חיזוי פרמטרים על פי ריכוזי ברזל וליפידים

207946229 שירלי אליעזר

2020 בדצמבר 31

חלק I

הקדמה

במסמך על פי ריכוזי על פרויקט סריקת פרמטרים פרויקט פי ריכוזי ברזל פי במסמך במסמך אציג את שלבי פרויקט חיזוי פרמטרים של פי עלפידים כפי שנקבעו במבחנות הניסוי.

כמו כן, אציג את שלבי חיזוי ריכוזי הברזל והליפידים על פי מערכת לומדת ונתוני התחלה.

חלק II

הצגת הנתונים

חיזוי על פי ריכוז ברזל

הגרפים הבאים מציגים חיזוי של הפרמטר הנבחר, כתלות בריכוז הברזל. שיטת החיזוי היא רגרסיה לינארית המשתמשת ב- cross validation כאשר שיטת החלוקה לדאטא־אימון ודאטא־מבחן היא leave one out.

חיזוי על פי ריכוז ברזל וכמות הליפיד

הגרפים הבאים מציגים חיזוי של הפרמטר הנבחר, כתלות בריכוז הברזל וכמות הליפיד. שיטת החיזוי היא רגרסיה לינארית המשתמשת ב- cross validation כאשר שיטת החלוקה לדאטא־אימון ודאטא־מבחן היא leave one out.

חיזוי על פי ריכוז ברזל, כמות הליפיד, אינטרקציה בין ריכוז ברזל וכמות ליפיד

בשלב זה, הרחבתי את המשתנים הבלתי תלויים לשלוש עמודות:

- 1. ריכוז ברזל כולל את סביבות הברזל הבאות:
 - FEא) ברזל חופשי ו־

- FE (1)
- Transferrin ') Ferritin (2)
 - 2. כמות ליפיד
 - 3. ריכוז ברזל * כמות ליפיד

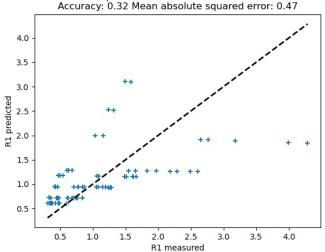
הרחבת הנתונים נעשתה על ידי הכפלה element wise של נתוני ריכוזי הברזל וכמות הליפיד. לאחר שיש בידינו את העמודה המייצגת נתונים אלו, ניתן לערוך רגרסיה לינארית מרובה, תוך שימוש בקרוס־ולידאציה, כנעשה קודם.

כעת, ניתן לראות שיפור בתוצאת R2 squared בחלק מן המשתנים התלויים:

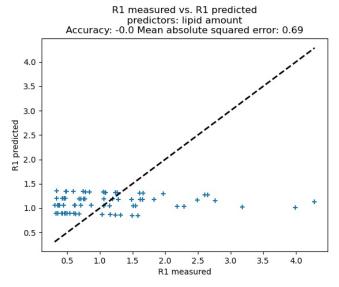
${ m FE}$ ברזל חופשי

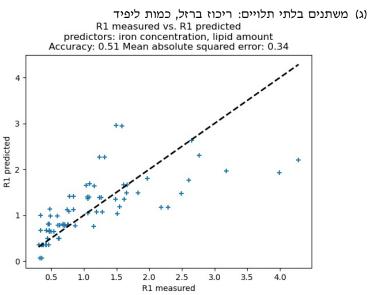
:R1 .1





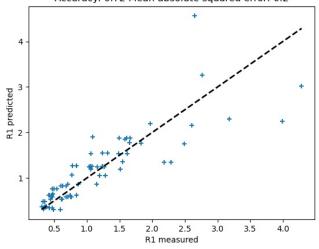
(ב) משתנה בלתי תלוי: כמות ליפיד





(ד) משתנים בלתי תלויים ריכוז ברזל, כמות ליפיד, ריכוז ברזל * כמות ליפיד

R1 measured vs. R1 predicted predictors: iron concentration, lipid amount, interaction Accuracy: 0.72 Mean absolute squared error: 0.2

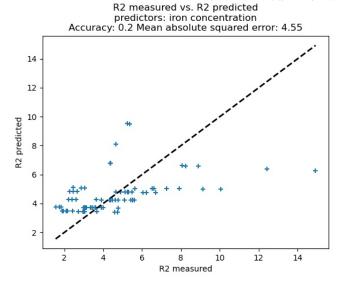


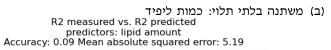
מסקנות: ניתן להראות מהגרפים לעיל כי לריכוז הברזל יש השפעה על R1 יותר מאשר לכמות הליפיד.

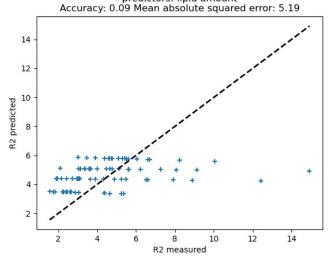
כמו כן, ניתן לראות שחיזוי R1 על ידי ריכוז הברזל וכמות הליפיד מניבה תוצאה R2_squared=0.32 מול $R2_squared=0.51$ ור טובה יותר מאשר כל אחד לחוד ($R2_squared=0.51$) ותוצאה זו משתפרת עוד יותר כאשר מחשיבים את האינטרקציה בין ריכוז הברזל וכמות הליפיד ($R2_squared=0.72$). על פי אלו, ניתן לקבוע כי המודל הלינארי המתאים לחיזוי R1 הוא רגרסיה לינארית מרובה עם אינטרקציה.

