

# מה משפיע על הצלחת סטארטאפים

פרויקט גמר בקורס - סדנת הכנה לפרויקט

ניתוח המידע: גבי לויס, אריאל סימן טוב, טל קליין, זהר עטר, ירדן צראף

10 ביולי 2024

# יש לכם רעיון לסטארטאפ??

על פי מחקרים שונים  
הדעה הרווחת כיום היא  
**שרק בין 1% ל- 10%**  
מהסטארטאפים בעולם  
מצליחים

מה הסיכוי  
של סטארטאפ  
להצליח כיום?

# נושא המחקר - מהן התכונות החשובות ביותר להצלחת סטארטאפים

## מטרת הניתוח:

לספק תובנות חשובות עבור יזמים ומשקיעים, על ידי זיהוי הגורמים הממלאים תפקיד מכריע בהצלחה של סטארטאפ.

יזמים יוכלו להתמקד בגורמים הספציפיים שהוערכו כחשובים.

משקיעים יוכלו להשתמש במידע זה כדי לשפר את בחירות ההשקעה שלהם, ולהעריך סטארטאפים ויזמים בצורה יעילה יותר.

## במסגרת שלב "סקירת הספרות" התמקדנו בשלושה מחקרים/עבודות קודמות, כבסיס לניתוח שלנו

03

### Summary of "Startup Success Prediction" Kaggle Project

#### גורמים שמשפיעים על הצלחה:

- ציוני דרך חשובים
- יכולת לייצר קשרים
- מספר סבבי גיוס
- כמות חברות מושקעות

02

### The Impact of Founder Personalities on Startup Success

#### גורמים שמשפיעים על הצלחה:

- תכונות אישיות כמו הרפתקנות, צניעות, אקטיביות
- גיוון צוותים
- צוות מייסדים עם תכונות משלימות

01

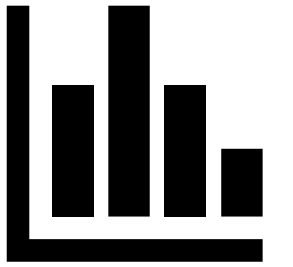
### Econometric Estimation of the Factors That Influence Startup Success

#### גורמים שמשפיעים על הצלחה:

- מיקום הסטארטאפ
- קידום שותפויות
- יכולת מסחרית
- גיל הסטארטאפ
- הכשרה טכנולוגית

**המידע כולל:** מעל ל- **920 סטארטאפים** שונים ברחבי ארה"ב.

**49 עמודות** המכילות פרטים כגון: מזהים, מידע גאוגרפי, מידע עסקי, מידע פיננסי, תחומי עיסוק, מידע על ההשקעות בסטארטאפ והאם הסטארטאפ נרכש או נסגר.



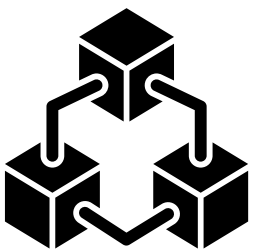
## **Data Cleaning**

- טיפולנו בערכים חסרים
- הסרנו תכונות כפולות ומיותרות
- המרנו עמודות רלוונטיות לפורמטים מתאימים



## **Feature Engineering – הוספנו את העמודות הבאות:**

- **גיל החברה:** מחושב מההפרש בין שנת ההקמה לשנת המימון האחרונה.
- **תדירות גיוס:** מספר סבבי גיוס חלקי גיל החברה.
- **יחס מימון לאבני דרך:** סך המימון שהתקבל למספר ציוני הדרך שהחברה השיגה.



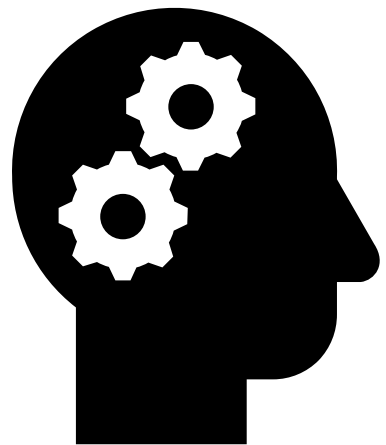
# תוכנית עבודה: איך נענה על שאלת המחקר?

