Regression neural networks. Application to speaker's age identification.

Anna Silnova

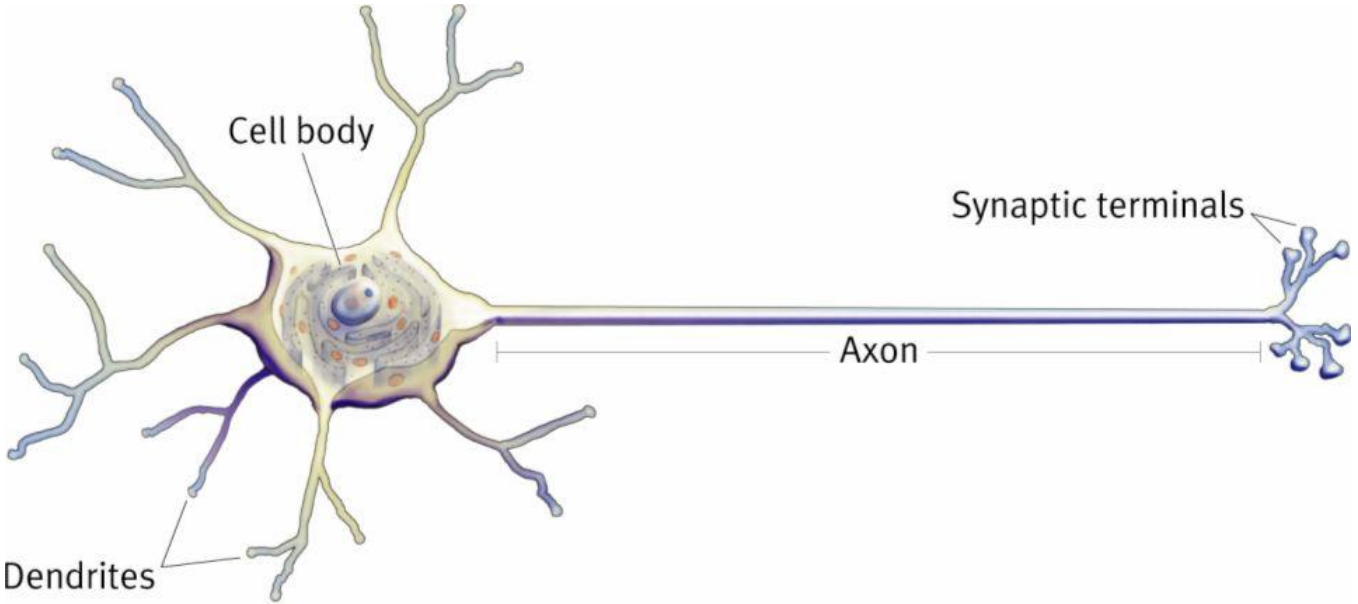
Brno, 17.12.2015

Regression problems

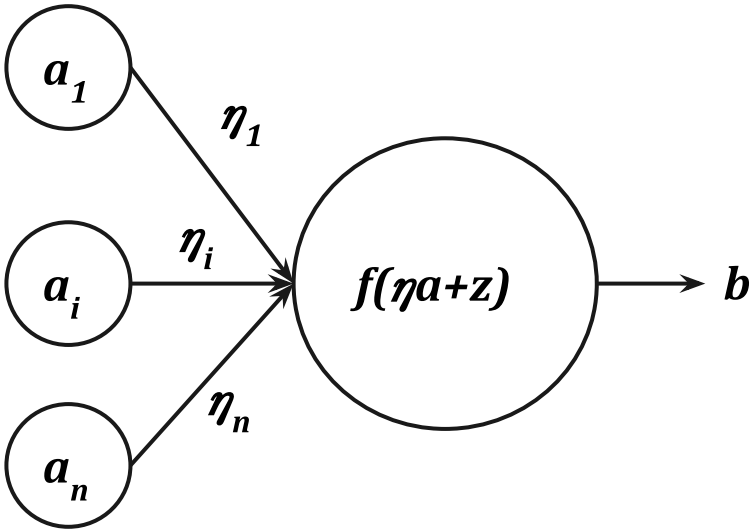
$$y=f(x,a)$$

- Linear
- Nonlinear
  - $f(x,a) = \Sigma a_i g_i(x)$, $g_i(x)$- nonlinear
  - $f(x,a)$ is nonlinear in respect to both x and a

