תירגול מספר 1 : סיווג בייסיאני אופטימלי

1 תקציר התאוריה

גישה סטטיסטית לסיווג המניחה כי כל הגדלים הינם מ"א בעלי פילוג הסתברות.

1.1 סימונים

Ω – מרחב סופי של קטגוריות (מחלקות). $\omega_i \in \Omega,\ i = 1,...,N$

X – מרחב הקלט (תבניות). $x \in X$.

$f : X \mapsto \Omega$ - מסווג כל $x \in X$ ל- $\omega \in \Omega$.

הנחה בסיסית – הגדלים הבאים ידועים :

א. $\left\{p(\omega_1), p(\omega_2), ..., p(\omega_N)\right\}$ - הסתברויות אפריוריות של הקטגוריות השונות.

ב. $p(x \mid \omega_i),\ i = 1,...,N$ - הסתברות הקלט בהינתן שהוא שייך לקטגוריה ω_i.

(עבור המקרה בו X מרחב <u>רציף</u> $p(x \mid \omega_i)$ מציין את <u>צפיפות ההסתברות</u>)

1.2 מסווג בייסיאני אופטימלי

תזכורת – נוסחת בייס (Bayes):

$$p(\omega_i \mid x) = \frac{p(x \mid \omega_i)p(\omega_i)}{p\left(x\right)} = \frac{p(x \mid \omega_i)p(\omega_i)}{\sum_i p(x \mid \omega_i)p(\omega_i)}$$

המסווג הבייסיאני האופטימלי נתון ע"י :

$$f(x) = \underset{i=1,...,N}{\arg\max}\, p(\omega_i \mid x)$$

כפי שהוכח בהרצאה, מסווג זה מביא למינימום את הסתברות השגיאה המותנית ואת הסתברות השגיאה הממוצעת.

2 שאלות

<u>שאלה מספר 1</u>

נתון: $X=\mathbb{R}^2$, $\Omega=\{\omega_1,\omega_2,\omega_3\}$.

הפילוג האפריורי הינו אחיד: $p(\omega_1)=p(\omega_2)=p(\omega_3)=\frac{1}{3}$.

פילוג הסתברות מותנה נתון ע"י $p(x\mid\omega_i)\sim N(\mu_i,\Sigma)$, כאשר

$$\mu_1=\begin{pmatrix}0\\3\end{pmatrix},\mu_2=\begin{pmatrix}3\\0\end{pmatrix},\mu_3=\begin{pmatrix}-3\\0\end{pmatrix}$$

$$\Sigma=\begin{pmatrix}1/2 & 0\\ 0 & 1/2\end{pmatrix}$$

א. מהו חוק ההחלטה הבייסיאני האופטימאלי? מהם תחומי ההחלטה במישור?

ב. האם עקום ההפרדה המתקבל במקרה הגאוסי תמיד זהה בצורתו? במה תלוי עקום זה?

<u>שאלה מספר 2</u>

נתון וקטור כניסה $x\in\{0,1\}^d$ כלומר $x=(x_1,...,x_d),\ x_i\in\{0,1\}$ (וקטורים בינאריים באורך d).

נגדיר,

$$p_i=p(x_i=1\mid\omega_0)$$
$$q_i=p(x_i=1\mid\omega_1)$$

וכן נתון כי כל רכיב בווקטור מוגרל בת"ס ברכיבים אחרים.

א. מהו חוק ההחלטה האופטימאלי?

ב. מה קורה אם עבור i מסוים $p_i=q_i$? מה משמעות התוצאה?

<u>שאלה מספר 3</u>

נתונות שתי קטגוריות ω_0,ω_1 בעלות הסתברות אפריורית זהה, וכן נתונים הפילוגים

$$p(x\mid\omega_0)=\frac{1}{\sqrt{2\pi}}\exp\{-(x-2)^2/2\},\qquad p(x\mid\omega_1)=\frac{1}{2\sqrt{2\pi}}\exp\{-(x-4)^2/8\}$$

מהו חוק ההחלטה הבייסיאני האופטימאלי?