## בחרנו לעבוד לפי מספר שלבים:

1. נבצע תהליך EDA במטרה להבין יותר טוב את הנתונים

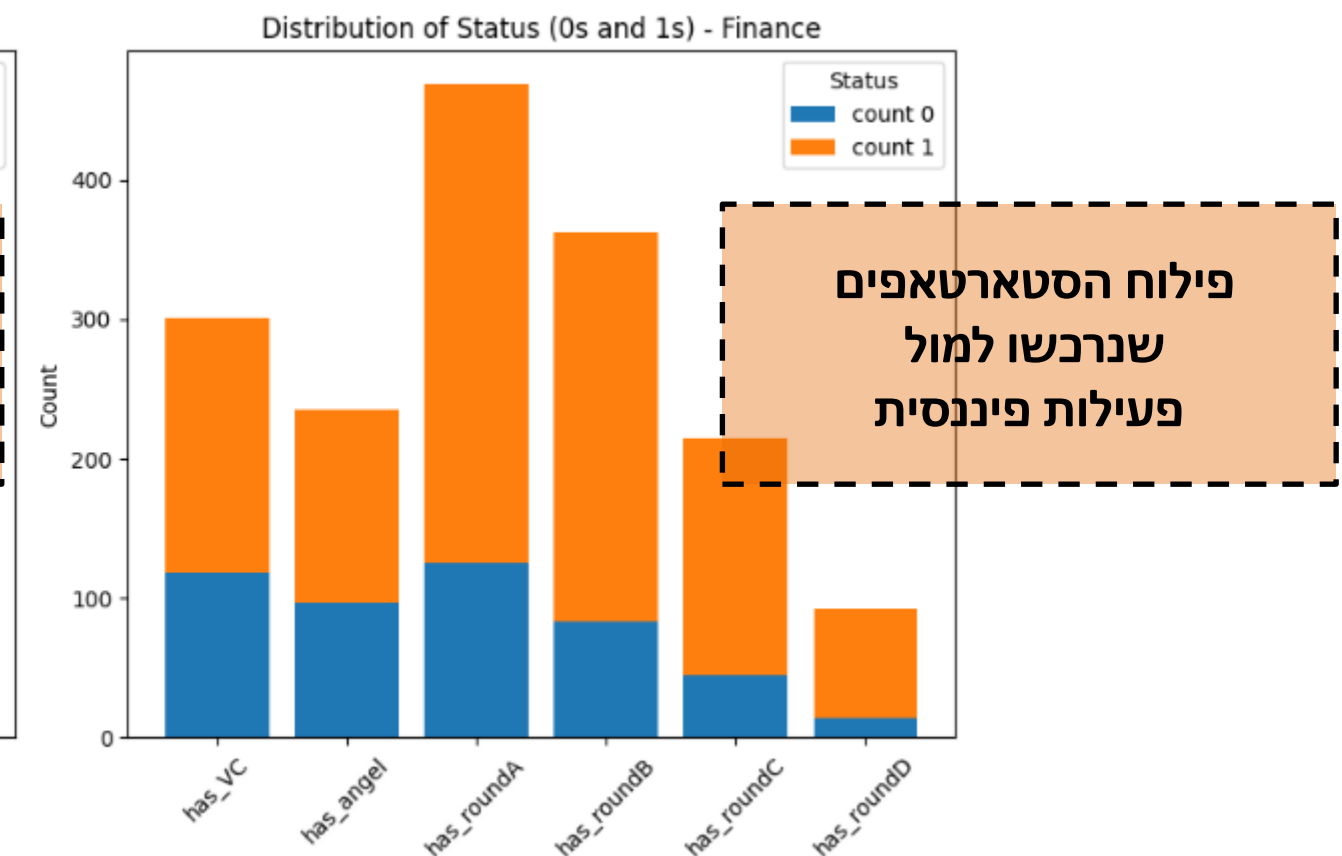
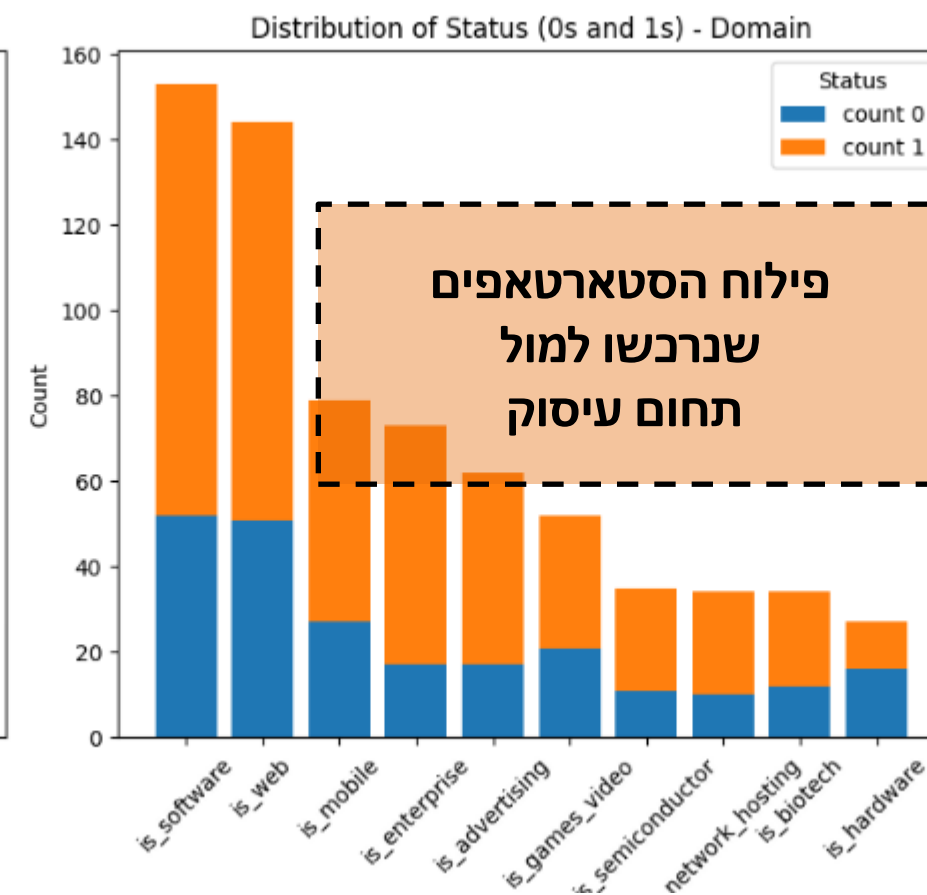
