

CLAIMS

1. A data processing device comprising:
a data processing unit for training a neural network; and
an encoding unit for generating encoded data in which model header information that identifies a model of the neural network, layer header information that identifies one or more layers of the neural network and indicates a configuration of each of the layers of the neural network, and layer by layer edge-weight information of each of edges belonging to each of the one or more layers identified by the layer header information are encoded.
2. The data processing device according to claim 1, wherein the encoding unit encodes the weight information of each of the edges belonging to each of the one or more layers, ~~on a bit plane by bit plane basis from higher bits~~, as a flag indicating whether or not a weight value of each of the edges is zero, and a weight of a corresponding one of the edges when the flag indicates that the weight value of the corresponding one of the edges is not zero, separately.
3. The data processing device according to claim 1 or 2, wherein the encoding unit encodes weight information of edges belonging to each of the one or more layers identified by the layer header information.

4. The data processing device according to claim 1 or 2, wherein the encoding unit encodes a difference between an edge weight value and a specific value.
5. The data processing device according to claim 1 or 2, wherein the encoding unit encodes the edge weight information as base encoded data and enhancement encoded data separately.
6. The data processing device according to claim 1 or 2, comprising a decoding unit for decoding the encoded data generated by the encoding unit, wherein
the data processing unit trains the neural network using information decoded by the decoding unit.
7. A data processing system comprising:
a first data processing device including:
a first data processing unit for training a neural network; and
an encoding unit for generating encoded data in which model header information that identifies a model of the neural network, layer header information that identifies one or more layers of the neural network and indicates a configuration of each of the layers of the neural network, and ~~layer-by-layer edge-weight information~~ of each of edges belonging to each of the one or more layers identified by the layer header information are encoded; and
a second data processing device including:

a decoding unit for decoding the encoded data generated by the encoding unit; and

a second data processing unit for creating the neural network using information decoded by the decoding unit, and performing data processing using the neural network.

8. The data processing system according to claim 7, wherein the encoding unit encodes information about a portion of the neural network up to an intermediate layer, and

the second data processing device performs data processing using, as a feature, data outputted from the intermediate layer of the neural network.

9. A data processing method comprising the steps of:

training a neural network by a data processing unit; and

generating, by an encoding unit, encoded data in which model header information that identifies a model of the neural network, layer header information that identifies one or more layers of the neural network and indicates a configuration of each of the layers of the neural network, and ~~layer-by-layer edge~~ weight information of each of edges belonging to each of the one or more layers identified by the layer header information are encoded.