



המחלקה להנדסת תעשייה וניהול

שם הפרויקט: מערכת לניהול לידיים באופן
Lead Management - חכם

הגדרת הפרויקט – Statement Of Work

שם הסטודנט:	דניאל לבקוביץ', טניה פילוזוף
מספר תעודת זהות:	314030735, 205789167
שם המנחה:	אבי עסיס
תאריך ההגשה:	14.7.21

תוכן עניינים

1. תקציר.....	5
1.1. תחום הפרויקט.....	5
1.2. תחום הנתונים.....	5
1.3. אוכלוסיית היעד.....	5
1.4. רקע לפרויקט.....	5
1.5. סוג הפתרון.....	5
1.6. תוצרי הפרויקט.....	6
1.7. ייחודיות הפרויקט.....	6
2. מטרות יעדים ומדדים.....	6
2.1. מטרת הפרויקט.....	6
2.2. יעדי הפרויקט.....	6
2.3. מדדי הצלחה.....	6
2.4. יעדי תהליך כריית המידע/ הפקת התובנות.....	7
2.4.1. תהליך כריית המידע.....	7
2.4.2. מדדי איכות לתוצאות כריית המידע.....	7
2.4.3. מדדי הצלחה למערכת שתשתמש בתוצרי תהליך כריית מידע.....	7
3. סקירה ותיאור מצב קיים בארגון.....	8
3.1. תיאור וניתוח של הסביבה הארגונית.....	8
3.2. הגורמים הפנימיים והחיצוניים בעלי השפעה על התנהלות הארגון.....	9
4. סקירת ספרות / סקר שוק.....	10
4.1. עולם הבעיה- לידים.....	10
4.1.2. תהליך ניהול לידים.....	10
4.1.3. ציון לידים- Lead Scoring.....	11
4.1.4. משפך שיווקי- Sales Funnel.....	11
4.2. טכנולוגיה ויישום.....	12
5. סקירה ותיאור מצב קיים בשוק (בארץ ובעולם).....	15
5.1. תיאור מצב קיים בשוק.....	15
5.2. פתרון הפרויקט מול ההיצע בשוק.....	15
5.3. Salesforce CRM and Einstein System.....	15
5.4. מערכת Hubspot.....	16
6. פירוט הדרישות והאילוצים ראשוניים.....	18
6.1. פירוט הדרישות ותיחום המערכת.....	18
6.1.1. דרישות איסוף הנתונים.....	18
6.1.2. דרישות תהליך הפקת התובנות.....	18
6.1.3. הכלים הטכנולוגיים שימשו אותנו ודרישותיהם.....	22

23	6.1.4. דרישות תהליך בדיקת איכות של התוצאות.....
23	6.1.4.1. מדדי דיוק ללמידת המכונה:.....
24	6.2. פירוט הדרישות למערכת המידע שתשתמש בתובנות שהטופקו מתהליך כריית המידע.....
24	6.2.1. טבלת שחקנים.....
25	6.2.2. טבלת בעלי עניין.....
25	6.2.3. Use Case Diagram.....
26	6.2.4. Class Diagram.....
26	6.2.5. דרישות פונקציונליות.....
26	6.2.6. דרישות לא פונקציונליות.....
27	7. הצגת חלופות מערכתיות אפשריות למימוש הפרויקט.....
27	7.1. הצגת חלופות.....
27	7.1.1. חלופה 1- מערכת ייעודית בפיתוח עצמי מלא.....
27	7.1.2. חלופה 2- רכישת תוכנת מדף וביצוע התאמות לצרכי הלקוח.....
27	7.1.3. חלופה 3- השארת המצב הקיים.....
28	7.2. הקריטריונים להשוואה בין החלופות.....
29	7.3. השוואה בין החלופות.....
31	8. ישימות הפרויקט.....
31	1.1. ישימות כלכלית.....
31	1.2. ישימות טכנולוגית.....
31	1.3. ישימות ארגונית.....
31	1.4. האמצעים / הכלים הנדרשים.....
32	9. פערי ידע שהסטודנט צריך להשלים.....
32	10. תוצרי הפרויקט.....
32	10.1. תיאור התוצר שיוגש בסוף הפרויקט.....
33	10.2. השלבים במימוש המערכת:.....
34	11. תכנית ניהול הפרויקט.....
34	11.1. תרשים עבודה- גאנט.....
35	11.2. WBS.....
36	11.3. ניתוח ניהול סיכונים.....
38	12. רשימת מקורות.....

אישור של המנחה להגשת SOW



Avi Assis
to me ▾

12:14 (8 hours ago) ☆ ↶ ⋮

שלום דניאל וטניה,

לאחר מעבר על המסמך, מאשר הגשת המסמך למערכת.
המשך הצלחה,

אבי עסיס
ת.ז. 054107842

—
Regards,

Avi Assis

1. תקציר

1.1. תחום הפרויקט

פרויקט זה מגיע מהתחום של אנליטיקה עסקית ומערכות המלצה, הוא עוסק בפיתוח של מערכת (אתר) שתוכל לטפל ברשומות של "לידים" (מכירות אופציונליות) ולמיין אותם לפי אחוזי הצלחת המכירה שלהם, ולחלק אותם לאנשי המכירות השונים. בנוסף, תספק למנהל המכירות תמונת מצב של המכירות בארגון וניתוחים שונים על עבודת אנשי המכירות.

1.2. תחום הנתונים

הנתונים שישמשו את המערכת יהיו נתונים עסקיים אשר התקבלו על ידי המשתמש או נתונים נצברים של הארגון בקובץ טקסט, הקובץ הזה יכיל "לידים" כלומר לקוחות פוטנציאליים שהראו התעניינות לגבי המוצר של החברה, המערכת תשתמש בקובץ הזה בשביל ליצור סיווג ומיון ותחזיר אותם למשתמש בשביל שיוכל להשתמש בהם.

1.3. אוכלוסיית היעד

הפרויקט מיועד לחברות קטנות ועד ארגונים גדולים בתחום הליסינג והשכרת הרכבים, עם זאת יכול גם להתאים לביטוחי הרכב וכל ארגון שיש לו מחלקת מכירות שעובדת עם רשימות לידים בצורה ישנה או לא ממוחשבת וללא מערכת המלצה חכמה. כתוצאה מאילוצי זמן, הפרויקט שלנו מתמקד במערכת התואמת למאפיינים של תעשיית השכרת הרכבים.

1.4. רקע לפרויקט

כיום יש המון ארגונים מתעשיית הרכב והליסינג (חברות פרטיות ועד ארגונים גדולים), שאופן ביצוע מכירת השכרת הרכב מתבצע בשיטה הישנה. שיטה ישנה משמעותה שכאשר אדם או חברה מתעניינים בשירות של השכרת רכב - הם משאירים את הפרטים שלהם באתר החברה או בטלפון, לאחר מכן אנשי המכירות השונים חוזרים בשיחת טלפון לאותם מתעניינים בשיטת פיפ"ו (First In First Out) במטרה לבצע מכירה של השכרת רכב. במצב האופטימלי, אותם אנשי מכירות מספיקים לחזור אל כל הלקוחות שהשאירו פרטים אך לעיתים קרובות מכירות רבות מתפספסות ונופלות עקב חוסר תיעוד ורישום של אותם לידים. הנקודה היא שאין איזשהו מעקב חכם אחרי המכירות הפוטנציאליות הללו, מיון של הרשומות, וחלוקתם לאנשי המכירות השונים לפי מאפיינים מסוימים של הלקוח ושל איש המכירות עצמו. את כל בעיות הללו אנחנו באים לפתור באמצעות המערכת החכמה שלנו.

1.5. סוג הפתרון

אנחנו נפתור את בעיות אלו באמצעות מערכת חכמה שתוכל לקבל המון מידע על לקוחות שונים ולמיין אותם לאנשי המכירות השונים, לפי מאפיינים של לקוח ומאפיינים של איש המכירות עצמו. דבר זה יגרום להעלאת אחוזי הצלחת סגירת המכירה, והגדלת הרווח של הארגון. בנוסף נספק אנליטיקות שונות למנהל המכירות על אחוזי ההצלחה במכירות השונות ועל הספקי המכירה של אנשי המכירות השונים.

1.6. תוצרי הפרויקט

אב טיפוס למערכת (אתר / אפליקציה) שיוכל לקבל קובץ גדול המכיל מידע של לקוחות, האב טיפוס יחזיר לאיש המכירות נתונים בזמן אמת לאחר מיון והנגשת המידע. המערכת תפעיל מספר אלגוריתמים בשביל לבצע את מיון זה, ובנוסף תיצור ממשק למנהל המכירות על בסיס הקובץ שהמערכת קיבלה עם מידע אנליטי למעקב אחר מצב המכירות בארגון.

1.7. ייחודיות הפרויקט

בשוק כיום קיימים מספר מערכות CRM חכמות המסייעות במעקב חכם ומיון של לידים לפי מאפיינים של לקוח, עליהן נרחיב במסמך זה בחלופות האפשריות. אך הייחודיות בפרויקט שלנו הינה שהארגון לא צריך להטמיע את המערכת שלנו יחד עם מערכות הארגון, לא נדרש לרכוש תוכנה חדשה, המערכת לא מסובכת להפעלה, ולא נדרש מהארגון שום שינוי בדרכי הפעולה שבהם הוא פועל כיום. לכן הפתרון שאנחנו מציעים הוא תוסף יעיל, מהיר, זול ופשוט להטמעה.

2. מטרת יעדים ומדדים

2.1. מטרת הפרויקט

בניית מערכת המלצות חכמה שתאפשר לאנשי מכירות לנהל לידים פוטנציאליים בצורה יעילה תוך מתן תמונת מצב של המכירות בכל זמן נתון עבור הדרג הבכיר.

2.2. יעדי הפרויקט

- 2.2.1. מקסום רווחי החברה מלידים קיימים.
- 2.2.2. יצירת התאמה אופטימלית של "לידים" (מכירות פוטנציאליות) לפי סיווג לאנשי המכירות העובדים בחברה.
- 2.2.3. ייעול זמני עבודה של אנשי המכירות- אנשי מכירות ישאפו ליצור קשר רק עם לידים עם אחוז הצלחה גבוה.
- 2.2.4. הרחבת כושר המעקב וקבלת ההחלטות של סמנכ"ל המכירות באמצעות יצירת דוחות כספיים על מצב הארגון באופן מיידי.
- 2.2.5. יצירת "ממשק משתמש גרפי" (GUI) בעל פונקציונליות פשוטה למשתמשי המערכת ליצירת סיווג לידים באופן מהיר.
- 2.2.6. שימוש במערכת התראות עבור אנשי המכירות למניעת איבוד לידים פוטנציאליים.

2.3. מדדי הצלחה

- 2.3.1. הפיכת 30% מהלידים ל"לידים חמים" אשר יהפכו למכירה.
- 2.3.2. הגדלת רווחי הארגון ב-20%.
- 2.3.3. מזעור איבוד לידים פוטנציאליים ב 80% באמצעות שימוש במערכת התראות.
- 2.3.4. הסתגלות מהירה של משתמשי המערכת למערכת החדשה וסיווג לידים בצורה מהירה בכ-20% מהמערכת הקודמת.

2.4. יעדי תהליך כריית המידע / הפקת התובנות

2.4.1. תהליך כריית המידע

המערכת שלנו בפרויקט זה תקבל קובץ טקסט של נתונים מהלקוח המכיל פרטים על כל הלקוחות שהתעניינו לגבי רכישת או השכרת רכב.

המשתמש יעלה את הקובץ הזה לאתר המערכת וכעת תהליך כריית המידע יתחיל.

בשלב הראשון יופעל אלגוריתם ממשפחת Unsupervised Learning שבאמצעות "פיצ'רים" של לקוח יידע לחלק את הלידים לקבוצות כאשר כל קבוצה יהיו לקוחות שהכי דומים אחד לשני בתכונות של לקוחות.

לאחר מכן הלידים יחולקו לאנשי המכירות לפי מאפיינים של אנשי המכירות בארגון (עובד חדש, מנוסה, טוב, בינוני וכו').

אנשי המכירות ישתמשו בקובץ המקוטלג ינסו לבצע מכירות ללידים ויסמנו בקובץ האם התבצעה מכירה, עכשיו יש לנו את התיוג של הלקוחות שסומנו על ידי אנשי המכירות, ובשביל לשפר את דיוק המודל אנשי המכירות יעלו שוב את הקובץ המתיוג והמערכת תשתמש הפעם באלגוריתם ממשפחת ה-Supervised Learning כי יש לו את תשובה האמיתית לפי מאפיינים מסוימים של לקוח, האם הוא מעוניין בהשכרת / רכישת רכב או לא.

לבסוף עם הנתונים שאספנו, נוכל לתת למנהל הסניף כלים ומטריקות שונות על מכירות העובדים וניצור Dash Board שיסכם את מצג המכירות בארגון.

2.4.2. מדדי איכות לתוצאות כריית המידע

2.4.2.1. בשלב הראשוני של המערכת קטלוג הלידים לפי שיטות קלסיפיקציה של Unsupervised Learning לפי 3 קבוצות איכות: לידים חמים, לידים בינוניים, לידים חלשים. המדד שלנו הוא קטלוג 100% מהלידים המגיעים למערכת.

2.4.2.2. בשלב השני לאחר שהמשתמש החזיר את קובץ הלידים "המתיוג" (האם הלקוח רכש / השכיר את הרכב או לא) סיווג בעל אחוז דיוק של 70% באמצעות אלגוריתם של Supervised Learning את קובץ הלידים הבא שיתקבל על ידי הלקוח.

2.4.3. מדדי הצלחה למערכת שתשתמש בתוצרי תהליך כריית מידע

2.4.3.1. חסכון זמן עבודה של אנשי המכירות באמצעות עבודה מהירה עם קובץ הלידים ב-20%.

