

עקרונות מנחים לפיתוח טכנולוגיות מבוססות למידת מכונה

הקדמה

טכנולוגיות בינה מלאכותית (Artificial Intelligence, AI) ולמידת מכונה (Machine Learning, ML) יש פוטנציאל לשנות את שירותי הבריאות על-ידי הפקת תובנות חדשות וחשובות מכמות הנתונים הגדולה שנוצרת במהלך אספקת שירותי הבריאות. טכנולוגיות אלו עושות שימוש באלגוריתמים כדי ללמוד מתנאי השימוש בעולם האמיתי, ובמצבים מסוימים עשויות לעשות שימוש במידע זה כדי לשפר את ביצועי המוצר. עם זאת, ישנם היבטים הדורשים התייחסות ייחודית בפיתוח טכנולוגיות אלו בשל המורכבות שלהן ובשל האופי המשתנה והמתפתח שלהן.

באחרונה פרסמו רשויות רגולטוריות מובילות בעולם, FDA, Health Canada וה-MHRA, עשרה עקרונות מנחים להתנהלות טובה בתחום לימוד מכונה - Good Machine Learning Practice (GMLP). עקרונות מנחים אלה יסייעו לקדם פיתוח ושימוש בטוח, יעיל ואיכותי במכשירים רפואיים אשר עושים שימוש בבינה מלאכותית ולמידת מכונה (AI/ML). ניתן דגש מיוחד לעקרונות ייחודיים ללמידת מכונה, כגון: מניעת הטיות (bias), מניעת התאמת יתר (over-fitting), שקיפות, מהימנות הנתונים, ושמירה על בטיחות בהתייחס לפרטיות וביטחון המידע.

עשרת העקרונות המנחים להלן נועדו להניח את היסודות להתנהלות טובה בתחום לימוד מכונה המתייחסת לאופי הייחודי של מוצרים אלה. הם גם יסייעו לטפח צמיחה עתידית אחראית בתחום זה המתקדם במהירות.

| עקרונות מנחים להתנהלות טובה בתחום לימוד מכונה | |
|--|--|
| שילוב מומחיות רב-תחומית לאורך כל מחזור החיים של המוצר | בניית מודל המותאם לנתונים הזמינים ומשקף את השימוש המיועד במוצר |
| שימוש בכלים מיטביים של הנדסת תוכנה ואבטחת מידע בבניית המודל | התמקדות בביצועים של צוות האדם-בינה מלאכותית |
| ייצוג ראוי של אוכלוסיית המטופלים המיועדת במסגרת משתתפי המחקר הקליני ומערכי הנתונים | שלב הבדיקה (testing) מדגים את ביצועי המכשיר בתנאים רלוונטיים מבחינה קלינית |
| בסיסי הנתונים המשמשים לאימון ולבחינה הינם בלתי תלויים | העברת מידע חיוני למשתמשים בצורה ברורה |
| בחירת נתוני ייחוס שפותחו בשיטות הטובות ביותר | ניטור ביצועי המודלים המוטמעים וניהול הסיכונים הכרוכים באימון מחדש |

משרד הבריאות הישראלי מתכוון לאמץ עקרונות אלו כקו מנחה במטרה לסייע לגורמים העוסקים בפיתוח טכנולוגיות המבוססות על לימוד מכונה (AI/ML), וזאת על-ידי התאמת שיטות עבודה טובות שהוכחו במגזרים אחרים והתאמתן כך שיהיו ישימות לטכנולוגיה רפואית ולמגזר הבריאות; עקרונות אלו יהוו בסיס לבנייה של שיטות חדשות ספציפיות לעולם הבריאות, אשר ימשיכו ויתעדכנו ככל שתחום המכשור הרפואי מבוסס למידת מכונה יתפתח. עבודה זו מתבצעת בתיאום, ותוך כדי דיאלוג מתמשך עם הרשויות הרגולטוריות המובילות בעולם.

