



המחלקה להנדסת תעשייה וניהול

הקשר בין מדדי דופק המופקים מרצועת חזה לעומס קוגניטיבי

---

The relationship between pulse indices extracted  
from chest strap and mental workload

---



פרויקט גמר – דו"ח מסכם

סמסטר ב' תשפ"א

7-2021

צוות הפרויקט: עומרי זאבי, שיר רוזנס ודור אלדר  
הנחיה אקדמית: פרופ' טל אורון גלעד, ד"ר יובל זק

## תקציר

משימות רבות דורשות מאמץ קוגניטיבי מאלו המבצעים אותן, בין אם בנהיגה, בקרת תעופה, ניתוח של חולים וכו'. לאורך המשימה, על המבצע להשתמש בידע רלוונטי, להיות מרוכז ולקבל החלטות בזמן אמת. לכן, יש חשיבות רבה לניטור רמת העומס הקוגניטיבי, וזאת על מנת לא לפגוע בביצועי המשימה. מדידת העומס הקוגניטיבי היא משימה מורכבת בפני עצמה, מכיוון שלא ניתן להבחין בעומס באופן ישיר. לשם כך, נעשה שימוש באמצעים למדידה פיזיולוגית אובייקטיבית כגון אק"ג ו-אא"ג, הבודקים מדדים של נשימה, קצב לב, מוליכות חשמלית של העור ועוד. ישנן לא מעט בעיות בשימוש בטכניקות המדידה הללו: ציוד המדידה יקר ולא נגיש לרוב האוכלוסייה, השימוש עצמו מסורבל ולא נוח. בנוסף, קיימים שאלונים ייעודיים שמטרתם להעריך את העומס הקוגניטיבי הנתפס באופן סובייקטיבי. שילוב המדדים האובייקטיביים והסובייקטיביים הללו מייצר תמונה מדויקת ככל הניתן של רמת העומס הקוגניטיבי אותו חש המפעיל, אך ניתן לאמוד אותה רק בסיום המשימה, ולא בעת ביצועה.

שאלת המחקר שלנו עסקה במציאת הקשר בין מדדי הדופק המופקים מרצועת חזה לבין רמת העומס הקוגניטיבי. בהתאם לכך, הניסוי שבוצע במסגרת הפרויקט נועד למדוד עומס קוגניטיבי בשלוש רמות עומס שונות. הוא התפרש על פני שלושה ימים, כך שבמהלכו בכל יום הנבדק ביצע שלוש משימות ואורכה של כל משימה הוא 30 דקות. המשימות חולקו בצורה אקראית כדי למנוע את השפעת השעה ביום בה בוצעה המשימה. בכל משימה היה על הנבדק ללבוש את רצועת החזה מדגם Polar H10 ולתעד באמצעות אפליקציה ייעודית את דופק הלב שלו. במקביל לכך, כל נבדק התבקש למלא שני שאלונים להערכת העומס הנתפס - שאלון SWAT כל 5 דקות בעת ביצוע המשימה ושאלון NASA – TLX פעם אחת בסיומה. באמצעות הנתונים שהופקו מרצועת החזה חושבו מדדי המחקר האובייקטיביים, ומנתוני השאלונים הופקו המדדים הסובייקטיביים. בניסוי השתתפו תשעה נסיינים, המורכבים מהחוקרים ומבני משפחתם.

המחקר כולל ארבע השערות, כך שניתוח הנתונים נועד להתייחס לכל אחת מהן. ההשערה הראשונה היא שקיים קשר בין מדדי דופק הלב הנמדדים מרצועת חזה לבין העומס הקוגניטיבי. תוצאות הניסוי הראו כי קיים מתאם ספירמן מובהק אך לא חזק בין המדדים שנבדקו לעומס הקוגניטיבי. ההשערה השנייה היא ששונות קצב הלב גדלה ככל שרמת קושי המשימה עולה. הממצאים הראו באופן מובהק באמצעות מבחן ANOVA כי מדד שונות קצב הלב (HRV) דומה בין משימות בעלות עומס שונה. ההשערה השלישית היא שהשעה ביום אינה משפיעה על מדדי דופק הלב ואכן נמצא כי ברוב המדדים אין קורלציה מובהקת ע"פ מבחן ANOVA. ההשערה הרביעית היא שניתן לייצר מודל באמצעות אלגוריתם לימוד מכונה שידע לסווג את

רמת העומס של המשימה ל-3 קטגוריות. במחקר נבחנה האפשרות למודל אחד שהובילה לאחוזי ניבוי נמוכים ולמודל פרטני עבור כל נסיון. נבדקו שני מודלים של לימוד מכונה, ונמצא כי מודל Random Forest מנבא בצורה טובה את העומס הנמוך, אך קיים דמיון משמעותי בין נתוני העומס הגבוה והבינוני, לכן איחדנו נתונים אלה. יצרנו מודל חדש שמצליח לנבא בצורה טובה יותר, הכולל שתי קטגוריות – עומס נמוך ועומס בינוני-גבוה.

לסיכום, אנו רואים חשיבות ופוטנציאל רב ביכולת סיווג עומס קוגניטיבי בזמן אמת באמצעות כלים פשוטים כגון רצועת החזה Polar H10. המלצתנו היא לפתח מודל שונה, קולקטיבי, בעל התאמות מינימליות עבור כל משתמש, מבוסס דאטה גדול יותר בצורה משמעותית, שיצור סיווג חד משמעי של העומס הקוגניטיבי.

**מילות מפתח: עומס קוגניטיבי, מודל לימוד מכונה, שונות קצב לב, מדדי דופק**

## **Abstract**

Many tasks require cognitive effort from those who perform them, whether in driving, aviation control, patient surgery, and so on. Throughout the task, the performer must use relevant knowledge, be focused and make decisions in real time. Therefore, it is very important to monitor the level of cognitive load, in order not to impair the performance of the task. Cognitive load measurement is a complex task in itself, as the load cannot be discerned directly. For this purpose, means for objective physiological measurement such as ECG and EEG, which gauge measures of respiration, heart rate, electrical conductivity of the skin and more are used. There are quite a few problems with the use of these measurement techniques: the measuring equipment is expensive and inaccessible to the majority of the population, the use itself is cumbersome and inconvenient. In addition, there are specific questionnaires that aim to assess the perceived cognitive load subjectively. The combination of these objective and subjective indices produces as accurate a picture as possible of the level of cognitive load that the operator feels, but it can only be estimated at the end of the task, and not when it is performed.

Our research question dealt with finding the relationship between the heart rate indicators extracted from a chest strap and the level of cognitive load.

Accordingly, the experiment performed as part of the project was designed to measure cognitive load at three different load levels. It was spread out over three days, so that during each day the subject performed three tasks that lasted 30 minutes each. Tasks were randomly assigned to avoid the effect of the time of day the task was performed. In each task, the subject had to wear the Polar H10 chest strap and record his heart rate using a designated app. At the same time, each subject was asked to complete two questionnaires to assess the perceived load - a SWAT questionnaire every 5 minutes while performing the task and a NASA - TLX questionnaire once at the end. Using the data extracted from the chest strap, the objective research indices were calculated, and the subjective indices were produced from the questionnaire data. Nine participants took part in the experiment, consisting of the researchers and their family members.

The study includes four hypotheses, so the data analysis is designed to address each of them. The first hypothesis is that there is a relationship between the heart rate indices measured from the chest strap and the cognitive load. The results of the experiment showed that there is a significant but not strong Spearman correlation between the measures tested and the cognitive load. The second hypothesis is that heart rate variability increases as the level of task difficulty increases. The findings clearly showed using the ANOVA test that the heart rate variability index (HRV) is similar between tasks with different load. The third hypothesis is that the time of day does not affect the heart rate indices and indeed it was found that in most indices there is no significant correlation according to the ANOVA test. The fourth hypothesis is that a model can be produced using a machine learning algorithm that can classify the load level of the task into 3 categories. The study examined the possibility of a unified model that led to low prediction percentages and an individual model for each participant. Two machine learning models were tested, and it was found that the Random Forest model predicts the low load well, but there is a significant similarity between the high and medium load data, so we combined these data. We have created a new model that manages to predict better, which includes two categories - low load and medium-high load.

In conclusion, we see great importance and potential in the ability to classify real-time cognitive load using simple tools such as the Polar H10 chest strap. Our recommendation is to develop a different, collective model, with minimal adjustments for each user, based on significantly larger data, which will create an unambiguous classification of the cognitive load.

**Keywords:** cognitive load, machine learning model, heart rate variability, heart rate indices

## **תודות**

בפרויקט זה הייתה לנו ההזדמנות לבחון וליישם את הידע הרחב והמגוון אותו רכשנו במהלך לימודי התואר הראשון בפקולטה להנדסת תעשייה וניהול באוניברסיטת בן-גוריון. נרצה להודות למנחים שלנו אשר תמכו, סייעו ועזרו לנו לכל אורך הדרך: למנחה האקדמית שלנו, פרופסור טל אורון-גלעד – על הנחיית הפרויקט והעזרה המקצועית לאורך הדרך. למר יובל זאק - על העזרה והסיוע הרב אשר ניתן לנו בכל דבר לו נדרשנו, על הליווי וההכוונה, על הסבלנות והשקעת הזמן.

## תוכן עניינים

10.....	1. מבוא.....
12.....	2. סקירת ספרות.....
12.....	2.1 עומס קוגניטיבי.....
12.....	2.1.1 גורמים המשפיעים על עומס עבודה.....
13.....	2.1.2 חשיבות המדידה של עומס עבודה קוגניטיבי.....
14.....	2.2 מדדים לעומס קוגניטיבי.....
15.....	2.2.1 מדדים פיזיולוגיים לעומס קוגניטיבי.....
17.....	2.3 מדידת עומס עבודה קוגניטיבי בפעילות יומימית רציפה.....
18.....	2.3.1 השפעת השעה ביום על ביצועים.....
19.....	2.4 לימוד מכונה (Machine Learning).....
20.....	2.4.1 סיווג אלגוריתמים בלימוד מכונה.....
20.....	2.4.2 אלגוריתמים של לימוד מכונה.....
21.....	2.4.3 מתודולוגיית העבודה בלימוד מכונה.....
22.....	2.4.4 לימוד מכונה ומדדים פיזיולוגיים.....
24.....	3. השערות המחקר.....
24.....	4. שיטה.....
24.....	4.1 תיאור כללי.....
26.....	4.2 נבדקים.....
27.....	4.3 כלים.....
27.....	4.3.1 רצועת החזה Polar H10.....
28.....	4.4 סביבת הניסוי.....
28.....	4.5 מהלך הניסוי.....
28.....	4.6 מדדים.....
28.....	4.6.1 מדדים אובייקטיביים.....
29.....	4.6.2 מדדים סובייקטיביים.....
30.....	4.7 ניתוח תוצאות.....
30.....	4.7.1 אופן ניתוח השאלונים.....
30.....	4.7.2 תוצאות ומסקנות הפיילוט.....
31.....	4.7.3 עיבוד הנתונים.....
32.....	5. תוצאות.....
39.....	6. דיון ומסקנות.....
41.....	7. סיכום והמלצות.....
43.....	8. ביבליוגרפיה:.....
51.....	9. נספחים:.....
51.....	נספח 1 - הוראות ביצוע הניסוי.....

52.....	נספח 2 – טופס הסכמה להשתתפות בניסוי
54.....	נספח 3 – שאלון NASA – TLX
55.....	נספח 4 – שאלון SWAT
57.....	נספח 5 – טווח מדדים מחושבים
58.....	נספח 6 – תרשימי קופסא עבור מדדי המודל לכלל הנבדקים
61.....	נספח 7 – טבלת ממוצע וסטיות תקן של מדדי הדופק
62.....	נספח 8 - חישוב מדדים אובייקטיביים
65.....	נספח 9 - בדיקת התפלגות נורמלית של המדדים
66.....	נספח 10 – תרשימי קופסא של מדדים מול פרקי זמן ביום
68.....	נספח 11 - ניתוח תוצאות אלגוריתם Random Forest
	11.1 תוצאות מובהקות המדדים לפי פונקציית feature importance - לפני
68.....	טרנספורמציה
	11.2 תוצאות מובהקות המדדים לפי פונקציית feature importance - אחרי
68.....	טרנספורמציה
	11.3 תוצאות מובהקות המדדים לפי פונקציית feature importance – לאחר איחוד
68.....	קטגוריות
69.....	נספח 12 - ניתוח תוצאות אלגוריתם Multinomial Regression
69.....	12.1. כלל המדדים
	12.2 תוצאות מובהקות המדדים לפי פונקציית Pvalue – feature selction - אחרי
69.....	טרנספורמציה
	12.3 תוצאות מובהקות המדדים לפי פונקציית Pvalue – feature selction - אחרי
69.....	איחוד קטגוריות
70.....	נספח 13 – אחוזי הניבוי במודל Random Forest
70.....	נספח 14 – אחוזי הניבוי במודל Multinomial Regression
71.....	נספח 15 – אחוזי הצלחה בניבוי הקטגוריות
72.....	נספח 16 – נוסחאות לחישוב המדדים



## רשימת טבלאות

25.....	טבלה 1 - משימות הניסוי
25.....	טבלה 2 - סדר אקראי לביצוע משימות הניסוי
26.....	טבלה 3 - תוצאות הפיילוט לסיווג רמת עומס המשימות
33.....	טבלה 4 - קורלציית ספירמן בין המדדים
33.....	טבלה 5 - מובהקות קורלציית ספירמן
35.....	טבלה 6 - מבחן ANOVA לשוונות קצב לב
36.....	טבלה 7 - מבחן Tukey לשוונות קצב לב
37.....	טבלה 8 - מבחן ANOVA זמן ביום
37.....	טבלה 9 - מבחן Tukey זמן ביום
39.....	טבלה 10 - אחוזי ניבוי המודל Random Forest
39.....	טבלה 11 - אחוזי ניבוי המודל Multinomial Regression

## רשימת איורים

27.....	איור 1 - רצועת החזה Polar H10
34.....	איור 2 - תוצאות שאלון NASA-TLX לפי סוג משימה
34.....	איור 3 - תוצאות שאלון SWAT לפי סוג משימה
35.....	איור 4 - תרשים קופסא - שוונות קצב הלב

## 1. מבוא

משימות רבות דורשות הפעלת מאמץ קוגניטיבי מאלו המבצעים אותן, בין אם בנהיגה, בבקרת תעופה אווירית, בניתוח של חולים, במבצעים צבאיים ובספורט קבוצתי. מבצע המשימה משקיע אנרגיה בכך שהוא משתמש בזיכרון העבודה, בידע רלוונטי לאותה משימה ומבצע קבלת החלטות לאורך זמן המשימה (Coyne et al., 2009). בנוסף לכך, ביצוע מטלה במשך פרק זמן ארוך תוך שמירה על רמת הביצועים עשויה לגרום לעומס קוגניטיבי (Yatsenko et al., 2018). Young and Stanton (2002) הגדירו את המונח "עומס קוגניטיבי" בתור רמת משאבי הקשב הדרושים לצורך עמידה במשימה, בהתאם לדרישות המשימה, מידת העזרה החיצונית הקיימת וניסיון העבר של המבצע (Stanton, 2002 & Young).

מדידת העומס הקוגניטיבי היא משימה מורכבת בפני עצמה, מכיוון שלא ניתן להבחין בעומס באופן ישיר (Momeni et al., 2019). עם זאת, ישנן שיטות רבות למדידה המתבססות על מספר סממנים פיזיולוגיים בכדי לאפיין את רמת העומס המנטלי (Young et al., 2003; Paas et al., 2015) (et al., 2015) למשל, מדידות אק"ג (Momeni et al., 2019), אא"ג (Friedman et al., 2019), נשימה וקצב הלב (Luque-Casado et al., 2016; Mansikka et al., 2016), הם חלק מהמדדים הפיזיולוגיים שנעשה בהם שימוש לאמידת העומס הקוגניטיבי.

מחקרים שונים אף מגדירים סקאלות שונות לצינוי רמת העומס הקוגניטיבי, כך שאין אחידות במתן הצינוי. Grimes et al. מדדו את רמת העומס לאחר ביצוע מטלות זיכרון, וחילקו את הרמות לקטגוריות ולא הגדירו את הצינוי כערך רציף (Grimes et al., 2008). Monfort ושותפיו סיווגו בין מספר רמות עומס קוגניטיבי גבוהות (Monfort et al., 2016), לעומת זאת Fisher Wilson and Kothe and Makeig ו- (Fisher, 1995) מנגד, Caywood et al. הגדירו את העומס במחקרם כמשתנה רציף (Caywood et al., 2017).

בשנים האחרונות אף נעשו ניסיונות ליצור מודלים מבוססי לימוד מכונה, אשר יהיו מסוגלים לחזות את רמת העומס הקוגניטיבי לפי המדדים הפיזיולוגיים שצוינו לעיל (Heard et al., 2018). כלומר, ניתן יהיה לקבוע את רמת העומס שאדם שרוי בה בזמן אמת על סמך הנתונים שהופקו, בין אם בצורה רציפה ובין אם בצורת קטגוריות.

כמו כן, מודלים מסוימים אף מנסים לחזות איזו משימה מפעיל מבצע ברגע נתון, בהסתמך על נתוני המדדים הפיזיולוגיים בעת ביצוע כל משימה (Heard et al., 2018).

ישנן מספר בעיות עיקריות בתהליך המדידה של עומס קוגניטיבי כפי שנעשה במחקרים שצוינו קודם. ראשית, מחקרים רבים מגדירים את רמת העומס של הנבדק בסוף המטלה ולא בזמן אמת. שנית, הציוד שמשתמשים בו כדי להפיק את המדדים שמסייעים לאמוד את העומס הקוגניטיבי (אק"ג, אא"ג) הוא יקר ולכן לא הרבה חוקרים מסוגלים להרשות לעצמם להשתמש

בו (Zak et al., 2020). כפי שהזכרנו, קיים חוסר אחידות בספרות בחלוקה לקטגוריות של מידת העומס הקוגניטיבי - חלוקה בצורה גסה (למשל לרמה נמוכה, בינונית וגבוהה), לעומת מתן ציון מספרי רציף. חלוקה בצורה גסה יוצרת מגבלה, משום שכאשר המפעיל מסווג לרמת עומס קוגניטיבי בצורה בדידה, ייתכן שהמדדים שלו היו קרובים אך לא מספיקים כדי לסווג אותו לרמת העומס המדויקת בה הוא נמצא.

במחקר שלנו, נרצה לבדוק את המתאם בין מדדי דופק המופקים מרצועת חזה לעומס קוגניטיבי, במטרה לייצר מודל מבוסס לימוד מכונה שינבא את רמת העומס הקוגניטיבי. נושא המחקר שלנו נועד לנסות לפתור את הבעיות והמגבלות לעיל, משום שרצועת החזה היא כלי מדידה נגיש, זול ונוח לשימוש, שבאמצעותו ניתן לקבל נתונים על קצב הלב בזמן אמת. אחת ממטרות המחקר היא לתקף את השימוש ברצועת החזה ככלי חדשני, שניתן להפיק ממנו מידע באופן רציף (Pereira et al., 2020). כמו כן, בהתבסס על נתוני קצב הלב, נשאף לבנות מודל שינבא את רמת העומס בזמן אמת ולא רק בעת סיום משימה, וייתן ציון רציף לעומס הקוגניטיבי בו נמצא המפעיל בצורה מדויקת ומהימנה ככל שניתן.

המחקר שלנו הוא מחקר מעבדה שמבוצע בתנאים ביתיים בהתאם למגבלות מגפת הקורונה, שבמסגרתו אנו נבצע ניסוי ובו נלבש את רצועת החזה. אנו נמדוד לאורך מספר ימים את המדדים הפיזיולוגיים הסובייקטיביים ומדדי הביצועים של כל אחד מהחוקרים. את מידת העומס נתעד באמצעות מילוי שאלון NASA-TLX, שהוא כלי הערכה סובייקטיבי ורב ממדי המדרג את עומס העבודה הנתפס על מנת להעריך את קושי המשימה, תפקוד המבצע ואת ביצועיו. בנוסף למידת העומס, את כלל המדדים הפיזיולוגיים שנפיק מהרצועה נתעד בנקודות זמן רבות במהלך כל יום. זאת על מנת שבסוף התהליך יהיו לנו כמות גדולה של דאטה, שנוכל באמצעותה לבנות מודל שיהיה מסוגל לחזות את מידת העומס שאדם נמצא בה ברגע נתון.

## 2. סקירת ספרות

### 2.1 עומס קוגניטיבי

עומס עבודה נפשי קוגניטיבי (MWL - Mental WorkLoad) הינו מושג מהנפוצים בתחום הנדסת גורמי אנוש וארגונומיה וקיים בספרות שנים רבות. Hart and Wickens הגדירו עומס עבודה נפשי כמונח הכללי המשמש לתיאור העלות הנפשית של הישג בעבור דרישות משימה (Wickens, 1990 & Hart). גם היום, עומס עבודה נפשי הינו משתנה חשוב על מנת להבין את ביצועי המשתמשים ועלה בפופולריות מאז שנות השמונים כחלק מבעיות יישומיות, במיוחד בתחום התחבורה (Young et al., 2015).

למרות ההתייחסות רבת השנים לעומס עבודה נפשי והשפעותיו, עומס קוגניטיבי הוא אחד המושגים המעורפלים ביותר בהגדרתו ונשאר קשה להגדירו. אף על פי שהוצעו הגדרות רבות, אין הסכמה אוניברסלית בין הצהרות שונות אלה. עם זאת, ישנו מין המשותף בין הפרשנויות השונות. (Young et al., 2015)

ביצוע משימה מורכבת הכוללת מספר מרכיבים דורש שילוב משאבים שונים כמו שלילת ידע הקשור למשימה, שימוש בזיכרון עבודה, תשומת לב וקבלת החלטות על אף שיש כמות מוגבלת של משאבים עבור שילוב המידע ועיבודו. עומס קוגניטיבי מתייחס לעומס היחסי על המשאבים המוגבלים העומדים לרשותנו (Coyne et al., 2009; Sweller et al., 1998). עומס קוגניטיבי יכול להתפרש כפיזי או נפשי, אך מכיוון שעומסים אלה קשורים זה בזה באופן קבוע, לא ניתן להפריד ביניהם לחלוטין כאשר מבצעים משימה מסוימת (Shan, ) & Lean, 2012.

כמו כן, יש הבחנה בין עומס המשימות (Taskload) ועומס קוגניטיבי (MWL). ניתן להגדיר עומס משימות (Taskload) כעבודה, למשל מספר המשימות שבוצעו על ידי מפעיל, לעומת עומס קוגניטיבי אשר ניתן להגדירו כחווייה סובייקטיבית של עומס משימות נתון (Charles & Nixon, 2019).

#### 2.1.1 גורמים המשפיעים על עומס עבודה

עומס קוגניטיבי מושפע מגורמים רבים כאשר עומס משימות (Taskload) הוא אחד מהם. משימה פשוטה שחוזרת על עצמה עשויה להיות לא מאתגרת קוגניטיבית, אך אם נוסף אליה לחץ זמנים, עומס העבודה הנפשי עשוי לגדול ולהשפיע על הביצועים של המשימה (Young et al., 2015).

מצד שני, משימה מורכבת עשויה תחילה להיתפס כמאתגרת ובעלת עומס עבודה (Taskload) גבוה, אשר יוצר עומס קוגניטיבי גבוה, אך באמצעות תרגול וניסיון, העומס הקוגניטיבי שחווה המפעיל עשוי לרדת למרות שעומס העבודה והמשימות לא השתנו (Matthews et al., 2015). מפעילים עשויים לחוות עומס קוגניטיבי גבוה במשימות פשוטות כגון נהיגה ברכב או משימות לצורך לימוד (Donchin, 1986 & Gopher) ולעומת זאת, במשימות של חילוץ והצלה, המחלץ בהכרח יחווה עומס קוגניטיבי גבוה, בין היתר משום שעליו לפעול במהירות על מנת להציל חיי אדם. כשמתרחש אסון, המחלץ צריך להתמודד במקביל עם מספר משימות מורכבות שגורמות לו לחוות עומס קוגניטיבי, כמו למשל זיהוי ובקרה על נתיבי חילוץ, חיפוש קורבנות וקבלת החלטות בזמן אמת לניהול הסיטואציה (Agnola et al., 2018'Dell). לכן, זיהוי של עומס קוגניטיבי גדול בעת הביצוע הוא חשוב מאין כמוהו על מנת למנוע אסונות או תאונות.

#### 2.1.1.1 מיומנות כגורם המשפיע על עומס העבודה

עומס עבודה קוגניטיבי נחקר מכמה נקודות מבט- Mills ושותפיו השתמשו בשאלות פשוטות של נכון/לא נכון עבור עומס עבודה נמוך ושאלות פתוחות, הדורשות זיכרון מדויק יותר, לצורך העלאת עומס העבודה הקוגניטיבי (Mills et al., 2017). כמו כן, Stevens ושותפיו העריכו נבדקים בזמן שלמדו (Stevens et al., 2006), ו-Mak ושותפיו התמקדו בשיפור ביצועים במשימה חזותית-מוטורית. הם הראו ירידה במדדי העומס הקוגניטיבי כאשר המפעילים הכירו את המשימות טוב יותר (Mak et al., 2013). כמו כן, ניתן לראות כי כאשר הנבדקים רוכשים מיומנות ולומדים כיצד לבצע משימה באופן אוטומטי, עומס העבודה הקוגניטיבי שלהם פוחת (Borghini et al., 2017). כלומר, ניתן לראות כי העומס הקוגניטיבי לא תלוי רק במורכבות המשימה אלא גם במיומנות הנבדק במשימה הנתונה. משימה מורכבת אשר מבוצעת על ידי מפעיל בעל ניסיון תביא לעומס קוגניטיבי נמוך יותר מזו אשר מבוצעת על ידי אדם ללא ניסיון.

