

Akusztikus jelek automatikus osztályozása beágyazott rendszerekben

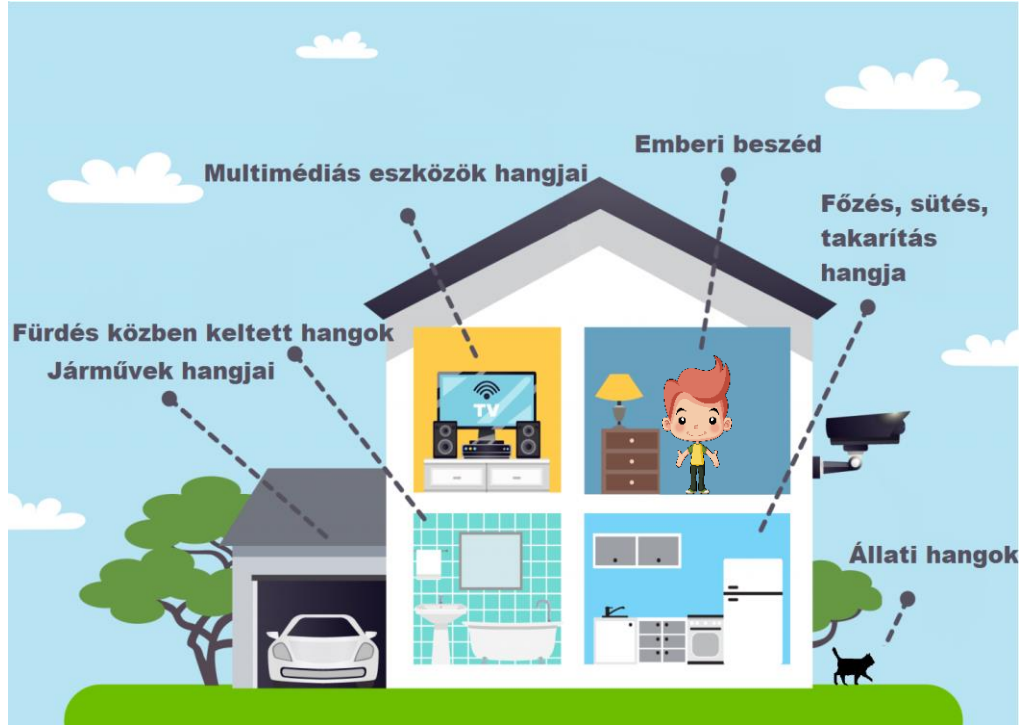
Ossik László

Konzulens: dr. Orosz György

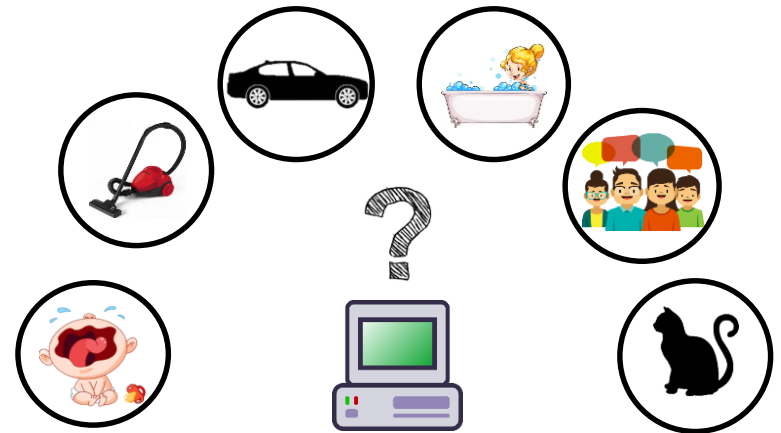


