



שם הפרויקט		מס' פרויקט
שימוש בנתוני עתק לשיפור התנובה - חקר מקרה גידול אבטיח לגרעינים		2022-01-171
מנחה שותף	מנחה אקדמי	
	פרופ' גלעד רביד	
חברי הצוות		
	עומרי גרנדה	
	omrigra@post.bgu.ac.il	

תקציר

במהלך הגידול החקלאי נאספים במערכות מידע שונות חלק מהנתונים הנוגעים לפעילויות שונות המבוצעות, כגון נתוני השקיה, בדיקות קרקע ומזיקים, דשנים ועוד. איסוף נתונים זה, בשילוב של נתונים חיצוניים, כדוגמת נתוני מזג אויר, משקעים, נתוני מערכות מידע גאוגרפיות הנוגעות לחלקות ונתונים נוספים, יכול לשמש בסיס לקבלת החלטות הנוגעות לגידול הנוכחי ולגידולים עתידיים. עבודות קודמות מצביעות על הפוטנציאל הגדול הגלום בניתוח נתוני עתק להפקת תובנות חקלאיות. על כן, שילוב של מספר בסיסי נתונים, ממספר מגדלים שונים, יאפשר להפיק תובנות חוצות מגדל לגבי אופי הגידול המיטבי. בגלל היקף הגידול וחוסר הידע בו, גידול אבטיח לגרעינים נבחר כמדגים לפרויקט זה ועליו נאספו הנתונים.

מטרת הפרויקט היא יצירת מנגנון לתהליכי ETL (Extract, Transform, Load) של נתוני גידול משתי מערכות ניהול גידולים נפוצות, וממערכות חיצוניות לבסיס נתונים אחד. על נתונים אלו ניתן יהיה בעתיד להריץ אלגוריתמי למידה שיפיקו תובנות לגבי אופי הגידול המיטבי. בשלב הראשון של הפרויקט, התקיימו פגישות עם גורמים מקצועיים לצורך הבנת הפרמטרים הנדרשים לאיסוף בכל חלקה. פרמטרים אלו יהיו המאפיינים של כל חלקה ובאמצעותם יהיה אפשר לבצע את תהליך הלמידה בעתיד.

לאחר מכן, התקיימו פגישות עם נציגים משני משקים שונים – משק "צבר קמה" ו"יד מרדכי". ממשקים אלו התקבלו נתונים ממגוון מקורות שונים - מערכות אוטומטיות לניהול המשק החקלאי, נתונים שנאספו בצורה ידנית ונתוני מעקב גידול חלקות. לנתונים אלו נוספו נתונים מהשירות המטאורולוגי ומהם יחד ניתן ליצור בסיס נתונים אחד ומגוון.

הקוד, שנכתב בשפת Python, לוקח כקלט את כל הנתונים הגולמיים שנאספו בשלב התכנון, ומתוכם מייצר את המאפיינים שהוגדרו לכל חלקה, על סמך הפרמטרים שנאספו בשלב איסוף הדרישות. בשלב כתיבת הקוד ניתן דגש רב על קוד גנרי, קריא וברור, וכזה שיתאים למספר מערכות לאיסוף נתונים. הפונקציות מחלצות מתוך הנתונים מאפיינים עבור כל חלקה בכל עונת גידול, בצורה אוטומטית כך שגם אם בעתיד יתווספו פעולות נוספות על החלקות, הפונקציה תוכל לחלץ גם פיצ'רים אלו. תוצרי הפרויקט העיקריים הם תהליכי עבודה וקוד הממיר את המידע ממערכות המגדל לבסיס נתונים אחד המשלב מידע מטאורולוגי, הכולל נתונים משני משקים על פני ארבע שנים.