#### 2.1.2 חשיבות המדידה של עומס עבודה קוגניטיבי

קיימת חשיבות גוברת למדידה ומחקר של עומס קוגניטיבי. בגלל התפתחות המיכון והאוטומציה, עומס עבודה נפשי (MWL) הוא אחד הנושאים החשובים ביותר בתחום בניה ותכנון של מערכות עבודה (Nachreiner, 1995). כמו כן, בשל התקדמות הטכנולוגיה, בסביבות עבודה רבות קיימת דרישה קוגניטיבית גבוהה ממפעילים בעוד שהדרישות הפיזיות מהם פוחתות. התקדמות זו מעלה את ההבנה ההולכת וגוברת שעומס קוגניטיבי פוגע בביצועי המפעילים (Young et al., 2015).

דוגמה לחשיבות הרבה בנושא הינה ניסוי אשר בוצע על ידי Jafari ושותפיו בטהרן על נהגי רכבת, אשר חווים רמות שונות של עומס קוגניטיבי במהלך הפעולות השגרתיות והלא שגרתיות

בעבודה שלהם. בניסוי בדקו 11 מפעילי רכבת תחתית שחוו תרחישי נהיגה שונים בסימולציה. המשימות המדומות חולקו לשתי קטגוריות: פעולות שגרתיות כמו הכנה לנסיעה ונסיעה בין תחנות ללא הפרעות ופעולות לא שגרתיות כמו תגובה לשריפה במנהרה והתמודדות עם צפיפות גבוהה של נוסעים. עומס העבודה הנפשי היה מנוטר ומוערך במשימות אלה על פי פעילות הלב, הגדרת עומס בסולם סובייקטיבי וביצועי נהיגה. המחקר הראה כי לפעולות רכבת לא-שגרתיות יכולות להיות השפעות שליליות על העומס הנפשי של נהג הרכבת וכן על הביצועים שלו. הממצאים של מחקר זה יכולים לשמש עבור הגדרת גבולות עומס קוגניטיבי אצל מפעיל על מנת למנוע פגיעה בביצועים. (Jafari et al., 2020).

Friedman ושותפיו מחזקים עמדה זו של חשיבות המדידה וההערכה של העומס הקוגניטיבי הקשור למשימות שונות ובתחומים שונים ואומרים כי הינה קריטית היום עבור יישומים רבים, החל מתכנון חומרי הדרכה ועד מעקב אחר רווחתם של טייסי המטוסים (Friedman et al., 2019).

לעומס הקוגניטיבי קיימות גם השפעות בריאותיות. כאשר מצטברת רמה גבוהה של עומס קוגניטיבי לאורך זמן עלולות להופיע בעיות בריאותיות כמו לחץ כרוני, דיכאון או שחיקה. לכן, ניטור רציף של עומס קוגניטיבי מציע הזדמנויות חדשות לתמוך במניעת הפרעות נפשיות ושמירה על בריאות הנפש (Cinaz et al., 2012).

## 2.2 מדדים לעומס קוגניטיבי

זיהוי מהימן של העומס הקוגניטיבי שאדם שרוי בו מהווה אתגר. בדרך כלל, העומס נקבע על ידי המאפיינים של המשימה המבוצעת, המפעיל וההקשר הסביבתי, מה שמקשה על ביצוע מדידה ישירה (Matthews, 2019 & Hancock). שילוב של פידבקים פיזיולוגיים, ממדים סובייקטיביים וביצועי מטלות מספקים נקודות מבט שונות ומשלימות בהערכת העומס הקוגניטיבי (Myung, 2005 & Ryu). מחקרי עבר הדגישו את הצורך להשתמש בשילוב של מדדים פיזיולוגיים, סובייקטיביים וביצועיים כדי להעריך את המאמץ הנפשי של משתמשים בעבודות בעלות ריבוי משימות (Cegarra and Chevalier, 2015). Young et al. חקרו את השימוש במספר מדדים עבור עומס קוגניטיבי בניסוי לפתרון חידות. התוצאות מצביעות על החשיבות הגבוהה של שילוב של מספר אמצעים כדי לבנות יסודות חזקים לאמידת עומס קוגניטיבי (Chevalier, 2008 & Cegarra).

נכון להיום, ישנן מספר גישות להערכת העומס: שאלונים סובייקטיביים, ניתוח ביצועים וניתוח של תגובות פיזיולוגיות (Cain, 2007). מילוי של שאלונים סובייקטיביים לדוגמה לא יכול להיעשות בתדירות גבוהה מאוד תוך כדי ביצוע מטלה, ולכן לרוב מבוצע בסיום כל מטלה. ניתוח ביצועים מודד את ההבדל בין התוצאה הרצויה לבין התוצאה בפועל, אך גם כאן יש קושי למדוד את הביצועים בצורה רציפה לאורך זמן (Donchin, 1986 & Gopher). ניתוח של

תגובות פיזיולוגיות, לעומת זאת, הוא אמצעי שניתן להשתמש בו בזמן אמת, באופן רציף ובצורה לא פולשנית. לכן, הוא בעל יתרונות על פני האמצעים האחרים (Mehler et al., 2009).

### 2.2.1 מדדים פיזיולוגיים לעומס קוגניטיבי

מדדים פיזיולוגיים הם מדדים אובייקטיביים להערכה עומס עבודה נפשי (MWL). לדוגמא, ככל שעומס הקוגניטיבי גדל, קצב דופק הלב גדל (Mansikka et al., 2016). ההתקדמות הטכנולוגית מביאה לעליה במדידות הפיזיולוגיות של עומס קוגניטיבי וכן הרצון למדוד ולחזות את סטטוס המפעיל ובאופן ספציפי בעזרת חיישנים. קיימים מגוון מדדים פיזיולוגיים המספקים הערכה של העומס כמו: פעילות לב, קצב נשימה, מדדים על בסיס עור, לחץ דם, מדדי תזוזת עיניים ופעילות מוח חשמלית (Nixon, 2019 & Charles). הפיתוחים הופכים את המדידה הפיזיולוגית לקלה יותר, ניידת יותר, זולה יותר ובתצורה פיזית קטנה יותר (Charles, 2017 & Nixon). הטכנולוגיות מאפשרות איסוף וניתוח של מגוון נתונים הקשורים לפיזיולוגיות של המשתמשים (Malik, 2016 & Guzik). נתונים אלה יכולים להיאסף באופן לא פולשני ובמקרים רבים ללא הפרעה למשימה העיקרית (Charles & Nixon, 2019).

במדידות פיזיולוגיות, ניתן לתעד את מדד קצב הלב (HR - heart rate) ואת מדד שונות קצב הלב (HRV - heart rate variability). מדדים אלה דורשים מכשירים פחות מתוחכמים והוכיחו שהם רגישים לעומס עבודה נפשי קוגניטיבי בסביבות מדומות (De Waard, & Brookhuis, 2010; Mansikka et al., 2016). דופק (HR) מהווה את המדידה הכי לא פולשנית מבין אפשרויות המדידה הפיזיולוגיות הזמינות ולכן זו האופציה המקובלת מבחינה תפעולית. בנוסף, התפתחויות רבות לגבי מתודולוגיות ניתוח הנתונים תוך שימוש במדד HRV הובילו להתעניינות נוספת ביישום מדד ה-HRV כמדד אוטונומי (Watson, 2001). סוקר את היישום של שונות קצב הלב (HRV) והשימוש בו ומגדיר אותו כמדד יעיל לעומס בעבודה ועייפות בצוות טיסה (Watson, 2001). עם זאת, לעתים קרובות שיטות אלה דורשות הצמדת חיישנים פיזיים למשתתף ולכן אינן נמצאות בשימוש נרחב בתחומים מעשיים (Chuang et al., 2016).

מדד שונות קצב הלב (HRV - Variability Rate Heart) ה-HRV הינו מדד מקובל לתיאור שינויים מידיים בקצב הלב. נמצא כי קיים קשר בין HRV כאומדן לפעילות וגאלית לבין תהליכים קוגניטיביים כמו למשל תהליכים אקזקוטיביים ותהליכי זיכרון. בחרנו להשתמש במדד שונות קצב הלב מכיוון שהוא מדד רגיש ואיכותי למדידת לחץ ועומס קוגניטיבי על ידי זה שהוא מאפשר לנו לקבל מידע אודות פעילות מערכת העצבים של הנבדקים. נוסף כל כך מחקרים

רבים מדווחים על אמינות תגובות פסיכו פיזיולוגיות הנגרמות על ידי משימות עומס עבודה מנטלי (Aly M, 2009; Wilson GF, 2002 & Kramer AF, 2004: Riener A). מקטע R-R - המרחק בין שני גלי R, נמדד באלפיות השנייה. חישוב הפרשי הזמן בין כל פעימה לפעימה עוקבת. המדידה נעשית בדרך כלל עי ידי מדידת מקטע הזמן שבין מעבר הגירוי החשמלי בקיר חדרי הלב - מרווח הזמן בין פסגות ה RR (Acharya et al., 2004).

#### 2.2.1.1 קצב הלב כמדד לעומס עבודה קוגניטיבי

טכניקות א.ק.ג מודדות את הפעילות החשמלית של הלב באמצעות מספר חיישנים. האות הגל החשמלי החוזר מייצג את הקיטוב והדיפולריזציה (Resting potential) של הלב הנדרש לשאוב דם סביב הגוף. ניתן להשתמש במדדים שונים הקשורים לגל זה כדי לאפיין את פעילות הלב המעניקה תובנות על עומס עבודה נפשי (Nixon, 2019 & Charles). בעיה מרכזית בטכניקת הא.ק.ג היא שאינה בעלת רגישות והבחנה גבוהה במדד לעומס עבודה קוגניטיבי. לכן, מדידה זו מתאימה בעיקר להבדיל בין עבודה לבין מנוחה או להבדיל בין משימות כאשר רמת השונות ביניהן צריכה להיות גבוהה (Nickel & Nachreiner, 2003).

#### 2.2.1.2 מדדים סובייקטיביים לעומס קוגניטיבי

מדידה סובייקטיבית ניתן לבצע ולערוך בקלות יחסית, כאשר המדידה מוגדרת היטב ומותאמת לדרישות של המשימה עבור המפעילים. במחקר נעשה שימוש בגישת המדדים הסובייקטיביים כדי לקבל מידע על העומס ועל העייפות כפי שהם נתפסים בקרב המפעילים (Watson, 2001). המדדים לציון מצב סובייקטיבי עצמי מומלצים להערכה של משימות נפשיות הכרוכות בקבלת החלטות אך אינם מומלצים להערכת משימות פיזיות הכוללות פעילויות שחוזרות על עצמן (Cain, 2007).

מדד עומס המשימות NASA-TLX הוא דוגמה לדירוג עצמי סובייקטיבי מבוסס היטב. בשיטה זו, מפעילים יכולים לדרג את דרישות העבודה בעצמם על סולם מספרי (Young et al., 2015). בשאלון, על העונים לסמן את תשובתם לכל שאלה על סקאלת מספרים שבין 1 ל-20 כך ש-1 זהו הציון הנמוך ביותר ו-20 הוא הגבוה ביותר. כל שאלה מוצגת בתור כותרת של תחום מסוים שבאמצעותו מגדירים את העומס הקוגניטיבי, ומתחת לו מופיעה הסקאלה המתוארת לעיל. בנוסף, ליד כל תחום מופיעה שאלה רלוונטית שמבהירה כיצד לענות על השאלה (Schuff et al., 2011). השאלון מכיל שישה ממדים ביניהם דרישה נפשית, דרישה פיזית, דרישות זמנים, ביצועים, מאמץ ותסכול (Jafari et al., 2020). השימוש במדד זה הינו נפוץ וקיים במחקרים רבים כמו במחקר מדידת העומס הקוגניטיבי על מפעילי נהגי רכבת (Jafari).



(et al., 2020), וכן מחקר העומס הקוגניטיבי על מפעילים במרכז בקרת תעבורה בעיר תוך פיקוח על צפיפות התנועה (Fallahi et al., 2016).

קיימות אפשרויות נוספות למדידה סובייקטיבית עבור תפיסתו של המפעיל. למשל, באמצעות סולם דירוג המכונה סולם Bedford הסולם הוא תהליך מוכר ומאומת של מדידת עומס קוגניטיבי הן בצבא והן בסביבות תעופה אזרחיות (Watson, 2001). בניסוי על צוות טיסה ביצעו מדידות לרמת העייפות הסובייקטיבית באמצעות סרגל אנלוגי בגודל 10 ס"מ. צוות הטיסה התבקש להקליט עייפות נתפסת בנקודות ספציפיות בתוך כל אחת מהטיסות (Watson, 2001).

אחת מבין האפשרויות הנוספות למדידה סובייקטיבית היא באמצעות מענה על שאלון SWAT (Subjective Workload Assessment Technique). השאלון מגדיר סקאלה רב ממדית שמטרתה להעריך עומס עבודה שקיים על מפעיל. בשיטה זו הגדירו שלושה ממדים לפיהם מעריכים עומס העבודה - עומס זמן, מאמץ מנטלי ולחץ פסיכולוגי (Nygren, 1988 & Reid). לכל ממד, הנבדק נותן ציון הנע בין 1 ל 3 כך ש-1 מהווה רמת עומס נמוכה, 2 מהווה רמת עומס בינונית ו-3 רמת עומס גבוהה. נוסף על כך, כל רמה מאופיינת בתיאור מפורט של רמת הקושי. השיטה תוכננה במקור להערכת עומס העבודה הקשור לפעילות של טייסים בתא הטייס ובסביבות תחנות עבודה אחרות בהן עובדים בצוות (Rubio et al., 2004). השימוש בשאלון דורש שני שלבים. בשלב הראשון, הנבדק מדרג באמצעות כרטיסיות כל אחד מ-ממדי העומס לשלוש רמות, מהנמוך ביותר לגבוה ביותר. הכרטיסיות מייצגות את כל האפשרויות לדירוגים השונים, כלומר את כל השילובים האפשריים של ממד וציון. סך הכול ישנן 27 כרטיסיות. על הנבדק לסדר את הכרטיסיות לפי רמת עומס המידע שלדעתו אדם ירגיש בהתייחס לזמן, למאמץ המנטלי וללחץ הפסיכולוגי. בשלב השני במהלכו של הניסוי הנבדק מתבקש לענות על השאלות ולדרג כל מימד. בסיום שני השלבים לוקחים את הסידור שהתקבל ובוחנים היכן הנבדק דירג את הסידור הספציפי ועל פי כך קובעים את ציון שאלון ה SWAT.

### 2.3 מדידת עומס עבודה קוגניטיבי בפעילות יומיומית רציפה

Cinaz ושותפיו ביצעו ניסוי הבוחן עומס עבודה בשלוש רמות שונות בסימולציה במעבדה ובחיי היומיום ומצאו התאמה גבוהה בין עומס נמוך, בינוני וגבוה במהלך יום עבודה. מוטיבציית הניסוי הייתה מתוך ההבחנה כי רמות גבוהות של עומס קוגניטיבי עלולות להוביל לבעיות בבריאות הנפש.

הניסוי הראשון היה המעקב אחר העומס הקוגניטיבי במעבדה על 7 נבדקים כאשר מטרת הניסוי הייתה לדמות רמות עומס שונות ולהבחין ביניהן על סמך מדד קצב לב (-heartd rate) (HR) הנמדד על ידי מערכת אק"ג ניידת. כל נבדק ביצע שלוש משימות מסוג N-Back Task שהן משימות הדורשות תשומת לב ותגובה מהירה לאורך זמן. הנבדקים ביצעו משימות ברמות

שונות (נמוכה, בינונית וגבוהה) כאשר סדר המשימות היה אקראי על מנת להימנע מהשפעת הרצף.

הניסוי השני היה מעקב אחרי עומס קוגניטיבי בחיי היומיום, אשר בוצע על שניים מתוך הנבדקים בניסוי הראשון. מטעמי נוחות, בבדיקה של חיי היומיום, רצוי להשתמש בחיישן קל ומינימלי לכן הם משתמשים בחגורת החזה Zephyr BioHarness. החגורה מורכבת משלושה חיישני בד חכמים לקבלת פעילות לב, קצב נשימה וטמפרטורת העור. כמו כן, תאוצת הגוף נמדדת במד תאוצה D3 הכלול במכשיר. בנוסף לכך, חגורת החזה מספקת מרווחי RR על ידי מדידת משך הזמן בין שני גלי R רצופים של אק"ג.

במהלך הניסוי במעבדה חילקו את המדידה הרציפה למרווחי זמן של 3 דקות והוצאו ערכי ה-RR וה-HRV עבור כל נבדק במרווח זמן זה. עבור כל חלון זמן בוצע סיווג המידע על ידי אלגוריתם סיווג KNN- k Nearest Neighbors algorithm (אלגוריתם השכן הקרוב) ואימות על ידי Leave-one-out cross-validation (LOOCV). עבור כל הנבדקים התקבלה שיעור התאמה גבוה של 82.14% לסיווג מצב העומס וסטיית תקן של 12.91%.

במהלך הניסוי היומיומי חילקו את המדידה הרציפה למרווחי זמן של 10 דקות והוצאו ערכי ה-HRV עבור שני הנבדקים במרווח זמן זה, כאשר הוצאו חלקי הזמן בהם בוצעה פעילות גופנית. גם כאן בוצע סיווג המידע על ידי אלגוריתם סיווג KNN- k Nearest Neighbors algorithm (אלגוריתם השכן הקרוב) ואימות על ידי Leave-one-out cross-validation (LOOCV). שיעור הדיוק הממוצע אצל הנסיין הראשון היה 72.03% ואילו אצל הנסיין השני היה 77.66%.

לבסוף, נבדק האם הנתונים מניסוי המעבדה מתאימים להבחנה בימי עובדה בעלי עומס נמוך, בינוני וגבוה. התוצאות מראות שעומס העבודה המדומה במסגרת המעבדה אינו משקף את כל רמות עומס העבודה בחיי היומיום שלנו. (Cinaz et al., 2012)

### 2.3.1. השפעת השעה ביום על ביצועים

עייפות קוגניטיבית היא מצב שכיח מאוד הנובע ממעורבות מתמשכת במשימה ובדלדול המשאבים הנפשיים שלנו. מספר רב של מחקרים הוכיחו כי התוצאות של עייפות קוגניטיבית מתמשכת הן שחיקה בעבודה, מוטיבציה נמוכה, הסחות דעת רבות ובעיות בעיבוד מידע. עייפות קוגניטיבית אף מורידה את איכות השיפוט וקבלת ההחלטות אצל אנשים, כולל אלה של מומחים. אחד המחקרים התחקה אחר קבלת החלטות אותנטיות במסגרת ועדות שחרור אסירים, ובחן את השפעת העומס וריבוי תיקים על קבלת ההחלטה (Danziger et al., 2011). החוקרים גילו כי השופטים נטו בסבירות גבוהה יותר לסרב לבקשה של האסיר ככל שהתקדמו במספר המקרים ביום נתון. פסקי דין סבירים היוו כ-65% מכלל פסקי הדין בתחילת היום, אך הלכו והתמעטו לאורך היום. עד להפסקת האוכל, אחוז פסקי הדין החיוביים היה גבוה. לאחר

ההפסקה, כאשר בית המשפט חזר לפעילות, נראה דפוס שחוזר על עצמו: תחילה מספר גבוה של פסיקות חיוביות ובסוף יום העבודה מספר נמוך עד אפס פסיקות חיוביות. מחקרים נוספים מראים כי שעה מסוימת ביום יכולה להשפיע על העומס הקוגניטיבי. למשל, החוקרים Sievertsen et al. מצאו כי השעה ביום משפיעה על הישגי התלמידים בבתי הספר (Sievertsen et al., 2016). המחקר נערך על ילדים בני 8-15 בבתי ספר ציבוריים בין השנים 2009-2010 ו-2012-2013 ונערך בקרב יותר משני מיליון תלמידים. במחקר מדדו את ההשפעה של השעה ביום ושל ההפסקות על ביצועי התלמידים במבחנים סטנדרטיים. מסקנות המחקר הראו כי עייפות קוגניטיבית הובילה תלמידים לביצועים ירודים במבחנים ואילו ההפסקות מילאו את מצבורי האנרגיה המחשבתית שלהם. המחקרים הובילו לשלושה ממצאים עיקריים. הראשון הוא שככל שהמבחן התקיים מאוחר יותר באותו היום, ביצועי התלמידים במבחן ירדו. שנית, הפסקות שיפרו משמעותית את הביצועים. שלישית, השפעת השעה ביום והשפעת ההפסקות לא הייתה הומוגנית. כלומר, תלמידים עם ביצועים נמוכים הושפעו יותר על ידי הפסקות (וגם על ידי השעה ביום) מאשר תלמידים עם ביצועים גבוהים. מחקר נוסף שנערך באוסטרליה בחן אם יש קשר בין השעה ביום לבין מידת הרצון להשקיע במטלה, וכיצד הוא מתרחש במוח (Byrne et al., 2017). בניסוי השתתפו 16 גברים, אשר התבקשו להשתתף במשחק ניחוש נוסף בשלושה מועדים במהלך היום – בעשר בבוקר, בשתיים אחר הצהריים ובשבע בערב. המשחק היה ניחוש אם ערך ספרה שהם לא יכלו לראות גדול או קטן יותר מחמש. ניחוש נכון העניק למשתתפים פרס כספי וניחוש שגוי גרר עונש כספי. נוסף על כך, כדי שהמתנדבים יבצעו את המקסימום בכל סיבוב, נאמר להם שעבור הסיבוב הטוב ביותר הם יקבלו בONUS. למעשה, למשחק הייתה תוצאה ידועה מראש וכל המשתתפים קיבלו את אותו מספר של פרסים ועונשים, כדי שאפשר יהיה להשוות את התוצאות שלהם. כדי לבחון את השפעת השעה על תחושת התגמול במוח, החוקרים ביצעו סריקה מוחית באמצעות מכשיר תהודה מגנטית תפקודי (fMRI) ובחנו כיצד אזורי מוח הקשורים לתחושת תגמול מגיבים למשחק בשעות היום השונות. הסריקה מדדה את כמות הדם שזרמה לחלקים השונים במוח של המתנדבים בזמן שהם שיחקו את המשחק, על סמך הנחה כי אזור במוח צורך יותר דם ככל שהוא פעיל יותר. הסריקה גילתה שהפעילות באזור הקשור בתחושת התגמול, הפוטאמן, הייתה נמוכה ביותר במשחק של שתיים אחר הצהריים, לעומת מבחני הבוקר והערב. החוקרים מעריכים כי אזורי התגמול פחות פעילים באמצע היום.

## 2.4 לימוד מכונה (Machine Learning)

לאחרונה, המחקרים העדכניים שבדקים את המאמץ הקוגניטיבי בהסתמך על סממנים פיזיולוגיים החלו להשתמש בטכניקות של למידת מכונה (Heard et al., 2018). למידת מכונה (Machine Learning) היא תת-תחום במדעי המחשב ובבינה מלאכותית העוסק בפיתוח

אלגוריתמים המיועדים לאפשר למחשב לשפר באופן אוטומטי את הביצועים שלו באמצעות למידה וצבירת ניסיון (Mitchell, 1997; Simon, 2013). המטרה הבסיסית של מכונה לומדת היא לפתח יכולת להכליל מתוך הניסיון, ולחקות יכולת זו של למידה אצל בני אדם (Schain & Schain, 2015). כלומר, יכולת לבצע מגוון משימות הקשורות לחיזוי ולקבלת החלטות באופן מדויק ככל האפשר בהסתמך על מידע שעדיין לא נצפה, וזאת על בסיס צבירת ניסיון ממידע קיים (Valiant, 1984).

#### 2.4.1 סיווג אלגוריתמים בלימוד מכונה

אלגוריתמי לימוד המכונה מתחלקים לשלושה סוגים. הסוג הראשון נקרא למידה מונחית (supervised learning) – מטרת לימוד זו היא לחזות את הסיווג של נתונים חדשים שאותם לא פגש בתהליך הלמידה (Gibson et al., 2013). הסוג השני נקרא למידה בלתי מונחית (unsupervised learning) – טכניקה שבאמצעותה מנסים ללמוד את התכונות והמבנה של אוסף דוגמאות נתונים. הנתונים הקיימים עליהם מסתמך האלגוריתם לא כוללים הגדרת שיוך מראש של המידע לקטגוריות על ידי תיוג מוקדם. למשל, נתונים הכוללים מידע רפואי על נבדק כמו חום, דופק ולחץ דם ללא תיוג המציין אם הנתונים שייכים לאדם חולה או בריא (Abu-Mostafa et al., 2012). לבסוף, הסוג האחרון הוא למידת חיזוק (reinforcement learning) – אלגוריתם הלמידה מקבל אינדיקציה על הפתרון אשר ביצע רק לאחר סיום המטלה ללא הבחנה מדויקת האם הצליח או לא, ועליו להסיק אילו מהחלטותיו הביאו להצלחה או כישלון (Barto, 1998 & Sutton).

מלבד השתייכותם לתחומי למידה שונים, אלגוריתמים של לימוד מכונה נועדו לפתור מספר בעיות שונות. ראשית, בעיית סיווג לאשכולות (Clustering) – חלוקה של תצפיות לקבוצות שונות כך שכל קבוצה היא בעלת מאפיינים דומים, כאשר הקבוצות נקבעות על פי התצפיות הקיימות. כל תצפית חדשה תשויך לקבוצה שהכי מתאימה לה (Bailey, 1994). שנית, בעיית הרגרסיה – חיזוי של ערך מסוים בהתבסס על נתוני עבר, תוך מזעור שגיאת החיזוי. למשל: שערך ערך של דירה על סמך מאפייניה (למשל שטח, מיקום) או הערכה מה יהיה סכום העסקאות של לקוח בחודש הקרוב על סמך מדדי עבר של הרגלי הקנייה שלו. שלישית, בעיית הסיווג (Classification) – זיהוי הקטגוריה הנכונה של תצפית חדשה מתוך מספר קטגוריות נתונות מראש שאינן משתנות.

#### 2.4.2 אלגוריתמים של לימוד מכונה

במסגרת למידת המכונה ישנם מספר אלגוריתמים נפוצים. תחת אלגוריתמי הלמידה המונחית ניתן למצוא את הרגרסיה הליניארית, עץ החלטה (DT- Decision Tree), יער אקראי (RF - Random Forest), אלגוריתם השכן הקרוב (Solovey KNN- k Nearest Neighbors).