2.4.3.2. עליית הכנסות הארגון ב-15% כתוצאה מסיווג נכון של הלידים.

3. סקירה ותיאור מצב קיים בארגון

3.1. תיאור וניתוח של הסביבה הארגונית

מה זה ליסינג?

הַחֶכֶר או לִיסינג היא שיטת מימון לרכישת ציוד. השיטה מוכרת בציבור בזכות השימוש בה למתן רכב צמוד לעובדים, אך משמשת גם לרכישת ציוד אחר. בהקשר זה קיימים גם תחומים של החכר תפעולי, מימוני והחכר לפרטיים. [1]

הליסינג בישראל

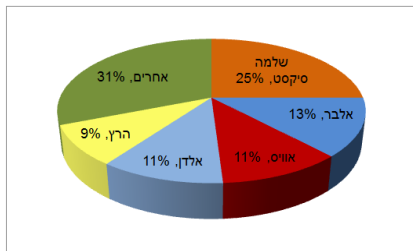
בארץ, מחזיקות חמש חברות הליסינג הגדולות במעלה מ-200 אלף מכוניות, כאשר שלמה SIXT אלדן תחבורה ואלבר הן שלוש החברות הגדולות והשולטות בתחום חברות הליסינג והשכרת הרכב.

"שלמה סיקסט" רשמה בשנת 2013 הכנסות של מעל 7 מיליארד שקל והעסיקה כ-1,300 עובדים.

בשנה זו החברה החזיקה בכ-77 אלף כלי רכב. אחריה בדירוג נמצאת אלדן, שרשמה הכנסות של כמעט 3 מיליארד שקל, העסיקה 744 עובדים והחזיקה בכ-34 אלף כלי רכב.

אלבר, המדורגת במקום השלישי, הכניסה כ-3 מיליארד שקל, העסיקה כ-1,100 עובדים והחזיקה ב-2013 כ-36 אלף כלי רכב.

חברות בולטות נוספות הן הרץ, אויס, קל אוטו, באדג'ט, ליסקאר, Enterprise-ו Lease4u.



תרשים 1: התפלגות ענף הליסינג התפעולי בישראל
במונחי היקף צי (אומדן מדרוג לספטמבר 2019)

המצב כיום מבחינה טכנולוגית

היום בחברות הליסינג בישראל, עובדי המכירות עובדים בצורה מיושנת ולא טכנולוגית. הלידים עוברים בצורה אוטומטית לאחראי המכירות דרך אפליקציית ה-WhatsApp מהפלטפורמות השונות (פייסבוק, אתר החברה, השארת פרטים בצורה טלפונית), והוא שולח כל ליד לפי התחום שלו והעיסוק העיקרי לאיש המכירות המתאים ומוסיף את הליד בצורה ידנית לקובץ Excel מאוחד. על בסיס סקר הספרות שביצענו ראינו כי דבר זה גורם לפספוס לידים רבים ותכנון לא יעיל של זמני העובדים ובנוסף הלידים לא מקבלים ציון ועדיפות ליצירת קשר.

הצורך של הארגון בסוף הפתרון

למרות גודל השוק של חברות הליסינג בישראל (סך ההכנסות השנתיות מפעילות ליסינג לשנת 2012 מוערך בכ-9 מיליארד ₪), לאחר שיחות עם אנשי מכירות בתחום הליסינג הופתענו לגלות שעדיין רוב חברות הליסינג היום עובדות בצורה "לא חכמה".

"לא חכמה" הכוונה לכך שאת הלקוחות הפוטנציאליים שהם משיגים דרך שיווק דיגיטלי

(רשתות חברתיות, פרסומות גוגל וכו') או דרך אתר החברה, הם פונים ללקוחות הללו בשיטת "פיפו" (First In First Out) ואין סינון או תיעדוף של לקוחות פוטנציאליים, כאן הפרויקט שלנו ייכנס לתמונה, ויגרום לתעשיית הרכב לעבוד בצורה חכמה יותר ורווחית יותר.

3.2. הגורמים הפנימיים והחיצוניים בעלי השפעה על התנהלות הארגון

(טבלה זו מתארת סניף של חברת ליסינג ולא את הארגון כולו)

הגורם		יחסי הגומלין עם הארגון
פנימיים	מנהל סניף	לכל חברות הליסינג ישנם כמה סניפים הפרוסים ברחבי הארץ, מנהל הסניף אחראי על ניהול העובדים, מעקב אחר המכירות, לעמוד ביעדי הסניף ועוד
	אנשי מכירות	אנשי המכירות של הסניף, תפקידם להיפגש עם לקוחות מזדמנים המגיעים לסניף, לבצע מכירות טלפוניות באמצעות לידים המגיעים דרך שיווק הדיגיטלי, להתאים את הרכב הנדרש ללקוח.
	מזכירת אדמיניסטרציה	קובעת פגישות לאנשי המכירות, מדפיסה ומתייגת חוזים שנחתמים בסניף, קובעת את ימי החזרת הרכבים ולקיחתם.
	נהג שינוע	שינוע רכבי הסניף - לקיחתם לשטיפה, תדלוקים, טיפולים, העברת טסטים, סידור ושינוע הרכבים במגרש הסניף, מתן שירות ללקוחות.
	מחלקת שיווק	אחראית על קידום מכירות ומיתוג הארגון . פרסום הארגון ברשתות חברתיות, ניהול האתר של חברת הליסינג, לארגן ולסדר את הלידים ולייצא אותם לאנשי המכירות.
חיצוניים	צרכנים פרטיים	רכישת רכב דרך חברת ליסינג, הכולל תשלום של מקדמה ותשלומים קבועים מדי חודש. בתום התקופה הרוכש יכול לבחור בין תשלום יתרת המחיר והעברת בעלות על הרכב לרוכש, מכירה של הרכב חזרה לחברת הליסינג ו/או רכישת רכב חדש במסלול דומה.
	צרכנים עסקיים	מעסיקים שרוצים לחסוך במשאבים בניהול צי הרכב ולבצע מיקור חוץ עבור צי הרכב שלהם, וגם עבור מי שרוצה לתת לעובדים הטבה משתלמת מאוד.
	ספקים	הספקים לחברות הליסינג, אחראים לספק את אלפי הרכבים לחברות, מדובר על חברות גדולות הנמצאות בחו"ל.

4. סקירת ספרות / סקר שוק

4.1. עולם הבעיה - לידים

4.1.1. ליד Lead

ליד הוא מושג בתחום ניהול המכירות אשר מייצג לקוח פוטנציאלי המתעניין בשירות או במוצר המוצע למכירה על ידי עסק, חברה או ארגון. פרטי ליד מכילים את פרטי יצירת הקשר עם הלקוח הפוטנציאלי המתקבלים על ידי אדם שפנה לאותו ארגון ישירות או אדם אשר השאיר פרטים ליצירת קשר בפלטפורמות השונות (דפי נחיתה, באגרים, רשתות חברתיות ואתרי הארגון).

פרטי המתעניין לרוב יכללו שם מלא, מס' טלפון וכתובת אימייל ליצירת קשר, וסיבת ההתעניינות.

נהוג לפלח ולסנן לידים על בסיס מאפיינים דמוגרפים וגיאוגרפים וגם לסווג את מטרת הפנייה של אותו הלקוח הפוטנציאלי כדי שיגיע לאיש המכירות המתאים והוא יצור קשר עם אותו הלקוח הפוטנציאלי וינסה להניב מהשיחה מכירה.

אנשי מכירות רואים את הלידים כמכירות פוטנציאליות היות והלקוח כבר הראה את התעניינותו בשירותי הארגון ולכן חשוב מאוד לנהל אותו וליצור קשר בחלון זמנים קצר ככל האפשר.

מטרת הפרויקט שלנו היא לעזור לאותם אנשי מכירות לא לאבד מכירות פוטנציאליות ולנהל אותם באופן חכם ויעיל על מנת להגדיל את אחוזי המכירות של אותו הארגון. [2]

4.1.2. תהליך ניהול לידים

Step 1: Lead Capture → Step 2: Lead enrichment and tracking → Step 3:

Lead qualification → Step 4: Lead distribution → Step 5: Lead nurturing

שלב 1- לכידת ליד- יצירת לידים יכולה להתרחש באופן אונליין כמו כן באופן אופליין - בצורה של מיילים, רשתות חברתיות, באגרים, פרסומות, אתרים ועוד.

הוספת הלידים למערכות ה-CRM (או בחברות פחות ממוחשבות לקובץ אקסל) אינו תהליך חכם במיוחד כאשר החשיפה היא רחבה. ניהול חכם של לידים יבטיח כי כל ליד יעבור מהמערכות השונות למערכת מרוכזת אחת- אשר ימנע פספוס מכירות פוטנציאליות. בנוסף, ניהול נכון יכתוב את מקור הליד- דבר שיעזור לדעת איזה קמפיין עושה את עבודתו בצורה מהימנה.

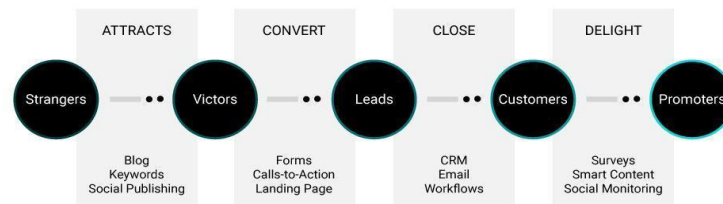
שלב 2- העשרת ליד- ברגע שליד נכנס למאגר הלידים, העשרת המידע עליו יעזור לאיש המכירות לבצע מכירה יותר אישית פרסונלית לאותו הליד. פרטים כמו - סוג משרה, חברה בה הליד עובד ופרופילים ברשתות החברתיות יעזרו לנתח את אותו הליד בצורה מקיפה יותר ויעזרו להבין את מטרת אותו לקוח עתידי ולהתאים לו את המוצר שהכי רלוונטי אליו עוד לפני שבוצע קשר עימו ולהבין מה הסיכוי שהליד יניב מכירה. מערכות יכולות להוסיף את הפעילות של הליד באתר עצמו- כולל דפים שביקר בהם, לינקים שהקליק, הורדת מוצרים למחשב האישי וכו'.

שלב 3- דירוג לידים- נהוג לדרג לידים בסולם של A,B,C,D או קטלוגם כליד "חם" או ליד "קר". על ידי קטלוג ודירוג נכון של לידים בהתבסס על העבודה, המיקום הגיאוגרפי, ההתנהגות שלהם באתר- אנשי המכירות יוכלו לדעת על איזה ליד עדיף להשקיע יותר זמן בכדי ליצור מכירה עבור אותו העסק.

שלב 4- חלוקת לידים- כאשר מדובר בעסק קטן סמנכ"ל המכירות מכיר את אנשי המכירות שלו והחלוקה של הלידים לאנשי המכירות השונים היא קלה. ברגע שהעסק גדל כך גם מאגר הלידים יגדל, השיטה הישנה של חלוקת לידים בצורה ידנית עלולה לגרום לבזבז זמן ביצירת קשר עם המכירה הפוטנציאלית. נהוג להשתמש במערכת מבוססת חוקים ומבוססת טריטוריות עבור כל

איש מכירות בכדי לארגן את הלידים בצורה יעיל, לקשר בין כל איש מכירות והשטח בו הוא מתמחה.

שלב 5- טיפוח הלידים- לצער הארגון לא כל ליד הופך להיות מכירה ממשית בין אם הליד עדיין נמצא בתהליך של חקירה אודות המוצר או עדיין בהתלבטות אם המוצר שהעסק מספק אכן מתאים לו. חשוב לשלוח מיילים רלוונטיים לאותו ליד על מנת לטפח את אותו הליד, להדגים את ערך המוצר\ השירות על ידי שליחת תוכן רלוונטי וקידום מכירות בתחומים אחרים. שלב זה עוזר בדחיפה בתהליך קבלת ההחלטות של הלקוחות הפוטנציאליים.



תרשים 2: מייצג את תהליך התפתחות הליד למכירה ממשית

4.1.3. ציון לידים - Lead Scoring

ציון לידים הוא תהליך שנעשה כדי להבין את חשיבותו ועדיפותו של ליד עבור אותו הארגון. ציון הלידים יעזור להבין מה הסיכוי של אותה המכירה הפוטנציאלית להיות מכירה ממשית. החשיבות של הליד נקבעת על פי מספר פרמטרים שנקבעו מראש, לדוגמה- מיקום הליד, תחום עבודתו (אם מדובר באדם פרטי), גיל וכולי. מודלי דירוג הלידים גם שמים דגש לדאטה שמקורה הוא מהאתר האינטרנט- מספר הקליקים שאותו הליד הקליק, מספר הפעמים שביקר באתרים השונים של הארגון או כמות הזמן ששוהה באותם האתרים. פעולה זאת בתהליך של ניהול הלידים עוזרת לארגון להבין את האופי של הלקוחות העתידיים שלו ונעזר בהם בעת ביצוע המכירה. במהלך ציון הלידים- ישנם אלגוריתמים רבים היודעים לזהות לידים שנראים מזויפים ואותו הדירוג עוזר לאנשי המכירות לא לפנות לאותם הלידים ולחסוך את זמנם. במהלך הפרויקט שלנו נשתמש בשיטת ציון לידים בכדי לדרג את המכירות הפוטנציאליות תוך שימוש אופטימלי בכל הדאטה אודות הליד. [3]

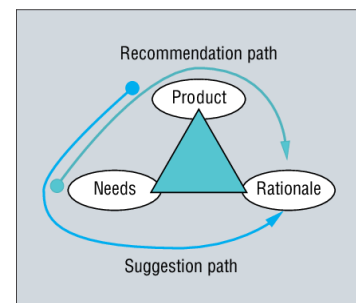
4.1.4. משפך שיווקי - Sales Funnel

תהליך שעובר אדם מסוים מהרגע שהוא פוגש בפעם הראשונה את המוצר או העסק ועד הרגע שהוא מבצע את הרכישה הראשונה בארגון. השלב הראשון הוא שלב החשיפה למוצר או לשירות של הארגון. השלב השני- שלב ההתעניינות והמחקר- לאחר שהאדם התרשם מהמוצר או מהשירות שנחשף אליו הוא יתחיל לברר עליו ויקיים השוואת מחירים עם הארגונים המתחרים, יבדוק יתרונות וחסרונות ופופולריות בשוק. השלב השלישי- שלב הבחירה- אחרי סיום החקירה והבדיקה הלקוח יצור קשר טלפוני או דרך המייל או ימלא טופס באתר או בדף הנחיתה של המוצר עם פרטי ההתקשרות שלו. השלב הרביעי- רכישה- הליד הפך להיות לקוח רשמי של הארגון כאשר החליט לרכוש את המוצר או את שירותי הארגון. השלב החמישי- שימור לקוחות- שליחת עדכונים על מבצעים או מוצרים חדשים של הארגון שעשויים לעניין את הלקוח ולעודד להמשיך ולרכוש את מוצרי או שירותי החברה.

