****תמונה שמכילה טקסט, גופן, צילום מסך, ירוק

תוכן שנוצר על-ידי בינה מלאכותית עשוי להיות שגוי.

**Data Evaluation Report**

**רום קדוש – 207691841**

**סנדרה זייגרמכר -206987521**

**תוכן עניינים**

1. **הערכת התוצאות .................................................................... 3-7**
2. **תהליך הסקירה ....................................................................... 7-8**
3. **הערכת התוצאות**

בשלב זה נערכה בחינה מקיפה של כלל התוצאות שהתקבלו מהמודלים שנבחנו במסגרת הפרויקט, מתוך מטרה לבחון האם תוצאות האלגוריתמים שהרצנו עומדות בקריטריונים שהוגדרו כיעדי הצלחה עסקיים עבור הפרויקט שלנו, בשלב זה שאפנו להבין בצורה עמוקה ולהביא לידי ביטוי את היעדים הארגוניים שקבענו מראש, וכן גיוס של תובנות עסקיות לצד שיקולים טכנולוגיים. הערכת המודלים נעשתה תוך שימוש בכלים אנליטיים, השוואתיים והסבריים, והובילה לבחירה מודעת ומנומקת של מודלים המאושרים לשלב הסופי בפרויקט.

**הצגת התוצאות וברירותן**

הצגת התוצאות בפרויקט בוצעה באופן שיטתי, שקוף ונגיש, תוך שימוש במבנה אחיד שמאפשר הבנה קלה והשוואה ישירה בין המודלים שנבדקו. עבור כל אחד מהמודלים נבנתה טבלה מפורטת הכוללת את ארבעת המדדים המרכזיים בעולם ניתוח הנתונים: דיוק כללי (Accuracy) רגישות (Recall) Precision ומדד F1. מדדים אלו איפשרו לנו לבצע הערכה כמותית מדויקת של איכות התחזיות, הן ברמת התחזית הכללית והן ברמת התחזית לכל קטגוריה של חומרת התאונה (קלה, קשה וקטלנית).

כדי לחזק את ההמחשה של פיזור התחזיות והטעויות האפשריות של כל מודל, צירפנו גם מטריצת בלבול (Confusion Matrix) לכל הרצה. מטריצות אלו איפשרו לנו לזהות בקלות אילו סוגי תאונות סווגו נכון ואילו הוסטו לקבוצות שגויות – ובכך לספק תובנה קריטית לגבי הרגישות של המודלים למחלקות חמורות ונדירות, כמו תאונות קטלניות. לדוגמה, ניכר בבירור כי מודל Random Forest המשופר הצליח לזהות כ-64% מהתאונות הקטלניות (כלומר 64 מתוך 100 זוהו נכון), בעוד שמודלים אחרים זיהו שיעור נמוך בהרבה.

תצוגת התוצאות נעשתה במקביל גם בטקסט הסברי, גם באמצעות גרפים חזותיים וגם בטבלאות השוואתיות אחידות בין כלל המודלים. הפלט הוגש בפורמט שמאפשר שילוב קל במסמכי הפרויקט, בדוחות העסקיים ובמצגות, במידת הצורך. כמו כן, סיכום כל תוצאה נעשה תחת כותרות ברורות, עם התייחסות למאפייני המודל, לפרמטרים שנבחרו, למגבלותיו ולחוזקותיו, באופן שמקנה לקורא גם אם אינו מומחה במדעי הנתונים – להבין את ממצאי הפרויקט והמסקנות בצורה ברורה.

לבסוף, כל אלו נבנו על בסיס נתונים אמין שעבר תהליך ניקוי, טיוב, והמרה, כמפורט בדוח ״Data Preparation Report", תוך טיפול בערכים חסרים, קידוד משתנים קטגוריאליים ובחירת תכונות מובהקות. כך בנינו רצף לוגי בין הבנת הנתונים, הכנתם, הרצת המודלים והצגת התוצאות – מה שמבטיח כי תוצרי הפרויקט אכן ניתנים להצגה בפני גורמים מקבלי החלטות, תוך שמירה על שקיפות, בהירות ודיוק.

**ממצאים חדשניים ויוצאי דופן**

אחד הממצאים המרכזיים והמשמעותיים שעלו מן הניתוחים הוא היכולת של מודל Random Forest בגרסתו המשופרת לזהות תאונות קטלניות – קטגוריה בעייתית במיוחד מבחינת חיזוי בשל שכיחותה הנמוכה. מודל זה הצליח להגיע ל־ Recall של 64% עבור תאונות קטלניות, בעוד שמודלים אחרים כמו Logistic Regression כלל לא הצליחו לזהות אף מקרה אחד (Recall = 0) ואילו מודלים אחרים כמו CatBoost ו-XGBoost זיהו רק 23% ו־36% בהתאמה. מדובר בקפיצת מדרגה של ממש, אשר יכולה להיות קריטית בעת קבלת החלטות במדיניות תחבורתית מניעתית.