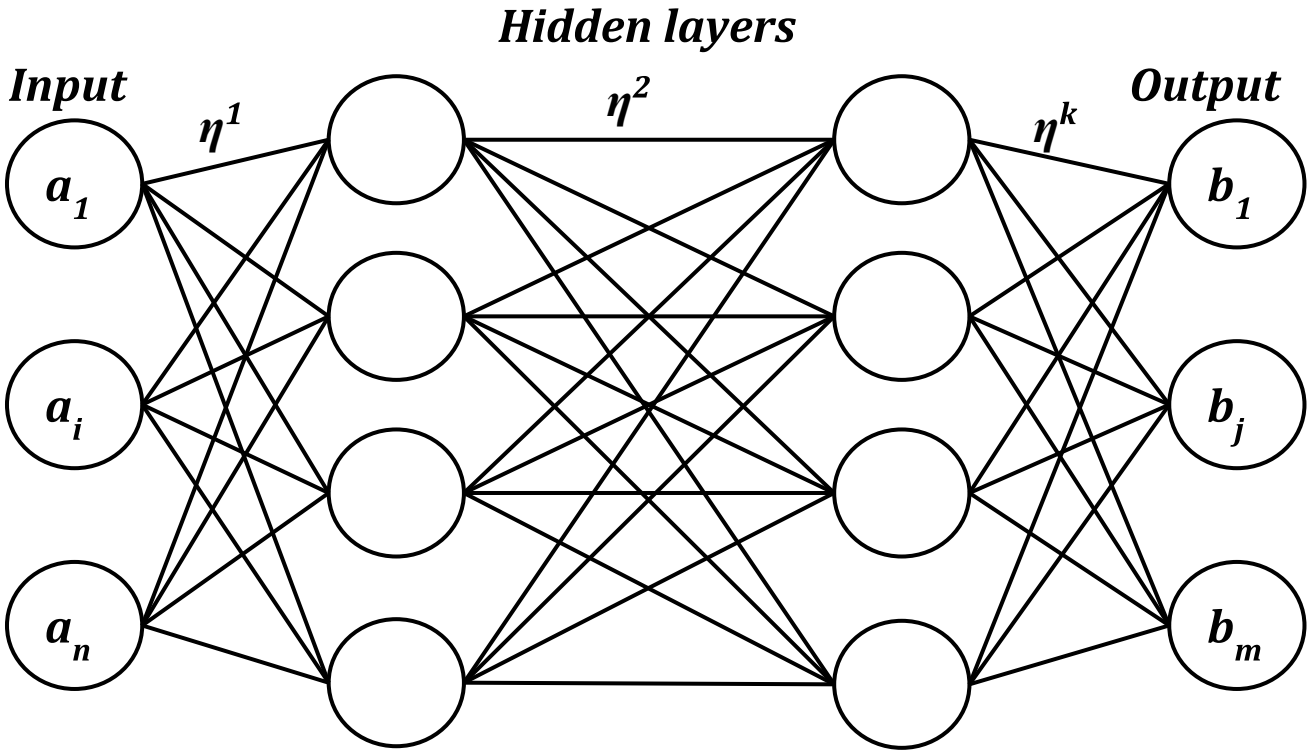
Biological inspiration



The artificial neuron



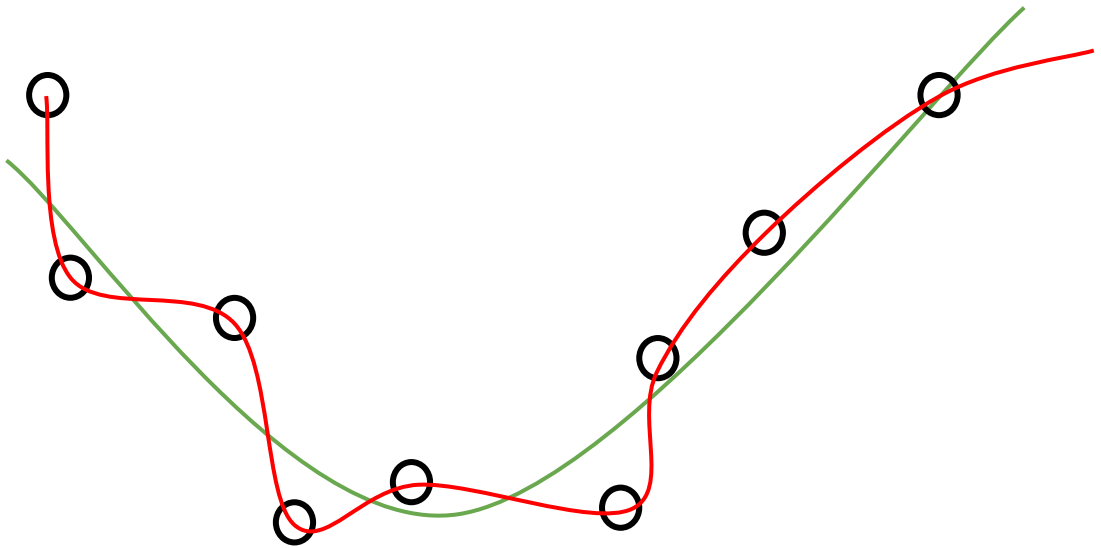
Multilayer perceptron



Network’s training

- Backpropagation algorithm
- Algorithms utilizing first-order derivative information of error surface
  - Stochastic gradient descent
- Algorithms utilizing higher-order derivative information of error surface
