

Algoritmy a struktury neuropočítačů

ASN – P10

Klasifikace a predikce *Classification, prediction*

Klasifikace - sdružování dat do tříd → počet tříd známe

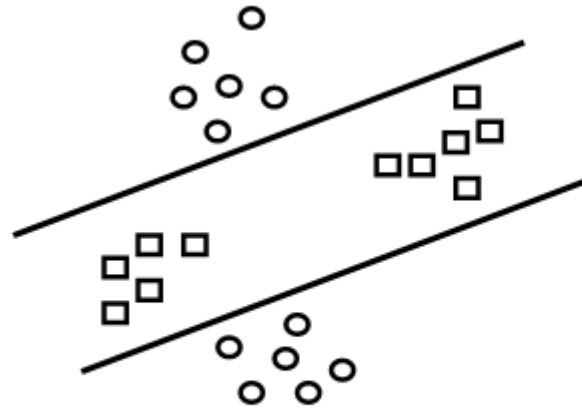
- parametry vyjadřují podobné vlastnosti objektu
- třídy = shluky (*clusters*) → shluková analýza
- typ klasifikace: a) s učitelem (*supervised*)
b) bez učitele (*unsupervised*)

Klasifikátory založené na umělých neuronových sítích

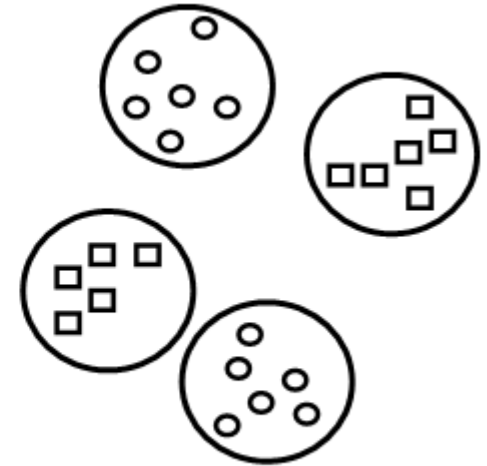
Klasifikace s učitelem - porovnání neznámých vzorů se všemi známými vzory

Kritérium: stupeň podobnosti nebo vzdálenosti v nějaké metrice

Klasifikace do 2 tříd



a)



b)

pomocí MLNN se sigmoidální aktivační funkcí

pomocí RBFN


Radial basic function

Vybrané metody:

- metoda K-nejbližších sousedů (*hranice nejsou jednoznačné*)
K-Nearest-Neighbor Method, KNN
- klasifikace vzorků podle úhlu
podobnost určuje „směr“ vektorů

směrový kosinus (*direction cosine*)

$$\cos \Theta_k = \frac{(x, x_k)}{\|x\| \|x_k\|}$$

nejlepší shoda (*best match*)

$$\cos \Theta_k = 1$$

apriorní informace - porovnání jednotlivých vzorků se všemi referenčními daty a určení nejlepšího z nich z hlediska podobnosti resp. vzdálenosti **vítěz v soutěži** (*winner*)

rozhodovací funkce (*discriminant function*)
způsob přiřazení jednotlivých vzorů do tříd

pro $x \in S_i$ **a** $j \neq i$ **je** $\delta_i(x) \geq \delta_j(x)$

hranice třídy, hranice rozdělovacího prostoru mezi třídami:

platí-li $\delta_i(x) - \delta_j(x) = 0$ pak pro $x \in S_c^i$ je

$$\delta_c(x) = \max_i \{ \delta_i(x) \}$$

Klasifikace bez učitele - vzájemné vztahy mezi vzory se určují shlukováním
- neznáme počet tříd

jednoduché metody shlukování - podle vzdálenosti

hierarchické metody shlukování (*hierarchical methods*)

