



סמינר בהנדסת
תוכנה 2020

נופר אלפסי

The Participants



51%

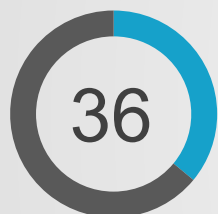
זוגות נשואים באושר



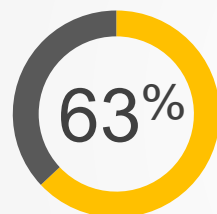
49%

זוגות שהיו נשואים
והתגרשו

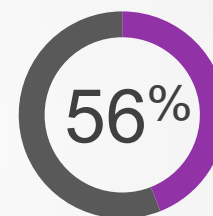
Participants Statistics



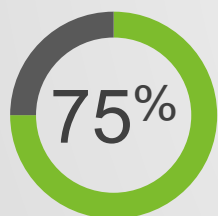
ממוצע גילאי
המשתתפים



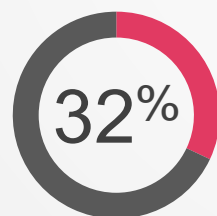
בעלי השכלה
אקדמאית



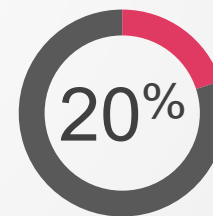
נישאו בנישואין
מסודרים (שידוך)



זוגות עם ילדים



מרוויחים מעל 9,000
ש"ח בחודש



מרוויחים מתחת ל-
5,000 בחודש

The Research

01

המחקר נערך בשנת
2018 בטורקיה.

02

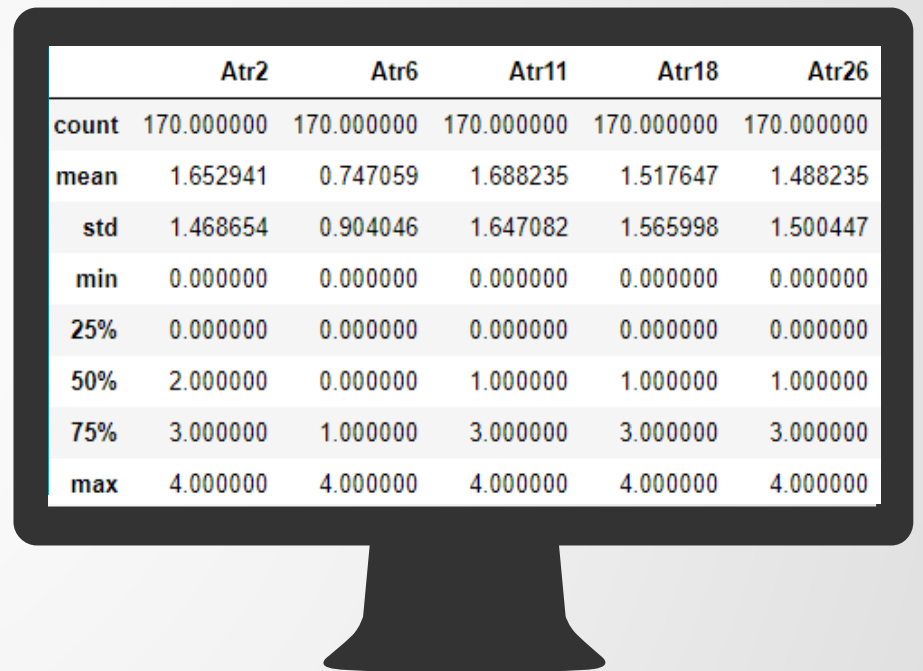
המשתתפים במחקר
תוחקרו בראיון פנים
מול פנים.

03

הזוגות הגרושים ענו
על השאלות בהתבסס
על זוגיותם כשהיו
נשואים.

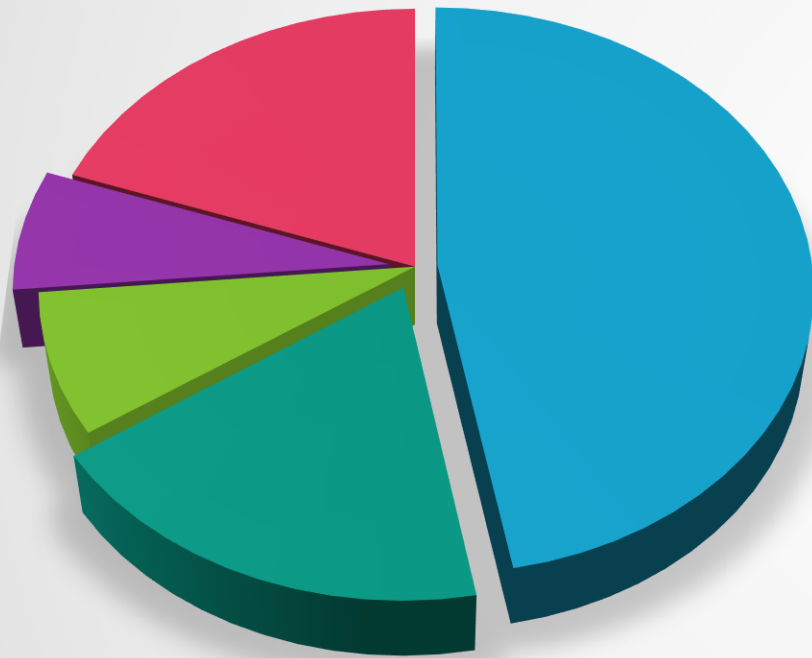
04

מבין הנשואים, נכללו
רק זוגות עם נשואים
מאושרים, ללא
מחשבה על גירושים.



	Atr2	Atr6	Atr11	Atr18	Atr26
count	170.000000	170.000000	170.000000	170.000000	170.000000
mean	1.652941	0.747059	1.688235	1.517647	1.488235
std	1.468654	0.904046	1.647082	1.565998	1.500447
min	0.000000	0.000000	0.000000	0.000000	0.000000
25%	0.000000	0.000000	0.000000	0.000000	0.000000
50%	2.000000	0.000000	1.000000	1.000000	1.000000
75%	3.000000	1.000000	3.000000	3.000000	3.000000
max	4.000000	4.000000	4.000000	4.000000	4.000000

