

הצעת פרויקט סוף קורס למידת מכונה 2025

שם המבצע:

ליאור בן ארי

מספר תעודת זהות:

208649145

שם המאמר עליו יבוסס הפרויקט:

Global Soil Salinity Prediction by Open Soil Vis-NIR Spectral Library

Zhou, Y. et al. (2022). *Global Soil Salinity Prediction by Open Soil Vis-NIR Spectral Library*. Remote Sensing, 14(21), 5627. <https://doi.org/10.3390/rs14215627>

נושא הפרויקט:

חיזוי מליחות בקרקע באמצעות Machine Learning בהתבסס על מאגר נתונים המכיל ספקטרוסקופיה בתחומי האור הנראה ואינפורה-אדום (Vis-NIR).

תיאור הפעולות אותן אני מתכנן לבצע:

במסגרת הפרויקט אני איישם את המחקר המתואר במאמר תוך שימוש ב-Machine Learning:

- איסוף נתונים מ-database: אני אשתמש במאגר נתונים בשם Open Soil Spectral Library (OSSL). המאגר מכיל דגימות קרקע רבות ובנוסף מדידות ספקטרוסקופיה בתחום 350nm-2500nm וערכי מוליכות חשמלית המייצגים את מליחות הקרקע (EC).
 - אבצע שימוש בשיטת ה-pre-processing עבורה לפי המאמר התקבלו תוצאות החיזוי הטובות ביותר. בנוסף, במידת הצורך ייבחנו שיטות pre-processing נוספות.
 - חיזוי באמצעות למידת מכונה:
- את חיזוי המליחות בקרקע אבצע באמצעות ארכיטקטורות Machine Learning בהתבסס על אלו שנבחנו במאמר:

- Random Forest
- PLSR
- Cubist
- XGBoost

בנוסף, ייבחנו השפעות של features נוספים מתוך ה-database כפי שתואר במאמר כדי לבדוק האם יש שיפור בתוצאות החיזוי וכן השפעה של שימוש ב-principal components על ידי PCA.

- אבצע חלוקה לסט אימון ובדיקה ואשתמש ב-cross validation או randomized search לטיוב פרמטרי המודלים ולבדיקת תוצאות החיזוי.
- אשתמש במדדים כמו R^2 ו-RMSE לבחינת תוצאות החיזוי.
- אבצע השוואה בין התוצאות אותן אקבל לבין התוצאות שהתקבלו במאמר.