

Keywords: machine learning, artificial neural networks, artificial intelligence, naive Bayes, Hopfield's model

Igor Kononenko  
Univerza v Ljubljani  
Fakulteta za elektrotehniko in računalništvo  
Tržaška 25, Ljubljana

## POVZETEK

V prispevku je opisana posplošitev Bayesovih nevronske mreže na zvezna stanja nevronov, analogno Hopfieldovi posplošitvi. Podane so posplošene prehodne funkcije in dokazana stabilnost izvajanja zvezne Bayesove nevronske mreže, ki temelji na verjetnosti in zvezne Bayesove nevronske mreže, ki temelji na razmerju verjetnosti. Opisane so prednosti zveznih Bayesovih nevronske mreže pred diskretnimi.

## CONTINUOUS BAYESIAN NEURAL NETWORKS

In the paper the Bayesian neural network model is generalized to continuous states of neurons, analogously to Hopfield's generalization. The generalized transition functions are given together with the proof of convergence of execution for the continuous Bayesian neural network based on probability and for the continuous Bayesian neural network based on probability ratio. The advantages of continuous Bayesian neural network models are discussed.

## 1 Uvod

V (Kononenko 1990) so opisane *diskretne* Bayesove nevronske mreže, ki temeljijo na verjetnosti (BNM-p), in *diskretne* Bayesove nevronske mreže, ki temeljijo na razmerju verjetnosti (BNM-odds). Celotna kombinacijska funkcija za izračun novega stanja  $S'_j$  in s tem izhoda nevrona v BNM-p je definirana za:

$$S'_j = D_j(P(S_j = 1 | S_1, \dots, S_N)) \quad (1)$$

kjer je

$$P(S_j | S_1, \dots, S_N) = P(S_j = 1) \prod_{S_i=1, (i \neq j)} \frac{P(S_j = 1 | S_i = 1)}{P(S_j = 1)} \quad (2)$$

verjetnost, da je j-ti nevron aktiven, pri danih indeksih aktivnih nevronov in

$$D_j(X) = \begin{cases} 1, & \text{če } X > P(S_j = 1) \\ S_j, & \text{če } X = P(S_j = 1) \\ 0, & \text{če } X < P(S_j = 1) \end{cases} \quad (3)$$

pragovna odločitvena funkcija. Pri tem so  $S_i, i = 1..N$  trenutna stanja vseh nevronov v mreži in  $P(S_j = 1)$  apriorna verjetnost aktivnega stanja j-tega nevrona. Če je izračunana verjetnost večja od apriorne verjetnosti, potem je novo stanje enako 1, če je enaka apriorni verjetnosti, potem se stanje nevrona ne spremeni, in če je izračunana verjetnost manjša od apriorne verjetnosti, potem

je novo stanje nevrona enako 0.

Celotna kombinacijska funkcija za izračun novega stanja  $S_j'$  in s tem izhoda nevrona v BNM-odds je definirana z:

$$S_j' = Dq_j(odds(S_j = 1|S_1, \dots, S_N)) \quad (4)$$

kjer je

$$\begin{aligned} odds(S_j|S_1, \dots, S_N) &= \\ &= odds(S_j = 1) \prod_{i \neq j} \frac{odds(S_j = 1|S_i = 1)}{odds(S_j = 1)} \end{aligned} \quad (5)$$

razmerje verjetnosti, da je  $j$ -ti nevron aktiven, pri danih indeksih aktivnih nevronov in

$$Dq_j(X) = \begin{cases} 1, & \text{če } X > odds(S_j = 1) \\ S_j, & \text{če } X = odds(S_j = 1) \\ 0, & \text{če } X < odds(S_j = 1) \end{cases} \quad (6)$$

pragovna odločitvena funkcija. Pri tem so  $S_i, i = 1..N$  trenutna stanja vseh nevronov v mreži in  $odds(S_j)$  apriorna verjetnost stanja 1  $j$ -tega nevrona deljena z apriorno verjetnostjo nasprotnega stanja. Če je izračunano razmerje verjetnosti večje od apriornega razmerja, potem je novo stanje enako 1, če je enako apriornemu razmerju, potem se stanje nevrona ne spremeni, in če je izračunano razmerje verjetnosti manjše od apriornega, potem je novo stanje nevrona enako 0.