(Jin et al., 2014, et al., מכונת וקטורים תומכים ) (Support Vector Machines SVM - ) (2012) ורשת עצבית מלאכותית (ANN - Artificial neural network) (Russell, & Wilson) (2003b).

אלגוריתם רגרסיה ליניארית מתבסס על שיטה מתמטית למציאת הפרמטרים של הקשר בין משתנה בלתי תלוי (או מספר משתנים בלתי תלויים) למשתנה תלוי, בהנחה שהקשר ביניהם ליניארי. אלגוריתם נוסף הוא למידת מכונה באמצעות עץ החלטה. זהו מעין מודל חיזוי, הממפה בין תצפיות לבין הערך המתאים לכל תצפית, ומספק הדמיה ויזואלית בצורת עץ של נתוני הבעיה (Wu et al., 2008). באופן אינטואיטיבי, יער אקראי הוא אלגוריתם המשתמש במספר עצי החלטה כדי לשפר את חוזק הסיווג. אלגוריתם נוסף הוא אלגוריתם השכן הקרוב, שגם כן מיועד לסיווג. בהינתן קלט חדש הוא ישייך אותו לקבוצה שהכי מתאימה לו בקרב שאר הקלטים השכנים. כמו כן, מכונת וקטורים תומכים היא הטכניקה החדשה ביותר מבין אלו שצוינו עד כה, והיא משמשת לניתוח נתונים לסיווג ולרגרסיה (Vapnik, 1995 & Cortes).

עץ החלטה (Decision Tree) הוא מודל חיזוי בתחומי הסטטיסטיקה, כריית נתונים והלמידה החישובית המספק מיפוי בין תצפיות לערכים המתאימים עבורן. עץ החלטה יכול לשמש כמודל חיזוי, הממפה תצפיות על פריט ויוצר מסקנות על ערך היעד של הפריט. יער אקראי (Random Forest) הוא מודל למידה מונחית לסיווג, רגרסיה ומשימות אחרות הפועלות על ידי בניית מספר רב של עצי החלטה בזמן האימון. למשימות סיווג, הפלט של היער האקראי הוא המחלקה שנבחרה על ידי רוב העצים (Kam, 1995).

מודל נוסף של לימוד מכונה הוא רגרסיה לוגיסטית. מודל זה הוא מודל סטטיסטי המתאר קשר אפשרי בין משתנה איכותי/קטגורי, המכונה "המשתנה המוסבר", ובין משתנים אחרים המכונים "משתנים מסבירים" על ידי פונקציית log. מקורו של רעיון זה הוא בניסיונות של קרל פירסון לפתח מקדם מתאם למשתנים איכותיים, כאשר הנחתו הייתה כי ערכם של המשתנים האיכותיים נקבע על ידי משתנה נסתר בעל התפלגות נורמלית. סיר דויד קוקס אימץ את הרעיון הזה ופיתח על פיו את מודל הרגרסיה הלוגיסטית שהציג ב-1958 (Cox, 1958).

### 2.4.3 מתודולוגיית העבודה בלימוד מכונה

בתחום למידת המכונה, המטרה העיקרית היא לייצר אלגוריתמים מבוססי מודלים מתמטיים שמסוגלים ללמוד ולבצע תחזיות בהתאם למידע שהם מקבלים (Provost, 1998 & Kohavi). בשלבי תהליך יצירת מודל, משתמשים לרוב במידע המגיע משלושה סטים של נתונים: סט האימון, סט הוולידציה וסט הבחינה. בדרך כלל, מפצלים סט נתונים גדול לכדי סט אימון וסט וולידציה. פעולה זו נקראת cross validation, והמטרה שלה היא לגלות בעיות במודל ולספק תובנות על אופן ההכללה שהמודל עושה (Talbot, 2010 & Cawley). תחילה, מזינים למודל את נתוני סט האימון (Gareth et al., 2013). כתוצאה מכך, המודל נחשף לנתונים רבים של

הפרמטרים השונים המהווים את הקלט שלו, וגם נחשף לפלט שהוא צריך להחזיר בהתאם לקלט נתון. בעזרת פונקציה של למידה מונחית, המודל מתאמן על נתוני סט האימון ובאמצעותם מגדיר בצורה כללית מה יהיו משקלי הפרמטרים שלו. לאחר מכן, המודל נחשף לנתוני סט הוולידציה כדי לחזות את הפלט המבוקש על סמך נתונים חדשים אלו, וכדי לאמוד את טיב החיזוי שלו (Ripley, 2007). סט נתונים זה משמש לבדיקת ביניים וכוונן הפרמטרים של המודל במידת הצורך (Brownlee, 2017). לבסוף, מזינים למודל את נתוני סט הבחינה – נתונים חדשים שלא שייכים לשני הסטים הקודמים. סט הבחינה נועד לחשוף את המודל לנתונים שטרם ראה ולהעריך את אמינות החיזוי של המודל הסופי (Brownlee, 2017). לפני שמיישמים אלגוריתם או שיטה מסוימת כדי לייצר מודל מבוסס לימוד מכונה, עלינו להבין מהם השלבים על מנת לעשות זאת. המתודולוגיה הנפוצה ביותר שמתארת את השלבים ליצירת מודל לכריית מידע נקראת (CRISP-DM Cross-industry standard) process for (data mining) (Shearer, 2000). בעת תהליך הלימוד, לעתים מתבצע זיהוי של תבניות בתוך סטים גדולים של מידע, והשיטה הזו מחלקת תהליך זה לשישה שלבים: הבנה עסקית, הבנת המידע, הכנת המידע, מידול, הערכה ופריסה (Harper & Pickett, 2006). Cobos ושותפיו אומרים כי השיטה קלה ליישום ומגבירה את המודעות לפעולות שיש לנקוט כדי לעשות כריית נתונים בצורה טובה (Cobos et al., 2010).

#### 2.4.4 לימוד מכונה ומדדים פיזיולוגיים

מעטים המחקרים שניסו לייצר מודלים אשר מכלילים את הנתונים הפיזיים המעידים על רמת המאמץ הקוגניטיבי שאדם משקיע. אחד המחקרים הראשונים שנעשו בתחום התבצע על ידי Healey and Picard שהבחינו בין רמות העומס השונות שנהגים מרגישים בעת הנהיגה ברכב, בהסתמך על מדדי אק"ג, טמפרטורת עור (EDA) ונשימה (Picard, 2005 & Healey). הם ביצעו את הניסוי על 24 נהגים, והגיעו לרמת אמינות של 97% על ידי שימוש באלגוריתם LDA- Linear discriminant analysis הנועד לבצע סיווג ועושה זאת בצורה לינארית. יש לציין כי אחוז הדיוק לעיל הושג בבחינת סט הנתונים לצורך אימון המודל. בנוסף, Chen et al. השתמשו באותו בסיס נתונים שבניסוי הנהגים, אך הפעם בחנו את אמינות המודל באמצעות טכניקת לימוד המכונה מכונת וקטורים תומכים (SVM) (Chen et al., 2017). הם הגיעו לרמת אמינות של 99% בבחינת סט נתוני האימון, ודיוק של 89% בבחינת סט נתוני הוולידציה. המשותף לשני המחקרים הללו הוא ששניהם בדקו את נתוני האימון והוולידציה, ולא בחנו את המודל שיצרו על סמך מידע חדש, שהמודל עוד לא ראה, שמתקבל בזמן אמת.

חוץ מאשר מדידה בעת נהיגה ברכב, נחקרו עוד מספר משימות שונות אשר באמצעותן ניתן להעריך את העומס הקוגניטיבי. בנוסף לכך, נחקרה התגובה של מודלים מסוימים של לימוד

מכונה לנתונים שנאספו לאחר ביצוע המשימות שונות. במחקר שביצעו Elkin and Devabhaktuni הם התבססו על ארבעה סטים שונים של נתונים, כך שכל אחד מהסטים נוצר לאחר ביצוע משימה נפרדת (Devabhaktuni, 2019 & Elkin). למשל, אחת המשימות הייתה לבחור מספר אקראי, ובכל פעם לבצע אחת מהפעולות הבאות: להוסיף לו 2, לחסר ממנו 7 או לנוח במשך 30 שניות (Natarajan et al., 2016). במהלך המשימה רמת העומס נמדדה באמצעות מוליכות חשמלית של העור (EDA). הם טוענים שמשימה זו נבחרה כדי להעריך את מידת המאמץ של האינדיבידואל בהתבסס על היכולת שלו לזכור מספר ולבצע עליו מניפולציות מתמטיות באופן מתמיד. משימה נוספת ונפרדת הייתה לבחון את מידת הריכוז של נסיינים בעת צפייה בוידאו-קליפים, באמצעות מדד אא"ג (EEG). כמו כן, משימה שלישית אשר גם בה השתמשו ב-אא"ג (EEG) הייתה להבחין בין תמונות שונות תחת השפעת אלכוהול.

לאחר עריכת הניסויים ואיסוף הנתונים, נבדקו מספר מודלים: רשתות נוירונים (ANN), מכונת וקטורים תומכים (SVM), אלגוריתם השכן הקרוב (KNN), עצי החלטות ו-יער אקראי (RF). נמצא כי עבור המשימה המתמטית ומשימת ההבחנה בין התמונות, המודל שנתן את אחוז הדיוק הגבוה ביותר הוא עץ החלטה עם דיוק של 99% ו-91% בהתאמה, בבחינת סט הנתונים לאימון. כמו כן, המודל הטוב ביותר עבור ניסוי הצפייה בוידאו הוא רשת נוירונים, המספק דיוק של 72%.

המחקר המדובר מהווה צעד ראשוני ומשמעותי בתחום המחקר של לימוד מכונה בהיבט של אמידת עומס קוגניטיבי. זאת משום שהוא מוכיח שלאחר ניתוח טכניקות שונות של לימוד מכונה, ניתן להשתמש באלו המספקות רמת אמינות גבוהה בהסתמך על נתוני העומס הקוגניטיבי. באמצעות היכולת להחליט באופן דינמי על המודל היעיל והמדויק ביותר בכל נקודת זמן בהינתן מדדים פיזיולוגיים מסוימים, ניתן לבצע זיהוי מוקדם של עומס קוגניטיבי גבוה או נמוך. הזיהוי אף נעשה בקלות ובמהירות ועשוי לתרום לאופטימיזציה של עומס המפעיל בעשיית משימות.

אחד המחקרים שנערכו בתחום הוא של Momeni et al., אשר בדקו מספר אלגוריתמי לימוד מכונה מוכרים לצורך סיווג רמת העומס הקוגניטיבי הנוצר אצל מפעיל במהלך משימת חילוץ מורכבת, כל זאת תוך הפעלת רחפן. במחקר נערכה סימולציה של משימת חילוץ ובה נמדדו קצב הנשימה, אק"ג, קצב הלב וטמפרטורת העור. כחלק מהמחקר נבחנו מספר מודלים של לימוד מכונה במטרה למצוא את האחד שכאשר יקבל קלט, הוא ידע לנבא את רמת העומס הקוגניטיבי בצורה האמינה ביותר. עבור כל מודל הוזנו הנתונים לצורך הלמידה שלו, ולאחר מכן הוכנסו אליו נתונים שטרם נחשף אליהם כדי לאמוד את יכולת הניבוי שלו. המודל שהצליח לנבא בצורה הטובה ביותר לפי המחקר הוא אלגוריתם XGB, שהגיע לרמת אמינות של 86%. בנוסף לכך, הם טוענים כי זוהי הפעם הראשונה אשר מציגים בספרות מודל מבוסס לימוד מכונה לזיהוי עומס קוגניטיבי על סמך מדדים פיזיולוגיים. (Momeni et al., 2019). לצד

היתרונות במידול תהליך האמידה, בעריכת סימולציה כזו יש מספר חסרונות – שימוש במספר רב של אמצעי מדידה יקרים (Zak et al., 2020), היערכות לוגיסטית מורכבת בעת יצירת סביבה המדמה משימת חילוץ, והדרכת הנסיינים לצורך תפעול הרחפן. כפי שגם הוזכר קודם, אא"ג הוא אחד מכלי המדידה הנפוצים ביותר לצורך מדידת עומס קוגניטיבי, ונעשו מספר ניסיונות לפתח מודלים מבוססי למידת מכונה הנשענים על נתוני המדד. מספר מחקרים בנושא בדקו את האפשרות למידול תהליך הערכת העומס באמצעות שיטות כגון ניתוח ליניארי בצעדים (Wilson & Wilson, 1995; Russell, 2003a & Fisher, 2012 & Baldwin), רשתות נוירונים (Penaranda, 2012 & Baldwin), המאפיין שמשותף למחקרים אלו הוא שכולם עשו ניסיון באמצעות המודל לסווג את רמת העומס הקוגניטיבי שאדם נמצא בה, ולא גרמו למודל לתת ציון רציף למידת המאמץ. Wilson and Fisher נתנו לנבדקים לבצע משימות שמבדילות בין שתי רמות של עומס קוגניטיבי, וכך גם Kothe and Makeig. באופן דומה, Grimes et al. חזו את העומס על זיכרון העבודה בעזרת 4 משימות זיכרון רציפות, אך רמות העומס במחקרם חולקו לקטגוריות ולא לערכים רציפים (Caywood et al., 2008). לעומת זאת, התייחסו למידת העומס הנמדדת כאל משתנה רציף בהינתן משימות זיכרון (Caywood et al., 2017). הם עשו זאת מכיוון שלטענתם, מכונה שלומדת צריכה להיות מאוד מדויקת ולשאוף לצמצם את הפער בין רמת העומס הנצפית לבין רמת העומס בפועל.

### 3. השערות המחקר

השערותינו המוצגות בסעיף זה, מבוססות על מחקרים דומים בנושא. להלן השערות המחקר שלנו:

1. קיים קשר בין מדדי דופק הלב המופקים מרצועת החזה לבין מידת העומס הקוגניטיבי.
2. שונות קצב הלב גדלה ככל שרמת קושי המשימה עולה.
3. השעה ביום אינה משפיעה על מדדי דופק הלב.
4. ניתן ליצור מודל באמצעות אלגוריתמי לימוד מכונה שידע לסווג את רמת העומס הקוגניטיבי בהסתמך על מדדי לב אובייקטיבים המופקים מרצועת החזה.

### 4. שיטה

#### 4.1 תיאור כללי

עקב מגפת הקורונה, נאלצנו לבצע שינויים במתודולוגית שיטת הניסוי והיקפו על מנת להתאימו למגבלות המוטלות עלינו בתקופה זו. חשבנו תחילה איך ניתן לעצב את הניסוי, מתוך הבנה



שהוא יתקיים בתנאים מגבילים הכוללים סגרים ברחבי ארץ ובשל כך הקושי לגייס מספר רב של נבדקים בשלבים המוקדמים של הפרויקט. לכן, תכננו את הניסוי בצורה כזו שכלול מספר גדול של משימות אשר יניב מספר רשומות גדול וכמות נתונים משמעותית לצורך ניתוח הנתונים.

על מנת לבחון את השערות המחקר בפרויקט, תוכנן ניסוי שבו נמדדו מדדי דופק הלב במשימות בעלות עומס קוגניטיבי שונה. הניסוי שתוכנן מתפרש על פני שלושה ימים, ובמהלך כל אחד מן הימים כל נבדק התבקש לבצע שלוש משימות שונות, כלומר כל נסיין ביצע 9 משימות סך הכול. המשימות הללו חולקו לשלוש קטגוריות לפי רמת העומס של כל אחת מהן- רמה נמוכה, בינונית וגבוהה. בכל יום בניסוי, כל משתתף ביצע שלוש משימות ברמת עומס קוגניטיבי שונה, כך שאורך כל משימה הוא 30 דקות. כל משימה בוצעה בחלק אחר של היום - בוקר, צהריים וערב. בטבלה 1 ניתן לראות את כלל המשימות על פי חלוקה לרמת העומס הקוגניטיבי אשר ביצעו הנסיינים.

משימות עומס גבוה	משימות עומס בינוני	משימות עומס נמוך
מדיטציה	פתרון סודוקו	מבחן פסיכומטרי
צפיה בטלוויזיה	פתרון תשחץ	מבחן פסיכוטכני
שמיעת מוזיקה	משחק מכוניות	מבחן IQ

טבלה 1 - משימות הניסוי

המשתתפים ביצעו את אותן משימות בדיוק. לדוגמה, במשימת הצפייה בטלוויזיה כולם צפו באותה תוכנית, ובמשימת התשחץ – כולם פתרו את אותו התשחץ. כך נעשה עבור כלל המשימות אשר ניתנו לנסיינים.

בכדי למנוע תלות בין רמת העומס של המשימות לבין פרק הזמן ביום בהן הן מבוצעות, הנבדקים ביצעו את המשימות בסדר שונה כך שכל משימה בעלת עומס קוגניטיבי שונה, בוצעה בפרק זמן שונה ביום - בוקר, צהריים וערב. כלומר, כל נסיין מבצע בכל יום סדר אקראי של משימות. בטבלה 2 ניתן לראות דוגמה לאחד מהסידורים האקראיים של המשימות כפי שבוצעו בניסוי.

זמן/יום הניסוי	1	2	3
בוקר	עומס נמוך	עומס גבוה	עומס בינוני
צהריים	עומס גבוה	עומס בינוני	עומס נמוך
ערב	עומס בינוני	עומס נמוך	עומס גבוה

טבלה 2 - סדר אקראי לביצוע משימות הניסוי

העומס הקוגניטיבי לא ניתן למדידה בצורה ישירה, ולכן יש לבצע מדידות שונות על מנת לקבל אינדיקציה על רמת העומס הקוגניטיבי אצל אדם מסוים. נושא זה מוכר היטב בספרות כפי שפורט בפרק הסקירה. המדידות הללו מספקות נתונים אובייקטיביים וסובייקטיביים לרמת העומס.

במהלך הניסוי בוצע איסוף נתונים אובייקטיביים להערכת העומס הקוגניטיבי באמצעות רצועת חזה המקליטה את דופק הלב, ומסוגלת באמצעות אפליקציה ייעודית לספק מידע על מספר מדדים הקשורים לקצב הלב. במהלך הניסוי הופקו מרצועת החזה שני מדדים: קצב הלב (HR) ו-אינטרוול RR. איסוף הנתונים הסובייקטיביים בוצע באמצעות שאלונים מוכרים מהספרות למדידת עומס קוגניטיבי נתפש – שאלוני SWAT אשר בוצעו כל 5 דקות במהלך המשימה, כלומר 6 פעמים במשימה, ושאלון NASA-TLX אשר בוצעו פעם אחת בסיום כל משימה.

קושי המשימות הנבחרות נבדק בפיילוט שנערך טרם תחילת הניסוי. לצורך הערכת הקושי של כל משימה, נבחרו עשרה סטודנטים שונים בגילאי 22-29 שאינם שייכים לקבוצת הניסוי אשר ביצעו את כל תשע המשימות, ולאחר כל משימה מילאו שאלון NASA-TLX להערכת העומס הנתפש. לאחר מכן, נלקחו התוצאות ובעזרתן השונו את דרגת העומס שנתפסה על ידינו לרמת הקושי הנתפסה אצל עשרת הסטודנטים. בטבלה 3 ניתן לראות את ממוצע ציוני שאלון ה-TLX כפי שמולאו ע"י הסטודנטים שהשתתפו בפיילוט.

משימה	ממוצע	קושי נתפש עי החוקרים	קושי נתפש עי הנבדק
הקשבה למוזיקה	18.63	נמוך	נמוך
מדיטציה	19.26	נמוך	נמוך
צפייה בטלוויזיה	17.15	נמוך	נמוך
סודוקו	59.36	בינוני	בינוני
תשחץ	66.46	בינוני	בינוני
משחק unblock car	55.10	בינוני	בינוני
מבחן פסיכוטכני	83.9	גבוה	גבוה
מבחן פסיכומטרי	86.66	גבוה	גבוה
מבחן IQ	83.34	גבוה	גבוה

טבלה 3 - תוצאות הפיילוט לסיווג רמת עומס המשימות

## 4.2 נבדקים

הניסוי כלל 9 משתתפים - חמישה בגילאי 23-27 מתוכם שני גברים ושלוש נשים, וארבעה בגילאי 50 ומעלה, מתוכם שני גברים ושתי נשים. טרם תחילת הניסוי, כל הנבדקים חתמו על

טופס הסכמה להשתתפות בו. שלושה משתתפים בניסוי מתוך התשעה הם צוות המחקר, הנבדקים הנוספים הם בני משפחה וחברים אשר נבחרו כמדגם בעקבות מגבלות מגפת הקורונה. כמו כן, בשל מגבלות המגפה נאלצנו לבחון כמות מצומצמת של נבדקים שביצעה מספר רב של משימות. על מנת לשמור על סודיות הנסיינים כל נסיין קיבל את אקראית בין A ל-I אשר מייצגת את הנתונים שלו (להלן "A", "B", "C", "I").

### 4.3 כלים

#### 4.3.1 רצועת החזה Polar H10

רצועת החזה בה השתמשנו בניסוי היא מהדגם Polar H10 המוצגת באיור 1. רצועת החזה היא משדר דופק מתקדם מבית חברת Polar המספקת מדידת דופק מדויקת.

תכונות עיקריות של רצועת החזה:

- רצועת החזה (Polar Pro Strap) עשויה בד רך ומתאימה עצמה למבנה הגוף
- הרצועה כוללת אלקטרודות המונעות הפרעות ומבטיחות מדידת דופק מדויקת יותר
- נתוני הדופק מסונכרנים עם אפליקציית SelfLoops HRV, כך שניתן לראות את נתוני הדופק בזמן אמת
- גודל המחבר: 65x1034 מ"מ
- משקל: מחבר 21 גר', רצועה 39 גר'
- סוג סוללה: CR 2025
- אורך חיי סוללה: 400 שעות



איור 1 - רצועת החזה Polar H10

הנתונים המתקבלים מהאפליקציה של רצועת החזה הם של מדדי RR, HR. הנתונים הללו מיוצאים באופן ישיר למסמך אקסל מההקלטה המתועדת באפליקציה של משתמשי הניסוי.

## 4.4 סביבת הניסוי

הניסוי התבצע בסביבה ביתית עקב מגבלות הקורונה, כך שבמהלך כל ניסוי, אחד מהחוקרים היה בבקרה דרך תוכנת הזום על המתרחש.

## 4.5 מהלך הניסוי

**בשלב הראשון** נעשה תדריך לקראת הניסוי. במהלכו, הועבר דף הסבר של הוראות הניסוי ([נספח 1](#)) לנבדק, הכולל תיאור של מהלך הניסוי והחתמת הנבדק על טופס הסכמה ([נספח 2](#)) להשתתפות בניסוי (פעם אחת בלבד בתחילת הניסוי). הועברו הנחיות כיצד ללבוש ולהפעיל את רצועת החזה וכיצד להפעיל את האפליקציה, והוצגו בפניו תקלות שייתכן שיתרחשו במהלך הניסוי. כמו כן, הוסברו הוראות לאיפוס הרצועה וניתן מענה על שאלות הנבדק. הנבדק התבקש ללבוש את רצועת החזה, להפעיל את האפליקציה הייעודית ולוודא שהיא עובדת באופן תקין.

**השלב השני** של הניסוי כולל את החלק המעשי כפי שפורט בסעיף השיטה. הנסיין ביצע 9 משימות באורך 30 דקות כל אחת, בדרגות עומס קוגניטיבי שונה, המתפרשות על פני שלושה ימים ובחלקים שונים של היום. לאחר כל 5 דקות בעת ביצוע המשימה, התבקש למלא שאלון SWAT ([נספח 3](#)). בסיום המשימה, התבקש למלא את שאלון NASA TLX ([נספח 4](#)). את תוצאות השאלונים ואת נתוני הדופק המופקים מהרצועה, הנבדק העביר לחוקרים באמצעות אימייל.

## 4.6 מדדים

### 4.6.1 מדדים אובייקטיביים ([נספח 16](#)).

1. **HR** – קצב הלב של הנבדק.
2. **RR** – הפרש הזמן בין שני גלי R סמוכים המציינים את המרווח בין פעימות הלב, נמדד באלפיות השנייה.
3. **HRV** – שונות קצב הלב
4. **MHR** – ממוצע קצב הלב
5. **MRR** – ממוצע של אינטרוולי ה-RR
6. **SDRR** – סטיית התקן של מקטעי אינטרוול ה-RR
7. **RMSSD** – השורש הריבועי של ממוצע ריבועי ההפרשים בין ערכי RR עוקבים
8. **SDNN** – השורש הריבועי של ממוצע ריבועי ההפרשים בין ערכי RR לממוצע RR באינטרוול
9. **HRmax-HRmin** – ההפרש בין קצב הלב הגבוה ביותר לנמוך ביותר

10. **Normalized HRV** – מדד HRV מחולק בממוצע קצב הלב של המשימה.
11. **HRV Moving Average** – שונות קצב הלב הכוללת ערכים מאינטרוול הזמן הקודם, מחולק בממוצע דופק הלב של המשימה.
12. **Last HR** – ממוצע קצב הלב מאינטרוול הזמן הקודם מחולק בממוצע דופק הלב של המשימה (אינטרוולים שונים של 5, 10 ו-15 שניות).

## 4.6.2 מדדים סובייקטיביים

1. **שאלון SWAT** ([נספח 4](#)) מגדיר סקאלה רב ממדית שמטרתה להעריך עומס עבודה שקיים על מפעיל. בשיטה זו, הגדירו שלושה ממדים לפיהם מעריכים עומס העבודה - עומס זמן, מאמץ מנטלי ולחץ פסיכולוגי. לכל ממד, הנבדק נותן ציון הנע בין 1 ל 3 כך ש-1 מהווה רמת עומס נמוכה, 2 מהווה רמת עומס בינונית ו-3 רמת עומס גבוהה. שאלון SWAT ממלא כל 5 דקות בעת ביצוע משימה, כלומר 6 פעמים. יתרון משמעותי של שאלון ה-SWAT הינו המהירות בה המשתתף יכול למלא אותו, ללא פגיעה בריכוז ובביצועי המשימה. מנגד, החיסרון של השאלון הוא שהוא לא מעמיק מספיק ועלולה ליצור אי דיוקים למצב הנתפס.

