**תרגול3**

**Model Evaluation**

משתמשים ב2 מתודות שונות על הtraining set:

רגרסיה לינארית- לא משנה איך נזיז את הישר, הוא אף פעם לא יתפוס את המגמה, אם היא לא לינארית, לכן למודל זה יש bias גדול בדרך כלל, אך זה מודל שלא משתנה יותר מדי בין datasets שונים, לדוגמה בין הtraining לtest-שונות נמוכה.

Curved line- מודל אשר יש לו פחות bias, יכול לתפוס את כל המגמה בין הנתונים, אך מודל זה רגיש לdataset שבו משתמשים. עבור datasets שונים נקבל curved line שונה- bias נמוך, שונות גבוהה.

**Underfit**-מודל שלא תופס את המגמה בצורה טובה-high bias, low variance

**Overfit**- כאשר המודל מותאם יותר מדי לdataset מסוים ולא יתאים לdatasets אחרים- Low Bias, High Variance

המטרה שלנו היא למזער את הטעות הכללית

טעות כללית=Err

-טעות כללית שתמיד קיימת

קיים יחס tradeoff בין מורכבות המודל לטעות הכללית

נדרש למצוא את הנקודה הספציפית בין גודל השונות והbias, בה בטעות הכללית היא הנמוכה ביותר.

לדוגמה ניתן להגביל את העומק של העץ, ניקח כמה עומקים שונה ונמדוד איזה עץ הוא האופטימלי.

Low bias and low variance- המודל יחזה את הנתונים באופן מדויק

Low bias and high variance-המודל נותן תוצאות יחסית מדויקות אבל משתנה מאוד כאשר נותנים לו להתאמן על datasets שונים.

High bias and low variance- המודל לא מדויק, אבל נותן תוצאות דומות עבור datasets שונים.

High bias and high variance- המודל לא מדויק ומשתנה מאוד כאשר נותנים לו להתאמן על datasets שונים.

**2 class confusion matrix**

Text

Description automatically generated

A picture containing table

Description automatically generated

בדוגמה המספרים בטבלה מייצגים כמה נתונים יש לנו בכל הגדרה.

Accuracy- מודד כמה חיזויים נכונים עשינו מתוך כלל החיזויים

Text

Description automatically generated with medium confidence

Sensitivity- כמה צדקנו בחיזוי חיובי

כמה חשבת שהם חיוביים וצדקת חלקי כל מה שהוא חיובי

כאשר יותר חשוב לזהות חיוביים נכון -נשתמש בSensitivity

Graphical user interface, diagram

Description automatically generated with medium confidence

Specifity- כמה ניחשת נכון שהם שליליים מתוך כל השליליים

כאשר יותר חשוב לזהות שליליים נכון -נשתמש ב specifty

Graphical user interface, text, application

Description automatically generated with medium confidence

Precision- כמה תצפיות ניחשת נכון שהן חיוביות חלקי כל מה שחשבת שהוא חיובי

Chart

Description automatically generated with medium confidence

F1 Score- ממוצע ממושקל בין recall ו precision

Text

Description automatically generated

יש הרבה מדדים לבדוק את דיוק המודל כיוון שיכולות להיות בעיות שונות, לדוגמה:

אם יש טעות בזיהוי classאחד יותר מאשר השני זה קשור לspecifity.

תלוי גם במה המטרה של המודל על פי מטרת המודל נבחר מדדים שונים להעריך את המודל.

Multi class confusion matrix

מסתכלים על הערכים של TP,TN,FP,FN על פי כל קלאס בנפרד. הTrue positive יהי התא של אותו קלאס גם בשורה וגם בטור(אלכסון אמצעי) והשאר יהיו בהתאם.

Table

Description automatically generated

**מדדים של רגרסיה**

**MSE-** סכום הריבועים של הטעות חלקי מספר התצפיות, משווים את הערך בין המודלים ובוחרים את המודל עם הMSE הנמוך ביותר

**MAE-** סכום הטעויות בערך מוחלט חלקי מספר התצפיות

**R Square**- מודד כמה שונות מתוך המשתנה התלוי אפשר להסביר באמצעות המודל, ככל שהוא יותר גבוה אז המודל יותר טוב. מחשבים את סכום ההפרש בין ערך התצפית האמיתי למה שניחשת שיהיה ערך התצפית **חלקי** סכום ההפרש בין ערך התצפית האמיתי לממוצע ערכי התצפיות האמיתיות. מפחיתים את התוצאה מ1

Text, letter

Description automatically generated