:R2 .2

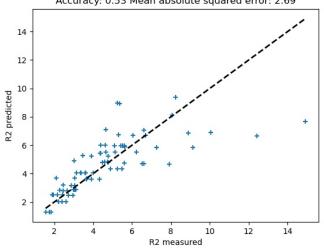
(א) משתנה בלתי תלוי: ריכוז ברזל





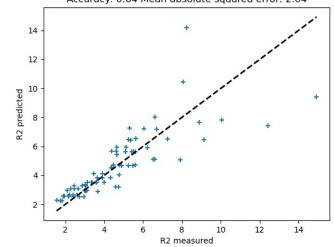


(ג) משתנים בלתי תלויים: ריכוז ברזל, כמות ליפיד R2 measured vs. R2 predicted predictors: iron concentration, lipid amount Accuracy: 0.53 Mean absolute squared error: 2.69



(ד) משתנים בלתי תלויים ריכוז ברזל, כמות ליפיד, ריכוז ברזל * כמות ליפיד

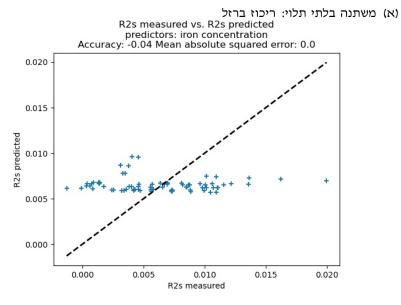
R2 measured vs. R2 predicted predictors: iron concentration, lipid amount, interaction Accuracy: 0.64 Mean absolute squared error: 2.04

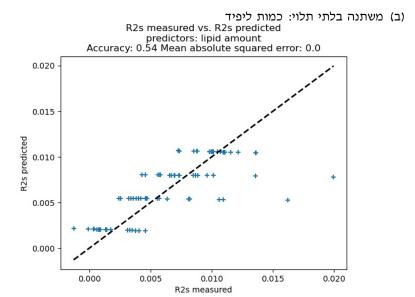


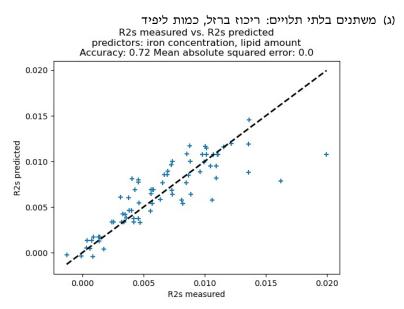
יותר מאשר R2 יותר הברזל ש השפעה לניכוז לעיל כי לעיל מהגרפים מסקנות: ניתן להראות מהגרפים לעיל כי לריכוז הברזל אותר מאשר לכמות הליפיד.

כמו כן, ניתן לראות שחיזוי $\mathrm{R}2$ על ידי ריכוז הברזל וכמות הליפיד מניבה תוצאה $m R2 \quad squared = 0.2$ מול $m R2 \quad squared = 0.53$) טובה יותר מאשר כל אחד לחוד ותוצאה או משתפרת עוד יותר כאשר משתפרת (R2 squared=0.09בין הברזל וכמות הליפיד (R2 =0.64). על פי אלו, ניתן לקבוע כי המודל הלינארי מרובה אינטרקציה. $\mathrm{R}2$ הוא רגרסיה לינארית מרובה עם אינטרקציה.

$:R2_{s}$.3







(ד) משתנים בלתי תלויים ריכוז ברזל, כמות ליפיד, ריכוז ברזל * כמות ליפיד

R2s measured vs. R2s predicted predictors: iron concentration, lipid amount, interaction Accuracy: 0.75 Mean absolute squared error: 0.0

0.020

0.015

0.005

0.000

0.005

0.010

0.015

0.000

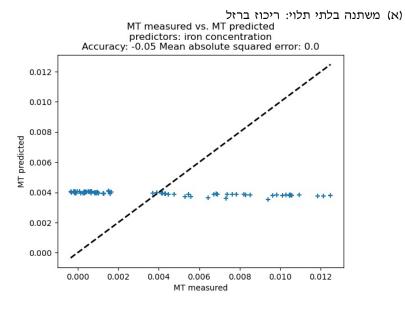
0.015

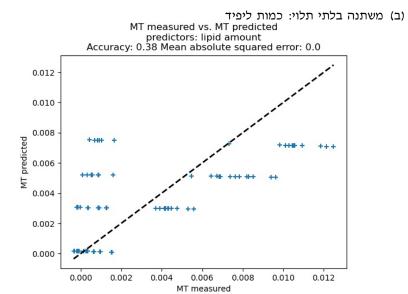
0.020

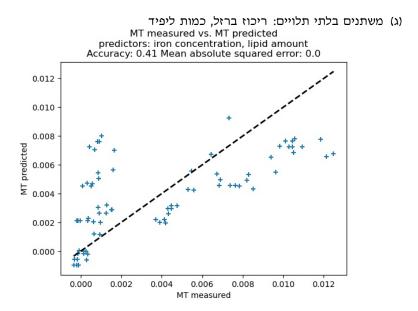
מסקנות: ניתן להראות מהגרפים לעיל כי לכמות הליפיד יש השפעה על R2s מסקנות: מאשר לריכוז הברזל.