4.2. טכנולוגיה ויישום

4.2.1. מערכת המלצה

מערכות המלצה הן כלי תוכנה וטכניקות המספקות הצעות לפריטים שישימשו את המשתמש, ההצעות המוצעות נועדו לתמוך בתהליך קבלת ההחלטות של המשתמש. פיתוח מערכות ההמלצה הוא מאמץ רב תחומי הכולל מומחים מתחומים רבים – כריית מידע, מערכות תומכות החלטה, בינה מלאכותית, סטטיסטיקה ועוד. המוטיבציה למערכות המלצה מוצגת בתרשים 1, הרעיון המכני שמקורו בצרכים (בעיות) ולאחר שמבינים את הצורך, האלגוריתם החכם של הממליץ יכול להציע פתרון בפשטות רבה. אומנם מערכות המלצה צריכות לאתר את הבעיה העיקרית של המשתמש, אך לא רק, מערכות המלצה יעילות הן אלה שיאפשרו למשתמש לחקור את מרחב האופציות שלו ובהתאם יתמכו בהעדפות השונות של המשתמש. רוב משתמשי האינטרנט כיום נתקלו אי פעם במערכת המלצה בדרך זו או אחרת, חנות ספרים וירטואלית שכאשר המשתמש יבחר ספר, תופיע לו הודעה "משתמשים שבחרו את הספר הזה, בחרו בנוסף"..
ניתן לראות, כי כדי שהמערכת תוכל לתת המלצות עליה לדעת פרטים אישיים על המשתמשים שלה.
במסגרת הפרויקט תפותח מערכת המלצה לאנשי מכירות מתחום הרכבים והליסינג על מנת לחסוך משאבים של זמן וכסף לארגון על ידיים עם סיכויי הצלחה נמוכים, ולהתאים לאנשי המכירות לידים בכדי שסיכויי המכירה שלה יעלו. [4]

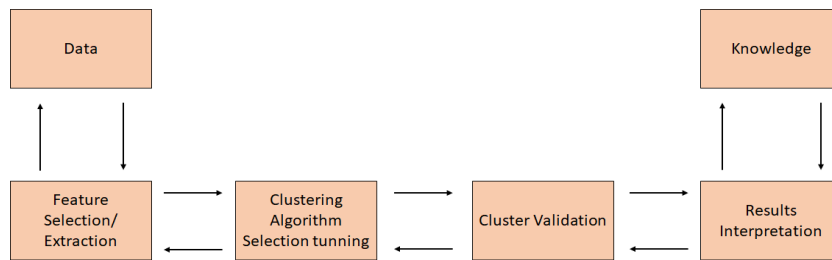


תרשים 3: הדגמה של מערכת המלצה

4.2.2. למידה בלתי מונחית- Unsupervised Learning

למידה בלתי מונחית (Unsupervised Learning) היא טכניקה בלמידה חישובית שבה מנסים ללמוד את התכונות והמבנה של אוסף דוגמאות נתונים כאשר הנתונים זמינים כפי שהם ללא תוספת תיוגים. ללא תוספת תיוגים כלומר שאנחנו לא יודעים את ה"תשובה האמיתית" ולאיזה קבוצה הדוגמאות הללו שייכות. רוב הפעמים בעולם האמיתי, כמו באתגר שהפרויקט שלנו עומד בו, המידע שלנו יגיע ללא תיוג לאיזו קבוצה הוא שייך, אז אנחנו נרצה לפתח אלגוריתם למידת מכונה שיוכל לסווג את המידע בצורה נכונה. האלגוריתם עושה זאת בכך שהוא מוצא במידע שלנו ערכים חופפים ומחלק אותם לקבוצות, בנוסף הוא מזהה ערכים "אנומליים" שלא שייכים לשום קבוצה ומאגד אותם יחדיו. באופן כללי המטרה היא למצוא במידע תכונות משותפות ולחלק אותם לאשכולות. בפרויקט שלנו אנחנו נרצה לחלק את המידע שלנו על לקוחות פוטנציאליים (לידים) ולחלק אותם לקבוצות של מי יש לו סיכוי גבוהה למכירה למשל או מי לא בעל סיכוי גבוהה למכירה אך פוטנציאל הרווח שלו הוא גבוה. [5]

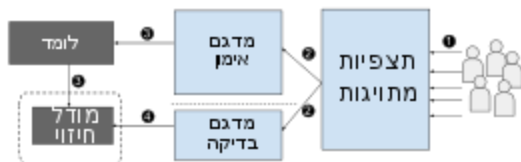
התהליך שהמידע עובר בלמידת מכונה בלתי מונחית מוצגת בתרשים 4.



תרשים 4: התהליך שהמידע עובר בלמידה בלתי מונחית

4.2.3. למידה מונחית (Supervised Learning)

למידה מונחית או למידה מפקחת (Supervised learning) היא טכניקה בענף למידת מכונה, המאפשרת לפתח מכונה או מערכת (בדרך כלל תוכנית מחשב) שלומדת לפתור בעיות על בסיס מאגר גדול של דוגמאות "פתורות". אלגוריתם למידה אופטימלי הוא אלגוריתם שהפונקציה הנלמדת על ידיו תוכל לחזות נכונה את התשובה גם בעבור דוגמאות שטרם נראו על ידי המערכת. מה שקורה מאחורי הקלעים שאלגוריתם מסוים לומד תכונות מסוימות ו"מתאמן עליהם" ולאחר מכן שניתן לו תכונה שהוא לא יודע מה התיוג שלה, לפי מה שהוא למד הוא יידע לאיזה קבוצה לשייך אותה. איך שזה יבוא לידי ביטוי בפרויקט שלנו בשלב הראשון אנחנו נשתמש בלמידה בלתי מונחית, וניתן לאנשי המכירות להשתמש בקובץ הלידים הממוין לנסות לבצע מכירות, לאחר שאנשי המכירות ישתמשו בקובץ הזה הם ייתנו לנו קובץ נוסף שמכיל את האם הם הצליחו לבצע מכירה או לא. אנחנו נלמד את המה מאפיין לקוח שהיה מעוניין במוצר, וכך נשפר את ההמלצות שלנו אף יותר. [6]



תרשים 5: דוגמה למערכת למידה מונחית – בשלב הראשון נאספות התצפיות. בשלב השני הן מחולקות למדגם אימון ומדגם מבחן, אחר כך הלומד (אלגוריתם הלמידה) משתמש בדוגמאות המתויגות במדגם האימון כדי ללמוד מודל חיזוי. באמצעות מדגם הבדיקה ניתן לבדוק את טיב המודל.

4.2.4. מערכת ניהול קשרי לקוחות - CRM

מערכת ה-CRM הינה מערכת לניהול קשרי לקוחות המאגדת את מכלול כלי הניהול הממוחשבים תחת מעטפת אחת ומציגה תמונה עדכנית ומלאה על הלקוח ופעילות הארגון בכל רגע נתון: כולל תהליכי מכירה, ניהול ושימור לקוחות, ניהול פרויקטים, הנהלת חשבונות ומלאי, ניהול משימות באופן חכם, שליחת דיוור לרשימות תפוצה, ניתוח של נתוני לקוחות באמצעות כלים של בינה עסקית וניהול תהליכי מכירה ומסעות שיווק ופרסום.

מערכת ה-CRM מדמה תמונה מלאה של הלקוח כולל מיזוג של כל פרטי הקשר לתוך ממשק אחיד ונגיש בקלות ונועדה לשפר את איכות הקשר בין עסק או ארגון לבין הלקוחות שלו.

ארגונים, חברות ומוסדות רבים עושים שימוש במערכת כדי למנף את הפוטנציאל העסקי מול הלקוחות, להגביר ולייעל את קצב העסקאות, לאגור מידע היסטורי להכרת הלקוחות, למנף את היכולות השרירותיות ולקבל תובנות עסקיות לטווח הארוך. לכל מערכת CRM יכולות להיות הגדרות ותכונות הייחודיות לה אך היתרון המשמעותי שלה המערכת היא אחסון המידע אודות הלקוחות והאפשרות לשלוף נתונים בצורה קלה ומהירה.

כאשר ארגון בוחר את מערכת ה-CRM המתאימה עבורו עליו להשוות בין המערכות השונות בשוק לפי התאמה לצרכים שלו- בין אם ניהול פרויקטים, מלאי, ביצוע מעקב זמנים או עבודת צוות או ניתוח מידע אודות העובדים שלו.

בנוסף, במערכת ה-CRM נעשה גם שימוש לטובת ניהול לידים לעסקים והמערכת מסונכרנת עם הנתונים על הלקוחות הקיימים. בפרויקט שלנו אנחנו עתידים לחקור את השימוש במערכות ה-CRM השונות, השימוש בהן לטובת ניהול לידים וההבדלים בין המערכות הקיימות בשוק לבין המערכת החכמה שאנחנו עתידים לפתח ולהבין את היתרונות והחסרונות של מערכת לניהול לידים אשר מתממשת עם מערכת ה-CRM.

4.2.5. שרת

תוכנת מחשב או כלי שמספק שירות למחשב אחר ולמשתמש בו, הלקוח. שרת הוא תוכנת מחשב או כלי המספק שירות למחשב אחר ולמשתמש בו, הלקוח.

נחלק את השרתים ל-3 סוגי שרתים עיקריים:

Web server - תוכנה המשתמשת בפרוטוקול HTTP (סט חוקים להעברת קבצים שונים דרך האינטרנט) המגישה את קבצי הHTML היוצרים עמודי אינטרנט. במקרה זה דפדפן האינטרנט מתפקד כלקוח המתקשר עם השרת.

File server - מחשב האחראי לאחסון המרכזי וניהול קבצי מידע כדי שלמחשבים אחרים ברשת המקומית תהיה גישה אל המידע.

Database server - אמצעי לאחסון מסודר של נתונים אשר ניתן לסווג במחשב לצורך אחזור ועיבוד. הגישה לנתונים והעיבוד שלהם מתבצעים על ידי תוכנות ייעודיות. הפופולריים בהם: MySQL, Oracle ו-Microsoft SQL Server. כל שרת משתמש במבנה ולוגיקת שאילתות משלו. שפת השאילתות SQL היא דומה יחסית ברוב שרתי בסיסי הנתונים.

במהלך הפרויקט שלנו אנו נשתמש בשירותים של שלושת סוגי השרתים השונים בשביל להעלות את קבצי הלידים שלנו, לעבד ולאחזר אותם ולהעלות אותם לעמודי האינטרנט בעזרת פרוטוקולי HTTP.

5. סקירה ותיאור מצב קיים בשוק (בארץ ובעולם)

5.1. תיאור מצב קיים בשוק

כיום בסוגיית ניהול הלידים בחברות הליסינג השונות בארץ ובעולם רוב החברות עובדות בצורה מיושנת. הלידים עוברים בצורה אוטומטית לאחראי המכירות דרך אפליקציית ה-WhatsApp מהפלטפורמות השונות (פייסבוק, אתר החברה, השארת פרטים בצורה טלפונית), והוא שולח כל ליד לפי התחום שלו והעיסוק העיקרי לאיש המכירות המתאים ומוסיף את הליד בצורה ידנית לקובץ Excel מאוחד. על בסיס סקר הספרות שביצענו ראינו כי דבר זה גורם לפספוס לידים רבים ותכנון לא יעיל של זמני העובדים.

5.2. פתרון הפרויקט מול ההיצע בשוק

ישנם פתרונות נוספים המציעים ניהול לידים חכם ויעיל אך דורשות הטמעה מלאה של אותה המערכת ויכולים להתאים לאנשי המכירות השונים.

אך בסוג הפתרון שלנו, היתרון הגדול הוא שאין צורך להתקין את המערכת כולה, אלא התוסף יאפשר ניהול נכון ויעיל וחלוקה לאנשי המכירות בצורה שתמנע פספוס של לידים תוך כדי שתוכל להתריע ולהזכיר לאותם אנשי מכירות לטפל בלידים הקיימים. בנוסף, המערכת תאפשר למנהל המכירות לנתח את התוצאות של אנשי המכירות השונים אשר יהווה ערך מוסף עבור המנהלים בכדי לבחון את הצלחת העובדים שלהם. המערכת אשר מספקת שירותי ניהול לידים נקראת Einstein System ובכדי להשתמש בכלים שלה, החברה נדרשת להשתמש גם בCRM תוכנה לניהול קשרי לקוחות.

