

Predictive AI – מה השתנה?

עדכוני לקוחות

כולנו נסחפנו בשנה האחרונה בסערת הבינה המלאכותית. בכל מקום שנפנה מדברים על בינה מלאכותית. מערכות בינה מלאכותית חדשות מושקות לחדשות לבקרים, חברות הממתגות עצמן כחברות בינה מלאכותית זוכות להמון תשומת-לב והשקעות עתק ונראה שתחת כל עץ רענן נערך כנס או סמינר בנושא בינה מלאכותית. גם בזירה המשפטית לא נפקד מקומה של הבינה המלאכותית – בחקיקה (שעדיין בחיתוליה כאשר המאמץ המשמעותי ביותר נעשה באיחוד האירופי – ראו [עדכון לקוחות אשר פרסמנו לאחרונה בנושא זה](#), בהליכי ליטיגציה חלוציים סביב הזירה המשפטית הטרייה והלא מוכרת וכמובן בפרסומים של משרדי עו"ד...

הסיבה לכל ההמולה היא הפריצה לחיינו בסערה של כלי הבינה המלאכותית היוצרת (Generative AI). כלים כגון ChatGPT, CoPilot, Midjourney ואחרים הדהימו אותנו ביכולותיהם ליצור טקסטים, תמונות ותוצרים אחרים. אלא שעם כל הרעש סביב ה-Generative AI, קצת שכחנו שיש קטגוריה נוספת של כלי בינה מלאכותית המלווים אותנו בהיבטים שונים של חיינו כבר מספר שנים – כלי הבינה המלאכותית הפרדיקטיבית (Predictive AI). כלי ה-Predictive AI הינם מערכות בינה מלאכותית לומדות אשר תפקידן לסייע לנו בקבלת החלטות באמצעות חיזוי אירועים ומגמות עתידיים. אלה משמשות ככלים תפעוליים ותומכי החלטה.

המערכת הפרדיקטיבית מנתחת מספר עצום של נתונים ומשתנים ועל בסיסם "חוזזה" הסתברויות להתנהגויות, לדפוסים ולאירועים עתידיים. במגזר העסקי היא עשויה לסייע בחיזוי העדפות צרכנים, תגובות שוק, יכולת החזר פיננסית, צפי ל"פיק במכירות" ועוד. במערכות בריאות היא עשויה לספק תחזיות רפואיות לפי בדיקות, מידע אישי והיסטוריה רפואית. ברכב אוטונומי היא יכולה לחזות מה הסיכוי שבמצב נתון תתרחש תאונה וכאשר אנחנו מנווטים באמצעות אפליקציית ניווט היא תחזה מהו הנתבי המהיר ביותר בהתחשב בתנועה ובנתוני הכביש. כמעט בכל תחום בחיים יש מערכות כאלה כיום או שנעשים מאמצים לשילובן.

למרות שמערכות Predictive AI נוכחות בחיינו כבר שנים רבות ומהוות כיום חלק אינטגרלי מחיי היומיום שלנו, השימוש בהן אינו נטול סיכון. מערכות אלה מעוררות סוגיות משמעותיות של שקיפות, הוגנות ואתיקה שרק היום מתחילות לקבל תשומת לב ראויה.

הסיכון העיקרי נובע מכך שמערכות אלה, אשר אופן פעולתן אמור להיות שקוף וידוע, ניתן להבנה ולהסבר (explainable) ותוצאות פעולתן ניתנות להצדקה (justifiable), פועלות לרוב כבתוך "קופסה שחורה" אשר את שקורה בתוכה איש אינו יודע להסביר. אפילו לאנשים שיצרו את מערכת ה-Predictive AI או למי שמפעילים אותה לא תמיד ברור איך, ועל סמך אילו נתונים, הגיעה המערכת לתוצאה כזו או אחרת. האמור נכון במיוחד כאשר מדובר במערכות המשמשות בסיטואציות רגישות, כגון בידי גורמי הממשל ואכיפת החוק.

כך למדה "על בשרה" משטרת ישראל לאחר שבפסק דין שפרסומו הותר לאחרונה, בית המשפט המחוזי הביע ביקורת על שימושה בכלי Predictive AI מסוים שאת אופן פעולתו והסיבות שבגינן הביא לתוצאת מסוימת איש לא ידע

להסביר.

בתיק צ"א 24474-01-22 **ברגר נ' מדינת ישראל** (11.09.22), מתח השופט דרויאן-גמליאל מבית המשפט המחוזי מרכז ביקורת על מערכת "פרופיילינג" שהמטרה מפעילה בנמל התעופה בן-גוריון, העושה שימוש במידע ממאגרים שונים ותפקידה לייצר רשימה של אנשים שלכאורה הסיכוי שיימצא שהם בלדרי סמים גבוה. על-סמך הפלט של המערכת, המטרה מעכבת נוסעים שמגיעים לנתב"ג ושמים נכלל ברשימה שיצרה המערכת ונערך להם לחיפוש כדי לברר האם הם מבריחים סמים. במהלך המשפט התברר כי איש במטרה לא ידע להסביר בדיוק כיצד מגיעה המערכת למסקנות, איך נקבע מי ייכלל ברשימה שמייצרת המערכת ואיזה משקל מעניק האלגוריתם שלה לכל נתון. עמימות זו העלתה חשש לשרירותיות ולהחלטות על ביצוע חיפוש שאיננו כדון, באופן העלול לפגוע בזכויות פרטיות ולשוויון של המעוכבים.

