

דוגמאות לריצה ללא התחשבות בsocial burden:

צבע כחול: data לפני תזוזה

צבע אדום: data אחרי תזוזה

+: נקודות עם תיוג 1

-: נקודות עם תיוג -1

כל אחת מהמטריצות 2×2 היא confusion matrix.

Y : תיוג אמיתי (ground truth)

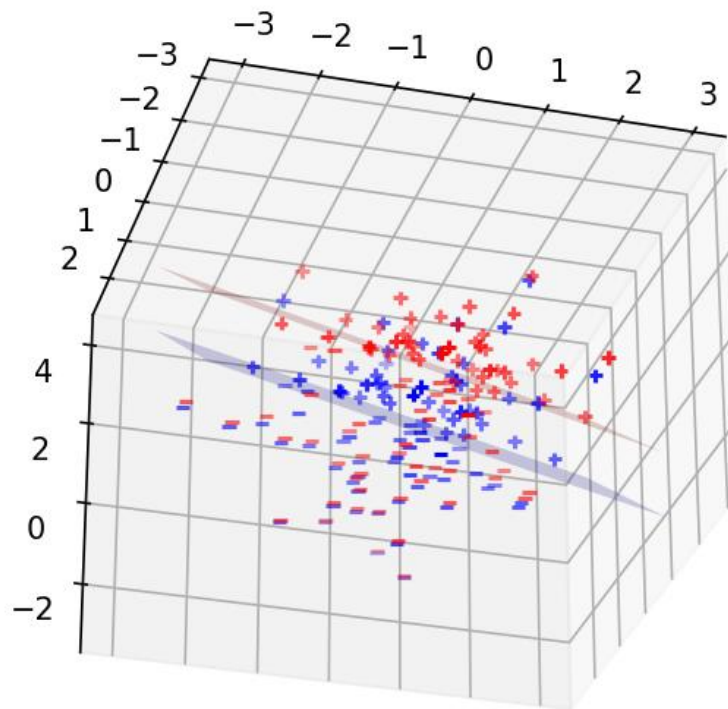
$F(x)$: למידה לא אסטרטגית על שחקנים לא אסטרטגים (מישור כחול בגרף)

$F(x')$: למידה לא אסטרטגית על שחקנים אסטרטגים

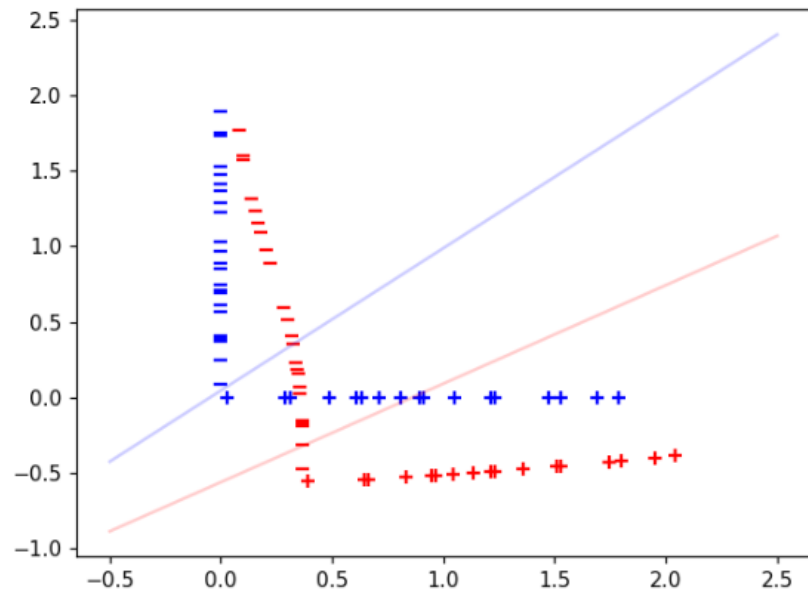
$F'(x)$: למידה אסטרטגית על שחקנים לא אסטרטגים

$F'(x')$: למידה אסטרטגית על שחקנים אסטרטגים (מישור אדום בגרף)