5.3. Salesforce CRM and Einstein System

Salesforce היא חברת תוכנה אמריקאית העוסקת בשירותי מחשוב ענן ומספקת שירותי תוכנה כשירות (SaaS) לניהול קשרי לקוחות (CRM) וכן פלטפורמת פיתוח (PaaS) לאפליקציות עסקיות (Force.com) עבור ארגונים, דרך רשת האינטרנט. החברה מספקת ללקוחותיה מערכות שונות, הניתנות להתאמה אישית לניהול תהליכי שיווק, מכירות ושירות בארגונים. בתוך תוכנת סילספורס יש מערכת אינשטיין שהיא המערכת שמתחרה בתוצר שאנו רוצים לפתח בפרויקט זה.

בשביל להטמיע את המערכת הזו צריך להטמיע את מערכת Salesforce שבתוכה נמצא מערכת אינשטיין, כלומר היא בעצם Feature נוסף במערכת.

מערכת "איינשטיין סילספורס" היא מערכת הבינה המלאכותית הראשונה המקיפה עבור מערכת ה-CRM של Salesforce.

תרשים 6: אילוסטרציה של מערכת "איינשטיין"



זוהי מערכת משולבת בינה מלאכותית (AI) שהופכת את פלטפורמת Customer Success לחכמה יותר.

מה שמייחד את המערכת הזו:

- לא צריך להכין את הדאטה ולעשות לו איזשהו שינוי, רק צריך להכניס אותו למערכת סיילספורס והוא יעבוד.
 - יש במערכת למידת מכונה אוטומטית "מרובת משתתפים" כלומר שהמודל למידת מכונה הנכון מתאים באופן אוטומטי לארגון שלך.
- המערכת הזו יכולה לחזות מכירות עתידיות, ואפשר להתאים את שיטת השיווק לפי צרכי הלקוח.

המערכת הזו יכולה לבצע כמה דברים:

1. לתת שירות לקוחות באמצעות "בוט" (Bot) מותאם ללקוח לפי מאפיינים של הלקוח.
2. לקטלג את הלידים לפי מאפייני לקוח לאנשי המכירות השונים.
3. ליצור ויזואליזציה של הנתונים ומצב המכירות בארגון.

איך קטלוג הלידים במערכת איינשטיין עובדת?

אלגוריתם "ניקוד מוביל" של איינשטיין (Einstein Lead Scoring) מבצע ניתוח של כל התכונות שמצורף לקובץ הלידים שהלקוח העלה למערכת Salesforce ולאחר מכן מנסה אלגוריתמים שונים של למידת מכונה כמו: רגרסיה לוגיסטית, Random Forest, נאיבי ביס. והמודל עם אחוז הדיוק הכי גבוהה לאותו דאטה המערכת אוטומטית בוחרת בו. [3]

כל חודש המערכת מבצעת את הניתוח הזה מחדש בשביל לראות שהמודל למידת מכונה שנבחר עדיין הכי רלוונטי.

מחיר המערכת: 150 אירו בחודש למערכת CRM ו-50 אירו בחודש בנוסף למחיר שאתה משלם על מערכת Salesforce, סה"כ 200 אירו בחודש.

אפשרות לשליטה על המידע שנחשף - המערכת מתממשקת למערכת ה CRM של Salesforce שבמערכת ה CRM על הלקוח לשתף ולסנכרן את כל המידע אודות הלקוחות, העובדים בחברה והעסקאות בכדי להגיע לפעולה מיטבית של המערכת והחברה לא יכולה לשלוט על מה נחשף ומגיע לחברת Salesforce.

5.4 מערכת HubSpot

מתמקדת בניהול קשרי לקוחות (מערכת CRM) אשר מספקת פתרונות טכנולוגיים שיווקיים חדשניים. חברת HubSpot בנתה מוניטין כחדשנית בתחומים כמו אוטומציה שיווקית וניהול קמפיינים ברשתות חברתיות. יחד עם זאת, למערכת אין אפשרויות ניתוח ואפשרות לייצר דוחות על סמך הנתונים של החברות שמתממשקות בשירותים של אותה המערכת. בחרנו להשוות גם את מערכת זו כיוון שגם היא מספקת ניתוח לידים יחד עם CRM הכול במערכת אחודה אחת. ניתן להשתמש גם בשירותים של המערכת בחינם אך לא מקבלים את כל האפשרויות, על מנת לקבל את כל השירותים- השירות יעלה \$3200 לחודש לחברה.

היות והמערכת צריכה לעבוד עם מערכת ה CRM- גם כאן לחברה אין שליטה על המידע שנחשף לחברת HubSpot.

טבלת השוואות:

תוכנה	Salesforce	HubSpot	המערכת שלנו
ניתוח נתונים וייצור דוחות לסגל הבכיר	V	X	V
ניתוח לידים	V	V	V
מחייבת התממשקות למערכת CRM פרטית	V	V	X
אפשרות לשליטה על המידע שנחשף	X	X	V
ממשק נוח למשתמש	V	V	V
גישה למערכת מכל מכשיר אלקטרוני	V	V	V
עלות לחודש	200€ לחודש	3200\$ לחודש	קנייה חד פעמית של המוצר בשווי 200\$

6. פירוט הדרישות והאילוצים ראשוניים

6.1 פירוט הדרישות ותיחום המערכת

6.1.1 דרישות איסוף הנתונים

במערכת שלנו הנתונים יתקבלו על ידי מנהלי המכירות בצורה הבאה: לקוחות שהתעניינו לגבי השכרת ליסינג ממלאים את הפרטים שלהם בפלטפורמות השונות כמו: פייסבוק, אתר אינטרנט של החברה, יצירת קשר טלפוני ועוד. עובדת החברה תאסוף את כל הלידים הפוטנציאליים הללו ותקבץ אותם לקובץ אחד. מנהל המכירות ייקח את הקובץ המכיל לידים פוטנציאליים, יתחבר למערכת ויעלה את הקובץ.

עכשיו המערכת תתחיל להפעיל אלגוריתמים בשביל למיין ולסנן לידים ולחלק אותם לאנשי המכירות השונים, בנוסף המערכת תחשב שדות שונים מהפרמטרים שהתקבלו ותבצע טרנספורמציה לנתונים כגון: תחשב את סיכויי ההצלחה למכירה של משתמש ותוסיף אותו כעמודה נוספת בקובץ, תחשב את שווי המכירה הפוטנציאלי ועוד. המערכת תיקח את הנתונים, תמיין אותם ותעלה אותם לקובץ חדש – קובץ הלידים הממוין, ותחזיר אותם למשתמש שעכשיו יוכל להתחיל לבצע מכירות ולסמן איזה ליד הפך למכירה.

6.1.2 דרישות תהליך הפקת התובנות

במערכת שלנו אנחנו נשתמש ב-2 סוגי משפחות של אלגוריתמים, משפחת ה-Supervised Learning ו-Unsupervised Learning.

השלב הראשון: הקובץ מועלה בפעם הראשונה למערכת – ללא התשובה האמיתית

כאשר קובץ הלידים יועלה בפעם הראשונה למערכת על ידי המשתמש, כאשר אנחנו עדיין לא יודעים מה התכונות של ליד בשביל להפוך למכירה, אנחנו נשתמש באלגוריתם של Unsupervised Learning בשביל לבצע מיון ראשוני, ניתן סקירה קצרה על כמה אלגוריתם שאנחנו עלולים להשתמש ובסוף אנחנו נבחר את האלגוריתם שיבצע מיון בצורה המדויקת והמהירה ביותר.

● K - means Clustering

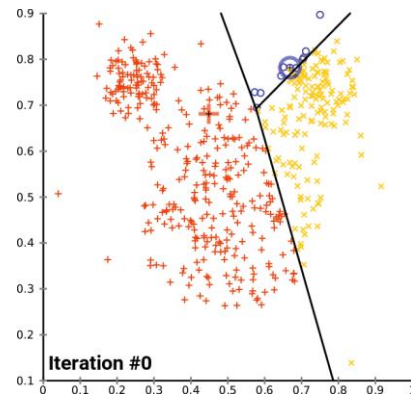
אלגוריתם k-מרכזים (k-means) הוא שיטה פופולרית עבור ניתוח אשכולות (Clustering) בכריית נתונים.

מטרתו לחלק את התצפיות ל-k אשכולות לפי מרכזי כובד (k-mean). כל תצפית משויכת לאחד מ"מרכזי הכובד" על ידי בחירה נכונה של מרכזי כובד ניתן לאתר את הקבוצות השונות.

נדרשות תצפיות רבות על מנת להשתמש במודל ותוספת של תצפיות חדשות עשויה לחייב חישוב חוזר.

מדובר באלגוריתם היריסטי שמשתמשים בו כדי לבצע חישובים שמובילים להתכנסות לפתרון שאינו בהכרח הפתרון הטוב ביותר.

זהו מודל סטטיסטי שאינו מתבסס על ידע מוקדם על הנתונים אלא רק על תצפיות בפועל.



תרשים 7: מדמה איך אלגוריתם K-Means Clustering משנה את המרכזים בכל איטרציה וכך מחלק את הנקודות לקבוצות.

● K-nearest Neighbours

על פי אלגוריתם KNN אנו יכולים לשער לאיזו קבוצה שייך נקודה מסוימת (נקודה שמייצגת ליד) על ידי כך שנבדוק מי השכנים שלה וכמה שכנים יש לאותה נקודה מכל סוג לפי עקרון זה ניתן לשער שכל שיש יותר נקודות מסוג מסוים באותו האזור שאנו בודקים, הנקודה הנבדקת שייכת לאותה קבוצה.

גודל השטח נקרא K והוא מספר המייצג את מספר הפריטים הכי קרובים לאותה נקודה שאנו בודקים. כאשר $K=1$ אנו בודקים מי הנקודה הכי קרובה לנקודה הנבדקת. כאשר $K=3$ אנו בודקים מי שלוש הנקודות הכי קרובות לנקודה הנבדקת. באופן זה ניתן לחזות האם הנקודה שאנו בודקים שייכת ללידים שיש להם סיכוי להניב מכירה או לקבוצה של לידים שאין להם סיכוי להניב מכירה. [5]

השלב השני: הקובץ מועלה בפעם השנייה למערכת – ידוע איזה ליד הפך למכירה או לא

כאשר הקובץ מועלה בפעם השנייה למערכת, זה לאחר שהוא השתמש במיון הראשוני שביצענו על הקובץ, אנשי המכירות התשמשו בקובץ הלידים כדי לבצע מכירות, והם סימנו בקובץ האם המכירה הצליחה או לא.

עכשיו יש בידינו את התשובה האמיתית, מה הופך ליד למכירה, ונוכל להשתמש בקובץ הממוין בשביל להשתמש באלגוריתמים ממשפחת ה-Supervised Learning. ניתן סקירה קצרה באלגוריתם בהם נשתמש הם:

● Logistic Regression

רגרסיה לוגיסטית היא רגרסיה בה המשתנה התלוי הוא בינארי – כלומר, יש בו רק שתי קטגוריות (כמו "כן" ו-"לא", או "גבר" ו-"אישה"), המוצגות באמצעות הערכים 0 ו-1.

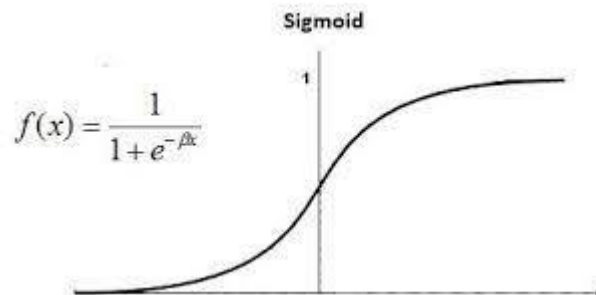
בשימוש במודל זה נקבל פונקציה של ההסתברות לערך 1 כפונקציה של המשתנים הבלתי-תלויים, כלומר: מה הסיכוי שהערך של המשתנה התלוי יהיה 1 כפונקציה של המשתנים הבלתי-תלויים.

נשתמש ברגרסיה לוגיסטית כאשר המשתנה המוסבר הוא בסולם קטגוריאלי בעל שני ערכים בלבד.

למשל, נבדוק כיצד מצב משפחתי וגובה המשכורת משפיעים על הבעלות על דירה. האם גובה משכורת או מצב משפחתי יכולים לנבא את הסיכוי שהנבדק בעל דירה?

הבעלות על דירה היא משתנה בינארי – במקרה זה, 0 מציין שהנבדק אינו בעל דירה ו-1 מציין שהנבדק בעל דירה, או במקרה שלנו 0 מציין שהלקוח לא רכש ליסינג ו-1 שהוא כן רכש ליסינג.

