**מבוא למחשוב ענן - סמסטר חורף התשפ"ה**

**תרגיל בית 3**- **עבודה בצוותי העבודה – צוות Tiger**

המשימה בתרגיל זה: סיום בניית המערכת, ביצוע קוד ריוויו לצוות אחר: איטרצית פיתוח מס 2

שימו לב: למטלה זו שלושה חלקים

חלק ראשון.עליכם להגדיר את בעלי התפקידים לאיטרציה זו (יש להחליף מתרגיל בית 2) יש לרשום מי מהנדס.ת המערכת, האחראי.ת בתרגיל זה.

מהנדס מערכת: בן זכאי:  
חלקות העבודה מול הצוות נעשתה בפגישה הראשונה שתואמה על ידי הסקראם מאסטר, ונוהלה בקובץ דוק אינטרנטי שהיה במעקב רציף של מהנדס המערכת, הסקראם מאסטר וה-QA לצורך קידום העבודה.  
הפגישות נעשו בזום פעמיים בשבוע ותואמו על ידי הסקראם מאסטר כדי להתאים בלוז הכולל, בנוסף חברי הצוות שהתעסקו יותר בפרונט קיימו 3 פגישות יעודיות לצורך קידום עבודה וכנ"ל גם חברי הצוות שעבדו על הבק קיימו 3 פגישות יעודיות לצורך קידום העבודה. טרם הצגת המצגת בוצעה פגישה בזום של כלל חברי הצוות לצורך הצגת כלל הצדדים של הקוד והסברים מפורטים כך שכל חבר צוות יכיר כל חלק בפרויקט.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **שם חבר הצוות** | **תפקיד בתרגיל זה** | **משימות שהוקצו** | **משימות שהושלמו** | **בדיקת קבלה (acceptance test) – האם הבדיקה עברה?** |
| נדיר | Scrum master  Frontend | זימון פגישות, חלוקת משימות, מעקב אחר משימות, עזרה בעיצוב העמודים, העברת עמוד אדמין לHTML, עבודה על הצגת גרפים של סטטיסטיקות. | הכל :) | כאשר משתמש מזין מונח חיפוש קיים המערכת תציג מידע רלוונטי וקישורים למונח מהאתר של Vercel. |
| דימה | Product  Backend | פיתוח זחלן להתאמה לקישורים מרובים, בניית אלגוריתם לתעדוף אינדקס, ריפקטור לקוד בבקאנד והצגתו לצוות הפרונט, וידוא של עמידה בדרישות הלקוח. | הכל :) | כאשר משתמש מזין מונח שאינו נמצא באינדקס המערכת תציג הודעת שגיאה ידידותית. |
| יובל | Backend | שינוי צורת חיבור ועדכון ל-DB לצורך ייעול זמנים, הפרדת מילים בזחלן, עבודה על אלגוריתם לתעדוף אינדקס, כתיבת מיקרו שירותים. | הכל :) | כאשר אדמין מוחק מונח מהאינדקס בעמוד האדמין, המונח הנמחק אינו מופיע יותר בתוצאות החיפוש. |
| אלדר | Frontend | ריפקטור של קוד בפרונטאנד, המרת וידג'טים לHTML עבור חיפוש מילים ועבודה על הצגת גרפים. | הכל :) | עמוד הסטטיסטיקה מציג נתונים נכונים, גם לאחר שינוי של האינדקס ע"י האדמין. |
| שי | Backend  QA | קישור אינדקס והעלאת המידע ל-DB, שיפור בניית אינדקס ויצירת פרי ופוסט פרוסס, עבודה על כתיבת מיקרו שירותים, וידוא בדיקות קבלה ומקרי קצה בקוד. | הכל :) | לעמוד האדמין רק משתמשים מורשים יכולים לגשת. ניסיון גישה ללא הרשאה יוביל להודעת שגיאה מתאימה. |
| רואי | Frontend | העברת צאט בוט לפורמט HTML בהצגה, ריפקטור לקוד של פרונט, העברה של כלל המסכים לפורמט HTML, עדכון עיצוב הפרונט. | הכל :) | חיפושים של מונחים עם אותיות קטנות או גדולות מחזירים את אותן תוצאות. |
| בן | User Interface | עיצוב ממשק נוח וקל לתפעול, חשיבה על זרימת התהליכים, יצירת אבטיפוס של מסכים לצורך בדיקת היתכנות לצוות הפרונט עבור מעבר מ-widgets ל-HTML, ליווי של צוות הפרונט. | הכל :) | הקישורים שמוצגים למשתמש מסודרים בסדר הגיוני לפי אלגוריתם מסודר. |

*חלק שני: בניית המערכת* (80 נקודות)

עליכם לסיים את בניית המערכת. בפרט, יש לוודא קיום של כלל האלמנטים הבאים:

1. כל הפיצ'רים שהתחלתם לממש בתרגיל בית 2 - יש לוודא שכל המסכים תקינים ועובדים כראוי, **בקולאב**,ולעבוד עם firebase לצורך שמירת נתונים .אין להשאיר קישורים או כפתורים למסכים שלא קיימים.