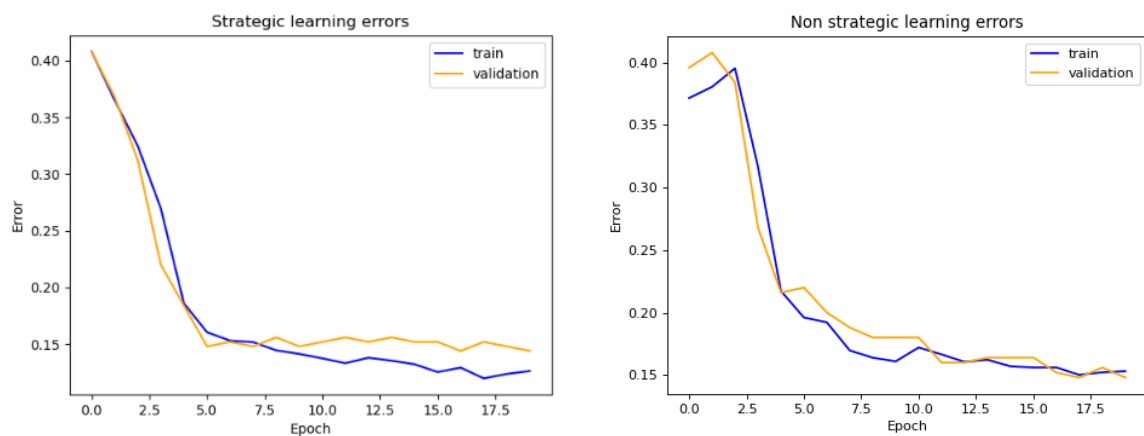
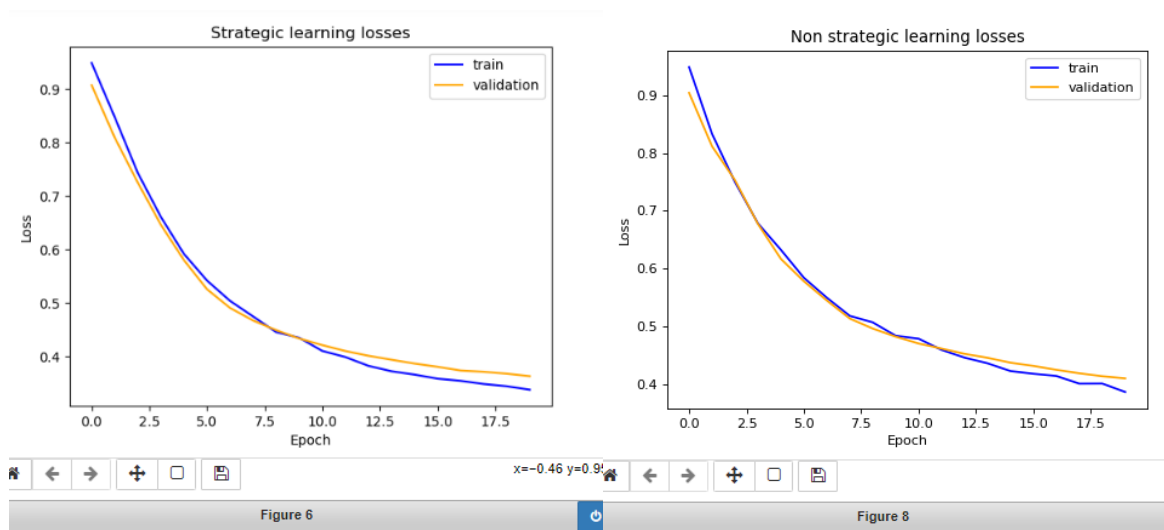
```
y vs f(x):  
[[55.  0.]  
 [ 0. 45.]]  
accuracy: 100.0  
y vs f(x'):  
[[37. 18.]  
 [ 0. 45.]]  
accuracy: 82.0  
y vs f'(x):  
[[55.  0.]  
 [34. 11.]]  
accuracy: 66.0  
y vs f'(x'):  
[[55.  0.]  
 [ 8. 37.]]  
accuracy: 92.0
```