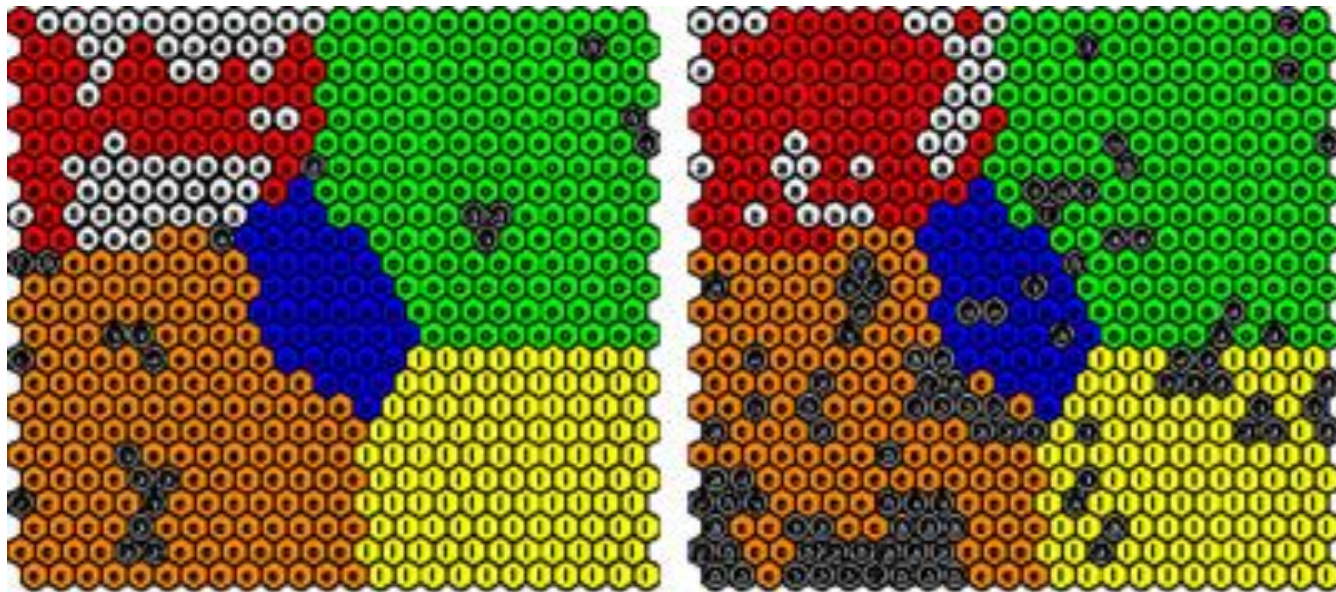
menší výpočetní náročnost

metody: **aglomerativní** - slučování podtříd (*merging methods*)

štěpící (rozdělující) - jednu velkou třídu dělí na menší podtřídy (*splitting methods*)

Příklady klasifikace

Umělé neuronové sítě - Kohonenovy samoorganizující se mapy

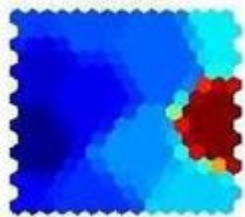


Zdravé dítě

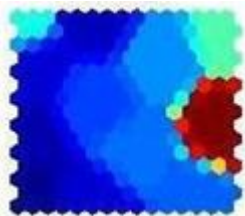
Nemocné dítě

Aplikace KSOM

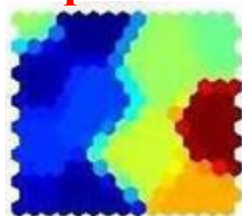
Max.úhel



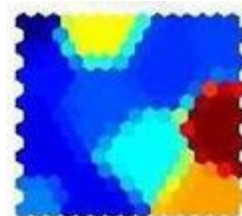
Min.úhel



**Perioda
opakování**

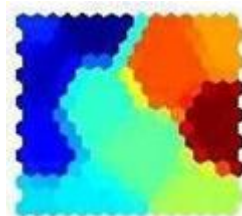


Frekvence

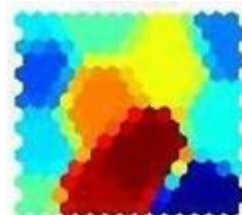
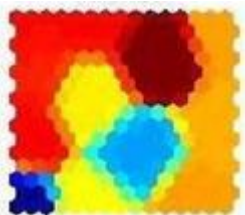


Flexe a extenze v lokti

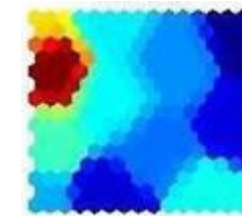
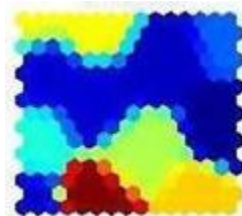
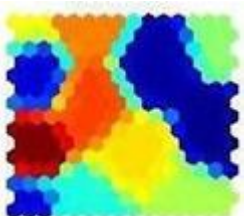
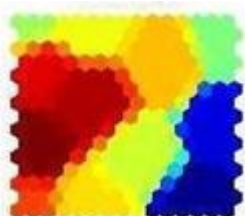
Žlutá = pacient s CMP



