**פרויקט שנת י"ב ImageToText:**

**פיצ'רים בפרויקט:**

**1. צילום תמונה באמצעות האפליקציה –** המערכת תצלם מהמשתמש תמונה שמכילה טקסט עבורה המערכת תמיר את הטקסט בתמונה לטקסט המוקלד בה.

**2. קליטת התמונה מהמשתמש –** המערכת תקלוט מהמשתמש תמונה שמכילה טקסט עבורה המערכת תמיר את הטקסט בתמונה לטקסט המוקלד בה.

**3. עריכת התמונה: חיתוך התמונה -**  המשתמש יוכל לחתוך את התמונה, לבחור את הערך שהוא רוצה שיומר על ידי המכונה

**4. עריכת התמונה: סיבוב התמונה -** המשתמש יוכל לסובב את התמונה

**5. עריכת התמונה : הבהרת התמונה -**  המשתמש יוכל להבהיר את התמונה במידת הצורך.

**6. שליחת התמונה לשרת –** לאחר שהמשתמש ערך את התמונה ובחר את החלק אותו הוא רוצה להמיר, הוא שולח אותה לשרת שמכיל את המכונה.

**7. דגימת התמונה כדי למצוא איזה גוון מבדיל בין הטקסט לרקע התמונה –** לאחר שהמשתמש שולח את התמונה למכונה, המכונה למעשה מבדילה בין הטקסט לבין רקע התמונה, שלב זה מבוצע בשביל שלב זיהוי האותיות בהמשך.

**8. זיהוי שורות הטקסט ויישור השורה כך שהאותיות יהיו אופקיות –** המכונה למעשה מזהה שורות טקסט ומיישרת אותם במידת הצורך , שלב זה פותר בעיות של דפים שסריקתם נעשתה בזווית, מה שמקשה על התוכנה לזהות את צורת האותיות בצורה נכונה.

**9. הפרדת השורות השלמות למילים על פי מרווחים –** המכונה מפרידה את השורות למילים לצורך זיהוי האותיות בהמשך.

**10. הפרדת השורות השלמות לאותיות על פי המילים –** המכונה מפרידה את המילים לאותיות לצורך זיהוי האותיות בהמשך.

**11. חיפוש בתוך מאגר המידע של האותיות –** המכונה מחפשת במאגר המידע באיזה אות מדובר.

**12. זיהוי האותיות בתמונה שהתקבלה –** המכונה לומדת תבניות של אותיות בכך שעוברת על עשרות אלפי דוגמאות של אותיות, בונה לעצמה תבניות של אותיות ובכך מזהה את התו ומעביר את התו לשרת.

**13. סיווג האותיות בתמונה שהתקבלה –** לאחר חיפוש במאגר והבנה באיזה אות מדובר, היא מסווגת את האות (אומרת באיזה אות או תו מדובר).

**14. זיהוי מקום של האות בתוך התמונה –** מסדרים את האות שסיווגנו לפי הסדר שהם היו בתמונה.

**15. הפיכת הסיווג לפונט קריא –** האותיות שסווגו הופכות לפונט קריא שניתן להבין.

**16. קבלת לינק לקריאה בפונט קריא –** לאחר סידור וסיווג הטקסט שהיה בתמונה, המערכת מעבירה את הטקסט שהומר לקובץ txt או word.

**17. תרגום הטקסט שזוהה בתמונה –** אם המשתמש בחר בתרגום הטקסט, הטקסט שהיה בתמונה בשפה האנגלית יתורגם לשפה העברית

**18. הרצת קוד שזוהה בתמונה –** אם המשתמש בחר בסריקת קוד, תמונת הקוד שצילם המשתמש בשפת python, מומרת לקוד וניתן להריץ אותה באמצעות המערכת.

