

מבחן בזיהוי דיבור 2-608-89

מועד ב', סמסטר ב' תשס"ג

משך הבחינה שעתיים וחצי.
 מותר להשתמש בכל חומר עזר.
 שאלה 4 (מודלים מרקוביים חבויים) היא שאלת חובה.
 יש לענות על לפחות 3 שאלות מתוך שאר השאלות.
 מותר לענות על יותר מ-4 שאלות (והטובות ייבחרו בשקלול הסופי).

1. נתונות שתי מערכות אימות דובר. ביצועי המערכות נמדדו בתנאים של דיבור ללא רעש (שקט) ובתנאים של דיבור רועש.
 להלן ביצועי המערכות:
 מערכת א':

$$\Pr(\text{דובר מתחזה, תנאי רעש} \mid \text{זיהוי נכון}) = 0.8$$

$$\Pr(\text{דובר לא מתחזה, תנאי רעש} \mid \text{זיהוי נכון}) = 0.8$$

$$\Pr(\text{דובר מתחזה, תנאי שקט} \mid \text{זיהוי נכון}) = 0.9$$

$$\Pr(\text{דובר לא מתחזה, תנאי שקט} \mid \text{זיהוי נכון}) = 0.9$$

מערכת ב': ביצועים זהים להם למערכת א'.

הנח כי שתי המערכות בלתי תלויות.

נבנה משתי המערכות מערכת משולבת אשר מקבלת את תוצאת שתי המערכות ומחליטה "מתחזה" אם 2 המערכות מחליטות "מתחזה", המערכת תחליט "לא מתחזה" אם 2 המערכות מחליטות "לא מתחזה", ולבסוף, במקרה של אי-הסכמה המערכת המשולבת תחליט "לא-ידוע".

כמו כן ידוע שהסיכוי האפרורי לתנאי רעש הוא 0.5:

$$\Pr(\text{תנאי רעש}) = \Pr(\text{תנאי שקט}) = 0.5$$

כמו כן ידוע שהסיכוי האפרורי לדובר מתחזה הוא 0.5:

$$\Pr(\text{דובר מתחזה}) = 0.5$$

- א. מהו הסיכוי, בהינתן שהדובר **מתחזה**, שהמערכת המשולבת תחליט שהדובר **איננו מתחזה**?
- ב. מהו הסיכוי, בהינתן שהדובר **איננו מתחזה**, שהמערכת המשולבת תחליט שהדובר **איננו מתחזה**?
- ג. מהו הסיכוי, בהינתן שהמערכת המשולבת מחליטה שהדובר **איננו מתחזה**, שהדובר **אכן איננו מתחזה**?

2. ידוע שקזינו מסוים משתמש לפעמים בקובייה חוקית ולפעמים בקובייה בלתי חוקית (מוטית). סוכן פרץ בלילה לקזינו, הטיל את הקובייה המוטית 6 פעמים וקיבל את ההטלות הבאות: 2,6,1,3,4,6

למחרת, את/ה מגיע לקזינו ורואה את סידרת ההטלות הבאות (של קובייה ספציפית):
6,2,1,6,3,5

הנחה: ההסתברות האפריורית לקובייה מוטית היא 50%.
א. חשבי את הסיכוי שהקובייה היא הקובייה המוטית ולא הקובייה החוקית תוך שימוש בגישת ML (נראות מירבית) לשיערוך פרמטרים.
ב. חשבי את הסיכוי שהקובייה היא הקובייה המוטית ולא הקובייה החוקית תוך שימוש בגישת ML אך תוך שימוש בפרמטר $\epsilon=1$.

3. עליך לתכנן מערכת זיהוי דיבור המבדילה בין המילה yes למילה no. תארי את המרכיבים השונים של המערכת, את שלב באימון ואת שלב הזיהוי.

4. נתון המודל המרקובי הבא:

מצבים: 1,2,3

תצפיות אפשריות: D,C,B,A

הסתברויות התחלה: 100% להתחיל במצב הראשון.

הסתברויות מעבר:

$$\Pr(1 \rightarrow 1) = \Pr(1 \rightarrow 2) = \Pr(2 \rightarrow 2) = \Pr(2 \rightarrow 3) = \Pr(3 \rightarrow 3) = \Pr(3 \rightarrow \text{exit}) = 0.5$$

Exit: סיום סידרת המצבים והתצפיות.

הסתברויות פליטה:

State 1: $\Pr(A)=\Pr(B)=\Pr(C)=\Pr(D)=25\%$

State 2: $\Pr(A)=\Pr(B)=40\%$ $\Pr(C)=\Pr(D)=10\%$

State 3: $\Pr(A)=\Pr(B)=\Pr(C)=\Pr(D)=25\%$

נתונה סידרת התצפיות הבאה: (משמאל לימין) CAABCD

א. מצא/י את סידרת המצבים המסתברת ביותר.

ב. מהי ההסתברות של סידרת התצפיות? (יש להתחשב רק בסידרת המצבים המסתברת ביותר).

5. הנחי כי ברשותך התוכנות הבאות: א. תוכנה להצפנת קבצים קוליים אשר מקבלת מפתח שהינו מספר טבעי בין 1 ל-99999, ומצפינה באמצעותו את הקובץ הקולי. ב. תוכנה לפענוח קובץ קולי מוצפן אשר מקבלת מפתח (בין 1 ל-99999) ומפעילה את הפונקציה ההופכית לפונקציית ההצפנה אם ורק אם המפתח זהה למפתח שבאמצעותו הוצפן הקובץ הקולי (אחרת מופעלת פונקציה אחרת), כלומר עבור המפתח הנכון נקבל את הקובץ

המקורי, ועבור המפתח הלא נכון נקבל קובץ שלא נשמע כמו קובץ דיבור (נשמע כמו אוסף צפופים).

עליך לתכנן אלגוריתם פריצה אשר בהינתן קובץ מוצפן, האלגוריתם ישחזר את הקובץ המקורי וזאת בלי שהמפתח ידוע מראש.
הסבר/י את האלגוריתם שתיכנת.