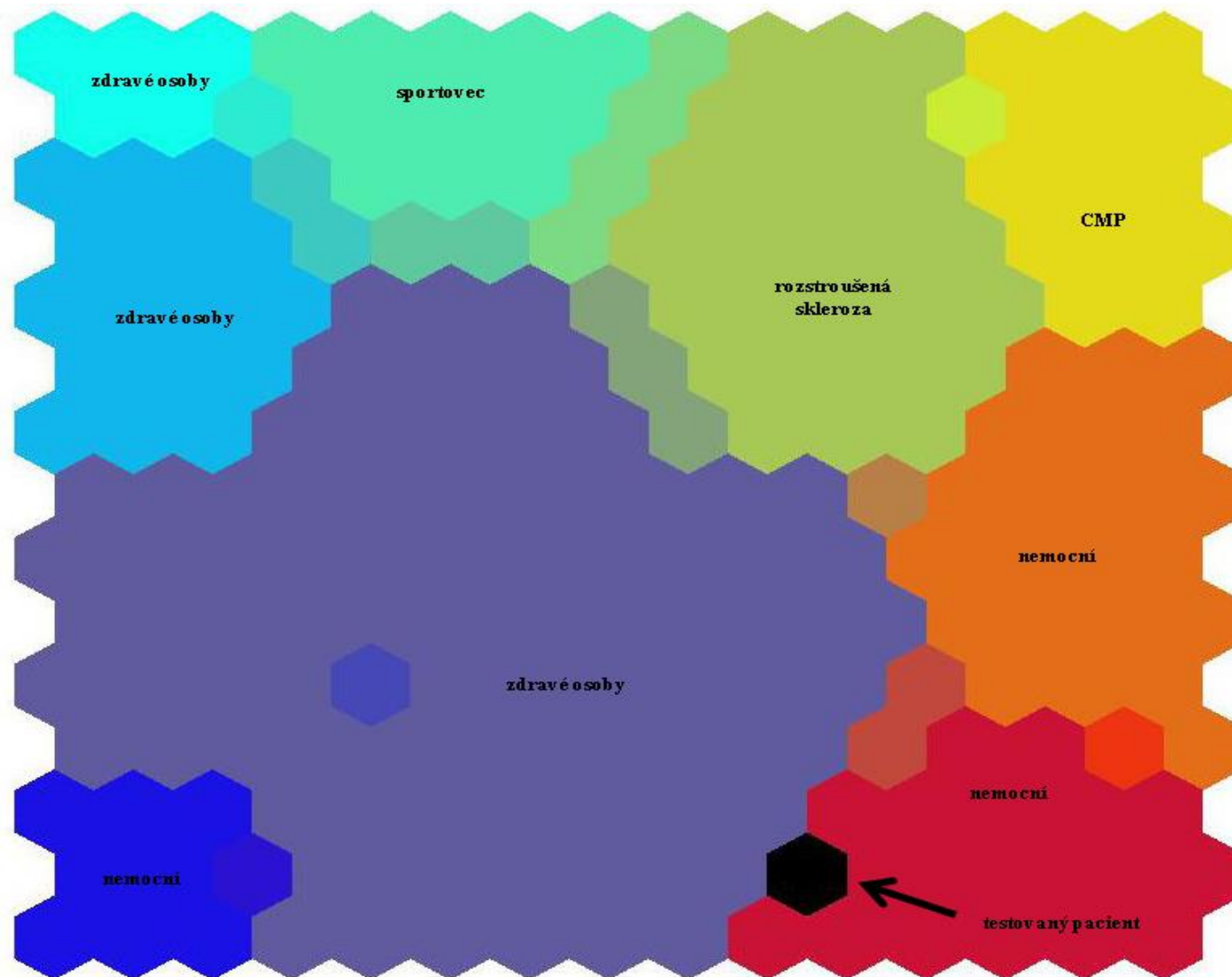
Flexe a extenze v zápěstí



**Flexe a extenze v zápěstí
s oporou**

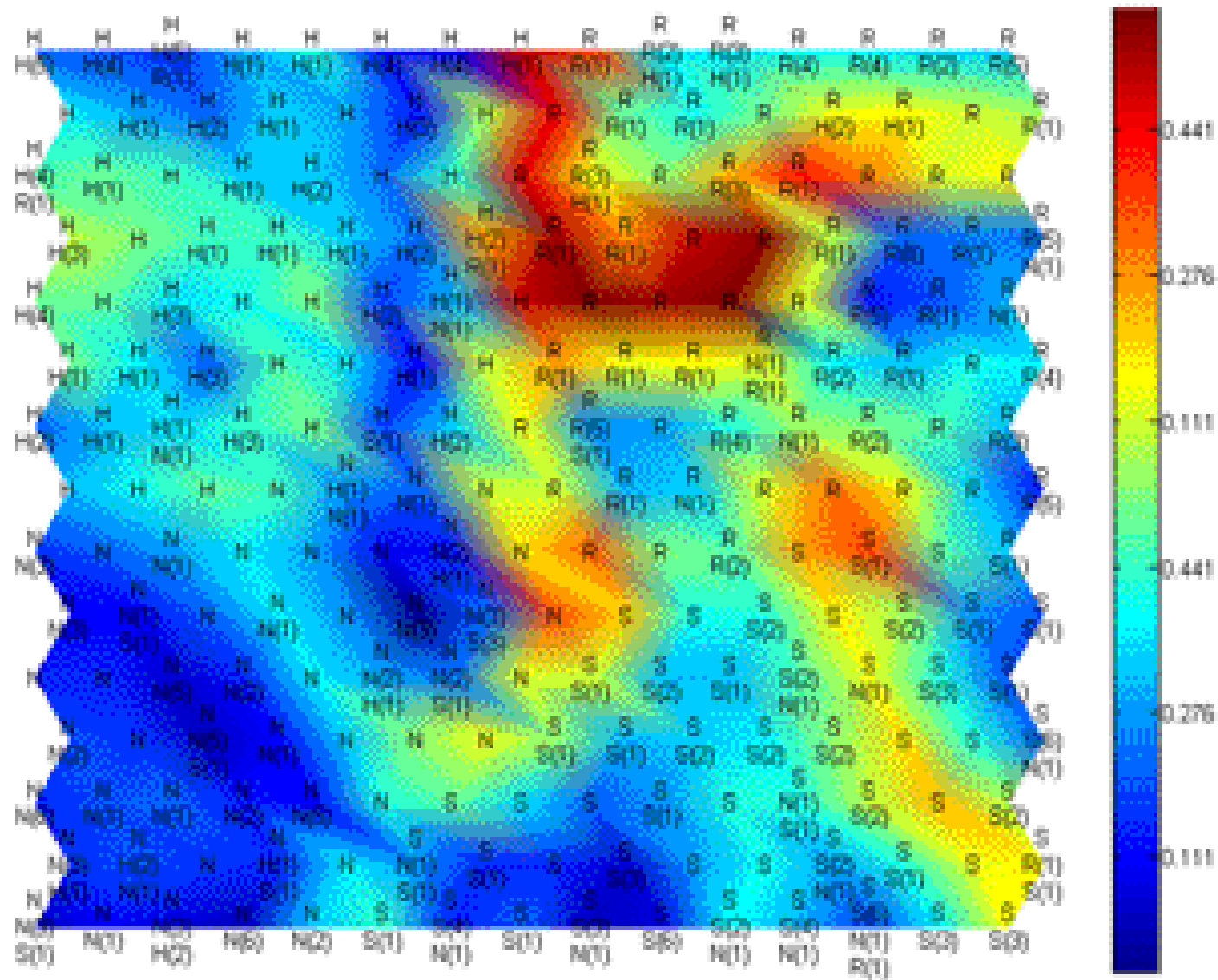


Klasifikace pohybu ruky



těžce
poškozená
hybnost ruky

Klasifikace emocí – Radost, Hněv, Smutek, Nuda



Predikce - předpověď, prognóza toho, co se stane nebo nestane v budoucnosti

- odhady opřené o vědeckou hypotézu nebo teorii

Predikce časových řad - predikujeme jednu nebo více hodnot

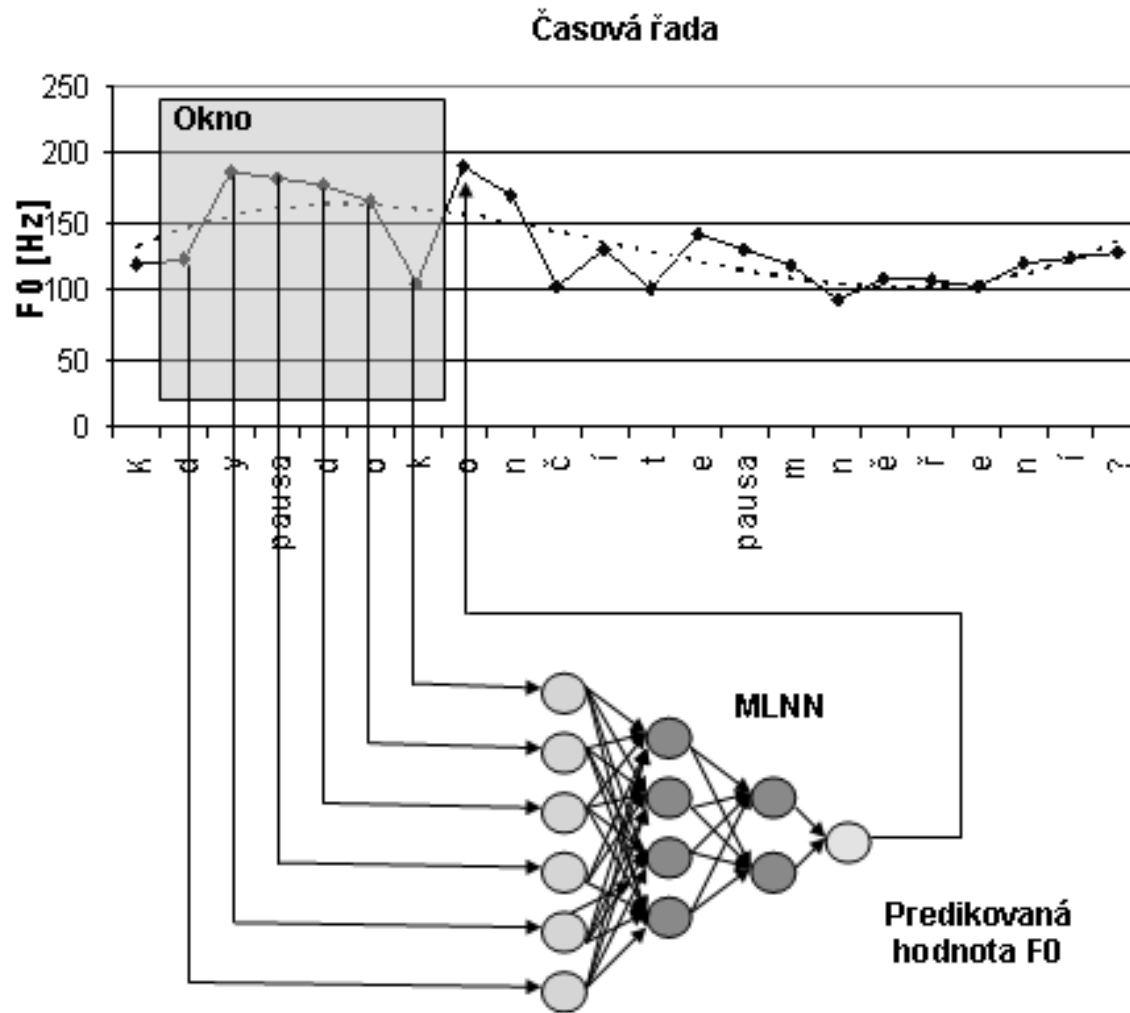
- predikujeme trend nějaké veličiny

klasifikace do 2 tříd → *veličina roste nebo klesá*

Postup

- 1) Analyzujeme fyzikální a matematickou povahu predikovaných veličin.
- 2) Provedeme výběr vstupních dat. Určíme požadovanou kvalitu prediktoru.
- 3) Specifikujeme požadavky na přesnost predikce.
- 4) Vybereme vhodný model neuronové sítě .
- 5) Určíme trénovací a testovací data .
- 6) Optimalizujeme navrženou architekturu.
