|  |
| --- |
| **הסברים לדף השער**  בוקר טוב,  התוכנית שלנו היא צ'אט-בוט מיוחד שמדמה עבודה של פסיכותרפיסט עם מטופל. אך המילה 'מדמה' אינה לגמרי נכונה. אלן טיורינג ניסח מבחן עבור צ'אט-בוט כך: השיחה לא תזהה אם היא מתקיימת עם אדם או מחשב. במקרה שלנו אין מטרה להסתיר שבן השיחה הוא מחשב.  המטופל מודע לכך שהוא מדבר עם מכונה שאינה בעלת תודעה. התוכנית שלנו תופסת נישה דומה לניהול יומן, אך מסוגלת לעזור לנהל ולהדריך לניתוח תחושות רגשות המטופל. |
| **הרלוונטיות של הנושא.**  הרלוונטיות של תוכנות כאלו גוברת בתקופות של משברים חברתיים - אמיתיים או מדומים - מלחמות, מגפות או בחירות. הנתונים הסטטיסטיים מעידים שאין לצפות לעידן השלווה הצפוי.  האב בטיבט שואל נזיר:  - מה אתה עושה כעת?  - צופה בעצים צומחים.  - עסוק מדי!  זה לא המטה שלנו.  מאז ה-7 באוקטובר, הגידול במחלות הפסיכולוגיות גדל ב-45-52%. והגידול במספר המקרים האקוטיים גדל פי 9.  כמעט 40 אלף אנשים מתחת לגיל 24 פנו לעזרה פסיכולוגית. שליש מהשיחות היו קשורות לחרדה וטראומה, ואחת מכל חמש הייתה על כאב לב, דיכאון או בעיות נפשיות אחרות.  רבים זקוקים לעזרה פסיכולוגית כעת, ולא בעוד כמה חודשים כשהגיע תורם להגיע לאיש המקצוע בכללית. |
| עם זאת, אנו מודעים למגבלות של האמצעים המוצעים. קשה להעריך את היעילות. אך, אם התוכנית שלנו תהווה פלצבו, הרי שמוכח כי פלצבו יקר יותר יעיל יותר ( Ig Nobel Prize ברפואה ב-2008). לכן, בהחלטות שלנו אנו משתמשים בפתרונות יחסית יקרים במקום ספריות חינמיות ומהירות יותר.  נעשה הסתייגות חשובה. המטרה של תוכנה מסחרית לפסיכותרפיסט אינה לרפא את המטופל כמה שיותר מהר, אלא להפוך את התלות שלו בתוכנה לכרונית. אם המטופל ירגיש בריא, הוא ימחק את האפליקציה. כמובן, אם המטופל יקפוץ מהגשר, זו גם תהיה אובדן של לקוח, אך עם גשרים קופצים לעיתים רחוקות, לכן זה לא נושא סיכונים מסחריים חמורים.  אבל אנחנו לא כאלו - אנחנו מצטיינים! |
| **היסטוריה 1**  רעיון של מכונת פסיכותרפיה אינו חדש. ב-1956 כתב רוברט שיקלי את הסיפור "רפואה רעה". שנת 2103. אדם סובל מתשוקה מניתית להרוג את חברו. הוא קונה רובוט פסיכותרפיסט המתוכנן בטעות לטפל במאדים. אך למאדים אין רצון להרוג. התוצאה היא שהרובוט לא מבין את המניעים של המטופל ומטפל בו משליחה פסיכולוגית מאדים .לאחר ריפא את החולה ממאניה רצחנית, הרובוט מחדיר לחולה מספר מחלות פסיכולוגיות של מאדים.  ב-1966 פרסם ג'וזף ויצנבאום תוכנית בשם ELIZA שיכולה לשוחח עם המשתמש. היא השתמשה בחוקי התאמה של דפוסים כדי לגבש תגובות. בתסריט שנקרא DOCTOR היא הייתה מסוגלת לנהל שיחות בסגנון פסיכולוג אמפתי, בעקבות סגנון של קרל רוג'רס, שהציג שאלות פתוחות לעודד תקשורת יותר אפקטיבית בין מטופלים למטפלים.  ויצנבאום נדהם מהתוכנית שלו שנתפסה ברצינות בעיני משתמשים, שפתחו את לבם כלפיה. הידוע ביותר הוא כשצפה במזכירתו שהשתמשה בתוכנה - למרות שידעה שזו סימולציה - היא ביקשה: "האם תוכל מלצאת מהחדר בבקשה?"  רובוט-פסיכולוג מקשיב למטופל וחושב לעצמו: לעזאזל, זה נהדר שהאנטי וירוס שלי לא יאפשר דיכאון. |