Important Parameters



■ 1st Qtr ■ 2nd Qtr ■ 3rd Qtr ■ 4th Qtr ■ 5th Qtr





We're just starting a fight before I know what's going on





I know my wife's basic concerns


Most of our goals are common


I know exactly what my wife likes


I know how my wife wants to be taken care of when she's sick


I share the same views with my wife about being happy



I enjoy traveling with my wife


We don't have time at home as partners


We have intense arguments


I know my wife's favorite food


I know exactly what my wife likes


Sometimes it's good for me to leave the house for a while



Possible Answers

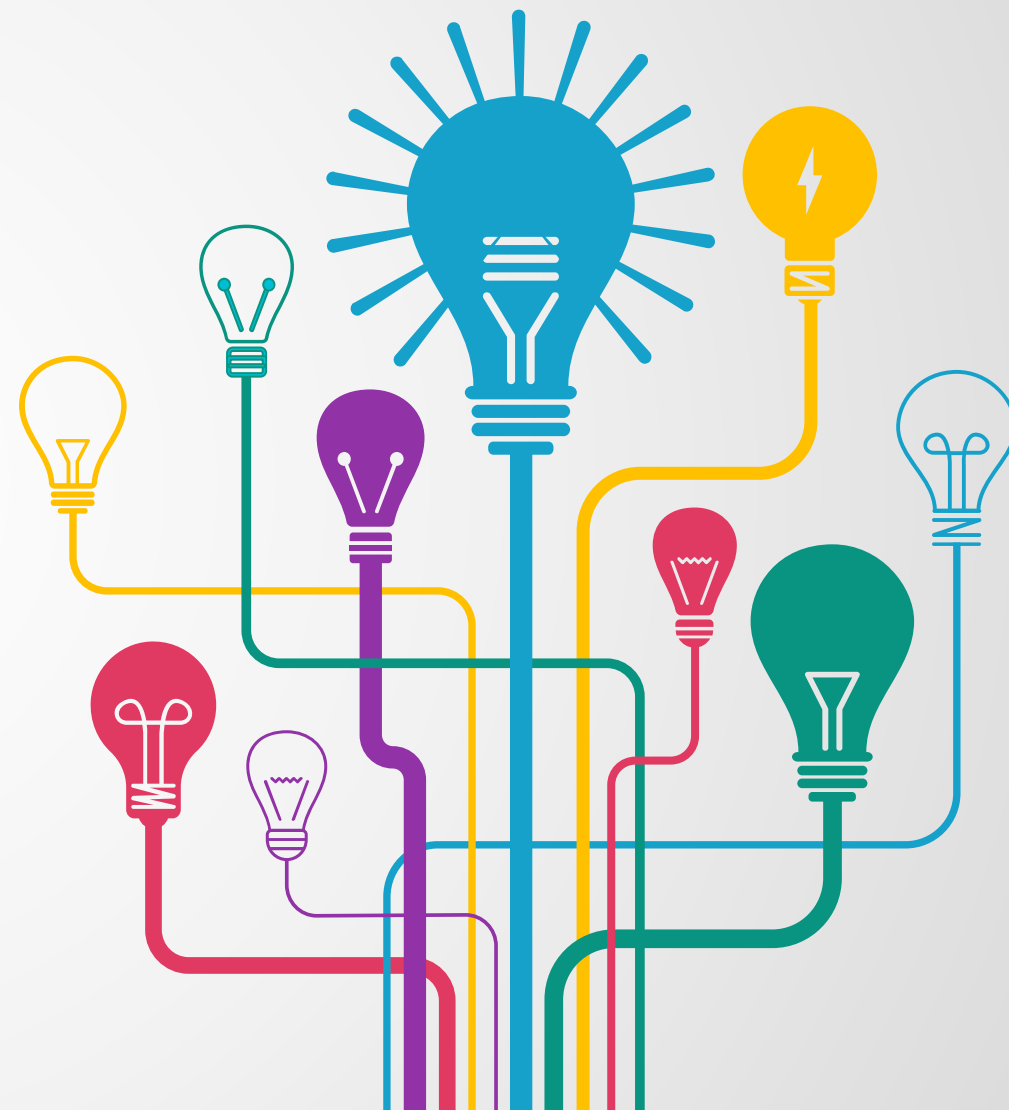


Hypothesis

במחקר שלנו נציג היפותזות לגבי השפעת פרמטרים שונים על סיכויי הגירושים של בני זוג.

אנו נבחן מאפיינים שונים במספר תחומים של מערכות יחסים של זוגות נשואים:

- כיצד הם באים לידי ביטוי בשיעורי הגירושים של הזוג?
- האם ניתן ניתן לנבא על פיהם גירושים עתידיים?



Research Questions

שאלת המחקר:

אילו מאפיינים או השקפות אישיות קשורים ביותר לגירושין?

- **אנו צופות כי השקפות חברתיות ופוליטיות מסוימות משפיעות על בריאות ואורך הנישואין.**
- **מאפיינים אלה יאפשרו לנו לחזות מי הסיכוי הטוב ביותר להתגרש או להישאר נשואים.**



Workflow

1

Dataset processing

אחרי קליטת המידע, נבחר את האלגוריתמים הטובים ביותר עבורנו

2

Algorithms selection

3

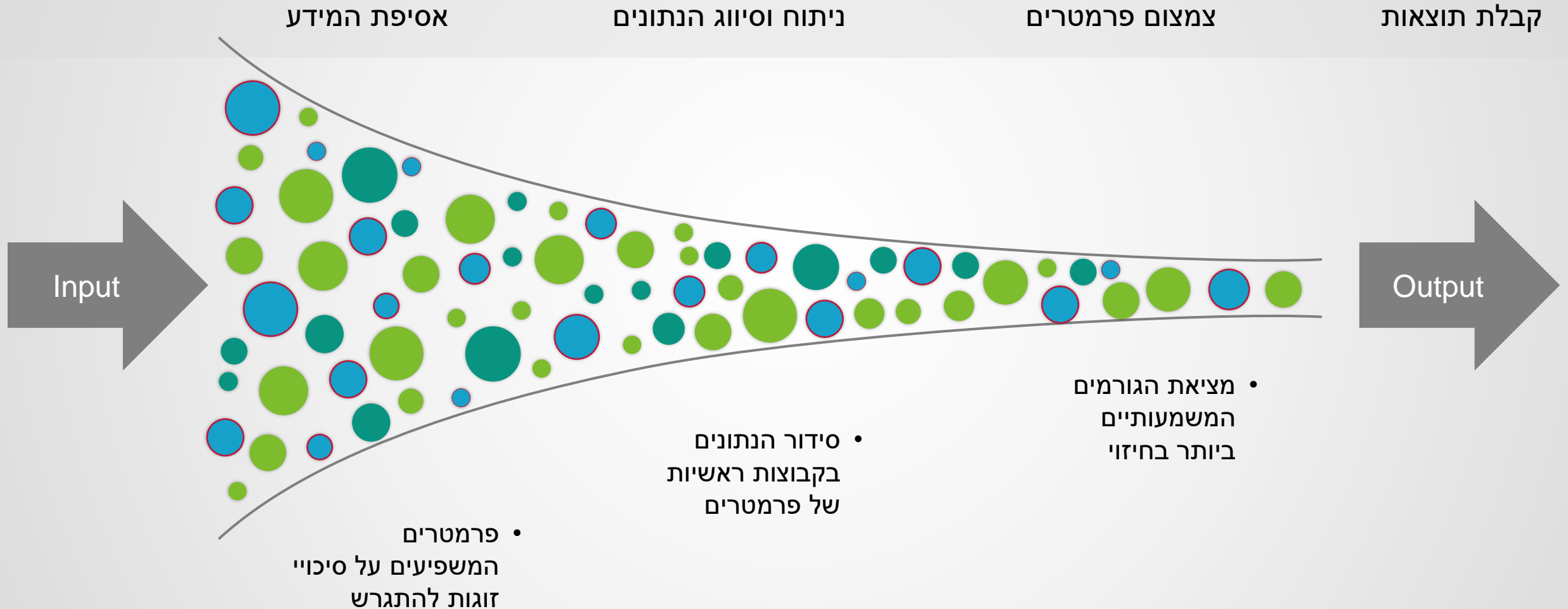
Features selection

לאחר בחירת האלגוריתמים, נבחר את התכונות המשמעותיות ביותר עבורנו

4

Run & get results

Our Model



Machine Learning

Supervised learning

מטרת הלמידה המונחית היא ללמוד על ההתפלגות המשותפת.

מתוך כך ניתן לנבא עבור אלמנט את התיוג שלו.

האלגוריתם "לומד" כלל המקשר בין תיאור הדוגמא לתווית המתאימה לה.

Classification

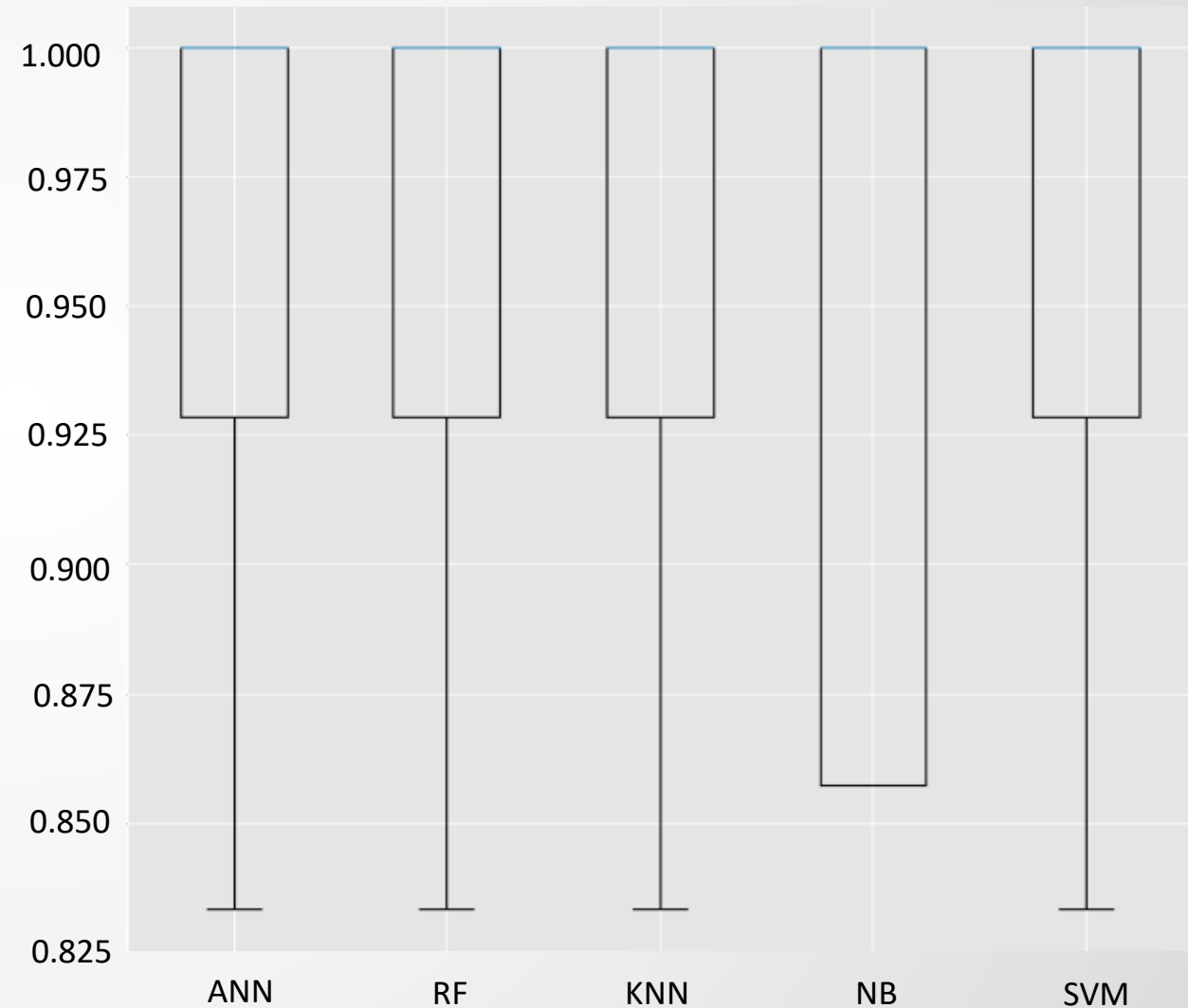
משמש לצורך חיזוי ערכים של המשתנה התלוי במצבים שונים

האלגוריתם יופעל על דוגמאות חדשות, שהתווית שלהן לא ידועה, ויחזה מהי התווית הנכונה לכל דוגמא.