יש לוודא שניתן להריץ את הפרויקט מהמחברת בלבד , ללא כל צורך בהטענה של קבצים נוספים או יציאה למערכת חיצונית (כגון FLASK). (30 נקודות)

תשובה 1:כל הפיצ'רים ממומשים ונמצאים בדף Colab יחיד בלינק:

[TigerVercel\_HW3.ipynb](https://colab.research.google.com/drive/1FIctTIYxFi1hUt1UKq2afl7wPgEtWITm?usp=sharing)

1. שימוש ב -microservices (תרגול 7)- עליכם לעשות שימוש לפחות בשניים. פרטו באילו שירותים עשיתם שימוש, ומה היתרון בשימוש בשירותים אלו. (10 נקודות)

תשובה 2:בחלק הראשון של הקוד (Backend), חילקנו את פעולות הזחילה שאנחנו מבצעים לארבעה מיקרו שירותים:

* **Upload object service:** מיקרו שירות שמאפשר לעדכן אובייקט ששמור במחלקה ולאחר שנעדכן אותו כרצוננו, הוא מעלה אותו לשרת Firebase. הוא מכיל בתוכו פונקצית העלאה, שבה מעלים את האובייקט השמור לענן, עם כל המילים שנשלפו, הלינקים שבהם המילים מופיעות, כמות המופעים שלהן, וה-score שלהן לכל דף.   
  היתרון בשירות זה הוא הפשטות בשמירת כל הנתונים, כיוון שאנחנו קוראים לו בכל פעם שאנחנו רוצים לשמור בו רשימת מילים והלינקים שבהם הן מופיעות, וכשנרצה, נקרא לפונקציה שמעלה לשרת.
* **Dictionary creation service:** מיקרו שירות שבנינו שמאפשר למשתמש ליצור מילון לכל המילים שקיימות בדף. כאשר נפנה למחלקה זו עם הקלט של כל המילים בדף, היא בונה לנו את המילון של כל המילים הקיימות, כמות המופעים שלהן, ואת המיקום (האינדקס) המדויק בהן הן נמצאות. בנוסף, יש לה פונקציה שמחשבת את ה-score של כל מילה בדף מסויים. מילה שמופיעה כמות רבה של פעמים ובתחילת הדף, תקבל ערך גבוה יותר ממילה שמופיעה הרבה בסוף הדף.

היתרון בשירות הוא שניתן לפנות למחלקה הזאת עם מערך של מילים, והיא תחזיר לנו מילון מסודר ועם כל המידע שנצטרך.

* **Pre-process service:** מיקרו שירות שמחזיק פונקציות לניקוי ועיבוד טקסט. מורכבת מפונקציה מרכזית שמשתמשת ב-5 נוספות. היא מקבלת מחרוזת של מילים, ויודעת לעבד את המחרוזת לרשימה של כל המילים לאחר ניקוי של סימני פיסוק, מספרים, ועוד.

היתרון בשירות זה הוא שהוא בסך הכל מקבל מחרוזת ארוכה, ומחזיר רשימה של מילים "נקיות" למשתמש המבקש.

* **Post-process service:** מיקרו שירות שלוקח את המידע ה-raw ומוסיף metadata עבור כל מונח (word count, page count). כלומר, עבור כל מונח שקיים באתר, נוסיף לו את כמות המופעים שלו בכל הדפים, וכמות הדפים בהם המונח נמצא.

ובנוסף הוא מחשב metadata של הסטטיסטיקות המעניינות (מידע הנחוץ במקרה של שינוי ידני של מסד הנתונים), והסטטיסטיקות עצמן.

היתרון של שימוש במיקרו שירות זה הוא פיצול והפרדת האחריות של עיבוד המידע, והוא מרכז פיתוח, תחזוקה ועדכון של הלוגיקה הזאת באופן בלתי תלוי במחלקות אחרות.

1. צטבוט שישולב במערכת (כפי שלמדתם בתרגול 8), הכולל מידע אשר מגיע מה - DB שלכם, ועונה בצורה אינטליגנטית לשאלות המשתמש. (15 נקודות).

תשובה 3:הצ'אטבוט משולב בפרויקט ונגיש למשתמש לאחר שהוא מחפש מונח ממסד הנתונים.

הוא מציע למשתמש הסבר קצר ומסכם על המונח אותו חיפש המשתמש ויכול בנוסף לענות על שאלות.

