פגישה עם הצבא

1. משתנים שאי אפשר להשתמש – combat exposure, exposure unit. כל השקובית של MilitaryExposure, Brigade
2. משתנים שאפשר – דמוגרפיים, VAR00011 – סכימה של שמונה, אפשר להשתמש במדדים של אשכנזיות או לא.
3. גנטיקה – אפשר להשתמש בשביל המאמר לא בשביל הצבא
4. בלי אינטרקציות
5. שאלונים –
6. לא להשתמש PTGI2
7. ליצור הבחנה ברורה בין ניבוי מזמן 1 לזמן 3
8. בלי זמן 2 בכלל
9. בלי zero
10. דוט פרוב רק זמן אחד –
11. לצוק איתן אין נתונים גנטיים, אולי חלק מהשאלונים חסרים.
12. נבדוק מה הם הפיצרים הכי תורמים לניבוי. נבדוק האם יש פיצר קריטי.
13. **רוצים לייצר מודל מפיצרים רזים** – בלי זמן 2, בלי גנטיקה. מכאן אפשר רק לרדת
14. זמן 2 – מוסיפים לו את זמן 2 – אפשר לעשות אנג'נירים, אפשר להוסיף את המדידה הנוספת, אפשר להוסיף את הדלטה, בלי אינטרקציות,
15. לא להתחייס לארועי לחימה בזמן הטירונות.
16. לא להשתמש ב- PTGI2
17. טווח של PCL – 40 וגם strict.
18. מאמר – בשם analysis -
19. מודל מספר 3 – full feature model
20. מודל מספר 1 – full features model from time 1
21. מודל מספר 2 – MVP
22. כל מודל אופטימיזציה משלו
23. דו"ח קצר לגבי ה-dropout - מבוא של חמש שורות -שאלת המחקר כאן הייתה האם ניתן לנבא עם משתנים שנאספים בשבוע הראשון של הטירונות מי לא יסיים אותה – רק אותם נתונים של פיצרים רזים.
24. אח"כ יש שיטה – לכתוב מסודר את השיטה
25. ואז תוצאות – לבדוק את ה-feature importance . לבדוק אם אפשר להעיף חלק מהשאלונים.
26. מסקנה – אנחנו מסוגלים לנבא מתוך כל 100 חיילים שימלאו את השאלונים האלה והאלה להצביע על 11 לצדוק ב-10 ולטעות ב-3