**למידת מכונה- פרויקט סיום**

**מבוא**

בעבודה זו בנינו מודל למידת מכונה המיועד לאפשר לרופאים לבצע פרדיקציה מדויקת בנוגע לגמילת מונשמים ממכונת הנשמה. נבדקים יכולים להיות מחוברים למכונת הנשמה מסיבות שונות, לדוגמה- מחלת ריאות, פגיעת ראש, או פציעות שונות. מטופלים שמונשמים במכונת ההנשמה מקבלים עזרה מכנית כדי לאפשר לריאות שלהם לעבוד ולחמצן להיכנס לגוף. עם זאת, חיבור מתמשך למכונות ההנשמה עלול להוביל לסיבוכים רפואיים שונים ולפגוע באיכות חיי המטופל. תהליך הגמילה ממכונת הנשמה הינו תהליך מורכב, הכולל ירידה הדרגתית מהתלות במכונות ההנשמה(Boles et al., 2007). ההחלטה לגבי התחלת גמילה מהנשמה תלויה ברופאים במחלקות, החלטה זו מתקבלת על בסיס ידע רפואי, אינטואיציה, ופרוטוקולים מובנים. עם זאת, ישנה שונות גבוהה בהצלחת הגמילה בין רופאים שונים. כלומר ההחלטה את מי מהמטופלים לגמול ומתי אינה עקבית בקרב הרופאים (בהתבסס על ראיונות עם הנהלת בית חולים שוהם) . על מנת לעזור לרופאים לקבל החלטה מושכלת בנינו מודל למידת מכונה החוזה את סיכויי הגמילה של כל מטופל ממכונת ההנשמה בהתבסס על ידע קודם. כדי שהרופאים ישתמשו במודל שלנו על המודל להיות הסברי, כלומר לאפשר לרופאים להבין את תהליך הפרדיקציה. כדי לאפשר זאת צמצמנו את מספר המנבאים במודל, והסברנו את חשיבות המנבאים בתוכו (Tonekaboni et al., 2019).

**שאלות המחקר**

**1.**  מהו מודל למידת המכונה ההסברי, שיצליח לנבא בצורה המדויקת ביותר את סיכויי הגמילה של המטופלים ממכונות ההנשמה.

**2.**  מהם המדדים המשפיעים ביותר במודל שנבחר.

**שיטה** סט הנתונים עליו נבצע את למידת המכונה הופק באמצעות mimic4 סט נתונים המכיל מידע רפואי מבית חולים בארה"ב ופתוח בצורה מקוונת, פרטים מסוימים בו שונו מטעמי אתיקה. מתוך mimic4 הפקנו את המדדים שנמצאו אינדיקטיבים לחיזוי סיכויי גמילה בספרות ובאמצעות ראיונות עם רופאים. בניתוח מקדים תייגנו את הנבדקים ל2 קטגוריות של הצלחה: נבדקים שהצליחו להיגמל ונבדקים שלא הצליחו להיגמל. התיוג נעשה גם כן בייעוץ רפואי, נבדקים שהצליחו לנשום יותר מ48 ללא מכונת הנשמה, ושלא מתו ב3 ימים לאחר הניתוק תויגו כנבדקים שהצליחו להיגמל ממכונת ההנשמה. סט הנתונים עליו ביצענו את למידת המכונה כלל נתונים מ15,999 מטופלים, ללא נתונים חסרים,  54 פיצרים  ותיוג ל2 קבוצות- הצלחה בגמילה וחוסר הצלחה בגמילה. בתור התחלה חילקנו את הנתונים שלנו לtest ולtrain על מנת לאפשר השוואה בין מודלים בהתבסס על מדדי הביצוע שלהם. סט נתוני האימון כלל 80 אחוז מהתצפיות. (12,800 תצפיות). סט הנתונים שלנו לא מאוזן מבחינת כמות הנבדקים שנגמלו ושלא נגמלו, כמות הנגמלים לפי סט האימון היה 0.88%. ראשית כל החלטנו לצמצם את מספר המנבאים שלנו כדי לקבל פרדיקציה הסברית יותר, ולהקל על הכח החישובי שיידרש בבניית המודלים. לשם כך בנינו elastic net המאפשרת את צמצום מספר המנבאים על ידי ניפוי מנבאים שאינם מוסיפים לניבוי. מטרת צמצום המנבאים היתה להגביר את הסבריות המודל, על אף שלפי יחס המנבאים תצפיות לא היה בכך צורך. כדי ליצור רשת אלסטית אופטימלית ביצענו cross validation באמצעות k-fold כדי למצוא את הרשת שתאפשר גמישות מבלי לפגוע בניבוי. לאחר שמצאנו את הרשת האופטימלית ביצענו ניבוי באמצעותה והוצאנו פיצ'רים שלא תרמו לניבוי מנתוני האימון.  
לאחר מכן יצרנו 3 מודלים של למידת מכונה. הקפדנו שהמודלים שאותם אנו יוצרים הינם מודלים שניתנים להבנה כדי שלרופאים יהיה קל יותר להאמין להם (וכך להימנע מקופסה שחורה ככל הניתן)

