

Findel

מגישים: דין מורביה ורן דוידאי

מנחה: ד"ר שירה צוקר

רקע: באינטרנט נמצאים מקורות מידע רבים. תלמידים הניגשים לחפש מידע עבור פרויקטים, שיעורים וידע חופשי, עשויים להיתקל לא פעם, במקורות מידע ואתרים שאינם אמינים, בהם קיים מידע מטעה. המערכת מציעה פתרון לבעיה על ידי חיפוש המשלב בין גוגל עבור מציאת אתרים, ויקיפדיה עבור מיפוי התוצאות לאתרים לימודיים, ובנתונים שהמשתמשים מזינים ישירות למערכת (דירוג אתרים, תגובות, הוספת קישורים לאתרים ותכנים).

תיאור המערכת:

דף הבית הינו דף החיפוש. לאחר הזנת ערך ולחיצה על חיפוש: המשתמש פונה לשרת, הבודק האם במסד נתונים קיימים אתרים, תגובות, ודירוגים עבור הערך המבוקש. אם אין נתונים, אזי השרת יבצע חיפוש בגוגל, ישמור את הערך והאתרים שקיבלנו מגוגל, וישלח למשתמש את האתרים. המכשיר של המשתמש, עליו פועלת אפליקציית React, יבדוק האם הערך קיים בוויקיפדיה. אם כן, יתחיל תהליך החיפוש: הוצאת העמוד והלינקים על הערך מויקיפדיה, web Scraping לאתרים ושליפת הטקסט מהם, חיפוש הלינקים בטקסטים בעזרת אלגוריתם rabin-karp*. בנוסף נחשב את jaccard similarity* בין כל העמוד ויקיפדיה לכל אתר ברשימה על מנת למצוא את האתרים הרלוונטיים ביותר. על פי התוצאות נמיין את האתרים, ונעזר במיין על מנת למצוא את התכנים הנוספים הנפוצים ביותר בשביל המשתמש. נשלח לשרת את התוצאות על מנת לשמור אותן, ולחסוך חישוב עודף עבור אותו תוכן. נציג למשתמש את התוצאות. אם קיימים נתונים על הערך ואם הינם עדכניים (על פי טווח זמן), אזי נשלח אותם למשתמש ביחד עם דירוג ותגובות המשתמשים האחרים. אם הנתונים אינם עדכניים, אזי נעשה חיפוש בדומה למקרה שאין נתונים, ונשלב את התוצאות ביחד עם הדירוג משתמשים.

אם אין תוכן ויקיפדיה, אזי נציג תוצאות גוגל ביחד עם דירוגי ותגובות המשתמשים.

אמצעים טכנולוגיים: צד שרת: node.js, צד לקוח: React, MongoDB, מוד נתונים: mongoDB.