המחלקה להנדסת תוכנה

קורס פרויקט גמר בהנדסת תוכנה– מספר 10051

**דו"ח בטא – סמסטר ב'**

|  |  |
| --- | --- |
| **פרטי הפרויקט** | |
| **סטודנט - שם: אביגייל הילה שרבף**  **מספר ת. זהות: 318631488**  **שם: ליאל לוי**  **מספר ת. זהות: 207045741** | **מנחה אקדמי - שם: ד"ר יהודה חסין** |
| **כותרת** פרויקט הגמר **בעברית:**  מציאת משפטים, מילים ואותיות בכתב יד של השפה הערבית. | |
| **כותרת** פרויקט הגמר **באנגלית:**  Finding handwritten sentences,words andletters ofthe Arabic language | |
| **דו"ח** | |
| **תיאור ההתקדמות בפרויקט:**  **לאחר הבנה שה-dataset שהגיע עם רשת הנוירונים מהאינטרנט אינו מספיק טוב עבור ה-data שלנו,**  **בנינו dataset מתוך החיבורים שברשותנו וזה גרם לאחוזי הדיוק לעלות.**  **לאחר ייצוב רשת הנוירונים המשכנו לשלב הבא בפרויקט שהוא כתיבת אלגוריתם קוף.**  **לאחר כתיבת האלגוריתם בעת בדיקתו נתקלנו בבעיה המחייבת שמירה וטעינה של רשת הנוירונים (דבר**  **שלא הצלחנו לפתור בשלבים מוקדמים יותר).**  **בנוסף, הרכבנו את רצף ריצת הקוד ושיפרנו את מבנה ויעילות קטעי הקוד הקודמים.** | |
| **מצב תוצרי הפרויקט**  **פרויקט מחקרי – פרוט בקצרה (3 משפטים) של התוצאות המרכזיות**  **התוצאות המרכזיות שאליהן הגענו הן:**  **ייצוב רשת הנוירונים והגעה לאחוזי דיוק טובים.**  **רשת הנוירונים משתמשת ב-dataset שהרכבנו וכתוצאה מכך אחוזי הדיוק עלו משמעותית.**  כתיבת אלגוריתם קוף (נדרשת בדיקה).  **כולם – יש תיעוד? יש יומן link?**  [לינק ליומן](https://trello.com/b/eWo8hygi) | |
| **התכנית לסיום (לו"ז, משימות, אבני דרך):**  **המשימות שנותרו לנו לבצע הן:**   1. **שמירה וטעינה של רשת הנוירונים.** 2. **כתיבת ה-auto-encoder.** 3. **שיפור אחוזי הדיוק במידת הצורך.** 4. **כתיבת מערכת וניסויים אחרונים.** | |
| **עדכון סיכונים בפרוייקט (מה סיכויים?, מה השלכות?):**   1. **בעיה בשמירה וטעינה של רשת הנוירונים.**   **במידה ולא נצליח לטעון את רשת הנוירונים נצטרך לכתוב מחדש את רשת הנוירונים**  **בפלטפורמה אחרת.**  **הסיכויים נמוכים שמצב זה יקרה.**   1. **בעקבות הבעיה של שמירה וטעינה של רשת הנוירונים לא נספיק להגיע לשלב כתיבת**   **ה-auto-encoder או שלא נספיק לסיים אותו.**  **השלכות דבר זה הם שלא נצליח להגיע להחלטה סופית של זהות המחבר.** | |
| **היבטים של הנדסת תוכנה (יישום והתנסות, ולקחים – מה והאם הייתי עושה אחרת):**   1. **בעקבות הפרויקט התמודדנו עם נושא שלם של רשתות נוירונים ואימון מודלים ללא שום רקע**   **קודם.**  **דבר זה עיכב אותנו לא מעט במהלך הפרויקט שכן זה נושא לא פשוט והוא מרכז הפרויקט.**  **הלקח שלנו הוא לקחת קורס אינטרנטי לפני תחילת הפרויקט, זה היה נותן לנו רקע**  **ובסיס לנושא זה.**   1. **בנוסף, עניין השפה (הפרויקט נעשה עבור חיבורים בשפה הערבית – לא שפת האם שלנו) ולכן**   **היינו צריכות להיעזר באנשים חיצוניים דוברי השפה ודבר זה עיכב את תהליך בניית**  **ה-dataset והוצאת נתונים מהחיבורים.**  **הלקח שלנו הוא להבין את מידת מעורבות השפה בפרויקט יותר לעומק וכך להיערך**  **בצורה טובה יותר (מציאת אנשים דוברי השפה וכ'ו..)** | |
| **ח ת י מ ו ת** | |
| **חתימת הסטודנט: אביגייל הילה שרבף**  **חתימת הסטודנט: ליאל לוי**  **תאריך:19/05/2021** | **חתימת המנחה האקדמי:**  **תאריך:** |