Strojno učenje – Auditorne vježbe

Bayesov teorem (Thomas Bayes, 1702.-1761.)

Bayesov teorem nam kaže kako prepraviti vjerovanja u svijetlu novih događaja. Bayesov teorem povezuje uvjetne i apriorne vjerojatnosti događaja.

Pokaz:
$P(AB) = P(A)\,P(B|A) = P(B)P(A|B)$

$P(B|A) = P(B)P(A|B)/P(A)$

$$P(H_i|A) = P(H_i)P(A|H_i)/P(A)$$

$P(H_i)$ – apriorna vjerojatnost hipoteze H_i
$P(H_i|A)$ – aposteriorna vjerojatnost hipoteze H_i
$P(A)$ – vjerojatnost nastupanja događaja A

Apriorna vjerojatnost hipoteze može biti «objektivna», kada se temelji na nekakvom stvarnom eksperimentu (pr. uzastopno bacanje novičića) ili «subjektivna» kada se temelji na vjerovanju.

Primjer.

U žari su 3 kuglice. Znamo da je svaka od njih bijele ili crne boje. Točan broj kuglica pojedine boje je nepoznat i pretpostavljamo da je svaka mogućnost jednako vjerojatna. Pretpostavimo 4 hipoteze:
H_i = { u žari se nalazi *i* bijelih kuglica }
$P(H_0) = P(H_1) = P(H_2) = P(H_3) = ¼$

Biramo kuglicu. Izvadili smo bijelu kuglicu. Kakve su sada vjerojatnosti hipoteza?
H_0 je nemoguća, to odmah zaključujemo.
$P(A|H_0) = 0$
$P(A|H_1) = 1/3$
$P(A|H_2) = 2/3$
$P(A|H_3) = 1$

Formula potpune vjerojatnosti

$P(B) = P(B|A_1)P(A_1) + \ldots + P(B|A_n)P(A_n)$
B – proizvoljan događaj
$A_1,\ldots, A_n$ – međusobno isključivi događaji ($P(X_i|X_j)=0,\ i \neq j$)

$P(A) = ¼ * 0 + ¼* 1/3 + ¼ * 2/3 + ¼* 1 = ½$ (formula potpune vjerojatnosti)
Računamo po Bayesovom teoremu:
$P(H_0|A) = 0$ (ovo smo odmah zaključili)
$P(H_1|A) = 1/6$
$P(H_2|A) = 1/3$
$P(H_3|A) = 1/2 \Rightarrow H_3$ je maksimalna aposteriorna hipoteza (h_{MAP}).

Maksimalna aposteriorna hipoteza je hipoteza s najvećom vjerojatnošću nakon što se dogodio određeni događaja. Najviše ćemo vjerovati upravo maksimalnoj aposteriornoj hipotezi.

Naivni Bayesov klasifikator

Problem klasifikacije vektora $x = (f_1,..., f_n)$ u skup kategorija C modeliramo pomoću uvjetnih vjerojatnosti.

$P(C|F_1, \ldots, F_n) = P(C)\, P(F_1,\ldots, F_n|C) / P(F_1,\ldots, F_n)$ (pomoću Bayesovog teorema)

Događaji su pojavljivanje određenih vrijednosti značajki $F_1, ..., F_n$, a hipoteze su pripadnosti kategorijama.

Računamo:

$P(F_1, \ldots, F_n) = \text{konst.}$

(nazivnik formule je konstantan za sve kategorije pa ga u smislu traženja h_{MAP} niti ne trebamo računati)

$P(C)\, P(F_1,\ldots, F_n|C) =$
$= P(C)\, P(F_1|C)\, P(F_2,\ldots, F_n|C,F_1)$
$= P(C)\, P(F_1|C)\, P(F_2|C)\, P(F_3,\ldots, F_n|C,F_1,F_2)$
$= ...$

Uvodimo tzv. **<u>naivnu pretpostavku</u>**:

$$P(F_i|C,F_j) = P(F_i|C),\ i \neq j$$

Ovo nazivamo **uvjetnom nezavisnošću događaja (značajki)**.