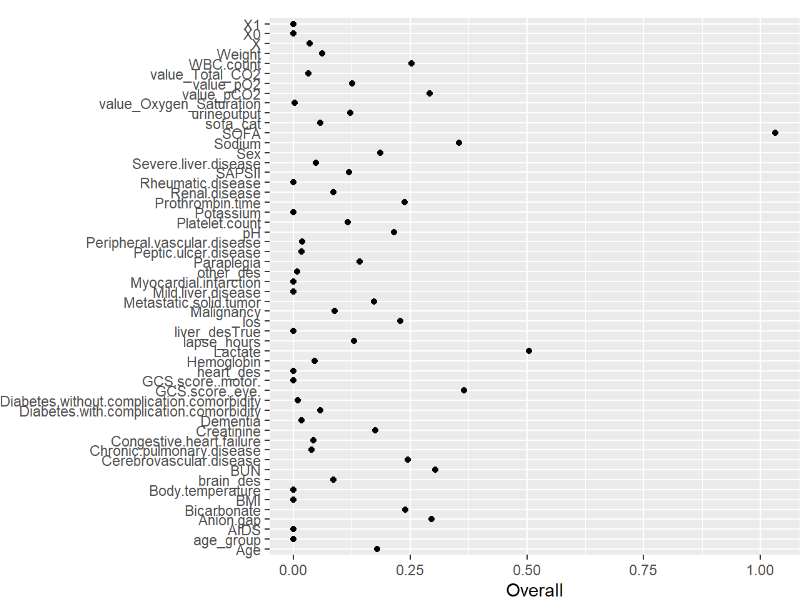
-        KNN -המתייג את הנתונים על פי הקרבה שלהם לתצפיות אחריות ליצירת אזורי תיוג.

-        Random Forest-יוצר אלפי עצי החלטה, ויוצר בניהם ממוצע ליצירת עץ החלטה אופטימלי

-        Logistics Regression- יוצר קו רגרסיה המסווגת את הנתונים ל2 קטגוריות.  
עבור שני המודלים הראשונים ביצענו cross validation בשיטת 10-fold, כדי למצוא את ההיפר פרמטרים עם הביצועים הטובים ביותר. לאחר מכן השווינו בין ביצועי המודלים על פי Accuracy שמראה כמה מהפרדיקציות סהכ היה נכונות, Balanced Accuracy שמראה את אחוזי הדיוק הממוצעים עבור כל קטגוריה (הצליח או לא הצליח בגמילה), כלומר מתחשב בהטייה הקיימת בנתונים עצמם, sensitivity שמראה כמה מתוך אלו שלא הצליחו להיגמל הצלחנו לנבא. וAUC שמראה את השטח תחת עקומת הROC.  לבסוף, בחנו את חשיבות הפיצ'רים בכל מודל כדי להבין מהם המדדים האינפורמטיביים ביותר עבור המודל הנבחר על מנת להגביר את ההסבריות שלו.

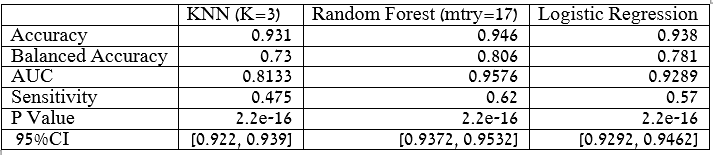
**תוצאות**הרשת האלסטית נמצאה כאופטימלית עבור ערכי coefficient שנמצאו על פי הרשת מוצגים בגרף 1, ניתן לראות כי ישנם 13 מנבאים עם  coefficient אפסי שהוצאו מהמודל. [גרף 1]

עבור מודל הKNN ערך K האופטימלי שנמצא הוא 3, עבור מודל הRF כמות הפרמטרים האופטימלית בכל פיצול הוא 17.

