**פרויקט סיום הולכה חשמלית בתאים**

דור מסאס 318218625, אברהם חרינביצקי 313133324

ניתוח דאטה:

1. תחילה כתבנו קוד שעובר על כל קבצי ה-.hea הנתונים ומוצא את שבעת האבחונים הכי נפוצים – אדם בריא ו-6 המחלות הנפוצות ביותר.
2. לאחר שיש לנו את המידע נחלק את הנתונים לתיקיות train, validation, test לפי יחס 70-15-15. בשלב זה נרצה לשמור על פילוג אחיד של המידע, לכן נחלק כל אחד מהקבצים באותו פילוג לכל תיקייה, תוך כדי שמירה על כמות מספקת של נתונים עבור על אבחון ומבלי ליצור bias. על מנת לספק זאת נגביל את מספר הקבצים הקיימים מכל אבחון ל-7500 קבצים לכל היותר.
3. מאחר ויכולים להיות מספר אבחונים לכל סיגנל נתייג את הדאטה באמצעות תגית one-hot, כלומר מילון/וקטור המייצג את האבחונים המתאימים לסיגנל, כך שהרשת תוכל ללמוד תוך התייחסות לכל התסמינים המתאימים במקביל.