Méréstechnika és
Információs Rendszerek
Tanszék

Bevezetés



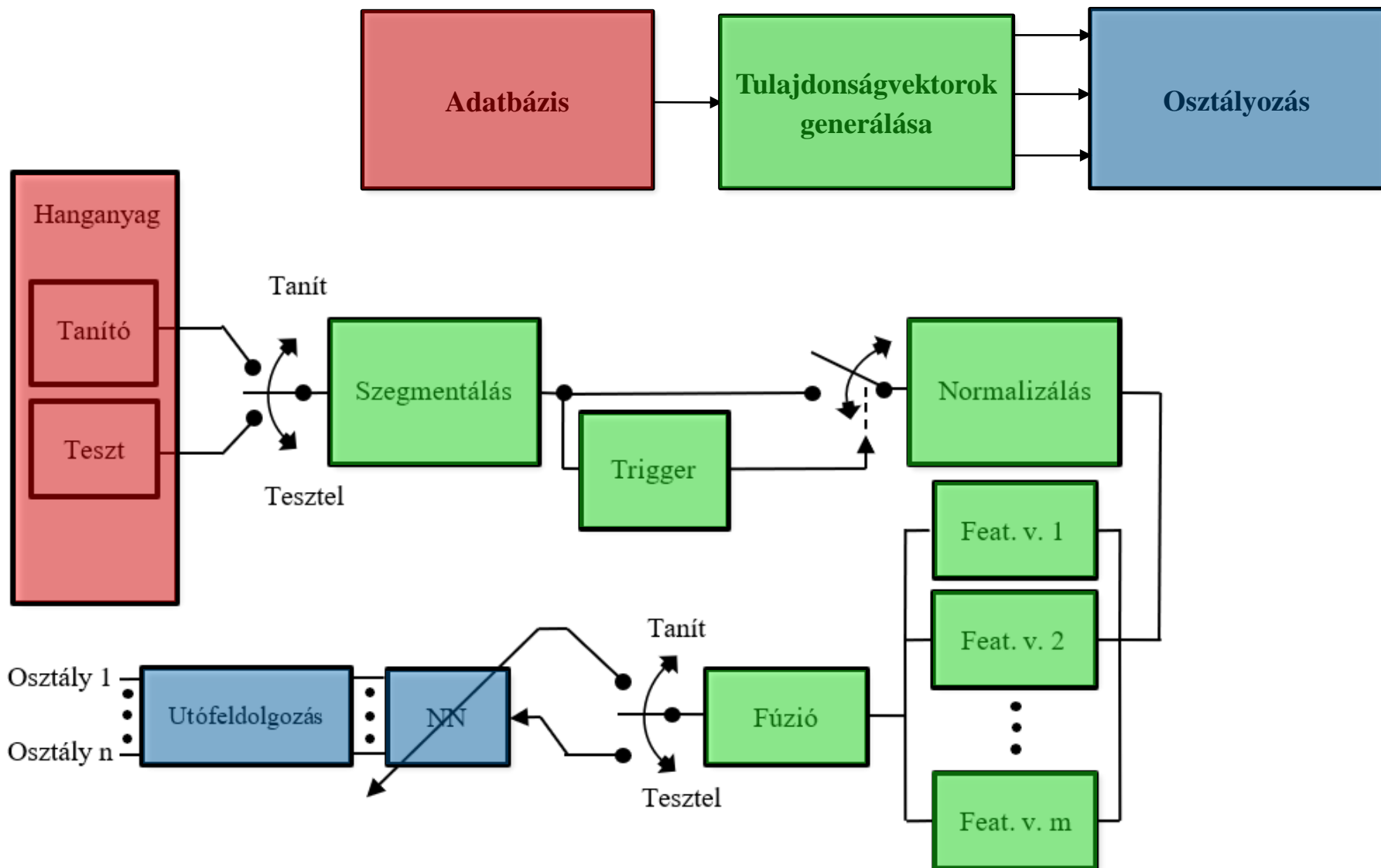
A rendszer feladata az otthoni környezetben előforduló hangforrások beazonosítása.