Sada gornji rastav postaje:

$$P(C)\, P(F_1,\ldots, F_n|C) = P(C)\, P(F_1|C) \ldots P(F_n|C) = P(C)\prod_{i=1}^{n} P(F_i = f_i \mid C)$$

Uvodimo pravilo Naivnog Bayesovog klasifikatora koje kaže da odaberemo onu kategoriju koja ima najveću aposteriornu vjerojatnost tj. h_{MAP}.

$$klasifikacija(f_1,...,f_n) = \arg\max_{c \in C}\left(P(c)\prod_{i=1}^{n} P(F_i = f_i \mid c)\right)$$

Prednosti Naivnog Bayesovog klasifikatora:
- jednostavan izračun (konstrukcija)
- radi i s malo primjera za učenje
- otporan na šum (nema izbacivanja hipoteza već samo mijenjanja vjerojatnosti)

Zadaci

Zadatak 1.

Poznate su apriorne vjerojatnosti hipoteza, a to su H_1 da osoba ima rak, odnosno H_2 da osoba nema rak, tj. P(rak) = 0,008 i P(–rak) = 0,992. Test na rak klasificira točno pozitivne slučajeve u 98% slučajeva, a negativne u 97% slučajeva, tj. vrijedi:

P(test+ | rak) = 0,98	P(test– | rak) = 0,02
P(test– | –rak) = 0,97	P(test+ | –rak) = 0,03

Pretpostavimo da je nekoj osobi test na rak dao pozitivan rezultat. Nađi maksimalnu aposteriornu hipotezu (h_{MAP}).

Rješenje.

$$h_{MAP} = \arg\max_{h \in H}\left(P(h)P(A \mid h)\right) = \arg\max_{h \in \{rak, -rak\}}\left(P(h)P(test+ \mid h)\right)$$

P(rak)P(test+|rak) = 0,008 * 0,98 = 0,00784
P(–rak)P(test+|–rak) = 0,992 * 0,03 = 0,02976

h_{MAP} = –rak
Vjerojatnije je da osoba nema rak.

Zadatak 2.

U tablici 1. zadani su primjeri za učenje koncepta „Ručak kakvog preferira Marko“. Koristeći naivni Bayesov klasifikator za ciljni koncept odredi klasifikaciju novog primjera: (svinjetina, povrće, mrkva, gljive). Koja je induktivna pretpostavka Naivnog Bayesovog klasifikatora?

	Meso	**Prilog**	**Salata**	**Juha**	**c(x)**
1.	svinjetina	povrće	zelje	gljive	da
2.	teletina	tijesto	mrkva	šparoge	ne
3.	teletina	povrće	mrkva	gljive	ne
4.	piletina	krumpir	mrkva	goveđa	da
5.	piletina	tijesto	zelje	gljive	da
6.	svinjetina	krumpir	cikla	gljive	da
7.	piletina	krumpir	mrkva	šparoge	da
8.	svinjetina	tijesto	mrkva	gljive	ne

Tablica 1. primjeri za učenje

Rješenje.

P(da) = 5/8
P(ne) = 3/8

klasifikacija(svinjetina, povrće, mrkva, gljive) =
= argmax P(c) P(svinjetina | c) P(povrće | c) P(mrkva | c) P(gljive | c)

Navedene uvjetne vjerojatnosti procjenjujemo iz tablice.

da ... 5/8 * 2/5 * 1/5 * 2/5 * 3/5 = 3/250
ne ... 3/8 * 1/3 * 1/3 * 3/3 * 2/3 = $\boxed{1/36}$

=> klasifikacija(svinjetina, povrće, mrkva, gljive) = ne

Induktivna pretpostavka N.B. klasifikatora je uvjetna nezavisnost pojavljivanja svih značajki.