- 7) Otestujeme.

Aplikace UNS: průběh základního kmitočtu F0 ve větě



Přístup: algoritmický, heuristický, aplikace UNS

schopnost adaptace na základě známých vzorů
možnost využití nelinearit



Aplikace: předpokládaný vývoj zdravotního stavu pacienta
(příklady) pohyb kursů měn v příštím časovém období
spotřeba el. energie v následujících hodinách
predikce hustoty provozu v dané dopravní lokalitě

Otázky: na jakou dobu budeme predikovat
jakým způsobem budou data předzpracována

Specifikace požadavků na přesnost: a) střední kvadratická chyba MSE
(*mean squared error*)

$$\text{MSE} = 1/N \sum (\text{predikce} - \text{skutečnost})^2$$

b) střední absolutní chyba MAE
(mean absolute error)


$$\text{MAE} = 1/N \sum | (\text{predikce} - \text{skutečnost}) |$$

c) střední relativní chyba MAPE
(mean absolute percentage error)

$$\text{MAPE} = 100/N \sum | (\text{predikce} - \text{skutečnost}) / \text{skutečnost} | \quad \%$$

Typy používaných UNS: MLNN s 1 skrytou vrstvou (alespoň)
RBF síť (radiální přenosová funkce)
KSOM

Nutnost užití reprezentativní trénovací a validační množiny

 **10% - 50%**

problém „stárnutí“ dat

složitý průběh, zašuměná data → větší datový soubor

Počet dat > 10 krát počet parametrů (doporučené heuristické pravidlo)

Optimalizace architektury

- **Klestění (**pruning**) velké sítě**
- **Přidávání neuronů k malé síti**
- **Citlivostní a toleranční analýza**
- **Test nečekaných událostí v reálném provozu**
- **Kdy je nutné síť přetrénovat ?**

Přirozená řeč jako zvukový signál

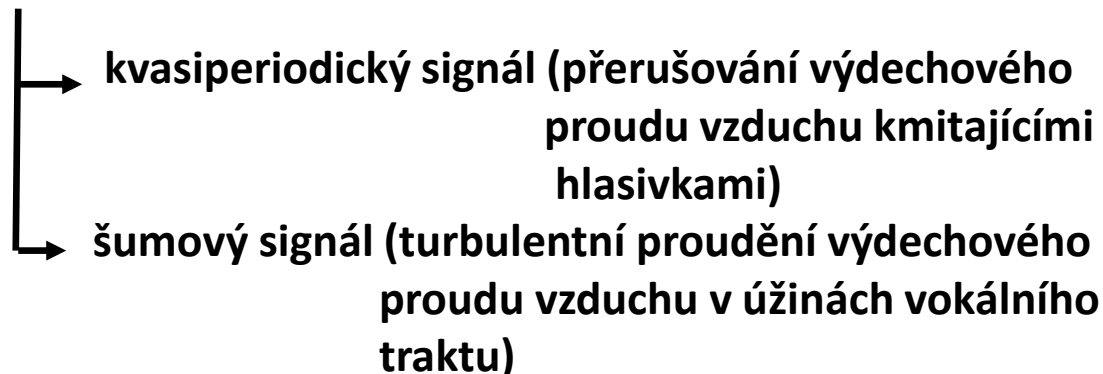
Řeč (speech) - komplex technických, fyziologických, fonetických a fonologických aspektů

Vznik zvukového signálu v mluvidlech = výsledek spolupůsobení jednoho nebo více zdrojů zvuku a filtračního systému.

