15. Bayesov klasifikator

Strojno učenje 1, UNIZG FER, ak. god. 2021./2022.

Jan Šnajder, vježbe, v3.0

1 Zadatci za učenje

1. [***Svrha: Razumjeti model Bayesovog klasifikatora i njegove komponente. Razumjeti što su to generativni modeli, kako se razlikuju od diskriminativnih te koje su njihove prednosti i njihovi nedostatci.***]

   (a) Definirajte model Bayesovog klasifkatora i navedite sve veličine koje se pojavljuju u definiciji modela. Objasnite zašto faktoriziramo brojnik. Objasnite ulogu nazivnika i objasnite kada ga možemo zanemariti.

   (b) Je li taj model parametarski ili neparametarski? Obrazložite odgovor.

   (c) Objasnite zašto Bayesov model nazivamo generativnim i opišite generativnu priču Bayesovog klasifikatora.

   (d) Objasnite razliku između generativnih i diskriminativnih modela te navedite prednosti jednih i drugih.

2. [***Svrha: Isprobati izračun maksimalne aposteriorne hipoteze i najvjerojatnije hipoteze uz minimizaciju rizika.***] Razmotrimo problem klasifikaciji neželjene el. pošte u klase *spam* ($y = 1$), *important* ($y = 2$) i *normal* ($y = 3$). Neka su apriorne vjerojatnosti tih klasa $P(y = 1) = 0.2$, $P(y = 2) = 0.05$ i $P(y = 3) = 0.75$. Za neku poruku el. pošte $\mathbf{x}$ izglednosti iznose $p(\mathbf{x}|y = 1) = 0.8$ i $p(\mathbf{x}|y = 2) = p(\mathbf{x}|y = 3) = 0.5$. Izračunajte aposteriorne vjerojatnost za svaku od klasa te maksimalnu aposteriornu hipotezu za primjer $\mathbf{x}$.

3. [***Svrha: Razviti intuiciju za model kontinuiranog Bayesovog klasifikatora.***]

   Izrađujemo Bayesov model za klasifikaciju primjera iz $\mathcal{X} = \mathbb{R}$ u tri klase. Učenjem na skupu primjera dobili smo sljedeće parametre modela: $P(y = 1) = 0.3$, $P(y = 2) = 0.2$, $\mu_1 = -5$, $\mu_2 = 0$, $\mu_3 = 5$, $\sigma_1^2 = 5$, $\sigma_2^2 = 1$, $\sigma_3^2 = 10$. Skicirajte funkcije gustoće vjerojatnosti $p(x|y)$, $p(x, y)$, $p(x)$ i $p(y|x)$.

4. [***Svrha: Razumjeti izvod modela kontinuiranog Bayesovog klasifikatora i osvježiti potrebno znanje matematike.***]

   (a) Krenuvši od izraza (4.29) iz skripte, izvedite model višedimenzijskog Bayesovog klasifikatora s kontinuiranim ulazima s dijeljenom i dijagonalnom kovarijacijskom matricom.

   (b) Napišite broj parametara ovog modela.

   (c) Objasnite zašto je izglednost faktorizirana u produkt univarijatnih razdioba, što odgovara pretpostavci o uvjetnoj nezavisnosti, premda značajke mogu biti nelinearno uvjetno zavisne.

5. [***Svrha: Razviti intuiciju za složenost modela kontinuiranog Bayesovog klasifikatora i shvatiti kako se problem u konačnici svodi na odabir optimalnog modela.***] Želimo izgraditi klasifikator za klasifikaciju brucoša u jednu od dvije klase: $y = 1 \Rightarrow$ *"Završava FER u roku"* i $y = 2 \Rightarrow$ *"Produljuje studij"*. Svaki je primjer opisan sa šest ulaznih varijabli: prosjek ocjena 1.–4. razreda (četiri varijable), bodovi državne mature iz matematike te bodovi državne mature iz fizike. Raspolažemo trima modelima: modelom $\mathcal{H}_1$ s dijeljenom kovarijacijskom matricom, modelom $\mathcal{H}_2$ s dijagonalnom (i dijeljenom) kovarijacijskom matricom i modelom $\mathcal{H}_3$ s izotropnom kovarijacijskom matricom.

(a) Koliko svaki od ova tri modela ima parametara?

(b) Za koji od ova tri modela očekujete da će najbolje generalizirati u ovom konkretnom slučaju (uzmite u obzir prirodu problema i očekivane odnose između značajki)? Zašto?