אימוץ הקווים המנחים אשר פותחו ע"י הרשויות המובילות בעולם והתאמתן למערכת הבריאות הישראלית מסייע גם בתהליך הרמוניזציה של המסגרת הרגולטורית לטכנולוגיות מבוססות למידת מכונה, ומאפשר לטכנולוגיות אלו ניווד גלובלי אחוד יותר, בטוח ומהיר יותר, לטובת שיפור שירותי הבריאות לאזרחי ישראל.

עשרת העקרונות המנחים:

1. **שילוב מומחיות רב-תחומית לאורך כל מחזור החיים של המוצר** לטובת בחינה כי מכשירים רפואיים מבוססי בינה מלאכותית בטוחים, יעילים ועונים על צרכים קליניים משמעותיים לאורך מחזור החיים של המוצר. בחינה זו דורשת הבנה מעמיקה של האופן בו עתיד המודל להשתלב בתהליך העבודה הקליני, לצד היתרונות הרצויים והסיכונים הנלווים למטופלים. גיוון אנושי בתהליך הפיתוח יכול לתת ביטוי למגוון דיסציפלינות מקצועיות, תפיסות חברתיות, וקבוצות באוכלוסייה.
2. **שימוש בכלים מיטביים של הנדסת תוכנה ואבטחת מידע בבניית המודל** תוך מתן תשומת לב לעקרונות הבאים: שיטות הנדסת תוכנה טובות, אבטחת איכות נתונים, ניהול ותיעוד נתונים, הגנה על פרטיות, ושיטות אבטחת סייבר מעודכנות. שיטות עבודה אלה כוללות תהליך תכנון וניהול סיכונים שיטתי שיאפשר לזהות ולתקשר כראוי החלטות ורציונל של תכנון, הטמעה וניהול סיכונים, כמו גם להבטיח את מהימנות ותקינות הנתונים.
3. **ייצוג ראוי של אוכלוסיית היעד במסגרת המחקר הקליני ומערכי הנתונים**: פרוטוקול איסוף הנתונים צריך להבטיח שמאפייני אוכלוסיית היעד (לדוגמה: גיל, מין, מגדר, מאפיינים סוציו-דמוגרפיים), מאפייני השימוש, ומדדי המודל, מיוצגים באופן ובהיקף מספקים במחקר הקליני ובמערכי הנתונים, כך שניתן יהיה להכליל את תוצאותיו באופן סביר לאוכלוסייה הרלוונטית. עקרון זה הינו חשוב כדי לנהל הטיית אפשריות, לקדם ביצועים הולמים הניתנים להכללה על פני אוכלוסיית היעד, להעריך את השימושיות ולזהות נסיבות שבהן המודל עשוי ליצר ביצועי חסר.
4. **בסיסי הנתונים המשמשים לאימון ולבדיקה צריכים להיות בלתי תלויים**: תהליכי הבחירה והתחזוקה של בסיסי הנתונים יבטיחו חוסר תלות. כל מקורות התלות הפרוטנציאליים, לרבות: מאפייני המטופלים, רכישת נתונים, ומאפייני אתר ההרצה, נלקחים בחשבון ומקבלים מענה, כדי להבטיח את עצמאות בסיסי הנתונים המשמשים לבדיקה ולבחינה.
5. **בחירת נתוני ייחוס שפותחו בשיטות הטובות ביותר**: שימוש בשיטות מקובלות ומיטביות לפיתוח מערך נתוני ייחוס (reference dataset), קבוצות לוגיות של ערכי קוד המהווים סטנדרט הנגיש למשתמשים שונים באופן מרוכז, כמו לדוגמה, קודי אבחנות), באופן שמבטיח איסוף נתונים רלבנטיים ומאופיינים היטב מבחינה קלינית, והבנת מגבלות נתוני הייחוס. ככל שניתן, נעשה שימוש במערכי נתוני ייחוס מקובלים בפיתוח ובדיקת מודלים, המקדמים ומדגימים את יכולת ההדירות (robustness) וההכללה (generalizability) של המודל בקרב אוכלוסיית היעד.
6. **בניית מודל המותאם לנתונים הזמינים העדכניים ומשקף את השימוש המיועד במוצר**: בניית המודל מותאמת לנתונים הזמינים ותומכת בהפחתה פעילה של סיכונים ידועים, כגון: התאמת יתר (overfitting), ירידה בביצועים וסיכוני אבטחה. היתרונות הקליניים והסיכונים הקשורים למוצר מובנים היטב, משמשים להפקת יעדי ביצוע קליניים משמעותיים לבחינה, ותומכים בכך שהמוצר יוכל להשיג את השימוש המיועד בו בבטחה וביעילות. השיקולים כוללים את ההשפעה של ביצועים בזירה המקומית ובזירה הבינלאומית ואת חוסר הוודאות/השונות בין הנתונים המוזנים למודל, הנתונים המופקים מהמודל, אוכלוסיית היעד ותנאי השימוש הקליניים. יש מקום להתייחס למימד הזמן של הנתונים הזמינים, מתוך הבנה כי נתונים אלו משקפים ניסיון קליני מתפתח ומשתכלל על-פני זמן.
7. **התמקדות בביצועים של צוות האדם-בינה מלאכותית**: כאשר למודל יש "בן אדם בתהליך" (human-in-the-loop), יש לתת דגש לביצוע של צוות האדם-בינה מלאכותית, תוך התייחסות לגורם השיקול והפרשנות האנושית במקום רק לביצוע של המודל עצמו. במסגרת ביצועים אלו יש מקום להתייחס לממשק העבודה של האדם עם המודל, וליכולת הפרשנות של האדם את המודל (interpretability).
8. **שלב הבדיקה (testing) מדגים את ביצועי המכשיר בתנאים רלוונטיים מבחינה קלינית**: פיתוח והוצאה לפועל של תכניות בדיקה מבוססות סטטיסטית בכדי להפיק מידע קליני רלבנטי על ביצועי המכשיר, ללא