כמו כן, ניתן לראות שחיזוי R2s על ידי ריכוז הברזל וכמות הליפיד מניבה תוצאה על העובר אחד הותר מאשר כל אחד לחוד ($R2_\mathrm{squared}{=}0.72$ מול $R2_\mathrm{squared}{=}0.72$ ור טובה יותר מאשר כל אחד לחוד ($R2_\mathrm{squared}{=}0.54$) ותוצאה זו משתפרת עוד יותר כאשר מחשיבים את האינטרקציה בין ריכוז הברזל וכמות הליפיד ($R2_\mathrm{squared}{=}0.75$). על פי אלו, ניתן לקבוע כי R2 הוא רגרסיה לינארית מרובה עם אינטרקציה.

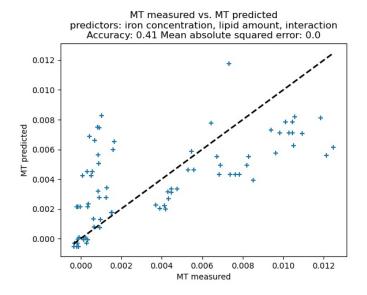
:MT .4







(ד) משתנים בלתי תלויים ריכוז ברזל, כמות ליפיד, ריכוז ברזל * כמות ליפיד

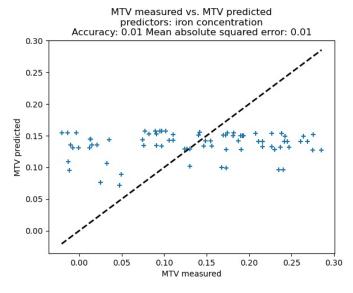


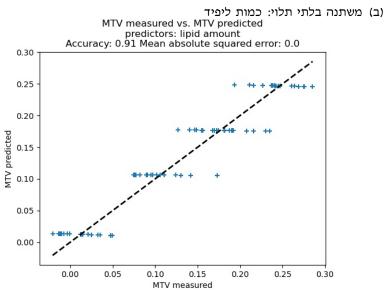
מסקנות: ניתן לראות מהגרפים לעיל כי לכמות הליפיד יש השפעה על MT מסקנות: מאשר לריכוז הברזל.

כמו כן, ניתן לראות שחיזוי $\rm MT$ על ידי ריכוז הברזל וכמות הליפיד מניבה תוצאה $\rm R2_squared=-0.05$ מול $\rm R2_squared=0.41$) טובה יותר מאשר כל אחד לחוד ($\rm R2_squared=0.48$). תוצאה זו נשארת אותו הדבר כאשר מחשיבים את האינטרקציה בין ריכוז הברזל וכמות הליפיד ($\rm R2_squared=0.41$). על פי אלו, ניתן לקבוע כי בין ריכוז הברזל וכמות הליפיד ($\rm MT$ הוא רגרסיה לינארית מרובה המסתמכת על הכרזל וכמות הליפיד.

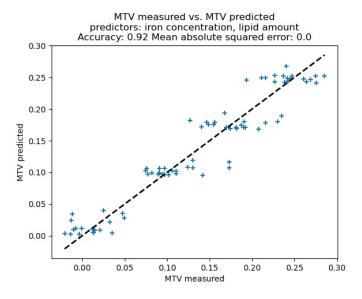
:MTV .5

(א) משתנה בלתי תלוי: ריכוז ברזל

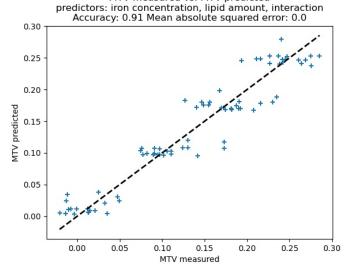




(ג) משתנים בלתי תלויים: ריכוז ברזל, כמות ליפיד



(ד) משתנים בלתי תלויים ריכוז ברזל, כמות ליפיד, ריכוז ברזל * כמות ליפיד MTV measured vs. MTV predicted

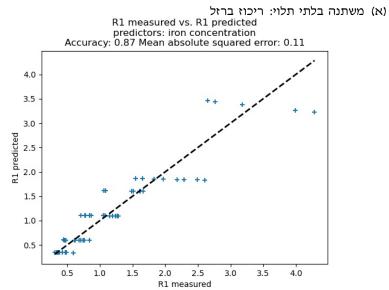


מסקנות: ניתן לראות מהגרפים לעיל כי לכמות הליפיד יש השפעה על MTV מסקנות: מאשר לריכוז הברזל.

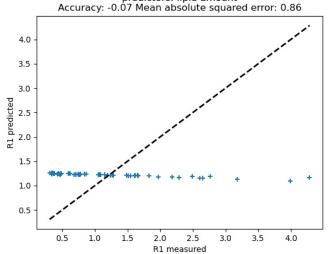
כמו כן, ניתן לראות שחיזוי MTV על ידי ריכוז הברזל וכמות הליפיד מניבה תוצאה טובה יותר מאשר כל אחד לחוד (R2_squared=0.01 מול R2_squared=0.02ו־ טובה יותר מאשר כל אחד לחוד (R2_squared=0.0.91). תוצאה זו נפגעת מעט כאשר מחשיבים את האינטרקציה בין ריכוז הברזל וכמות הליפיד (R2_squared=0.91). על פי אלו, ניתן לקבוע כי המודל הלינארי המתאים לחיזוי MTV הוא רגרסיה לינארית מרובה המסתמכת על ריכוז הברזל וכמות הליפיד.