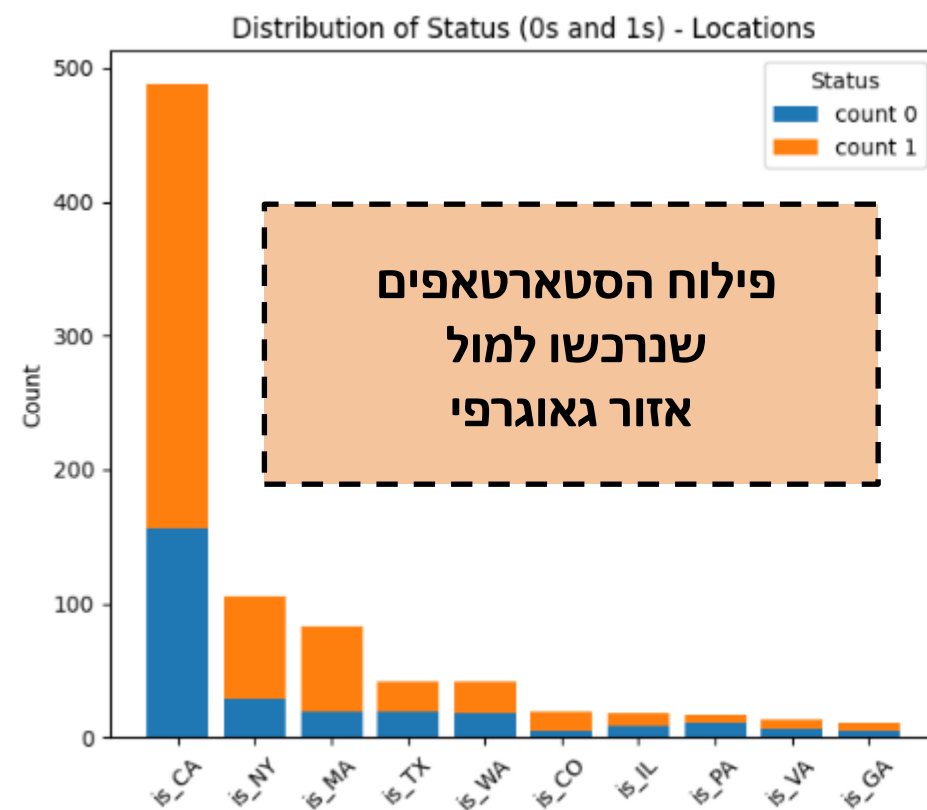
2. נרצה להבין באיזה מודל סיווג עדיף להשתמש כדי להשיג את התוצאות המדויקות ביותר.  
על כן, נערוך השוואה בין מודלים ושיטות ניתוח שונות.