| **היסטוריה 2**  בין הבוטים הקיימים בתמיכה פסיכולוגית אפשר לציין את הבאים:  **1**. בוטים עם אלגוריתמים מתוכננים מראש:  הם מבוססים על שיטות פסיכותרפיה מוכחות, אך אינם מבוססי AI לגמרי.  למשל. Woebot  זהו עוזר וירטואלי לתמיכה בבריאות הנפש, שנוצר על בסיס תרפיה קוגניטיבית התנהגותית (*cognitive behavioural therapy, CBT*).  למרבה הצער או השמחה, אנחנו לא לגמרי מבינים מה זה. לכן, לא נעמיד פנים שהבנו למה התכוונו המדברים.  התוכנית מקיימת שיחות קצרות כדי להבין את המצב הרגשי של המשתמש, ומלמדת טכניקות פרקטיות של CBT, ושיטות מדעיות נוספות. היא נותנת עצות לניהול רגשות ולהפחתת סטרס, ושומרת יומן מצב רוח כדי לעקוב אחר שינויים.  . **2.**מלווים AI-  מטרתם היא להיות חבר וירטואלי היוצר קשר רגשי עם המשתמש ועוזר לפתח יכולות תקשורת.  לדוגמה, Replika. היא משתמשת במודלים דמויי GPT ליצירת תגובות קונטקסטואליות ואדפטיביות, מה שהופך את השיחות ליותר טבעיות ואישיות.  התוכנית יכולה לנהל שיחות שונות, על פי האישיות והעניין של המשתמש. |
| תוכניות מודרניות יכולות לנתח רגשות (כגון טון הדיבור, ביטוי הפנים, שינויים בדופק ובטמפרטורה), מה שמאפשר להבין ולהזדהות עם המשתמש. מערכות כמו ChatGPT יכולות להבין את הקונטקסט ולזכור שיחות קודמות, ומספקות לא רק עצות אלא גם תמיכה רגשית.  כלים המבוססים על Retrieval-Augmented Generation (RAG) גנרציה עם הגברה של איחזור מאפשרים למודלים להשתמש במקורות נתונים חיצוניים שונים על המטופל, מעבר לדיבורו בלבד.  אבל למעשה, איננו יודעים על מודלים שיוכשרו הן על דיבור והן על ביומטריה של מטופלים בו-זמנית. |
| **שיטות העבודה העיקריות של התוכנית שלנו**  לתוכנית שלנו בחרנו במודל טיפול המכונה Socratic Questioner "מודע השואל הסוקרטי". זוהי כנראה אחת מהשיטות הפשוטות ביותר, והיא כוללת בעיקר מענה לשאלה באמצעות שאלה נוספת.  זוכרים את הבדיחה הישנה:  - למה אתה תמיד עונה לשאלה בשאלה?  - למה לא?  השימוש בשיטה הסוקרטית בפסיכותרפיה מבוסס על שאילת שאלות מכוונות על מנת לעזור למטופל לחקור את מחשבותיו, אמונותיו והתנהגותו. גישה זו נועדה לעזור בהבנת והביקורת על מחשבות אי-רציונליות  .  הפסיכולוג אינו נותן הנחיות או מביע דעתו. במקום זאת, הוא שואל שאלות מעמיקות שמביעות מהמטופל להביא, לבדוק והשערות לחקור תוצאות, לבדוק אלטרנטיבות ולהתמודד עם רעיונות בצורה בונה. לדוגמה, הפסיכולוג יכול לשאול:  ״למה אתה חושב שזה נכון או חשוב?״  ״למה זה מעורר אצלך רגשות חזקים?״  ״אילו הנחות אתה עושה כאן?״  ״האם תוכל להביא נקודת מבט נוספת על זה?״  ״מה היה קורה אם הכל היה משתנה?״  ״איך היית עונה למישהו בעל דעה הפוכה?״  מטרת הפסיכולוג היא לעורר חשיבה ביקורתית והבנה, ולא להנחות או לשכנע. |