(c) Nacrtajte skicu funkcije empirijske pogreške i pogreške generalizacije i naznačite na njoj točke koje označavaju navedenim trima modelima.

(d) Kako biste u praksi odredili koji ćete model upotrijebiti?

2 Zadatci s ispita

1. (T) Bayesov klasifikator definirali smo na sljedeći način:

$$h_j(\mathbf{x};\boldsymbol{\theta}) = P(y=j|\mathbf{x}) = \frac{p(\mathbf{x}|y)P(y)}{\sum_{y'} p(\mathbf{x}|y')P(y')}$$

Neka je broj klasa veći od dva, $K > 2$, a značajke neka su realni brojevi, $\mathbf{x} \in \mathbb{R}^n$. **Koje teorijske distribucije ćemo koristiti za $P(y)$ i $P(\mathbf{x}|y)$?**

A Kategoričku distribuciju za $P(y)$ i Gaussovu distribuciju za $P(\mathbf{x}|y)$

B Bernoullijevu distribuciju za $P(y)$ i Gaussovu distribuciju za $P(\mathbf{x}|y)$

C Kategoričku distribuciju za $P(y)$ i za $P(\mathbf{x}|y)$

D Gaussovu distribuciju za $P(y)$ i multinulijevu distribuciju za $P(\mathbf{x}|y)$

2. (T) Bayesov klasifikator definiran je kao

$$h(\mathbf{x};\boldsymbol{\theta}) = \text{argmax}_y\ p(\mathbf{x}|y)P(y)$$

Po čemu se vidi da je ovo generativan, a ne diskriminativan model?

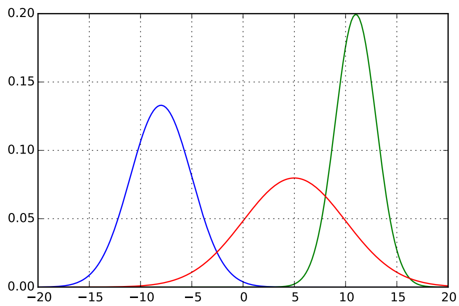
A Modelira vjerojatnost primjera i oznaka, budući da je, na temelju pravila umnoška, umnožak $p(\mathbf{x}|y)P(y)$ jednak zajedničkoj vjerojatnosti $p(\mathbf{x}, y)$

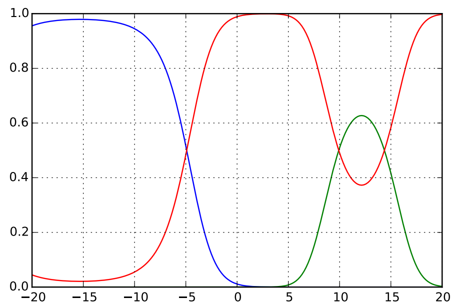
B Zajedničku vjerojatnost primjera i oznaka, $p(\mathbf{x}|y)P(y)$, faktorizira u dva faktora te zanemaruje nazivnik $p(\mathbf{x})$, koji je ionako konstantan za svaku klasu y

C Parametre distribucija $p(\mathbf{x}|y)$ i $P(y)$, a time indirektno i parametre aposteriorne distribucije $P(y|\mathbf{x})$, računa MAP-procjeniteljem, čime sprječava prenaučenost

D Primjer $\mathbf{x}$ klasificira prema MAP-hipotezi, dakle u klasu koja maksimizira aposteriornu vjerojatnost oznake, $p(y|\mathbf{x})$, koja je proporcionalna zajedničkoj vjerojatnosti primjera i oznaka, $p(\mathbf{x}, y)$

3. (P) Koristimo Gaussov Bayesov klasifikator kako bismo riješili troklasni klasifikacijski problem. Procijenjene gustoće vjerojatnosti za izglednosti klasa su $p(x|y=1) = \mathcal{N}(-8, 3)$, $p(x|y=2) = \mathcal{N}(5, 5)$ i $p(x|y=3) = \mathcal{N}(11, 2)$. Na slikama ispod prikazane su izglednosti klasa (lijeva slika) i aposteriorne vjerojatnosti dobivene Bayesovim pravilom (desna slika):





S obzirom na ova dva grafikona, što su najizglednije vrijednosti za apriorne vjerojatnosti klasa?