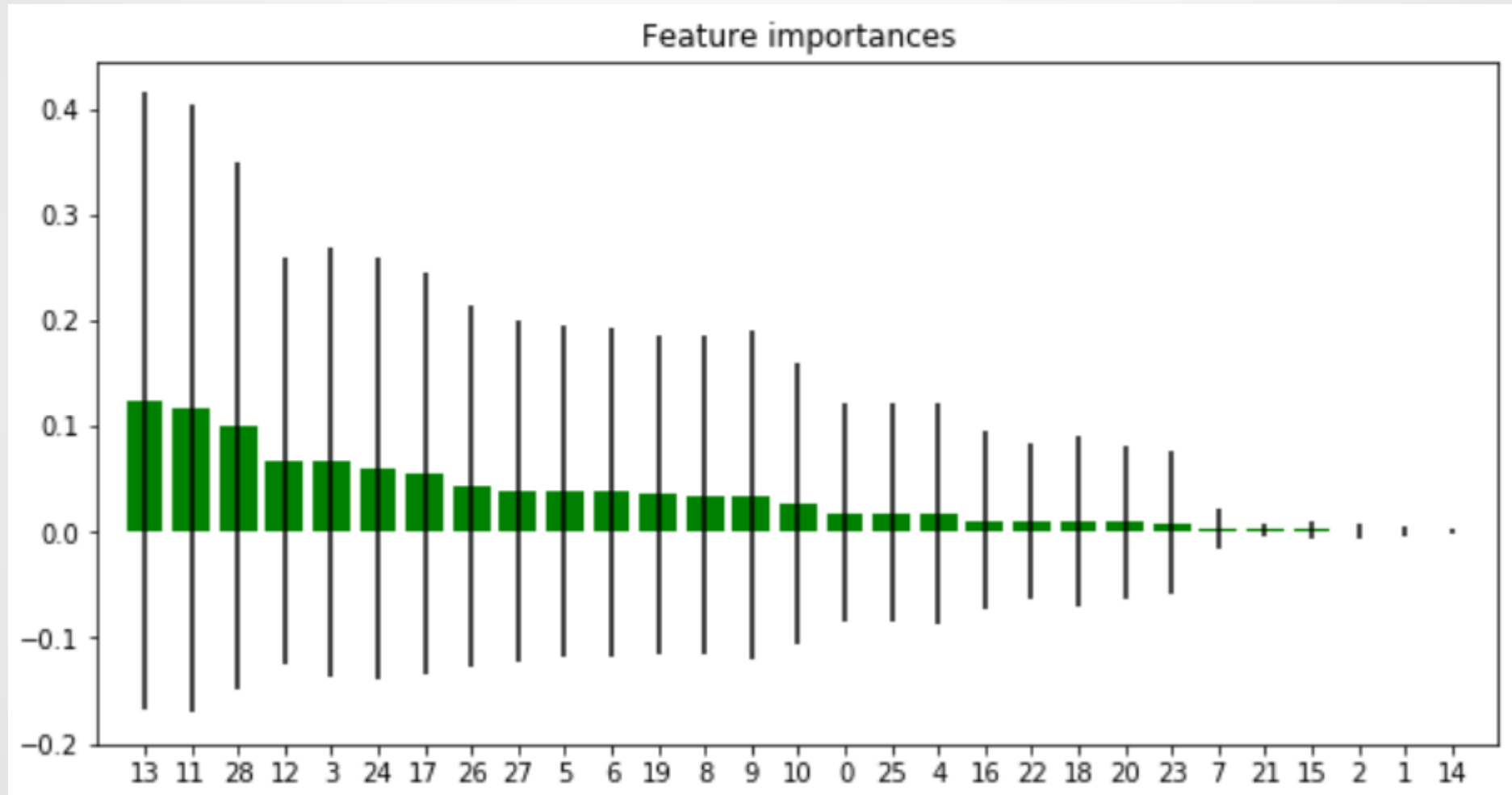
Step 1 – Algorithms Comparison

Results:

ANN: 0.960317
RF: 0.960317
KNN: 0.960317
NB: 0.952381
SVM: 0.960317



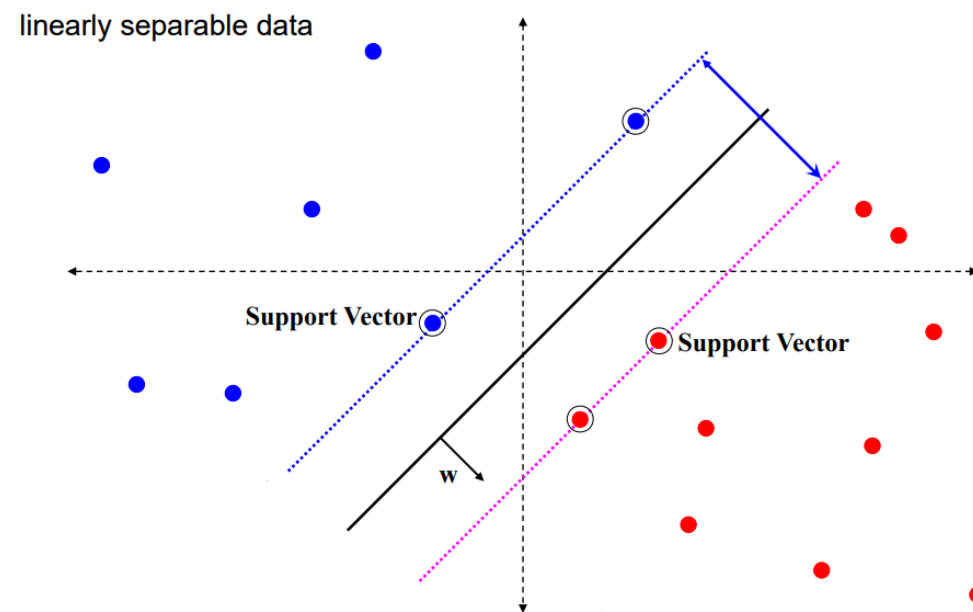
Step 2 – Features of Importance



Algorithms

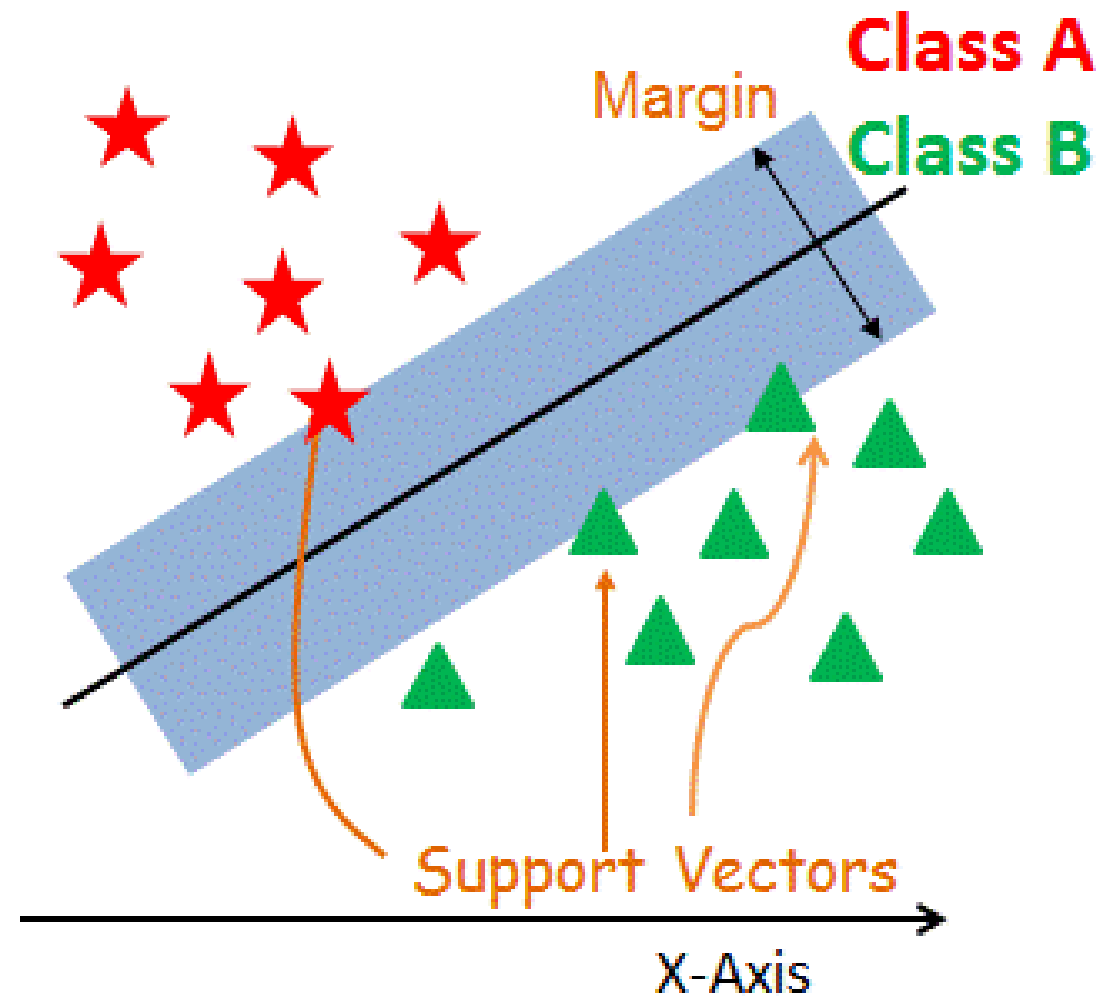
SVM – Support Vector Machine

- כל נקודה באוסף הנתונים מיוצגת על ידי וקטור בגודל קבוע.
- בגישה של SVM מחלקים את הנקודות (קלט) במרחב הווקטורי, באמצעות על מישור (על-מישור הוא מרחב $n-1$ ממדי בתוך מרחב n ממדי) כך שהמרווח, המרחק, בין אותו על-מישור המחלק את הנקודות לבין הנקודות הממוקמות הכי קרוב אליו, יהיה המרחק המקסימלי האפשרי.
- ה-SVM בונה על מישור שהוא המפריד הלינארי.
- מטרת ה-SVM היא להגדיל את המרחק בין המפריד הלינארי (מרחק שוליים), ולמעשה למצוא את המפריד בעל השוליים הרחבים ביותר.
- דוגמאות האימון המתלכדות עם מישורי השוליים נקראות וקטורים תומכים, ומכאן נגזר שם האלגוריתם.



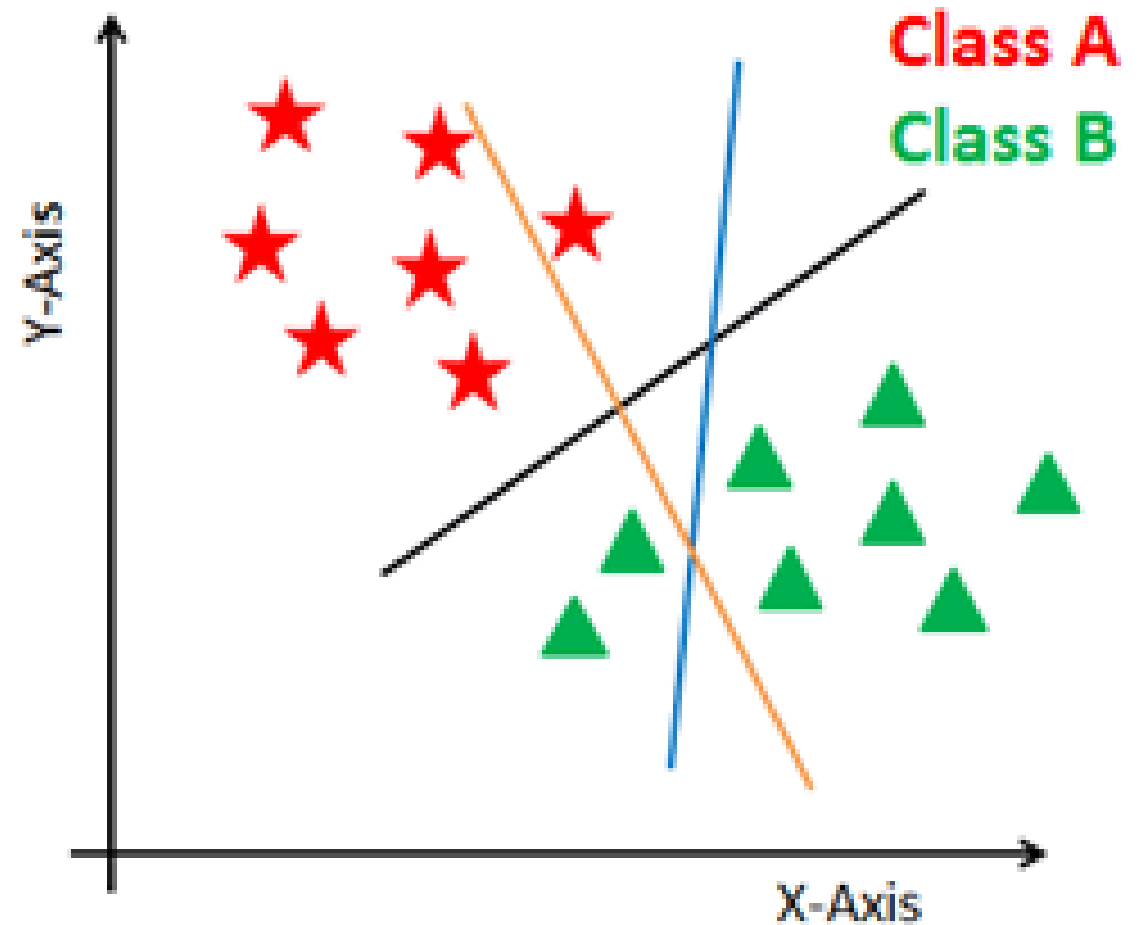
SVM – Support Vector Machine

- באלגוריתם זה אנו מתווים כל פריט נתונים כנקודה בחלל n ממדי (n הוא מספר התכונות) כאשר הערך של כל תכונה הוא הערך של קואורדינטות מסוימות.
- לאחר מכן, נבצע סיווג על ידי מציאת מסווג בין שתי מחלקות (נשואים וגרושים).
- האלגוריתם "לומד" כלל המקשר בין תאור הדוגמא לתוית המתאימה לה.
- האלגוריתם יופעל על דוגמאות חדשות, שהתוית שלהן לא ידועה, ויחזה מהי התוית הנכונה לכל דוגמא.



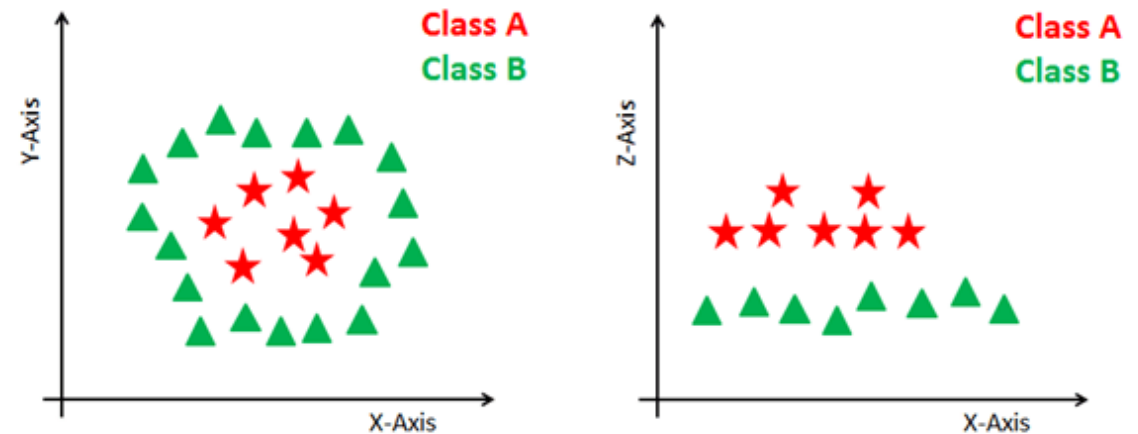
SVM – Advantages

- מסוגי SVM מציעים דיוק טוב ומבצעים ניבוי מהיר יותר בהשוואה לאלגוריתם של Naïve Bayes.
- הם משתמשים בפחות זיכרון מכיוון שהם משתמשים בתת קבוצה של נקודות אימון בשלב ההחלטה.
- SVM עובד טוב עם מרווח הפרדה ברור ועם מרחב ממדי גבוה.