בסופו של דבר הפלט של הרגרסיה על כל משתנה תהיה הסתברות בין 0 ל-1 שאפשר לתרגם אותה מה הסתברות שהלקוח ירכוש את הליסינג או לא, אם המספר שיצא לנו הוא מעל 50% אנחנו נמיר אותו ל-1, ואם הוא מתחת אנחנו נחזה שהוא 0. [7]



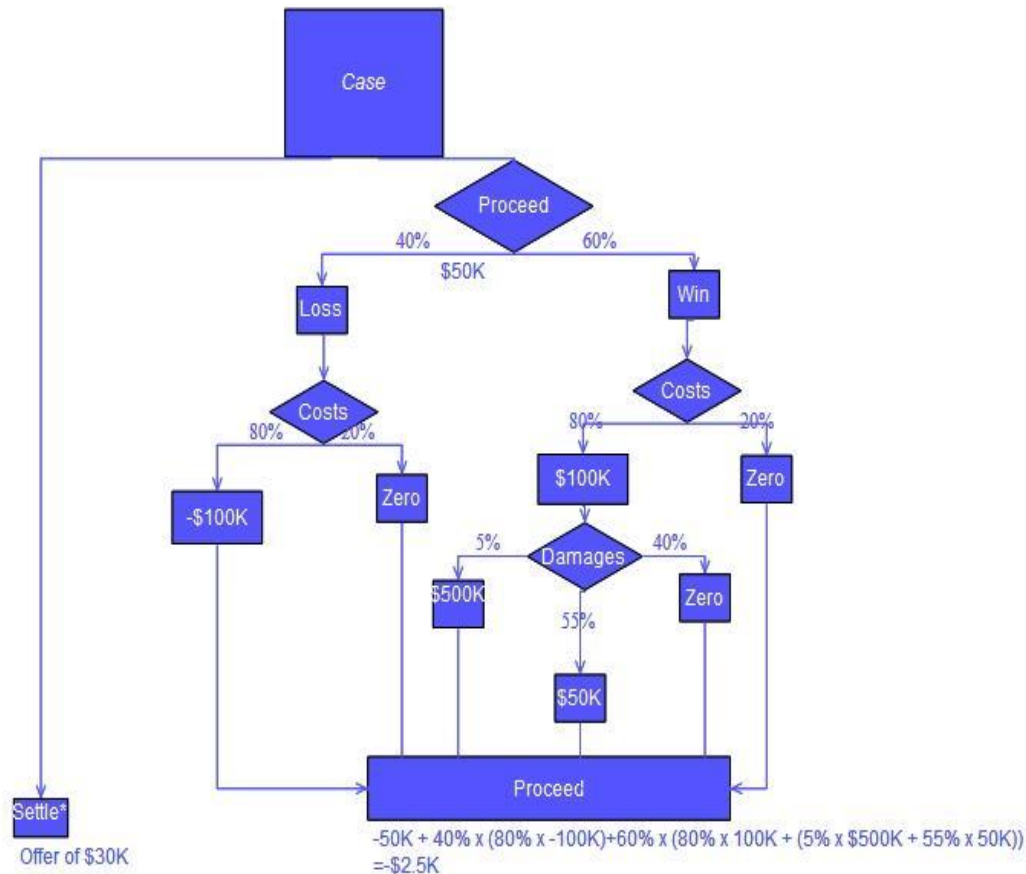
תרשים 8: דוגמה של הפונקציה הסיגמואידית שמשמשת את הרגרסיה הלוגיסטית

• Decision Tree

עץ החלטה מממש פונקציה שמקבלת דוגמה של הבעיה, ומחזיר סיווג כלשהו. כל צומת בעץ מהווה בחינה של אלמנט מסוים בדוגמה, וכל קשת מתאימה לערך שאלמנט זה יכול לקבל. עלים בעץ מתאימים לאחד הסיווגים האפשריים.

המעבר על העץ מתבצע כך: בכל צומת נבדק אלמנט כלשהו, ומתבצע מעבר על תת-העץ שאליו מובילה הקשת שמתאימה לערך של האלמנט הנבדק, הדבר נמשך עד שתת-העץ הוא עלה בלבד.

הסיווג המתאים לעלה זה הוא הסיווג שהדוגמה מקבלת לפי הפונקציה שהעץ מממש, למעשה, עץ החלטה משמש מודל חיזוי שנותן פלט עבור תצפית כלשהי בעלת מאפיינים שונים (קרי – קלט מסוים), וזאת בהסתמך על מידע אודות תצפיות קודמות וערך המטרה שלהם (קרי: סט למידה). השימוש בעצי החלטה הינו נפוץ בתחומי כריית נתונים, למידה חישובית וסטטיסטיקה.



תרשים 9: דוגמה לאיך עץ החלטה עובד מאחורי הקלעים

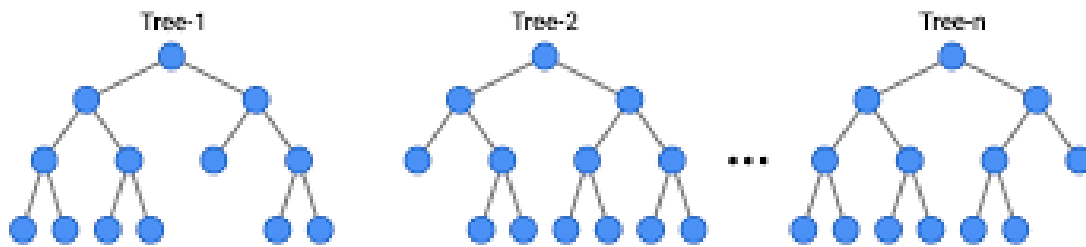
Random Forest •

יער אקראי, כפי ששמו מרמז עליו, מורכב איך לא מעצי החלטה. העצים נוצרים לרוב על ידי דגימה מתוך המאפיינים או מתוך התצפיות (כלומר, באמצעות גישת ה-bagging). כל אחד מהעצים נותן תוצאה לא-אופטימלית (Suboptimal) אך באופן כללי, על פי רוב החיזוי בדרך זו משתפר.

גישה נוספת ליצירת יער אקראי היא להגריל באופן מקרי את רמות הסף (Thresholds) המשמשות עבור המאפיינים, במקום לחפש את רמת הסף האפשרית הטובה ביותר. זה יכול להיות יעיל מבחינת חישובית שכן מציאת רמת הסף האופטימלית של המאפיין בכל צומת עלולה לקחת זמן רב.

ניתן לחשב את חשיבותו של כל אחד מהמאפיינים השונים ביער אקראי באמצעות הרווח הממוצע המשוקלל מהמידע (אשר נמדד על בסיס מדד אנטרופי או על פי מדד ג'יני) עם משקולות שפרופורציונאליות למספר התצפיות שנלקח בצומת מסוים. [9]

EXAMPLES



תרשים 10: מראה איך היער האקראי יוצר קבוצה של עצי החלטה ולבסוף על פי רוב התוצאות שהתקבלו בעצי החלטה הוא מבצע סיווג.

Naive Bayes●

סיווג באסימני נאיבי בלמידת מכונה הוא אוסף שיטות סיווג, המבוססות על חוק בייס ועל ההנחה ה"נאיבית" שאין תלות בין תכונות האובייקטים המסווגים כאשר כבר ידוע סיווגם. שיטות סיווג אלה נחקרו כבר משנות ה-1950 ולמרות ה"נאיביות" של ההנחה, הניבו מאז תוצאות יפות, בעיקר בתחום של סיווג טקסט לקטגוריות (למשל סיווג דואר אלקטרוני ל"ספאם" או "לא-ספאם"). היתרון של שיטות סיווג בייס נאיבי היא היכולת שלהם להתרחב בקלות כאשר מוסיפים להם עוד נתונים. יתרון של הסיווג הבייס הנאיבי הוא שהוא דורש רק מספר קטן של נתוני אימון כדי לאמון את הפרמטרים הנדרשים לסיווג. [10]

$$P(c|x) = \frac{P(x|c)P(c)}{P(x)}$$

Likelihood Class Prior Probability
 ↓ ↓
 Posterior Probability Predictor Prior Probability

$$P(c|X) = P(x_1|c) \times P(x_2|c) \times \dots \times P(x_n|c) \times P(c)$$

תרשים 11: הנוסחה שמשמשת את הסיווג בייסאני

6.1.3 הכלים הטכנולוגיים שימשו אותנו ודרישותיהם

Python 6.1.3.1

הכלי המרכזי בו נפתח את המערכת יהיה בשפת התכנות פייתון. פייתון הינה שפת קוד הידועה כקלה לשימוש, בזכות התחביר הנוח שלה. בנוסף, פייתון ידועה שהיא שפה שנוחה לשימוש בעולם ה-Data Science מכיוון שיש בה המון ספריות מובנות שהופכות את תהליך כריית המידע לפשוט ומהיר. ספריות עיקריות שבהם נעשה שימוש הם: scikit-learn, Pandas, NumPy, Matplotlib, Seaborn, Django.

דרישות ל-Python: תוכנת PIP, Anaconda, Version 3.6+.

6.1.3.2 Django Web development

דגננו היא ספרייה בתוך השפה פייתון, אך נכתוב עליה ככלי נפרד מכיוון שזהו הכלי העיקרי שבו נעשה שימוש בשביל לפתח את האתר של המערכת.

דג'נגו נחשבת הפלטפורמה הכי טובה לפיתוח טכנולוגיות אינטרנט והמון חברות גדולות משתמשות בחבילה הזו.

הסיבה שבחרנו בחבילה זו מכיוון שאנחנו הולכים לפתח את האלגוריתם שלנו בשפת פייתון אז יהיה נוח לקשר בין האתר לאלגוריתם באמצעות חבילה זו.

המערכת הזו מתאימה לפיתוח יישומי אינטרנט בעלי עומס רב, זאת הודות לשיטה שבה היא משתמשת שנקראת "ארכיטקטורת המסגרת", בנוסף, אפשר לעשות קישור בין דג'נגו לתוכנת בסיסי נתונים כמו SQL שנצטרך לעשות קישור זה בשביל לשמור מידע על החברה שבה תעשה שימוש במערכת שלנו. דרישות ל-, Django: Python, Deploying Django in production, HTML, CSS.

6.1.3.3 MySQL

MySQL היא מסגרת ניהול נתונים ברשת החברתית בקוד פתוח אשר נמצאת בידי Oracle Corporation שאחראית על הפיתוח מסד הנתונים וניתן להשתמש בה תחת רישיון ציבורי כללי של GNU או היתר עסקי סטנדרטי שנרכש מ-Oracle.

MySQL - הוא DBMS נוח ומבוסס ערך, יש לו יתרונות רבים כאשר המרכזי שבהם הוא תכליתיות, כלומר הוא תכליתי וניתן להעביר אותו דרך שרתים רבים.

למרות ששפת MySQL נחשבת לעיתים קרובות כבלתי ראויה לשמש במסדי נתונים של ארגונים תקציביים או שטחים מסוימים בממשל, היא הפכה למסד הנתונים החברתי הראשי באזורים רבים בעולם המלומד, כולל עבודה במחקרים ניסיוני ושימוש בה בהדרכות רבות. דרישות ל-MySQL: MySQL Work Bench.

6.1.4 דרישות תהליך בדיקת איכות של התוצאות

6.1.5 מדדי דיוק ללמידת המכונה:

● Recall - ממדין את הפריטים באוכלוסיית המדגם לאחת משתי קבוצות, בשאיפה להיות דומה עד כמה שניתן לסיווג הנכון שלהם.

● Accuracy - מדד הידוע להעריך באופן כללי עד כמה המודל הוא מדויק, כאשר מחלקים את הדאטה ל-Test ו-Training.

מדדים כמותיים (משתנים רציפים):

● MAE - הוא מדד של שגיאות (בערך מוחלט) בין תצפיות מזווגות המבטאות את אותה תופעה. במונחי כמויות.

● RMSE - מדד זה לוקח את MAE ומעלה אותו בריבוע, נותן "עונש" גדול על שגיאה גדולה יותר. במונחי כמויות.

● MAPE - שגיאת אחוז מוחלטת מדד לדיוק החיזוי של שיטת חיזוי בסטטיסטיקה. מדד במונחי אחוזים.

