**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?

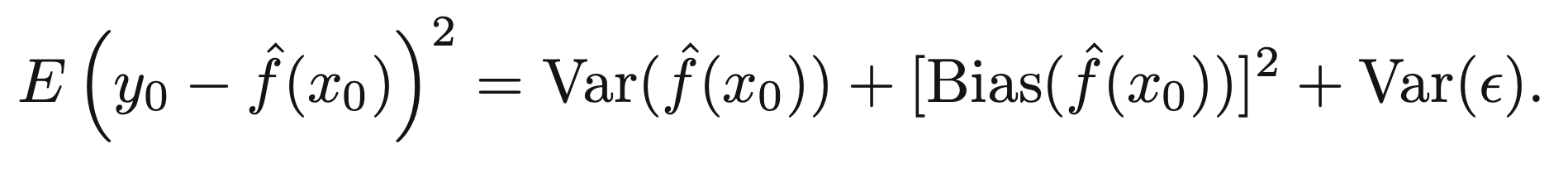
הגישה הפרמטרית באה לשערך את פונקציית f בעזרת שיערוך פרמטרים, לדוגמא W\*X + b , כאן צריך לשערך את W ואת b. היתרון שלה הוא שכדי לשערך את f צריך רק למצוא את הפרמטרים הנכונים שמקטינים את ה loss הרלוונטי. החסרון בה הוא שאנו מניחים את המבנה של פונקציית f, לא בטוח שהמבנה שבחרנו הוא המבנה הנכון אם יש בכלל מבנה.

הגישה הא-פרמטרית לא מניחה כזה מבנה, היא פשוט מנסה לשערך את פונקציית f הכי טוב שאפשר. היתרון שלה הוא שהיא לא מניחה מבנה מסויים מה שיכול לשפר את התוצאות של ה predictor, החסרון שלה הוא שהיא בדרך כלל דורשת המון דאטא.

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

היתרונות של גישה גמישה היא שהגישה מקטינה את השגיאה של ה training data ככל שהגישה יותר גמישה, העניין הוא שזה לא בהכרח מקטין את השגיאה על הטסט, ולפעמים אם נשתמש בשיטה גמישה מדי השגיאה על הטסט אף עלולה לעלות. במילים אחרות יש חשש ל overfitting.

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



פורמלית, הטריידאוף בין שונות להטייה מתייחס לפירוק של התוחלת בריבוע של השגיאה של הפונקצייה ששיערכנו f, אם נסתכל על הנוסחא למעלה, נראה שכדי למזער את השגיאה, נרצה למזער את גם את השונות וגם את ההטייה. הבעיה היא שכשאנחנו מקטינים אחד השני עולה, זאת אומרת שאנחנו רוצים למצוא איזשהו אמצע בניהם שימזער באופן כולל את השגיאה הכוללת.

במילים יותר פשוטות, ככל שהמודל יותר גמיש הוא יוצר יותר שונות אבל מקטין את ההטייה (overfitting), לעומת זאת כשהמודל פחות גמיש השונות יותר קטנה אבל ההטייה יותר גדולה (underfitting).