**19. העברת קובץ הטקסט ל PDF-** אם המשתמש בחר בהעברת קובץ הטקסט לpdf, הטקסט שהומר ועבר לקובץ txt או word מועבר לקובץ pdf**.**

**20. שינוי הטקסט בתמונה עצמה –** אם המשתמש בחר בשינוי הטקסט בתמונה עצמה, יהיה ניתן לסדר ולשנות את הטקסט בתמונה עצמה.

**טכנולוגיות בפרויקט :**

|  |  |  |
| --- | --- | --- |
| **פיצ'ר / תהליך:** | **טכנולוגיות ושפות תכנות:** | **משאבים נדרשים ושירותים חיצוניים:** |
| קליטת התמונה מהמשתמש | נכתוב אפליקציה לאנדרואיד ( נכתוב אותה בשפת פייתון, ובאפליקציה נקלוט את התמונה מהמשתמש. ) | לא נדרשים כרגע. |
| עריכת התמונה שהתקבלה מהמשתמש | נשתמש בספריית opencv בפייתון לשם עריכת התמונה | לא נדרשים כרגע. |
| שליחת התמונה לשרת | נשתמש בפייתון בsocket/scapy | לא נדרשים כרגע. |
| דגימת התמונה כדי למצוא איזה גוון מבדיל בין הטקסט לרקע התמונה | באמצעות ספריית opencv, נשתמש  בפונקציה cnvtColor ונהפוך את התמונה לצבע שחור-אפור-לבן | לא נדרשים כרגע. |
| (?) זיהוי שורות הטקסט ויישור השורה | באמצעות ספריית opencv בפייתון. | לא נדרשים כרגע. |
| (?) הפרדת השורות השלמות לאותיות ומילים על פי מרווחים | באמצעות ספריית opencv בפייתון. | לא נדרשים כרגע. |
| זיהוי האותיות בתמונה שהתקבלה | נשתמש במודל שנקרא CNN שהוא השימושי ביותר לזיהוי תמונות, נכתוב את המודל באמצעות ספרייה שנקראת TensorFlow2 ,ספרייה של למידת מכונה שכתובה ב-Python | נשתמש בפייתון בסביבת google collab ( אנו עוד נשקול האם יש מקום לעשות שימוש בסביבת Jupyter). |
| חיפוש בתוך מאגר המידע של האותיות | המכונה שנבנה לזיהוי האותיות בתמונה תצטרך לחפש במאגר המידע EMNIST את האות אותה זיהתה. | נשתמש בפייתון בסביבת google collab ( אנו עוד נשקול האם יש מקום לעשות שימוש בסביבת Jupyter). |
| סיווג האותיות בתמונה שהתקבלה. | לאחר שהאות שזוהתה נמצאה במאגר המידע EMNIST , המכונה תצטרך לסווג אותה לפי האות שנמצאה במאגר. | נשתמש בפייתון בסביבת google collab ( אנו עוד נשקול האם יש מקום לעשות שימוש בסביבת Jupyter). |
| (?) זיהוי מקום של האות בתוך התמונה | נשתמש בפייתון, ייתכן שזה בר ביצוע באמצעות מתמטיקה פשוטה של חישוב מקומות של האותיות. | לא נדרשים כרגע. |
| קבלת לינק לקריאה בפונט קריא | נשתמש בפייתון ונעביר את הטקסט אל תוך קובץ word | לא נדרשים כרגע. |
| תרגום הטקסט שזוהה בתמונה | נשתמש בפייתון בsocket/scapy , ונשלח את הטקסט אל השרת של google Translate | לא נדרשים כרגע. |
| הרצת קוד שזוהה בתמונה | נשתמש בפייתון , ונכתוב interpreter באפליקציה | לא נדרשים כרגע. |
| העברת קובץ הטקסט  ל PDF | נעביר את קובץ הword אל קובץ pdf | לא נדרשים כרגע. |