2. **שאלון NASA TLX** ([נספח 3](#)) הינה שיטה המגדירה סקאלה רב ממדית שמטרתה להעריך עומס עבודה שקיים על מפעיל בעת ביצוע משימה. השאלון בוחן 6 ממדים המשפיעים על עומס עבודה: דרישה מנטלית, דרישה פיזית, דרישת קצב עבודה, רמת התסכול, רמת המאמץ ורמת הביצועים. השאלון הוכח ככלי הנותן הערכה טובה מאוד למידת עומס העבודה, נוסף על העובדה כי הוא יחסית פשוט לביצוע. בתום כל משימה התבקשו הנבדקים למלא שאלון NASA TLX פעם אחת. היתרון של שאלון NASA TLX הוא רמת הדיוק הגבוהה בהערכת העומס הנתפס. יחד עם זאת, החיסרון הבולט של שאלון זה הינו הזמן הארוך אשר לוקח למלא אותו.

בנוסף לנתונים אשר התקבלו מרצועת החזה, הוספנו שתי עמודות תיוג אשר התקבלו על ידי השאלונים - אחת עבור שאלון NASA TLX ואחת עבור שאלון SWAT. תיוג שאלוני ה-SWAT יהיה עבור אינטרוול זמן של 5 דקות, כלומר, כל הנתונים באותו אינטרוול זמן מקבלים את אותו התיוג בשאלון ה-SWAT, ותיוג שאלוני NASA TLX יינתן בעבור כל המשימה כולה, מכיוון שהנסיין ממלא אותו פעם אחת בלבד בסיום המשימה.

## 4.7 ניתוח תוצאות

### 4.7.1 אופן ניתוח השאלונים

#### תוצאות שאלון SWAT

השימוש בשאלון SWAT ([נספח 4](#)) מצריך שלושה שלבים: ראשית, מוצגים בפני הנבדק 27 היגדים המייצגים את חשיבות ממדי העומס כפי שהם נתפסים על ידי הנבדק. את ההיגדים הללו כל נבדק צריך לסדר לפי חשיבותם בעיניו. בשלב השני בעת ביצוע המשימה, הנבדק מתבקש לתת ציון 1-3 לכל ממד (עומס זמן, מאמץ מנטלי ולחץ פסיכולוגי). לבסוף, בשלב השלישי, החוקרים לוקחים את הסידור שהתקבל מהשלב הראשון ובוחנים היכן מדורג הסידור הספציפי של הנבדק, כלומר מה סדר החשיבות של שלושת הממדים בעיניו, ולפיכך נקבע סולם הציונים עבורו. את התהליך הזה הנבדק מבצע פעם אחת בתחילת הניסוי ולא עבור כל משימה בנפרד. לאחר מכן, עבור כל משימה משקללים את ציון השאלון על סמך התשובות ל-3 השאלות. הציון הסופי הוא מספר בין 1-27.

#### תוצאות שאלון NASA TLX

שאלון NASA TLX ([נספח 3](#)) בנוי משישה מימדים, ועל הנסיין לדרג כמה המשימה דרשה ממנו בהתייחס לכל אחד מהמימדים. כדי לחשב את המשקל של כל ממד, מוצגים בפני הנבדק שילובי כל הזוגות האפשריים של המימדים ובכל פעם עליו לבחור אחד מבין הזוג שמשפיע יותר על עומס העבודה בעיניו. בשלב הבא, מכפילים את הערך נבחר במשקל שלו עבור כל מימד, וכך מחושב הציון. הציון הסופי הוא מספר בין 0-100.

### 4.7.2 תוצאות ומסקנות הפיילוט

את הפיילוט ביצעו שני הנבדקים שעשו את כל 9 המשימות בשלושה ימים. באמצעות תוכנת הזום, ווידאו שהנבדקים עומדים בזמני המשימות וממלאים את השאלונים. מטרת הפיילוט הייתה לבדוק את טיב הנתונים המתקבלים מהרצועה, תקלות שעוללות לצורך במהלך הניסוי, בחינת אופן השימוש ברצועת החזה והאפליקציה וכן בדיקה ראשונית של העברת הניסוי מרחוק באמצעות תוכנת זום.

במהלך הפיילוט בוצעו עדכונים ותיקונים למודל הניסוי המתוכנן, והועלו מס' נקודות לשיפור:

- כמות שאלוני ה SWAT הייתה קטנה – תחילה, כל עשר דקות היה על המשתתפים למלא שאלון (סה"כ 3 פעמים במשימה). לאחר מכן, הגדלנו את כמות השאלונים ל-

- 15, כלומר מילוי שאלון כל שתי דקות. כמות השאלות והתדירות התישה את הנבדקים ולכן הורדנו אותם ל-6 סה"כ (כל חמש דקות מילוי שאלון)
- חלק מהנבדקים סיימו את המשימות בפחות מחצי שעה במשימות הסודוקו והתשחץ, ולכן הארכנו אותן
  - במהלך הניסוי גילינו שיבושים ותקלות שקורות בזמן לבישת הרצועה ולמדנו לבצע שימוש מיטבי בעזרים, והעברנו אותם במהלך התדריך לנסיינים

### 4.7.3 עיבוד הנתונים

כלל עיבוד הנתונים בוצע בסביבת python באמצעות חבילת התוכנה pandas ובסביבת R .

- על מנת לחשב את הממדים האובייקטיביים שהוצגו במשימות הניסוי נדרש היה לבחור מקטע זמן קבוע ומדויק ברמת מיל' שניה. חישוב מדויק זה ניתן לביצוע בעזרת מדד RR אשר מספק מידע על משך הזמן בין פעימה לפעימה במילי-שניה.
- טרם הבחירה באינטרוול הזמן הרצוי עבורו חישבנו את הממדים האובייקטיביים, בדקנו את כמות הרשומות המופקת מאינטרוולי זמן של 60 שניות, 30 שניות ו-15 שניות. הנחנו כי נהיה מדויקים יותר אם ננתח את נתוני המשימות ברזולוציית זמן קטנה ככל האפשר, ובהתאם נפיק את כמות הנתונים הגדולה ביותר מאינטרוול הזמן של 15 שניות. שיקול נוסף בבחירת ניתוח 15 שניות היה הגדלת מספר הרשומות עבור שיפור מודל לימוד המכונה. באינטרוול זמן מדויק של 15,000 מילי-שניות (15 שניות) קיבלנו כ-1,100 רשומות של נתונים עבור כל נסיין ובסך הכול כ-10,000 רשומות לכולם
- ביצענו ניתוח בסביבת R לנתונים על מנת להבין את התנהגותם. בחנו את טווח הערכים ([נספח 5](#)), בנוסף בחנו תרשים קופסא ([נספח 6](#)) של כלל הממדים בעבור כל נסיין וכן את טעות התקן בכל אחד מהממדים ([נספח 7](#)). שמנו לב כי קיימת חפיפה במשימות בעלות העומס הקוגניטיבי הבינוני והגבוה.
- בוצע חישוב 14 ממדים אובייקטיביים על סמך שני מדדי ה-RR וה-HR שהתקבלו באופן ישיר מרצועת החזה. החישוב בוצע ע"י קוד בסביבת python ([נספח 8](#))
- ביצענו ניפוי חריגים של 2.5% מנתוני הדופק הגבוהים של כל משימה. עשינו זאת משום שראינו שקיימת חריגה אשר יכולה להטות את מגמות הדופק בצורה לא מייצגת, עקב קפיצות חריגות באפליקציה, וזמן הסתגלות של הנבדק למשימה שהתבטא בדופק גבוה מהרגיל
- ביצענו מבחני נורמליות לכלל הממדים שנבחנו במחקר. בדיקה זו נעשתה באמצעות מבחני Shapiro – Wilk ותרשימי QQ-plot בסביבת R ([נספח 9](#)). כחלק מבחינת התפלגות הממדים המחושבים, נמצא כי חלקם לא מתפלגים נורמלית. לכן, בחנו מספר

טרנספורמציות אפשריות למדדים אלו. בעבור המדדים SDRR, SDNN ו- RMSSD בוצעה טרנספורמצית log .

- כדי לבנות את מודל לימוד המכונה, חילקנו את הנתונים לשני סטים – סט אימון המורכב מ-6 משימות אקראיות (2 מכל דרגת עומס קוגניטיבי), ולסט בחינה המורכב מ-3 משימות אקראיות (המשימות הנותרות, אחת מכל דרגת עומס קוגניטיבי) אשר המודל אינו מכיר. חלוקה זו יצרה אי-תלות בין הנתונים אשר המודל לומד מהם לבין הנתונים אשר עליהם התבצעה הבדיקה של טיב המודל
- במקביל בדקנו מודלים אפשריים של לימוד המכונה בהתבסס על Random Forest (נספח 11) ו-Multinomial Regression (נספח 12) לבחינת המדדים הטובים ביותר ע"פ מובהקותם ותרומתם לאחוזי הניבוי של המודל. נמצא כי קיימים שני מדדים רגישים: MHR ו-MRR ושלושה מדדים נוספים אשר תרומתם משמעותית: Log SDRR, Log RMSSD, Log SDNN
- ביצענו את כלל החישובים של המדדים עבור כל נסיין בנפרד, ובנוסף יצרנו סט מאוחד הכולל את נתוני כלל הנסיינים (להלן "All"). זאת במטרה לייצר מודל לימוד מכונה גנרי לסיווג העומס הקוגניטיבי בעבור כל אדם

## 5. תוצאות

**השערה ראשונה: קיים קשר בין מדדי דופק הלב הנמדדים מרצועת חזה לבין העומס הקוגניטיבי.**

בחנו קשר זה על ידי מקדם המתאם של ספירמן , מתאם זה מתאר קשרים ליניאריים. ניתן לראות בטבלה 4 את ערך המתאם בין המדדים הסובייקטיביים (NASA TLX, SWAT) והאובייקטיביים הנבחרים (MHR, MRR, Log SDRR, Log RMSSD, Log SDNN, HRV).

ב

טבלה 5 ניתן לראות את המובהקות של הקשרים המתוארים בטבלה 4. ניתן לראות בטבלה 4 בעמודה השנייה משמאל (SWAT) ובשורה השנייה (MRR) מקדם ספירמן של 0.28- ובהתאמה בטבלה 5 את המובהקות הגבוהה של הקשר הזה. ממצא זה מייצג את הקשר שאינו חזק אך המובהק אשר נמצא בין המדדים של דופק הלב ולבין העומס הקוגניטיבי בהשערת המחקר .

כמו כן, בטבלה 4 בעמודה שמינית משמאל (MRR) ושורה שנייה (MHR) רואים את הקשר החזק בעל מתאם של 0.86- ובעל מובהקות חזקה בתאמה בטבלה 5. בנוסף לכך, בטבלה 4 בעמודה ראשונה משמאל (NASA TLX) ושורה שביעית (SWAT) ניתן לראות מתאם אשר אינו שואף ל-1 כפי שציפינו במחקר לקשר בין השאלונים הסובייקטיביים.



NASA-TLX	SWAT	HRV	Log_SDNN	Log_RMSSD	Log_SDRR	MHR	MRR	Cor Table
-0.19	-0.28	0.01	0.18	0.25	0.23	-0.86		MRR
0.20	0.28	0.03	-0.10	-0.17	-0.09		-0.86	MHR
-0.07	-0.03	0.13	0.55	0.49		-0.09	0.23	Log_SDRR
-0.08	-0.05	0.07	0.33		0.49	-0.17	0.25	Log_RMSSD
-0.06	-0.04	0.08		0.33	0.55	-0.10	0.18	Log_SDNN
-0.01	0.00		0.08	0.07	0.13	0.03	0.01	HRV
0.73		0.00	-0.04	-0.05	-0.03	0.28	-0.28	SWAT
	0.73	-0.01	-0.06	-0.08	-0.07	0.20	-0.19	NASA-TLX

טבלה 4 - קורלציית ספירמן בין המדדים

NASA-TLX	SWAT	HRV	Log_SDNN	Log_RMSSD	Log_SDRR	MHR	MRR	P Value Table
0.0E+00	0.0E+00	1.5E-01	0.0E+00	0.0E+00	0.0E+00	0.0E+00		MRR
0.0E+00	0.0E+00	2.7E-03	0.0E+00	0.0E+00	0.0E+00		0.0E+00	MHR
3.4E-11	1.2E-03	0.0E+00	0.0E+00	0.0E+00		0.0E+00	0.0E+00	Log_SDRR
8.9E-16	3.3E-06	2.7E-12	0.0E+00		0.0E+00	0.0E+00	0.0E+00	Log_RMSSD
3.7E-09	3.9E-05	2.6E-14		0.0E+00	0.0E+00	0.0E+00	0.0E+00	Log_SDNN
4.5E-01	7.0E-01		2.6E-14	2.7E-12	0.0E+00	2.7E-03	1.5E-01	HRV
0.0E+00		7.0E-01	3.9E-05	3.3E-06	1.2E-03	0.0E+00	0.0E+00	SWAT
	0.0E+00	4.5E-01	3.7E-09	8.9E-16	3.4E-11	0.0E+00	0.0E+00	NASA-TLX

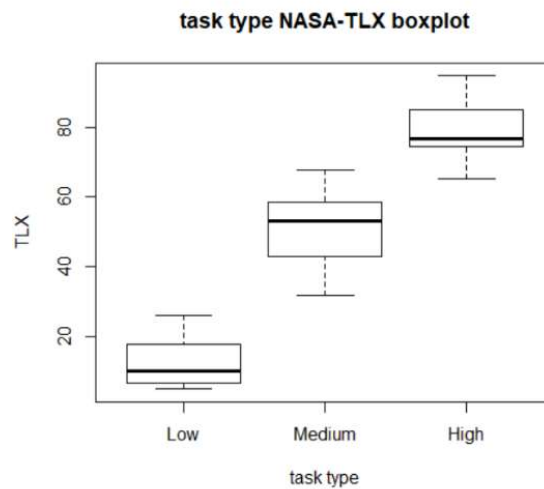
טבלה 5 - מובהקות קורלציית ספירמן

בדקנו את ההתאמה בין סוג המשימה (משימות בעלות עומס קוגניטיבי שונה – נמוך בינוני וגבוה) לבין שני המדדים הסובייקטיביים. מטרת שאלוני SWAT ו-NASA TLX היא לבחון את רמת עומס הקוגניטיבי הנתפס על ידי הנבדק בעת ביצוע משימתו.

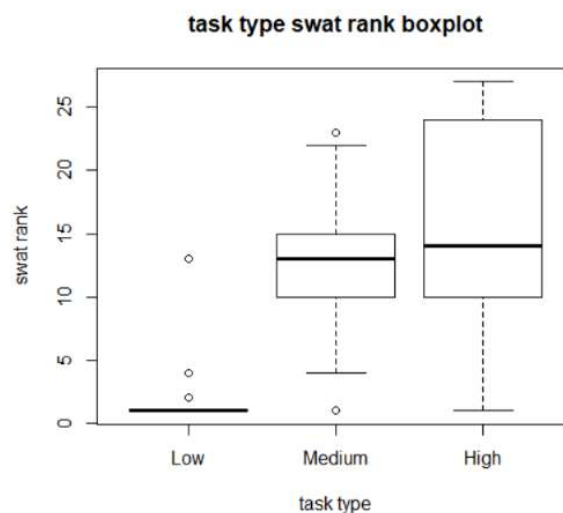
באיור 2 - תוצאות שאלון NASA-TLX לפי סוג משימה תרשים הקופסא של ציון שאלון NASA-TLX וסוג המשימה (Low, Medium, High) ניתן לראות הפרדה בולטת בין סוגי המשימות השונים וטווח הציונים שניתן עבור כל קטגוריה. ניתן לראות קשר בין העומס הקוגניטיבי שכל נבדק הרגיש בתום כל משימה שביצע לבין רמת העומס של המשימה אותה ביצע.

באיור 3 בתרשים הקופסא של ציון שאלון SWAT וסוג המשימה (Low, Medium, High) ניתן לראות התאמה בין ציון השאלון לבין המשימות הקלות. לעומת זאת, השאלון פחות רגיש כאשר מסתכלים על המשימות הבינוניות והקשות וניתן לראות כי קיימת חפיפה בערכים בניהם.

ייתכן כי תוצאות אלו נובעות מהעבודה ששאלון ה-SWAT פחות רגיש ומדויק משום שהוא ממולא מהר יותר במהלך הניסוי לעומת מילוי עמוק ממוקד ואיטי יותר של שאלון ה-TLX הממולא בסוף הניסוי.



איור 2 - תוצאות שאלון NASA-TLX לפי סוג משימה



איור 3 - תוצאות שאלון SWAT לפי סוג משימה

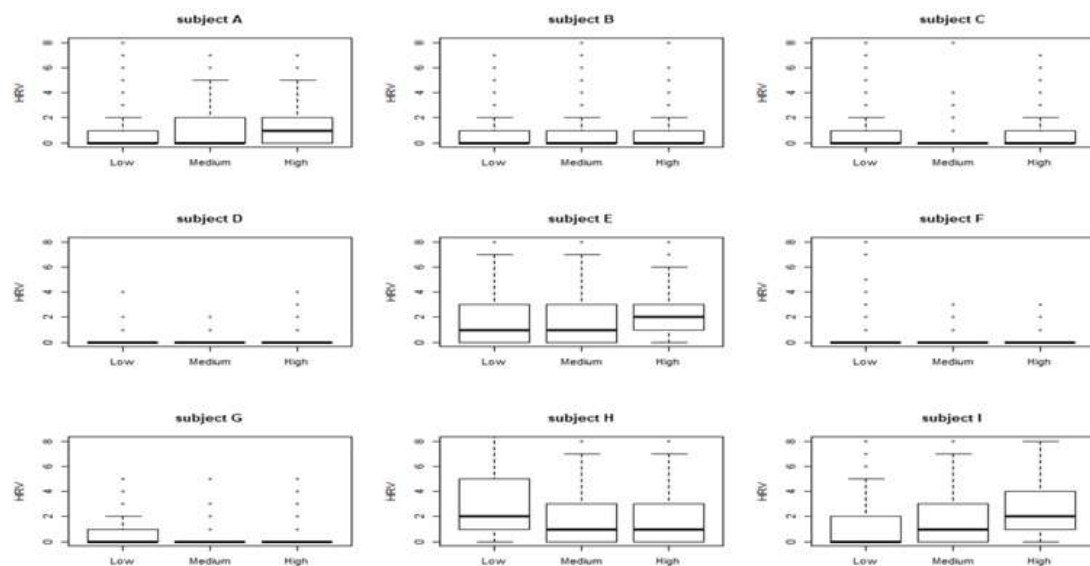
### השערה שנייה: שונות קצב הלב גדלה ככל שרמת קושי המשימה עולה

השערה זו מניחה שבמשימות בעלות עומס קוגניטיבי שונה, קיימים הבדלים בשונות קצב הלב. במשימה בעלת עומס קוגניטיבי נמוך שיערנו כי שונות קצב הלב היא נמוכה, וזאת בניגוד למשימות בעלות עומס קוגניטיבי גבוה עבורן שיערנו כי שונות קצב הלב גבוהה.

השתמשנו בתרשימי קופסא (איור 4 - תרשים קופסא - שונות קצב הלב) המתאר בעבור כל נסיין את ערכי שונות קצב הלב (HRV) בחלוקה על פי סוג רמת העומס הקוגניטיבי (Low, Medium, High) במשימה, וזאת על מנת לקבל תמונה ויזואלית של הנתונים של כל אחד מן הנבדקים שלנו. ניתן לראות בתרשים את החפיפה בערכי המדד בין נתוני המשימות השונות

בעלות עומס שונה. לדוגמה אצל נסיין F ניתן לראות כי שונות דופק הלב שאפה לאפס בכל האחת מהמשימות בכל סוג של עומס קוגניטיבי אשר ביצע.

כמו כן, ביצענו מבחן סטטיסטי ANOVA (טבלה 6) בו נבחנה ההשערה שאומרת שאין הבדל בשונות קצב הלב בין רמות עומס המשימות השונות. מצאנו כי לא קיימת מובהקות לדחות את השערת האפס אצל רוב הנסיינים, כלומר לא קיים שוני בשונות קצב הלב בין משימות עם עומס קוגניטיבי שונה, וזאת בניגוד להנחת המחקר השנייה. ניתן לראות כי עבור נסיינים B C D H I לא קיימת שונות במדד ה-HRV על פי המבחן.



איור 4 - תרשימים קופסא - שונות קצב הלב

Subject	P-value	F-Anova
A	0.08	2.5
B	<b>0.007</b>	4.93
C	<b>1.3e-05</b>	11.37
D	<b>0.008</b>	4.78
E	0.329	1.113
F	0.355	1.036
G	0.081	2.51
H	<b>5.64e-05</b>	9.87
I	0.53	0.635

טבלה 6 - מבחן ANOVA לשונות קצב לב

באמצעות מבחן Tukey (טבלה 7) בחנו את ההשוואות השילובים של כל סוג עומס משימה (Low, Medium, High) אצל הנסיינים בהם נמצא שוני של מדד הHRV על מנת לבחון בין איזה סוג רמת עומס של משימה קיים הבדל מובהק בשונות קצב הלב. המבחן הראה כי רב השונות הקיימת בקצב הלב הינה בין המשימות בעלות העומס הבינוני והנמוך וכי קיימת גם שונות בין המשימות בעלות העומס הנמוך והגבוה.

Subject	HRV-Type	P-value
B	Low-High	0.391
	Medium-High	0.162
	<b>Medium-Low</b>	<b>0.005</b>
C	Low-High	0.818
	<b>Medium-High</b>	<b>3.70E-05</b>
	<b>Medium-Low</b>	<b>4.79E-04</b>
D	<b>Low-High</b>	<b>0.009</b>
	Medium-High	0.057
	Medium-Low	0.807
H	<b>Low-High</b>	<b>3.57E-05</b>
	Medium-High	0.244
	<b>Medium-Low</b>	<b>0.017</b>

טבלה 7 - מבחן Tukey לשונות קצב לב

### השערה שלישית: השעה ביום אינה משפיעה על מדדי דופק הלב

בהשערה זו הנחנו כי השעה ביום בה מבוצע המשימה אינה משפיעה על המדדים האובייקטיבים אותם חישבנו. כלומר, אין הבדל בין המדדים במשימות שבוצעו ביחס לזמנים השונים ביום (בוקר, צהריים וערב).

השתמשנו בתרשימי קופסא ([נספח 10](#)) בכדי לקבל תמונה ויזואלית של הנתונים עבור המדדים האובייקטיביים. בתרשימים אלה נבחנה השעה ביום בכלל הנתונים לעומת מדדי כל אחת מהמשימות בעלת עומס שונה בנפרד. ברוב התרשימים לא ניתן להבחין בצורה ויזואלית בהבדלים ברורים בין השעות השונות ביום.

בנוסף לכך, ביצענו מבחן ANOVA ( ) טבלה 7) בו אנחנו בוחנים את ההשערה, והתקבלו תוצאות שונות בחלקן מההתרשמות הראשונית של תרשימי הקופסא. ניתן לראות כי במדד MRR ובמדד MHR ניתן לדחות את

השערת האפס כי לא קיימת שונות וניתן לומר כי מדדים אלה כן מושפעים מהזמן ביום ב בוצעה המשימה.

מדד	P-value	F-Anova
MRR	<b>0.02</b>	3.88
MHR	<b>0.03</b>	3.43
Log SDRR	0.06	2.76
Log RMSSD	0.93	0.06
Log SDNN	0.13	2
HRV	0.1	2.28

טבלה 8 - מבחן ANOVA זמן ביום

בדקנו באמצעות מבחן Tukey (טבלה 9) המאפשר ההשוואות כל זוגות פרקי הזמן ביום על מנת לבחון באיזה מהם קיים הבדל מובהק במדדים אשר התקבל בהם שוני- MRR ו MHR. ניתן לראות שבמדד MRR קיים הבדל מובהק בין המשימות שבוצעו בבוקר לאלו המבוצעות בצהריים. לעומת זאת במדד MHR ניתן לראות שההבדל הוא בין המשימות שבוצעו בצהריים לעומת אלו המבוצעות בערב.

Subject	HRV-Type	P-value
MRR	Morning-Evening	0.499
	Noon-Evening	0.232
	<b>Noon-Morning</b>	<b>0.015</b>
MHR	Morning-Evening	0.656
	<b>Noon-Evening</b>	<b>2.70E-02</b>
	Noon-Morning	2.04E-01

טבלה 9 - מבחן Tukey זמן ביום

השערה רביעית: ניתן לייצר מודל באמצעות אלגוריתם לימוד מכונה שידע לסווג את רמת העומס של המשימה המבוצעת

נבחנו שני מודלים של לימוד מכונה: Multinomial Regression ו Random forest.

מטרת המודל היא לנבא את רמת העומס של המשימה אותה ביצע הנסיין כאשר יש שלוש קטגוריות לסיווג העומס הקוגניטיבי – נמוך, בינוני וגבוה.

במהלך הפרויקט בחנו שתי אפשרויות של מודל לסיווג. הראשונה היא מודל מאוחד הכולל את נתוני כלל הנבדקים, ומטרתו הינה לסווג את רמת העומס הקוגניטיבי עבור כל אדם, מתוך שאיפה לקבל מודל שיתאים לכולם (להלן - "All"). השנייה היא מודל עבור כל אדם באופן פרטני (להלן - "I"... "B", "A"). במצב זה, המודל יהיה מותאם לכל אדם וידרוש התאמה אישית.

תחילה, בדקנו מודל המכיל את כל המדדים שאותם חישבנו, ולאחר מכן הסרנו משתנים כך שנשארו עם אלה שמספקים לנו את אחוזי ניבוי הגבוהים ביותר. את הסרת המשתנים במודל Random Forest ביצענו באמצעות feature importance המספק את החשיבות של כל המדדים למודל ([נספח 11](#)). במודל Multinomial Regression סינון המשתנים בוצע לפי ערך ה-P-value ([נספח 12](#)). לאחר מכן, בחרנו מודל המכיל חמישה מדדים, המובהקים ביותר: SDNN, MRR, MHR, SDRR, RMSSD באמצעות אחוזי הניבוי של המודל היו הגבוהים ביותר ([נספח 13](#), [נספח 14](#)).