Kiindulási pont

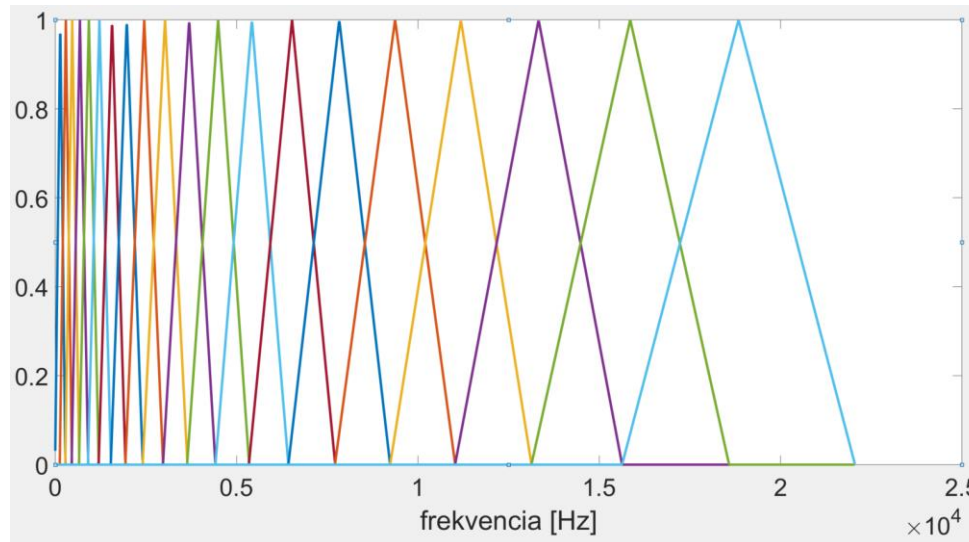
- Korábban megvalósításra/kialakításra került:
 - Adatbázis.
 - Hardveres rendszer kialakítása.
 - Tulajdonságvektorokat generáló és azokat feldolgozó szoftver Matlabban/Pythonban.
- Cél:
 - Rekurrens neurális hálók megismerése ezen belül is az LSTM (Long Short Term Memory).
 - LSTM alkalmazása, megfelelő paraméterhalmaz megtalálása.

Megvalósított szoftver rendszerterve



Mel Frequency Cepstral Coefficient

- Alacsonyabb frekvenciákon lényegesebb információk találhatóak.
- Cepstrum:
 - *Egy jel teljesítmény cepstruma* $= |F^{-1}\{\log(|F\{f(t)\}|^2)\}|^2$
- Az MFCC a Mel spektrum Cepstruma.



LSTM felépítése

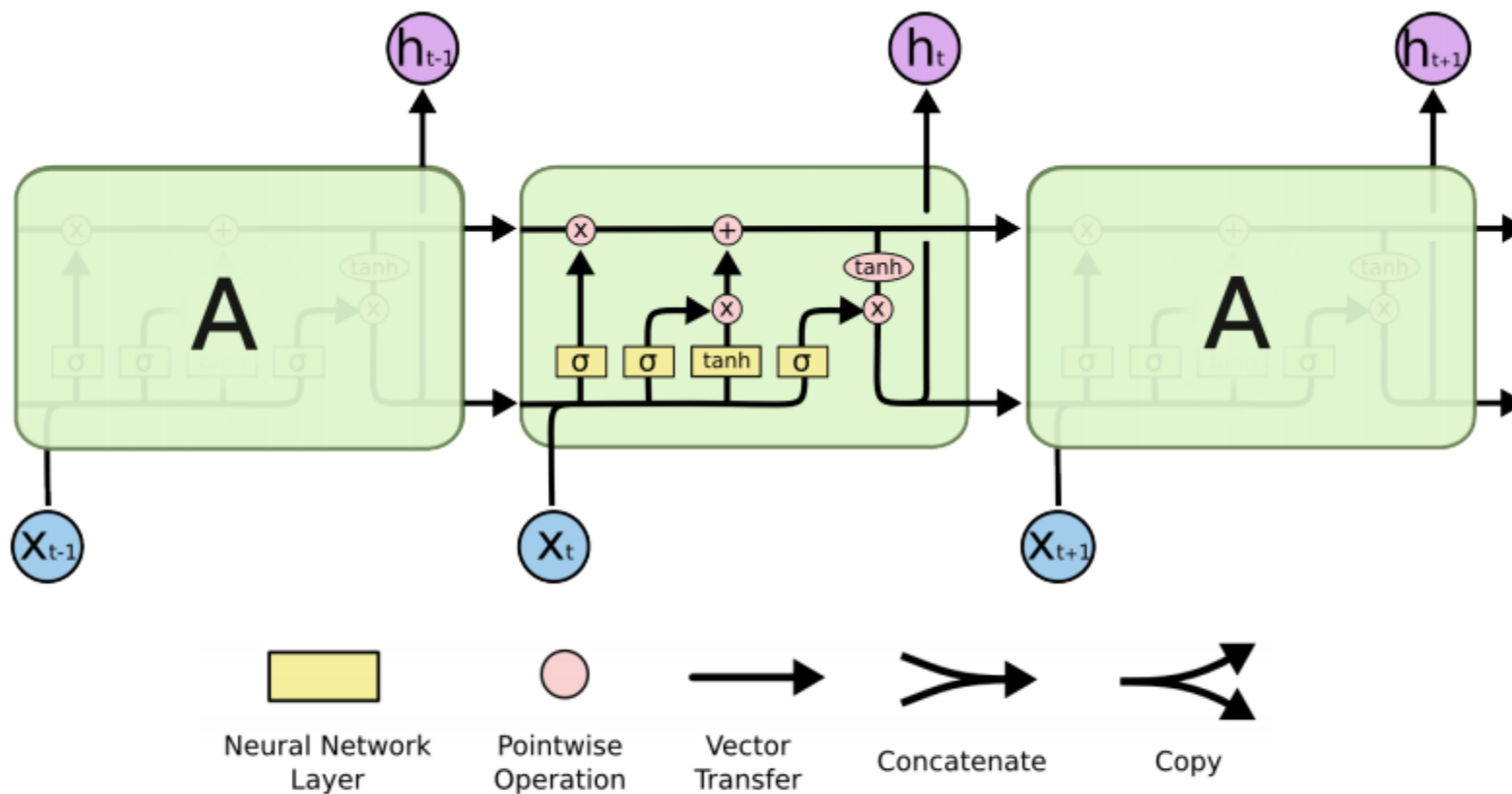
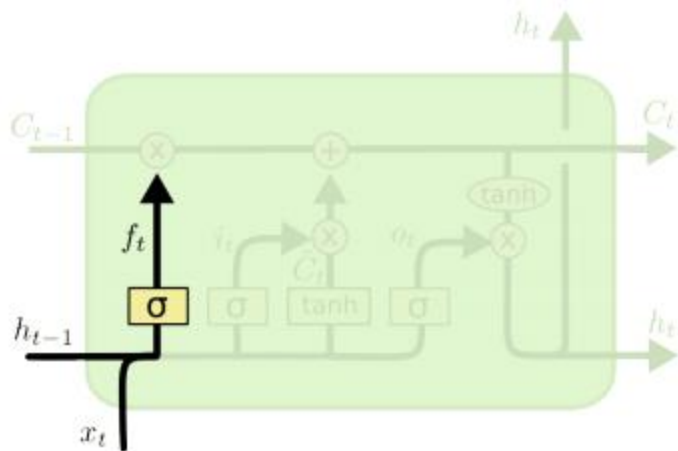


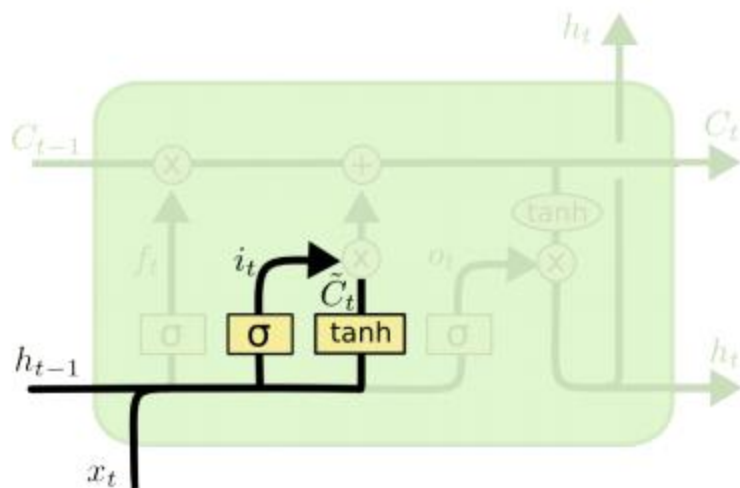
Fig: Christopher Olah

LSTM felépítése



Forget gate: megtartjuk nem tartjuk (meg)