Ker stanja BNM-p in BNM-odds lahko zavzemajo samo diskretni vrednosti 0 in 1, jim pravimo *diskretni*. V (Kononenko 1990) je dokazana stabilnost asinhronnega izvajanja za obe vrsti mrež iz poljubnega začetnega stanja, z uporabo funkcij podobnosti odvisnih samo od stanja mreže. Le te z asinhronim izvajanjem monotono padajo. Ker je možnih stanj diskretne nevronske mreže končno mnogo, nam to garantira prihod v fiksno točko v končnem številu iteracij. Oglejmo si še enkrat funkciji podobnosti za BNM-p in BNM-odds. Naslednja funkcija definira podobnost med aktivacijskimi nivoji nevronov in njihovimi trenutnimi stanji v BNM-p:

$$Sim(S_1, \dots, S_N) =$$

$$= \prod_{i, S_i=1} \frac{P(S_i = 1|S_1, \dots, S_N)}{P(S_i)} \quad (7)$$

Podobnost je večja, če je aktivacijski nivo (izračunana verjetnost, da je nevron aktiven) večji od apriorne verjetnosti, da je nevron aktiven in obratno. Za BNM-odds je funkcija podobnosti enaka, le da produkt teče preko vseh nevronov (ni pogoja  $S_i = 1$ ).

V tem prispevku je podana generalizacija večsmernih BNM na zvezna stanja nevronov analogno Hopfieldovemu (1984) zveznemu modelu nevronske mreže. Namesto diskretnih stanj 0 in 1 je stanje nevrona v zvezni BNM predstavljeno z realno vrednostjo na intervalu  $[0,1]$ . Stanje nevrona bo proporcionalno verjetnosti, da je nevron trenutno aktiven. V naslednjem razdelku je na kratko opisan Hopfieldov zvezni model nevronske mreže. V nadaljnjih dveh razdelkih sta opisana zvezna modela večsmerne BNM-p in BNM-odds. V zadnjem razdelku prikazane prednosti in slabosti zveznih modelov.

## 2 Hopfieldov zvezni model nevronske mreže

Hopfield (1984) za zvezni model uporablja isto topologijo kot pri diskretni nevronske mreži, t.j. vsak nevron je povezan z vsakim. Kombinacijska funkcija diskretnega modela je podana z enačbo (Hopfield 1982):

$$A_j = \sum_{S_i=1, i \neq j} T_{ji} + I_j = \sum_{i \neq j} S_i T_{ji} + I_j \quad (8)$$

V Hopfieldovem modelu so  $T_{ji}$  elementi spominske matrike, dobljene kot vsota zunanjih produktov učnih vzorcev, ki so popravljeni tako, da so vrednosti komponent 0 spremenjene v -1.  $I_j$  predstavlja konstanten vhod v  $j$ -ti nevron. Model BNM brez povratnih povezav ( $i \neq j$ ) ustreza spominski matriki z ničelno diagonalo ( $T_{ii} = 0$  za vse  $i = 1 \dots N$ ).  $S_i$  ima lahko vrednosti 0 in 1 kot v originalnem Hopfieldovem modelu, čeprav učno pravilo Hopfieldovega modela uporablja vrednosti -1 in 1.

Kombinacijska funkcija zveznega Hopfieldovega modela je analogna. Dinamika zveznega modela je opisana z naslednjima enačbama (Hopfield 1984):

$$C_j \frac{du_j}{dt} = \sum_{i, i \neq j} T_{ji} S_i - \frac{u_j}{R_j} + I_j = A_j - \frac{u_j}{R_j} \quad (9)$$

$$S_j = g_j(u_j) \quad (10)$$

kjer je  $S_i$  stanje oziroma izhod  $i$ -tega nevrona,  $u_j$  vhod (aktivacijski nivo)  $j$ -tega nevrona in  $I_j, R_j, C_j$  konstante.  $g_j$  je izhodna funkcija, ki definira relacijo med vhomom in izhodom, in je odvedljiva in sigmoidna z asimptotama 0 in 1. Iz enačbe (9) sledi, da je hitrost spremembe  $u_j$  proporcionalna razliki med trenutnim in novim  $u_j$ , izračunanem po enačbi (8). V originalni Hopfieldovi (1984) enačbi je izpuščen pogoj  $i \neq j$ , ker za Hopfieldovo pravilo učenja velja, da so vsi  $T_{ii} = 0$ .