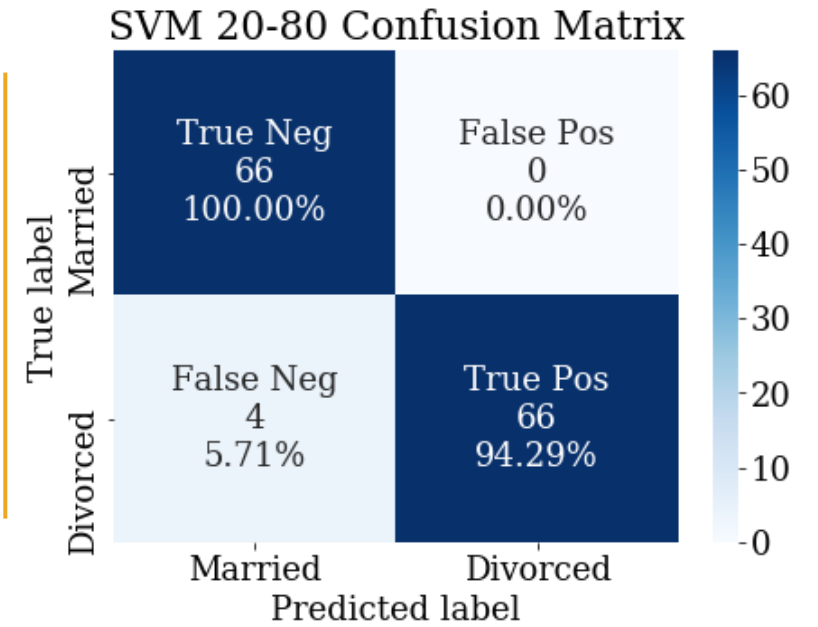
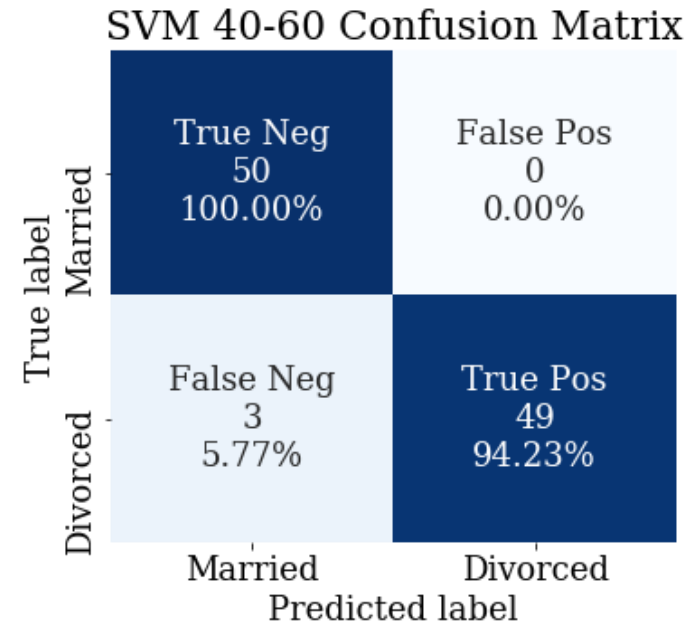
SVM – Disadvantages

- SVM אינו מתאים למערכות נתונים גדולות בגלל זמן האימונים הגבוה שלו וזה לוקח גם יותר זמן באימונים בהשוואה ל-Naive Bayes.
- SVM רגיש לסוג הגרעין שבו הוא משתמש.



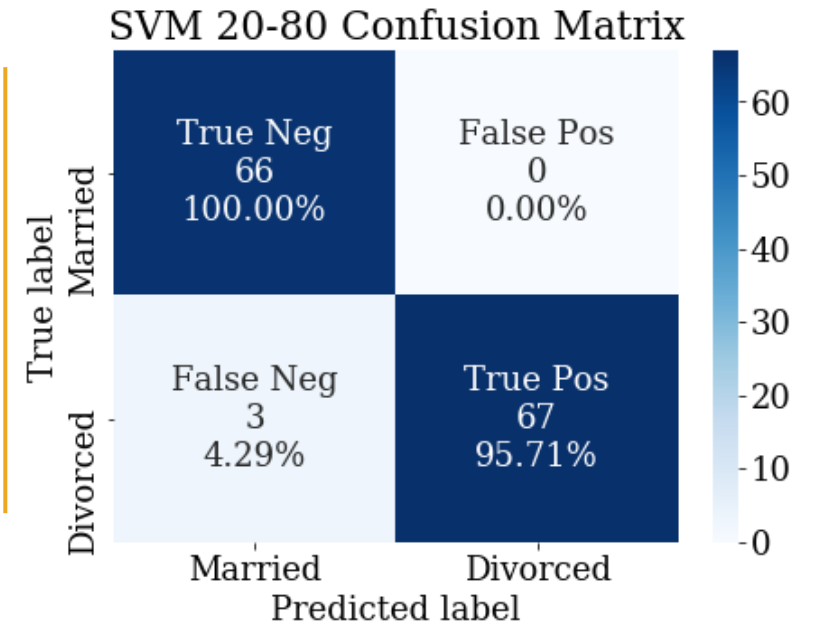
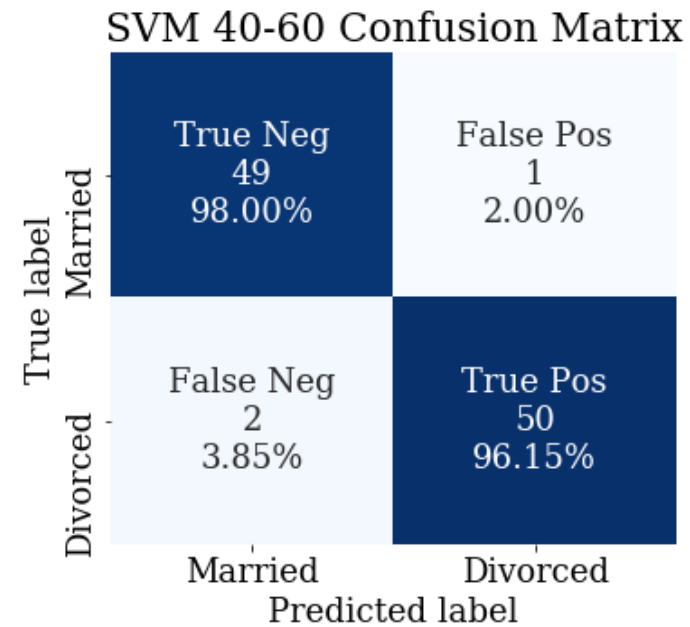
SVM

All 54 Features



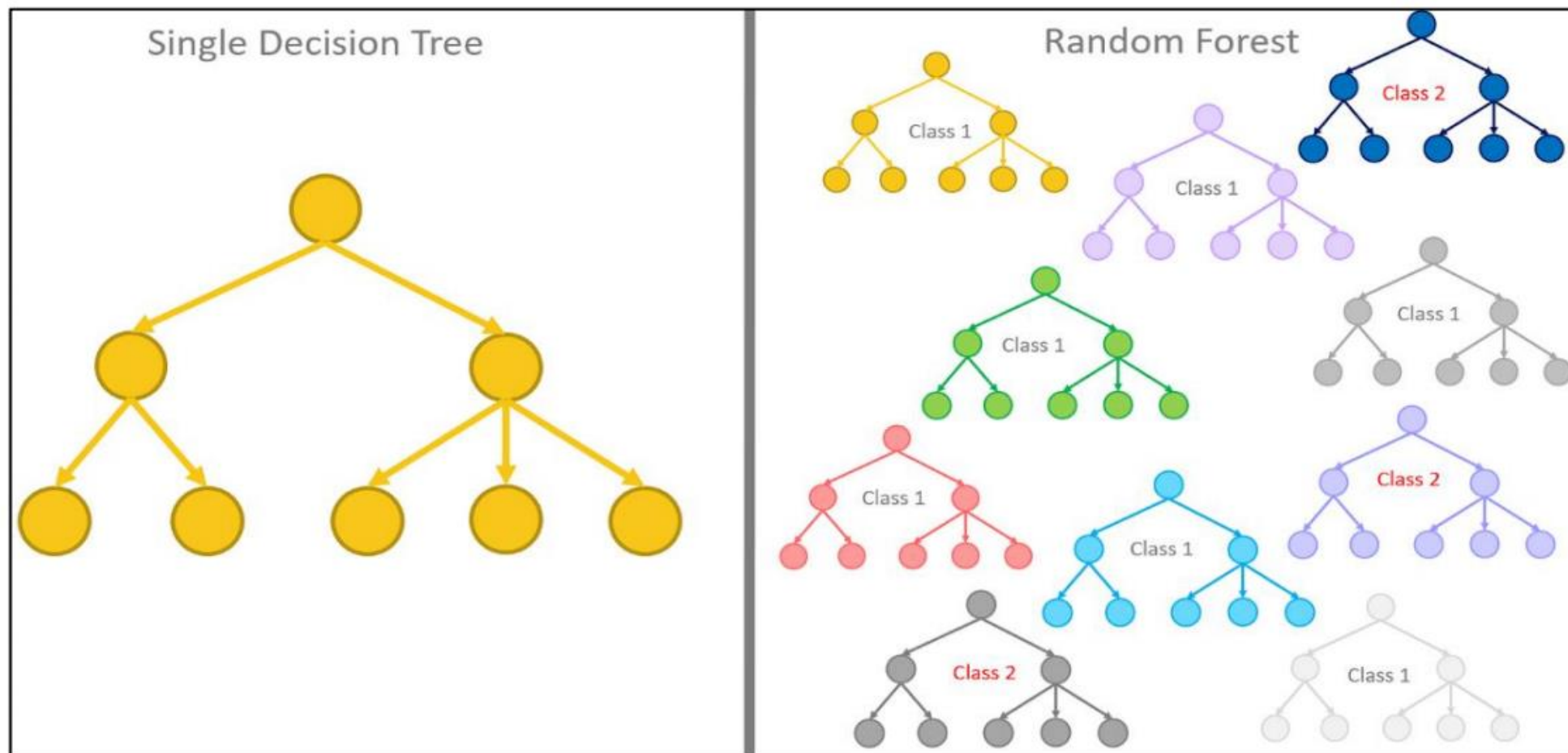
SVM

6 Selected Features



SVM – Support Vector Machine

The Code...