6.2. פירוט הדרישות למערכת המידע שתשתמש בתובנות שהופקו בתהליך כריית המידע

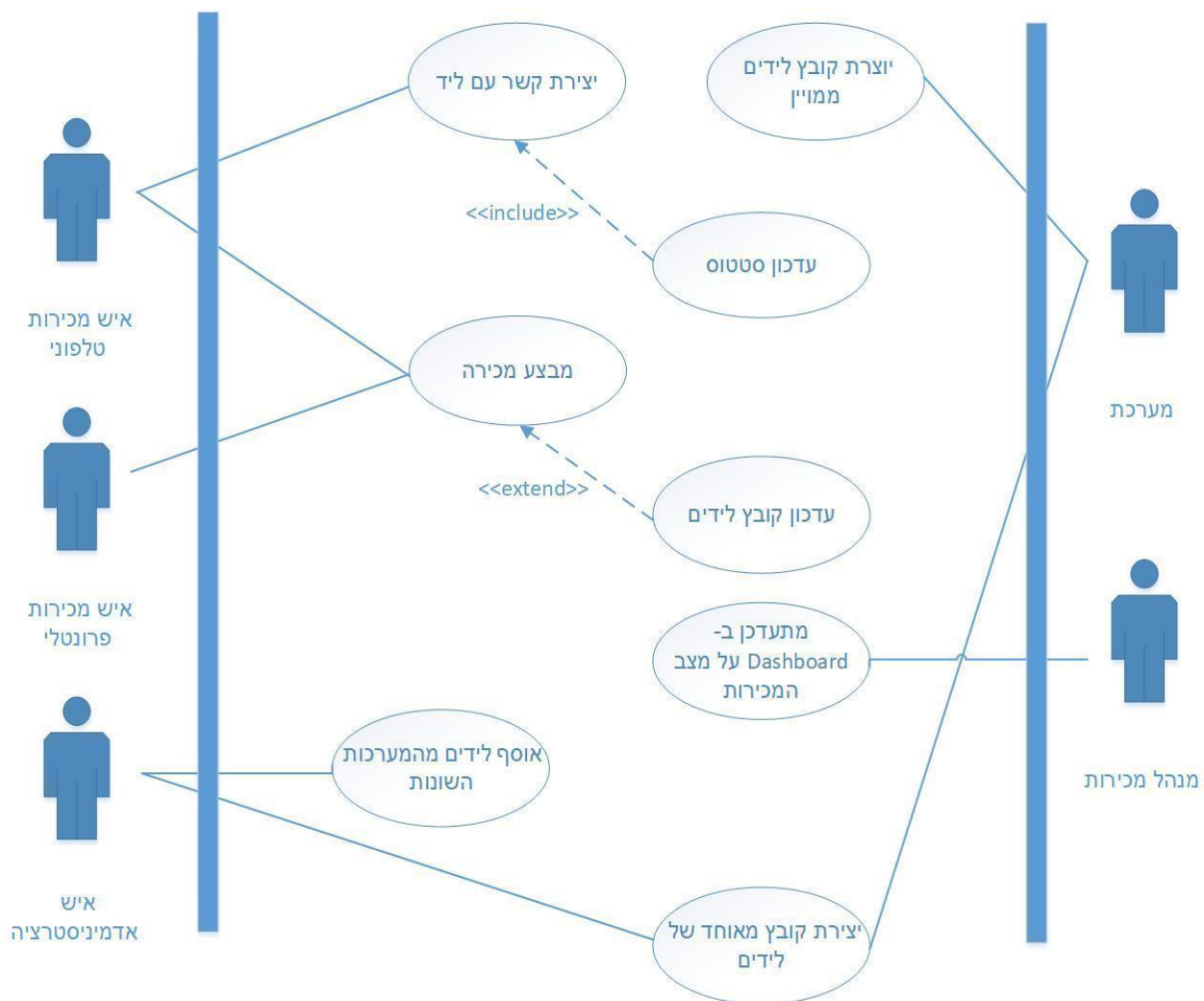
6.2.1. טבלת שחקנים

שחקן	תפקיד מול המערכת	ערך המופק מהמערכת
איש מכירות טלפוני	משתמש בקובץ הלידים המקוטלג בשביל לבצע מכירות, מעדכן את קובץ הלידים את סטטוס הלקוח (האם הצליח לבצע מכירה או לא)	הלידים מותאמים במיוחד לאותו איש מכירות כך שסיכויי ההצלחה מהמכירה עולים, חוסך זמן עבודה על ניסיון מכירה ללידים בעלי סיכויי הצלחה נמוכים.
איש מכירות פרונטלי	מקבל מאיש המכירות הטלפוני פגישות של אנשים שמעוניינים בליסינג, מעדכן את קובץ הלידים האם הצליח לבצע מכירה	מקבל מאיש המכירות הטלפוני המון פגישות של אנשים שמעוניינים בהשכרת הרכבים, חוסך זמן עבודה על ניסיון מכירה ללידים בעלי סיכויי הצלחה נמוכים.
איש אדמיניסטרציה	אסיפת לידים מהרשתות החברתיות והאתר של החברה והעלאת למערכת הלידים.	מארגן את הלידים בצורה מסודרת, מחלק את הלידים לאנשי המכירות השונים.
מנהל סניף	מקבל מטריקות שונות על מצב המכירות בארגון באמצעות Dashboard מובנה.	לבצע ניתוחים סטטיסטיים על הארגון, מעקב מצב המכירות של אנשי המכירות השונים, העלאת אחוזי המכירות של הסניף, הגדלת יעילות העובד.

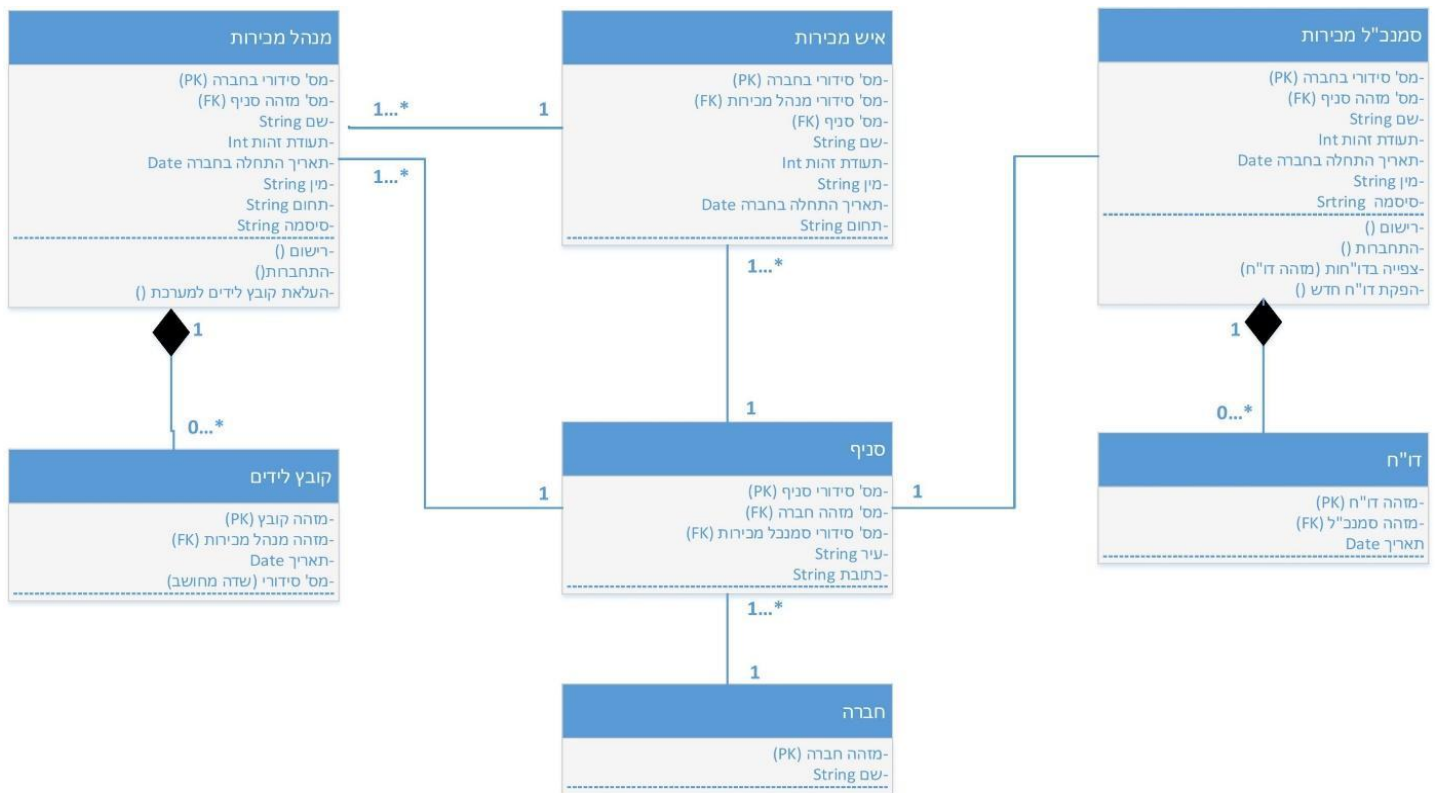
6.2.2. טבלת בעלי עניין

שם בעל העניין	תפקיד מול המערכת	ערך המופק מהמערכת
צרכן	באמצעותו המערכת מנתחת את המאפיינים השונים של אותו צרכן.	מקבל איש מכירות שמותאם לו מבחינת המאפיינים של אותו איש מכירות
מנהל הארגון	לעודד את שימוש המערכת בשאר הסניפים של הארגון	מקבל רווח כספי מהתייעלות המכירות ובנוסף התייעלות העובדים בסניף.

Use Case Diagram 6.2.3



Class Diagram 6.2.4



6.2.5 דרישות פונקציונליות

- המערכת תאפשר הרשמה ראשונית באופן ידני לסמנכ"ל ומנהל המכירות בסניף.
- המערכת תאפשר להגדיר את מבנה הארגון ומספר אנשי המכירות.
- המערכת תאפשר לקבל קבצי CSV ולהחזיר אותם למשתמש.
- המערכת תציג דוחות כספיים על פי הקבצים שהועלו למערכת לסמנכ"ל המכירות.
- המערכת תבצע התראות למשתמש על לידים שלא שומשו.
- המערכת תמין ותסנן את הקובץ שהתקבל על ידי המשתמש (סמנכ"ל המכירות).
- המערכת תבצע אלגוריתם ממשפחת Unsupervised Learning על הקובץ הראשון שהועלה למערכת.
- המערכת תבצע אלגוריתם ממשפחת Supervised Learning על הקובץ השני והלאה שהועלה למערכת.

6.2.6 דרישות לא פונקציונליות

- המערכת תהיה נגישה לשימוש רק ב-Web.
- המערכת תהיה נגישה לכל מערכת עם גישה לאינטרנט וכונן זיכרון.
- המערכת תאפשר התחברות רק למנהלי המכירות וסמנכ"ל המכירות.
- המערכת תאפשר העלאת קובץ אחד בכל שימוש (ולא העלאת כמה קבצים ביחד).
- לכל משתמש יהיה שם משתמש וסיסמה ייחודיים.
- רק לסמנכ"ל המכירות תהיה גישה לדוחות הכספיים במערכת.
- המערכת לא תהיה מערכת מידע למשתמש, אלא רק כלי למיון קבצים קיימים והצגת נתונים מהקבצים הללו.

7. הצגת חלופות מערכתיות אפשריות למימוש הפרויקט

7.1. הצגת חלופות

7.1.1. חלופה 1- מערכת ייעודית בפיתוח עצמי מלא

פיתוח מלא של מערכת המלצה ייחודית עבור חברות הליסינג וכתובת אלגוריתם המבוסס על הדרישות הפונקציונליות של הפרויקט. לטובת בניית המערכת ייעשה שימוש בשפת Python- שפת תכנות דינמית שתאפשר לנתח את הנתונים אשר יאוחסנו בשרת ייעודי. ניתוח המידע וביצוע האנליזות יבוצעו גם על ידי שפת Python בסביבות פיתוח התומכות בשפה זו.

יתרונות חלופה 1:

- חוסר תלות בגורם חוץ לצורכי תפעול ושיפור המערכת.
- מידת התאמה של המערכת לצרכים גבוהה מאוד.
- שימוש בסביבת עבודה מוכרת, חינוכית, בעלת שם עולמי ובעלת ספריות מובנות לביצוע ניתוח הנתונים.
- גמישות גבוהה לשינויים תוך כדי פיתוח ושילוב שינויים עתידיים במערכת.

חסרונות חלופה 1:

- הקמת מערכת ופיתוח שלה תדרוש זמן רב.
- תפעול שוטף ותחזוקה עצמאית של המערכת.
- ייתכנו עלויות נוספות עקב ייעוץ מאנשי תוכנה מקצועיים לטובת מתן אבטחת המערכת.

7.1.2. חלופה 2- רכישת תוכנת מדף וביצוע התאמות לצרכי הלקוח

חלופה זו כוללת פיתוח של מערכת בשילוב של מערכות קיימות. מערכת Einstein הינה תוסף למערכת ה-CRM של חברת Salesforce והן יחד מספקות פלטפורמה לניהול קשרי לקוחות עבור ארגונים. מערכת ה-Einstein היא מערכת בינה מלאכותית המשמשת כמערכת המלצה ומאפשרת קטלוג לידים. יהיה צורך בביצוע שינויים והתאמות במערכת לטובת התאמה לדרישות הפרויקט- כגון חלוקה של הלידים לאנשי המכירות לפי פיצ'רים והצגת האנליזות עבור סמנכלי המכירות.

יתרונות חלופה 2:

- מערכות מקצועיות המוכיחות את עצמן כבר שנים רבות במגוון תחומים.
- תחזוקה ותמיכה טכנית מהספק.
- הטמעה מהירה בארגון.

חסרונות חלופה 2:

- עלויות גבוהות. על הארגון יהיה לרכוש את שירותי ה-CRM עבור השימוש בתוסף של Einstein של חברת Salesforce.
- כל המידע של הארגון יצטרך לעבור למערכת ה-CRM כיוון שיש צורך בהתממשקות מלאה וסנכרון בין שתי המערכות.
- חשיפת גורם זר לנתונים של הארגון.

7.1.3. חלופה 3- השארת המצב הקיים

בחברות הליסינג בארץ עובדים באמצעים לא טכנולוגיים- הכוללים שימוש בטבלאות Excel ותקשורת בשימוש של WhatsApp. בחלופה זו לא מתבצע שינוי בתהליכי הארגון. חברות הליסינג בארץ עובדות בצורת FIFO, כאשר נכנסים פרטים אודות ליד חדש הם מועברים על ידי סמנכלי המכירות באפליקציית ה-WhatsApp לאיש המכירות הרלוונטי, והוא יוצר עמו קשר. כל פרטי הלידים מוזנים ידנית לקבצי Excel לצורך תיעוד אך לא מקבלים תיעוד מסוים או מעקב אחר הסטטוס של הלקוח הפוטנציאלי.

יתרונות חלופה 3:

- חלופה חסכונית, לא דורשת אמצעים שעולים כסף.
- אין צורך בהקמת תשתיות בשביל מערכת או שרתים מיוחדים.
- אין צורך בהכשרות עובדי הארגון עבור תפעול מערכת חדשה.
- אין סכנה של חשיפת עובדי חוץ למידע הרגיש של החברה.

חסרונות חלופה 3:

- תהליך עבודה של אנשי המכירות לא יעיל ועלול להוביל לבזבז זמן.
- עלול לגרום פספוס לידים או יצירת קשר לא מהירה מספיק עבור לקוח פוטנציאלי.
- חוסר מעקב אחרי עבודת אנשי המכירות אשר עלול לגרום ל"אבטלה סמויה".
- אי מקסום רווחי החברה עקב עבודה לא יעילה ורצף עבודה לא סדיר.

7.2 הקריטריונים להשוואה בין החלופות

7.2.1. התאמה לדרישות הפונקציונליות של המערכת

מהות: פרמטר זה הוא החשוב ביותר מבין הקריטריונים. הקריטריון בוחן את התאמת החלופות השונות לדרישות המערכת בצורה שהמערכת תקיים את כל התהליכים העסקיים והפונקציונליים הנדרשים. משקל (22%): הצורך בהתאמת המערכת לדרישות הלקוח הינו גבוה ומהותי מאוד לטובת יישום התהליכים העסקיים של הארגון.

דירוג: 10- התאמה גבוהה, 0- התאמה נמוכה.