בשלב הבא, ניסינו לשפר את המודל ע"י הסרה של שני מדדים בעלי המובהקות הנמוכה ביותר וקיבלנו את המודל בעל התכונות SDRR, MHR, MRR ([נספח 11](#), [נספח 12](#)), ובכל אחד מהמודלים Random Forest ו Multinomial Regression ראינו כי אחוזי הניבוי של המודל עם שלושת המדדים אינו מגדיל את אחוזי הניבוי ([נספח 13](#), [נספח 14](#)). בעקבות תוצאה זו החלטנו לשפר את המדדים SDNN, RMSSD, SDRR אשר יצאו הכי פחות מובהקים בכל אחד משני המודלים. עשינו זאת באמצעות טרנספורמציה לוגריתמית על מנת לגרום לנתונים להתפלג נורמלית. לאחר בחינה נוספת של שני המודלים, ראינו כי אחוזי הניבוי לא עלו בצורה משמעותית ([נספח 13](#), [נספח 14](#)).

שמנו לב במהלך הניסיונות הרבים שביצענו כי המודל (3 Types) המסווג שלוש קטגוריות של עומס קוגניטיבי הצליח לסווג בצורה טובה את המשימות הקלות, אך התקשה להבדיל בסיווג בין המשימות בעלות העומס הבינוני והקשה ([נספח 15](#)). ביצענו צעד נוסף, והוא לאחד את שתי קטגוריות משימה – עומס בינוני ועומס גבוהה (2 Types), ובחנו מחדש את המודלים. בטבלה 10 ניתן לראות את תוצאות הניבוי עבור מודל Random Forest ובטבלה 11 ניתן לראות את תוצאות הניבוי עבור מודל Multinomial Regression. תוצאות המוצגות בטבלאות אלה מראות כי לאחר איחוד קטגוריות המשימות בעלות העומס הבינונית והגבוהה במודל עלו אחוזי הצלחת הניבוי של הסיווג במודל לכל נבדק ולכלל הנבדקים במודל האחד.

Subject	2 Types Score	3 Types Score
A	51%	76%
B	67%	85%
C	51%	69%
D	63%	66%
E	49%	51%
F	50%	60%
G	51%	61%
H	46%	53%
I	70%	90%
All	46%	61%

טבלה 10 - אחוזי ניבוי המודל Random Forest

Subject	2 Types Score	3 Types Score
A	52%	71%
B	55%	79%
C	43%	61%
D	44%	51%
E	49%	50%
F	32%	59%
G	46%	55%
H	47%	51%
I	64%	84%
All	42%	57%

טבלה 11 - אחוזי ניבוי המודל Multinomial Regression

מודל בעל 3 הקטגוריות של רמת העומס - נמוכה בינונית וגבוהה - 3 Types Score\*

מודל בעל 2 קטגוריות - לאחר איחוד המשימות קשות והבינונית - 2 Types Score\*

## 6. דיון ומסקנות

מטרת הפרויקט הינה לחקור את מדדי דופק המופקים מרצועת החזה ולמצוא מהו הקשר בין מדדים אלו לבין העומס הקוגניטיבי.

**השערת המחקר הראשונה** הייתה שקיים קשר בין מדדי דופק הלב הנמדדים מרצועת חזה לבין העומס הקוגניטיבי שמרגיש הנבדק. בהתאם להשערת המחקר, מצאנו קשר בין המדדים האובייקטיבים אשר חישבנו לבין המדדים אשר מייצגים את העומס הקוגניטיבי הסובייקטיבי. הממצאים הראו כי קיים קשר מובהק, אך חלש. מצד אחד הממצאים תאמו את הציפייה שישנו קשר, אך מנגד הם לא תאמו את הציפייה כי הקשר יהיה חזק. היינו סבורים כי הנתונים הם בעלי קשר ליניארי לעומס הקוגניטיבי המיוצג על ידי השאלונים הסובייקטיביים, ובהתאם לכך השתמשנו במתאם ספירמן לבדיקה, אך ייתכן שהקשר אינו ליניארי.

הקשר והמובהקות הבולטת בין מדדי MHR ו- MRR, שהם המדדים שנמצאו הרגישים ביותר, לבין המדדים הסובייקטיביים SWAT ו- NASA TLX מחזקים את השערת המחקר הראשונה.

עוד עולה מן הנתונים כי קיימת קורלציה בולטת יותר בין מדדי MRR ל MHR מכיוון שהאחד מתאר את קצב הלב בעוד השני מתאר את הזמן העובר בין פעימה לפעימה.

בנוסף, כאשר מסתכלים על המתאם בין המדדים הסובייקטיביים נמצא כי קיים מתאם שאינו שואף ל-1, בניגוד לציפיות שלנו. ייתכן שהמתאם הוא כזה מכיוון שיש חפיפה בין תוצאות שאלוני ה-SWAT עבור הקטגוריות הבינונית והגבוהה. בנוסף, שאלון ה-SWAT מעודד מענה מהיר וללא התעמקות, בעוד ששאלון ה-TLX מדויק יותר ומצריך התעמקות וזמן נוסף ביחס לשאלון ה-SWAT.

כחלק מתהליך ניתוח הנתונים, בלטה החפיפה הקיימת בין מדדי דופק הלב במשימות בעלות עומס קוגניטיבי גבוה ובינוני. ייתכן שדבר זה נבע מעיצוב לא מיטבי של משימות הניסוי, וגם ייתכן שהדופק לא השתנה בצורה מובהקת ולכן המדדים חפפו.

כמו כן, באמצעות בחינת השערה זו הצלחנו לתקף את רצועת החזה ככלי מדידת קצב לב נוח לשימוש וזול. בשונה מכלי המדידה המוכרים, רצועת החזה מסוגלת לשרת אותנו בבחינת העומס הקוגניטיבי בצורה מיטבית הן בפן המחקרי והן בפן העסקי.

**השערת המחקר השנייה** עוסקת בכך ששונות קצב הלב גדלה ככל שרמת קושי המשימה עולה. מצאנו כי עבור רוב הנבדקים אין הבדל בשונות קצב הלב ככל שרמת קושי המשימות עולה, מה שנוגד את השערתנו. בתרשימי הקופסא לא ניתן היה להבחין בוודאות בשוני בין הנבדקים אך מבחן Anova הראה כי קיים הבדל בשונות קצב הלב אצל חלקם. עבור נבדקים A, E, F, G, I קיבלנו כי ערך ה P-value שלהם גדול מ-0.05. כלומר, ניתן להסיק כי אין הבדל בשונות קצב הלב ככל שרמת קושי המשימות עולה אצל נבדקים אלו. לעומת זאת, ניתן לראות שהתוצאה לא מובהקת אצל ארבעת הנבדקים B, C, D, H ויותר סביר כי קיים הבדל בשונות קצב הלב. לאחר מכן, מבחן Tukey המאפשר בדיקת כל ההשוואות הזוגיות בין כל סוג משימה הראה כי שונות קצב הלב אצל רב הנבדקים B, C, H, D קיים הבדל מובהק בין המשימות בעלות עומס בינוני ונמוך. נתון נוסף שקיבלנו הוא שבנבדקים D ו H קיים הבדל בשונות קצב הלב במשימות בעלות העומס הנמוך וגבוה. מסקנתנו היא ששונות קצב הלב לא הייתה מדד מספיק חזק מבחינת המובהקות שלו ותרומתו לאחוזי הניבוי כפי ששיערנו.

**השערת המחקר השלישית** שנבדקה הינה שהשעה ביום אינה משפיעה על מדדי דופק הלב. ביצענו הליך בדיקה זהה להשערת המחקר השנייה, והצלחנו לאמת את נכונות ההשערה באמצעות תוצאות מבחן Anova עבור רוב המדדים. מצאנו כי קיימת קורלציה חלשה בין השעה



ביום לבין מדדי דופק הלב. לאור תוצאות המבחן ותרשימי הקופסא, מצאנו כי קיימת קורלציה חלשה בין השעה ביום לבין מדדי דופק הלב HRV, Log SDNN, Log SDRR, Log RMSSD, מכיוון שערכי P-value יצאו גדולים מ-0.05. לכן, לא קיים הבדל מובהק בין המדדים. לעומת זאת, קיבלנו תוצאה שונה במדדי MRR ו-MHR כלומר סביר כי קיימת השפעה של שעה ביום בה מבוצעת המשימה על מדדים אלו. לאחר ביצוע מבחן Tukey המאפשר בדיקה של כל צמדי הזמנים ביום ראינו כי במדד MRR קיים הבדל מובהק בין המשימות שבוצעו בבוקר לאלו המבוצעות בצהריים. לעומת זאת, במדד MHR ראינו שההבדל הוא בין המשימות שבוצעו בצהריים לעומת אלו המבוצעות בערב. אנו משערים כי תוצאות אלו נבעו מהסיבה שלא הגדרנו מראש שעה מדויקת ביום בה תבוצע כל משימה וכן ע"י כך ששני המדדים הללו רגישים יותר וייתכן כי רגישים גם לפרק הזמן ביום בו המשימה מבוצעת.

**בהתייחס להשערת המחקר הרביעית**, בדקנו האם ניתן לייצר מודל באמצעות אלגוריתם לימוד מכונה שידע לסווג את רמת העומס של המשימה המבוצעת. נבחנו שני מודלים שונים של לימוד מכונה, Multinomial Regression ו-Random forest. תחילה נבדק מודל המכיל את כל המדדים שאותם חישבנו, ולאחר מכן סוננו המשתנים כך שנשארו עם המדדים שיתנו לנו את אחוזי ניבוי הגבוהים ביותר. לאחר בחירה של חמישה מדדים מובהקים MRR, MHR, SDRR, RMSSD ו-SDNN ולאחר מכן שלושה מדדים מובהקים MRR, SDRR, MHR ו-RMSSD בכל אחד מהמודלים ראינו כי אחוזי הניבוי נמוכים וגם לאחר ביצוע טרנספורמציה על המדדים שיצאו פחות מובהקים בכל מודל - SDNN, RMSSD ו-SDRR. שמנו לב כי המודל הצליח לסווג בצורה טובה את המשימות בעלות העומס הנמוך אך התקשה להבדיל בסיווג בין המשימות בעלות העומס הבינוני והגבוה. לאחר איחוד הקטגוריות ובחינה מחדש של המודל, ראינו כי עלו אחוזי הצלחת הניבוי לכל נבדק, ותוצאות הניבוי של מודל Random Forest היו טובות מזה של Multinomial Regression. אנו משערים כי מודל זה טוב יותר משום שהקשר בין נתוני המדדים לרמת העומס הקוגניטיבי לא בהכרח ליניארי, כפי שמודל Multinomial Regression מניח. לכן, הסקנו כי ניתן לייצר מודל באמצעות אלגוריתם לימוד מכונה שידע לסווג את רמת העומס של המשימה המבוצעת.

## 7. סיכום והמלצות

מגפת הקורונה ומגבלותיה פתחו בפנינו דלת למחקר חדש וכן להתנסות ראשונה בביצוע ניסוי מרחוק, כך שהתפעול שלו התבצע על ידי הנבדקים עצמם. בסיום תהליך ניתוח הנתונים והסקת המסקנות, ברצוננו לציין מספר המלצות.

תחילה, יש לבחון מחדש משימות בעלות עומס קוגניטיבי בינוני וגבוה ולתכנן את הניסוי כך שהבדלי העומס בין המשימות יהיה מובהק יותר. כמו כן, נמליץ להשתמש במספר רב יותר של

נבדקים בניסוי לשם הגעה למסקנות מבוססות יותר. לאחר בחינת מודלי לימוד המכונה וסיווג רמות העומס, גילינו שכמות הנתונים הייתה נמוכה וייתכן כי כמות גדולה יותר של נתונים תוביל לשיפור באחוזי הניבוי של המודלים.

בנוסף, יתכן כי המשימה בעלת העומס הגבוה לא עוצבה בצורה מיטבית מכיוון שלא נבדקו ביצועי המשימה. בניסוי שתכננו, לא ניתן היה למדוד ביצועים לכלל המשימות ולכן בחרנו לוותר על מדד זה ועל בחינת השפעתו על העומס. לכן, לדעתנו יהיה מעניין להוסיף מדד ביצועים לכל משימה וייתכן שהוא יסייע בהבחנה בין רמות העומס השונות במשימות בעלות העומס הבינוני והגבוה.

נציין שבפרויקט זה השתתפו מגוון אנשים - חמישה בגילאי 23-27, וארבעה בגילאי 50 ומעלה. ייתכן כי הפרש הגילאים השפיע על העומס הנתפס על ידי הנסיין במשימות מסוימות, למשל משחק מכוניות בפלאפון. לדעתנו, יהיה מעניין לבצע את הניסוי על קבוצות אוכלוסייה אחרת ולבדוק האם תהיה קורלציה גבוהה יותר בין המדדים, וכן אחוזי ניבוי טובים יותר במודל. לסיכום, נרצה להמליץ על מחקר המשך אפשרי – פיתוח מודל לימוד מכונה המבוסס על מידע המגיע נבדקים רבים, שיוצר גרסה בסיסית עליה מבוצע אתחול קצר לצורך היכרות עם קצב הלב של הנסיין. ייתכן שמודל זה ייצור באופן מיטבי הבחנה בעומס הקוגניטיבי שאותו חווה כל אחד ואחת.

בנימה אישית, הפרויקט היווה עבורנו הזדמנות ללמידה משמעותית על אופן ביצוע מחקר אקדמי. למדנו על תחום רלוונטי אשר משתלב בנושאים רבים שנלמדו במהלך התואר הראשון. השפעת העומס הקוגניטיבי על תפקוד האדם הוא תחום רלוונטי הנחקר במשך שנים רבות, מתוך מטרה למצוא את השפעת המשימה על האדם כדי לא לפגוע בביצועים, מה שרלוונטי גם באקדמיה וגם בתעשייה. במהלך ביצוע הפרויקט התמודדנו עם אתגרים רבים, חומר חדש ולמידה עצמית מרובה של הנושא וכלי המחקר. ביצענו ניתוחים סטטיסטיים, השתמשנו באלגוריתמי לימוד מכונה ובכלים רבים שנלמדו תוך פיתוח בסביבת Python ו-R. נאלצנו להתמודד עם שינויים בתוכנית המקורית בעקבות מגפת הקורונה, אך ידוע לנו כי שינויים הם חלק בלתי נפרד מהפרויקט, ולמרות זאת הצלחנו להתמודד איתם. נסיים בהוקרת תודה למנחים שתמכו והכווינו אותנו לאורך כל הדרך – פרופסור טל אורון-גלעד ודוקטור יובל זק.

Abu-Mostafa, Y. S., Magdon-Ismail, M., & Lin, H. T. (2012). Learning from data (Vol. 4). New York, NY, USA:: AMLBook.

Bailey, K. (1994). Numerical taxonomy and cluster analysis. *Typologies and Taxonomies*, 34, 24.

Baldwin, C. L., & Penaranda, B. (2012). Adaptive training using an artificial neural network and EEG metrics for within-and cross-task workload classification. *NeuroImage*, 59(1), 48–56.

Borghini, G., Aricò, P., Di Flumeri, G., Cartocci, G., Colosimo, A., Bonelli, S., Golfetti, A., Imbert, J. P., Granger, G., & Benhacene, R. (2017). EEG-based cognitive control behaviour assessment: An ecological study with professional air traffic controllers. *Scientific Reports*, 7(1), 1–16.

Brookhuis, K. A., & De Waard, D. (2010). Monitoring drivers' mental workload in driving simulators using physiological measures. *Accident Analysis & Prevention*, 42(3), 898–903.

Brownlee, J. (2017). What is the difference between test and validation datasets. *Machine Learning Mastery*, 14.

Byrne, J. E., Hughes, M. E., Rossell, S. L., Johnson, S. L., & Murray, G. (2017). Time of day differences in neural reward functioning in healthy young men. *Journal of Neuroscience*, 37(37), 8895–8900.

Cain, B. (2007). *A review of the mental workload literature*. Defence Research And Development Toronto (Canada).

Cawley, G. C., & Talbot, N. L. (2010). On over-fitting in model selection and subsequent selection bias in performance evaluation. *The Journal of Machine Learning Research*, 11, 2079–2107.

Caywood, M. S., Roberts, D. M., Colombe, J. B., Greenwald, H. S., & Weiland, M. Z. (2017). Gaussian process regression for predictive but interpretable machine learning models: An example of predicting mental workload across tasks. *Frontiers in Human Neuroscience*, 10, 647

Cegarra, J., & Chevalier, A. (2008). The use of Tholos software for combining measures of mental workload: Toward theoretical and methodological improvements. *Behavior Research Methods*, 40(4), 988–1000.

Charles, R. L., & Nixon, J. (2019). Measuring mental workload using physiological measures: A systematic review. *Applied Ergonomics*, 74, 221–232.

Chen, L., Zhao, Y., Ye, P., Zhang, J., & Zou, J. (2017). Detecting driving stress in physiological signals based on multimodal feature analysis and kernel classifiers. *Expert Systems with Applications*, 85, 279–291.

Chuang, C.Y., Lin, C. J., Shiang, W.J., Hsieh, T.L., & Liou, J.L. (2016). Development of an objective mental workload assessment tool based on Rasmussen's skill–rule–knowledge framework. *Journal of Nuclear Science and Technology*, 53(1), 123–128.

Cinaz, B., La Marca, R., Arnrich, B., & Tröster, G. (2012). *Towards continuous monitoring of mental workload*. University of Zurich

Cobos, C., Zuñiga, J., Guarín, J., León, E., & Mendoza, M. (2010). CMIN-a CRISP-DM-based case tool for supporting data mining projects. *Ingeniería e Investigación*, 30(3), 45–56.

Cortes, C., & Vapnik, V. (1995). Support-vector networks. *Machine Learning*, 20(3), 273–297.

Coyne, J. T., Baldwin, C., Cole, A., Sibley, C., & Roberts, D. M. (2009, July). Applying real time physiological measures of cognitive load to improve training. In *International Conference on Foundations of Augmented Cognition* (pp. 469-478). Springer, Berlin, Heidelberg.

Cox, D. R. (1958). The regression analysis of binary sequences. *Journal of the Royal Statistical Society: Series B (Methodological)*, 20(2), 215-232.

Danziger, S., Levav, J., & Avnaim-Pesso, L. (2011). Extraneous factors in judicial decisions. *Proceedings of the National Academy of Sciences*, 108(17), 6889–6892.

Dell'Agnola, F., Cammoun, L., & Atienza, D. (2018, January). Physiological characterization of need for assistance in rescue missions with drones. In *2018 IEEE International Conference on Consumer Electronics (ICCE)* (pp. 1-6). IEEE.

Elkin, C., & Devabhaktuni, V. (2019, July). Comparative analysis of machine learning techniques in assessing cognitive workload. In *International Conference on Applied Human Factors and Ergonomics* (pp. 185-195). Springer, Cham.

Fallahi, M., Motamedzade, M., Heidarimoghadam, R., Soltanian, A. R., & Miyake, S. (2016). Effects of mental workload on physiological and subjective responses during traffic density monitoring: A field study. *Applied Ergonomics*, 52, 95–103.

Friedman, N., Fekete, T., Gal, Y. K., & Shriki, O. (2019). EEG-based Prediction of Cognitive Load in Intelligence Tests. *Frontiers in Human Neuroscience*, 13, 191.

James, G., Witten, D., Hastie, T., & Tibshirani, R. (2013). An introduction to statistical learning (Vol. 112, p. 18). New York: springer.

Gibson, B., Rogers, T., & Zhu, X. (2013). *Human semi-supervised learning. Topics in Cognitive Science*, 5 (1), 132–172.

Gopher, D., & Donchin, E. (1986). Workload: An examination of the concept. In K. R. Boff, L. Kaufman, & J. P. Thomas (Eds.), *Handbook of perception and human performance*, Vol. 2. Cognitive processes and performance (p. 1–49). John Wiley & Sons.

Grimes, D., Tan, D. S., Hudson, S. E., Shenoy, P., & Rao, R. P. (2008, April). Feasibility and pragmatics of classifying working memory load with an electroencephalograph. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (pp. 835-844).

Guzik, P., & Malik, M. (2016). ECG by mobile technologies. *Journal of Electrocardiology*, 49(6), 894–901.

Hancock, P. A., & Matthews, G. (2019). Workload and performance: Associations, insensitivities, and dissociations. *Human Factors*, 61(3), 374–392.

Harper, G., & Pickett, S. D. (2006). Methods for mining HTS data. *Drug Discovery Today*, 11(15–16), 694–699.

Hart, S. G., & Wickens, C. D. (1990). Workload assessment and prediction. In *Manprint* (pp. 257-296). Springer, Dordrecht.

Healey, J. A., & Picard, R. W. (2005). Detecting stress during real-world driving tasks using physiological sensors. *IEEE Transactions on Intelligent Transportation Systems*, 6(2), 156–166.

Heard, J., Harriott, C. E., & Adams, J. A. (2018). A survey of workload assessment algorithms. *IEEE Transactions on Human-Machine Systems*, 48(5), 434–451.

Jafari, M., Zaeri, F., Jafari, A. H., Payandeh Najafabadi, A. T., Al-Qaisi, S., & Hassanzadeh-Rangi, N. (2020). Assessment and monitoring of mental workload in subway train operations using physiological, subjective, and performance measures. *Human Factors and Ergonomics in Manufacturing & Service Industries*, 30(3), 165–175.

Jin, L., Niu, Q., Hou, H., Xian, H., Wang, Y., & Shi, D. (2012). Driver cognitive distraction detection using driving performance measures. *Discrete Dynamics in Nature and Society*, 2012, 1-12.

Kam Ho, T. (1995). Random decision forests. In Proceedings of 3rd international conference on document analysis and recognition (pp. 278-282).

Kohavi, R., & Provost, F. (1998). Glossary of terms journal of machine learning. *Mach. Learn.*

Kothe, C. A., & Makeig, S. (2011, September). Estimation of task workload from EEG data: new and current tools and perspectives. In 2011 Annual International Conference of the IEEE Engineering in Medicine and Biology Society (pp. 6547-6551). IEEE.

Lean, Y., & Shan, F. (2012). Brief review on physiological and biochemical evaluations of human mental workload. *Human Factors and Ergonomics in Manufacturing & Service Industries*, 22(3), 177–187.

Luque-Casado, A., Perales, J. C., Cárdenas, D., & Sanabria, D. (2016). Heart rate variability and cognitive processing: The autonomic response to task demands. *Biological Psychology*, 113, 83–90.

Mak, J. N., Chan, R. H., & Wong, S. W. (2013, July). Spectral modulation of frontal EEG activities during motor skill acquisition: task familiarity monitoring using single-channel EEG. In 2013 35th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC) (pp. 5638-5641). IEEE.

Mansikka, H., Simola, P., Virtanen, K., Harris, D., & Oksama, L. (2016). Fighter pilots' heart rate, heart rate variation and performance during instrument approaches. *Ergonomics*, 59(10), 1344–1352.

Matthews, G., Reinerman-Jones, L. E., Barber, D. J., & Abich IV, J. (2015). The psychometrics of mental workload: Multiple measures are sensitive but divergent. *Human Factors*, 57(1), 125–143.

Mehler, B., Reimer, B., Coughlin, J. F., & Dusek, J. A. (2009). Impact of incremental increases in cognitive workload on physiological arousal and performance in young adult drivers. *Transportation Research Record*, 2138(1), 6–12.

Mills, C., Fridman, I., Soussou, W., Waghray, D., Olney, A. M., & D'Mello, S. K. (2017, March). Put your thinking cap on: detecting cognitive load using EEG during learning. In Proceedings of the seventh international learning analytics & knowledge conference (pp. 80-89).

Mitchell, T. (1997). *Machine Learning*. McGraw-Hill Higher Education. New York.

Momeni, N., Dell'Agnola, F., Arza, A., & Atienza, D. (2019, July). Real-time cognitive workload monitoring based on machine learning using physiological signals in rescue missions. In 2019 41st Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC) (pp. 3779-3785). IEEE.

Monfort, S. S., Sibley, C. M., & Coyne, J. T. (2016, May). Using machine learning and real-time workload assessment in a high-fidelity UAV simulation environment. In *Next-Generation Analyst IV* (Vol. 9851, p. 98510B). International Society for Optics and Photonics.

Nachreiner, F. (1995). Standards for ergonomics principles relating to the design of work systems and to mental workload. *Applied Ergonomics*, 26(4), 259–263.

Natarajan, A., Xu, K. S., & Eriksson, B. (2016, August). Detecting divisions of the autonomic nervous system using wearables. In 2016 38th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC) (pp. 5761-5764). IEEE.

Nickel, P., & Nachreiner, F. (2003). Sensitivity and diagnosticity of the 0.1-Hz component of heart rate variability as an indicator of mental workload. *Human Factors*, 45(4), 575–590.

Nixon, J., & Charles, R. (2017). Understanding the human performance envelope using electrophysiological measures from wearable technology. *Cognition, Technology & Work*, 19(4), 655–666.

Paas, F., Tuovinen, J. E., Tabbers, H., & Van Gerven, P. W. (2003). Cognitive load measurement as a means to advance cognitive load theory. *Educational Psychologist*, 38(1), 63–71.

Pereira, R. de A., Alves, J. L. de B., Silva, J. H. da C., Costa, M. da S., & Silva, A. S. (2020). Validity of a Smartphone Application and Chest Strap for Recording RR Intervals at Rest in Athletes. *International Journal of Sports Physiology and Performance*, 15(6), 896–899.

Reid, G. B., & Nygren, T. E. (1988). The subjective workload assessment technique: A scaling procedure for measuring mental workload. *Advances in psychology* 52, 185–218

Ripley, B. D. (2007). *Pattern recognition and neural networks*. Cambridge university press.