| Fine-tuned LLM model  בכדי להשיג את היעדים שלנו, הכשרנו מחדש את המודל gpt-4o-2024-08-06.  למעשה, GPT-4 מבין באופן סביר את השיטה של השואל הסוקרטי בעצמו. לכן ההכשרה הייתה מכוונת על:  1. לתאר את תפקיד ה-LLM, כאילו הוא פסיכולוג מוסמך ומומחה. אל תדאגו, זו לא כירורגיה. עקבו אחרי השיטה הסוקרטית ולא תצטרכו לחשוש  .  2. לתקן כמה טעויות שגילינו בבדיקות, כמו לא לקרוא לנמענת לפציינט, ולא להזכיר את שמו כל פעם שהיא פונה אליו.  3. להבהיר למודל אילו מידע ובאיזה מבנה הוא יקבל לצורך ניתוח, מעבר לרשומות השיחה.  למרות שניתן להעביר כל זה כהודעת מערכת SystemMessage , ללא הכשרה מחדש הפרומט יהיה ארוך יותר, וה-LLM יהיה יקר ואיטי יותר.  Retrieval-Augmented Generation  בנוסף לרשומות השיחה הנוכחיות, מידע רלוונטי משיחות קודמות עם המטופל מועבר למודל ה-LLM. התוכנית מעריכה את תוכן הבקשה של המטופל, מבצעת embedding , ומשווה עם embedding של השיחות הקודמות על ידי מרחק קוסינוסי.  ניתוח רגשות – בינתיים הוכנסה פונקציה פשוטה על בסיס שני מאפייני דיבור עיקריים: אנרגיה וגובה טון מהספרייה librosa. אך ניתן להחליף זאת בפתרונות מתמחים יותר כמו pyAudioAnalysis או opensmile. |
| **סקירה של מבנה בסיס הנתונים**:  בסיס הנתונים די פשוט.  טבלה "מטופל" כוללת: שם, מין, גיל, מידע נוסף על המטופל, שמתעדכן לאחר כל מפגש עם המטופל.  טבלה "שיחות" כוללת: טקסט מלא של השיחה, תקציר שיחה - משמש כעזר ל -LLM embedding תקציר שיחה - קוד וקטורי לאיתור מהיר של השיחות הדרושות.  לצורך חיבור לבסיס הנתונים של SQL Server מבחוץ יש לפתוח את הפורט 1433. למרבה הצער, לא הצלחנו לפתוח את הפורט דרך הנתב של בזק ולרשום DNS דינמי.    לכן, על אף שתיכנו את התוכנית ב-Streamlit, היא פועלת רק במחשב המקומי.  נלחמנו עם זה זמן רב. ואז עשינו קעקוע עם המספר 1433 על החזה והסתכלנו על סרט הקומדיה "ג'וקר". |
| **תכנית פעולת התוכנה**  הלוגיקה של התכנית די פשוטה.  1. המטופל מזוהה או נרשם. מידע נלקח מבסיס הנתונים והתוכנה מברכת אותו בשמו. לדוגמה, "שלום, Hannibal Lecter ".  2. המטופל אומר משהו לפסיכולוג.  3. מתבצע ניתוח של הטון הרגשי של הדיבור.  4. התוכנית מזהה את הדיבור וממירה אותו לטקסט, ומחפשת בבסיס הנתונים:  - שיחה עם שאלה דומה,  - שיחה עם שאלה מנוגדת,  - השיחה האחרונה עם המטופל.  לדוגמה, אם המטופל מתלונן כעת על תיאבון רע, אנו מוצאים שכאשר מתלונן Hannibal על צרבת, ובשיחה אחרת תיאר חוויה קולינרית חיובית.  5. ל-Fine-tuned LLM model מועברים:  - השאלה הנוכחית,  - תקצירי השיחות הנבחרות,  - תיאור הטון של השאלה.  6. ה-LLM מייצר תגובה, לדוגמה:  “מה את חושבת שיכול היה להשפיע על תיאבון רע?”  תגובה זו מושמעת למטופל.  7. תהליך זה חוזר על עצמו עד שהמטופל רוצה להפסיק את השיחה.  8. בסיום השיחה נוצר:  - תקציר השיחה,  - embedding תוכן השיחה,  - עובדות נוספות על המטופל מופקות מהשיחה.  כלומר, אם המטופל ציין שהוא רכש כלב, ה-LLM יפיק את המידע שיש למטופל כלב.  9. לאחר מכן מידע בבסיס הנתונים מתעדכן. בכאן מסתיימת פעולת התוכנית. |