7.2.2. גמישות לשינויים

מהות: פרמטר זה בוחן את הגמישות לשינויים במערכת. היכולת העתידית בכל חלופה לפתח ולשלב שינויים קטנים כגדולים במערכת הקצה. משקל (16%): הצורך בגמישות עבור שינויים עתידיים הינו גבוה ומהותי בפרויקט זה.

דירוג: 10 – גמישות גבוהה, 0 – גמישות נמוכה

7.2.3. הטמעה

מהות: פרמטר זה בא לבחון את הטמעת המערכת והשימוש בה. בפרמטר זה משוקללת גם כן התאמת המערכת לשיטות העבודה הקיימות, הצורך בהכשרה והדרכות, קלות התפעול ונוחות הטמעה של המערכת בארגון. משקל (15%): חווית המשתמש עבור מערכת זו הינה חשובה מאוד. רק מערכת אשר תאפשר תפעול נוח ופשוט לכוח אדם בעל הכשרה מינימלית, תוכל להשתלב בארגון ולמלא את ייעודה.

דירוג: 10 – הטמעה נוחה, 0 – הטמעה לא נוחה

7.2.4. מורכבות הפיתוח

מהות: פרמטר זה בוחן את מורכבות הפיתוח בכל אחת מהחלופות השונות. בפרמטר זה משוקללים הסיכונים השונים בפיתוח, אשר עתידים להשפיע על משך ועלויות הפיתוח. משקל (9%): קיימת חשיבות מסוימת למורכבות הפיתוח, בעיקר בשל ההשפעה הפוטנציאלית על משך ועלויות הפיתוח.

דירוג: 10 – מורכבות נמוכה, 0 – מורכבות גבוהה

7.2.5. משך הפיתוח וזמן מסירה ללקוח

מהות: פרמטר זה בוחן את משך הזמן הצפוי לפיתוח המערכת בכל אחת מהחלופות השונות. משקל (8%): קיימת חשיבות נמוכה למשך הפיתוח כיוון וכיום הארגונים פועלים ללא המערכת.

דירוג: 10 – משך קצר, 0 – משך ארוך

7.2.6. עלויות

מהות: פרמטר זה בוחן את העלויות הצפויות לפיתוח המערכת בכל אחת מהחלופות השונות. עלויות הפיתוח כוללת את הפרויקט על כלל שלביו. משקל (17%): עלויות הפיתוח הינן חשובות מאוד. פרויקט בעלות גבוהה מדי העולה על התועלת המופקת לא יוכל להתבצע ותוכניתו לא תצא לפועל. בנוסף, המערכת באה לייעל תהליכים עסקיים המתרחשים בארגון, אם העלויות יהיו גבוהות- התועלת תהיה נמוכה.

דירוג: 10 – עלות נמוכה, 0 – עלות גבוהה

7.2.7. תחזוקה

מהות: פרמטר זה בוחן את אחזקה הטכנית העתידית בכל אחת מהחלופות השונות. משקל (13%): קיימת חשיבות לבחינת מול האחזקה העתידית של המערכת. המערכת עתידה לפעול לאורך שנים ארוכות ויש להבטיח כי ניתן יהיה ניתן לתחזק אותה בצורה טובה ולהבטיח את פעולתה.

דירוג: 10 – תחזוקה פשוטה, 0 – תחזוקה מורכבת

7.3. השוואה בין החלופות

משקל ל (%)	חלופה 1- פיתוח עצמי		חלופה 2- תוכנת מדף		חלופה 3- מצב קיים		
	ציון	ציון משוקלל	ציון	ציון משוקלל	ציון	ציון משוקלל	
22	9	1.98	7	1.54	2	0.44	התאמה לדרישות הפונקציונליות
	נוכל לבצע התאמה מרבית	נוכל לבצע התאמה חלקית	נוכל לבצע התאמה מינימלית				
16	9	1.44	5	0.8	0	0	גמישות לשינויים
	גמישות מלאה. המערכת תתוכנן ותבנה תוך הקמת בסיס לצרכים נוספים של הארגון	גמישות חלקית. לא ניתן לבצע שינויים משמעותיים בתוכנת המדף.	אין אפשרות לגמישות.				
15	8	1.2	4	0.85	9	1.35	הטמעה
	המערכת תהיה קלה להטמעה וחווית המשתמש תהיה מיטבית.	יהיה צורך לבצע הרבה שינויים לטובת הטמעת המערכת הקיימת וסנכרון מלא. הכשרות עובדי	אין צורך בהטמעה של מערכת חדשה והתפעול נוח עבור העובדים.				

		החברה ייקחו הרבה זמן.					
0.9	10	0.54	6	0.36	4	9	מורכבות הפיתוח
אין סיכון בחלופה זאת או מורכבות		הפיתוח בחלופה זו הינו במורכבות בינונית		מורכבות הפיתוח יחסית גבוהה ובעלת סיכונים רבים			
0.8	10	0.4	5	0.32	4	8	משך הפיתוח
אין זמן פיתוח		זמן פיתוח בינוני		זמן פיתוח רב			
1.7	10	0.86	4	1.19	7	17	עלויות
אין עלות		עלות פיתוח גבוהה		עלות פיתוח בינונית			
1.3	10	0.52	4	0.91	7	13	תחזוקה
אין צורך בתחזוקה		נוכל להשתמש בצוות התמיכה של התוכנת מדף אך נצטרך עדיין להעסיק צוות טכני שמבין בקוד ובשינויים שנעשו		הפרויקט יקצה צוות טכני מלא שיוכל לתת תמיכה לארגון			
6.49		5.51		7.4			ציון

החלופה הנבחרת, היא חלופה 1 מערכת ייעודית בפיתוח עצמי מלא, אשר קיבלה את הציון 7.4 הכי גבוה מבין כל החלופות בסה"כ בכל הקריטריונים שהגדרנו כחשובים. כמו כן היא קיבלה גם את הציון הגבוה ביותר מבין החלופות בקריטריון "התאמה לדרישות הפונקציונליות" שכמו שנאמר זהו הקריטריון החשוב ביותר אשר עשוי להביא להצלחת הפרויקט.

8. ישימות הפרויקט

8.1. ישימות כלכלית

הפרויקט שאנחנו מתכננים הוא בעל ישימות כלכלית גבוהה, וזה מכיוון שפיתוח האלגוריתם של מיון הלידים נעשה בשפת התכנות פייתון שזוהי שפה חנמית, החבילות שנמצאות בשפה עצמה הם בחינם גם כן, בנוסף גם מסד הנתונים והעברת המידע מהלקוח אלינו נעשה הכל באינטרנט לכן שאין עלויות על תהליכים אלה. העלויות היחידות שישנם הם תשלום לספק שרת אינטרנט חיצוני שבו אנחנו נעלה את אפליקציית האינטרנט שלנו וגם התשלום על שימוש בשרת הוא תשלום זניח לפי שימוש. לכן הפרויקט שלנו הוא זול ובר השגה עם תוכנות וכלים שונים הנמצאים באינטרנט לשימוש הרחב.

8.2. ישימות טכנולוגית

כמו שצינו בסעיף הקודם, מכיוון שהתוכנות, שפות התכנות ושרת אינטרנט נמצאים בשימוש הרחב באינטרנט יש ישימות טכנולוגית גבוהה. רוב האלגוריתמים שאנחנו נעשה בהם שימוש הם אלגוריתמים שפיתחו אותם בספריות שונות, אנחנו נצטרך להתאים את האלגוריתמים לשימושים שלנו שזה לחזות ליד שיהפוך למכירה, אך ישנם המון תיעודיים ועזרה מהקהילה על אלגוריתמים שבהם נעשה שימוש. ישנה אי וודאות לגבי האם האלגוריתם יצליח לבצע חיזוי מוצלח לגבי הלידים והאם הממשק ללקוח יהיה נגיש באינטרנט ויענה על הצרכים, אז על אי וודאות זו נוכל לענות רק איך שנתחיל לפתח את האלגוריתם עצמו.

8.3. ישימות ארגונית

היתרון הגדול ואולי העיקרי של הפרויקט שלנו, שהארגון לא נדרש לשום שינוי בתהליכים היומיוניים בהם הוא עושה כיום, מכיוון שהמערכת שלנו תשמש ככלי המלצה ללידים ולא כמערכת CRM או מערכת מידע, כלומר הארגון יוכל להמשיך להשתמש במערכות בהם הוא עושה כיום, ולא נדרש לשום שינוי מיוחד, כל מה שהוא יצטרך לעשות במידה וירצה בכך, הוא לעלות את קובץ הלידים שקיים כבר בארגון שאותו הם מחלקים לאנשי המכירות השונים לאתר האינטרנט שלנו, ואנחנו תוך דקות ספורות נספק לו קובץ ממין חדש.

8.4. האמצעים / הכלים הנדרשים

הכלים הנדרשים על מנת לבנות את המערכת הזו הוא שימוש בשרת אינטרנט חיצוני שאליו נעלה את אפליקציית הרשת שלנו, רישיון בתוכנת MySQL בשביל שיהיה לנו מערכת מידע נגישה שבה נוכל לשמור מידע על הלקוחות שלנו (חברות הליסינג).

כל שאר האמצעים הטכנולוגיים הם אמצעים שנגישים לחינם באינטרנט וישנן קהילות רבות התומכות בתוכנות אלה ומסייעות בשימוש נכון וחכם בהן.

9. פערי ידע שהסטודנט צריך להשלים

מועד	אופן השלמת הפער	תיאור הפער	
לפני מסירת דוח הביניים.	למידה מעמיקה ממקורות באינטרנט.	השלמת ידע בשפת תוכנה ליצירת האלגוריתם והמערכת- נדרשת השלמה וחזקת הידע הקיים בשפת Python.	1.
שנה רביעית סמסטר א'	לקיחת קורס "מבוא למערכות המלצה".	העמקת ידע בעולם מערכות ההמלצה.	2.
לפני מסירת דוח הביניים.	קורס ייעודי באתר הלמידה העצמי Udemey.	למידה עצמית על חבילת Django	3.
שנה שלישית סמסטר ב'	במהלך קורס "טכנולוגיות אינטרנט".	למידה על שפות תכנות לפיתוח אתרי אינטרנט (HTML, CSS,) (JAVASCRIPT).	4.
לפני מסירת דוח הביניים.	למידה מעמיקה ממקורות באינטרנט.	למידת אלגוריתמים בתחום הקלסיפיקציה, Unsupervised Learning ו-Supervised Learning.	5.
לפני מסירת דוח סופי.	למידה מעמיקה ממקורות באינטרנט ונלמד חלקית בקורס "טכנולוגיות אינטרנט".	למידה ושימוש על שרת אינטרנט חיצוני שאליו נעלה את אפליקציית הרשת.	6.
לפני מסירת דוח סופי.	למידה מעמיקה ממקורות באינטרנט.	שימוש והשלמת ידע בשפת SQL והתממשקות MYSQL עם שפת התכנות Python.	7.
לפני מסירת דוח סופי.	ייעוץ מעובדים מתחום חברות ההשכרה והליסינג.	הרחבת ידע על שוק השכרות הרכב ולמידה מה מאפיין לקוחות שמעוניינים ברכישת ליסינג מחברות אלו	8.

10. תוצרי הפרויקט

10.1 תיאור התוצר שיוגש בסוף הפרויקט

התוצר שיוגש בסוף הפרויקט תהיה אפליקציית אינטרנט שתשמש מנהלי מכירות מעולם הליסינג (רכבים), המערכת הזו תשב על שרת אינטרנט בה אנשים יוכלו להתחבר מכל מערכת זמינה של הלקוח (מחשב נייד, לפטופ, טלפון סלולרי) עם גישה לאונליין. מנהלי המכירות יעלו את קובץ הלידים היומי שלהם למערכת שלנו, ואנחנו נבצע מיון וחלוקת הלידים לאנשי המכירות השונים לפי המאפיינים של החברה (מספר אנשי מכירות וכו') בנוסף, לאחר קבלת מדדים על ביצוע המכירות אנחנו נבנה Dashboard עם תובנות עסקיות למנהלי המכירות שיראו את מצב המכירות בחברה. לכן לסיכום, המערכת שלנו תהיה מערכת המלצה, שבאמצעות אלגוריתמים של למידת מכונה תוכל לאפיין מה הופך ליד למכירה ותדע איך לחלק את הלידים השונים לפי מאפיינים של איש מכירות.

10.2. השלבים במימוש המערכת:

שלב א': השלב הראשון בבניית המערכת הוא לאפיין את הדאטה שאנחנו נקבל מחברת הליסינג, כלומר אנחנו נקבל מידע שמכיל בתוכו סוגים שונים של לידים כלומר סוגים שונים של לקוחות ואנחנו נצטרך להחליט איזה מידע תורם לנו בשביל להחליט מה יהפוך ליד למכירה, מה הם הסוגים השונים של ה-Data Types של העמודות בקובץ, ובכללי לעשות Data Pre-processing לקובץ שלנו בשביל שנוכל לממשק אותו עם האלגוריתם.

שלב ב': לאחר אפיון המידע שנקבל מהלקוח, נתחיל לעבוד על פיתוח האלגוריתם בשפת התכנות Python, אנחנו נשתמש בכל מיני אלגוריתמים למידה מכונה שידועים בשוק כמו: K-Means Clustering, KNN, Logistic Regression, Decision Tree, שהם נגישים בספריות החינמיות בשפת התכנות, בשביל לסווג את הלידים וליצור מודל שיקבל פרמטרים על לקוח ויוכל לחזות לנו מה יהפוך ליד למכירה.

שלב ג': לאחר שבנינו את האלגוריתם בשפת Python, נתחיל לעבוד על פיתוח אפליקציית האינטרנט שלנו באמצעות חבילת Django פייתון, נבנה אתר עובד שבו הלקוח יוכל לעלות את הקובץ הלידים שלו ולקבל קובץ לידים ממיון בחזרה, אנחנו נשתמש גם בעיצוב אתרים בשפות פיתוח אינטרנט כמו CSS, JAVA SCRIPT, HTML בשביל להפוך את האתר ליותר נוח ועם עיצוב יפה למשתמש.