בין נקודות הביקורת שהעלה השופט על המטרה ועל אופן שימושה במערכת:

- לא ברור איזה משקל יחסי מעניק האלגוריתם שבמערכת לכל נתון – השוטרים יודעים באילו נתונים עושה המערכת שימוש אך לא את המשקל היחסי שניתן לכל אחד מהם בפעולת המערכת, הפועלת כ"קופסה שחורה".
- האלגוריתם "מלמד את עצמו" וקובע באופן עצמאי ושאינו שקוף את המשקל שייתן לכל נתון – לאיש אין יכולת להתערב ולהשפיע באשר למשקל הנתונים.
- אין איש שיוודע בדיוק כיצד המערכת פועלת – ישנם מי שיש להם ידיעה כללית אך לא נמצא מי שידע להסביר את אופן פעולתה של המערכת בוודאות ולעומק.
- נמצא כי במערכת קיימים נתונים אשר הקשר בינם לבין חיזוי עבריינות סמים איננו מובן ("חוסר רלוונטיות") ולא ברור מדוע הנתונים הללו הוזנו למערכת.
- אין הנחיה באשר לכלילה או גריעה של נתונים מן המערכת.
- במערכת לא נעשו עדכונים ולא ברור אם נערך לה תיקוף בעבר או בהווה – השימוש במערכת החל לפני מספר שנים כפיילוט שמעולם לא הוכרז סיומו ולא בוצעו הסקת מסקנות ושיפורים בעקבותיו.

את הביקורת על המטרה בעניין זה ניתן לקבץ לשתי הקטגוריות של explainability (איש אינו יודע איך המערכת מגיעה לתוצאות אותן היא מספקת ולא יודע להסביר את אופן פעולתה) ו-justifiability (עקב העדר היכולת להסביר איך המערכת עובדת לא ניתן לתקף ולהצדיק את מסקנותיה).

אמנם פסק הדין עוסק בביקורת על גוף ציבורי, אשר מטבעו כפוף לכללים מחמירים יותר של סבירות, שקיפות ואתיקה, אך אין ספק שהעקרונות שעלו בו הינם בעלי השפעה גם על המגזר הפרטי וחשוב לשים אליהם לב. האמור נכון במיוחד ביחס לפעולתן של מערכות במגזר העסקי אשר להן השלכות על זכויותיהם של אנשים (למשל מי זכאי לקבל משכנתא ומי לא, מי זכאי שיוצעו לו מוצרים כאלה ולא אחרים) והן יוצרות סיכונים ברורים לחברות שפועלות ומסתמכות עליהן.

אז מה עושים?

1. ככל שניתן, חשוב לעשות שימוש ולהעמיד לרשות לקוחות רק מערכות Predictive AI שאנו יודעים להסביר את

אופן פעולתן ולהצדיק את הגעתן לתוצאות כאלה או אחרות.

2. יש להקפיד על נהלים והנחיות ברורים באשר להוספת או גריעת נתונים מהמערכת ולוודא שהמערכת לא משקלת נתונים שאינם רלוונטיים.
3. ידיעת הנתונים אינה מספיקה, חשוב שתהיה יכולת להסביר איזה משקל ניתן לכל נתון וכיצד הוא משוקלל יחד עם יתר הנתונים – ורצוי שיהיה מישהו בארגון שמבין את המערכת לעומק ויודע להסביר בדיוק כיצד היא פועלת ומגיעה למסקנות.
4. מערכות Predictive AI הן כלי חשוב ורמתן עולה מדי שנה – אך הן לא מושלמות. חשוב לשמור בן אנוש בתהליך כדי לתקף את התוצאות ואת אופן ההגעה אליהן.
5. חשוב לבצע תיקוף ועדכון של המערכת באופן תדיר ולשמור את הממצאים ואת התיעוד הטכני שלה.
6. ככל שמערכת Predictive AI מופעלת כפיילוט, חשוב לציין זאת באופן ברור וגלוי ולקבוע בבירור מתי מסתיים הפיילוט ומה קורה לאחר מכן (למנוע מצב בו השימוש במערכת ממשיך ללא שהוסקו מסקנות מהפיילוט ויושמו לקחים ושיפורים נדרשים).

אנו מסייעים ללקוחותינו בצמצום הסיכונים המשפטיים הנובעים מהטמעת מערכות בינה מלאכותית חדשות, בבניית והחלת מדיניות AI לחברה ובהבנת ההיבטים המשפטיים של הפעלת והנגשת מערכות כאלה. פנו אלינו ונשמח לסייע.

תודה להילה דוידי על הסיוע בהכנת נייר עדכון זה.

עדכון זה נועד לספק מידע כללי ותמציתי בלבד. והוא אינו מהווה ניתוח מלא או שלם של הסוגיות הנידונות, אינו מהווה חוות דעת משפטית או ייעוץ משפטי ואין להסתמך עליו.

אנשי קשר



נטע שקד שטדלר
שותפה



ד"ר זיו רוטנברג
שותף