\mathbf{FE}

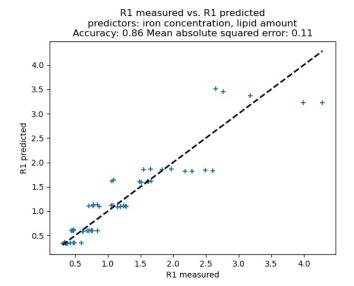
:R1 .1



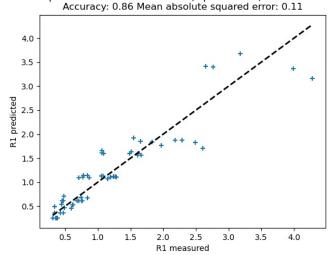
(ב) משתנה בלתי תלוי: כמות ליפיד R1 measured vs. R1 predicted predictors: lipid amount Accuracy: -0.07 Mean absolute squared error: 0.86



(ג) משתנים בלתי תלויים: ריכוז ברזל, כמות ליפיד



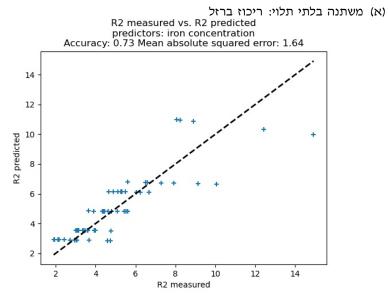
(ד) משתנים בלתי תלויים ריכוז ברזל, כמות ליפיד, ריכוז ברזל * כמות ליפיד R1 measured vs. R1 predicted predictors: iron concentration, lipid amount, interaction

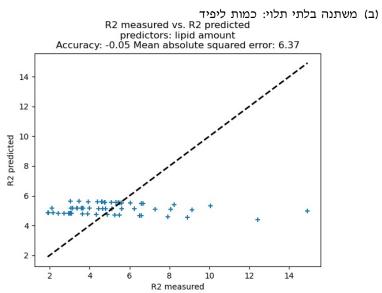


מסקנות: ניתן להראות מהגרפים לעיל כי לריכוז הברזל יש השפעה על m R1 יותר מאשר לכמות הליפיד.

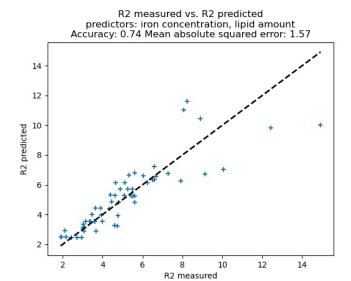
כמו כן, ניתן לראות שחיזוי R1 על ידי ריכוז הברזל מניבה תוצאה טובה יותר מאשר ריכוז הברזל וכמות הליפיד יחד ($R2_\mathrm{squared}{=}0.86$ מול $R2_\mathrm{squared}{=}0.87$ חיכוז הברזל וכמות הליפיד יחד כאשר מחשיבים את האינטרקציה בין ריכוז הברזל ותוצאה זו נשארת אותו הדבר כאשר מחשיבים את האינטרקציה בין ריכוז הברזל וכמות הליפיד. על פי אלו, ניתן לקבוע כי המודל הלינארי המתאים לחיזוי R1 הוא רגרסיה לינארית על פי ריכוז הברזל.

:R2.2

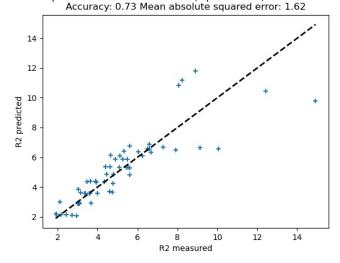




(ג) משתנים בלתי תלויים: ריכוז ברזל, כמות ליפיד



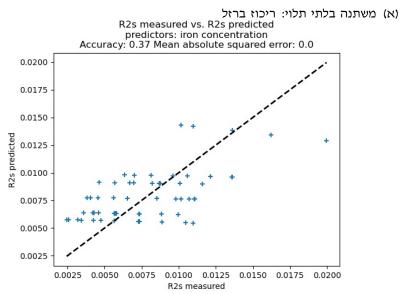
(ד) משתנים בלתי תלויים ריכוז ברזל, כמות ליפיד, ריכוז ברזל * כמות ליפיד R2 measured vs. R2 predicted predictors: iron concentration, lipid amount, interaction

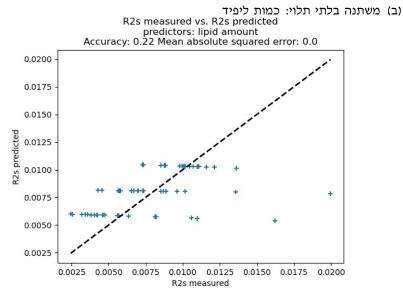


מסקנות: ניתן להראות מהגרפים לעיל כי לריכוז הברזל ש השפעה על R2 יותר מאשר לכמות הליפיד.