ממצא זה בולט במיוחד על רקע מגבלות מבניות של רוב המודלים הנבחנים, אשר נוטים "להתמקד" בקטגוריה הנפוצה ביותר (תאונות קלות) ולהזניח את התחזית עבור הקטגוריות הנדירות יותר. בפרויקט זה, בו הנתונים היו לא מאוזנים באופן טבעי – 72% מהתצפיות הן תאונות קלות – נדרשה רגישות מודל ייחודית שמצליחה להתגבר על האיזון הלקוי באמצעות פרמטרים אשר תורמים ישירות לשיפור ביצועי המודל ולהתאמתו למטרות הפרויקט. הפרמטר הראשון הוא איזון בין מחלקות, אשר נועד לתת משקל גבוה יותר לקטגוריות נדירות כמו תאונות קשות וקטלניות. ללא איזון זה, המודל היה עלול להטות את התחזיות לטובת הקטגוריה הנפוצה ביותר – תאונות קלות, וכך לפגוע ביעד המרכזי של הפרויקט שלנו. הפרמטר השני הוא מינימום דגימות בכל עלה, אשר מכתיב את מספר התצפיות המינימלי הנדרש כדי שעץ ההחלטה יקבל החלטה מסוימת. הגדרה זו מאפשרת למודל להימנע מהסתמכות על מקרים חריגים או בודדים, ובכך משפרת את היכולת להכליל ולזהות תבניות אמיתיות. הפרמטר השלישי הוא מספר העצים ביער, אשר שולט בכמות עצי ההחלטה המרכיבים את המודל. הגדלת מספר העצים תורמת ליציבות גבוהה יותר, מפחיתה את השפעת הרעש, ומשפרת את רמת הדיוק הכללית של התחזיות. השילוב בין שלושת הפרמטרים הללו תרמו לבניית מודל מאוזן, יציב ומותאם היטב למטרות זיהוי תחזיתי של תאונות חמורות.

ממצא נוסף שראוי להדגיש הוא שמודל זה לא רק שיפר את יכולת הזיהוי של המקרים הקריטיים ביותר, אלא גם הצליח לשמור על ביצועים חזקים ואחידים גם בקטגוריות האחרות, Recall של 50% עבור תאונות קשות ו־85% עבור תאונות קלות. המשמעות של כך היא שהמודל מצליח להציע פתרון מאוזן – לא רק לזהות נכון את התאונות הקטלניות, אלא גם לא "לפשל" בקטגוריות אחרות.

תובנה מרכזית שעלתה מתוך השוואת כלל המודלים היא כי לעיתים דווקא מודל בעל דיוק כללי טיפה נמוך יותר עשוי להיות המתאים יותר – וזאת כאשר הוא מצטיין בזיהוי טוב יותר של קבוצות היעד הקריטיות. בפרט, בפרויקטים בעלי מטרה מנחה מובהקת, כמו איתור מוקדם של סיכון חיי אדם בכבישים, יש להעדיף מודלים שמפגינים רגישות גבוהה לאירועים חמורים, גם אם הדבר בא על חשבון הדיוק הכללי. הבחירה במודל מסוג זה אינה טכנית בלבד, אלא אסטרטגית: מערכות ניתוח מידע בתחום הבטיחות התחבורתית צריכות להעדיף זיהוי מדויק של מקרי קצה – גם במחיר פגיעה מסוימת בביצועים הסטטיסטיים הכלליים. גישה זו מדגישה את החשיבות של תיעדוף Recall בקבוצות הסיכון על פני מקסום Accuracy - במיוחד כאשר מטרת המערכת לבסוף היא לתמוך בקבלת החלטות מונעת ולא רק בניתוח רטרוספקטיבי של מידע.

**דירוג המודלים על פי רלוונטיות עסקית**

בהתבסס על כלל הממצאים, ניתן לבצע דירוג מפורט של המודלים שנבחנו בפרויקט לפי מידת התאמתם למטרה העסקית המרכזית — הפחתת תאונות חמורות (קשות וקטלניות) בישראל ב־15%. דירוג זה מתחשב לא רק במדדים כמותיים כגון דיוק ו-Recall אלא גם ברלוונטיות המעשית של כל מודל לזיהוי נכון של קטגוריות מסוכנות, שעל בסיסן ניתן לגבש המלצות מונעות.