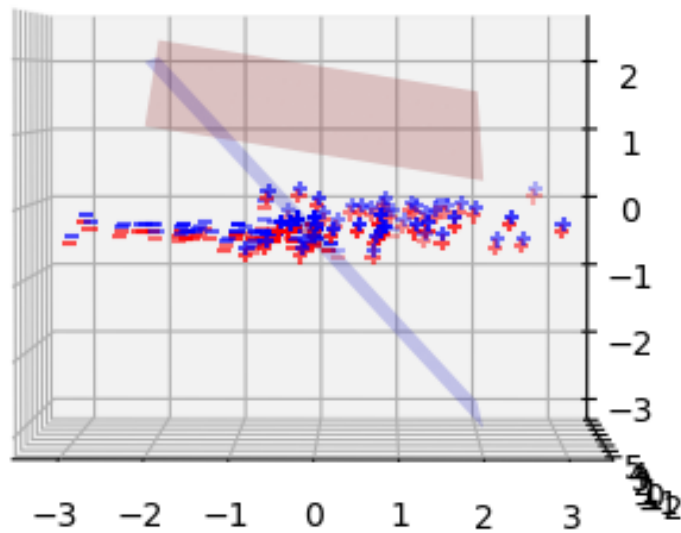
y vs $f(x)$:
 $\begin{bmatrix} 57.5 & 0. \\ 0. & 42.5 \end{bmatrix}$
 accuracy: 100.0
 y vs $f(x')$:
 $\begin{bmatrix} 32.5 & 25. \\ 0. & 42.5 \end{bmatrix}$
 accuracy: 75.0
 y vs $f'(x)$:
 $\begin{bmatrix} 57.5 & 0. \\ 20. & 22.5 \end{bmatrix}$
 accuracy: 80.0
 y vs $f'(x')$:
 $\begin{bmatrix} 55. & 2.5 \\ 0. & 42.5 \end{bmatrix}$
 accuracy: 97.5



עקומת למידה (train&validation losses and errors)

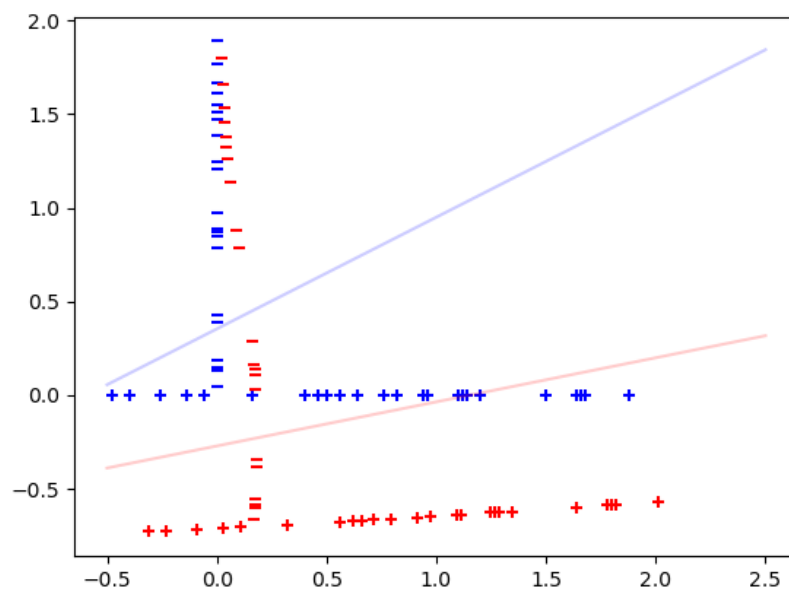


דוגמה לריצה עם איבר רגולריזציה ענק ($\lambda = 100$):



המישור פשוט מרחף מעל כל הנקודות ומסווה את כולם כדי שה *burden* יהיה 0.