שלב ד': בשלב זה אנחנו נעבוד במקביל גם על בסיס הנתונים שיתממשק עם האתר בשביל לשמור מידע על הלקוח ואנחנו גם נתחיל לעבוד על השלב האחרון במערכת שהוא ליצור Dashboard למנהלי המכירות בשביל לקבל תובנות עסקיות על מצב המכירות ועל אנשי המכירות, גם את לוח צג זה אנחנו נבנה באמצעות שפת התכנות פייתון שיש בה ספריות עם גישה נוחה ליצור גרפים וויזואליזציה של דאטה.

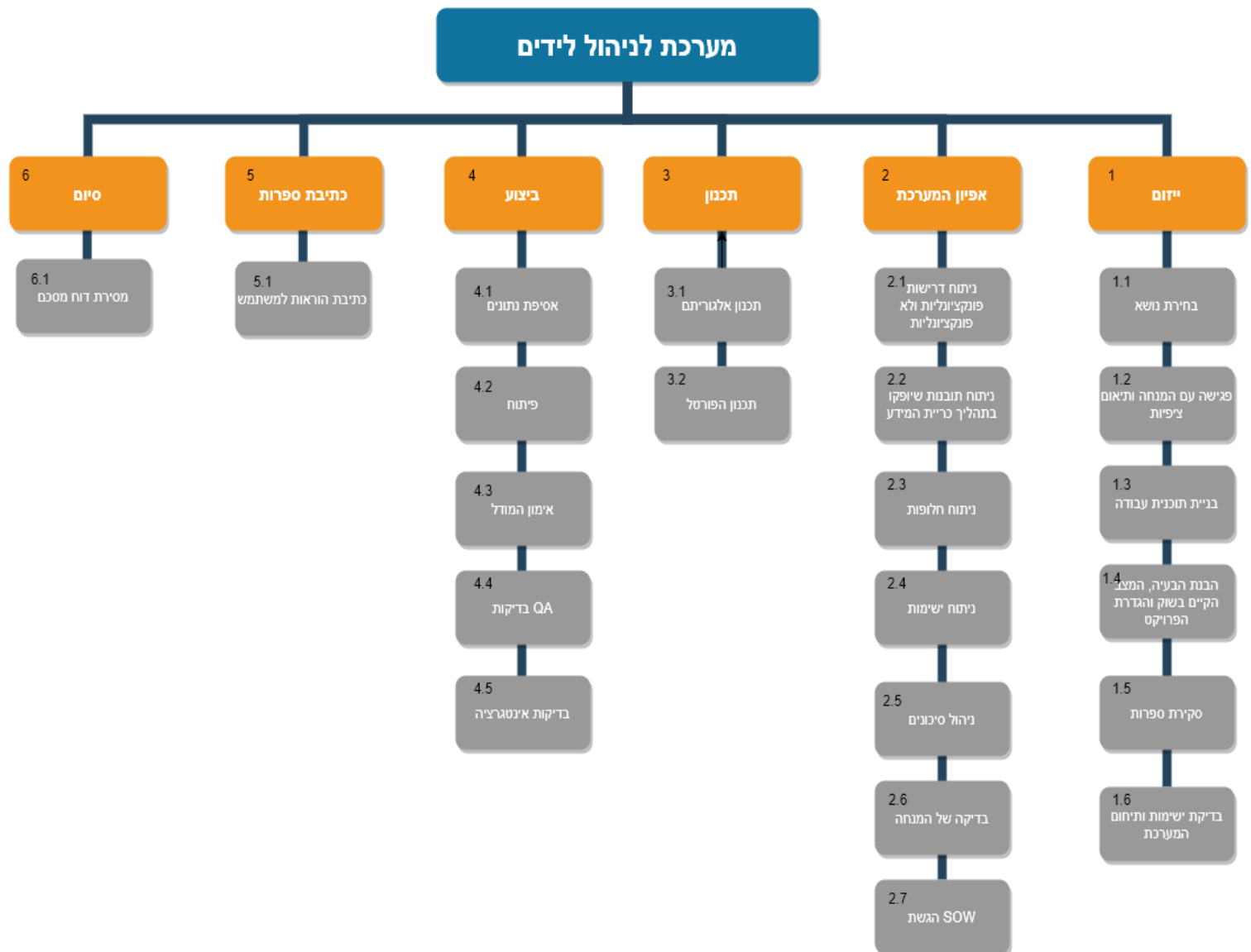
האב טיפוס של הפרויקט יהיה פיתוח אלגוריתם שישמש כמערכת המלצה עבור חברות ליסינג בסביבת הפיתוח Python שבה יהיה Dashboard - שאליו מנהל מחלקת המכירות יוכל להעלות את קובץ הלידים והמערכת תדרג את הלידים ותחלק אותם לאנשי המכירות השונים לפי הפרמטרים המתאימים. האב טיפוס של המערכת יספק אפשרות להפיק דו"חות עבור הדרג הבכיר של החברה כולל ניתוח רווחים של אנשי המכירות.

11. תכנית ניהול הפרויקט

11.1. תרשים עבודה- גאנט

ID	WBS	Task Mode	Task Name	Duration	Start	Finish	Predecessors	H2	2021	H1	H2	2022	H1	H2
1	1		ייזום	52 days	Mon 01/02/21	Sun 11/04/21								
2	1.1		בחירת נושא	5 days	Mon 01/02/21	Fri 05/02/21								
3	1.2		פגישה עם המנחה	1 day	Mon 08/02/21	Mon 08/02/21	2							
4	1.3		אישור הפרויקט במחלקת פרויקטים	5 days	Tue 09/02/21	Mon 15/02/21	3							
5	1.4		בניית תוכנית עבודה	12 days	Tue 16/02/21	Tue 02/03/21	4							
6	1.5		הבנת הבעיה הגדרת הבעיה, והבנת המצב הקיים	11 days	Wed 03/03/21	Wed 17/03/21	5							
7	1.6		סקירת ספרות	4 days	Thu 18/03/21	Tue 23/03/21	6							
8	1.7		בדיקת שיטות ותיחום המערכת	5 days	Thu 04/03/21	Wed 10/03/21								
9	1.8		הכנה להצגת הנושא במטלת אמצע	10 days	Thu 11/03/21	Wed 24/03/21	8							
10	1.9		הצגת מטלת אמצע	1 day	Sun 11/04/21	Sun 11/04/21	9							
11	2		אפיון המערכת	50 days	Mon 12/04/21	Fri 18/06/21	10							
12	2.1		ניתוח דרישות פונקציונליות ולא פונקציונליות	4 days	Mon 12/04/21	Thu 15/04/21	10							
13	2.2		ניתוח תובנות שיופיקו בתהליך כריית	3 days	Fri 16/04/21	Tue 20/04/21	12							
14	2.3		ניתוח חלופות	4 days	Wed 21/04/21	Mon 26/04/21	13,12							
15	2.4		ניתוח שיטות	3 days	Fri 16/04/21	Tue 20/04/21	12							
16	2.5		ניהול סיכונים	2 days	Wed 21/04/21	Thu 22/04/21	15							
17	2.6		SOW הצגת	9 days	Fri 23/04/21	Wed 05/05/21	16							
18	2.7		SOV-עריכת מסמך ה	18 days	Thu 06/05/21	Mon 31/05/21	17							
19	2.8		בדיקה של המנחה	14 days	Tue 01/06/21	Fri 18/06/21	18							
20	2.8.1		בדיקה של המנחה	6 days	Tue 01/06/21	Tue 08/06/21	18							
21	2.8.2		קונים לאחר בדיקה	8 days	Wed 09/06/21	Fri 18/06/21	20							
22	3		SOW הגשת	18 days	Mon 21/06/21	Wed 14/07/21	21							
23	4		הגשת דוח התקדמות	25 days	Mon 16/08/21	Fri 17/09/21								
24	5		תכנון	31 days	Mon 18/10/21	Mon 29/11/21								
25	5.1		תכנון האלגוריתם	25 days	Mon 18/10/21	Fri 19/11/21	23							
26	5.2		תכנון הפרוטל	31 days	Mon 18/10/21	Mon 29/11/21	23							
27	6		הגשת דוח ביניים	29 days	Tue 30/11/21	Fri 07/01/22	26							
28	7		ביצוע	51 days	Tue 30/11/21	Tue 08/02/22	26							
29	7.1		אסיפת נתונים	15 days	Tue 30/11/21	Mon 20/12/21	26							
30	7.2		פיתוח	13 days	Tue 21/12/21	Thu 06/01/22	29							
31	7.3		אימון המודל	8 days	Fri 07/01/22	Tue 18/01/22	30							
32	7.4		בדיקות	7 days	Wed 19/01/22	Thu 27/01/22	31							
33	7.5		בדיקות אינטגרציה	8 days	Fri 28/01/22	Tue 08/02/22	32							
34	8		כתיבת ספרות	8 days	Wed 09/02/22	Fri 18/02/22	33							
35	8.1		כתיבת הוראות למשתמש	8 days	Wed 09/02/22	Fri 18/02/22	33							
36	9		סיום	66 days	Mon 21/02/22	Mon 23/05/22	35							
37	9.1		מעבר לפני הגשה למנחה	12 days	Mon 21/02/22	Tue 08/03/22	35							
38	9.2		בדיקת מנחה	12 days	Wed 09/03/22	Thu 24/03/22	37							
39	9.3		שיפורים לפני הגשה	22 days	Fri 25/03/22	Mon 25/04/22	38							
40	9.4		הגשת ספר פרויקט	20 days	Tue 26/04/22	Mon 23/05/22	39							

WBS .11.2



11.3. ניתוח ניהול סיכונים

להלן טבלת הסיכונים הממשיים העלולים להיווצר בפרויקט. כל סיכון מקבל ציון בטווח של 1-5 המבטא את חומרת הסיכון ואת הסבירות להתרחשותו.

שם הסיכון	תיאור הסיכון	חומרה (1-5)	סבירות (1-5)	פעילות לנטרול	ל"ז לביצוע
פערי ידע הדרושים להקמת המערכת	ייתכן כי יהיו קיימים פערי ידע הנובעים מחוסר ניסיון בפיתוח וכתיבת אלגוריתמים	5	4	התייעצות עם גורמים רלוונטים והמנחה המלווה להשלמת הפערים. למידה מעמיקה בקורסי התמחות או למידה עצמית של החומר הנדרש.	החל מהגשת SOW ועד סוף הפרויקט
התנגדות עובדים לשינוי המצב הקיים	עובדי חברות הליסינג עלולים להביע חשש מהטמעת התוסף בשל חשש מחידוש או קשיים טכנולוגיים.	2	4	יש צורך לוודא כי הסגל הבכיר והזוטרי בחברה מבין את היתרונות וההתקדמות הטכנולוגית אשר התוסף מביא עמו יחד עם הבנת הצרכים האישיים של העובדים.	בקרה לאורך הטמעת התוסף
אי עמידה בלוחות זמנים, איחור בהגשת מסמכים	עיכוב של אחד או יותר מאבני הדרך של הפרויקט.	4	3	ניהול הפרויקט על פי תרשים הגאנט המכיל בתוכו זמני תגובה לתרחישים לא צפויים, שיח קבוע בין חברי צוות הפרויקט והצבת יעדים. עדכון שוטף ממחלקת הפרויקטים.	בקרה לאורך כל שלבי הפרויקט

מורכבות יתר של הפרויקט (בהגדרתו)	אפיון לא נכון של הפרויקט והעמסת דרישות פונקציונליות ולא פונקציונליות עלול להקשות על יצירת אב-טיפוס	4	2	תיאום ציפיות של צוות הפרויקט מול המנחה, הגדרה כללית וייעוץ בנושא האב טיפוס המצופה.	עד ה-14.7.21 תאריך הגשה מסמך SOW
אלגוריתם לא תקין	פעולת אלגוריתם בצורה לא תקינה שלא תסווג נכון את המידע הרצוי.	5	2	התייעצות עם המנחה וגורמים רלוונטיים, ניתוח האלגוריתם לעומק ושינוי בהתאם.	מתחילת הפיתוח וכתובת הקוד

רשימת מקורות .12

1. Gila, A. (2011). The Impact of Company Cars Taxation on Travel Demand. *The Economic Quarterly*.
2. Caron, M., Ishan, M., & Mairal, J. (2020). Unsupervised Learning of Visual Features by Contrasting Cluster Assignments. *Cornell University*.
3. How Einstein Lead Scoring Works. (2016). *Salesforce*, 2.
4. Jeannette, P., & Matthew, W. (2020). Collaborative intelligence: How human and artificial intelligence create value along the B2B sales funnel. *Business Horizons*, 403-414.
5. Jianping, G., Hongxing, M., & Weihua, O. (2019). A generalized mean distance-based k-nearest neighbor classifier. *Expert Systems with Applications*, 356-372.
6. Nygård, R. (2019). AI-Assisted Lead Scoring. *Faculty of Social Sciences, Business and*, 78.
7. Priya, R. (2017). Common pitfalls in statistical analysis: Logistic regression. *Perspect Clin Res.*, 148-151.
8. Shani, G. (2010). Evaluating Recommendation Systems. *Recommender Systems Handbook*, 257-297.
9. Thais, M. O., & Pedro Santoro, P. (2012). How Many Trees in a Random Forest? *Lecture Notes in Computer Science*, 154-168.
10. Vangelis, M., Ion, A., & Georgios, P. (2006). Spam Filtering with Naive Bayes - Which Naive Bayes? *CEAS 2006 - The Third Conference on Email and Anti-Spam*, 2006.
11. Zhi-Hua, Z. (2017). A brief introduction to weakly supervised learning. *National Science Review*, 44–53.