

Semestrální zkouška CZR/ZRE, řádný termín, 14.5.2007, skupina A

Login:

Podpis:

1. Při výpočtu koeficientů prediktoru pomocí Levinsona-Durbina vyšly koeficienty prediktoru 3. řádu takto:

$$a_1^{(3)} = 0.1145, \quad a_2^{(3)} = -0.0470, \quad a_3^{(3)} = -0.271$$

Ve 4. kroku vyšel $k_4 = 0.0859$. Vypočtete koeficient $a_3^{(4)}$

A: $a_3^{(4)} = -0.051$

B: $a_3^{(4)} = 0.051$

C: $a_3^{(4)} = -0.261$

D: $a_3^{(4)} = 0.261$

2. IIR filtr prvního řádu:

$$H(z) = \frac{1}{1 - 0.65z^{-1}}$$

měl minulý vzorek výstupu $y[-1] = 4$. Jaká je odezva tohoto filtru naprázdno, tedy pro $x[n] = 0$ pro $n \geq 0$? Zajímáme se pouze o vzorky $y[n]$ pro $n \geq 0$.

A: $y[0] = 2.6$, pak nějaké další výstupy.

B: $y[0] = -2.6$, pak nějaké další výstupy.

C: $y[n] = 2.6$ až do $n = \infty$.

C: $y[n] =$ samé nuly až do $n = \infty$.

3. Řečový signál má (od $n = 0$) hodnoty: $x = [1 \ 2 \ 3 \ 4 \ 5 \ 6 \ 7 \ 8]$. Rámce mají délku $N = 4$. Vypočtete normalizovaný cross-korelační koeficient $CCF(2)$ pro rámec, který začíná na $zr = 2$.

A: $CCF(2) = 0.7844$

B: $CCF(2) = 0.9844$

C: $CCF(2) = 1.1844$

D: $CCF(2) = 1.3844$

4. Která z technologií zpracování řeči nejvíce podléhá standardizaci ?

A: Rozpoznávání mluvcího

B: Audiovizuální rozpoznávání řeči

C: Identifikace jazyka

D: Kódování

5. V algoritmu hodnocení kvality kódované řeči PESQ (Perceptual Evaluation of Speech Quality) najdeme funkční blok "Time align and equalize" (časové srovnání a ekvalizace). Tento blok je důležitý zvláště u kodérů pro IP telefonii, protože:

A: Mluvcí používají méně kvalitní mikrofony a sluchátka než u klasické telefonní sítě.

B: Jedná se o silně multilinguální prostředí.

C: Kompenzuje se tak výpadek paketů či jejich příchod v nesprávném pořadí.

D: Kompenzuje se tak tzv. základní audio-šum Internetu (ekvivalent kosmického záření).

6. Přenosová funkce dlouhodobého prediktoru (LTP) je $-bz^{-L}$. V případě správného fungování LTP má L hodnotu:

- A: periody základního tónu ve vzorcích.
- B: periody základního tónu v ms.
- C: počtu formantů.
- D: obvykle 10.

7. V kodéru RPE-LTP (GSM Full Rate) je signál po průchodu krátkodobým LPC filtrem a dlouhodobým LTP filtrem kvantován:

- A: Není kvantován.
- B: Tento signál se v kodéru vůbec nenachází.
- C: Pomocí kódové knihy.
- D: Pomocí sekvence pravidelně se opakujících impulsů.

8. Jaká je funkce perceptuálního filtru $W(z)$ v kódování CELP ?

- A: Potlačení spektra chyby v regionech se silnými frekvenčními složkami řečového signálu.
- B: Potlačení spektra chyby v regionech bez přítomnosti silných frekvenčních složek řečového signálu.
- C: Potlačení harmonické složky v řečovém signálu (násobky frekvence základního tónu).
- D: "Vybělení" spektra chybového signálu.

9. K výpočtu DTW máte k dispozici neúplnou mřížku lokálních vzdáleností **D** a neúplnou mřížku částečných kumulovaných vzdáleností **G** (bez hodnot ∞ v nultém řádku a nultém sloupci a hodnoty 0 vlevo dole).

		D				
reference		10	5	2		
		test				

		G				
reference						
		20	15	18		
		16	10	12		
		8	13	5		
		test				

Doplňte poslední řádek mřížky **G**:

- A: 30 20 14
- B: 30 25 14
- C: 30 20 19
- D: 30 25 19

10. Jaký je důvod zavedení hodnot ∞ do levého sloupce a spodního řádku matice částečných kumulovaných vzdáleností u DTW ?

- A: Inicialisace (seed) generátoru náhodných čísel při výpočtu vzdálenosti mezi dvěma vektory.
- B: Zamezení začátku srovnání na jiném místě, než je začátek referenční a testovací promluvy.
- C: Hodnoty ∞ jsou nutné pro normalizaci (scaling) vypočtených částečných kumulovaných vzdáleností.
- D: Pomocí těchto hodnot se určuje složitost algoritmu DTW.

-
11. Ergodický HMM (EHMM) je charakteristický tím, že má všechny vysílací stavy navzájem propojené. Máme k dispozici EHMM s 5-ti vysílacími stavy, který má zpracovat promluvu o 5 vektorech. Určete, které ze stavových sekvencí X jsou povolené:

$$X_1 = [1 \ 2 \ 2 \ 2 \ 2 \ 2 \ 7] \quad X_2 = [1 \ 6 \ 5 \ 4 \ 3 \ 2 \ 7] \quad X_3 = [1 \ 2 \ 3 \ 4 \ 5 \ 6 \ 7] \quad X_4 = [1 \ 6 \ 3 \ 6 \ 3 \ 6 \ 7]$$

A: žádná ze sekvencí není povolena

B: pouze X_3 je povolena

C: povolena je pouze X_1 – proces musí zůstat pouze v jednom vysílacím stavu.

D: všechny sekvence jsou povolené.

12. Při trénování 6-stavového HMM (4 vysílací stavy) jsme zjistili, že state occupation functions 45. akustického vektoru mají pro jednotlivé vysílací stavy následující hodnoty:

$$L_2(45) = 0.5, \quad L_3(45) = 0.1, \quad L_4(45) = 0.05, \quad L_5(45) = 0.6$$

A: Takové hodnoty značí přetrénování modelu.

B: Algoritmus je zřejmě naprogramován chybně.

C: Vše je v pořádku, vysoké hodnoty u dvou stavů znamenají, že se zřejmě jedná o akusticky podobný úsek (např. 'a' ve slově "avia").

D: Model je natrénován, algoritmus je možné ukončit.

13. Skrytý Markovův model má 4 stavy, z toho 2 vysílací. Obsahuje Gaussovky s diagonálními kovariančními maticemi. Stav 2 má následující vektor středních hodnot a vektor směrodatných odchylek: $\boldsymbol{\mu} = [1, 1]^T$ a $\boldsymbol{\sigma} = [4, 6]^T$ (pro úsporu místa jsou vektory zapsány jako řádkové a pak transponovány). Určete hodnotu funkce hustoty rozdělení pravděpodobnosti ("vysílací pravděpodobnost") pro vektor: $\mathbf{o} = [0, 0]^T$

A: $b_2(\mathbf{o}) = 0.0003$

B: $b_2(\mathbf{o}) = 0.0023$

C: $b_2(\mathbf{o}) = 0.0043$

D: $b_2(\mathbf{o}) = 0.0063$

14. Skrytý Markovův model má 4 stavy, z toho 2 vysílací. Jeho matice logaritických přechodových pravděpodobností je:

$$\mathbf{A} = \begin{bmatrix} -\infty & 0 & -\infty & -\infty \\ -\infty & -0.11 & -2.30 & -\infty \\ -\infty & -\infty & -0.22 & -1.60 \\ -\infty & -\infty & -\infty & -\infty \end{bmatrix}$$

Model má na vstupu sekvenci 17 vektorů. Při výpočtu logaritické Viterbiho pravděpodobnosti jsou v posledním platném čase $t = 17$ ve vysílacích stavech následující hodnoty tokenů (piv): $token_2(17) = -123.45$, $token_3(17) = -234.56$. Jaká je celková hodnota logaritické Viterbiho pravděpodobnosti?