  - Broyden–Fletcher–Goldfarb–Shanno algorithm

Overfitting prevention

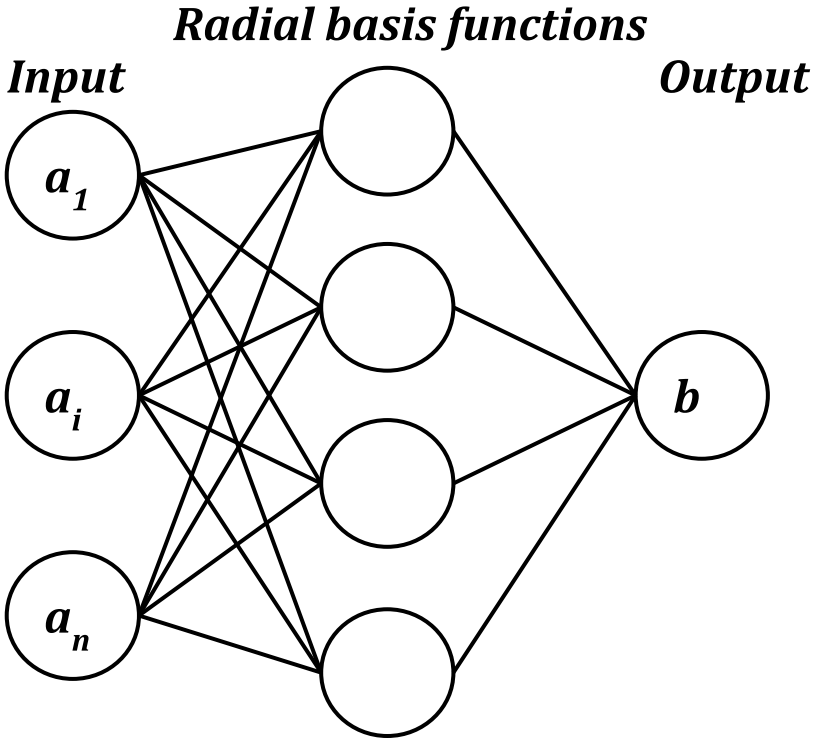


- Early stopping
- Dropout
- l_2-regularization

Drawbacks of MLP

- Large amount of data is needed
- For a reasonable performance great deal of training is needed
- The great amount of time is consumed for training

