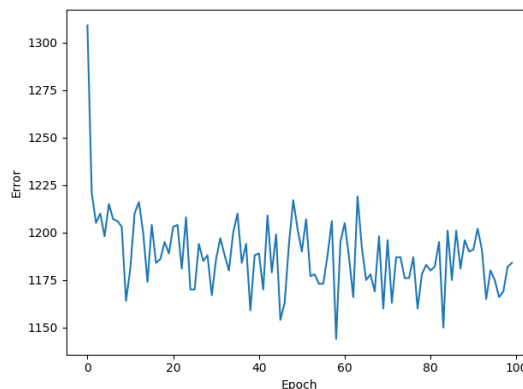


למידת מכונה תרגיל 2 – דו"ח

על מנת לבחור את הפרמטרים אשר יביאו לצמצום אחוזי השגיאה ביצעתי תחילה חלוקה של המידע ל-training set (כ-80%) ו-validation set (כ-20%).

לאחר הצבה של מספר ערכים רנדומליים שמתי לב ששינוי פרמטר ETA אינו מביא לשינוי באחוזי הדיוק במסווגים Preceptron ו-PA. ב-SVM העלתו גרמה לירידה באחוזי הדיוק משום כך בחרתי להשאיר את ערכו בכולם לערך ההתחלתי של 0.01.

לאחר מכן ניסיתי לשפר את אחוזי הדיוק ע"י בחירת הפרמטר epochs המתאים לכל מסווג. תחילה ניסיתי למצואו ע"י גרף על סט האימונים המתאר את הקשר בין אחוז השגיאה לepoch כדי למצוא נקודה מינימלית לפני overfitting. אולם הגרף שיצא לא הציג נקודה כזאת אלא עליה וירידה חוזרת של מספר השגיאות.



לכן ניסיתי דרך שונה חישובתי בלולאה את אחוז הדיוק עבור כל epoch ולקחתי את אילו שנתנו את האחוזים הכי גבוהים, יצרתי עוד סטי אימון וולדיציה כדי לבדוק את התוצאות מהסט הראשון ובחרתי בסוף את התוצאות שהניבו אחוזים גבוהים בכלל הסטים.

את הלמבדה חישובתי באופן דומה אולם שמתי לב שהעלת ערכה גורמת לירידה באחוזי הדיוק ולכן ההרצה הייתה בסדר יורד.