3. נמצא אילו תכונות הכי משפיעות על משתנה ההחלטה לפי המודל שבחרנו.  
בהתאם למאפיינים הרלוונטיים שנמצאו, נרחיב ניתוח מפורט עבור כל אחד מהם.

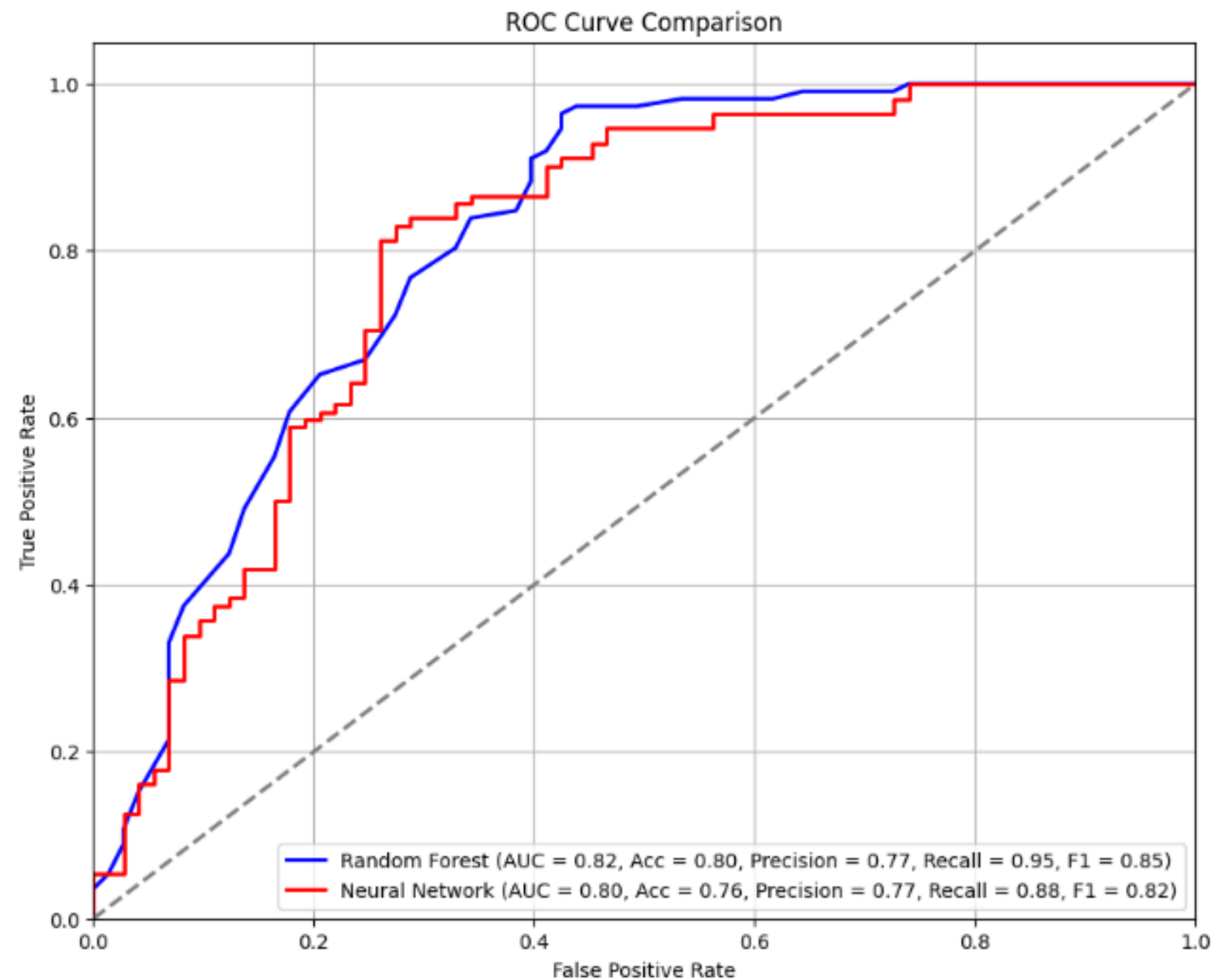


## תובנות מרכזיות מתהליך ה-EDA:

- כ- 65% מהחברות הסטארטאפ בדאטה-סט נרכשו.
- ריכוז גבוה של סטארטאפים בקליפורניה, במיוחד בסן פרנסיסקו.
- מרבית הסטארטאפים הם בתחומי התוכנה והאינטרנט.
- סטארטאפים רבים לא נרכשו, למרות שהייתה השתתפות משמעותית בהון סיכון והשקעות אנג'לים (angel investments) בהם.
- סטארטאפים עשו בדרך כלל 2 עד 4 סבבי גיוס.



# שלב 2: השוואת מודלים ובחירת המודל הטוב ביותר



הרצנו 2 מודלים על הנתונים:

**Neural Network** ו- **Random Forest**  
במטרה להשיג את התוצאות הטובות ביותר

מסקנות עיקריות:

ניתן לראות שמודל ה- **Random Forest**  
מוביל לתוצאות מדויקות יותר, לפי כלל המדדים:

Accuracy, Recall, F1 Score, Precision

ועקומת ה- ROC\_AUC,

— Random Forest (AUC = 0.82, Acc = 0.80, Precision = 0.77, Recall = 0.95, F1 = 0.85)

— Neural Network (AUC = 0.80, Acc = 0.76, Precision = 0.77, Recall = 0.88, F1 = 0.82)



# שלב 3: ניתוח חשיבות התכונות על משתנה ההחלטה לפי המודל שנבחר

## בסיכום תוצאות 2 השיטות לקביעת חשיבות תכונות, התכונות שנמצאו הכי חשובות הן:

- קשרים עסקיים עם חברות אחרות (גבוה-מגדיל)
- היחס בין המימונים לציוני הדרך בחברה (נמוך-מגדיל)
- גיל החברה בציון הדרך האחרון (נמוך-מגדיל)
- גיל החברה בגיוס הראשון (גבוה-מקטין)
- ותק חברה בשוק

## נראה כי 2 השיטות מצביעות על כך שהתכונות עם הכי פחות השפעה הן

- מיקום גאוגרפי של החברה
- תחום העיסוק של הסטארטאפ

