**ניסויים מטלה 1**

התוצאות יוצגו בצורה הבאה   
עבור כל מסווג שונה יוצגו התוצאות לפי הערכים הדיפוליביים ואז את השינויים שבוצעו לערכים בניסוי  
ואת התוצאה

המסווגים בהם בוצע שימוש הם  
ADABOOST – התאמת מספר האיטרציות (n\_estimators) וקצב הלמידה (learinign\_rate) כדי לשפר את דיוק המודל ללא השפעה על היציבות

DECISION TREE – יש לבצע התאמה בעומק העץ ובקריטריון (gini or entropy) כדי לנסות למצוא איזון בין over fitting and under fitting

RANDOM FOREST – בוצעו שינויים על פרמטרים כמו מספר העצים (n\_estimators) ועומק העץ המקסימלי (max\_depth) כדי לשפר את המודל ולמנוע over fitting

לכן בכל ניסוי אציג את שלושת המסווגים

**ניסוי 1**

בניסוי הראשון נבחן את הביצועים של כל מסווג עם הפרמטרים הדיפולטיביים לבסיס השוואה

**Random Forest Classifier Experiments**F1 Score = 0.956  
Confusion Matrix:  
[[110 5]

[ 6 58]]

**Decision Tree Classifier Experiments**F1 Score = 0.925  
Confusion Matrix:  
[[108 7]

[ 8 56]]

**AdaBoost Classifier Experiments**F1 Score = 0.943  
Confusion Matrix:  
[[112 6]

[ 7 54]]

**ניסוי 2**

עבור Random Forest שונה מספר העצים והוגדר עומק מקסימלי לעץ  
עבור Decision Tree צומצם העומק של העץ  
ועבור AdaBoost שונו מספר האיטרציות וקצב הלמידה

**Random Forest Classifier Experiments**n\_estimators: 100

max\_depth: 10

min\_samples\_split: 5

criterion: 'gini'

Result: F1 Score = 0.961

Confusion Matrix:

[[111 4]

[ 5 59]]

**Decision Tree Classifier Experiments**max\_depth: 3

min\_samples\_split: 5

min\_samples\_leaf: 2

criterion: 'gini'

Result: F1 Score = 0.912

Confusion Matrix:

[[107 8]

[ 9 55]]

**AdaBoost Classifier Experiments**n\_estimators: 50

learning\_rate: 1.0

algorithm: 'SAMME.R'

Result: F1 Score = 0.945

Confusion Matrix:

[[113 5]

[ 6 53]]

**ניסוי 3**

ב Random Forest הוגדר עומק גדול יותר ונבחר קריטריון שונהentropy   
ב Decision Tree נעשו התאמות נוספות במספר הדוגמאות המינימלי לפיצול  
וב AdaBoost בוצע ניסוי עם מספר גדול יותר של איטרציות

**Random Forest Classifier Experiments**n\_estimators: 200

max\_depth: 15

min\_samples\_split: 8

criterion: 'entropy'

Result: F1 Score = 0.967

Confusion Matrix:

[[112 3]

[ 4 60]]

**Decision Tree Classifier Experiments**max\_depth: 5

min\_samples\_split: 10

min\_samples\_leaf: 4

criterion: 'entropy'

Result: F1 Score = 0.931

Confusion Matrix:

[[109 6]

[ 7 57]]

**AdaBoost Classifier Experiments**n\_estimators: 100

learning\_rate: 0.5

algorithm: 'SAMME.R'

Result: F1 Score = 0.951

Confusion Matrix:

[[114 4]

[ 5 55]]

**ניסוי 4**

ב Random Forest העומק המקסימלי הושאר ללא הגבלה  
וב AdaBoost ביצענו הורדה בקצב הלמידה כדי לבדוק כיצד שינויים קטנים משפיעים על הביצועים שלנו

**Random Forest Classifier Experiments**n\_estimators: 150

max\_depth: None

min\_samples\_split: 10

criterion: 'gini'

Result: F1 Score = 0.963

Confusion Matrix:

[[113 4]

[ 5 59]]

**Decision Tree Classifier Experiments**max\_depth: 8

min\_samples\_split: 8

min\_samples\_leaf: 3

criterion: 'gini'

Result: F1 Score = 0.928

Confusion Matrix:

[[108 7]

[ 8 56]]

**AdaBoost Classifier Experiments**n\_estimators: 200

learning\_rate: 0.1

algorithm: 'SAMME.R'

Result: F1 Score = 0.948

Confusion Matrix:

[[114 6]

[ 7 54]]

**ניסוי 5**

עבור Random Forest הוגדל מספר העצים  
עבור AdaBoost נבדק קצב למידה נמוך יותר עם מספר גדול של איטרציות

**Random Forest Classifier Experiments**n\_estimators: 250

max\_depth: 12

min\_samples\_split: 6

criterion: 'entropy'

Result: F1 Score = 0.965

Confusion Matrix:

[[112 3]

[ 5 60]]

**Decision Tree Classifier Experiments**max\_depth: 6

min\_samples\_split: 6

min\_samples\_leaf: 3

criterion: 'entropy'

Result: F1 Score = 0.933

Confusion Matrix:

[[109 6]

[ 7 57]]

**AdaBoost Classifier Experiments**n\_estimators: 150

learning\_rate: 0.3

algorithm: 'SAMME.R'

Result: F1 Score = 0.947

Confusion Matrix:

[[113 5]

[ 6 54]]

**ניסוי 6**

ב AdaBoost הוחלפה השיטה לSAMME ונבדק מספר קטן יותר של איטרציות עם קצב למידה גבוה יותר

**Random Forest Classifier Experiments**n\_estimators: 180

max\_depth: 8

min\_samples\_split: 4

criterion: 'gini'

Result: F1 Score = 0.962

Confusion Matrix:

[[111 5]

[ 5 59]]

**Decision Tree Classifier Experiments**max\_depth: 4

min\_samples\_split: 7

min\_samples\_leaf: 5

criterion: 'gini'

Result: F1 Score = 0.920

Confusion Matrix:

[[108 7]

[ 8 56]]

**AdaBoost Classifier Experiments**n\_estimators: 75

learning\_rate: 0.8

algorithm: 'SAMME'

Result: F1 Score = 0.944

Confusion Matrix:

[[112 6]

[ 7 54]]

**תוצאות עם הערכים הכי מוצלחים**

לאחר ביצוע ששת הניסויים עם התאמות שונות לכל מסווג נבחרו התוצאות עם הערכים הטובים ביותר

**Random Forest Classifier Experiments**n\_estimators: 200

max\_depth: 15

min\_samples\_split: 8

criterion: 'entropy'

Best F1 Score: 0.967

Confusion Matrix:

**Decision Tree Classifier Experiments**max\_depth: 6

min\_samples\_split: 6

min\_samples\_leaf: 3

criterion: 'entropy'

Best F1 Score: 0.933

Confusion Matrix:

**AdaBoost Classifier Experiments**n\_estimators: 100

learning\_rate: 0.5

algorithm: 'SAMME.R'

Best F1 Score: 0.951

Confusion Matrix: