בינה מלאכותית 236501

תגביל בית 3

שאלה 2:

הוכחה , האינטואציה היא ש ID3 מסתמך על פונקצית IG שהיא רק מחושבת עבור פיצר ספיצי ולכן לא משנה אם מנרמלים פיציר זו או לא , אם אני אוכיח שאני אקבל אותו פיצול לצומת לא משנה אם ה פיצר מנורמל או לא אז סיימתי כי זה אומר שיש לי אותו ערך של IG.

באופן יותר פורמלי:

תהי פיצר רציפה כלשהיא ,

שאלה 5:

א)נגדיר את מסווג המטרה הבא לכל :

נגדיר את קבוצת הדוגמאות הבאה:

נשים לב שעץ ID3 יבחר קודם כל לפצל לפי (הוא יבחר את הממוצע בין -1 ו 1 ל 0 שזה 0 ויקבל שני צמתים סופיים), כי פיצול זה יניב אי וודאות של 0 , לעומת זאת אם הוא החליט לפצל לפי יקח את הממוצע שזה 4 ויקבל אי וודאות יותר מ 0 (כי ישאר לו צומת עם שני בנים עם סיווגים שונים (0,6) ו (1,4)).אז יחליט לפצל לפי  *עם סף 0 שזה בדיוק פונקצית המסווג לכן הוא תמיד יפגע.*

יהא  *כלשהו , נחלק למקרים:  
אם , ניקח את הנקודה (1.00000000001, 4) אז KNN יניב תוצאה + כי (1,4) היא הנקודה הכי קרובה עליה , אבל זו טעות כי הסיווג האמיתי הוא -.*

*אם : ניקח (2,0) , KNN תמיד יחזיר - כי הרוב הוא - אבל בפועל היה צריך להחזיר + כי לכן גם פה יטעה.   
(הראינו שלכל K , קיימת דוגמה שעבורה KNN יטעה והוכחנו ש זהה לפונקצית המטרה).  
\*נקודה אחרונה: הנחתי שלא צריך להתיחס למקרים בהם k זוגיים כי זה פשוט לא כזה חכם.*

ב)נגדיר את מסווג המטרה הבא לכל :

(אנטואיציה לפונקציה: אם הנקודה יותר קרובה ל (1,1) מאשר (2,2) אז תחזיר 0 , אחרת תחזיר 1)

נגדיר את קבוצת הדוגמאות הבאה:

לגבי ID3 : לא משנה עבור איזו פיצר הוא יפצל הוא יקבל אותם תוצאת שזה שני צמתים ובכל צומת דגימה בודדת (כי הערכים של הפיצרים השונים שווים) *לכן הוא יבחר לפצל לפי עם סף 1.5 (כי האינקס שלו יותר גדול). אז הוא יטעה עבור דוגמה (0,2), כי הוא יסווג אותה כחיובית אבל בפועל היא שלילית כי היא יותר קרובה לדוגמא (1,1).  
מרחק בין (1,1) ל (2,0) (השורש לא משנה):*

*מרחק בין (2,2) ל (2,0) (השורש לא משנה):*

*לכן היסווג האמיתי של נקודה (0,2) הוא שלילת כי היא יותר קרובה ל (1,1) אבל הוא יסווג אותה כחיובית היא ערך ה שלה יותר גדול מ 1.5.*

*לגבי KNN: נראה כי KNN תמיד פוגע עבור , לכל נקודה K מחזיר סיווג שלילי אם הנקודה יותר קרובה ל (1,1) מאשר (2,2)אחרת הוא יחזיר סיווג חיובי. זה בדיוק מה פונקצית מסווג המטרה עושה.*

ג)נגדיר את מסווג המטרה הבא לכל :

נגדיר את קבוצת הדוגמאות הבאה:

*לגבי ID3 : הוא יסווג את כל הדוגמאות שליליות בפרט עבור אבל זה לא נכון כי עבור דוגמא זו הסיווג האמיתי שלה הוא חיובי כי .*

*דוגמה שבה יכשל KNN: גם (1,1) , כי KNN יסווג את כל הדוגמאות כשליליות בגלל שהרוב שקרוב לו (רק אחד) הוא שלילי , אבל סיווג דוגמא זו הוא חיוב ולכן גם הוא יטעה.*

*ד)*נגדיר את מסווג המטרה הבא לכל :

נגדיר את קבוצת הדוגמאות הבאה:

*לגבי ID3 הוא יפצל לפי עם סף 0. כי עבורו האי וודאות היא 0 , לעומת זה אם הוא פיצל לפי עם סף 0 אז הוא ישאר עם אותה צומת כלומר אי וודאות יותר מ 0 , ולכן הוא יבחר לפצל עבור עם סף 0. לכן הוא יהיה בדיוק כמו פונקצית מסווג המטרה ולא יטעה אף פעם.*

*לגבי KNN: ניקח , ונראה שלכל מתקיים שהסיווג של KNN נכון.*

*עבור :  
נראה שהמרחק בין (-1,0) לבין (v1,v2) יותר קטן מהמרחק בין (1,0) לבינו , אם הוכחנו את זה אז זה מראה ש 1NN תמיד פוגע נכון עבור דוגמאות שהסיווג האמיתי שלהם הוא חיובי.*

*לכן עבור , 1NN תמיד יבחר אותו כחיובי כי תמיד הוא יהיה יותר קרוב לוקטור  
(-1,0).   
המקרה הוא דואלי עבור ולכן גם פונקצית ה 1NN תהיה שווה לפונקצית מסווג המטרה.*