A $P(y=1)=0.1, P(y=2)=0.7, P(y=3)=0.2$

B $P(y=1)=P(y=2)=P(y=3)=\frac{1}{3}$

C $P(y=1)=P(y=2)=0.4, P(y=3)=0.2$

D $P(y=1)=P(y=2)=0.1, P(y=3)=0.8$

4. (P) Gaussovim Bayesovim klasifikatorom rješavamo problem klasifikacije u $K=10$ klasa sa $n=5$ značajki. Prisjetite se da kod Gaussovog Bayesovog klasifikatora uvođenjem odgovarajućih pretpostavki na kovarijacijsku matricu Σ možemo utjecati na broj parametara modela a time onda i na složenost modela. Razmatramo tri modela s kovarijacijskim matricama u koje smo ugradili sljedeće pretpostavke:

$\mathcal{H}_1$: Značajke nisu korelirane, no imaju različite varijance unutar klase i između klasa

$\mathcal{H}_2$: Značajke nisu korelirane, imaju jednaku varijancu unutar svake klase, no različitu za svaku klasu

$\mathcal{H}_3$: Između značajki postoje korelacije, ali se one ne razlikuju između klasa

Neka '$\supset$' označava relaciju "složeniji od", a neka '$>$' označava relaciju "ima više parametara od". **Što možemo zaključiti o složenosti i broju parametara za gornja četiri modela?**

A $\mathcal{H}_1 > \mathcal{H}_3 > \mathcal{H}_2$, $\mathcal{H}_1 \supset \mathcal{H}_2$

B $\mathcal{H}_1 > \mathcal{H}_2 > \mathcal{H}_3$, $\mathcal{H}_1 \supset \mathcal{H}_2 \supset \mathcal{H}_3$

C $\mathcal{H}_3 > \mathcal{H}_1 > \mathcal{H}_2$, $\mathcal{H}_1 \supset \mathcal{H}_2$

D $\mathcal{H}_3 > \mathcal{H}_1 > \mathcal{H}_2$, $\mathcal{H}_3 \supset \mathcal{H}_2 \supset \mathcal{H}_1$

5. (N) Na skupu označenih primjera u ulaznome prostoru dimenzije $n=3$ treniramo Gaussov Bayesov klasifikator za klasifikaciju primjera u $K=2$ klase, uz pretpostavku dijeljene kovarijacijske matrice. Model je definiran kao

$$h_j(\mathbf{x}) = \ln p(\mathbf{x}, y)$$

Prisjetimo se da je izglednost klase s oznakom $y=j$ kod Gaussovog Bayesovog klasifikatora definirana multivarijantnom Gaussovom gustoćom vjerojatnosti:

$$p(\mathbf{x}|y=j) = \frac{1}{(2\pi)^{n/2}|\boldsymbol{\Sigma}_j|^{1/2}} \exp\Big\{-\frac{1}{2}(\mathbf{x}-\boldsymbol{\mu}_j)^{\mathrm{T}}\boldsymbol{\Sigma}_j^{-1}(\mathbf{x}-\boldsymbol{\mu}_j)\Big\}$$

gdje je Σ_j matrica kovarijacije za klasu j. Treniranjem modela dobili smo sljedeće procjene za parametre:

$$\hat{\mu}_1 = 0.2 \qquad \hat{\boldsymbol{\mu}}_1 = (1, 0, -2) \qquad \hat{\Sigma}_1 = \begin{pmatrix} 5 & 2 & 4 \\ 2 & 5 & 3 \\ 4 & 3 & 6 \end{pmatrix}$$

$$\hat{\mu}_2 = 0.8 \qquad \hat{\boldsymbol{\mu}}_2 = (2, -1, 5) \qquad \hat{\Sigma}_2 = \begin{pmatrix} 6.25 & -0.5 & -1 \\ -0.5 & 1.25 & -0.75 \\ -1 & -0.75 & 3.5 \end{pmatrix}$$

Iz ovoga smo zatim procijenili dijeljenu kovarijacijsku matricu $\hat{\Sigma}$ definiranu kao težinski prosjek kovarijacijskih matrica $\hat{\Sigma}_j$, $j=1,2$. Zanima nas klasifikacija modela za primjer $\mathbf{x}=(0,0,0)$. **Koliko iznosi predikcija modela za klasu $y=1$ za taj primjer, $h_1(\mathbf{x})$?**

A -6.885 B $+0.002$ C -4.819 D -6.429