Princip:

akustická soustava vokálního traktu

buzení dvěma druhy akustického signálu



Výška hlasivkového tónu nikdy nebývá konstantní.

základní tón (hlasivkový) F0 (*fundamental frequency*)

formantové kmitočty (superpozice vyšších harmonických k F0 a rezonance dutin hlasového traktu)

Kolísání základního tónu řeči F0 → **jitter**



závisí na duševním (emocionálním) stavu mluvčího.

Šumový signál v úžinách vokálního traktu:

turbulentní proudění vzduchových částic

(po překročení určité rychlosti proudění, dané

tvarem úžiny nebo překážky)

Zdroj šumu - hlasové ústrojí → šum čistý

→ šum superponovaný k hlasivkovému tónu

- na překážkách v nadhrtanových dutinách

Modelování: předpoklad širokopásmového šumu v oblasti nad 800 Hz, amplitudový spád o 6 dB/oktávu (pod kmitočtem 800 Hz a v oblasti nad 10 kHz)

Vokální trakt lze aproximovat soustavou dutin a průchodů.

Akustická soustava plynule měnící tvar a rozměry svých dutin, úžin, průchodů a závěrů.

Dva výstupy do vnějšího prostoru; akustické signály se ve vnějším prostoru skládají.

Hlas = složený zvuk, nejhlubší složka – základní tón F0

Kmitočtový rozsah hlasivkového tónu: 40 - 680 Hz

průměr podle [PAL94]: 100-150 Hz muži, 200-300 Hz ženy

Maximální dynamika intenzity běžné řeči: 50 dB

Souvislá řeč plynulý tok samohlásek, narušovaný souhláskovými změnami v místech

Souvislá řeč

plynulý tok samohlásek, narušovaný
souhláskovými změnami v místech
samohláskových spojů

Dvoukakanálového řízení tvorby řeči

Akustická skladba jednotlivých hlásek souvislé řeči je dále modifikována v rámci delších časových úseků (slova, taktu, větného úseku, věty) tzv. modulačními faktory, jimiž se vytvářejí prozodické rysy promluvy - modifikace časové, intenzitní a melodické

Fonetický popis (popis zvukové formy)

základní jednotka popisu - elementární zvukový segment
volba délky segmentu ovlivňuje další zpracování řeči

Vlastnosti: segment musí být

dostatečně krátký

schopen tvorby dalších jednotek (řetězením - *concatenation*)

znám vztah k jazyku jako lingvistické formě

lze kombinovat segmenty o různých velikostech

Základní jednotka - hláska (foném, morfém)

grafické znázornění fonému, zvuková varianta fonému

Hláska jako segment je vytvořena uměle, zvukový signál rozložíme přirozeným způsobem pouze na slabiky. Rozklad na hlásky se učíme při výuce psaní. Při poslechu řeči ji nevnímáme jako složenou z hlásek, ale z delších úseků.

Při zpracování akustického signálu dovedeme určit hlásky pouze tehdy, známe-li, co bylo řečeno.

Suprasegmentální charakteristiky

změny trvání, intenzity a F_0

změny prozodických parametrů v segmentech delších než 1 slabika uplatňují se - ve slovech

- v přízvukových taktech (*rytmická jednotka, skupina slabik s jedním přízvukem*)

- ve větných úsecích

- ve větách ...zde totožné s větným úsekem

Hrad je skoro rozbořen.

'hratje 'skoro 'rozbořen

Hledání zákonitostí - velmi obtížné (*komplexnost problému, často přípustných více realizací*)

Charakteristiky

Slovní přízvuk - vliv na prozodické parametry:

F0 trvání (*duration*) intenzita
↓ ↑ ↓

melodický pohyb hlasu na konci mluveného textu

systemem **kadencí (melodémů)** spolu s dynamickou složkou slouží

- k odlišení věty ukončené od věty (úseku), která bude (po pauze) dále pokračovat
- k odlišení různých druhů vět (otázky zjišťovací od ostatních druhů vět)
- k odlišení vět neutrálních (bez citového zabarvení) od vět citově zabarvených

!!! citové zabarvení může měnit obsah věty

Prozodie

suprasegmentální jevy: slabika, přízvuk, tón, intonace (melodie),
frázování, rytmus

- **Barva hlasu** (témbr) je přímým obrazem fyziologického stavu mluvčího.
- Zpracování řeči můžeme posuzovat ze tří pohledů:

- **akustického**, percepčního a lingvistického.