Random Forest משופר – המקום הראשון

המודל המשופר של Random Forest דורג כמודל הטוב ביותר בפרויקט, וזאת בשל היכולת המרשימה לזהות 64% מהתאונות הקטלניות (Recall=0.64), לעומת 4% בלבד בגרסה הבסיסית. למרות ירידה קלה בדיוק הכללי (Accuracy=0.66), מדובר בהקרבה מבוקרת ונכונה כאשר המטרה היא לזהות תאונות חמורות בזמן אמת, ולא רק לחזות נכון את הקטגוריה הנפוצה ביותר (תאונות קלות). יתרה מכך, גם זיהוי התאונות הקשות (Recall=0.44) נשמר ברמה גבוהה יחסית, ונתוני Precision מצביעים על כך שהמודל יודע להבדיל היטב בין הקטגוריות. המודל הצליח לאזן בין Recall גבוה לקטגוריות הקריטיות לבין יציבות בכל רמות התחזית, ולכן הוא נבחר כמענה האופטימלי ליעדי הפרויקט.

Gradient Boosting Classifier (GBC) - המקום השני

מודל GBC הציג את רמת הדיוק הגבוהה ביותר (Accuracy=0.77) ואת ביצועי התחזית הטובים ביותר לתאונות קלות (Recall=0.94) עם זאת, המודל סבל מ-Recall נמוך במיוחד לתאונות קטלניות (רק 4%) ומעט טוב יותר לתאונות קשות (36%). כלומר, המודל מתאים כאשר המטרה היא אופטימיזציה כללית של התחזיות ולאו דווקא מיקוד בזיהוי אירועים מסכני חיים. לכן, למרות הביצועים הגבוהים במדדים הכלליים, המודל פחות מתאים ליעדי הפרויקט הממוקדים בזיהוי מוקדם של תאונות חמורות ומתן המלצות.

XGBoost משופר – המקום השלישי

הגרסה משופרת של מודל XGBoost כללה שימוש בפרמטר scale\_pos\_weight שהותאם לכל קטגוריה, דבר שהוביל לשיפור ניכר בזיהוי תאונות קטלניות Recall=0.36, ופי 9 מהגרסה הבסיסית. עם זאת, חלה ירידה ניכרת בדיוק הכללי (Accuracy=0.66), מה שמעיד על מגבלה מסוימת ביכולת המודל לשמור על יציבות בזיהוי כלל הקטגוריות. למרות זאת, העובדה שהמודל שיפר משמעותית את זיהוי הקטגוריות הקריטיות מציבה אותו במקום השלישי והמכובד.

CatBoost משופר – המקום הרביעי

מודל CatBoost המשופר הציג ביצועים מאוזנים ויציבים בכלל הקטגוריות, כולל Recall של 0.23 עבור תאונות קטלניות ו־0.50 לתאונות קשות. היתרון המרכזי של המודל הוא ביכולת הייצוב שלו ובכך שאינו דורש קידוד נוסף, אך החיסרון הבולט הוא היעדר התמקדות בזיהוי הקטגוריה הקריטית ביותר — תאונות קטלניות. לכן, למרות שביצועיו טובים יחסית, המודל אינו מוביל מבחינת התאמה ליעדי הפרויקט שלנו.

Logistic Regression בסיסי – המקום החמישי

המודל מציג את הדיוק הכללי הגבוה ביותר (Accuracy=0.76), אך נפל לחלוטין בזיהוי תאונות קטלניות (Recall=0.00) כלומר, אף תאונה קטלנית לא זוהתה בפועל. גם תאונות קשות זוהו ברמה חלקית בלבד (Recall=0.33) תוצאה זו ממחישה את המגבלות של מודלים ליניאריים פשוטים בבעיות סיווג עם class imbalance משמעותי, ולכן המודל לא מתאים בכלל לפרויקט שנועד להציל חיים.

**התאמה למטרות הארגוניות**

לאור הדרישה העסקית המרכזית שהוגדרה בשלב הראשון של הפרויקט – איתור מוקדם ואמין של תאונות חמורות וקטלניות, תוך הצגת מערכת ניבוי פרשנית שניתן להטמיע בפועל – ניתן לקבוע באופן חד-משמעי כי המודל Random Forest עונה בצורה מלאה על מטרות הפרויקט SafeRoads.

המודל מספק שילוב אידיאלי בין יציבות, פרשנות, ביצועים גבוהים ומיקוד בקטגוריות הקריטיות. ביצועיו עקביים לאורך ריצות שונות, אינו דורש שינוי דרסטי במבנה הדאטה, ומתאים לשילוב בדשבורדים או מערכות ניתוח יישומיות בגופי ממשל, רשויות מקומיות או משרד התחבורה.