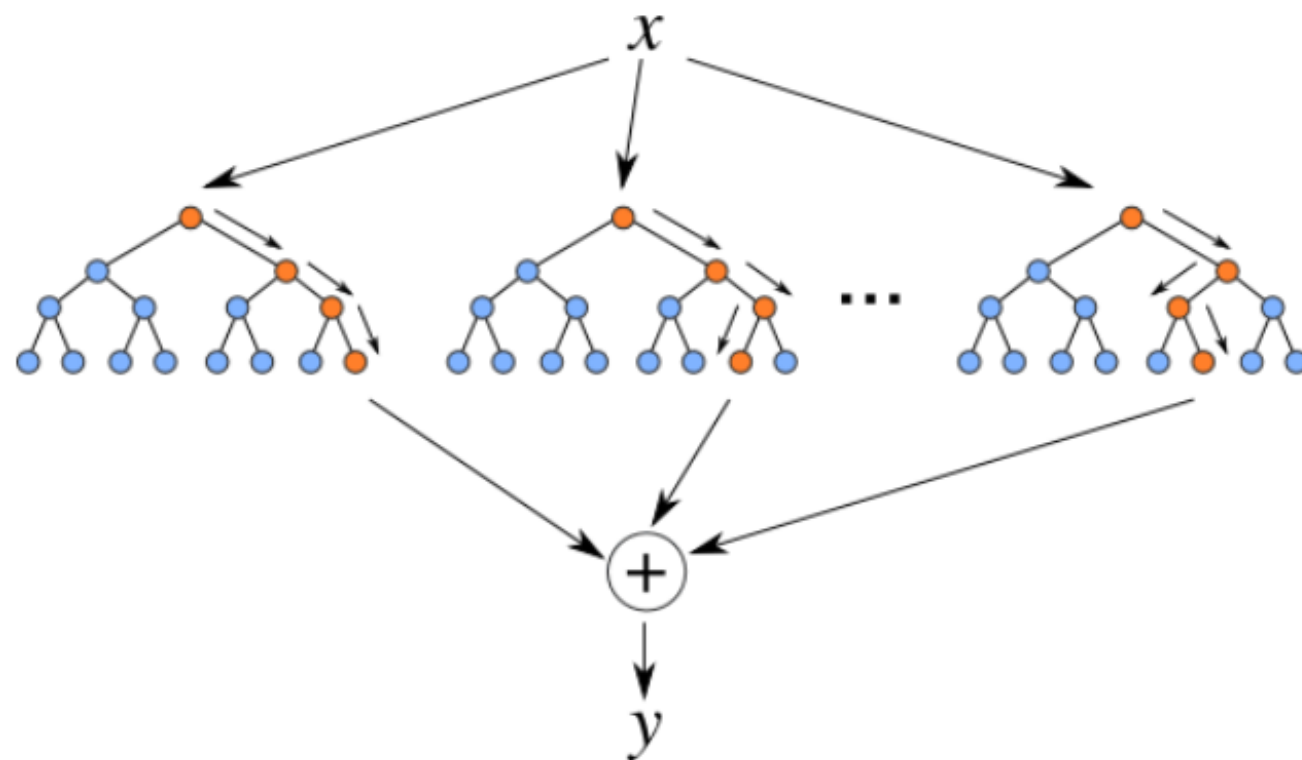


Random Forest

- יער הוא שילוב של כמה עצים על מנת לקבל תוצאה טובה יותר עבור בעיית הסיווג, במקום לאמן עץ אחד נצטרך לאמן כמה.

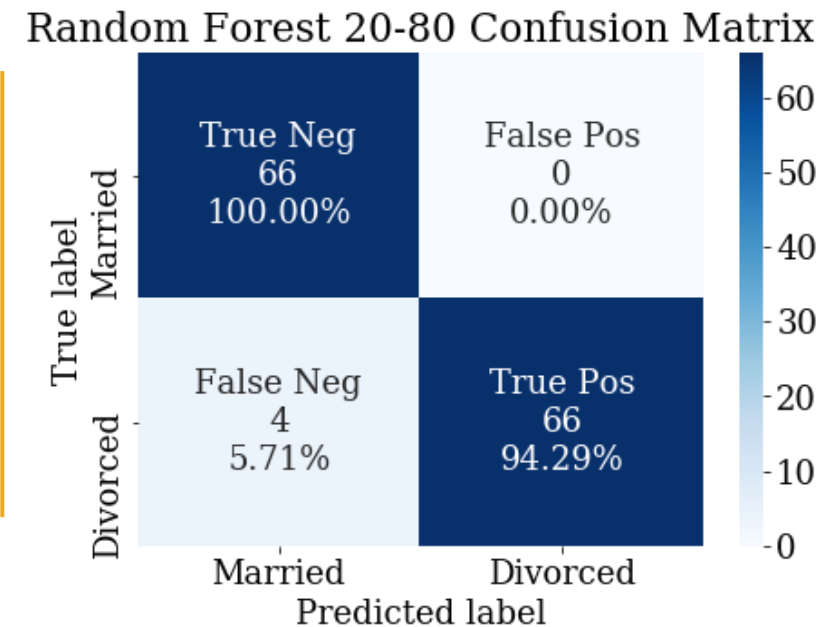
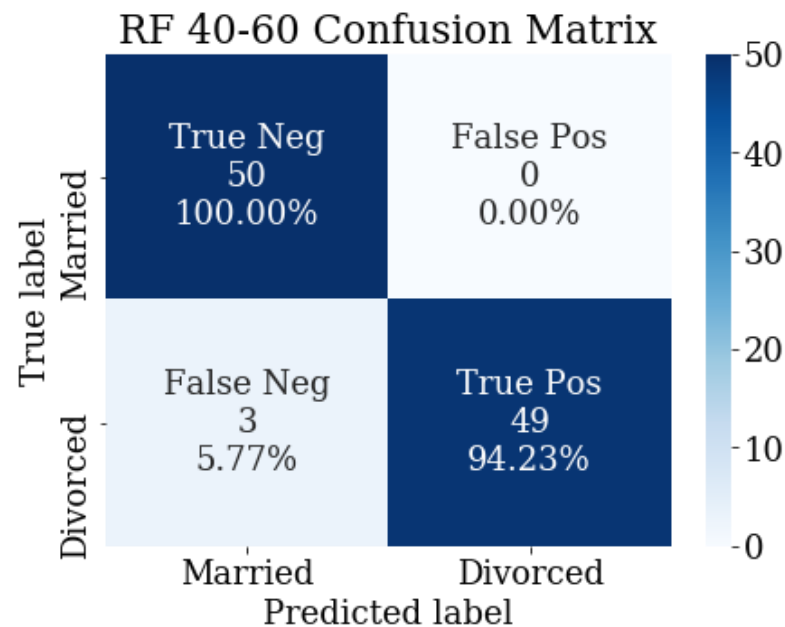
Random Forest

- עץ החלטה נבנה על ידי פיצול המידע לתיבת קבוצות, כל פעם בהסתמך על שדה אחד של המידע. תהליך זה מבוצע באופן רקורסיבי עד שכל תתי קבוצה מכילה רק איברים מאותו הסוג, או עד שפיצול נוסף לא יביא לתוצאות טובות יותר.
- זהו מסווג אשר מורכב ממספר עצי סיווג אקראיים. כל עץ סיווג נבנה תוך שימוש בתתי רשימה של מאפיינים שנבחרו באקראי מתוך כלל המאפיינים.