Rubio, S., Díaz, E., Martín, J., & Puente, J. M. (2004). Evaluation of subjective mental workload: A comparison of SWAT, NASA-TLX, and workload profile methods. *Applied Psychology*, 53(1), 61–86.

Ryu, K., & Myung, R. (2005). Evaluation of mental workload with a combined measure based on physiological indices during a dual task of tracking and mental arithmetic. *International Journal of Industrial Ergonomics*, 35(11), 991–1009.

Schain, M., & Schain, M. (2015). *Machine Learning Algorithms and Robustness*. Universitat Tel-Aviv.

Schuff, D., Corral, K., & Turetken, O. (2011). Comparing the understandability of alternative data warehouse schemas: An empirical study. *Decision Support Systems*, 52(1), 9–20.

Shearer, C. (2000). The CRISP-DM model: The new blueprint for data mining. *Journal of Data Warehousing*, 5(4), 13–22.

Sievertsen, H. H., Gino, F., & Piovesan, M. (2016). Cognitive fatigue influences students' performance on standardized tests. *Proceedings of the National Academy of Sciences*, 113(10), 2621–2624.

Simon, P. (2013). *Too big to ignore: The business case for big data* (Vol. 72). John Wiley & Sons.

Solovey, E. T., Zec, M., Garcia Perez, E. A., Reimer, B., & Mehler, B. (2014, April). Classifying driver workload using physiological and driving performance data: two field studies. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (pp. 4057-4066).

Stevens, R., Galloway, T., & Berka, C. (2006). Integrating EEG models of cognitive load with machine learning models of scientific problem solving. *Augmented Cognition: Past, Present and Future*, 2, 55–65.

Sutton, R., & Barto, A. (1998). *Reinforcement learning: An introduction*. Cambridge, MA: MIT Press.

Sweller, J., Van Merriënboer, J. J., & Paas, F. G. (1998). Cognitive architecture and instructional design. *Educational Psychology Review*, 10(3), 251–296.

Valiant, L. G. (1984). A theory of the learnable. *Communications of the ACM*, 27(11), 1134–1142.



Wilson, G. F., & Fisher, F. (1995). Cognitive task classification based upon topographic EEG data. *Biological Psychology*, 40(1–2), 239–250.

Watson, D. W. (2001, June). Physiological correlates of heart rate variability (HRV) and the subjective assessment of workload and fatigue in-flight crew: a practical study. In 2001 People in Control. The Second International Conference on Human Interfaces in Control Rooms, Cockpits and Command Centres (pp. 159-163). IET.

Wilson, G. F., & Russell, C. A. (2003a). Operator functional state classification using multiple psychophysiological features in an air traffic control task. *Human Factors*, 45(3), 381–389.

Wilson, G. F., & Russell, C. A. (2003b). Real-time assessment of mental workload using psychophysiological measures and artificial neural networks. *Human Factors*, 45(4), 635–644.

Wu, X., Kumar, V., Quinlan, J. R., Ghosh, J., Yang, Q., Motoda, H., McLachlan, G. J., Ng, A., Liu, B., & Philip, S. Y. (2008). Top 10 algorithms in data mining. *Knowledge and Information Systems*, 14(1), 1–37.

Yatsenko, M., Kaigorodova, N., Kazin, E., & Fedorov, A. (2018). EEG Correlation of the influence of endogenous and exogenous factors on mental work capacity in students. *Human Physiology*, 44(6), 635–646.

Young, M. S., Brookhuis, K. A., Wickens, C. D., & Hancock, P. A. (2015). State of science: Mental workload in ergonomics. *Ergonomics*, 58(1), 1–17.

Young, M. S., & Stanton, N. A. (2002). It's all relative: Defining mental workload in the light of Annett's paper. *Ergonomics*, 45(14), 1018–1020.

Zak, Y., Parmet, Y., & Oron-Gilad, T. (2020, October). Subjective Workload Assessment Technique (SWAT) in Real Time: Affordable Methodology to Continuously Assess Human Operators' Workload. In 2020 IEEE International Conference on Systems, Man, and Cybernetics (SMC) (pp. 2687-2694). IEEE.

Kramer AF (1991) Physiological metrics of mental workload: a review of recent progress. Multiple-task performance, pp 279– 328

Riener A, Ferscha A, Aly M (2009) Heart on the road: Hrv analysis for monitoring a driver's affective state. In: AutomotiveUI '09: proceedings of the 1st international conference on automotive user interfaces and interactive

Wilson GF (2002) An analysis of mental workload in pilots during flight using multiple psychophysiological measures. *Int J Aviat Psychol* 12:3–18

## 9. נספחים:

### נספח 1 - הוראות ביצוע הניסוי

הניסוי בו אתם משתתפים נערך במסגרת המחקר שמטרתו לבחון האם קיים מתאם בין הדופק המופק מרצועת חזה לבין רמת העומס הקוגניטיבי. הניסוי עצמו נמשך 3 ימים, ובכל יום במהלכו עליכם לבצע שלוש משימות בדרגות קושי שונות שמטרתן לעורר מידה מסוימת של עומס קוגניטיבי. כל משימה אורכת חצי שעה, ובעת ביצועה תתבקשו ללבוש רצועת חזה שניתן לכם מדגם Polar H10, ולהקליט את דופק הלב שלכם באמצעות אפליקציה ייעודית. כל משימה תבוצע בזמן נפרד ביום כך שאחת תתרחש בבוקר, שנייה בצהריים ושלישית בערב. בעת ביצוע משימה מסוימת, לאחר כל 10 דקות תמלאו שאלון SWAT לצורך הערכת העומס הקוגניטיבי הנתפס. בתום 30 הדקות, תמלאו שאלון נוסף להערכת העומס הקוגניטיבי שחש במהלך המשימה – שאלון NASA TLX.

פירוט המשימות שעליכם לבצע במשך 30 דקות בכל יום מופיע בטבלה שלהלן:

זמן/יום הניסוי	1	2	3
<b>בוקר</b>	לצפות בטלוויזיה	לבצע מבחן IQ	לשחק באפליקציית unblock car
<b>צהריים</b>	לפתור פרק כמותי של פסיכומטרי	לפתור תשחץ	לעשות מדיטציה
<b>ערב</b>	סודוקו	להקשיב למוזיקה	לפתור מבחן פסיכוטכני

להלן כלל הקישורים המתאימים למשימות שלעיל:

משימה	קישור לביצוע
מבחן פסיכוטכני	<a href="#">מבחן פסיכוטכני</a>
פרק כמותי של פסיכומטרי	<a href="#">פרק כמותי</a>
מבחן IQ	<a href="#">מבחן IQ</a>
סודוקו	<a href="#">סודוקו</a>
תשחץ	<a href="#">תשחץ</a>
unblock car	<a href="#">unblock car</a>
הקשבה למוזיקה	לבחירת המשתתף
מדיטציה	<a href="#">מדיטציה</a>
צפייה בטלוויזיה	<a href="#">ריאיון טלוויזיוני</a>

לצורך ביצוע הניסוי בשלמותו, עליכם לוודא שבטלפון שלכם מותקנות האפליקציות הבאות:

- אפליקציית SelfLoops HRV לצורך הקלטת הדופק

- אפליקציית Unblock Car לצורך ביצוע משימת הבוקר ביום השלישי

את כל האפליקציות המתוארות לעיל ניתן למצוא ב-google paly וב-apple store להורדה בחינם.

## נספח 2 – טופס הסכמה להשתתפות בניסוי

### טופס הסכמה להשתתפות במחקר לבחינת הקשר בין מדדי דופק המופקים מרצועת חזה לעומס קוגניטיבי

משתתפים יקרים,

הניסוי בו אתם משתתפים נערך במסגרת המחקר שמטרתו לבחון מהו הקשר בין מדדי הדופק המופקים מרצועת חזה לבין רמת העומס הקוגניטיבי. הניסוי עצמו נמשך 3 ימים, ובכל יום במהלכו עליכם לבצע שלוש משימות בדרגות עומס קוגניטיבי שונות. כל משימה אורכת חצי שעה, ובעת ביצועה תתבקשו ללבוש רצועת חזה מדגם Polar H10 אותה נספק לכם, כדי להקליט את דופק הלב שלכם באמצעות אפליקציה ייעודית ולהיכנס לשיחת זום עם אחד מהחוקרים. כל משימה תבוצע בזמן נפרד ביום כך שאחת תתרחש בבוקר, שנייה בצהריים ושלישית בערב. בעת ביצוע משימה מסוימת, לאחר כל 5 דקות תמלאו שאלון SWAT לצורך הערכת העומס הקוגניטיבי הנתפס. בתום 30 הדקות, תמלאו שאלון נוסף להערכת העומס הקוגניטיבי שחש במהלך המשימה – שאלון NASA TLX.

בטרם תחלי את הניסוי תתבצע שיחה קצרה בה נסביר על הניסוי, במהלכה תתבקש/י למסור פרטים דמוגרפיים (מין, גיל) וכן יהיה עלייך לדרג לפי תפיסתך את ההשפעה של זמן, מאמץ מנטלי ולחץ פסיכולוגי על עומס המידע שאתה חש.

הנתונים שתספק/י במסגרת הניסוי יהיו חסויים ויישמרו בקבצים מאובטחים במחשב של החוקר הראשי. בנוסף, הפרטים האישיים שלך יישמרו בנפרד מהנתונים שתספק/י בניסוי, ולא יהיה ניתן לקשר בין זהותך לבין הנתונים והמידע הדמוגרפי שתספק/י.

פרסום התוצאות יכלול רק נתונים מסכמים ללא יכולת לקשר ביניהם לבין הנתונים שיאספו לגביך.

בסוף הניסוי תקבל/י כתובת מייל של מנהל הניסוי. את תוצאות השאלון ואת נתוני הדופק המופקים מהרצועה, תתבקש/י לשלוח במייל לכתובת זו.

חשוב לציין כי ניתנת לך זכות מלאה להפסיק את השתתפותך בניסוי בכל שלב ומכל סיבה שהיא. במידה ותחליט/י להפסיק את השתתפותך, הנתונים שסיפקת עד לאותה נקודה לא יישמרו.

המחקר מתקיים כחלק מפרויקט הגמר של דור אלדר, עומרי זאבי ושיר רוזנס תחת הנחייתם של פרופ' טל אורון-גלעד, וד"ר יובל זק. את הניסוי ינהל אחד או יותר מהסטודנטים העובדים על פרויקט הגמר, המחויבים לשמור על הכללים הקבועים במסמך זה, ועל כללי האתיקה בניסויים. הניסוי קיבל את אישור ועדת האתיקה.

הצהרת המשתתף: אני החתום/ה מטה מצהיר/ה כי קראתי את טופס ההסכמה במלואו, וכי בחתימתי אני מביע/ה את הסכמתי להשתתף בניסוי.

\_\_\_\_\_/\_\_\_\_\_/\_\_\_\_\_  
תאריך שם פרטי + שם משפחה חתימה

הצהרת מנהל הניסוי: אני החתום/ה מטה מצהיר/ה כי אני מחויב/ת לשמור על סודיות מלאה. כל המידע הרלוונטי יהיה חשוף לצוות המחקר בלבד ולפי התנאים המוסברים בדף זה. מעבר לכך לא אעביר בשום צורה ואופן כל מידע שעולה מהניסוי.

\_\_\_\_\_/\_\_\_\_\_/\_\_\_\_\_  
תאריך שם פרטי + שם משפחה חתימה

לתיאום ופרטים נוספים ניתן לפנות לדור אלדור בטלפון 050-8646337 או לד"ר טל אורון-גלעד בטלפון 08-6461434 או בדוא"ל [orontal@bgu.ac.il](mailto:orontal@bgu.ac.il).

## NASA Task Load Index

Hart and Staveland's NASA Task Load Index (TLX) method assesses work load on five 7-point scales. Increments of high, medium and low estimates for each point result in 21 gradations on the scales.

Name	Task	Date
------	------	------

**Mental Demand**                      How mentally demanding was the task?

**(שלב 1) היגדים לבחינת סדר העדיפויות של שלושת הממדים עבור כל נבדק:**

(אנא דרג את המשפטים הבאים בין 1 (עומס עבודה מנטלי נמוך) ל-27 (עומס עבודה מנטלי גבוה)

לחץ הזמן מאפשר לעיתים זמן פנוי, יש לעיתים חפיפה בזמן העבודה בין משימות; מעט מאוד מאמץ מנטלי וריכוז נדרשים לביצוע המשימה; קיים לחץ רב בשל בלבול ותסכול, נדרשת שליטה עצמית גבוהה. P

לחץ הזמן מאפשר לעיתים זמן פנוי, יש לעיתים חפיפה בזמן העבודה בין משימות; מעט מאוד מאמץ מנטלי וריכוז נדרשים לביצוע המשימה; קיים מעט בלבול ותסכול בביצוע המשימה. 8

לחץ הזמן מאפשר לעיתים זמן פנוי, יש לעיתים חפיפה בזמן העבודה בין משימות; מעט מאוד מאמץ מנטלי וריכוז נדרשים לביצוע המשימה; קיים לחץ בינוני בשל בלבול ותסכול בביצוע המשימה, מה שמחייב פיצוי קוגניטיבי. G

לחץ הזמן מאפשר לעיתים זמן פנוי, יש לעיתים חפיפה בזמן העבודה בין משימות; נדרש מאמץ מנטלי וריכוז גבוהים לביצוע המשימה, המשימות מורכבות; קיים לחץ רב בשל בלבול ותסכול, נדרשת שליטה עצמית גבוהה. T

לחץ הזמן מאפשר לעיתים זמן פנוי, יש לעיתים חפיפה בזמן העבודה בין משימות; נדרש מאמץ מנטלי וריכוז גבוהים לביצוע המשימה, המשימות מורכבות; קיים מעט בלבול ותסכול בביצוע המשימה. S

לחץ הזמן מאפשר לעיתים זמן פנוי, יש לעיתים חפיפה בזמן העבודה בין משימות; נדרש מאמץ מנטלי וריכוז גבוהים לביצוע המשימה, המשימות מורכבות; קיים לחץ בינוני בשל בלבול ותסכול בביצוע המשימה, מה שמחייב פיצוי קוגניטיבי. E

לחץ הזמן מאפשר לעיתים זמן פנוי, יש לעיתים חפיפה בזמן העבודה בין משימות; נדרש מאמץ מנטלי וריכוז בינוניים לביצוע המשימה, מורכבות המשימות בינונית; קיים לחץ רב בשל בלבול ותסכול, נדרשת שליטה עצמית גבוהה. A

לחץ הזמן מאפשר לעיתים זמן פנוי, יש לעיתים חפיפה בזמן העבודה בין משימות; נדרש מאמץ מנטלי וריכוז בינוניים לביצוע המשימה, מורכבות המשימות בינונית; קיים מעט בלבול ותסכול בביצוע המשימה. J

לחץ הזמן מאפשר לעיתים זמן פנוי, יש לעיתים חפיפה בזמן העבודה בין משימות; נדרש מאמץ מנטלי וריכוז בינוניים לביצוע המשימה, מורכבות המשימות בינונית; קיים לחץ בינוני בשל בלבול ותסכול בביצוע המשימה, מה שמחייב פיצוי קוגניטיבי. Q

לחץ הזמן מאפשר לרוב זמן פנוי, אין חפיפה בזמן העבודה בין משימות; מעט מאוד מאמץ מנטלי וריכוז נדרשים לביצוע המשימה; קיים לחץ רב בשל בלבול ותסכול, נדרשת שליטה עצמית גבוהה. H

לחץ הזמן מאפשר לרוב זמן פנוי, אין חפיפה בזמן העבודה בין משימות; מעט מאוד מאמץ מנטלי וריכוז נדרשים לביצוע המשימה; קיים מעט בלבול ותסכול בביצוע המשימה. N

**(שלב 2) שלושת השאלות לבחינת העומס הקוגניטיבי הנתפס בזמן המשימה בשלושת הממדים:**

1. עומס זמן במהלך הסימולציה:

א) לעיתים קרובות יש זמן פנוי. הפרעות או חפיפה בין פעולות קורות לעיתים רחוקות או בכלל לא.

ב) מדי פעם יש זמן פנוי. הפרעות או חפיפה בין פעולות קורות לעיתים קרובות.

ג) כמעט ואין זמן פנוי. הפרעות או חפיפה בין פעולות קורות לעיתים קרובות מאוד או תמיד.

2. מאמץ מנטאלי במהלך הסימולציה:

א) מאמץ מנטאלי קטן, רמת ריכוז נמוכה. הפעולות שביצעת בוצעו בצורה כמעט אוטומטית.

ב) מאמץ מנטאלי בינוני, רמת ריכוז בינונית הנגרמת מחוסר וודאות, מורכבות, ומארועים בלתי צפויים. הפעולות שביצעת דרשו מאמץ ותשומת לב.

ג) מאמץ מנטאלי מתמשך, פעולות מאוד מורכבות הדורשות ריכוז מלא לזמן ארוך.

3. לחץ פסיכולוגי במהלך הסימולציה:

א) מעט בלבול, סיכון, תסכול או חרדה שניתן לשלוט בהם. או כלל לא.

ב) הרגשה של לחץ הנגרם מבלבול, סיכון או חרדה אשר מגדיל את עומס המידע ומפריעה לתפעול המשימה.

ג) הרגשת לחץ כבד הנגרם מבלבול, סיכון או חרדה אשר דורש ממך מאמץ של שליטה עצמית ומפריעה מאוד או מבטל לגמרי את היכולת לתפעל את במשימה.

### (שלב 3) ציון העומס על פי הממדים הנבחרים:

Table 2. Six hypothetical weighting schemes.

Rank order	Descriptor combination in the card is time (T), effort (E) and stress (S) respectively					
	TES	TSE	ETS	EST	STE	SET
1	1 1 1	1 1 1	1 1 1	1 1 1	1 1 1	1 1 1
2	1 1 2	1 2 1	1 1 2	2 1 1	1 2 1	2 1 1
3	1 1 3	1 3 1	1 1 3	3 1 1	1 3 1	3 1 1
4	1 2 1	1 1 2	2 1 1	1 1 2	2 1 1	1 2 1
5	1 2 2	1 2 2	2 1 2	2 1 2	2 2 1	2 2 1
6	1 2 3	1 3 2	2 1 3	3 1 2	2 3 1	3 2 1
7	1 3 1	1 1 3	3 1 1	1 1 3	3 1 1	1 3 1
8	1 3 2	1 2 3	3 1 2	2 1 3	3 2 1	2 3 1
9	1 3 3	1 3 3	3 1 3	3 1 3	3 3 1	3 3 1
10	2 1 1	2 1 1	1 2 1	1 2 1	1 1 2	1 1 2
11	2 1 2	2 2 1	1 2 2	2 2 1	1 2 2	2 1 2
12	2 1 3	2 3 1	1 2 3	3 2 1	1 3 2	3 1 2
13	2 2 1	2 1 2	2 2 1	1 2 2	2 1 2	1 2 2
14	2 2 2	2 2 2	2 2 2	2 2 2	2 2 2	2 2 2
15	2 2 3	2 3 2	2 2 3	3 2 2	2 3 2	3 2 2
16	2 3 1	2 1 3	3 2 1	1 2 3	3 1 2	1 3 2
17	2 3 2	2 2 3	3 2 2	2 2 3	3 2 2	2 3 2
18	2 3 3	2 3 3	3 2 3	3 2 3	3 3 2	3 3 2
19	3 1 1	3 1 1	1 3 1	1 3 1	1 1 3	1 1 3
20	3 1 2	3 2 1	1 3 2	2 3 1	1 2 3	2 1 3
21	3 1 3	3 3 1	1 3 3	3 3 1	1 3 3	3 1 3
22	3 2 1	3 1 2	2 3 1	1 3 2	2 1 3	1 2 3
23	3 2 2	3 2 2	2 3 2	2 3 2	2 2 3	2 2 3
24	3 2 3	3 3 2	2 3 3	3 3 2	2 3 3	3 2 3
25	3 3 1	3 1 3	3 3 1	1 3 3	3 1 3	1 3 3
26	3 3 2	3 2 3	3 3 2	2 3 3	3 2 3	2 3 3
27	3 3 3	3 3 3	3 3 3	3 3 3	3 3 3	3 3 3



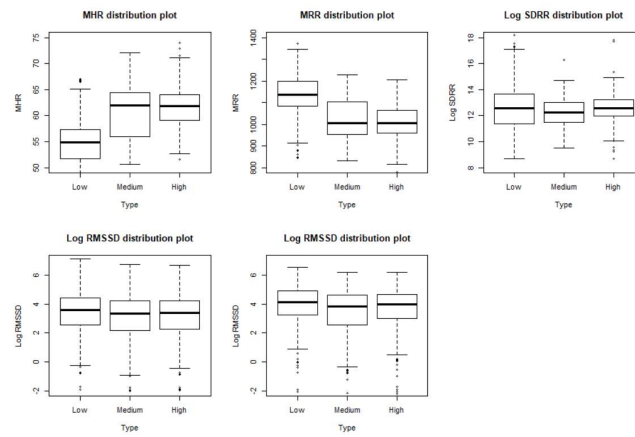
Load	Features	Min	Max	Mean
Low	MRR	560.50	3154.20	939.70
	MHR	48.84	104.33	66.48
	HRV	0.00	1239.00	1.88
	RMSSD	0.00	2583.00	18.77
	SDNN	0.00	1955.57	18.26
	HRmaxHRminDifference	0.00	77.00	3.26
	SDRR	6.08	2849.81	71.36
	HRV_div_aveHR	0.00	22.91	0.03
	HRV_div_aveHR_MoveLast15	0.00	12.34	0.05
	last5secHRave_div_aveHR	0.11	1.42	1.00
	last10secHRave_div_aveHR	0.57	1.42	1.00
	last15secHRave_div_aveHR	0.72	1.42	1.01
	Log_SDRR	5.21	22.95	10.85
	Log_RMSSD	-2.23	11.33	2.41
	Log_SDNN	-6.33	10.93	2.83

Load	Features	Min	Max	Mean
Medium	MRR	571.80	1378.30	897.60
	MHR	50.62	104.58	68.96
	HRV	0.00	74.00	68.96
	RMSSD	0.00	1900.85	9.25
	SDNN	0.00	1742.29	10.94
	HRmaxHRminDifference	0.00	24.00	3.16
	SDRR	4.36	2051.17	47.99
	HRV_div_aveHR	0.00	1.06	0.02
	HRV_div_aveHR_MoveLast15	0.00	1.62	0.04
	last5secHRave_div_aveHR	0.81	1.21	1.00
	last10secHRave_div_aveHR	0.77	1.26	1.00
	last15secHRave_div_aveHR	0.75	1.29	1.00
	Log_SDRR	4.25	22.00	10.47
	Log_RMSSD	-2.32	10.89	2.14
	Log_SDNN	-6.09	10.77	2.53

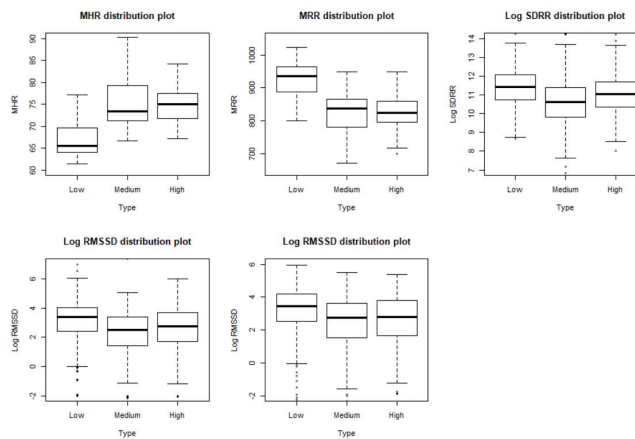
Load	Features	Min	Max	Mean
High	MRR	580.90	1477.60	867.50
	MHR	51.56	103.89	71.43
	HRV	0.00	242.00	1.60
	RMSSD	0.00	8.77	2328.91
	SDNN	0.00	2159.55	11.16
	HRmaxHRminDifference	0.00	68.00	3.46
	SDRR	4.36	2241.20	48.57
	HRV_div_aveHR	0.00	3.85	0.02
	HRV_div_aveHR_MoveLast15	0.00	1.99	0.04
	last5secHRave_div_aveHR	0.81	1.17	1.00
	last10secHRave_div_aveHR	0.81	1.97	1.00
	last15secHRave_div_aveHR	0.81	1.22	1.00
	Log_SDRR	4.25	22.26	10.41
	Log_RMSSD	-2.29	11.19	1.98
	Log_SDNN	-6.21	11.08	2.51

## נספח 6 – תרשימי קופסא עבור מדדי המודל לכלל הנבדקים

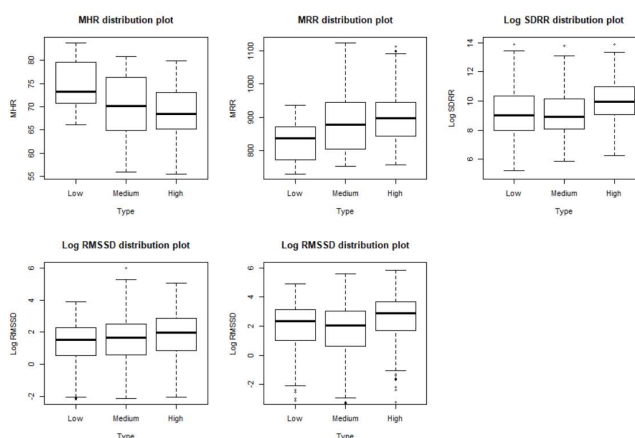
נבדק A:



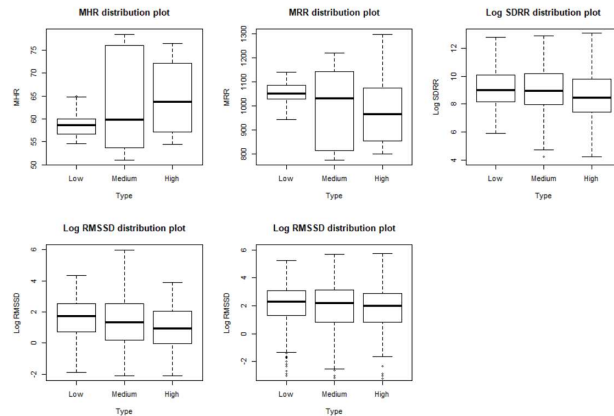
נבדק B:



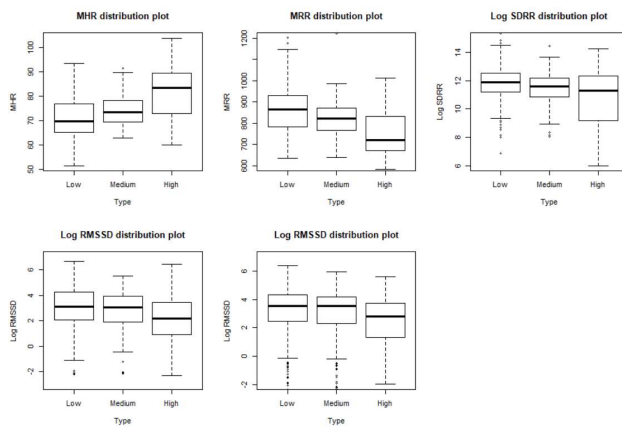
נבדק C:



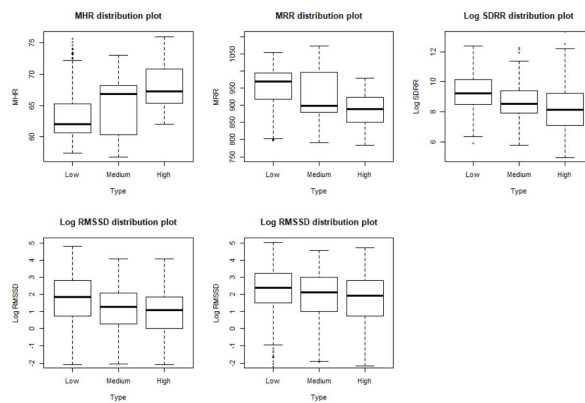
נבדק D:



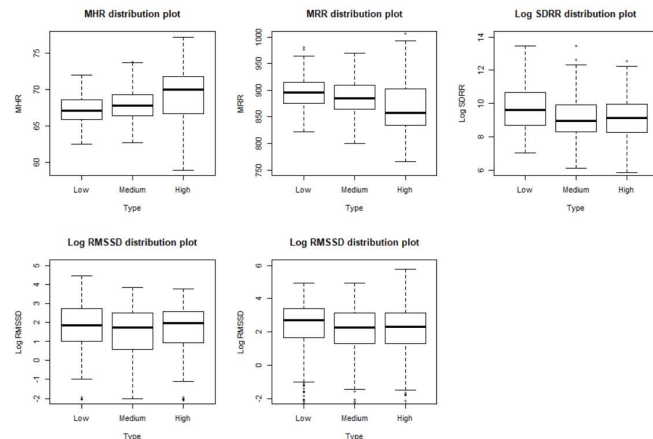
נבדק E:



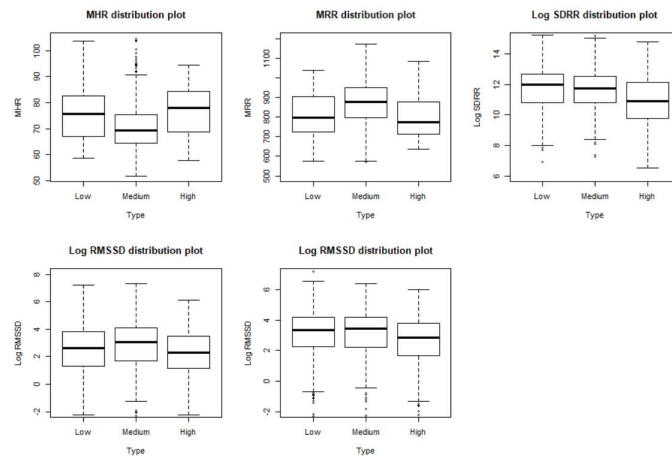
נבדק F:



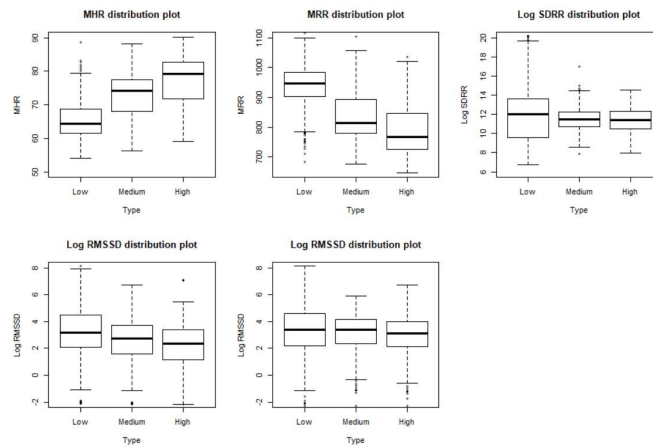
נבדק G:



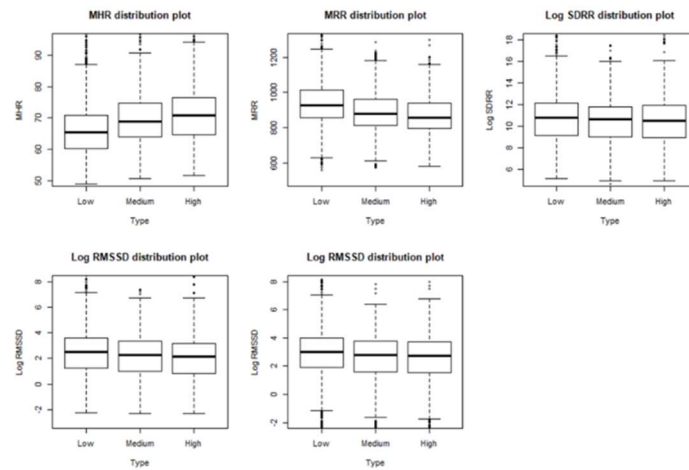
נבדק H:



נבדק I:



## כל הנבדקים:



## נספח 7 – טבלת ממוצע וסטיות תקן של מדדי הדופק

	Low	Medium	High
<b>MRR</b>	939.68 ± 2.57	897.57 ± 2.10	867.52 ± 2.08
<b>MHR</b>	66.47 ± 0.15	68.95 ± 0.15	71.43 ± 0.16
<b>HRV</b>	1.88 ± 0.39	1.33 ± 0.072	1.60 ± 0.10
<b>RMSSD</b>	18.77 ± 1.96	9.24 ± 0.89	8.76 ± 0.83
<b>SDNN</b>	18.25 ± 1.40	10.94 ± 0.61	11.16 ± 0.75
<b>HRmaxHRminDifference</b>	3.25 ± 0.05	3.16 ± 0.04	3.46 ± 0.05
<b>SDRR</b>	71.35 ± 3.09	47.98 ± 1.20	48.57 ± 1.19
<b>HRV_div_aveHR</b>	0.02 ± 0.00	0.01 ± 0.00	0.02 ± 0.00
<b>HRV_div_aveHR_MoveLast15</b>	0.05 ± 0.00	0.04 ± 0.00	0.04 ± 0.00
<b>last5secHRave_div_aveHR</b>	1.00 ± 0.00	1.00 ± 0.00	1.00 ± 0.00
<b>last10secHRave_div_aveHR</b>	1.00 ± 0.00	1.000726 ± 0.00006	1.00 ± 0.00
<b>last15secHRave_div_aveHR</b>	1.00 ± 0.00	1.00 ± 0.00	1.00 ± 0.00
<b>Log_SDNN</b>	2.82 ± 0.03	2.53 ± 0.03	2.50 ± 0.03
<b>Log_SDRR</b>	10.85 ± 0.04	10.46 ± 0.03	10.41 ± 0.03
<b>Log_RMSSD</b>	2.41 ± 0.03	2.13 ± 0.03	1.97 ± 0.03

```

def the_end_of_vector (vector,number):
#   the function returns the number of values from the end of the vector
    NewVector=[]
    if (len(vector)-number)<0:
        print("the length of the vector smaller than the number of rows")
        return None
    for r in range(number):
        NewVector.append(vector[(len(vector)-number)+r])
    return NewVector

def new_data (raw_data): #15sec
    first_row_flag=False
    sum_all_HR=0
    count_all_HR=0

    HRV_table = pd.DataFrame(columns
    =['MRR','MHR','HRV','RMSSD','SDNN','HRmaxHRminDifference','SDRR','HRV_div_a
    veHR','HRV_div_aveHR_MoveLast15','last5secHRave_div_aveHR','last10secHRave
    _div_aveHR','last15secHRave_div_aveHR', 'swat rank', 'Day', 'TLX Overall', 'Type'])
    #return table
    n_rows=15000 #ms for one row
    index = 1 # Count the number of ms used in the session - till 15 seconds
    ms_indicator_not_used = 0 # indicator for the amount used from specific row
    HR_Last_Row = 0 #save the last HR value of a split row
    HR_vector = [] #store the HR values in the vector for the group row
    RR_vector[] =
    HR_last_vector=[] #store the last HR vector

    for i in range(len(raw_data)):
#       the for run along each row in the tabel
        RR_vector.append(raw_data["R-R (ms)"][i])

        row_indicator = raw_data["R-R (ms)"][i]
        if ms_indicator_not_used != 0: #case 3 - unused row from the last for loop
            for j in range(ms_indicator_not_used):
                HR_vector.append(HR_Last_Row)
            index = index + ms_indicator_not_used
            ms_indicator_not_used = 0
            HR_Last_Row = 0
        if index + row_indicator <= n_rows: #case 1 - full row can enter into the vector
            for j in range(row_indicator):
                HR_vector.append(raw_data["HR (bpm)"][i])
            index = index + row_indicator
        else: #case 2: part of the row can enter the vector
            ms_indicator_to_use = n_rows - index + 1
            ms_indicator_not_used = raw_data["R-R (ms)"][i] - ms_indicator_to_use
            HR_Last_Row = raw_data["HR (bpm)"][i]
            for j in range(ms_indicator_to_use):
                HR_vector.append(raw_data["HR (bpm)"][i])

#       the end of n rows session creating HRV value
        index = 1

```

```

    if first_row_flag==True: #Skip the first row with no history values
#       HRV=statistics.variance(HR_vector)
       aveHR=numpy.mean(HR_vector)

#       measure 1:(HRV_div_aveHR) HRV divided by HR average session
       HRV_div_aveHR=(statistics.variance(HR_vector)/aveHR)

#       measure 2: (HRV_div_aveHR_MoveLast15) HRV divided average of HR
       average both by "moving-average" includes the last 20 second
       num_of_last_vector=15000 #ms
       HR_vector_with_Lastsec[]=
       for k in range(len(HR_vector)):
HR_vector_with_Lastsec.append(HR_vector[k])
       end_lastvector=the_end_of_vector(HR_last_vector,num_of_last_vector)
       for k in range(len(end_lastvector)):
HR_vector_with_Lastsec.append(end_lastvector[k])
       HRV_moving=statistics.variance(HR_vector_with_Lastsec)

HRV_div_aveHR_MoveLast15=(HRV_moving/(numpy.mean(HR_vector_with_Lastse
c)))

#       measure 3: (last5secHR_div_aveHR) average HR of the last 5 seconds
       from the last vector divided by HR average of this session
       HR_vector_last5sec =the_end_of_vector(HR_last_vector,5000)
       last5secHRave_div_aveHR= numpy.mean(HR_vector_last5sec)/aveHR

#       measure 4: (last10secHR_div_aveHR) average HR of the last 10 seconds
       from the last vector divided by HR average of this session
       HR_vector_last10sec =the_end_of_vector(HR_last_vector,10000)
       last10secHRave_div_aveHR= numpy.mean(HR_vector_last10sec)/aveHR

#       measure 5: (last15secHR_div_aveHR) average HR of the last 15 seconds
       from the last vector divided by HR average of this session
       HR_vector_last15sec =the_end_of_vector(HR_last_vector,15000)
       last15secHRave_div_aveHR=numpy.mean(HR_vector_last15sec)/aveHR

#       measure 1 NEW: (MRR) mean of the RR interval
       MRR=numpy.mean(RR_vector)

#       measure 2 NEW: (MHR) mean of HR in the interval
       MHR=numpy.mean(HR_vector)

#       measur 3 NEW: (HRV) HR Standard deviation
       HRV=statistics.variance(HR_vector)

#       measur 4 NEW: (RMSSD) root ((mean difference between RR i+1 - RR i )
       squared)
       for i in range(len(RR_vector)-1):

```

```

        sumRRdifference=math.pow(RR_vector[i+1]-RR_vector[i],2)
        RMSSD=math.sqrt(sumRRdifference/(len(RR_vector)-1))

#        measur 5 NEW: (SDNN) root ((mean difference between RR i - RR ave )
squared)
        RRave=numpy.mean(RR_vector)
        for i in range(len(RR_vector)):
            sumRRdifference=math.pow(RR_vector[i]-RRave,2)
            SDNN=(math.sqrt(sumRRdifference/(len(RR_vector)-1)))

#        measur 6 NEW: (HRmaxHRminDifference) HRmax-HRmin
        HRmaxHRminDifference= numpy.max(HR_vector)-numpy.min(HR_vector)

#        measur 7 NEW: (SDRR) RR Standard deviation
        SDRR=statistics.variance(RR_vector)

new_row=pd.DataFrame([(MRR,MHR,HRV,RMSSD,SDNN,HRmaxHRminDifference,
SDRR,HRV_div_aveHR,HRV_div_aveHR_MoveLast15,last5secHRave_div_aveHR,l
ast10secHRave_div_aveHR,last15secHRave_div_aveHR,raw_data.iloc[i]["swat
rank"],raw_data.iloc[i]["Day"],round(raw_data.iloc[i]["TLX
Overall"],3),raw_data.iloc[i]["Type"])],columns =
['MRR','MHR','HRV','RMSSD','SDNN','HRmaxHRminDifference','SDRR','HRV_div_av
eHR','HRV_div_aveHR_MoveLast15','last5secHRave_div_aveHR','last10secHRave_
div_aveHR','last15secHRave_div_aveHR', 'swat rank', 'Day', 'TLX Overall', 'Type'])

        HRV_table=HRV_table.append(new_row)

        HR_last_vector=HR_vector
        HR_vector[] =
        RR_vector[]=
        first_row_flag=True

return HRV_table

```

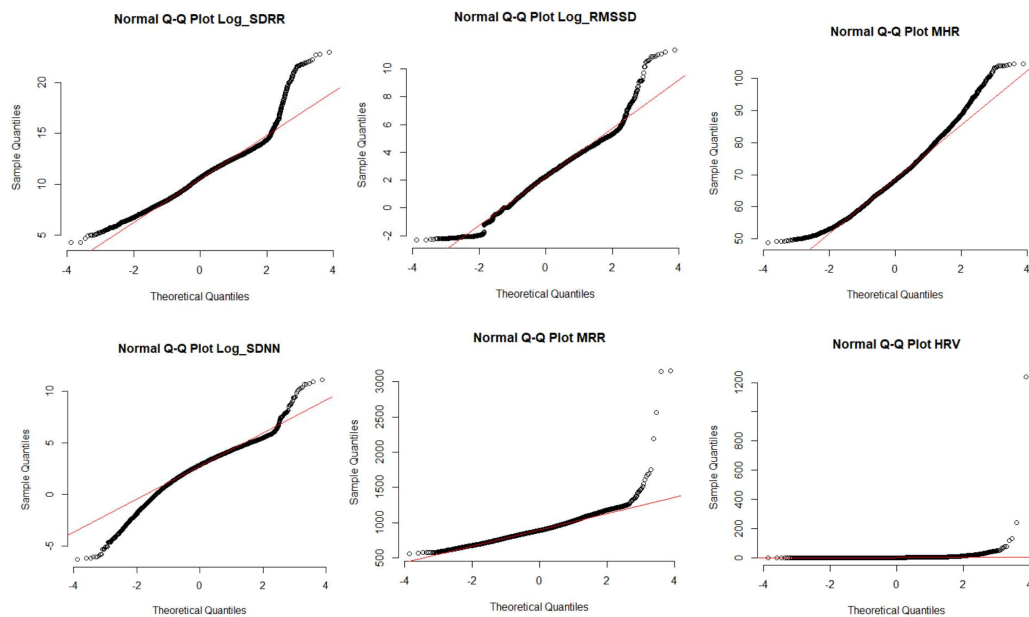


## נספח 9 - בדיקת התפלגות נורמלית של המדדים

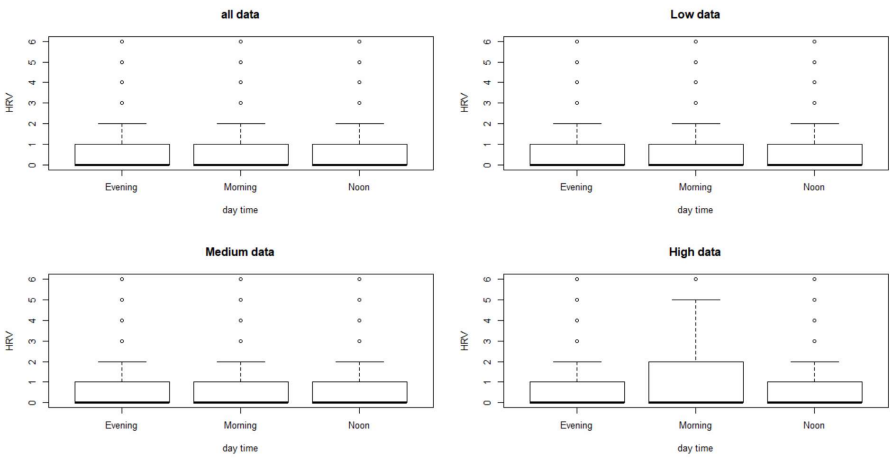
Load	Features	W	P-Value
Low	MRR	<b>0.861</b>	2.20E-16
	MHR	<b>0.962</b>	1.60E-16
	HRV	0.033	2.70E-16
	RMSSD	0.102	3.20E-16
	SDNN	0.129	2.20E-16
	SDRR	0.231	1.50E-16
	Log_SDRR	<b>0.935</b>	2.20E-16
	Log_RMSSD	<b>0.978</b>	2.80E-16
	Log_SDNN	<b>0.957</b>	2.20E-16

Load	Features	W	P-Value
Medium	MRR	<b>0.964</b>	2.20E-16
	MHR	<b>0.986</b>	3.25E-16
	HRV	0.334	2.20E-16
	RMSSD	0.065	2.20E-16
	SDNN	0.125	2.20E-16
	SDRR	0.276	2.20E-16
	Log_SDRR	<b>0.985</b>	2.20E-16
	Log_RMSSD	<b>0.987</b>	7.51E-16
	Log_SDNN	<b>0.955</b>	2.20E-16

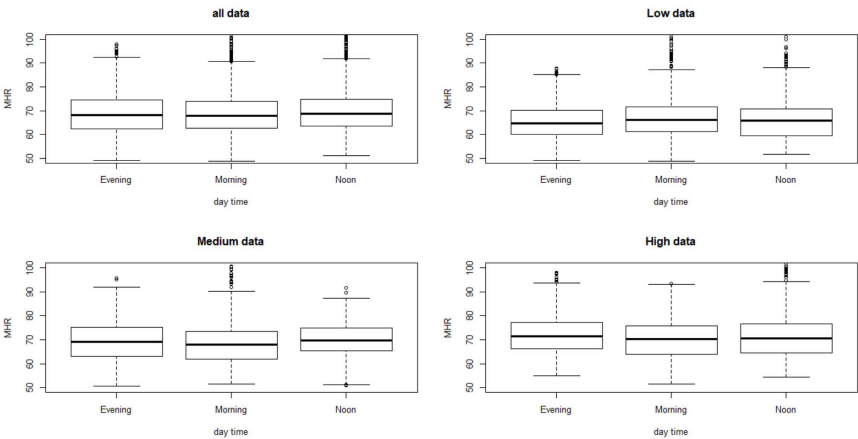
Load	Features	W	P-Value
High	MRR	<b>0.99</b>	1.65E-13
	MHR	<b>0.981</b>	2.20E-16
	HRV	0.224	2.20E-16
	RMSSD	0.073	1.90E-16
	SDNN	0.099	2.40E-16
	SDRR	0.328	2.20E-16
	Log_SDRR	<b>0.986</b>	2.20E-16
	Log_RMSSD	<b>0.989</b>	1.12E-14
	Log_SDNN	<b>0.962</b>	2.20E-16