הוא מציע את השילוב האופטימלי בין רגישות לתאונות קטלניות, דיוק ביצועים בכלל הקטגוריות, ויכולת הפעלה מעשית בפרויקט שנועד לזהות מוקדים מסוכנים בכבישים ולתרום ליעד המרכזי שלנו - צמצום של 15% בתאונות חמורות בישראל תוך שנתיים מרגע ההטמעה.

**שאלות עסקיות נוספות שעלו מתוך התוצאות**

בהתבסס על תוצאות המודלים והתובנות שעלו מתהליך המחקר והפיתוח שלנו, התעוררו מספר שאלות עסקיות אשר חשבנו שהן ראויות להמשך חקירה ומענה, במיוחד בשלב היישום והפריסה הארגונית של המערכת. שאלות אלו עוסקות לא רק ברמת הדיוק של המודלים, אלא באינטגרציה עם מערכות קיימות, ביכולת ההרחבה, בעמידות לאורך זמן, ובתרומה ישירה למקבלי ההחלטות.

שאלה מרכזית אחת היא: כיצד ניתן לשלב את המודל עם מערכות קיימות של משרד התחבורה, עיריות והרשויות המקומיות, כדי לבצע ניטור בזמן אמת של מוקדי סיכון?

כיום קיימים כלי GIS עירוניים, מערכות חיישנים בצמתים, ותשתיות של דיווחים משטרתיים – שילוב של המודל שלנו עם מקורות נתונים אלה עשוי להעצים את האפקטיביות של התחזיות, ולאפשר קבלת החלטות יומיומית מבוססת דאטה, בזמן אמת. שילוב כזה ידרוש ממשק API ויכולת קלט דינמית, אך עשוי לשנות את פני ניהול הבטיחות בדרכים בישראל.

שאלה נוספת שעלתה היא: האם המודל שומר על ביצועיו גם כשמשתנים תנאי הסביבה – כמו עונת השנה, אזור גאוגרפי שונה, או שינוי תשתיתי פתאומי?

מדובר באתגר מהותי לכל מערכת חזויה, שכן תאונות עשויות להיות מושפעות ממזג אוויר, חגים, תנועה משתנה או פיתוחים תחבורתיים. לכן עולה צורך בבדיקת Robustness חוסן המודל: הפעלת המודל על תתי־נתונים (למשל לפי מחוזות או לפי חודשי שנה) אשר תוכל לשפוך אור על יציבותו של המודל לאורך זמן ואפילו בהינתן ויהיו שינויים אקולוגיים כאלו ואחרים.

לבסוף, שאלה אסטרטגית לטווח ארוך היא: האם ניתן להרחיב את המערכת לתחזיות נוספות כגון עומסי תנועה, סכנת החלקה, זיהוי אירועי קצה (כמו תאונות שרשרת) או השפעות של מזג האוויר? יישום כזה ידרוש שילוב עם נתוני תחבורה בזמן אמת, נתוני תחזיות אקלים, הוספת משתנים חדשים למודל הקיים – אך יוכל להפוך את הכלי שלנו ממערכת יחסית סטטית וממוקדת לתשתית תחזית מקיפה ומתקדמת ברמת מדינה.

כל השאלות הנ״ל מהוות בסיס להמשך מחשבה ופיתוח של הפרויקט שלנו ולפריסת המערכת כפתרון להיקף תאונות הדרכים אשר מהווה נושא ציבורי בוער ורלוונטי מתמיד, ומתמקדות לא רק ביכולות הניתוח, אלא בתרומה של המודל לקבלת החלטות אמיתית ברמה מערכתית, בזמני אמת, ולמניעת תאונות – תוך עמידה באתגרי העולם האמיתי.

לסיכום, לאחר הערכת כלל המודלים שנבחנו, הוחלט לכלול בדוח הסופי את מודל Random Forest בגרסתו המשופרת כמועמד המרכזי ליישום בפועל. מודל זה נמצא כעונה בצורה הטובה ביותר הן על קריטריוני ההצלחה האנליטיים והן על קריטריוני ההצלחה העסקיים שהוגדרו בפרויקט, בראשם היכולת לתרום להפחתת תאונות חמורות בישראל ב־15% תוך שנתיים כמצוין למעלה. בעוד שמודלים אחרים הציגו תוצאות טובות בחלק מהמדדים, רק מודל זה הצליח להציג איזון עקבי, רגישות גבוהה לקטגוריות הקריטיות, ויציבות across categories. לכן, מודל זה הוא המודל המומלץ להטמעה בשלב הפריסה הממשי של המערכת, וישמש את בעלי העניין ככלי מבוסס נתונים אשר נותן מידע המאפשר זיהוי מוקדם של מוקדי סיכון, תכנון מושכל של מדיניות תשתיות, ותיעדוף פעולות התערבות בעלות פוטנציאל להצלת חיים.

