR	Bayes Search	5	100	208.490	2.080
LR	Grid Search	5	160	1.410	0.010
XGBoost	Bayes Search	1	100	152.680	1.530
XGBoost	Grid Search	1	1296	76.890	0.060
XGBoost	Bayes Search	2	100	258.110	2.580
XGBoost	Grid Search	2	1296	584.910	0.450
XGBoost	Bayes Search	3	100	269.510	2.700
XGBoost	Grid Search	3	1296	472.370	0.360
XGBoost	Bayes Search	4	100	395.650	3.960
XGBoost	Grid Search	4	1296	1929.640	1.490
XGBoost	Bayes Search	5	100	242.160	2.420
XGBoost	Grid Search	5	1296	301.170	0.230

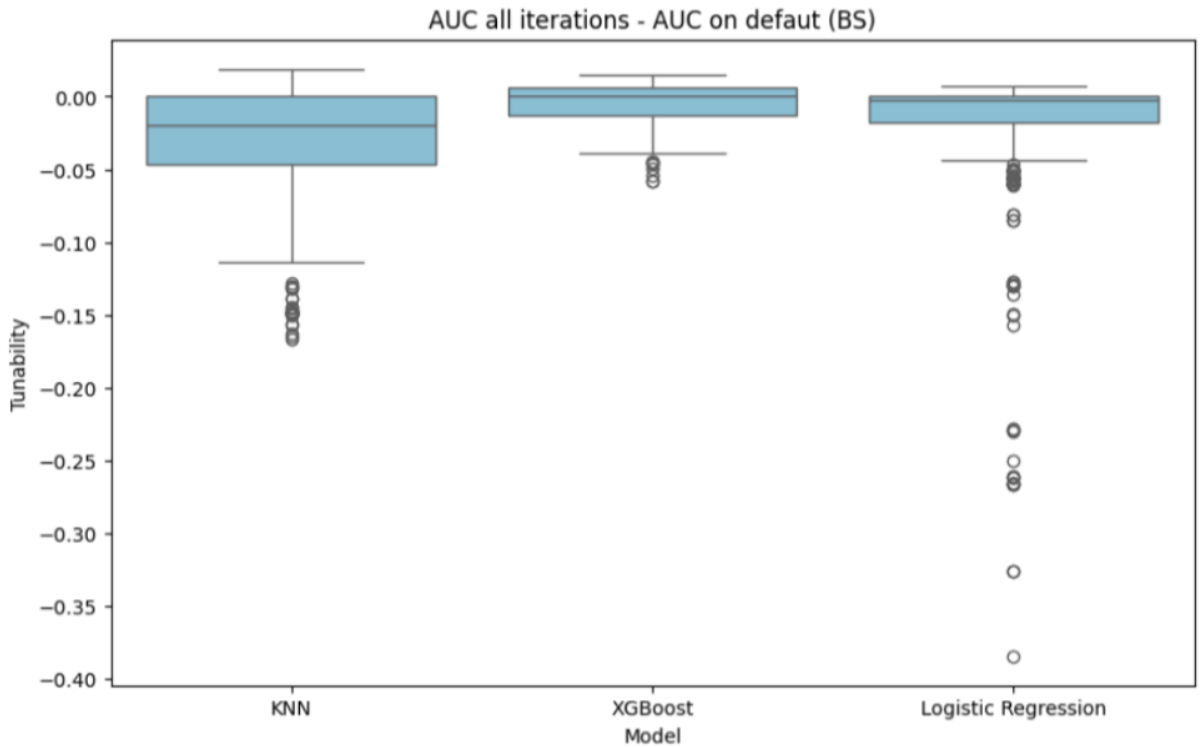


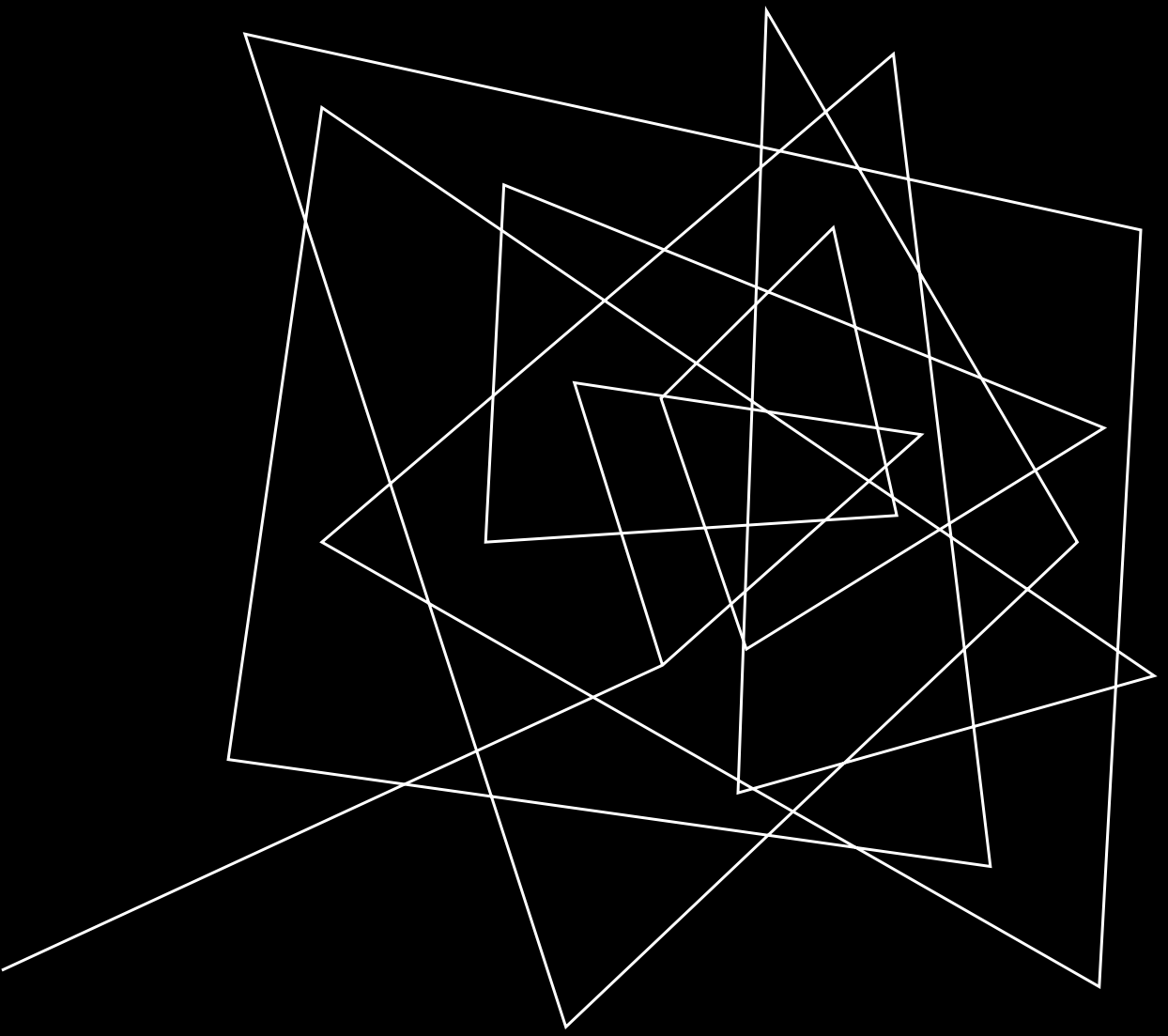
MODEL TUNABILITY


Tunability plot - Grid Search
Tunability
0.0175
0.0150
0.0125
0.0100
0.0075
0.0050
0.0025
0.0000
KNN
XGBoost
Logistic Regression
Model


Tunability plot - Bayes Search
Tunability
0.0175
0.0150
0.0125
0.0100
0.0075
0.0050
0.0025
0.0000
KNN
XGBoost
Logistic Regression
Model


AUC all iterations - AUC on defaut (GS)
Tunability
0.00
−0.05
−0.10
−0.15
−0.20
−0.25
−0.30
−0.35
−0.40
KNN
XGBoost
Logistic Regression
Model


AUC all iterations - AUC on defaut (BS)
Tunability
0.00
−0.05
−0.10
−0.15
−0.20
−0.25
−0.30
−0.35
−0.40
KNN
XGBoost
Logistic Regression
Model



THANKS FOR ATTENTION