מאמר ביולוגיה – דברים שצריך לחקור:

**Automated prediction of HIV drug resistance from genotype data**

להבין את המשפט:

המשפט:  
emergence of drug resistance mutations diminishes the effectiveness of drug therapy for HIV/AIDS  
  
הסבר: הופעתן של מוטציות עמידות לתרופות מפחיתה את היעילות של טיפול תרופתי ב-HIV/AIDS

המשפט: A unified encoding of protein sequence and structure was used as the feature vector for predicting

הסבר: קידוד מאוחד של רצף ומבנה חלבון שימש כווקטור התכונה לחיזוי

genotypic – הסבר: הרצף DNA – רצף של AGCT - **מציין את ההרכב הגנטי של יצור מסוים**

Phenotypic – הסבר: תצוגה ויזואלית של DNA - את התבטאותן של תכונות היצור  
דוגמא לשניהם: לדוגמה, "עיניים כחולות"; זהו תיאור של פנוטיפ. הגנוטיפ של אותו אדם (בהקשר לצבע עיניו) הוא מכלול הגנים המשתתפים בקביעת צבע העיניים.

Random Forest – אלגוריתם לבחירה יעילה של דאטה – כלומר נרכיב מספר מסויים של עצים – כאשר לכל אחד נזריק דאטה שונה – או מספר רנדומלי שונה – לאחר מכן נשאל באמצעות קובץ הtest שלנו את השאלות – ומי שצדק בכמות המקרים הגדולה ביותר יהיה המודל שלנו – העץ שלנו.  
  
Unified sequence\structure encoding – הסבר:  
  
PR – הסבר:  
  
RT – הסבר:   
  
Random Forest – הסבר:  
  
  
לבדוק את האתר: <http://apollo.cs.gsu.edu/~bshen/html/index.html>.

ההבדל בין regression לבין classification – הסבר:  
  
novelty – הסבר: שינוי?  
  
amino acid – הסבר: חומצת אמינו – 3 DNA – אמינו אחד - רצף אמינו זה חלבון