**תרגול רשתות נוירונים (לפי הסרטון של StatQuest)**

לדוגמה יש לנו תרופה שמרפאה מחלה, במינון גבוה ונמוך התרופה לא הייתה אפקטיבית (0) במינון בינוני היא כן (1).

\*אפשר להשתמש בתרגיל בית 3 ברשת נוירונים (יש בsklearn) כדי להגיע לדיוק הרבה יותר גבוה.

רשת נוירונים יכולה להתאים את עצמה לפרבולות ולא רק קוים ישרים. רשתות נוירונים בנויות ממספר שכבות כאשר השכבה הראשונה מקבלת את הקלט והשכבה האחרונה מוציאה את הפלט.

כאשר מבצעים סיווג אז מקבלים כפלט את ההסתברות של תצפית להיות שייכת לכל קלאס, נשייך את התצפית לקלאס עם ההסתברות הכי גבוהה.אם זה רגרסיה נקבל מספר.

על הקשרים בין הצמתים יש מספרים אשר מהווים את השיפוע ax+b בין הצמתים.

הצמתים השונים יכולים לבצע פונקציות אקטיבציה שונות על הנתונים:

Softplus,relius,sigmoid- ניתן להגדיר בהיפר פרמטרים. לכל צומת יכול להיות פונקציית אקטיבציה שונה. מנסים פונקציות אקטבציה שונות עד שמקבלים את רמת הדיוק הכי גבוהה.

במקרה שלנו, המינון של התרופה הוא בין 0 ל1. נתחיל מ0 בצומת הראשון, מחשבים את ערך שמתקבל מהמעבר לצומת הבא ומוסיפים את ערך הbias, מפעילים את הפונקציה של הצומת הבא, במקרה זה softplus. התוצאה שקיבלנו היא 2.25. כרגע זה החיזוי שקיבלנו עבור מינון שהוא 0.

כאשר נשים תצפית אחרת(מינון=0.1) נבצע את אותו תהליך ונקבל ערך אחר. נעבור כך על כל התצפיות ונקבל עקומה לכמה התרופה אפקטיבית במינונים שונים. לא בהכרח משתמשים כל הערכים של פונקציית האקטיבציה, אלא רק בחלק, לכן ניתן ליצור מספר עקומות שונות ולפי החלק שמשתמשים בפונקציית אקטיבציה. מכפילים במשקל של הקשת הבאה לצומת הבא ומקבלים עקומה חדשה. בנוסף שולחים את כל התצפיות לnode אחר ומשתמשים במשקלים אחרים מה שנותן עקומה אחרת. סוכמים את העקומה של 2 הצמתים ומקבלים פרבולה אשר חוזה את יעילות התרופה למינונים שונים.

**שאלות של רשת נוירונים:**

סיווג:

שכבת הקלט- כמות הנוירונים היא כמות הפיצ'רים שמתקבלים

שכבת הפלט -כמות הנוירונים היא מספר המחלקות

רגרסיה:

שכבת הקלט- כמות הנוירונים היא כמות הפיצ'רים שמתקבלים

שכבת הפלט- נוירון אחד שנותן את התוצאה.

**חזרה למבחן**

מבחן עם חומר פתוח, אפשר מחשבון

חלק מהשאלות במבחנים קודמים במודל לא רלוונטיות- החלקים של בינה עסקית (החלק הראשון עד 2018)

יש 5-6 בכל מבחן, במבחן יש בונוס בדרך כלל

המבחן הוא שעתיים וחצי

יש שאלות תאוריה -מה זה supervised, unsupervised, clustering, למה צריך סט מבחן, איך בונים מודל וכו' (10 נקודות לשאלה)

מטריצות בלבול- לחשב accuracy, specificity ל2 קלאסים או יותר

תהליך בניית מודל- מקבלים דטא, להסביר על pre processing, להסביר מה צריך לעשות מבחינת הכנה של נתונים לכל אלגוריתם, איך כל אלגוריתם עובד מבחינה רעיונית.

**בעץ חייבים לעשות השלמה של ערכים חסרים**

**KNN עובד יותר טוב עם scaling**

קיבלנו אחוז דיוק גבוה בtrain ואחוז דיוק נמוך בtest- יש overfitting, להסביר מה אפשר לעשות.

**Loss functions**

בSVM רוצים שהמרחק ללא נתונים יהיה כמה שיותר גדול- מינימום גלובלי

בK means -סכום המרחקים בריבוע בין התצפיות בקלאסטר למרכז הקלאסטר

בtSNE- ההסתברות בין זוג נקודות בממד הגבוה לבין ההסתברות בממד הנמוך צריכה להיות כמה שיותר דומה- מינימום גלובלי

עצים- ממזערים entropy או gini

SVD/PCA- הייצוג בעזרת האלגוריתם יהיה כמה ישותר קרוב לדטא המקורי- יתפוס כמה שיותר שונות

**שאלה שחוזרת על עצמה**

Adaboost,random forest , bagging- מה משותף לשלושת האלגוריתמים: הם super classifiers- משתמשים בהרבה weak learns ומבצעים אגריגציה ביניהם כדי לקבל תוצאות יותר מדויקות.

למה כדי להשתמש באחד מאלו ולא עץ החלטה?

הם עושים אגריגרציה של כמה weak learns ולכן מקבלים תוצאה יותר מדויקת בדרך כלל מאשר עץ אחד.

Adaboost משתמש בweak learns הקודמים כדי להגיע לתוצאה עם הlearner הנוכחי, ה2 האחרים

Adaboost וbagging יכולים להשתמש בweak learns שהם לא עצים

Random forest יכול להשתמש רק בעצים

**איך ממשים bagging עם svm?**

לוקחים dataset ועושים עליו svm

לוקחים dataset אחר ועושים עליו svm..

לפי הרוב\ ממוצע ביניהם בוחרים את התוצאה

**אילו בעיות יכולות להיות כאשר משתמשים בrandom forest עם svm?**

Random forest עובד רק על עצים

כאשר עושים הגרלה על הפיצ'רים בrandom forest אז מקבלים רק חלק מהפיצ'רים כל פעם, ואז נקבל עצים אשר כל פעם מחלקים באמצעות svm לפי פיצ'רים שונים

ניתן לפתור את זה באמצעות שימוש בadaboost במקום ואז כל עץ יחלק עם svm לפי פיצ'ר אחד.

**אילו ביעות יהיו לשימוש בadaboost עם svm?**

Adaboost נותן משקלים לכל weak learner ובsvm אין משקלים

ניתן לפתור את זה באמצעות לתת משקל לכל נקודה ואז להשתמש בsvm.