תלות במערך נתוני האימון. השיקולים כוללים את אוכלוסיית היעד, תתי-קבוצות חשובות, סביבה קלינית ושימוש על ידי צוות האדם-בינה מלאכותית, תשומות מדידה וגורמים מתערבים (confounders) פוטנציאליים.

9. העברת מידע חיוני למשתמשים בצורה ברורה: מתן גישה למשתמשים למידע ברור, בהקשר רלבנטי, ואשר מותאם לאוכלוסיית היעד (כגון: ספקי שירות או מטופלים), במטרה לבסס אמון של המשתמשים במודל. מידע זה כולל את השימוש המיועד של המוצר והתוויות לשימוש, ביצועי המודל עבור תתי-קבוצות רלבנטיות, מאפייני הנתונים אשר שימשו לאימון ובדיקת המודל, סטנדרטים ופרוטוקולים מקצועיים ששימשו להנחות המודל, הנתונים הנדרשים להזנה, מגבלות ידועות במודל, פרשנות ממשק משתמש, הסבר מותאם על ממצאי המודל (explainability), ואינטגרציה של המודל בתהליך העבודה הקלינית, כולל תהליך האימות (validation) הנדרש לפני שימוש ראשוני. יידוע המשתמשים בשינויים ועדכונים במכשיר שמקורם מניטור ביצועים בעולם האמיתי (Real World Evidence), הבסיס לקבלת החלטות כאשר זמין ואמצעים לתקשר חששות לגבי המוצר לגורם המפתח.

10. ניטור ביצועי המודלים המוטמעים וניהול הסיכונים הכרוכים באימון מחדש: יש להבטיח כי למודלים מוטמעים יש את היכולת להיות מנוטרים בשימוש ב"עולם האמיתי" תוך התמקדות בשימור או שיפור במדדי הבטיחות ובביצועים. בנוסף, כאשר מודלים מאומנים מעת לעת או באופן מתמשך לאחר ההטמעה, קיימות בקרות מתאימות לניהול סיכונים של התאמת יתר, הטיה לא מכוונת או ירידה בביצועי המודל (לדוגמה, סחף של מערכי נתונים, Data drift), אשר עלולה להשפיע על הבטיחות והביצועים של המודל כפי שהוא משמש את צוות האדם-בינה מלאכותית. יש מקום לבחון במסגרת ניהול הסיכונים את תוכנית ההטמעה, גם בהינתן הפעלה בסביבות שונות.