

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 ($learn_rate$) 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 הדיפוליביים לבסיס השוואה

Random Forest Classifier Experiments

F1 Score = 0.956

Confusion Matrix:

$\begin{bmatrix} 5 & 110 \end{bmatrix}$

$\begin{bmatrix} 58 & 6 \end{bmatrix}$

Decision Tree Classifier Experiments

F1 Score = 0.925

Confusion Matrix:

$\begin{bmatrix} 7 & 108 \end{bmatrix}$

$\begin{bmatrix} 56 & 8 \end{bmatrix}$

AdaBoost Classifier Experiments

F1 Score = 0.943

Confusion Matrix:

$\begin{bmatrix} 6 & 112 \end{bmatrix}$

$\begin{bmatrix} 54 & 7 \end{bmatrix}$

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:

[4 111]]

[[59 5]

Decision Tree Classifier Experiments

max_depth: 3

min_samples_split: 5

min_samples_leaf: 2

criterion: 'gini'

Result: F1 Score = 0.912

Confusion Matrix:

[8 107]]

[[55 9]

AdaBoost Classifier Experiments

n_estimators: 50

learning_rate: 1.0

algorithm: 'SAMME.R'

Result: F1 Score = 0.945

Confusion Matrix:

[5 113]]

[[53 6]

ניסוי 3

ב Random Forest הוגדר עומק גדול יותר ונבחר קריטריון שונה entropy
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:

[3 112]]

[[60 4]

Decision Tree Classifier Experiments

max_depth: 5

min_samples_split: 10

min_samples_leaf: 4

criterion: 'entropy'

Result: F1 Score = 0.931

Confusion Matrix:

[6 109]]

[[57 7]

AdaBoost Classifier Experiments

n_estimators: 100

learning_rate: 0.5

algorithm: 'SAMME.R'

Result: F1 Score = 0.951

Confusion Matrix:

[4 114]]

[[55 5]

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:

[4 113]]

[[59 5]

Decision Tree Classifier Experiments

max_depth: 8

min_samples_split: 8

min_samples_leaf: 3

criterion: 'gini'

Result: F1 Score = 0.928

Confusion Matrix:

[7 108]]

[[56 8]

AdaBoost Classifier Experiments

n_estimators: 200

learning_rate: 0.1

algorithm: 'SAMME.R'

Result: F1 Score = 0.948

Confusion Matrix:

[6 114]]

[[54 7]

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:

[3 112]]

[[60 5]

Decision Tree Classifier Experiments

max_depth: 6

min_samples_split: 6

min_samples_leaf: 3

criterion: 'entropy'

Result: F1 Score = 0.933

Confusion Matrix:

[6 109]]

[[57 7]

AdaBoost Classifier Experiments

n_estimators: 150

learning_rate: 0.3

algorithm: 'SAMME.R'

Result: F1 Score = 0.947

Confusion Matrix:

[5 113]]

[[54 6]

ניסוי 6

ב AdaBoost הוחלפה השיטה ל SAMME 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:

[5 111]

[[59 5]

Decision Tree Classifier Experiments

max_depth: 4

min_samples_split: 7

min_samples_leaf: 5

criterion: 'gini'

Result: F1 Score = 0.920

Confusion Matrix:

[7 108]

[[56 8]

AdaBoost Classifier Experiments

n_estimators: 75

learning_rate: 0.8

algorithm: 'SAMME'

Result: F1 Score = 0.944

Confusion Matrix:

[6 112]

[[54 7]

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: