

Paslėptas Markovo Modelis

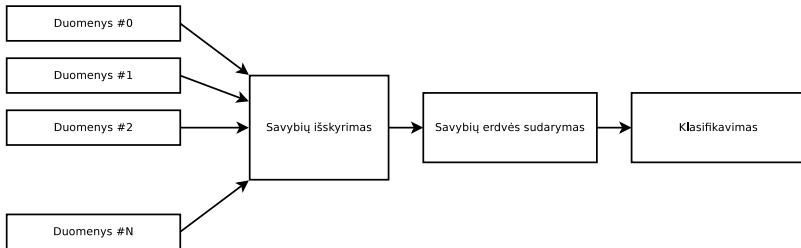
Ižanga

Maksim Norkin, ISK-08

Vilniaus Gedimino Technikos Universitetas
Elektronikos fakultetas

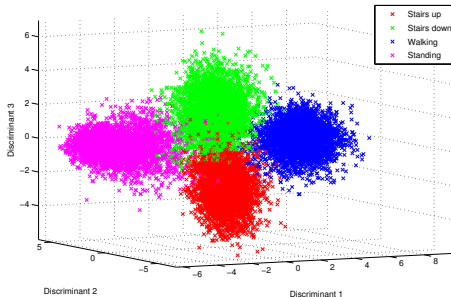
2011 m. gruodžio 17 d.

Dirbtinio intelekto sistemos apžvalga



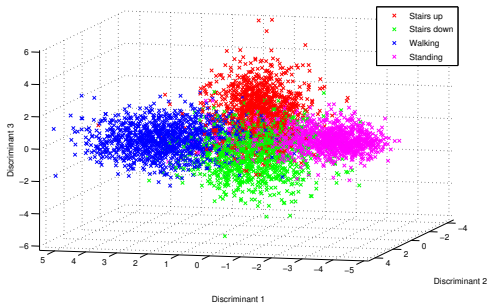
Savybių erdvė I

- Duomenų grupės savybių erdvėje gerai viena nuo kitos atskirtos;



Savybių erdvė II

- Duomenų grupės savybių erdvėje blogai viena nuo kitos atskirtos;



Mašininis apmokymas I

- Vadovu paremtas;
 - Klasifikavimas;
 - Duomenys yra sužymėti;
- Vadovu neparemtas;
 - Grupavimas;
 - Duomenys yra visiškai nesužymėti;
- Dalinai vadovu paremtos;
 - Maža duomenų dalis yra sužymėti, didelė dalis duomenų yra nesužymėti;

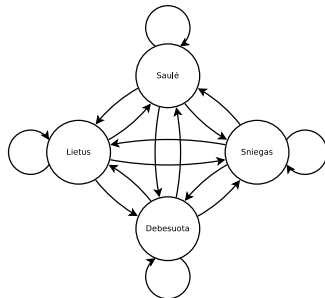
Diskretinis Markovo procesas I

Stebimas Markovo Modelis

$$P[q_t = S_j | q_{t-1} = S_i, q_{t-2} = S_k, \dots] \quad (1)$$

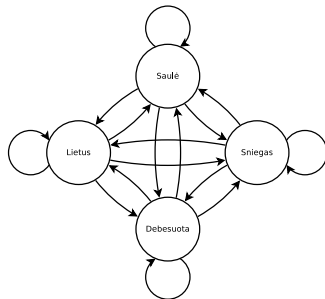
$$= P[q_t = S_j | q_{t-1} = S_i] \quad (2)$$

$$a_{ij} = P[q_t = S_j | q_{t-1} = S_i] \quad (3)$$



Diskretinis Markovo procesas II

- Ką mes galime nuveikti su duotu modeliu, žinodami perėjimo tikimybes?
 - Spėti ateinančių dienų orus;
 - Spėti kiek dienų truks saulėtą orą;



Diskretinis Markovo procesas III

- Monetos metimo eksperimentas;
- Žinomas tik eksperimento rezultatas;

$$\mathbf{O} = O_1 O_2 O_3 O_4 \cdots O_T = \mathfrak{H} \mathfrak{H} \mathfrak{H} \mathfrak{T} \mathfrak{H} \cdots \mathfrak{T} \quad (4)$$

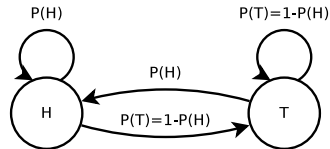
- Klausimas - kokia Markovo grandimi mes galima pakartoti sekantį eksperimentą?

