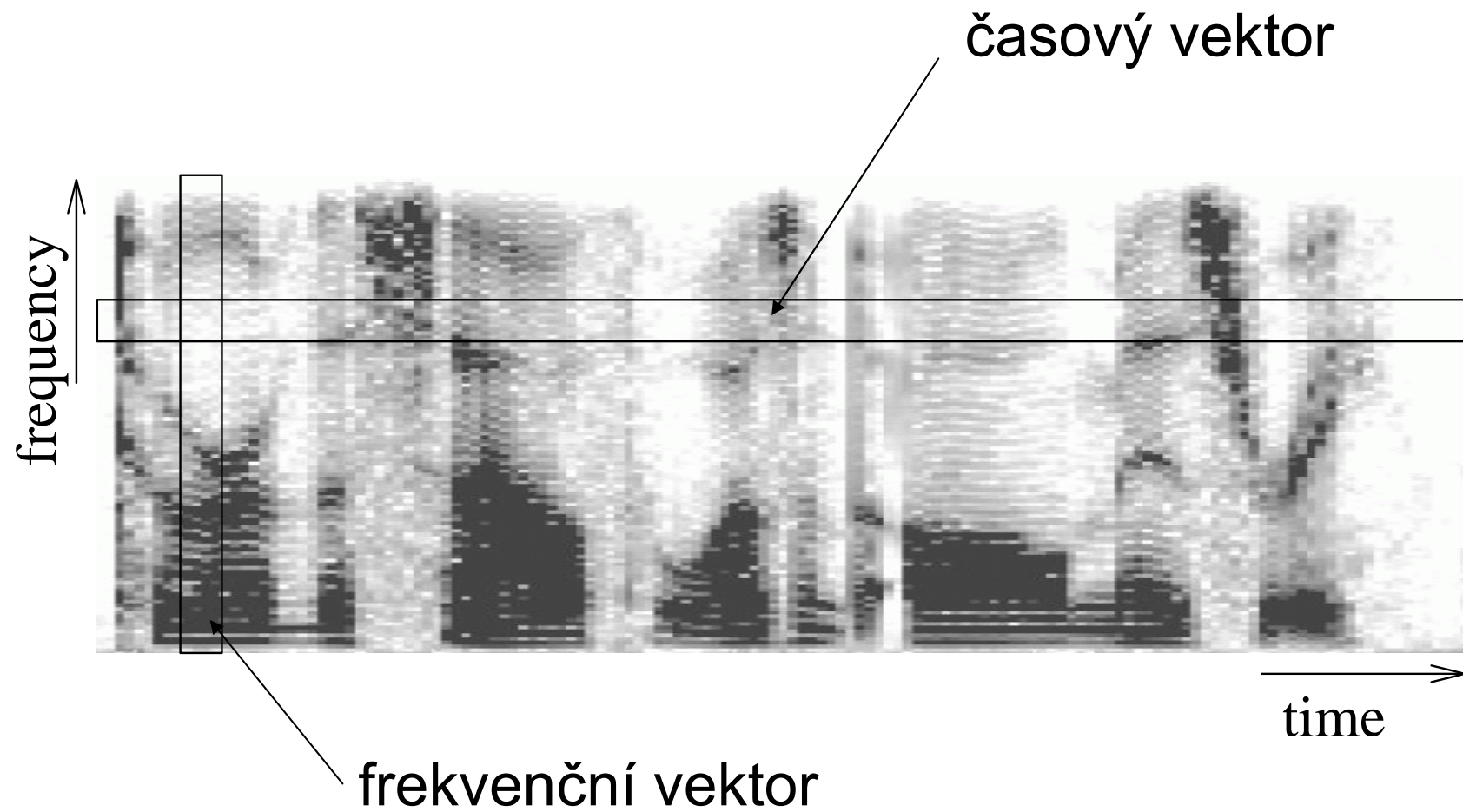


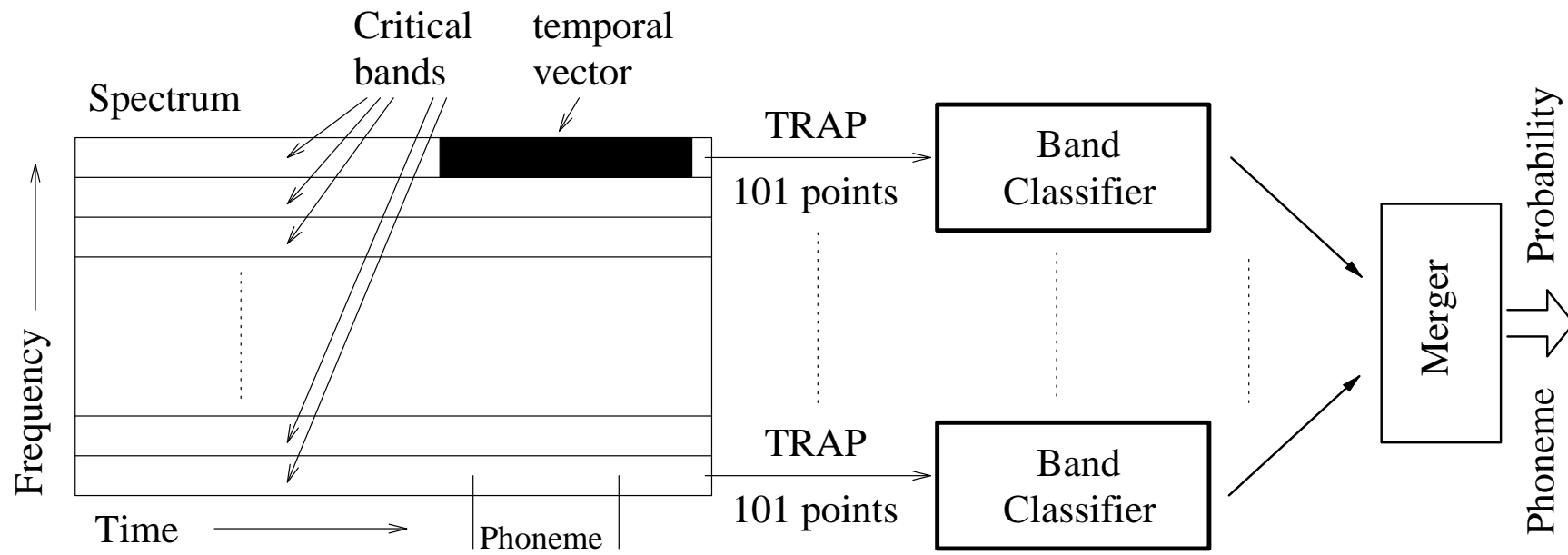
# TempoRAI Patterns (TRAPs)

- dlouhé časové trajektorie

# Spektrogram

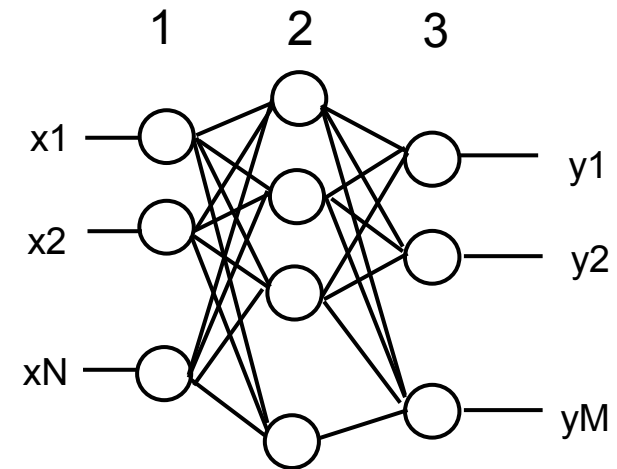


# TRAP systém



# Band classifier a merger

- jsou použity dopředné neuronové sítě (bez zpětných vazeb)
- třívrstvá neuronová síť:
  1. vrstva – reprezentuje vstup
  2. vrstva – sigmoida
  3. vrstva – softmax
- softmax - zajišťuje pravděpodobnostní výstup