1. מהם ה -KPI הרלוונטיים בפרויקט שלכם? הסבירו? (10 נקודות)

תשובה 4:המערכת שלנו עומדת בשלושה Key Performance Indicators מרכזיים:

* **User experience:** הממשק של המערכת שלנו מינימליסטי, מאורגן, פשוט מאוד לשימוש.
* **Performance:** שליפת הנתונים בעת פנייה לשרת על ידי המשתמשים מתבצעת בצורה מהירה מאוד, מסתכמת בזמני ריצה של ms700-900 בסה"כ.
* **Security:** למשתמשים רגילים אין גישה למסד הנתונים, אלא רק בתהליך הפנייה לצורך החיפוש. משתמש פשוט אינו יכול לגשת לגישת המנהל ולשנות מידע במסד. רק מנהל יכול להתחבר לגישת המנהל ע"י שם משתמש וסיסמא ייחודיים.

1. שקיפות אלגוריתמית – כיצד הבהרתם למשתמשים את האלגוריתמים והנתונים הנאספים בקוד שלכם? (5 נקודות)

תשובה 5:אנו מציגים למשתמש שקיפות אלגוריתמית ע"י הסבר מופשט של צורת העבודה של המערכת שלנו ביחס לאתר שאנחנו עובדים מולו (Vercel) ומציגים את התוצרים למשתמש.



1. אתגרים שעלו במהלך העבודה , וכיצד התמודדתם איתם (לדוגמא- DB גדול מדי, יצירת ויזואליזציה, מבנה הקוד וכו) (10 נקודות)

תשובה 6:נתקלנו בשני קשיים מרכזיים במהלך תכנון וביצוע הפרויקט:

* **Frontend:** במטלה 2 החלטנו לבצע אינטגרציה בין קוד HTML לבין וויג'טים של פייתון. הקוד היה רץ בצורה טובה אבל העיצוב היה מינימלי ולא מאורגן היטב.

במהלך העבודה על תרגיל בית 3, נתקלנו בבעיה בשילוב בין ה-html והוויג'טים וגילינו שלא ניתן לבצע אינטגרציה נקייה בין שתי המערכות בלי לוותר על איכות הניראות של הפרויקט. לכן, החלטנו לשנות את כל הקוד הישן ולהמיר את הוויג'טים ל-HTML, דבר שלקח זמן רב אבל השתלם לנו לבסוף.

* **Backend:** נדרש זמן לחשוב על אלגוריתם לדירוג המילים בכל דף. בנינו 5 נוסחאות מתמטיות לחישוב התיעדוף של המילים אבל ללא הצלחה. לבסוף, ביצענו סיעור מוחות באחת מפגישות הצוות והגענו לאלגוריתם יחסית פשוט שנותן ערך גבוה יותר למילה ככל שהיא נמצאת גבוה יותר בדף, ומגדילה את הערך הזה עוד יותר אם היא מופעה כמה פעמים בדף. האלגוריתם הזה והחישוב שלו מפורטים לעומק בתיק המתכנת, בחלק 1: Backend, במחלקת DictionaryCreationService, פונקציה get\_score\_of\_word\_in\_page.

*חלק שלישי : סגירת הפרויקט \_(20 נקודות)*

בחלק זה תכינו מסמכים המתארים את הפרויקט שלכם:

1. יש לבנות תיק למתכנת הכולל את שמות כל הקבצים המרכזיים, פונקציות מרכזיות, קטעי קוד/תבניות עיצוב מעניינים שהשתמשתם בהם.

תשובה: **מצורף בתיקיית ההגשה.**

1. יש לבנות תיק למשתמש , הכולל הסבר כללי על המערכת , פירוט מסכים, מעברים בין מסכים והסבר על טעויות אפשרויות.

תשובה: **מצורף בתיקיית ההגשה.**

1. התייחסו בתיק המשתמש לשקיפות אלגוריתמית – כיצד הבהרתם למשתמשים את האלגוריתמים והנתונים הנאספים בקוד שלכם?

תשובה: **מצורף בתיקיית ההגשה.**

1. יש להכין סרטון קצר של 30-60 שניות, המתאר את השימוש במערכת. הסרטון משמש כ – elevator pitch  למערכת שלכם, כלומר יש לכלול בו הסבר מקצועי ועם זאת שיווקי , המדגיש את האלמנטים המיוחדים של המערכת שבניתם. יש להגיש את הסרטון בפורמט mp4..  
   תשובה: **מצורף בתיקיית ההגשה.**

הוראות הגשה:

1. יש להגיש את התרגיל בצוותים, בתיקיית ה –GIT שלכם, וכן בתיקייית התרגיל ב moodle. **חובה לכלול קישור לתיקיית הגיט בקובץ במוודל.**
2. יש להגיש במוודל קובץ זיפ הכולל קובץ וורד ובו מענה לשאלות, וקישור ל- notebook ובו הקוד שלכם (יש לוודא שהקישור פומבי ונגיש).
3. כותרתו של הקובץ תהיה HW3\_TEAMNAME
4. שימו לב כי כל העבודות חייבות להיות שונות זו מזו.אנו מריצות תוכנה לבדיקת עבודות זהות. עבודות שייראו דומות ייפסלו ויינתן עליהן ציון 0.

בהצלחה!