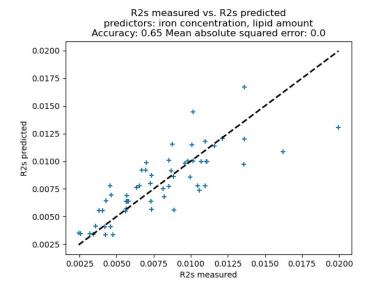
כמו כן, ניתן לראות שחיזוי R2 על ידי ריכוז הברזל וכמות הליפיד מניבה תוצאה $R2_\mathrm{squared}{=}0.73$ מול $R2_\mathrm{squared}{=}0.73$ ור טובה יותר מאשר כל אחד לחוד ($R2_\mathrm{squared}{=}0.74$) מול $R2_\mathrm{squared}{=}0.05$) ותוצאה זו פוחתת מעט כאשר מחשיבים את האינטרקציה בין ריכוז הברזל וכמות הליפיד ($R2_\mathrm{squared}{=}0.73$). על פי אלו, ניתן לקבוע כי המודל הלינארי המתאים לחיזוי R2 הוא רגרסיה לינארית מרובה על פי ריכוז הברזל וכמות הליפיד.

 $:R2_{s}$.3

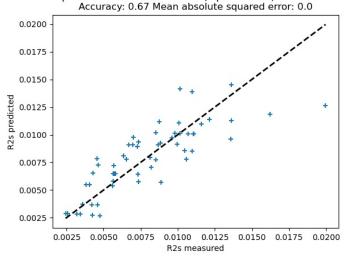




(ג) משתנים בלתי תלויים: ריכוז ברזל, כמות ליפיד



(ד) משתנים בלתי תלויים ריכוז ברזל, כמות ליפיד, ריכוז ברזל * כמות ליפיד R2s measured vs. R2s predicted predictors: iron concentration, lipid amount, interaction



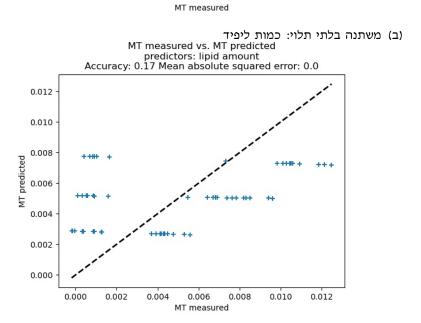
מסקנות: ניתן להראות מהגרפים לעיל כי לריכוז הברזל יש השפעה על m R2s יותר מאשר לכמות הליפיד .

כמו כן, ניתן לראות שחיזוי R2s על ידי ריכוז הברזל וכמות הליפיד מניבה תוצאה טובה יותר מאשר כל אחד לחוד ($R2_\mathrm{squared}{=}0.65$) מול $R2_\mathrm{squared}{=}0.65$) ותוצאה זו משתפרת עוד יותר כאשר מחשיבים את האינטרקציה ($R2_\mathrm{squared}{=}0.22$) ותוצאה זו משתפרת עוד יותר כאשר מהברזל וכמות הליפיד ($R2_\mathrm{squared}{=}0.67$). על פי אלו, ניתן לקבוע כי R2s הוא רגרסיה לינארי מרובה עם אינטרקציה.

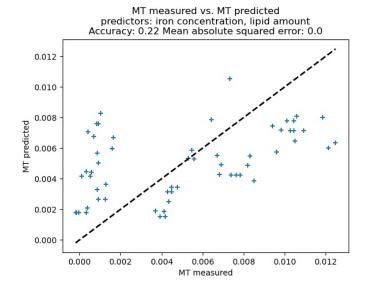
:MT .4

ארנה בלתי תלוי: ריכוז ברזל (א) MT measured vs. MT predicted predictors: iron concentration Accuracy: 0.01 Mean absolute squared error: 0.0

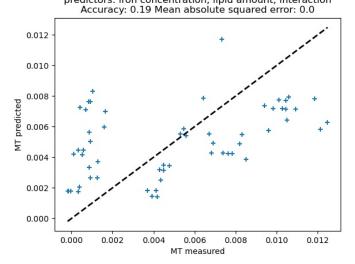
0.012 - 0.006 - 0.006 - 0.006 - 0.006 - 0.008 - 0.010 - 0.012



(ג) משתנים בלתי תלויים: ריכוז ברזל, כמות ליפיד



(ד) משתנים בלתי תלויים ריכוז ברזל, כמות ליפיד, ריכוז ברזל * כמות ליפיד MT measured vs. MT predicted predictors: iron concentration, lipid amount, interaction

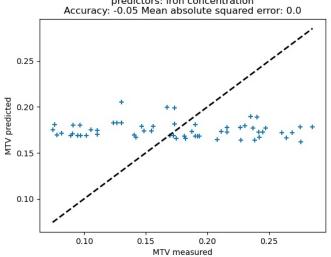


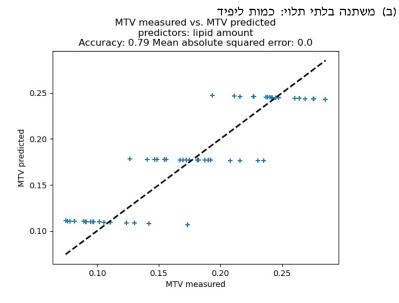
מסקנות: ניתן לראות מהגרפים לעיל כי לכמות הליפיד יש השפעה על MT יותר מאשר לריכוז הברזל.