A: $\log P^*(\mathbf{O}|M) = -234.78$

B: $\log P^*(\mathbf{O}|M) = -123.67$

C: $\log P^*(\mathbf{O}|M) = -236.16$

D: $\log P^*(\mathbf{O}|M) = -125.05$

15. Rozpoznávač izolovaných slov má rozpoznávat 10 číslovek, které mají dohromady 16 různých výslovnostních variant (např. "čtyři", "štyry", atd.). Každá varianta má svůj vlastní 20-stavový HMM. Kolik bude potřeba vyhodnocovat Viterbiho pravděpodobností.

A: 1

B: 16

C: 18

D: 20

16. Systémem rozpoznávání spojitě řeči s velkým slovníkem (LVCSR) byl zpracován záznam zpráv. Chceme vyhledat všechny výskyty slova “dýně”. Toto slovo však není obsaženo ve slovníku rozpoznávače a slovník není možné měnit.

A: Budeme sledovat likelihood jednotlivých rozpoznávaných slov a kdekoliv bude nízká, necháme uživatele poslechnout tento segment, zda zde není obsaženo hledané slovo.

B: Zadáme do Google hledané slovo a uděláme si statistiky slov ve vrácených odkazech. Pokud najdeme shluk takových slov ve výstupu rozpoznávače, určíme zde místo možného výskytu hledaného slova.

C: Přeházíme v hledaném slově fonémy a některé přidáme, např. na “hnědý”. Toto slovo budeme pak vyhledávat.

D: Spustíme rozpoznávač několikrát na různých architekturách CPU, především s různými hodinovými frekvencemi.

17. Fonetické znalosti o místu artikulaci, znělosti/neznělosti hlásky, atd. se při rozpoznávání spojitě řeči pomocí kontextově závislých modelů:

A: Používají pro natrénování modelu místa artikulace, znělosti/neznělosti, atd.

B: Nepoužívají nikde.

C: Používají při tvorbě otázek, které pokládáme při vázání (tying) jednotlivých stavů u kontextově závislých modelů.

D: Používají pro prvotní odhady přechodových pravděpodobností v modelech.

18. Dynamické koeficienty (Δ , $\Delta\Delta$) se v rozpoznávání řeči používají pro:

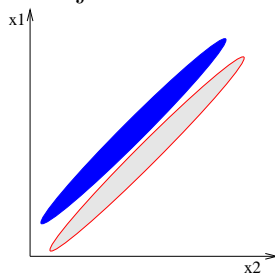
A: Větší vytížení procesoru na základě kartelové dohody s výrobcí CPU (rozpoznávače by jinak vyžadovaly příliš málo výkonné procesory).

B: Dekorelaci základních parametrů.

C: Zachycení frekvenčního kontextu mezi jednotlivými pásmy při výpočtu MFCC.

D: Zachycení časového kontextu okolo právě zpracovávaného vektoru.

19. Řečová data charakterizujeme dvouprvkovými vektory parametrů. Chceme klasifikovat 2 třídy, jejichž vektory mají ve 2D prostoru následující rozdělení.



Na datech natrénujeme lineární diskriminační analýzu (LDA). Jaký bude směr jejího prvního bázového vektoru ?

A. $\mathbf{b}_1 = [-1, 1]^T$

B. $\mathbf{b}_1 = [1, 1]^T$

C. $\mathbf{b}_1 = [1, -1]^T$

D. $\mathbf{b}_1 = [0, 1]^T$

20. Pro stejné vektory jako v předcházejícím příkladu použijeme výpočet bází pomocí principal component analysis (PCA). Každý vektor promítneme do **prvního** bázového vektoru PCA, takže výsledkem bude skalár. Tento budeme klasifikovat. Takový klasifikátor:

A. Bude velmi špatný - přesnost se bude pohybovat maximálně okolo 50% správně klasifikovaných vektorů.

B. Bude velmi dobrý - téměř všechny vektory budou klasifikovány dobře.

C. Zadání je nesmysl - klasifikátor, který pracuje s jediným skalárem, není možné zkonstruovat.

D. Pro data, jejichž rozdělení má tvar elipsy, není možné zkonstruovat žádný klasifikátor.