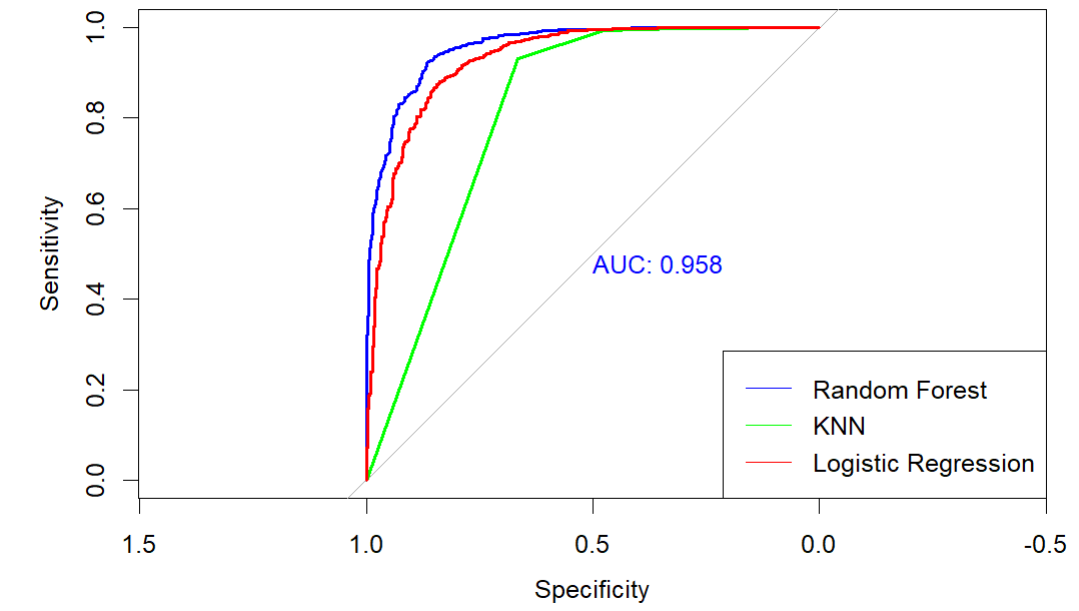


**גרף 1**. Coefficient של כל מדד על פי הelastic net

מדדי הביצוע עבור כל מודל מפורטים בטבלה 1 ועקומות הROC בגרף 2:



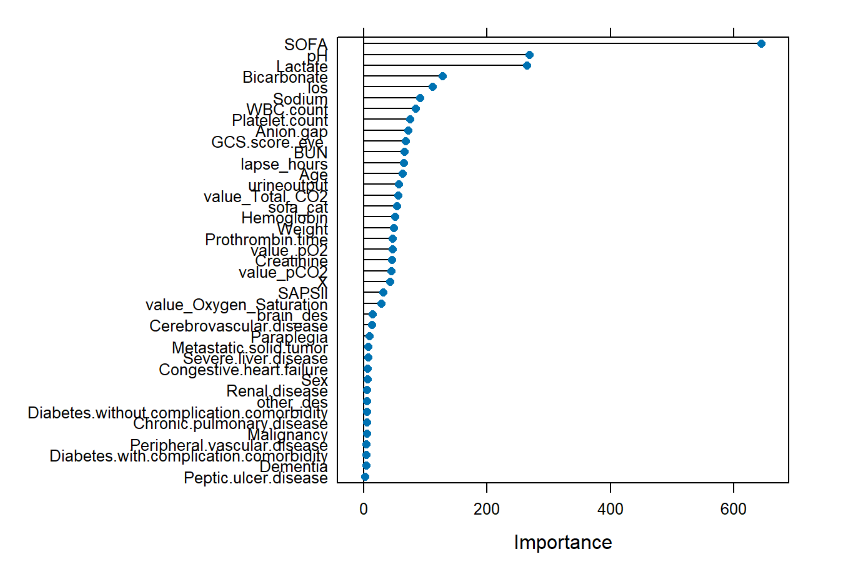
**טבלה 1**. מדדי הביצוע של כל מודל, כולל מובהקות ורווח סמך



**גרף 2**. השוואת עקומות הROC של המודלים השונים

כפי שניתן לראות ביצועי מודל הRandom Forest ביצע תחזיות מדויקות יותר בכל 3 המדדים שבחרנו לבחון.   
חשיבות הפרמטרים עבור המודל הנבחר (Random Forest) מפורטים בגרף 3.

**מסקנות ודיון:**בעבודה זו ביצענו ניבוי עבור נתונים אמיתיים של מונשמים במכונות ההנשמה.ביצענו שיטות של צמצום מנבאים, בחינת היפר פרמטרים אופטימאליים והשוואת מודלים כדי לבצע פרדיקציה מדויקת ככל הניתן למונשמים ממכונות  הנשמה. לאחר מכן בדקנו את חשיבות הפרמטרים כדי לבחור את הפרמטרים האינדקטיביים ביותר על מנת להגביר את ההסבריות של המודל הנבחר. המודל שביצע את הפרדיקציה המדויקת ביותר היה מודל Random Forest, עם מספר פיצולים של 17 בכל פיצול של העץ. מתוך סך התשובות שהמודל הפיק 95% מהתחזיות של המודל היו מדויקות (Accuracy). תחזיות מדויקות אלו יכולות לסייע לצוות הרפואי בבתי החולים לקבל החלטות נבונות יותר לגבי תחילת תהליך גמילה עם מטופלים, מה שיאפשר הקצאת משאבים נכונה יותר בבי"ח, ורפואה טובה יותר למטופלים המונשמים. כמו כן נמצא שארבעת המדדים האינדיקטיבים ביותר לתחזית זו הם: sofa, ph, lactate, bicarbonate

**גרף 3**. חשיבות המנבאים לפי מודל Random Forest

**ביבליוגרפיה**

Boles, J. M., Bion, J., Connors, A., Herridge, M., Marsh, B., Melot, C., ... & Welte, T. (2007). Weaning from mechanical ventilation. European Respiratory Journal, 29(5), 1033-1056.‏

Tonekaboni, S., Joshi, S., McCradden, M. D., & Goldenberg, A. (2019, October). What clinicians want: contextualizing explainable machine learning for clinical end use. In *Machine learning for healthcare conference* (pp. 359-380). PMLR.‏