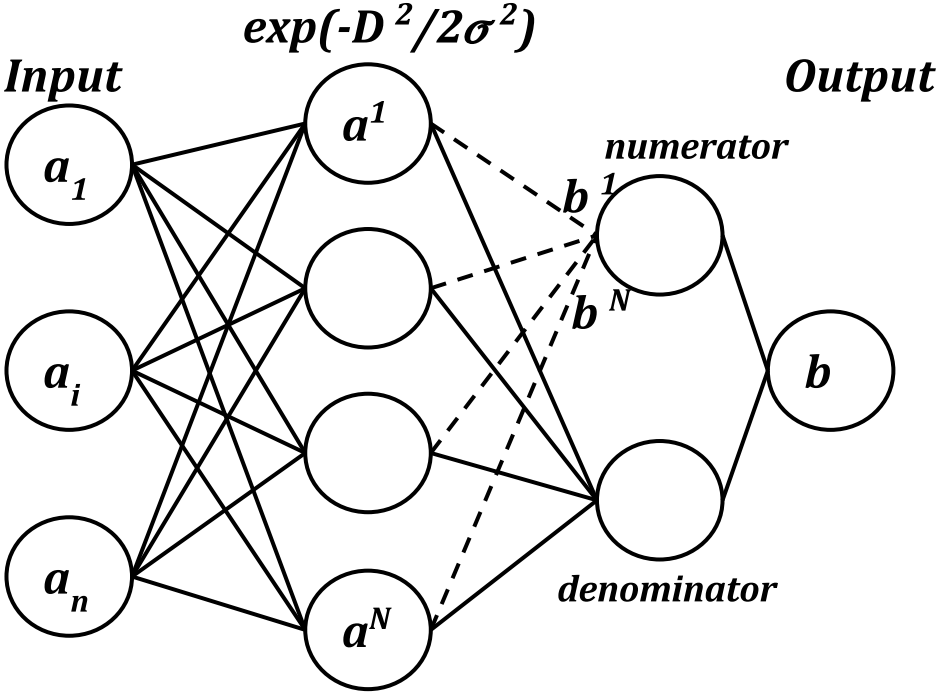
Radial basis function networks



$$b(a)=\Sigma\eta_j\varrho(||a-c_j||),$$

$$\varrho(||a-c_j||)=exp(-\beta_j||a-c_j||)$$

General regression NN



$$b(a)=\frac{\Sigma b^{j}\exp(-D^{2}_{j}/2\sigma^{2})}{\Sigma\exp(-D^{2}_{j}/2\sigma^{2})}$$

$$D^{2}_{j}=(a-a^{j})^{T}(a-a^{j})$$

Examples of regression problems solved by NN

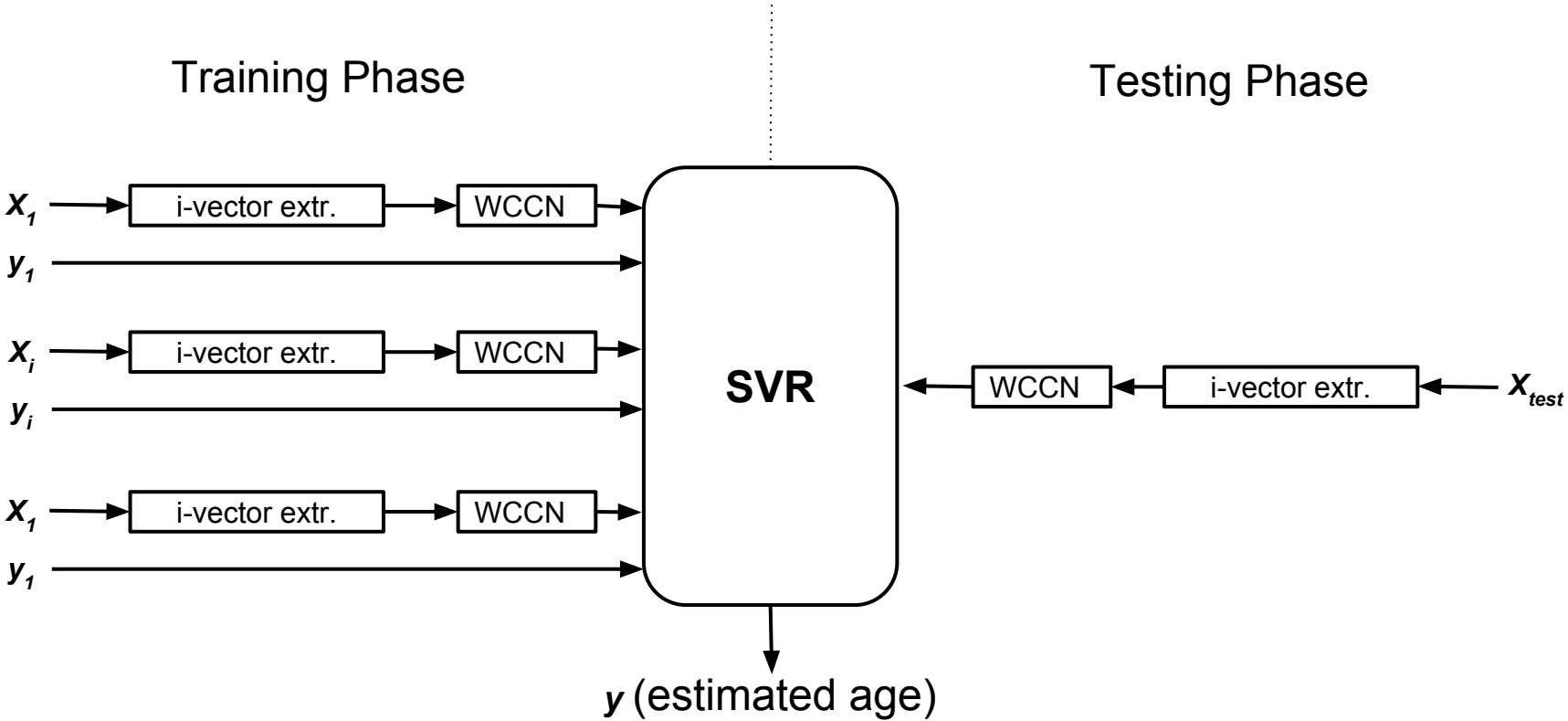
- Foreign exchange forecasting and trading
- Forecasting plant disease by leaf wetness prediction
- Flow forecasting

