

Multi-task Learning with a Shared Foundation Model

Abstract

A foundation neural network is trained to perform a first computational task. The foundation model has a number of layers, each including a number of functions defined by a set of numerical parameters, and the sets of parameters are trained to teach the foundation neural network the first computational task. Typically, each function receives an input vector (i.e. a plurality of input values), and generates an output vector (i.e. a plurality of output values). The foundation neural network is adapted to form an adapted neural network. In the adapted neural network, for at least one of these functions, a linear transformation is applied to the output (and/or input) values of the function. To learn the second computational task, parameters defining the linear transformation are trained, using a training database of examples of the second computational task, while substantially not changing the numeral parameters defining the functions.

Fig. 2