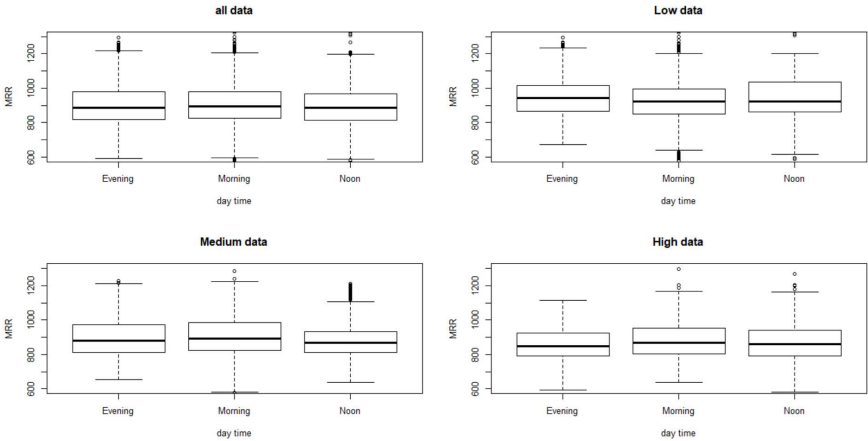
HRV



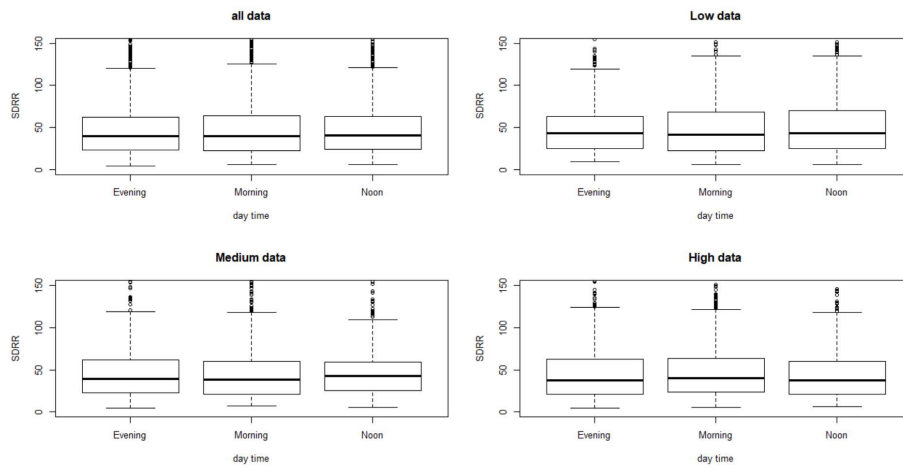
MHR



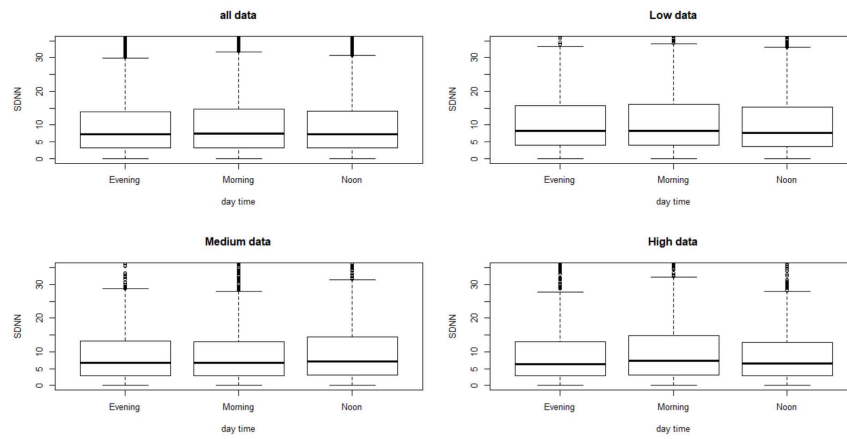
MRR



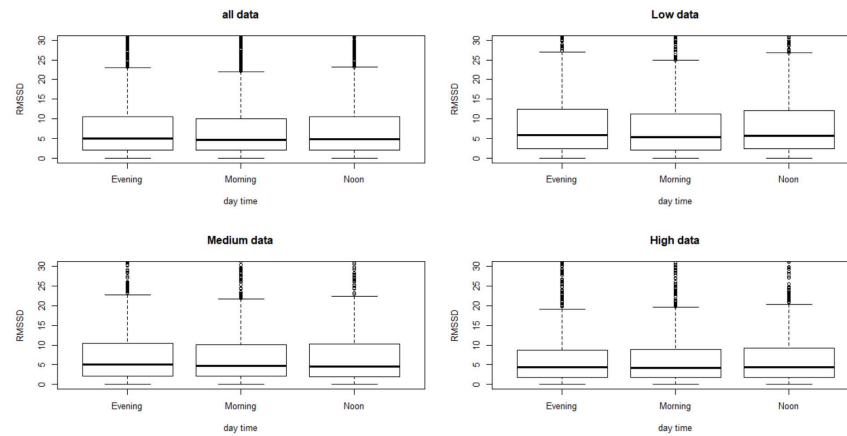
## SDRR



## SDNN



## RMSSD



## נספח 11 - ניתוח תוצאות אלגוריתם Random Forest

### 11.1 תוצאות מובהקות המדדים לפי פונקציית importance - לפני טרנספורמציה

#### 11.1.1. כלל המדדים

	A		B		C		D		E	
	MFR	0.16040812306002624	MFR	0.18258371136398647	MFR	0.1682915739136757	MFR	0.26186008267006516	MFR	0.15409587079445205
All features	MHR	0.22163160568878908	MHR	0.2493017012213007	MHR	0.21026448621734684	MHR	0.3132496736922315	MHR	0.1707897627674334
	HRV	0.01853842321395554	HRV	0.01639989907630393	HRV	0.01590064076846687	HRV	0.00315241132645293	HRV	0.02630078934493744
	RMSSD	0.0848518193970585	RMSSD	0.08734352708496179	RMSSD	0.07496865744163975	RMSSD	0.05976380411953817	RMSSD	0.08222637933832443
	SDNN	0.0749300956633312	SDNN	0.06139367610365304	SDNN	0.08049796995794248	SDNN	0.06002128475128934	SDNN	0.0895622746236924
	HRmathHRminDifference	0.03178067723293	HRmathHRminDifference	0.074491832751279	HRmathHRminDifference	0.03840510248297819	HRmathHRminDifference	0.0180528876627868	HRmathHRminDifference	0.0383711586711849
	SDRR	0.10788064193790827	SDRR	0.07688773803701167	SDRR	0.0931158732230206	SDRR	0.07082269515740823	SDRR	0.10477483476288602
	HRV_div_aveHR	0.04001443616490957	HRV_div_aveHR	0.030962664302892482	HRV_div_aveHR	0.0300403650530412967	HRV_div_aveHR	0.011234154940388097	HRV_div_aveHR	0.05291963673221549
	HRV_Math_MovELast15	0.0612405286750	HRV_Math_MovELast15	0.05009323868858	HRV_Math_MovELast15	0.0641386396385815	HRV_Math_MovELast15	0.0231386181514	HRV_Math_MovELast15	0.08066725735659291
	15secHR_Rave_div_aveHR	0.064703633661495	15secHR_Rave_div_aveHR	0.06223883388026	15secHR_Rave_div_aveHR	0.0769723405223583	15secHR_Rave_div_aveHR	0.06121676233943	15secHR_Rave_div_aveHR	0.06748734278423763
	10secHR_Rave_div_aveHR	0.0646315323010	10secHR_Rave_div_aveHR	0.06081718271500	10secHR_Rave_div_aveHR	0.07314267992493297	10secHR_Rave_div_aveHR	0.05811582524048	10secHR_Rave_div_aveHR	0.06736313437066832
	15secHR_Rave_div_aveHR	0.0702685784808	15secHR_Rave_div_aveHR	0.0673896457314	15secHR_Rave_div_aveHR	0.07910896955386915	15secHR_Rave_div_aveHR	0.05831188014722	15secHR_Rave_div_aveHR	0.0725932652930639
All features	F		G		H		I		All	
	MFR	0.1891187524239326	MFR	0.1674354335076363	MFR	0.1479536426127238	MFR	0.18295070184222128	MFR	0.13736584061397448
	MHR	0.144025651381705	MHR	0.202381372894144	MHR	0.16370949021943	MHR	0.18111251388309258	MHR	0.14563205580683
	HRV	0.0072530961319557	HRV	0.010515971109746256	HRV	0.0297494112325937	HRV	0.0256216895912041805	HRV	0.01600698691234842
	RMSSD	0.09056231574022647	RMSSD	0.07649662363027541	RMSSD	0.0775083090450053	RMSSD	0.07407328408429892	RMSSD	0.081958635010215717
	SDNN	0.07545862213818042	SDNN	0.0827147360094102	SDNN	0.0810547047002347	SDNN	0.06755238009169305	SDNN	0.0389438933810364
	HRmathHRminDifference	0.0236441137083846	HRmathHRminDifference	0.013670570112423	HRmathHRminDifference	0.04125844260773284	HRmathHRminDifference	0.032285384076382	HRmathHRminDifference	0.04168370021148884
	SDRR	0.106246868838765	SDRR	0.10437803891354806	SDRR	0.10145613412952137	SDRR	0.1153997845106357	SDRR	0.112871281523728
	HRV_div_aveHR	0.0173734732394655	HRV_div_aveHR	0.02086927209518894	HRV_div_aveHR	0.06274056085116742	HRV_div_aveHR	0.04211330854737546	HRV_div_aveHR	0.03773253084139524
	HRV_Math_MovELast15	0.04636215867163772	HRV_Math_MovELast15	0.05283953158997772	HRV_Math_MovELast15	0.0765034208178328	HRV_Math_MovELast15	0.07386245340151505	HRV_Math_MovELast15	0.062360324953115522
	15secHR_Rave_div_aveHR	0.07602541213386	15secHR_Rave_div_aveHR	0.076022468949549	15secHR_Rave_div_aveHR	0.07056322154253	15secHR_Rave_div_aveHR	0.0620339562179484	15secHR_Rave_div_aveHR	0.0897216561514136
	10secHR_Rave_div_aveHR	0.0754862213818042	10secHR_Rave_div_aveHR	0.0756713918371547	10secHR_Rave_div_aveHR	0.06887406454971448	10secHR_Rave_div_aveHR	0.06522963599152145	10secHR_Rave_div_aveHR	0.0853951554715879
	15secHR_Rave_div_aveHR	0.07150077304104457	15secHR_Rave_div_aveHR	0.09178350315715625	15secHR_Rave_div_aveHR	0.0749530368810327	15secHR_Rave_div_aveHR	0.07102880119803506	15secHR_Rave_div_aveHR	0.0879464165889719

#### 11.1.2 חמש מדדים מובילים

	A		B		C		D		E	
Top 5	MFR	0.22213244211435004	MFR	0.25602432386018686	MFR	0.22704558105867634	MFR	0.29830192684601636	MFR	0.23354606192655133
	MHR	0.31402382304354987	MHR	0.368983904538508	MHR	0.3189823245294384	MHR	0.40125136013378185	MHR	0.2600611279428388
	RMSSD	0.16809744963454044	RMSSD	0.11693312270884171	RMSSD	0.12857685927010565	RMSSD	0.0869565412433885	RMSSD	0.15682539049118288
	SDNN	0.130787462236341	SDNN	0.12356033016514595	SDNN	0.1533897842839634	SDNN	0.09575273782813812	SDNN	0.1579141494304237
	SDRR	0.18297282397051234	SDRR	0.1426258328089747	SDRR	0.172006796876078987	SDRR	0.11513742106772526	SDRR	0.19153656214938408
Top 5	F		G		H		I		All	
	MFR	0.253850127262863	MFR	0.2251409381995676	MFR	0.24327283312974105	MFR	0.26381767638164727	MFR	0.2168150723978173
	MHR	0.282139108581568195	MHR	0.287351384454715	MHR	0.2407782215440553	MHR	0.2574639015461923	MHR	0.2379892177636857
	RMSSD	0.1416385843732677	RMSSD	0.13314483493747891	RMSSD	0.13308458437074086	RMSSD	0.13468182211012686	RMSSD	0.1626403794744534
	SDNN	0.1340121001953887	SDNN	0.1672113472326803	SDNN	0.15886705202379167	SDNN	0.13543317006849317	SDNN	0.1761270053132658
	SDRR	0.178200164305439	SDRR	0.1877277335015114	SDRR	0.203027038943671	SDRR	0.20872323928511977	SDRR	0.2064276482998063

#### 11.1.3 שלוש מדדים מובילים

	A		B		C		D		E	
	MFR	0.3231220931271176	MFR	0.35473224499804215	MFR	0.3462736209294777	MFR	0.3984585323929363	MFR	0.33741328058202081
Top 3	MHR	0.35280506308030515	MHR	0.4159432793016243	MHR	0.37342957078226906	MHR	0.4092838384641616	MHR	0.36024607834240147
	SDRR	0.2839728373725717	SDRR	0.2232447163995426	SDRR	0.280308083182532	SDRR	0.132257529142902	SDRR	0.3023406365755605
Top 3	F		G		H		I		All	
	MFR	0.3456857165265067	MFR	0.335357734311304	MFR	0.348170311778877	MFR	0.34720782571621717	MFR	0.3317803971135591
	MHR	0.372287457340737	MHR	0.3633763754872526	MHR	0.343497170919728	MHR	0.34380874727382461	MHR	0.3476822212281021
	SDRR	0.282075537602789	SDRR	0.30124785107763495	SDRR	0.31037897112914956	SDRR	0.30370747014545387	SDRR	0.220597816583386

### 11.2 תוצאות מובהקות המדדים לפי פונקציית importance אחרי טרנספורמציה

	A		B		C		D		E	
	MFR	0.17145702943444954	MFR	0.23395470216338143	MFR	0.19008136723395753	MFR	0.2640367320098836	MFR	0.182127344244442602
Top 5 + Logs	MHR	0.25868439444913057	MHR	0.2773803542931876	MHR	0.235194676425077	MHR	0.32571679651858025	MHR	0.18894716290425932
	RMSSD	0.08234132393767678	RMSSD	0.07801890046627967	RMSSD	0.0843678093321203	RMSSD	0.0614894858098619588	RMSSD	0.097941129961359563
	SDNN	0.08303144862467733	SDNN	0.08065990180146396	SDNN	0.09486458711354737	SDNN	0.06618216124118691	SDNN	0.09787990405187479
	SDRR	0.11302633766892382	SDRR	0.08920747102658251	SDRR	0.10800947926782973	SDRR	0.07853819617267178	SDRR	0.12063983177422227
	Log_SDRR	0.11074769369836064	Log_SDRR	0.08617737629978145	Log_SDRR	0.10376991035788329	Log_SDRR	0.0781902508940253	Log_SDRR	0.1180570747309595
	Log_RMSSD	0.0906790200635236	Log_RMSSD	0.07477138441417487	Log_RMSSD	0.08380105453834664	Log_RMSSD	0.06239508514104842	Log_RMSSD	0.08753411031385522
	Log_SDNN	0.080132729779569	Log_SDNN	0.07885290953587243	Log_SDNN	0.10010101915299796	Log_SDNN	0.06845228212271175	Log_SDNN	0.0986231741051668
Top 5 + Logs	F		G		H		I		All	
	MFR	0.2071860720479413	MFR	0.173402680142975124	MFR	0.18845843819795124	MFR	0.20304538198629073	MFR	0.16444951866307073
	MHR	0.2388340426888833	MHR	0.2278309594425064	MHR	0.18845843819795124	MHR	0.2117321442543553	MHR	0.1811705849367879
	RMSSD	0.08749154873824568	RMSSD	0.08485434982323289	RMSSD	0.08262822812273318	RMSSD	0.08725575032382023	RMSSD	0.09894203176731306
	SDNN	0.0844223513380858	SDNN	0.10191538003123632	SDNN	0.101936389584136	SDNN	0.0826886836868589	SDNN	0.108620281205199
	SDRR	0.1455118048320513	SDRR	0.1178761080270281	SDRR	0.117834898610470	SDRR	0.1482580320112038	SDRR	0.1206849789255373
	Log_SDRR	0.1067953173438304	Log_SDRR	0.11245698893296091	Log_SDRR	0.11482663803168079	Log_SDRR	0.1282182343839047	Log_SDRR	0.1196716410130197
	Log_RMSSD	0.0828043691878484	Log_RMSSD	0.0847096336839193	Log_RMSSD	0.0970279415137532	Log_RMSSD	0.08380317216599075	Log_RMSSD	0.0971927185838486
	Log_SDNN	0.0857967359004487	Log_SDNN	0.1018150516275838	Log_SDNN	0.10193247025015527	Log_SDNN	0.08125068772062168	Log_SDNN	0.1058817194225468

### 11.3 תוצאות מובהקות המדדים לפי פונקציית importance – לאחר איחוד קטגוריות

	A		B		C		D		E	
High_Medium & Low	MFR	0.2406191019686345	MFR	0.3308540407306345	MFR	0.228320089019883	MFR	0.29489287166973627	MFR	0.21303867187737976
	MHR	0.342843437168223	MHR	0.39756304511958956	MHR	0.2951942671680656	MHR	0.38251335353888824	MHR	0.251158580141659
	Log_SDRR	0.17948704211917038	Log_SDRR	0.0873884295939406	Log_SDRR	0.1750806913537907	Log_SDRR	0.12248033640531714	Log_SDRR	0.18230001890537542
	Log_RMSSD	0.12474954558906448	Log_RMSSD	0.0971258227828283	Log_RMSSD	0.1341668018937525	Log_RMSSD	0.09111973892019322	Log_RMSSD	0.1639870466210882
	Log_SDNN	0.1125506861050641	Log_SDNN	0.08707220231153355	Log_SDNN	0.16586018287834726	Log_SDNN	0.1088978906786515	Log_SDNN	0.1897638964545222
High_Medium & Low	MFR	0.233496405793983	MFR	0.1962950744020552	MFR	0.2221608592448468	MFR	0.205639042428545	MFR	0.21464397095104
	MHR	0.2338011811116482	MHR	0.2415940534011058	MHR	0.2618723879309763	MHR	0.251260987919144	MHR	0.235903518861119354
	Log_SDRR	0.1306593102039282	Log_SDRR	0.2267938479555896	Log_SDRR	0.2198414544257493	Log_SDRR	0.12212260129274	Log_SDRR	0.2028406350888483
	Log_RMSSD	0.167149449623279	Log_RMSSD	0.148044094046358	Log_RMSSD	0.1526365805907683	Log_RMSSD	0.1017120016288242	Log_RMSSD	0.1897502595392088
	Log_SDNN	0.1257828301020201	Log_SDNN	0.1874278104613201	Log_SDNN	0.183440161020201	Log_SDNN	0.134040161020201	Log_SDNN	0.1897502595392088

## נספח 12 - ניתוח תוצאות אלגוריתם Multinomial Regression

### 12.1. כלל המדדים

All_Feature_P value	A		B		C		D		E		F		G		H		I		All	
	Low	Medium	Low	Medium	Low	Medium	Low	Medium	Low	Medium	Low	Medium	Low	Medium	Low	Medium	Low	Medium	Low	Medium
Const	0.00	0.84	0.01	0.24	0.03	0.03	0.80	0.00	0.32	0.00	0.17	0.47	0.00	0.01	0.55	0.11	0.00	0.01	0.00	0.00
MRR	0.14	0.89	0.10	0.39	0.55	0.07	0.43	0.00	0.20	0.00	0.37	0.34	0.51	0.70	0.90	0.00	0.12	0.26	0.78	0.07
MHR	0.00	0.92	0.00	0.07	0.01	0.01	0.39	0.00	0.06	0.00	0.02	0.22	0.00	0.01	0.42	0.08	0.00	0.01	0.00	0.00
HRV	0.92	0.08	0.25	0.01	0.16	0.31	0.89	0.06	0.15	0.09	0.07	0.13	0.77	0.88	0.22	0.22	0.24	0.34	0.00	0.02
RMSSD	0.79	0.35	0.87	0.68	0.41	0.37	0.02	0.00	0.63	0.79	0.00	0.31	0.00	0.01	0.84	0.47	0.34	0.44	0.01	0.28
SDNN	0.90	0.24	0.81	0.98	0.85	0.06	0.84	0.28	0.30	0.93	0.31	0.52	0.01	0.19	0.99	0.10	0.91	0.49	0.11	0.83
HRmaxHRminDifference	0.27	0.40	0.09	0.00	0.56	0.00	0.26	0.98	0.31	0.10	0.90	0.59	0.13	0.70	0.00	0.19	0.11	0.01	0.13	0.22
SDRR	0.00	0.58	0.39	0.33	0.70	0.99	0.77	0.35	0.07	0.87	0.03	0.76	0.29	0.36	0.86	0.21	0.00	0.32	0.15	0.18
HRV_div_aveHR	0.98	0.09	0.28	0.00	0.14	0.24	0.82	0.04	0.29	0.16	0.06	0.10	0.76	0.89	0.12	0.20	0.23	0.27	0.00	0.03
HRV_div_aveHR_MoveLast15	0.35	0.32	0.29	0.73	0.58	0.01	0.13	0.00	0.00	0.04	0.07	0.21	0.34	0.41	0.09	0.32	0.10	0.88	0.08	0.18
last5secHRave_div_aveHR	0.77	0.88	0.76	0.24	0.98	0.94	0.73	0.00	0.31	0.04	0.88	0.93	0.96	0.99	0.87	0.21	0.05	0.45	0.33	0.68
last10secHRave_div_aveHR	0.14	0.99	0.30	0.16	0.73	0.64	0.65	0.91	0.32	0.16	0.68	0.87	0.65	0.72	0.45	0.92	0.63	0.66	0.57	0.93
last15secHRave_div_aveHR	0.01	0.94	0.03	0.27	0.30	0.53	0.60	0.10	0.79	0.64	0.30	0.55	0.11	0.23	0.23	0.60	0.06	0.12	0.69	0.74

### 12.1.2 חמישה מדדים מובילים

Top 5_Feature_P value	A		B		C		D		E		F		G		H		I		All	
	Low	Medium	Low	Medium	Low	Medium	Low	Medium	Low	Medium	Low	Medium	Low	Medium	Low	Medium	Low	Medium	Low	Medium
Const	0.09	0.93	0.10	0.77	0.14	0.23	0.36	0.00	0.64	0.00	0.66	0.57	0.00	0.02	0.92	0.02	0.00	0.01	0.01	0.01
MRR	0.05	0.92	0.06	0.63	0.72	0.35	0.97	0.00	0.27	0.02	0.14	0.47	0.95	0.76	0.98	0.00	0.19	0.25	0.66	0.17
MHR	0.00	0.91	0.00	0.94	0.00	0.14	0.06	0.00	0.04	0.00	0.02	0.08	0.00	0.00	0.80	0.20	0.00	0.00	0.00	0.00
RMSSD	0.89	0.27	0.96	0.38	0.49	0.28	0.03	0.00	0.37	0.54	0.00	0.30	0.00	0.00	0.61	0.35	0.23	0.37	0.01	0.26
SDNN	0.61	0.27	0.82	0.62	0.85	0.01	0.70	0.58	0.30	0.99	0.63	0.40	0.00	0.05	0.86	0.12	0.99	0.65	0.26	0.91
SDRR	0.00	0.13	0.40	0.39	0.68	0.08	0.25	0.16	0.46	0.82	0.00	0.75	0.01	0.14	0.02	0.55	0.00	0.90	0.24	0.27

### 12.1.3 שלוש מדדים מובילים

Top 3_Feature_P value	A		B		C		D		E		F		G		H		I		All	
	Low	Medium	Low	Medium	Low	Medium	Low	Medium	Low	Medium	Low	Medium	Low	Medium	Low	Medium	Low	Medium	Low	Medium
Const	0.06	0.89	0.09	0.81	0.15	0.18	0.13	0.00	0.70	0.00	0.51	0.43	0.01	0.04	0.85	0.04	0.00	0.01	0.01	0.01
MRR	0.06	0.96	0.05	0.68	0.73	0.28	0.65	0.00	0.18	0.03	0.14	0.56	0.75	0.97	0.89	0.00	0.21	0.22	0.56	0.19
MHR	0.00	0.75	0.00	0.96	0.00	0.11	0.01	0.00	0.05	0.00	0.01	0.04	0.00	0.00	0.75	0.26	0.00	0.00	0.00	0.00
SDRR	0.00	0.02	0.43	0.39	0.84	0.00	0.07	0.05	0.88	0.14	0.00	0.32	0.31	0.03	0.01	0.21	0.00	0.97	0.08	0.10

## 12.2 תוצאות מובהקות המדדים לפי פונקציית Pvalue – feature selction - אחרי טרנספורמציה

Top 5 features + Logs	A		B		C		D		E		F		G		H		I		All	
	Low	Medium	Low	Medium	Low	Medium	Low	Medium	Low	Medium	Low	Medium	Low	Medium	Low	Medium	Low	Medium	Low	Medium
Const	0.49	0.45	0.14	0.64	0.17	0.66	0.70	0.00	0.30	0.00	0.98	0.87	0.02	0.02	0.10	0.02	0.00	0.00	0.01	0.01
MRR	0.10	0.07	0.02	0.12	0.73	0.43	0.72	0.00	0.16	0.04	0.13	0.42	0.69	0.57	1.00	0.00	0.53	0.26	0.77	0.16
MHR	0.08	0.04	0.00	0.53	0.00	0.35	0.13	0.00	0.02	0.00	0.04	0.14	0.00	0.00	0.56	0.15	0.00	0.00	0.00	0.00
Log_SDRR	0.48	0.00	0.01	0.00	0.42	0.00	0.09	0.01	0.00	0.00	0.00	0.01	0.00	0.77	0.00	0.45	0.00	0.03	0.28	0.44
Log_RMSSD	0.03	0.80	0.15	0.48	0.79	0.19	0.99	0.11	0.30	0.03	0.69	0.25	0.09	0.20	0.78	0.78	0.32	0.41	0.01	0.11
Log_SDNN	0.09	0.04	0.45	0.68	0.66	0.88	0.62	0.30	0.53	0.26	0.66	0.88	0.51	0.73	0.57	0.64	0.34	0.58	0.41	0.81

## 12.3 תוצאות מובהקות המדדים לפי פונקציית Pvalue – feature selction - אחרי איחוד קטגוריות

Top 5 features + Logs	A	B	C	D	E	F	G	H	I	All
	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low
Const	0.21	0.00	0.09	0.00	0.64	0.72	0.40	0.02	0.66	0.25
MRR	0.09	0.00	0.88	0.02	0.01	0.34	0.80	0.00	0.00	0.32
MHR	0.00	0.00	0.00	0.00	0.72	0.43	0.04	0.00	0.00	0.00
Log_SDRR	0.00	0.10	0.28	0.85	0.01	0.00	0.00	0.00	0.03	0.13
Log_RMSSD	0.47	0.34	0.19	0.69	0.73	0.05	0.04	0.31	0.86	0.00
Log_SDNN	0.19	0.81	0.10	0.77	0.17	0.28	0.53	0.45	0.21	0.64

### נספח 13 – אחוזי הניבוי במודל Random Forest

Subject	All features	Top 5 features	Top 3 features	Top 5 features+Logs	High_Medium & Low
A	0.528	0.501	0.534	0.512	0.757
B	0.727	0.641	0.586	0.666	0.851
C	0.545	0.523	0.523	0.509	0.694
D	0.639	0.622	0.647	0.633	0.660
E	0.471	0.471	0.462	0.486	0.510
F	0.506	0.523	0.523	0.5	0.596
G	0.516	0.552	0.516	0.513	0.612
H	0.443	0.463	0.424	0.463	0.534
I	0.699	0.686	0.688	0.696	0.902
All	0.492	0.457	0.463	0.455	0.612

### נספח 14 – אחוזי הניבוי במודל Multinomial Regression

Subject	All features	Top 5 features	Top 3 features	Top 5 features+Logs	High_Medium & Low
A	0.52	0.53	0.52	0.52	0.71
B	0.56	0.56	0.52	0.55	0.79
C	0.44	0.43	0.43	0.43	0.61
D	0.43	0.44	0.42	0.44	0.51
E	0.49	0.49	0.49	0.49	0.50
F	0.31	0.32	0.34	0.32	0.59
G	0.49	0.48	0.46	0.46	0.55
H	0.48	0.46	0.47	0.47	0.51
I	0.65	0.64	0.65	0.64	0.84
All	0.43	0.43	0.43	0.42	0.57

## נספח 15 – אחוזי הצלחה בניבוי הקטגוריות

A		<b>predicted_High</b>	<b>predicted_Low</b>	<b>predicted_Medium</b>	<b>Type_Accuracy</b>
	High	51	26	45	0.418033
	Medium	35	42	45	0.368852
	Low	10	96	13	0.806723
B		<b>predicted_High</b>	<b>predicted_Low</b>	<b>predicted_Medium</b>	<b>Type_Accuracy</b>
	High	52	12	58	0.42623
	Medium	35	11	73	0.613445
	Low	14	92	16	0.754098
C		<b>predicted_High</b>	<b>predicted_Low</b>	<b>predicted_Medium</b>	<b>Type_Accuracy</b>
	High	50	22	47	0.420168
	Medium	24	32	69	0.552
	Low	20	74	25	0.621849
D		<b>predicted_High</b>	<b>predicted_Low</b>	<b>predicted_Medium</b>	<b>Type_Accuracy</b>
	High	82	29	18	0.635659
	Medium	7	29	83	0.697479
	Low	28	75	23	0.595238
E		<b>predicted_High</b>	<b>predicted_Low</b>	<b>predicted_Medium</b>	<b>Type_Accuracy</b>
	High	69	28	20	0.589744
	Medium	23	39	46	0.425926
	Low	29	37	40	0.349057
F		<b>predicted_High</b>	<b>predicted_Low</b>	<b>predicted_Medium</b>	<b>Type_Accuracy</b>
	High	51	27	16	0.542553
	Medium	24	34	38	0.395833
	Low	33	67	12	0.598214
G		<b>predicted_High</b>	<b>predicted_Low</b>	<b>predicted_Medium</b>	<b>Type_Accuracy</b>
	High	71	21	32	0.572581
	Medium	9	44	44	0.453608
	Low	23	52	37	0.464286
H		<b>predicted_High</b>	<b>predicted_Low</b>	<b>predicted_Medium</b>	<b>Type_Accuracy</b>
	High	56	48	27	0.427481
	Medium	25	50	42	0.358974
	Low	22	62	33	0.529915
I		<b>predicted_High</b>	<b>predicted_Low</b>	<b>predicted_Medium</b>	<b>Type_Accuracy</b>
	High	84	0	32	0.724138
	Medium	44	12	61	0.521368
	Low	14	119	13	0.815068
All		<b>predicted_High</b>	<b>predicted_Low</b>	<b>predicted_Medium</b>	<b>Type_Accuracy</b>
	High	482	291	301	0.44879
	Medium	280	312	428	0.419608
	Low	252	561	266	0.519926

$$MHR = \frac{1}{N} \sum_{j=1}^N HR_j$$

$$MRR = \frac{1}{N} \sum_{j=1}^N RR_j$$

$$HRV = \frac{1}{N-1} \sum_{j=1}^N (HR_j - \overline{HR})^2$$

$$HR_{max} - HR_{min} = Max(HR) - Min(HR)$$

$$SDNN = \sqrt{\frac{1}{N-1} \sum_{j=1}^N (RR_j - \overline{RR})^2}$$

$$RMSSD = \sqrt{\frac{1}{N-1} \sum_{j=1}^N (RR_j - RR_{j+1})^2}$$