

## קורס 67800 תשע"ג

### בחינה מועד ב'

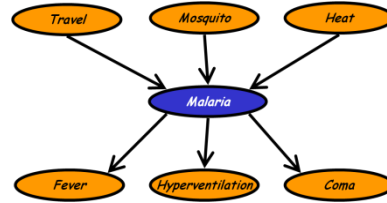
הנחיות:

- בבחינה חמש שאלות. עליכם לענות על שלוש מתוכן (משקל כל שאלה 33 נקודות).
- לא ייבדקו יותר משלוש שאלות כך שאם פתרתם שאלות נוספות עליכם לבחור אילו שאלות ייבדקו ולסמן זאת בצורה ברורה.
- עליכם לנמק היטב כל תשובה ולהוכיח באופן מדויק כל טענה שנדרשתם להוכיח. תשובה נכונה ללא נימוק ו/או דרך לא תזכה אתכם בנקודות.
- עליכם לענות על כל השאלות הבאות באופן עצמאי. תלמיד שיתפס מעתיק, או שיהיה קיים חשש כי העתיק יועבר לטיפול רשויות האוניברסיטה.
- אין להשתמש בחומר עזר מכל סוג שהוא.
- (1 נקודות) ודאו כי מספר הזהות שלכם רשום על מחברת הבחינה.

משך הבחינה: שלוש שעות

בהצלחה !

1. לפניכם רשת בייסיאנית אשר מתארת משתנים המעורבים במחלת המלריה. הניחו כי אי-התלויות המקודדות ברשת ורק הן מתקיימות בהתפלגות האמיתית (כלומר הרשת היא  $p$ -map של ההתפלגות). חוקר אשר לא מודע לקיומה של מחלת המלריה ו/או למבנה של הרשת, רוצה ללמוד מבנה של מודל מעל המשתנים הנצפים: הגורמים והסימפטומים (ללא משתנה המחלה עצמו). לאורך כל השאלה הניחו כי ההתפלגות האמפירית של דוגמאות האימון גם היא  $p$ -map של ההתפלגות האמיתית.



א. (11 נקודות) אם משתמשים בפונקציית ציון הנראות, מה תהיה הרשת שהחוקר ילמד? נמקו היטב.

ב. (11 נקודות) האם קיים אוסף דוגמאות (סופי או אינסופי, לבחירתכם) כך שאם נלמד מבנה בעזרת פונקציית ציון BIC נקבל מבנה שונה מ(א). אם התשובה חיובית, הסבירו היטב מדוע. אם התשובה שלילית, תארו באופן איכותי את התפלגות של הדוגמאות והסבירו מדוע ילמד מבנה שונה.

ג. (11 נקודות) האם תשובתכם ל(ב) תשתנה עבור בחירה מסוימת של פריור דיריכלה ושימוש בפונקציית הציון הבייסיאנית. נמקו את תשובתכם.

2. נתונה רשת מרקובית.

א. (9 נקודות) נניח שלרשת מבנה של עץ פורש. מה יהיה רוחב העץ של הרשת אם נוסיף לה קשת בודדת? הוכיחו.

ב. (12 נקודות) נניח שלרשת מבנה של עץ פורש ושלמבנה זה מוסיפים שתי קשתות. האם ניתן לומר בוודאות מה יהיה רוחב העץ של הרשת לאחר ההוספה? אם התשובה חיובית הוכיחו. אחרת, הראו 2 דוגמאות בהן מתקבל רוחב עץ שונה.

ג. (12 נקודות) נתון גרף  $G$  עם רוחב עץ  $K$ . נאמר ש  $J$  Junction Tree של  $G$  הוא חלק (smooth) אם כל קליקה בו היא בגודל  $K+1$  והחפיפה בין כל שתי קליקות סמוכות בעץ היא בדיוק  $K$ . הראו שתמיד קיים Junction Tree חלק.