- **Akustická hladina** - parametry, které jsou měřitelné a matematicky popsatelné.
- **Percepční hladina** souvisí s vnímáním promluvy posluchačem.
- **Lingvistický pohled** se zabývá posloupností abstraktních jednotek (fonémů, grafémů,...)

Příznaková prozodie → přítomnost emocí a postojů

Přízvukový takt - skupina slabik s jedním přízvukem

(slovo, předložka + slovo)

→ základní rytmická jednotka české prozodie

časové trvání taktu - tendence vytvářet stejně dlouhé

jednotky (*v češtině*)

$$10 \log T(n) / T(3) = - 4.05 \log (n/3)$$

trvání taktu o 3 hláskách

Větný úsek - skupina přízvukových taktů

nauka o významu slov

souvisí s délkou promluvy, syntaxí, sémantikou a kontextem

syntaxe - vzájemný vztah symbolů (gramatika)

sémantika - vztah symbolů k realitě (pomáhá stanovit smysl věty)

Větný přízvuk - mírné zesílení slovního přízvuku

Větná melodie - systém kadencí (melodémů)

rozlišení typu vět

Pauza - členění souvislé řeči na menší jednotky = *frázování*

!!! Správné umístění je nutné ke správnému pochopení promluvy.

Před pauzou se často protahuje koncová slabika (o 60% - 100%)

Tempo - počet slabik / min resp. počet slov / min
počet a délka pauz

Segmentální a suprasegmentální popis řeči

hierarchické skládání melodie

vrstvy :

- koartikulace hlásek ve slabikách
- slovo (takt) s přízvukem
- větný úsek s kontrastem mezi slovy (takty) = slovní důraz
- věta s melodémy (rozlišení typů vět)
- promluva s různými styly a emocionálními postoji
- individuální charakteristiky mluvčích

Pozn.: obdobné pro intenzitu a trvání segmentu

Akustický popis samohlásek (vowels) – vokálů

- Pohyb hmotnějších částí mluvidel, pomalejší změny spektra.
- Počet samohláskových formantů je závislý na tvaru vlny základního tónu (na způsobu kmitání hlasivek).
- Kmitočty formantů se u jednotlivých samohlásek odlišují a jsou pro ně charakteristické.
 - **Změny prvního formantu** odpovídají změnám **vertikálního** pohybu jazyka (**vysoké** samohlásky mají nízký první formant a naopak)
 - **Změny druhého formantu** odpovídají změnám **horizontálního** pohybu jazyka (**zadní** samohlásky mají nízký druhý formant a přední samohlásky vysoký)
- Trvání - odlišení samohlásky krátké a dlouhé

i u
e o
a

F1 a F2, částečně F3 charakterizují hlásky (akustické spektrum)

Průměrné hodnoty [PAL94] – podle několika pramenů, hodnoty nejsou jednoznačné:

<i>Formant/hlásky</i>	<i>pásmo F1</i>	<i>pásmo F2</i>
<i>i,í</i>	<i>300-500</i>	<i>2000-2800</i>
<i>e,é</i>	<i>480-700</i>	<i>1560-2100</i>
<i>a,á</i>	<i>700-1100</i>	<i>1100-1500</i>
<i>o,ó</i>	<i>500-700</i>	<i>850-1200</i>
<i>u,ú</i>	<i>300-500</i>	<i>600-1000</i>

F5, F6,.. Závisí na individuálních vlastnostech člověka a na konkrétní podobě F0.

F2/F1 podílí se na charakteristice spektra vokálů

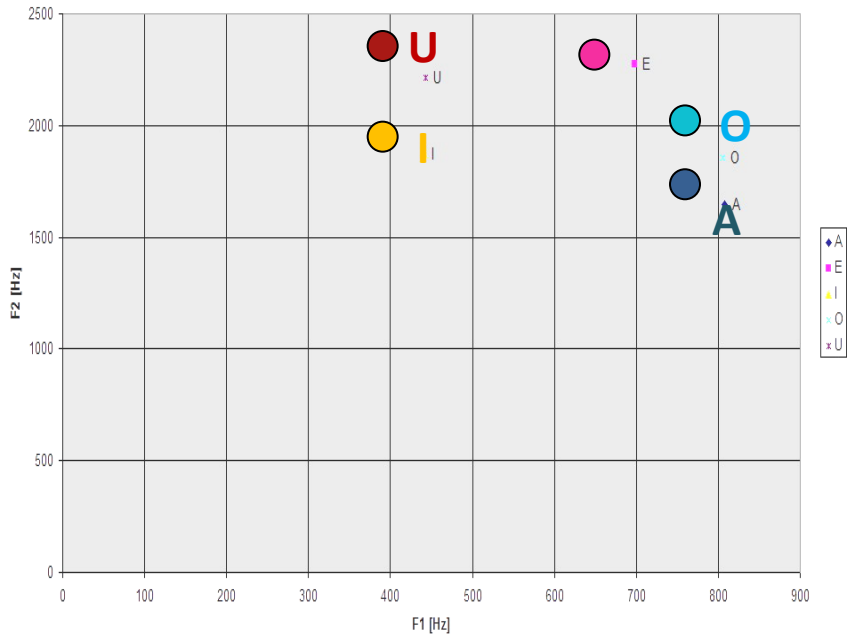
Vokalický (artikulační) trojúhelník
pro klasifikaci vokálů podle pohybů jazyka