כמו כן, ניתן לראות שחיזוי MT על ידי ריכוז הברזל וכמות הליפיד מניבה תוצאה $R2_\mathrm{squared}{=}0.01$ מול $R2_\mathrm{squared}{=}0.02$ ו־ טובה יותר מאשר כל אחד לחוד ($R2_\mathrm{squared}{=}0.22$) מול $R2_\mathrm{squared}{=}0.17$). תוצאה זו פוחתת מעט כאשר מחשיבים את האינטרקציה בין ריכוז הברזל וכמות הליפיד ($R2_\mathrm{squared}{=}0.19$). על פי אלו, ניתן לקבוע כי המודל הלינארי המתאים לחיזוי MT הוא רגרסיה לינארית מרובה המסתמכת על ריכוז הברזל וכמות הליפיד.

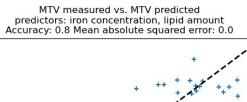
:MTV .5

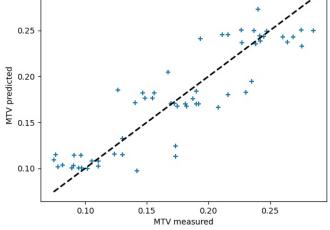
(א) משתנה בלתי תלוי: ריכוז ברזל MTV measured vs. MTV predicted predictors: iron concentration Accuracy: -0.05 Mean absolute squared error: 0.0



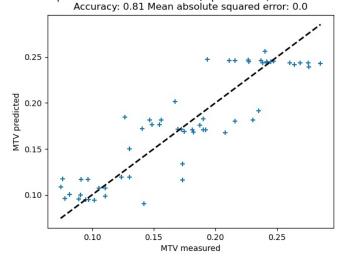


(ג) משתנים בלתי תלויים: ריכוז ברזל, כמות ליפיד





(ד) משתנים בלתי תלויים ריכוז ברזל, כמות ליפיד, ריכוז ברזל * כמות ליפיד MTV measured vs. MTV predicted predictors: iron concentration, lipid amount, interaction

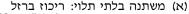


מסקנות: ניתן לראות מהגרפים לעיל כי לכמות הליפיד יש השפעה על MTV יותר מאשר לריכוז הברזל.

כמו כן, ניתן לראות שחיזוי MTV על ידי ריכוז הברזל וכמות הליפיד מניבה תוצאה $m R2 \;\; squared = -0.05$ מול R2 $m R2 \;\; squared = 0.8$ טובה יותר מאשר כל אחד לחוד תוצאה או משתפרת מעט כאשר מחשיבים את תוצאה ו תוצאה . (R2 m squared = 0.79בין הברזל וכמות הליפיד (R2 =0.81). על פי אלו, ניתן לקבוע כי המודל הלינארי המתאים לחיזוי MTV הוא רגרסיה לינארית מרובה המסתמכת על ריכוז הברזל וכמות הליפיד עם אינטרקציה ביניהם.

Transferrin Ferritin

:R1 .1

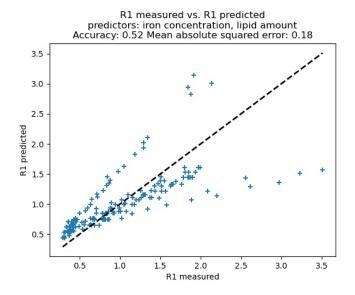


(א) משתנה בלתי תלוי: ריכוז ברזל R1 measured vs. R1 predicted predictors: iron concentration Accuracy: 0.5 Mean absolute squared error: 0.19 3.5 3.0 2.5 R1 predicted 1.0 0.5 0.5 1.0 1.5 2.0 2.5 3.0 3.5

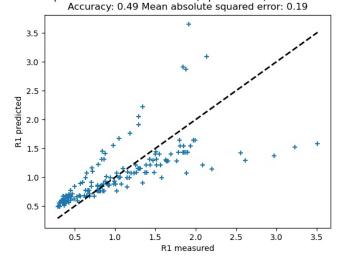
R1 measured

(ב) משתנה בלתי תלוי: כמות ליפיד R1 measured vs. R1 predicted predictors: lipid amount Accuracy: -0.01 Mean absolute squared error: 0.38 3.5 3.0 2.5 R1 predicted 1.0 0.5 0.5 1.0 1.5 2.0 2.5 3.0 3.5 R1 measured

(ג) משתנים בלתי תלויים: ריכוז ברזל, כמות ליפיד



(ד) משתנים בלתי תלויים ריכוז ברזל, כמות ליפיד, ריכוז ברזל * כמות ליפיד R1 measured vs. R1 predicted predictors: iron concentration, lipid amount, interaction



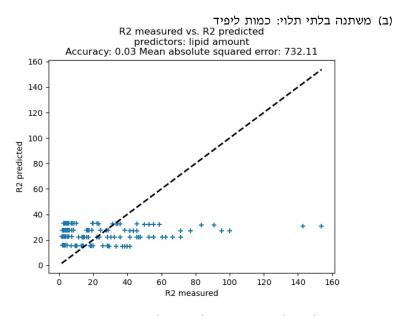
מסקנות: ניתן להראות מהגרפים לעיל כי לריכוז הברזל יש השפעה על R1 יותר מאשר לכמות הליפיד.