3. בשאלה זו נבחן שימוש ברשת מרקובית לביצוע הפרדה של אובייקט מהרקע (סגמנטציה). לכל פיקסל נשייך משתנה בינארי עבורו ערך 1 יציין שייכות הפיקסל לאובייקט וערך 0 לרקע. נגדיר פקטורים על פני פיקסלים שכנים כך ש-  $\phi(X_i, X_j) = \alpha_{i,j}$  אם שני הפיקסלים מקבלים את אותו הערך ו-1 אחרת. עבור מודל זה נגדיר גרף צברים (cliques) בו קיים צבר לכל פקטור ולכל משתנה, וכל צבר התואם לפקטור מחובר בקשת לצברים של המשתנים המופיעים בפקטור. נרמה להשתמש באלגוריתם Loopy-Max-Product (להזכירכם, אלגוריתם זה זהה ל Loopy-Belief-Propagation כאשר מבצעים מקסימום במקום סכימה).

א. (10 נקודות) בהנחה שאין במודל פקטורים נוספים ושכל ההודעות מאותחלות להיות אחידות. חשבו את ההודעה הראשונה הנשלחת על ידי האלגוריתם משתנה לאחד משכניו. למה תתכנס הריצה.

ב. (12 נקודות) מהי ההשמה (או השמות) שמקבלות הסתברות מקסימלית במודל הנתון? האם תוכלו לחשב אותן על ידי מיקסום  $b_i(x_i) \propto \prod_{k \in N(i)} m_{ki}(x_i)$  כאשר  $m$  הן ההודעות של האלגוריתם. הציעו שיטה אחרת למציאת המקסימום.

ג. (6 נקודות) כעת נוסף למודל פקטורים בינאריים (לא אחידים) לכל פיקסל בנפרד המגדירים (למשל על סמך צבע הפיקסל) את העדפת הפיקסל להיות חלק מהאובייקט או הרקע. חיזרו על חישוב ההודעה מ-(א).

ד. (5 נקודות) עבור המודל מסעיף ג, חישוב מה צריכים להיות הפקטורים כדי שהאלגוריתם יחזיר תשובה נכונה. הסבירו את תשובתכם.

4. ג

a. הראו ש  $X \perp Y, Z | W \Rightarrow X \perp Z | W$  &  $X \perp Y | W, Z$ .

b. הראו שבגרפים לא מכוונים קיום התכונות  $I_{LM}(G)$  גורר את קיום התכונות  $I_{pair}(G)$ .

c. נתון פילוג  $p(X_1, X_2, X_3)$  על שלושה משתנים בינאריים, שבו  $X_1 = X_2 = X_3$  והפילוג השולי על כל  $X_i$  הוא אחיד (כלומר הסתברות 0.5 לכל תוצאה). מצאו גרף

על משתנים אלו כך ש  $I(p) \supseteq I_{LM}(G)$  אך  $I(p) \not\supseteq I_{pair}(G)$ .

5. נתייחס למודל Naïve Bayes עם משתנה מחלקה C ושתי תצפיות  $X_1, X_2$ . הניחו פרמטריזציה מלאה של המודל. נניח כעת שב  $N_1$  תצפיות ראיתם רק את  $X_1$ , ב  $N_2$  תצפיות ראיתם רק את  $X_2$ , וב  $N_{1,2}$  תצפיות ראיתם את שניהם. נסמן את הפילוגים האמפיריים המתאימים ב  $p_{D,1}(X_1), p_{D,1}(X_2), p_{D,1,2}(X_1, X_2)$ .

a. כתבו את פונקציית הנראות שתוצו למקסם במקרה זה, ואת פונקציית ה Q המתאימה.

b. גזרו את העדכונים של EM עבור מקרה זה.

c. כעת הניחו שהפרמטרים "קשורים" כך ש:  $\theta_1(x_1 | c) = \theta_2(x_2 | c)$ . גזרו את עדכוני EM.