**שאלות על הציר הטכנולוגי :**

**1. באיזה טכנולוגיות אנחנו הולכים להשתמש בפרויקט? מה הטכנולוגיה המובילה/העיקרית?**

**תשובה:** בפרויקט שלנו אנחנו הולכים להשתמש בשני טכנולוגיות מרכזיות לכל הפרויקט. ראשית, טכנולוגיית עיבוד תמונה , שמטרתה היא לקלוט את התמונה מהמשתמש, לערוך אותה ומכאן לשלוח אותה אל השרת , שם היא תנקה את התמונה מרעשים , תהפוך אותה לשחור לבן כך שיהיה קל יותר לזהות ולהפריד בין אות לרקע, לאחר מכן היא תזהה שורות הטקסט ותיישר השורה כך שהאותיות יהיו אופקיות במידת הצורך, ולאחר מכן תפריד את השורות לאותיות ומילים לצורך זיהוי האותיות בהמשך. בנוסף, הטכנולוגיה העיקרית היא טכנולוגיית למידת המכונה וזיהוי האותיות וסיווגן. אנו נשתמש בטכנולוגיית למידת מכונה כדי שהמכונה באמצעות מאגר מידע תלמד את התבניות של האותיות לצורך זיהוי האות עצמה ולאחר מכן סיווג האות , כלומר לומר באיזה אות מדובר.

**2. בתוך הטכנולוגיה העיקרית איזה חלקים נראה לכם שנצטרך ללמוד? (לפעמים זו טכנולוגיה / שפת תכנות מאוד גדולה ורחבה וצריך להתמקד בחלק מסוים שלה או בספרייה ספציפית).**

**תשובה:** בטכנולוגיית עיבוד התמונה נשתמש בספריית opencv בפייתון.

בטכנולוגיית למידת המכונה נשתמש בספריית tensorflow2 (keras) בפייתון. מאגר המידע שנשתמש הוא EMNIST.

**3. מה אתם חושבים שרמת הידע שלכם בכל טכנולוגיה?**

**תשובה:** רמת הידע שלנו בטכנולוגית הפרויקט הן בסיסיות אך המטרה שלנו במהלך הפרויקט וגם בפרק זמן זה הוא ללמוד את הטכנולוגיות לעומק ולהבין כיצד להשתמש בהן. כמו כן, חקרנו על הטכנולוגיות והבנו כיצד יש לממש אותם וכיצד הן עובדות ואנו מבינים את הטכנולוגיות.

**4. איך אתם חושבים שנכון לכם ללמוד את החלקים שאתם לא מכירים? (התשובה יכולה להיות שונה לכל איש צוות וזה בסדר, כל אחד לומד בצורה שונה).**

**תשובה:** כדי ללמוד את החלקים שאנחנו לא מכירים נצטרך לחקור את החלק לעומק, כמו כן נדאג להיות בקשר רציף עם הרש"צ/ מנטור בכדי לדעת כמה העמקה לבצע בחלק מהנושאים. כמו כן, נעבור על פרויקטים מוכנים בחלקים שאנחנו לא מכירים ונלמד ונבין אותם, נצפה בסרטונים מסבירים ונתנסה בכדי להתקדם.

**5. האם יש התקנות של כלים/תוכנות/ספריות שאתם חושבים שיכולים להיות רלוונטיים?**

**תשובה:** נצטרך להתקין את הספריות opencv לצורך עיבוד התמונה וtensorflow2 לצורך למידת המכונה.

**6. האם יש תחומי ידע שלא קשורים לשפת תכנות שרצוי ללמוד?**

**תשובה:** תחומי ידע שלא קשורים לשפת תכנות אותם נרצה ללמוד זהו מודל CNN, שהיא השימושית ביותר לזיהוי של תמונות ובאמצעות שימוש במודל זה נבנה ונלמד את המכונה.

**קישורים שיכולים לעזור:**

**1. לצורך פיתוח האפליקציה :**

<https://realpython.com/mobile-app-kivy-python/>

**2. לצורך הפרויקט עצמו :**

<https://www.pyimagesearch.com/2020/08/17/ocr-with-keras-tensorflow-and-deep-learning/>

<https://www.pyimagesearch.com/2020/08/24/ocr-handwriting-recognition-with-opencv-keras-and-tensorflow/>

**3. לצורך עיבוד התמונה :**

<https://www.learnopencv.com/image-alignment-feature-based-using-opencv-c-python/>

**4. לצורך למידת המכונה :**

<https://reshetech.co.il/machine-learning-in-the-browser/how-to-develop-hand-written-digits-recognition-web-app>

<https://towardsdatascience.com/build-a-handwritten-text-recognition-system-using-tensorflow-2326a3487cd5>

**הDatasets שלנו הם :** MNIST ו- KAGGLE