**2. תהליך הסקירה**

לאחר סיום שלבי הפרויקט, ניתנה לנו הזדמנות חשובה לבחון את התהליך בכללותו – על ההצלחות, הקשיים והתובנות שהצטברו לאורך הדרך. אחת ההצלחות המשמעותיות ביותר הייתה היכולת ליישם תהליך מדעי נתונים שלם, משלב הבנת הבעיה העסקית ועד לשלב הערכת המודלים, תוך שימוש במסגרות מתודולוגיות מובנות אשר נתבקשנו לעבוד לפיהם (CRISP-DM ). הודות לתכנון מקדים והבנה עמוקה של מטרות הפרויקט כפי שפורטו בדוחות לאורך שלבי הייזום, הגדרנו יעדים מדידים אשר תורגמו לקריטריונים אנליטיים ברורים. ההגדרות האלו איפשרו לנו לבחון את הצלחת המודלים באופן עקבי, תוך חיבור ישיר בין הדרישות הביצועיות לבין מדדים כמותיים שנבחנו במהלך שלב המידול.

יחד עם זאת, הפרויקט כלל גם אתגרים לא מבוטלים. האתגר המרכזי היה התמודדות עם דאטה לא מאוזן – מצב שבו קטגוריית התאונות הקטלניות מופיעה לעיתים נדירות ביחס לשאר. מצב זה אילץ אותנו למצוא ולהשתמש בטכניקות שונות ומתקדמות בשביל לטפל בחוסר תאימות והוביל להבנה מעמיקה של חשיבות הטיפול באי־איזון נתונים כבר בתחילת התהליך. ערכים חסרים והוספת משתנים בשביל לשמור על שלמות הנתונים היוו גם אתגר עבורנו, אבל לאחר עבודה מעמיקה התגברנו על כך. כמו כן, בחירת המודל המתאים לא הייתה טריוויאלית – שכן לא נמדדה רק על פי דיוק כולל, אלא על פי תרומתו הישירה למטרה הארגונית אשר מנחה אותנו לאורך כל חיי הפרויקט - הצלת חיים וצמצום תאונות. התובנה העיקרית שעלתה לנו כתוצאה מכך היא ההכרה בכך שמודלים טובים נמדדים לא רק לפי מדדי סטטיסטיקה, אלא בהקשר בו הם מופעלים.

מהפרויקט למדנו רבות על חשיבות הרציפות בין שלבי הפרויקט – כאשר ההחלטות שהתקבלו בשלבים מוקדמים יותר (למשל בשלב הבחירה במשתנים גיאוגרפיים וסביבתיים) השפיעו ישירות על הצלחת המודל הסופי. בנוסף, התחזק אצלנו הצורך בגמישות אנליטית – לדעת לא רק לכתוב קוד, אלא להבין מתי לשלב כלים מתקדמים וחדשניים ואפילו איך לפענח מטריצות בלבול לצורך קבלת החלטות מושכלות. מנגד גם להעשיר את הידע שלנו בנוגע לתאונות דרכים באמצעות קריאת כתבות ונתונים המפורסמים באינטרנט - אלו חיזקו לנו את החיבור לפרויקט ובאמת לרצות לבצע אותו בצורה הטובה ביותר. בפרויקטים הבאים, נקפיד להשקיע זמן רב יותר ומוקדם יותר בבדיקת איכות הנתונים ובמיפוי הטיות קיימות בדאטה. כמו כן, היינו משלבת כלים חיצוניים לניתוח גיאו־מרחבי ולמידת זמן אמת, מה שהיה יכול להעשיר את הפלט האנליטי ואת התובנות מן המתרחש.

לסיכום, הפרויקט שיפר באופן ניכר את יכולותינו בתור מדעניות נתונים ובכלל בתחומי ניתוח נתונים, הבנת בעיות עסקיות, ובחירת פתרונות אנליטיים בעלי ערך ממשי – כל אלו ועוד מהווים בסיס מוצק לפרויקטים מתקדמים יותר שנעשה בעתיד. הבנו בעיקר כי כל פרויקט מדעי הנתונים הוא גם הזדמנות ללמידה והשתפרות מתמשכת כל הזמן.