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

**תרגיל בית 2** -– **עבודה בצוותי העבודה**

מועד הגשה: 28.5.25

המשימה בתרגיל זה: בניית מסכים מרכזיים במערכת

שימו לב: למטלה זו שלושה חלקים

*חלק ראשון : בניית מערכת (10 נקודות)*  
  
כיצד נעשתה החלוקה מול הצוות :   
חלוקת המשימות הוקצתה לאחר הניסיון מהתרגיל בית הראשון, סטודנטים חזקים יותר קיבלו משימות קשות יותר, סטודנטים השולטים בבקנד ובמימוש אלגוריתמים קיבלו את משימת מנוע החיפוש ואלו ששולטים בעיצוב קיבלו הכנת המסכים וממשק המשתמש.  
  
האם היה ממשק בין חברי הצוות:   
כן, הייתה עבודת צוות טובה בין רוב חברי הצוות ועבודה טובה על המטלות. סטודנטים חזקים יותר לקחו על עצמם משימות ששאר הסטודנטים לא הצליחו לבצע.

המשימות מולאו במלואן על ידי כל חברי הצוות.

|  |  |  |
| --- | --- | --- |
| **איטרציה 1** | | |
| **שם חבר הצוות ותפקיד בתרגיל זה** | **משימות שהוקצו** | **משימות שהושלמו** |
| ניב אורן Scrum master ,Project manager | חלוקת העבודה בין חברי הצוות, מימוש האינדקס, מימוש מנוע החיפוש | all |
| יוני גרינברג *backend developer* | מימוש האינדקס, מימוש מנוע החיפוש | all |
| אלי רובינוב  *frontend developer,procduct manager* | מימוש מנוע החיפוש, מימוש המסכים | all |
| עמנואל אלקובי  UI,QA | מימוש המסכים וממשק המשתמש | all |
| שחר נחום | לא הוקצה תפקיד – מילואים | None |
| אדהם עלואן  Frontend developer | הוספת פיצ'ר למערכת | None, done by Eli |

בניית אינדקס (20 נקודות)

באיטרציה זו עליכם לבנות את מסד הנתונים שמכיל את האינדקס של המילים המשמעותיות באתר [mqtt.org](http://mqtt.org), הפרוטוקול איתו אנו עובדים לקבלת מידע מהחיישנים.

מבנה האינדקס צריך להיות אחיד לכל הקבוצות , ולכלול **לפחות** את השדות הבאים  **( אין לשנות את שמות השדות!)**:

|  |  |
| --- | --- |
| שם השדה | הסבר |
| term | term |
| DocIDs | רשימת קישורים לדפים המכילים את ה- term, ממוספרים לפי בחירתכם |

משימות:

1. ממשו את האינדקס בקולאב.  
  
  
2.רשמו בצורה מפורשת את רשימת ה stop words שבחרתם, ונמקו מדוע בחרתם במילים אלו.   
  
custom\_stop\_words = {

    'a', 'an', 'the', 'and', 'or', 'to', 'from', 'of', 'in', 'on', 'at', 'for', 'by', 'with', 'as',

    'this', 'that', 'these', 'those', 'it', 'its', 'be', 'is', 'are', 'was', 'were', 'has', 'have', 'had',

    'not', 'no', 'yes', 'true', 'false', 'null', 'none', 'n/a',

    'temperature', 'humidity', 'pressure', 'sensor', 'value', 'data', 'reading',

    'device', 'system', 'status', 'msg', 'message', 'update', 'packet', 'received',

    'unit', 'type', 'time', 'timestamp', 'date', 'id', 'code', 'level', 'output', 'input',

    'low', 'high', 'normal', 'current', 'average', 'min', 'max', 'warning', 'alert', 'ok', 'error',

    'up', 'down', 'left', 'right', 'start', 'stop', 'run', 'check', 'set', 'get',

    'if', 'then', 'else', 'when', 'while', 'until'

}

**נימוק לבחירת רשימת ה־Stop Words:**

ברשימת ה־stop words שבחרנו כללנו:

* **מילות תפקוד באנגלית כללית** כגון: the, is, at, to, from
* **מילות עזר דקדוקיות** כגון: this, that, has, was, be
* **ערכים לוגיים וטכניים חסרי משמעות חיפושית** כגון: true, false, null, none
* **שמות שדות טכניים החוזרים בכל הודעה מ־MQTT** כמו:
  + temperature, humidity, pressure, sensor, value, data
  + שמות טכניים כלליים כמו device, status, packet, timestamp, id
* **תיאורים כלליים או חסרי משמעות הקשרית** כמו:
  + low, high, ok, normal, error, warning, average, max, min

**מטרת הסינון:**

לסנן החוצה מילים שכיחות או טכניות שחוזרות כמעט בכל הודעה ואינן מוסיפות ערך מוסף למנוע החיפוש.  
מילים אלו יוצרות "רעש" באינדקס ומונעות חיפוש אפקטיבי של מונחים ייחודיים ובעלי משמעות.

3. ציינו האם השתמשתם ב stem/lemmatization לצורך בניית האינדקס.

