# תשובות – מבוא

1. גישות פרמטריות מניחות שמבנה הפונקציה שנרצה לשערך היא מצורה מסוימת (למשל הצורה הבאה: ) ולאחר מכן מנסות למצוא ערכים למשתנים בעזרת מאפטם (Optimizer) כלשהו (כמו least squares). היתרון בשיטות פרמטריות הוא שיש לחוקר אפשרות להכניס הנחה מקלה על המידע, שמאפשרת להשתמש במודל סטטיסטי פשוט יותר. בנוסף, במקום לחפש פונקציה מצורה כלשהי במרחב אנחנו מחפשים מספר קבוע של פרמטרים. החיסרון הוא שבמידה ו לא מהצורה שהנחנו שהיא הביצועים של הפונקציה שנעריך עלולים להיות לא אופטימליים (שכן הנחת הבסיס שהמידע מתנהג על פי פונקציה מהצורה הזו לא בהכרח נכונה) – שים לב שגם רשת היא גישה פרמטרית לכן לפעמים ההנחה שאתה עושה על המציאות היא מאוד נמוכה, והשליטה בסיבוכיות המודל באופן ישיר יכולה להוות יתרון חזק למשל בהגבלת מספר השכבות למניעת overfitting.

גישות א-פרמטריות לא מניחות משהו על מבנה אלא פשוט מנסות למצוא פונקציה בהתאם לפונקציית המטרה. היתרון בשיטות אלה הוא שהן לא עושות הנחות על המידע לפני האימון, ולכן גמישות יותר ויכולות לבטא טווח רחב יותר של התפלגויות של המידע (קח בחשבון שגישה פרמטרית יכולה להיות מאוד מסובכת גם כגון רשת – לכן ההגדרה הזאת לא תמיד אומרת הרבה).. החיסרון הוא שבגלל שמרחב החיפוש גדול מאד, יש דרישה להרבה דוגמאות שונות של על מנת לקבל תוצאה טובה (לא בהכרח שיש צורך בהרבה דוגמאות כמו שייקח למודל זמן רב להתאמן – למשל רשת, גם קטנה, עדיין צריכה כמות גדולה יותר של דוגמאות לרוב על פני RF).

1. שיטות גמישות מאפשרות להעריך פונקציה מצורות יותר מגוונות. היתרון שבשיטות גמישות הוא שיש להם פוטנציאל (דגש על פוטנציאל וחוסר מגבלות – לא בהכרח שזה יהיה כך) לתת תוצאות יותר טובות. חסרון של שיטות גמישות הוא שמודלים גמישים מקשים על הבנה של הקשר בין משתנה/פיצ'ר כלשהו למשתנה המטרה (ממש לא נכון. RF כדוגמא נגדית קלה). חיסרון נוסף הוא שיש סכנה ל overfitting למידע שיש לנו (נכון. בגלל שהן לא מוגבלות לתצורה כלשהי של פונקציה, יש להן סיכוי יותר חזק להגיע לOverfitting על מידע הtrain [דוגמא קלה מאוד לזה זה מודל לא פרמטרי טיפש מאוד שפשוט זוכר את כל הדוגמאות וחוזה אותן במדויק בtrain – קצת דומה לKNN עם k=1]).
2. Bias-variance tradeoff. Variance של מודל מתאר כמה הפונקציה שנחזה תשתנה בעקבות שימוש ב training set שונה ו bias מתאר כמה המודל שלנו תואם את התפלגות התצפיות בפועל (שים לב שיש שגיאה שאפשר ממש למדוד variance אבל יותר קשה למדוד bias שכן יש גם שגיאות תיוג / מדידה שלא ניתן לתקן). לרוב, מודל פשוט מדי יסבול מ bias גבוה ומודל גמיש מדי יסבול מ variance גבוה. היות ופשטות המודל תמיד נמצאת ביחס הפוך מגמישות המודל, כך גם bias ו variance. דוגמה לכך היא שימוש ברגרסיה לינארית על מנת לחזות פונקציה שבפועל יותר מסובכת (כמו random של python [כן? עם RF אתה תצליח לחזות random? – האמירה נכונה הדוגמא ממש לא]). הדבר יגרום למודל עם bias גבוה.