Diskretinis Markovo procesas IV

- Vienos monetos grandis;

$$\mathbf{O} = \mathfrak{H}\mathfrak{H}\mathfrak{T}\mathfrak{T} \cdots \mathfrak{T} \quad (5)$$

$$\mathbf{S} = HHTT \cdots T \quad (6)$$



Diskretinis Markovo procesas V

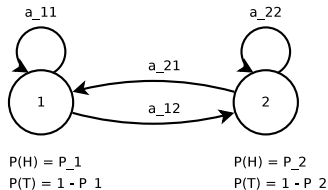
- Dviejų monetų grandis;

$$\mathbf{O} = \mathfrak{H}\mathfrak{H}\mathfrak{T}\mathfrak{T} \dots \mathfrak{T}$$

$$\mathbf{S} = 1221 \dots 1$$

(7)

(8)

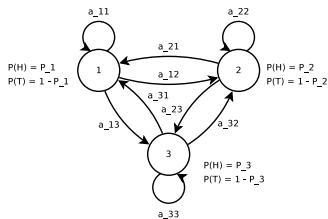


Diskretinis Markovo procesas VI

- Tryjų monetų grandis;

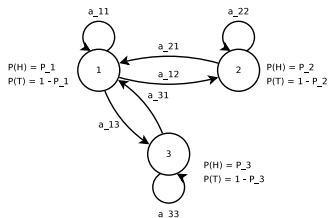
$$\mathbf{O} = \mathfrak{H}\mathfrak{H}\mathfrak{T}\mathfrak{T} \cdots \mathfrak{T} \quad (9)$$

$$\mathbf{S} = 1323 \cdots 1 \quad (10)$$



Diskretinis Markovo procesas VII

- Markovo grandis projektuojama remiantis konkrečiomis žiniomis apie modeliujama sistemą;
- Anksčiau paminėtos grandinės yra pilnai sujungtos (ergodinės);
- Grandinės struktūra yra labai lanksti ir nebūtinai turi būti pilnai sujungta;



Paslėptas Markovo modelis I

Modelis aprašomas:

- N - būsenų skaičius (arba grandines mazgų skaičius);
- M - būsenos stebėjimo simbolių skaičius (arba įėjimo dimensijų skaičius);
- $A = \{a_{ij}\}$ - būsenų perėjimo tikimybių matrica;

$$a_{ij} = P[q_{t+1} = S_j | q_t = S_i], \quad 1 \leq i, j \leq N \quad (11)$$

$$\sum_{j=1}^N a_{ij} = 1 \quad (12)$$

Paslėptas Markovo modelis II

- $B = \{b_j(k)\}$ - stebėjimo tikimybės pasiskirstymas, esant būsenai j

$$b_j(k) = P[v_k \text{ at } t | q_t = S_j], \quad 1 \leq j \leq N, \quad 1 \leq k \leq M \quad (13)$$

- $\pi = \{\pi_i\}$ pradinis būsenų tikimybių pasiskirstymas

$$\pi_i = P[q_1 = S_i] \quad (14)$$

- Trumpai, Paslėptas Markovo modelis aprašomas:

$$\lambda = (\pi, A, B); \quad (15)$$

Viterbi kelias I

- Viterbi kelio algoritmas yra naudojamas, norint sužinoti kokios sistemos būsenos $Q = \{q_1 q_2 q_3 \cdots q_T\}$ yra labiausiai tikėtinos, turint stebėjimo rezultatus $O = \{O_1 O_2 \cdots O_T\}$
- Reikia įsivesti naują parametą, kuris nusakytų didžiausia būsenos tikimybę laiko momentu t

$$\delta_t(i) = \max_{q_1, q_2, \dots, q_{t-1}} P[q_1 q_2 \cdots q_t = i, O_1 O_2 \cdots O_t | \lambda] \quad (16)$$

Viterbi kelias II

■ Iniacija

$$\delta_1(i) = \pi_i b_i(O_1), \quad 1 \leq i \leq N \quad (17)$$

$$\psi_1(i) = 0 \quad (18)$$

■ Rekursija

$$\delta_t(j) = \max_{1 \leq i \leq N} [\delta_{t-1} a_{ij}] b_j(O_t), \quad 2 \leq t \leq T, \quad 1 \leq j \leq N \quad (19)$$

$$\psi_t(j) = \arg \max_{1 \leq i \leq N} [\delta_{t-1}(i) a_{ij}], \quad 2 \leq t \leq T, \quad 1 \leq j \leq N \quad (20)$$

Viterbi kelias III

- Nutraukimas

$$q_T^* = \arg \max_{1 \leq i \leq N} [\delta_T(i)] \quad (21)$$

- Kelio atstatymas

$$q_t^* = \psi_{t+1}(q_{t+1}^*), \quad t = T - 1, T - 2, \dots, 1 \quad (22)$$

- Plačiai naudojamas garso atpažinimo sistemose;
- Galima klasifikuoti duomenis, kurie savybių erdvėje yra visiškai neatsiskyre;
- Yra daug plėtinių (Hierarchinis Paslėptas Markovo modelis, ...);
- Matlab implementacija parašyta Kevin Murphy, 1998 metais (palaiko ne tik diskretinius, bet ir vientisus stebėjimus);
- Geras straipsnis gilinti savo žinias parašytas Lawrence R. Rabiner, 1989 metais;

Ačiū už dėmesį!

Klausimai?