**שימוש ב־Stemming לצורך בניית האינדקס:**

במהלך בניית האינדקס השתמשנו ב־**Stemming** מסוג **PorterStemmer** (ממימוש של ספריית NLTK).  
המטרה הייתה לאחד מונחים בעלי שורש זהה אך בצורות כתיב שונות — לדוגמה:

* temperature - temperatur
* reading, readings - read
* message, messages - messag

השימוש ב־Stemming עוזר לצמצם כפילויות מיותרות באינדקס ולרכז תוצאות תחת שורש אחיד, גם אם המילה הופיעה בצורת יחיד, רבים או בהטיה אחרת.

לא נעשה שימוש ב־**Lemmatization**, מאחר והוא דורש משאבים כבדים יותר וזמן עיבוד ארוך, שאינם נחוצים במסגרת פרויקט זה.

*חלק שני: בניית מסכים להצגה בכיתה (50 נקודות)*

בחלק זה תכינו 4 מסכים, אותם תציגו לחבריכם בפעילות שתתבצע בכיתה.

המסכים צריכים לכלול (לפחות) (20 נקודות):

מסך מנהל למערכת, מסך שאילתא למנוע החיפוש (הזנה והצגת תוצאות), מסך סטטיסטיקות מעניינות.

בשלב זה נדרש לממש במלואם את בניית מסך המנהל ומסך השאילתא. ניתן את התוצאות להציג כרגע עם data חלקי.מומלץ לממש ככל הניתן גם את מנוע החיפוש , כפי שלמדתם בתרגול 6 😊

את מסך הסטטיסטיקות עליכם לממש בצורה בסיסית. עם זאת כמובן ,שכל תוספת שתחליטו עליה, תוביל להערכה גבוהה יותר של המשימה.

בשבוע ההרצאות של 19-20.5.25 תציגו את המערכות שבניתם. המפגש יתנהל במתכונת סטודיו – כל צוות מגיעה **במלואו** לאחד המועדים עם לפטופ, כל הצותים מציגים במקביל. הסטודנטים מסתובבים בין הצוותים, **מתנסים** במערכת,וממלאים משוב .

ההצגה ומילוי המשוב הם **חובה.**

אנא השתבצו בהקדם לאחת מקבוצות ההרצאה על מנת לוודא שקיים איזון בין הצוותים:

https://docs.google.com/spreadsheets/d/1-grJWvFQGtMkCJ8lg9v9ZSYULBvA4ixaEskBCZglIiE/edit?gid=0#gid=0

לאחר ההצגה תקבלו באופן אנונימי את המשובים של חבריכם, וכן את המשוב שלנו.משימות:

1. התייחסו ל -8 כללי הזהב של שניידרמן (הוצגו בתרגול). כיצד המערכת שלכם מבטאת אותם? (5 נקודות)
2. יש להגיש את הטבלה הבאה , תוך התיחסות למשובים שקיבלתם (5 נקודות):

|  |  |  |
| --- | --- | --- |
| הערת משוב | האם התבצע שינוי באפליקציה בעקבות ההערה? | נימוק |
|  |  |  |
|  |  |  |
|  |  |  |
|  |  |  |

1. יש לרשום את ציון ה SUS של המערכת שלכם.מה מעיד הציון?(5 נקודות)
2. הגדירו 3 מדדים להצלחת המערכת (הרצאה 3).(5 נקודות)
3. הציגו דיאגרמת ארכיטקטורה של המערכת שלכם. הסבירו באיזה סוג ארכיטקטורה השתמשתם (הרצאה 7), ופרטו את חלקי הקוד ההמתיחסים לכל חלק בארכיטקטורה.(10 נקודות)

*חלק שלישי : פיצ'ר לבחירתכם* (10 נקודות)

הוסיפו פיצ'ר מעניין למערכת, אשר לא נדרש מכם, לבחירתכם.שימו לב - הכוונה לפיצ'ר פונקציונאלי (ולא עיצובי, או שימוש בשרת כדי להעלות לענן את האתר).

כתבו מספר משפטים להסבר התוספת, וציינו היכן בקוד הוא ממומש. כמו כן הסבירו כיצד הוא מתבטא בחלק המוצג למנהל.

**הוראות הגשה:**

1.ש להגיש במודל קובץ זיפ הכולל קובץ וורד ובו מענה לשאלות, וקישור ל- notebook ובו הקוד שלכם (יש לוודא שהקישור פומבי ונגיש). **אין לבצע שינויים במחברת לאחר ההגשה!**

**2.** הקוד צריך לרוץ במלואו מהמחברת בלבד. לא יתקבלו הגשות הכוללות הרצה באתר חיצוני (בפרט slack), או צורך להעלות קבצים למחברת על מנת שתרוץ. הגשות כאלו יקבלו ציון אפס על מרכיב הקוד

3.יש להגיש את התרגיל בצוותים, בתיקיית ה –GIT שלכם (צרפו קישור), וכן בתיקייית התרגיל ב moodle.כותרתו של הקובץ תהיה HW2\_TEAMNAME

4. שימו לב כי כל העבודות חייבות להיות שונות זו מזו. עבודות שייראו דומות ייפסלו ויינתן עליהן ציון 0.

בהצלחה!