Automatic speaker's age identification

$S_{tr} = (X_1, Y_1), \ldots, (X_P, Y_P)$, X_p and Y_p - the p-th speech utterance and its age label

The goal is to create a system, which will predict, for an unseen utterance X_{test}, its label Y_{test} accurately.

State-of-the-art baseline approach



Neural network system description

- i-vectors
- WCCN
- NN-backend
  - MLP with single hidden layer (1024 neurons)
  - Minimum squared error objective function
  - SGD training algorithm
  - l_2 -regularization

Results

	MAE (females)	MAE (males)
Baseline method	5.75	6.65
NN-based method	**5.49**	**6.35**

Mean absolute error (MAE) in years for female and male speakers of NIST SRE 2008, NIST SRE 2010 datasets

Thank you!

References (1)

- C. M. Bishop, “Pattern recognition and machine learning,” 2006.
- D. E. Rumelhart, G. E. Hinton, and R. J. Williams, “Learning representations by backpropagating errors,” *Nature*, vol. 323, pp. 533–536, 1986.
- R. Caruana, S. Lawrence, and L. Giles, “Overfitting in neural nets: Backpropagation, conjugate gradient, and early stopping,” *Advances in Neural Information Processing Systems*, vol. 13, pp. 402–408, 2001.
- N. Srivastava, G. Hinton, A. Krizhevsky, I. Sutskever, and R. Salakhutdinov, “Dropout: A simple way to prevent neural networks from overfitting,” *Journal of Machine Learning Research*, vol. 15, pp. 1929–1958, 2014.
- A. Krogh and J. Hertz, “A simple weight decay can improve generalization,” *Advances in Neural Information Processing Systems*, vol. 4, pp. 950–957, 1992
- D. Sprecht ,”A General Regression Neural Network”, *IEEE Transactions on Neural Networks*, vol.2-6, pp. 568-576, 1991
- H. K. Cigizoglu ,”Generalized regression neural network in monthly flow forecasting”, *Civil Engineering and Environmental Systems,* vol. 22-2, pp. 71-81, 2005

References (2)

- Y. Chtioui, S.Panigrahi, L. Francl, “A generalized regression neural network and its application for leaf wetness prediction to forecast plant disease”, *Chemometrics and Intelligent Laboratory Systems,* vol. 48-1, pp. 47–58, 1999
- A. Chen , M.T. Leung, “Regression neural network for error correction in foreign exchange forecasting and trading”,*Computers & Operations Research,* vol.31, pp.1049-1068, 2004
- N. Dehak, P. J. Kenny, R. Dehak, P. Dumouchel, and P. Ouellet, “Front-end factor analysis for speaker verification,” *IEEE Tansactions on audio, speech, and language processing*, vol. 19, pp. 788–798, 2011.
- M. H. Bahari, M. McLaren, H. V. hamme, and D. van Leeuwen, “Speaker age estimation using i-vectors,” *Engineering Applications of Artificial Intelligence*, vol. 34, pp. 99–108, 2014.