sítě

$$\sum_{i=1}^{N_{out}} y_i = 1$$

# Trénování neuronových sítí

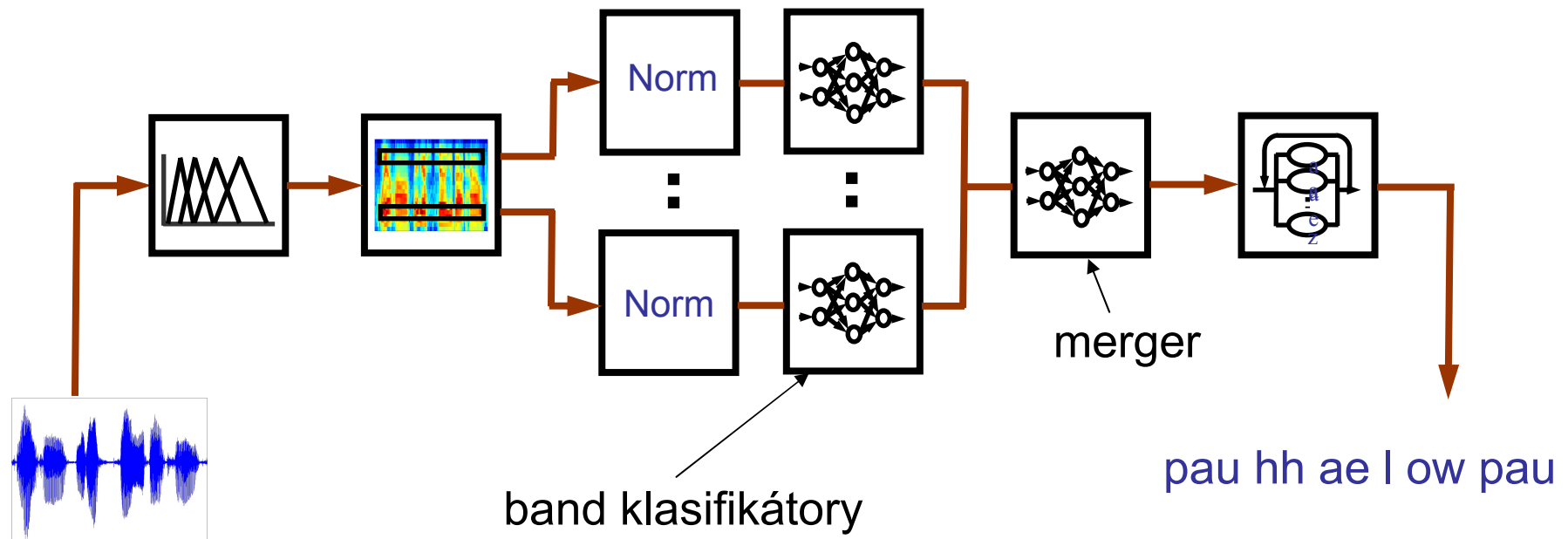
- back propagation
- trénovací, validační a testovací set
- všechny sety obsahují popis na úrovni rámců (víme ke kterému fonému rámeček náleží)
- vyhodnocení úspěšnosti – FER (Frame Error Rate)
- nejprve trénování klasifikátorů pro pásma, poté jsou trénovací data „prohnána“ těmito klasifikátory a následně je trénován merger

# Použití TRAPs systému

- 1) Přímou pro odhad posterior pravděpodobností fonémů
- 2) Jako příznaky do klasických rozpoznávačů (založených na skrytých Markovových modelech)

# TRAPs systém pro rozpoznávání fonémů

- TRAPs systém je použit k odhadu posterior pravděpodobností fonému
- Následuje dekodér, založený na Viterbi algoritmu, který převádí matici posterior pravděpodobností na řetězec fonémů

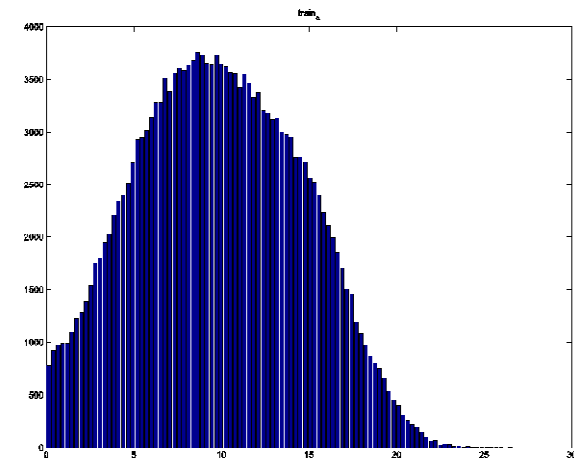
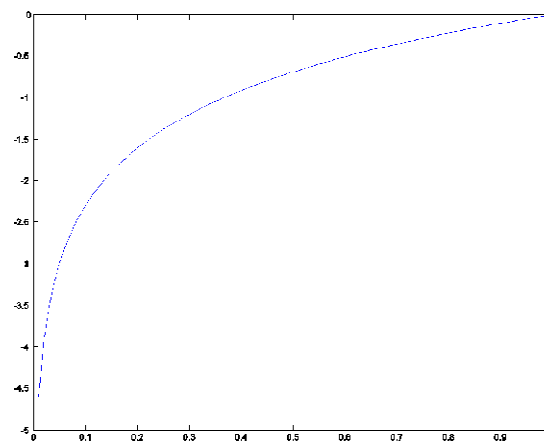
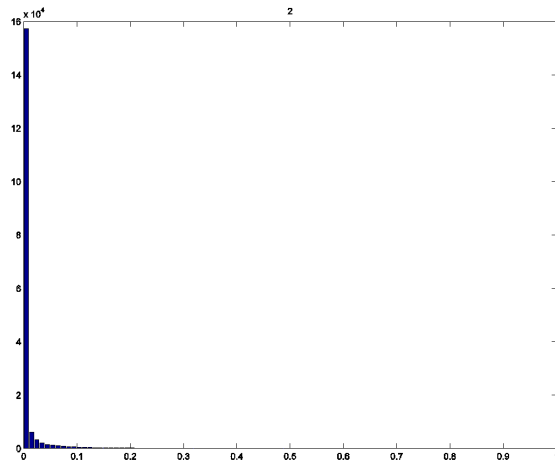


# Příznaky do klasického rozpoznávače

- snažíme se aby měly normální rozdělení a byly dekorelované



- Použití logaritmu





Fonémový rozpoznávač pro češtinu  
založený na TRAPs

<http://www.fit.vutbr.cz/speech/sw/phnrec>