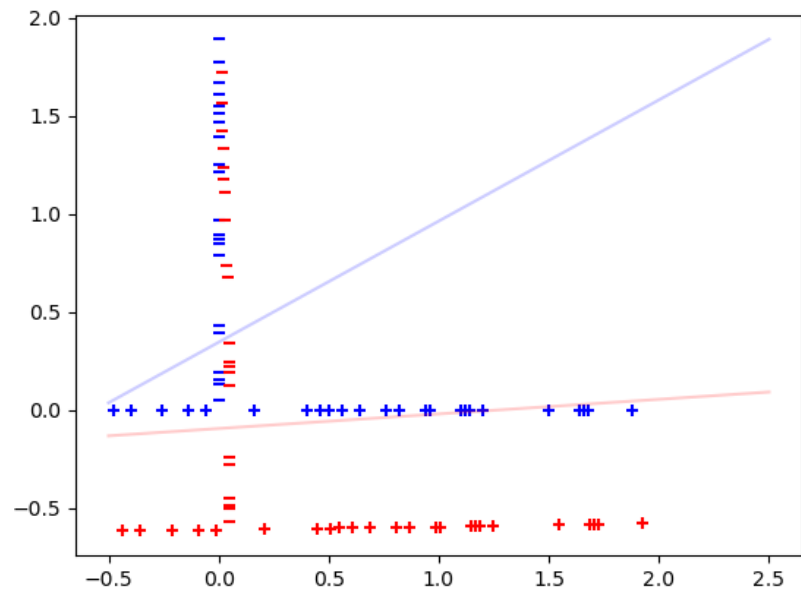
דוגמה לשיפור משמעותי ב *social burden* בתמורה לירידה לא משמעותית ב *accuracy*:



$$\lambda = 1$$

accuracy = 86.66 %

average social burden = 0.105



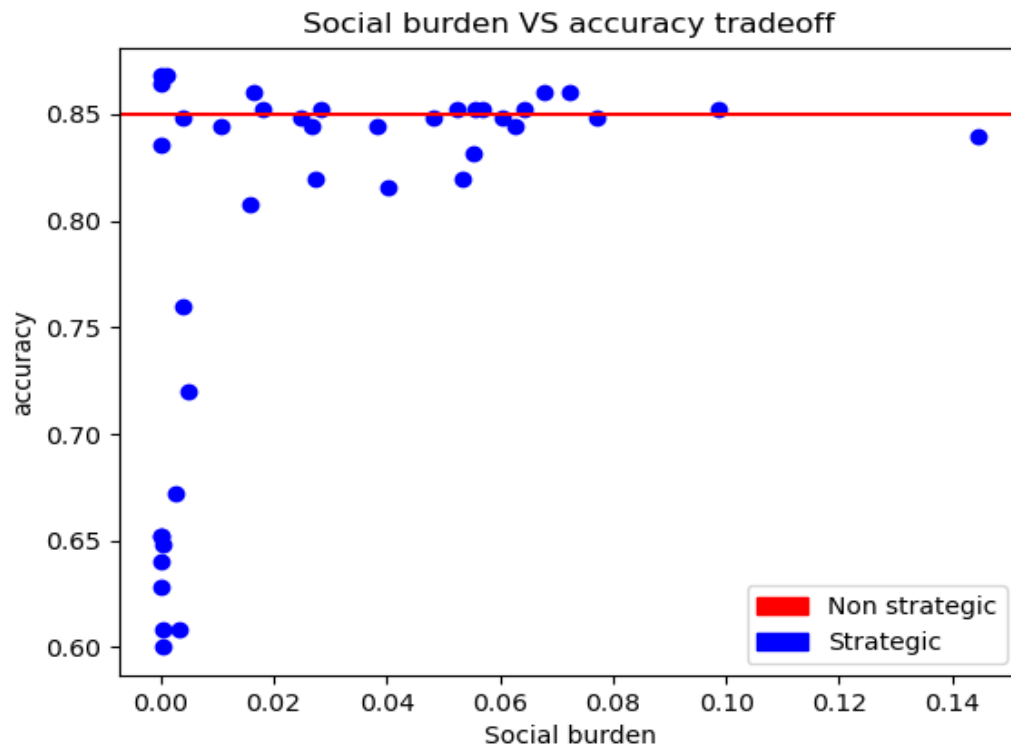
$\lambda = 3$

accuracy = 86.66 %

average social burden = 0.003

:Social burden VS accuracy tradeoff

logscale of λ : $0 \leq \lambda \leq 20$



:Utility VS accuracy tradeoff

logscale of λ : $0 \leq \lambda \leq 20$