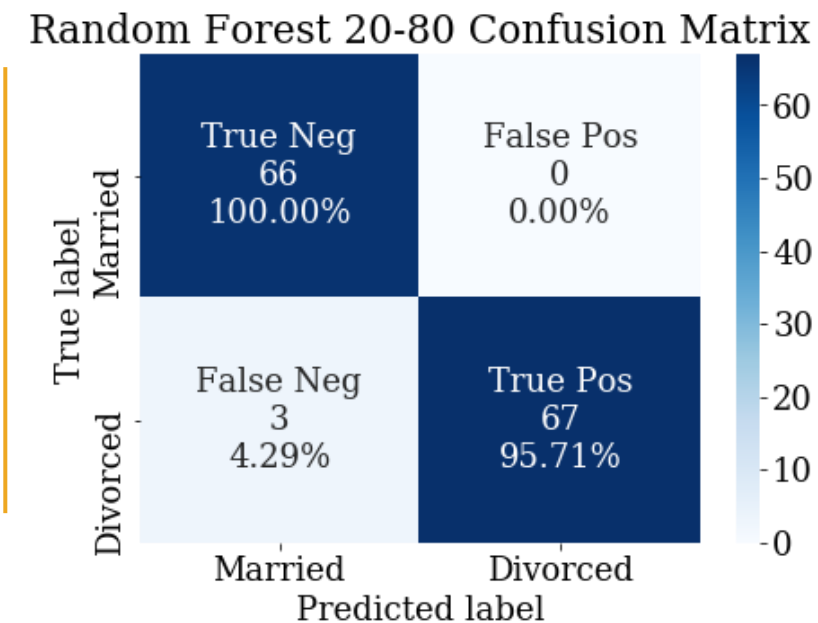
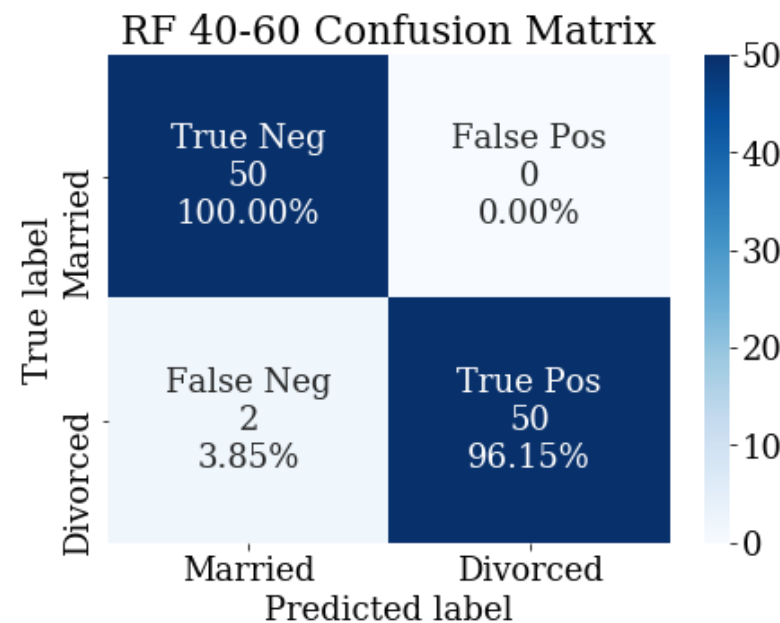
Random Forest

All 54 Features



Random Forest

6 Selected Features

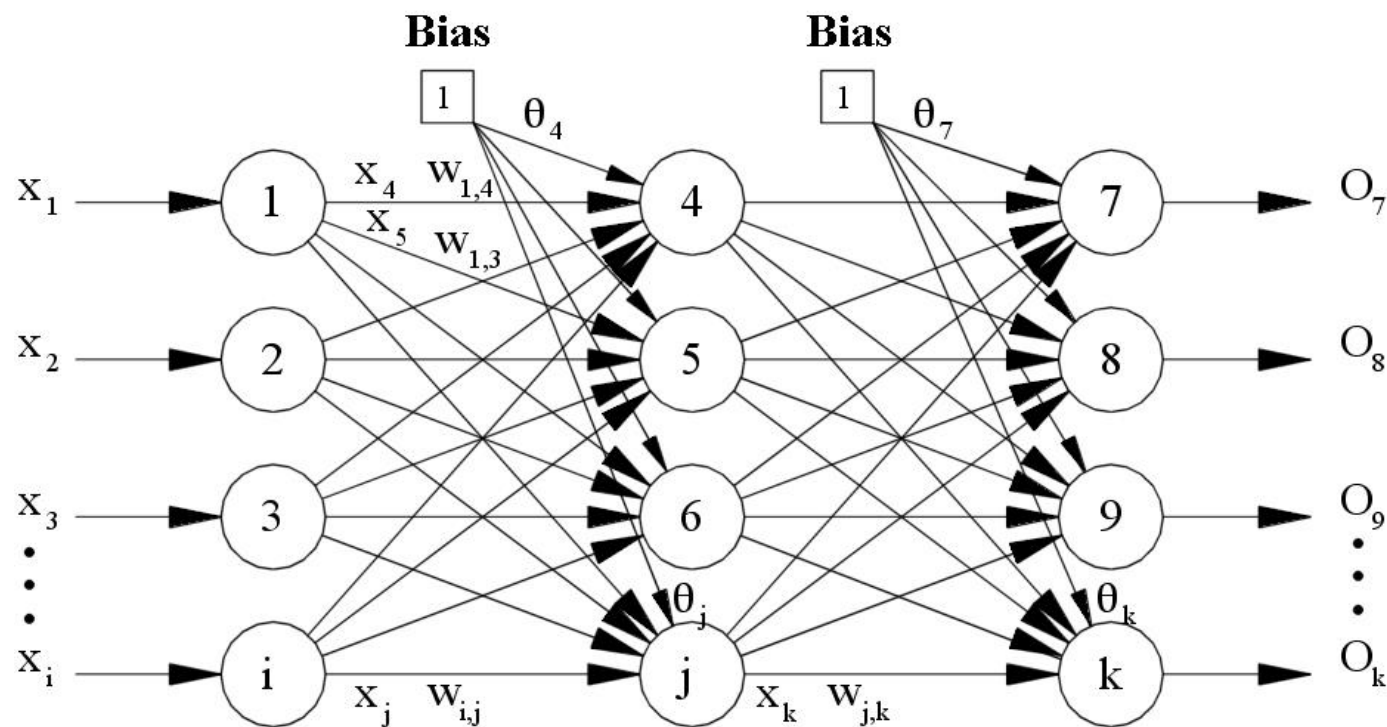


Random Forest

The Code...

Backpropagation

- אלגוריתמים זה קשור לרשת נוירונים.
- ניתן למשל לאמן רשת נוירונים לפתרון בעיית סיווג.
- המשקלים הם הפרמטרים של הרשת, בהתחלה נבחרים באקראי ולאחר מכן משתנים לפי השגיאה שחושבה בכל איטרצית אימון.



Artificial Neural Network

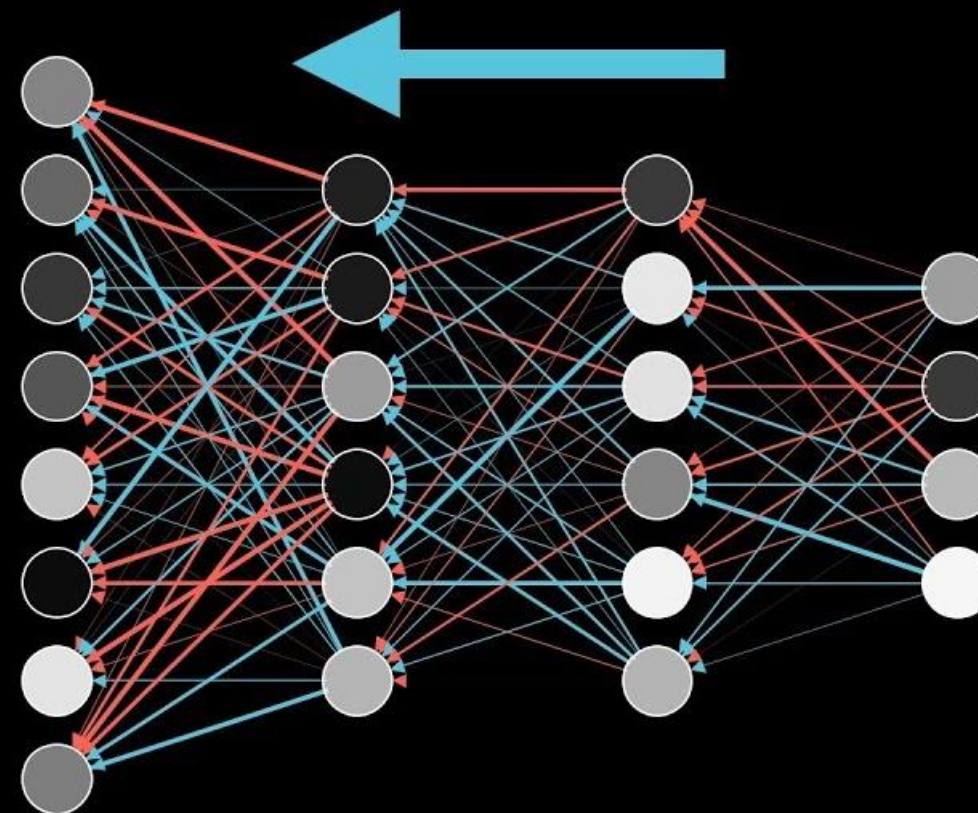
- רשת נוירונים היא כלי בסיסי בתחום הלמידה הממוחשבת. עקרון הלמידה מנסה לחקות את אופן הלימוד בתוך המוח האנושי, למרות שסודות המוח האנושי עדיין לא ידועים לגמרי.
- ניתן למשל לאמן רשת נוירונים לפתרון בעיית סיווג: בעיה שבה יש להחליט האם הקלט נופל תחת קטגוריה מסוימת או לא.



