**פרק 0 : מבוא.**

קרא מתוך ISLR את פרק 2, עמודים 15-42, וענה על השאלות הבאות:

1. הסבר מהן הגישות הפרמטריות והא-פרמטריות בלמידה סטטיסטית.

מהן היתרונות והחסרונות של כל גישה?

בלמידה סטטיסטית פרמטרית, צורת המודל מצוינת אפריורית עם ידע מוקדם. גישה זו כוללת שני שלבים עיקריים:

נניח צורה פונקציונלית, כלומר נחליט על מודל מתמטי לתיאור הקשר בין מנבאים למשתנה היעד, כגון רגרסיה לינארית, רגרסיה לוגיסטית וכו'.

שלב שני נלמד פרמטרים של המודל (למשל, מקדמים, היפר פרמטרים) מהנתונים. היתרונות של פרמטרית הם פשטות מודלים פרמטריים פשוטים יותר להבנה ולפרשה, במיוחד בממדים נמוכים.

ביעילות הם דורשים פחות נקודות נתונים כדי להעריך את הפרמטרים מכיוון שהמודל מוגדר על ידי מספר סופי של פרמטרים ומבחינת חישוב אימון מודלים פרמטריים הוא פחות אינטנסיבי מבחינה חישובית בשל המבנה המוגדר מראש שלהם.

החסרונות של גישה זו הם מפרט שגוי אם המודל הנבחר אינו תופס כראוי את הקשר האמיתי, זה יכול להוביל לתוצאות מוטות (למשל, שימוש במודל ליניארי עבור נתונים לא ליניאריים) והסתמכות יתר על הנחות דיוק התוצאות תלוי במידה רבה בתקפותן של הנחות המודל.

לעומת זאת גישה לא פרמטריות אינה מניחה צורה פונקציונלית ספציפית של הקשר בין מנבאים למשתנה היעד. במקום זאת, שיטות אלו שואפות ללמוד את המבנה הבסיסי ישירות מהנתונים.

למשל K-Nearest Neighbors (KNN) התחזיות נעשות על סמך השכנים הקרובים ביותר במערך הנתונים ועצי החלטה.

היתרונות הם גמישות שיטות לא פרמטריות יכולות לדגמן קשרים מורכבים ושרירותיים יותר בנתונים ושיטות אלו מנחות פחות הנחות לגבי התפלגות הנתונים או הצורה הפונקציונלית של הקשר.

החסרונות של גישה א-פרמטרית הם בכך שמודלים לא פרמטריים דורשים כמויות גדולות של נתונים כדי לבצע חיזויים מדויקים מכיוון שהם אינם מסתמכים על מבנים מוגדרים ובשל הגמישות שלהם, הם נוטים יותר להתאמת יתר, במיוחד עם נתונים רועשים.

1. מהם היתרונות והחסרונות של שיטות גמישות (flexible methods) בלמידת מכונה?

שיטות גמישות בלמידת מכונה מתייחסות למודלים שיכולים להתאים את המבנה שלהם כך שיתאים יותר לנתונים, וללכוד קשרים ודפוסים מורכבים. דוגמאות כוללות עצי החלטה, רשתות עצביות ושיטות אנסמבל כמו יערות אקראיים והגברת גרדיאנט.

היתרונות של שיטות גמישות יכולות ללכוד קשרים מורכבים ולא ליניאריים בנתונים, מה שהופך אותם מתאימים לבעיות שבהן מודלים פשוטים יותר נכשלים.

על ידי התאמת הנתונים באופן הדוק יותר, שיטות אלו מניבות לעתים קרובות ביצועי חיזוי גבוהים יותר, במיוחד במשימות כמו זיהוי תמונות, עיבוד שפה טבעית ותחומים מורכבים אחרים.

לבסוף, הם יכולים לטפל בסוגים שונים של נתונים (קטגוריים, מספריים, טקסט, תמונות) ולהתאים את המבנה שלהם על סמך מערך הנתונים.

החסרונות של שיטות גמישות בשל יכולתן להתאים באופן הדוק לנתונים, שיטות גמישות נוטות להתאמת יתר, במיוחד עם מערכי נתונים קטנים או רועשים. זה מפחית את ביצועי ההכללה שלהם על נתונים בלתי נראים.

שיטות אלה דורשות לעתים קרובות כמויות גדולות של נתונים כדי לבצע ביצועים טובים. ללא מספיק נתונים, הם עלולים לא ללכוד דפוסים ויחסים אמיתיים.

בנוסף, שיטות גמישות רבות כוללות כוונון של היפרפרמטרים רבים (למשל, מספר שכבות ברשת עצבית, עומק של עץ החלטות), מה שעלול לקחת זמן ומורכב ושיטות גמישות יכולות להיות רגישות לרעש ולחריגים במערך הנתונים, מה שעלול להוביל לירידה בביצועים.

1. הסבר בפירוט מהו ה-bias-variance tradeoff תוך מתן דוגמא.

ה-bias-variance tradeoff הוא מושג בסיסי בלמידה סטטיסטית ולמידת מכונה המתארת ​​את הפשרה בין היכולת של מודל ללכוד במדויק דפוסים בנתונים (הטיה נמוכה) לבין יכולתו להכליל לנתונים בלתי נראים (שונות נמוכה). איזון ההטיה והשונות הוא המפתח להשגת ביצועי ניבוי מיטביים.

לדוגמה, רגרסיה פולינומית. התאמת מודל לניבוי משתנה יעד

y מתוך תכונה x, כאשר הקשר האמיתי אינו ליניארי.

הטיה גבוהה (לא התאמה):

מודל ליניארי (y=ax+b) פשוט מכדי לתפוס את הקשר האמיתי.

למודל יש שגיאה שיטתית (הטיה גבוהה) אך רגישות נמוכה לנתוני האימון (שונות נמוכה).

שונות גבוהה (התאמת יתר): פולינום מדרגה 10 מתאים לכל נקודת נתונים בצורה מושלמת, כולל רעש.

המודל לוכד רעש בנתוני האימון (שונות גבוהה) אך אינו מצליח להכליל.

דגם מאוזן הוא פולינום מדרגה שלישית לוכד את המגמה הכללית מבלי להתאים רעש, משיג איזון טוב בין הטיה לשונות.