$$f_t = \sigma(W_f \cdot [h_{t-1}, x_t] + b_f)$$



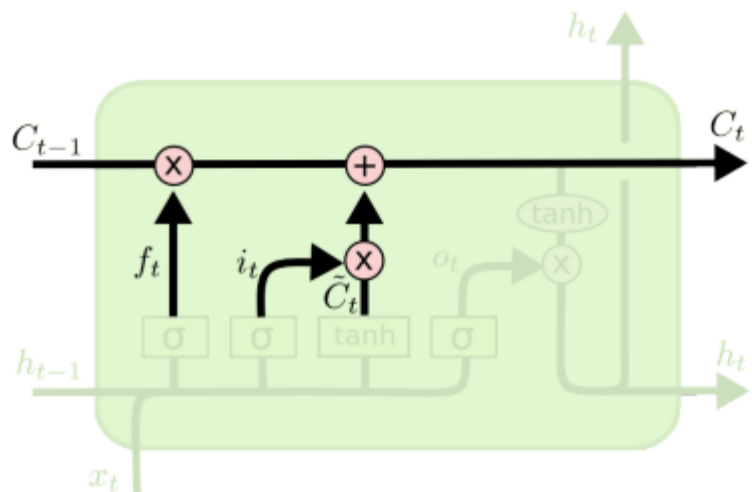
Input gate: szelekció és új állapot

$$i_t = \sigma(W_i \cdot [h_{t-1}, x_t] + b_i)$$

$$\tilde{C}_t = \tanh(W_C \cdot [h_{t-1}, x_t] + b_C)$$

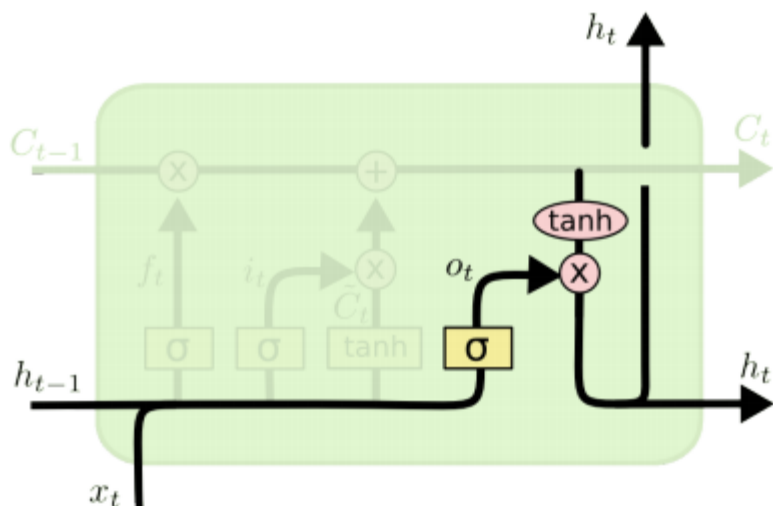
Fig: Christopher Olah

LSTM felépítése



Maga az új állapot

$$C_t = f_t * C_{t-1} + i_t * \tilde{C}_t$$



Kimenet: új állapot és kimenet

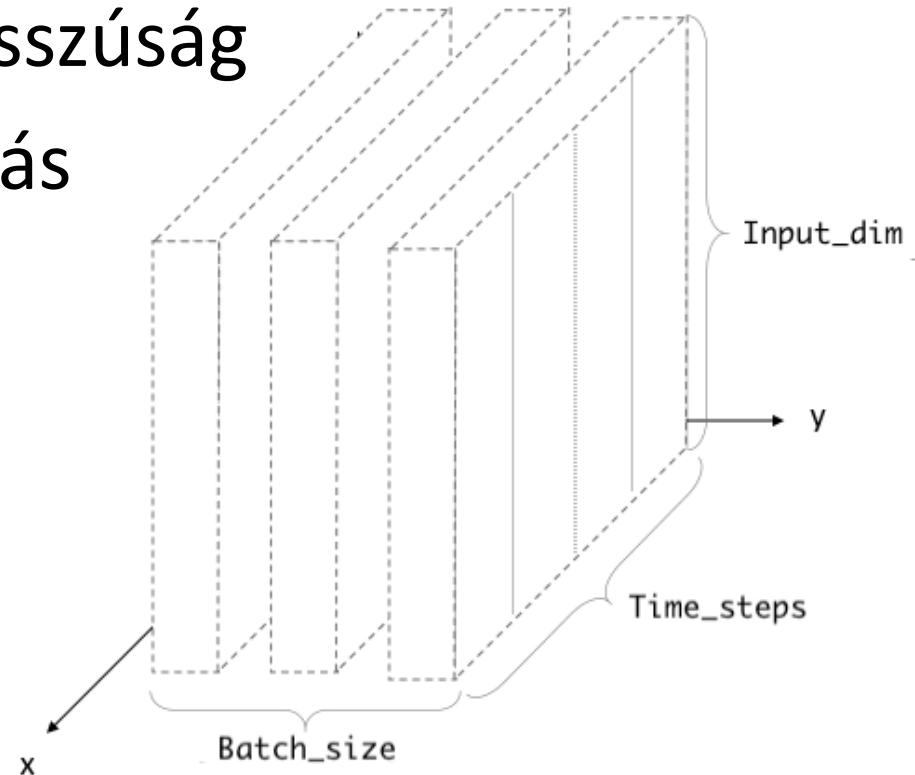
$$o_t = \sigma(W_o [h_{t-1}, x_t] + b_o)$$