Stabilnost zveznega modela pokaže Hopfield tako, da definira "energijsko" funkcijo stanja sistema:

$$\begin{aligned} E &= -\frac{1}{2} \sum_j \sum_{i, i \neq j} S_j S_i T_{ji} - \sum_j I_j S_j \\ &\quad + \sum_j \frac{1}{R_j} \int_0^{S_j} g_i^{-1}(S) dS \\ &= -E1 + \sum_j \frac{1}{R_j} \int_0^{S_j} g_i^{-1}(V) dV \end{aligned} \quad (11)$$

ki s časom monotono pada, ker velja:

$$\frac{dE}{dt} \leq 0 \quad (12)$$

in

$$\frac{dE}{dt} = 0 \rightarrow \forall j : \frac{dS_j}{dt} = 0$$

### 3 Zvezna Bayesova nevrnska mreža, ki temelji na verjetnosti

Informacijski prispevek  $i$ -tega nevrona za aktivnost  $j$ -tega nevrona je v diskretni BNM-p

dobimo z logaritmiranjem enačbe (2) in je podan z  $\log_2 \frac{P(S_j|S_i)}{P(S_j)}$ . Bolj splošna definicija informacijskega prispevka je podana z

$$S_i \times \log_2 \frac{P(S_j|S_i)}{P(S_j)}$$

kjer  $S_i$  predstavlja trenutno stanje  $i$ -tega nevrona.  $S_i$  lahko zavzema poljubno vrednost na intervalu  $[0,1]$  in ustreza verjetnosti, da je  $i$ -ti nevron aktiven. Posplošen logaritem enačbe (2) vsebuje vsoto po vseh nevronih in ne samo po aktivnih nevronih:

$$\begin{aligned} -\log_2 P(S_j|S_1, \dots, S_N) &= \\ -\log_2 P(S_j) - \sum_{i \neq j} \left( S_i \times \log_2 \frac{P(S_j|S_i)}{P(S_j)} \right) \end{aligned} \quad (13)$$

Z antilogaritmiranjem dobimo kombinacijsko funkcijo zvezne BNM-p, ki je posplošena enačba (2):

$$P(S_j|S_1, \dots, S_N) = P(S_j) \prod_{i \neq j} \left( \frac{P(S_j|S_i)}{P(S_j)} \right)^{S_i} \quad (14)$$

Posplošitev pogojne verjetnosti  $P(S_j|S_i = 1)$ , pri negotovi evidenci, da je  $S_i = 1$ , je podana z  $S_i \times P(S_j|S_i = 1)$  (Ihara 1987). Ta posplošitev vodi do drugačne definicije enačbe (14), kjer lahko več nevronov predstavlja isti atribut. Enačba (14) je smiselna posplošitev, če se vplivi različnih nevronov, čeprav mogoče predstavljajo isti atribut, obravnavajo neodvisno.

Če v enačbi (13) namesto  $-\log_2 P(S_j|S_1, \dots, S_N)$  pišemo  $A_j$ , namesto  $-\log_2 P(S_j)$  pišemo  $I_j$  in namesto  $\frac{P(S_j|S_i)}{P(S_j)}$  pišemo  $T_{ji}$ , dobimo enačbo (8), t.j. kombinacijsko funkcijo Hopfieldovega modela. Zato lahko za opis dinamike zvezne BNM-p uporabimo enačbi (9) in (10). Dokaz stabilnosti je ekvivalenten Hopfieldovemu dokazu (glej razdelek 2). Posplošena funkcija podobnosti (7) na zvezna stanja je torej:

$$Sim(S_1, \dots, S_N) = \prod_j \left( \frac{P(S_j | S_1, \dots, S_N)}{P(S_j)} \right)^{S_j} \quad (15)$$