כמו כן, ניתן לראות שחיזוי R1 על ידי ריכוז הברזל וכמות הליפיד מניבה תוצאה אותר מאשר כל אחד לחוד ($R2_\mathrm{squared}{=}0.52$) מול $R2_\mathrm{squared}{=}0.52$ ו־ טובה יותר מאשר כל אחד לחוד ($R2_\mathrm{squared}{=}-0.01$) ותוצאה זו פוחתת מעט כאשר מחשיבים את האינטרקציה בין ריכוז הברזל וכמות הליפיד ($R2_\mathrm{squared}{=}0.49$). על פי אלו, ניתן לקבוע כי המודל הלינארי המתאים לחיזוי R1 הוא רגרסיה לינארית מרובה המסתמכת על ריכוז הברזל וכמות הליפיד.

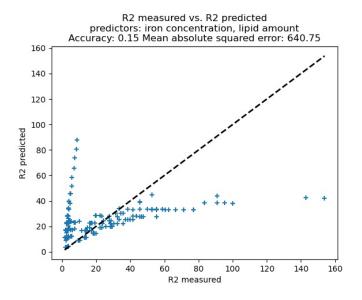
:R2 .2

(א) משתנה בלתי תלוי: ריכוז ברזל R2 measured vs. R2 predicted predictors: iron concentration Accuracy: 0.1 Mean absolute squared error: 678.02 R2 predicted

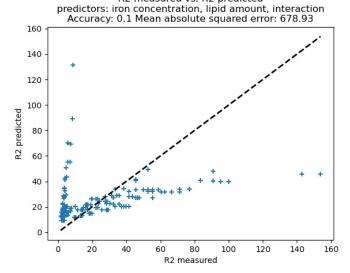
R2 measured



(ג) משתנים בלתי תלויים: ריכוז ברזל, כמות ליפיד



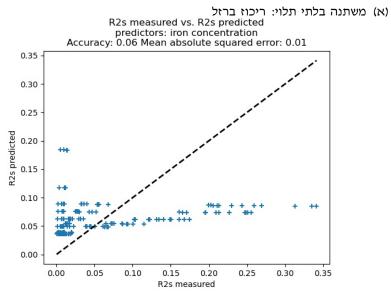
(ד) משתנים בלתי תלויים ריכוז ברזל, כמות ליפיד, ריכוז ברזל * כמות ליפיד R2 measured vs. R2 predicted

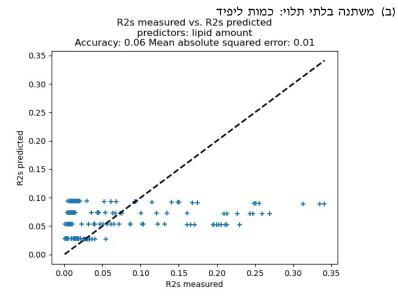


מסקנות: ניתן להראות מהגרפים לעיל כי לריכוז הברזל יש השפעה על $\mathrm{R}2$ יותר מאשר לכמות הליפיד.

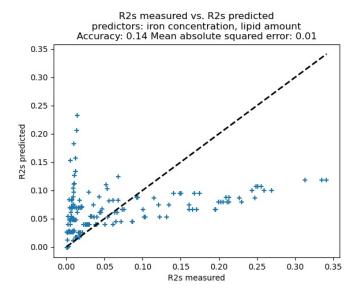
כמו כן, ניתן לראות שחיזוי R2 על ידי ריכוז הברזל וכמות הליפיד מניבה תוצאה $R2_\mathrm{squared} = 0.15$ מול $R2_\mathrm{squared} = 0.15$ ור טובה יותר מאשר כל אחד לחוד ($R2_\mathrm{squared} = 0.03$) ותוצאה זו פוחתת מעט כאשר מחשיבים את האינטרקציה בין ריכוז הברזל וכמות הליפיד ($R2_\mathrm{squared} = 0.03$). על פי אלו, ניתן לקבוע כי המודל הלינארי המתאים לחיזוי R2 הוא רגרסיה לינארית מרובה המסתמכת על ריכוז הברזל וכמות הליפיד.

 $:R2_{s}$.3

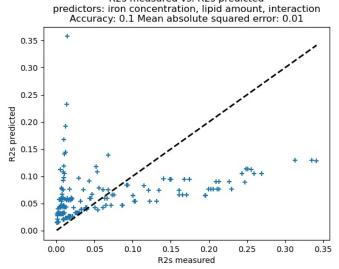




(ג) משתנים בלתי תלויים: ריכוז ברזל, כמות ליפיד



(ד) משתנים בלתי תלויים ריכוז ברזל, כמות ליפיד, ריכוז ברזל * כמות ליפיד R2s measured vs. R2s predicted predictors: iron concentration, lipid amount, interaction



m R2s מסקנות: ניתן להראות מהגרפים לעיל כי לכמות הליפיד יש השפעה דומה על כמו לריכוז הברזל.

כמו כן, ניתן לראות שחיזוי R2s על ידי ריכוז הברזל וכמות הליפיד מניבה תוצאה טובה יותר מאשר כל אחד לחוד ($R2_squared=0.06$ מול ו־ $R2_squared=0.14$) ותוצאה זו פוחתת מעט כאשר מחשיבים את האינטרקציה בין ריכוז הברזל וכמות הליפיד ($R2_squared=0.1$). על פי אלו, ניתן לקבוע כי המודל הלינארי המתאים לחיזוי $R2s_squared=0.1$ הוא רגרסיה לינארית מרובה המסתמכת על ריכוז הברזל וכמות הליפיד.

