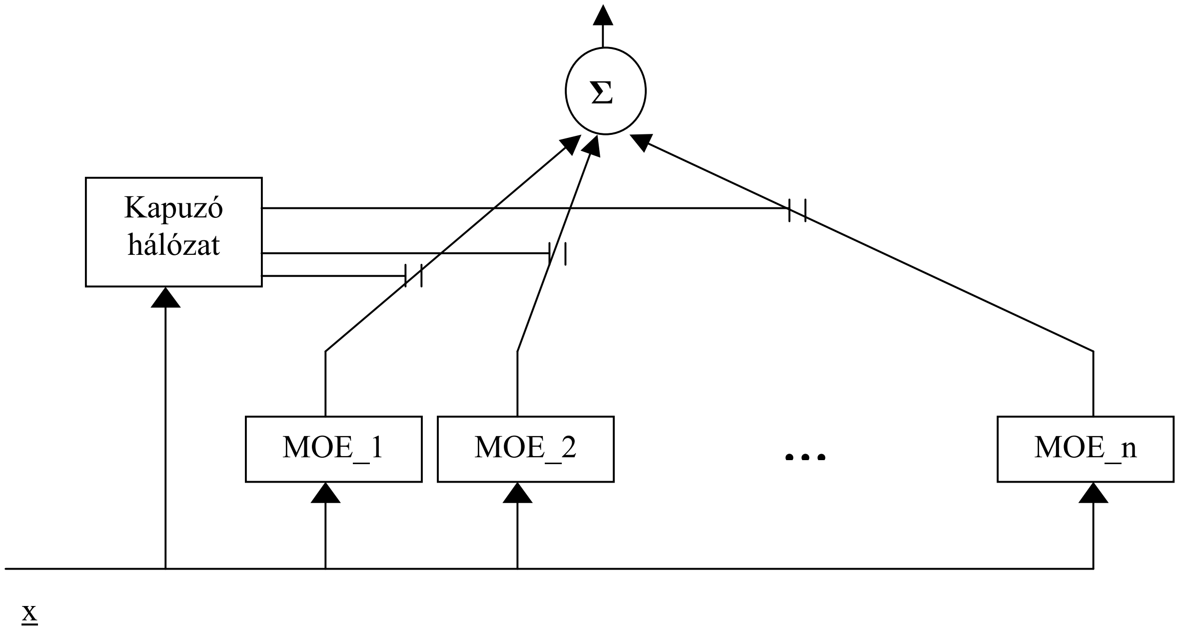
Szakértő együttesek (Mixture of Experts, MOE) vizsgálata

Hibrid rendszerek és 3D képfeldolgozás labor



Alapvető összefüggések:

A. Szakértők:

1. *i*. szakértő kimenete: $\mu_i = f(\mathbf{\Theta}_i^T \mathbf{x})$ lineáris szakértő esetén: $\mu_i = \mathbf{\Theta}_i^T \mathbf{x}$
2. asszociált zaj: $P(y^{(l)} | x^{(l)}, \Theta_j)$ Gauss zaj esetén: $\frac{1}{\sqrt{2\pi}\sigma_j} \cdot e^{-\left[(y-\mu_j)^2 / 2\sigma_j^2\right]}$
3. kapuzó súlyok: g_i
4. "*a posteriori*" valószínűség: $h_i^{(l)} = P(i^{(l)} | x^{(l)}, y^{(l)}) = \frac{g_i \cdot P(y^{(l)} | x^{(l)}, \Theta_j)}{\sum_j g_j \cdot P(y^{(l)} | x^{(l)}, \Theta_j)}$
5. Szakértő paramétereinek tanítása: $\Delta\Theta_i = +\rho \cdot \frac{\partial L}{\partial \Theta_i} = \rho \cdot \frac{\partial L}{\partial \mu_i} \cdot \frac{\partial \mu_i}{\partial \Theta_i}$
6. $\frac{\partial L}{\partial \mu_i} = h_i \cdot \frac{\partial}{\partial \mu_i} \ln P(y^{(l)} | x^{(l)}, \Theta_j)$ -------- gauss --------> $h_i \cdot \frac{y - \mu_i}{\sigma_i^2}$
7. $\frac{\partial \mu_i}{\partial \Theta_i}$ ------(lin.)------> $\underline{\mathbf{x}}$ ----($\mathrm{sigm}(s) = \frac{1}{1+e^{-s}}$)----> $\mathrm{sigm}(s) \cdot (1 - \mathrm{sigm}(s)) \cdot \underline{\mathbf{x}}$
8. pl.: $\frac{\partial L}{\partial \Theta_i} = \frac{h_i^{(l)}(y^{(l)} - \mu_i^{(l)})\mathbf{x}^{(l)}}{\sigma_i^2}$

B. Kapuzó hálózat:

1. $\xi_i = f(\underline{\mathbf{v}}_i^T \underline{\mathbf{x}})$

2. "*a priori*" valószínűség: $g_i^{(l)} = P(i^{(l)} \mid x^{(l)}, v^{(l)}) = \dfrac{e^{\xi_i}}{\sum_j e^{\xi_j}}$

3. Kapuzó hálózat paramétereinek tanítása: $\Delta v_i = +\rho \cdot \dfrac{\partial L}{\partial v_i} = \rho \cdot \dfrac{\partial L}{\partial \xi_i} \cdot \dfrac{\partial \xi_i}{\partial v_i}$

4. $\dfrac{\partial L}{\partial \xi_i} = \dfrac{\partial}{\partial \xi_i} \ln \sum_j g_j \cdot P(y^{(l)} \mid x^{(l)}, \Theta_j)$ ---------------> $h_i - g_i$

5. $\dfrac{\partial \xi_i}{\partial v_i}$ (lineáris kapuzó hálózat esetén) = $\underline{\mathbf{x}}$

6. pl.: $\Delta v_i = \rho \cdot (h_i - g_i) \cdot \mathbf{x}$