#### 4 Zvezna Bayesova nevronska mreža, ki temelji na razmerju verjetnosti

Težo evidence stanja  $S_i$   $i$ -tega nevrona za aktivnost  $j$ -tega nevrona v diskretni BNM-odds dobimo z logaritmiranjem enačbe (5) in je podana z  $\log_2 \frac{odds(S_j=1|S_i)}{odds(S_j=1)}$ . Bolj splošna definicija teže evidence je podana z

$$S_i \times \log_2 \frac{odds(S_j = 1 | S_i = 1)}{odds(S_j = 1)} + (1 - S_i) \times \log_2 \frac{odds(S_j = 1 | S_i = 0)}{odds(S_j = 1)}$$

kjer  $S_i$  predstavlja trenutno stanje  $i$ -tega nevrona.  $S_i$  lahko zavzema poljubno vrednost na intervalu  $[0, 1]$  in ustreza verjetnosti, da je  $i$ -ti nevron v stanju 1 (oziroma 1 minus verjetnosti, da je nevron v stanju 0). Posplošena kombinacijska funkcija (5) za zvezne BNM-odds je sledeča:

$$\begin{aligned} odds(S_j | S_1, \dots, S_N) &= \\ &= odds(S_j = 1) \prod_{i \neq j} \left[ \left( \frac{odds(S_j = 1 | S_i = 1)}{odds(S_j = 1)} \right)^{S_i} \times \right. \\ &\quad \left. \left( \frac{odds(S_j = 1 | S_i = 0)}{odds(S_j = 1)} \right)^{1-S_i} \right] \quad (16) \end{aligned}$$

Ustrezna posplošena funkcija podobnosti je pri zvezni BNM-odds različna od funkcije podobnosti zvezne BNM-P:

$$\begin{aligned} Sim(S_1, \dots, S_n) &= \\ &= \prod_j \left[ \left( \frac{P(S_j = 1 | S_1, \dots, S_n)}{P(S_j = 1)} \right)^{S_j} \times \right. \\ &\quad \left. \left( \frac{P(S_j = 0 | S_1, \dots, S_n)}{P(S_j = 0 | S_1, \dots, S_n)} \right)^{1-S_j} \right] \quad (17) \end{aligned}$$

Če enačbo (16) logaritmiramo in namesto  $-\log_2 odds(S_j | S_1, \dots, S_n)$  pišemo  $A_j$ , namesto  $-\log_2 odds(S_j)$  pišemo  $I_j$  in namesto  $-\log_2 \frac{P(S_j=X|S_i=Y)}{P(S_j=X)}$  pišemo  $T_{ji}^{XY}$ , dobimo:

$$A_j = I_j + \sum_{i \neq j} (S_i(T_{ji}^{11} - T_{ji}^{01}) + (1 - S_i)(T_{ji}^{10} - T_{ji}^{00})) \quad (18)$$

Če zgoraj definirani aktivacijski nivo nevrona  $A_j$  uporabimo v enačbi (9) dobimo enačbo za opis dinamike zvezne BNM-odds. Ker nova enačba ni enaka enačbi za zvezno BNM-P, dokaz stabilnosti ni tako očiten.

Iz dokaza stabilnosti Hopfieldovega modela podanega v razdelku 2 je razvidno, da je za stabilnost pri enačbi (11) zadosten pogoj:

$$\frac{dE_1}{dt} = \sum_j \left( \frac{dS_j}{dt} A_j \right) \quad (19)$$

Če v enačbi (11) uporabimo za  $E_1$  naslednji izraz:

$$E_1(S_1, \dots, S_N) = \sum_j I_j S_j +$$

$$+ \frac{1}{2} \sum_j \sum_{i \neq j} (S_j S_i T_{ji}^{11} + S_j (1 - S_i) T_{ji}^{10} +$$

$$(1 - S_j) S_i T_{ji}^{01} + (1 - S_j)(1 - S_i) T_{ji}^{00}) \quad (20)$$

potem je pogoj (19) izpolnjen, ker velja:

$$\begin{aligned} \frac{dE_1}{dt} &= \sum_j \left( \frac{dE_1}{dS_j} \frac{dS_j}{dt} \right) \\ &= \frac{1}{2} \sum_j \left( \frac{dS_j}{dt} A_j \right) + \frac{1}{2} \sum_i \left( \frac{dS_i}{dt} A_i \right) \\ &= \sum_j \left( \frac{dS_j}{dt} A_j \right) \quad (21) \end{aligned}$$

S tem je stabilnost zveznega modela BNM-odds dokazana.