:MT .4

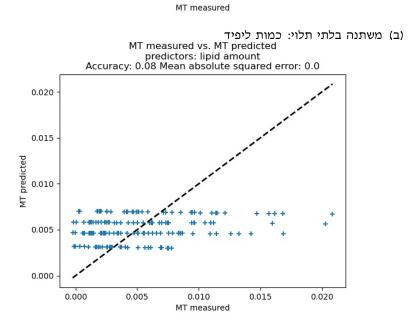
0.010

0.015

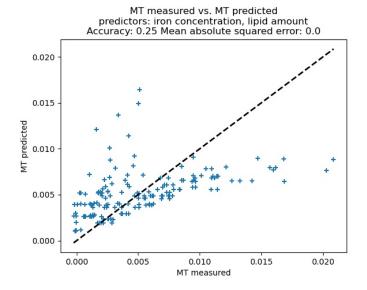
0.020

0.000

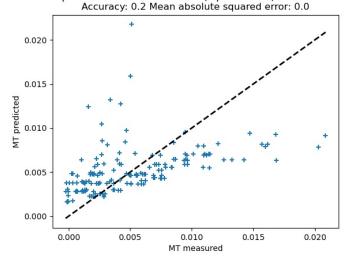
0.005



(ג) משתנים בלתי תלויים: ריכוז ברזל, כמות ליפיד



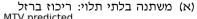
(ד) משתנים בלתי תלויים ריכוז ברזל, כמות ליפיד, ריכוז ברזל * כמות ליפיד MT measured vs. MT predicted predictors: iron concentration, lipid amount, interaction

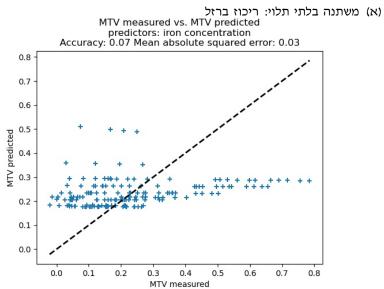


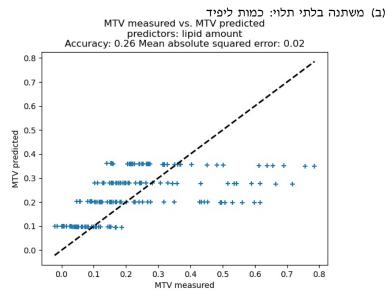
מסקנות: ניתן להראות מהגרפים לעיל כי לריכוז הברזל יש השפעה על $\mathrm{R}1$ יותר מאשר לכמות הליפיד.

כמו כן, ניתן לראות שחיזוי R1 על ידי ריכוז הברזל וכמות הליפיד מניבה תוצאה $R2_\mathrm{squared}{=}0.14$ טובה יותר מאשר כל אחד לחוד ($R2_\mathrm{squared}{=}0.25$) מול $R2_\mathrm{squared}{=}0.08$ ($R2_\mathrm{squared}{=}0.08$) ותוצאה זו פוחתת מעט כאשר מחשיבים את האינטרקציה בין ריכוז הברזל וכמות הליפיד ($R2_\mathrm{squared}{=}0.2$). על פי אלו, ניתן לקבוע כי המודל הלינארי המתאים לחיזוי MT הוא רגרסיה לינארית מרובה המסתמכת על ריכוז הברזל וכמות הליפיד.

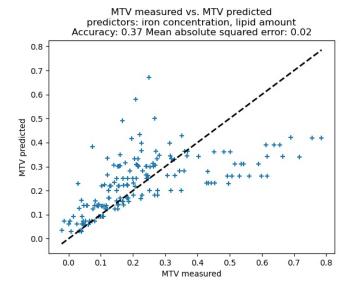
:MTV .5



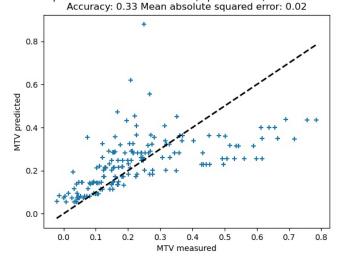




(ג) משתנים בלתי תלויים: ריכוז ברזל, כמות ליפיד



(ד) משתנים בלתי תלויים ריכוז ברזל, כמות ליפיד, ריכוז ברזל * כמות ליפיד MTV measured vs. MTV predicted predictors: iron concentration, lipid amount, interaction



מסקנות: ניתן לראות מהגרפים לעיל כי לכמות הליפיד יש השפעה על MTV מסקנות: מאשר לריכוז הברזל.

כמו כן, ניתן לראות שחיזוי MTV על ידי ריכוז הברזל וכמות הליפיד מניבה תוצאה עובה יותר מאשר כל אחד לחוד $R2_squared=0.07$ מול $R2_squared=0.37$ ור טובה יותר מאשר כל אחד לחוד ($R2_squared=0.26$). תוצאה זו נפגעת מעט כאשר מחשיבים את האינטרקציה בין ריכוז הברזל וכמות הליפיד ($R2_squared=0.33$). על פי אלו, ניתן לקבוע כי המודל הלינארי המתאים לחיזוי MTV הוא רגרסיה לינארית מרובה המסתמכת על ריכוז הברזל וכמות הליפיד.