Back-Propagation

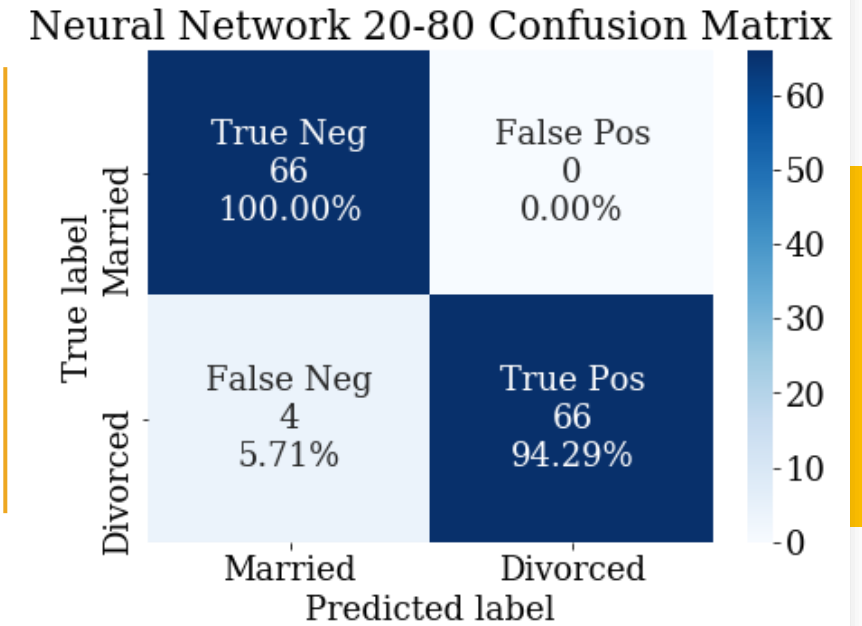
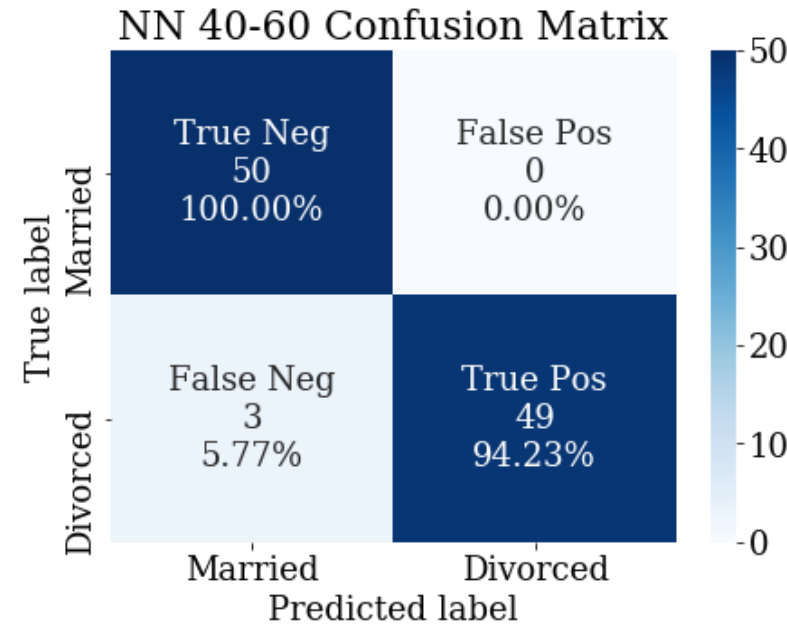
- האלגוריתם משווה את הפלט המתקבל לפלט הרצוי בכך שהוא מחשב את השגיאה.
- את השגיאה גוזרים לפי כל משקל ומוסיפים את הנגזרת לכל משקל לקראת האיטרציה הבאה.
- הדרך היעילה לחשב את הנגזרות היא לחשב אותן מהמשקלים בשכבה האחרונה וללכת אחורה עד לשכבה הראשונה.

Backpropagation



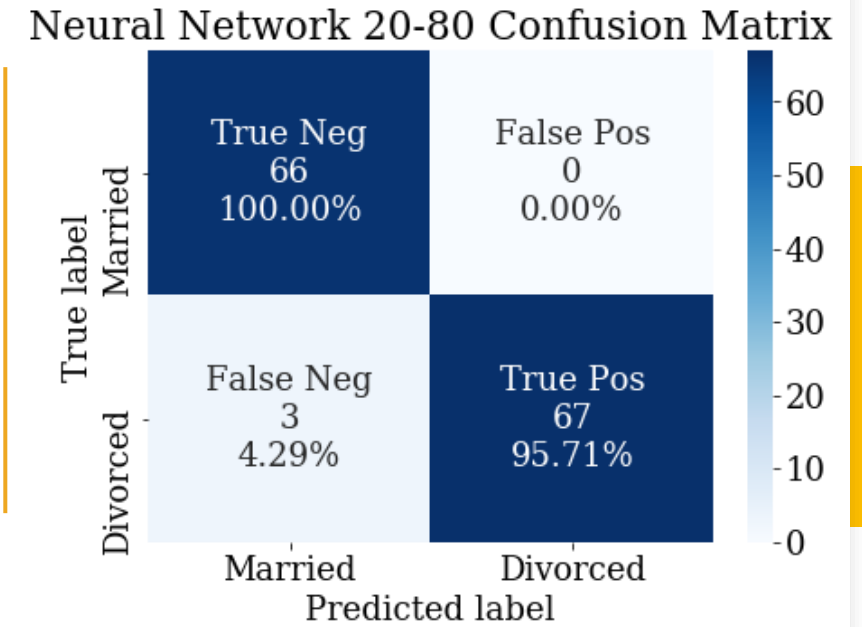
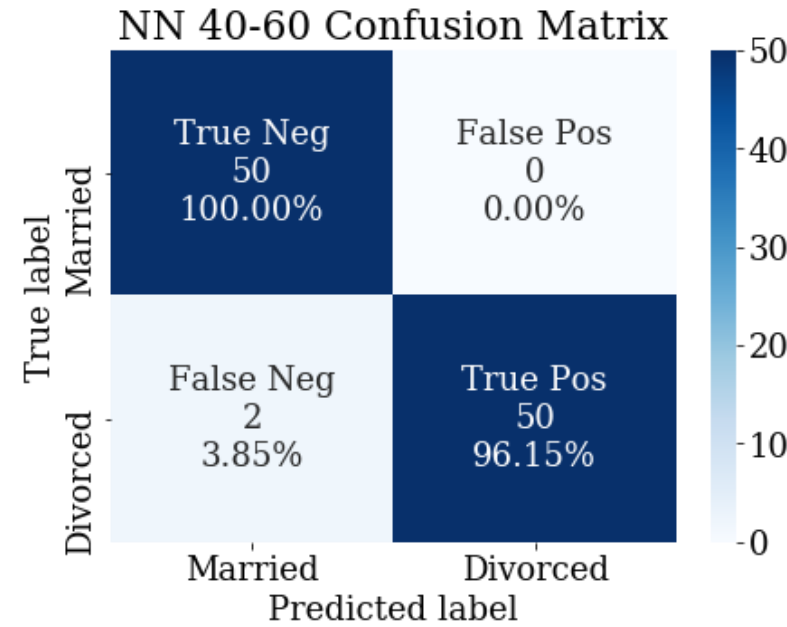
ANN

All 54 Features



ANN

6 Selected Features



Backpropagation

The Code...



Ethical Considerations

- הפרויקט שבחרנו יכול להשפיע על מערכות יחסים בחיים האמיתיים, תלוי באופן השימוש בתוצאות, במיוחד אם יועצי נישואין יכללו זאת ככלי בעבודתם. למרבה הצער, מישהו עשוי להיות מושפע לשקול גירושין שלא היו צריכים או שבני זוג אולי לא יתחתנו שחשבו על זה, אם התחזית מצביעה על כך שנישואיהם לא סבירים שיימשך.
- מצד שני, זה יכול לספק תמיכה בשירותי הכרויות המספקים את מי שמחפש קשרים ארוכי טווח על ידי זיהוי גורמים אישיים התורמים לאריכות החיים של הנישואין.

The End