## 5 Zaključki

Bayesove nevronske mreže so lahko diskretne ali zvezne ter lahko temeljijo na verjetnosti ali na razmerju verjetnosti. Razlika med Hopfieldovim modelom in BNM je v učnem pravilu. Za učenje BNM zadošča osnovno Hebbovo pravilo (Hebb 1949), za katerega je precej neurofiziološke evidence, da se po tem pravilu učijo biološki nevroni. Uteži na sinapsah ustrezajo apriornim verjetnostim, da sta povezana nevrna aktivna. Hopfieldov model uporablja posplošeno Hebbovo pravilo. Uteži na sinapsah ustrezajo razliki med verjetnostjo, da sta povezana nevrna v istem stanju, in verjetnostjo, da sta povezana nevrna v različnem stanju. Zato uteži Hopfieldovega modela vsebujejo manj informacije kot uteži pri BNM. Kompleksnost učenja in izvajanja je pri obeh modelih enaka, po klasifikacijski pravilnosti pa BNM daleč presega Hopfieldov model (Kononenko 1990).

Za opis stanja modela Hopfield uporablja energijsko funkcijo, ki je analogna funkciji podobnosti za opis stanja BNM. Logaritem funkcije podobnosti lahko interpretiramo kot entropijo ali informacijsko vsebino sistema. Iz tega sledi, da je fiksna točka stanje sistema z lokalno minimalno entropijo.

Prednost zveznih BNM je v večji fleksibilnosti predstavitve znanja. Vhodni podatki so lahko "mehki" in nezanesljivi. Dani primer ima lahko več vrednosti za isti atribut, ki jim je možno določiti zaupanje. Tudi odgovor zvezne BNM je zato lahko bolj fleksibilen.

BNM-P ima prednost pred BNM-odds, ker se lahko entropija direktno posploši na več kot dva disjunktna dogodka. BNM-P je zato bolj prožna in računsko manj zahtevna. BNM-odds je primerna samo za binarne attribute in razrede in je manj primerna za predstavitev manjkajočih podatkov. V diskretni BNM-odds je lahko manjkajoči podatek predstavljen z apriornim razmerjem verjetnosti (*odds*) samo med učenjem, medtem ko je v zvezni BNM-odds lahko predstavljen z apriornim

razmerjem verjetnosti, tako med učenjem kot med izvajanjem. Interpretacija izvajanja BNM-P je bolj naravna (prispevek informacije v bitih) kot interpretacija BNM-odds (teža evidence). Po drugi strani pa BNM-odds potrebuje bistveno manj nevronov za predstavitev istega problemskega prostora kot BNM-P.

Z nadaljnjimi raziskavami bo potrebno eksperimentalno testiranje zveznih BNM. Potrebno bo preveriti prednosti zveznih BNM na realnih problemih. Poleg bolj fleksibilne predstavitve znanja omogočajo zvezne BNM, z *asinhronim* izvajanjem, ki služi kot *globalna informacija*, bolj zanesljivo delovanje. Namreč zaradi asinhronosti bodo nevroni, katerih spremembe stanj so bolj zanesljive, prej spremenili svoje stanje, kar bo vplivalo na spremembe stanj ostalih nevronov, preden le-ti uspejo spremeniti svoje stanje. Spremembe niso diskretne (torej niso drastične) ampak majhne z majhnimi vplivi in ne prehudimi posledicami.

## Reference

- Hebb, D.O. (1949) *The Organization of Behavior*. New York: Wiley.
- Hopfield J.J. (1982) Neural networks and physical systems with emergent collective computational abilities. *Proc. of the National Academy of Sciences* 79:2554-2558.
- Hopfield J.J. (1984) Neurons with graded response have collective computational properties like those of two-state neurons. *Proc. of the National Academy of Sciences* 81:4586-4590.
- Ihara J. (1987) Extension of conditional probability and measures of belief and disbelief in a hypothesis based on uncertain evidence. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, Vol. PAMI-9, no.4, pp. 561-568.
- Kononenko, I. (1990) Bayesove umetne nevronske mreže, *Informatika*, letnik 14, št. 3, str.72-86.