	<i>přední</i>	<i>centrální střední</i>	<i>zadní</i>
<i>vysoké</i>	<i>i</i>		<i>u</i>
<i>středové</i>		<i>e</i>	<i>o</i>
<i>nízké</i>		<i>a</i>	

Vlivem spojování hlásek v souvislý tok řeči (koartikulace) a vlivem realizace suprasegmentálních rysů (v delších segmentech) se trvání jednotlivých samohlásek mění v širokém rozmezí (při normálním tempu v poměru až 1:5), a proto záleží i při rozlišování samohlásek dlouhých a krátkých na relativních poměrech trvání. V průměru jsou dlouhé samohlásky 1,5 - 2 krát delší než samohlásky krátké; při tempu asi 120 slov/minutu se trvání krátkých samohlásek pohybuje v rozmezí 40 až 160 ms (s průměrem 100 ms), trvání dlouhých v rozmezí 80 až 320 ms (s průměrem 200 ms).

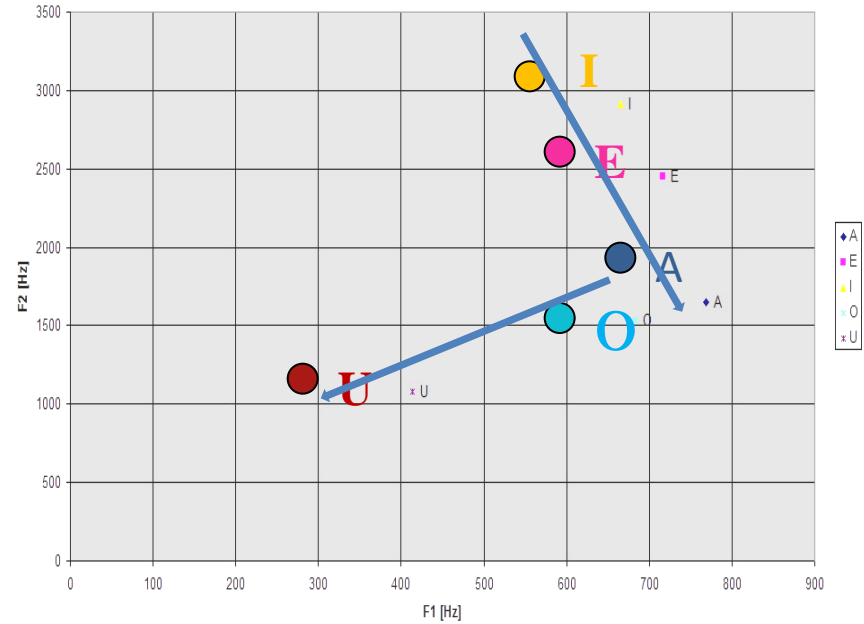
Ověření – vokální trojúhelník

50 frekvence MEAN: F1 - F2



Nemocné dítě

50 frekvence optimalizovaná pro každou samohlásku: F1 - F2



Zdravé dítě

Vokální trojúhelník



Akustický popis souhlásek (*consonants*) - konsonant

Vznikají vytvořením překážky výdechovému proudu

Frikativy: *f v s z š ž x h ř r l j* (*vibranty, laterály*)

Afrikáty: *c č*

Explozívy (okluzívy): *p b t d t' d' k g*

Nazály (nosovky): *m n ň*

Znělé: *b d d' g v h z ž m n ň ř*

Neznělé: *p t t' k f ch s š c č (ř)*

Sonorní (slabikotvorné): *r l j*

Percepčně závažné charakteristiky samohlásek

nejvýznamější jsou první dva až tři formanty v pořadí F2, F1, F3
vyšší formanty (pro všechny samohlásky prakticky stejné, obsahují
značnou část informace o barvě hlasu mluvčího

Percepčně závažné charakteristiky souhlásek

pro většinu souhlásek je významný začátek následující samohlásky
pro *s, z, š, ž, x, ř, c, č* je významný jejich charakteristický šum
pro *j* jsou významné sousední samohlásky
pro *h* se objevuje šum u sousedních samohlásek
pro *l* je závažný přechod na hranici mezi souhláskou a samohláskou