$$h_t = o_t * \tanh(C_t)$$

Fig: Christopher Olah

LSTM bemeneti mátrix

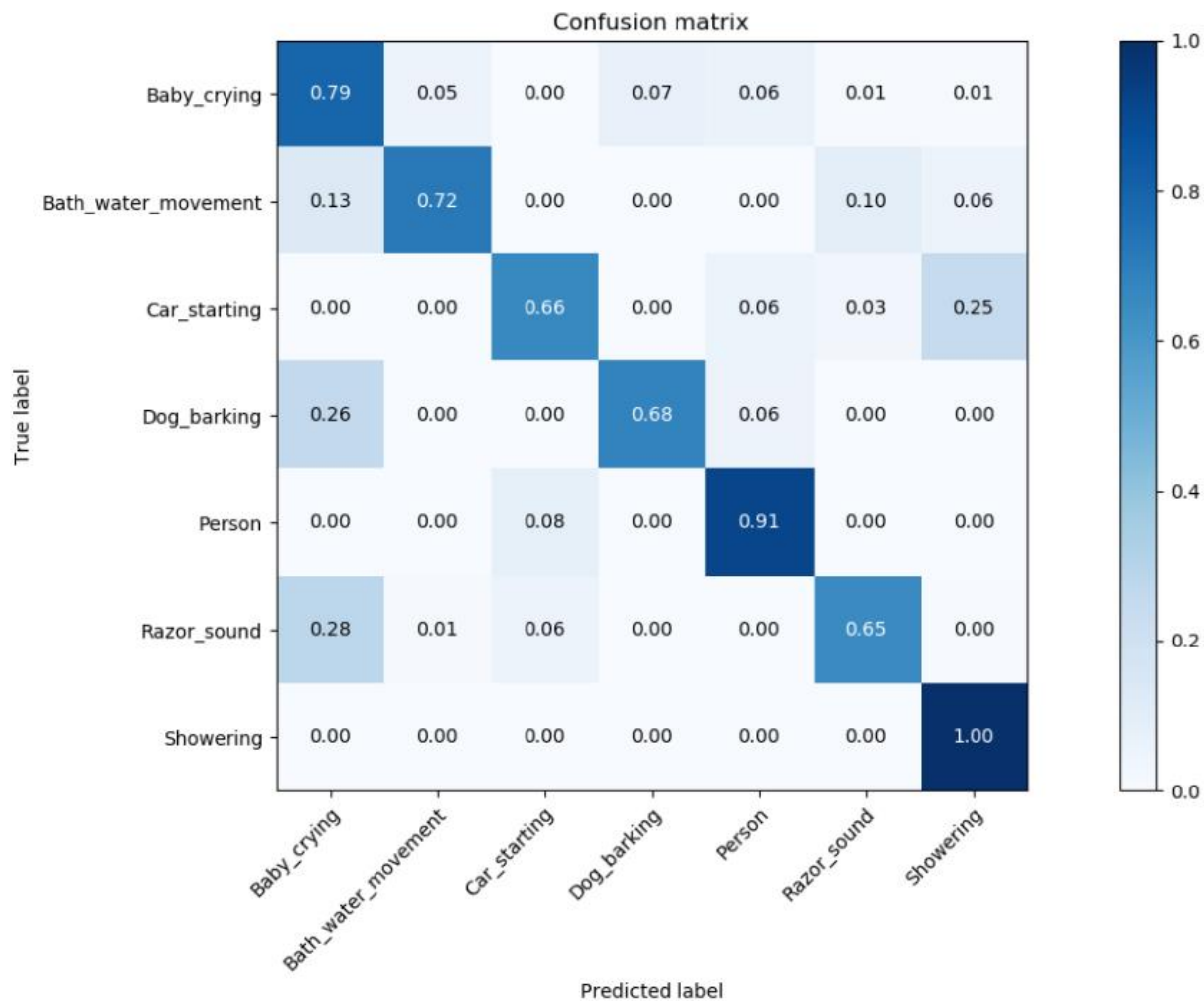
- MFCC tulajdonságvektorokat
- A gyorsabb tanulhatóság érdekében batch-eket alakítunk ki.
- Input_dim: szegmens hosszúság
- Batchek között átlapolódás



Paraméterek

- Hangolt paraméterek:
 - LSTM rétegek száma: 1
 - LSTM cellák kimeneti vektorának mérete: 36
 - Bemeneti mátrix Time_steps hosszúsága: 25
 - Bemeneti mátrix átlapolódásának mértéke: 24
 - Tanulási sebesség: 0.0001
 - Dropout: 0.3
 - Epoch: 20

Az LSTM-el kapott confusion matrix



Összefoglalás, kitekintés

- Szoftver sikeresen végrehajtja az osztályozást.
- Átlagosan olyan 77% a felismerési arány.
- Azon osztályok, amelyek rosszabban teljesítettek kevesebb mintával is rendelkeznek.
- Továbbfejlesztési lehetőségek:
 - A kialakított rendszerrel történő integráció megvalósítása
 - Adatbázis növelése (ez egy kritikus pont)
 - Paraméterek további hangolása

Köszönöm a figyelmet!