

CLAIMS

1. A system comprising:
a plurality of artificial intelligence processors;
one or more controllers; and
memory having stored thereon program code executable by the one or more controllers and the plurality of artificial intelligence processors, the program code causing the system to:

detect a computing error in a first artificial intelligence processor of a plurality of artificial intelligence processors during a first processing iteration of data from a data set;

eliminate the error from the first artificial intelligence processor;

and

load a model in one or more of the artificial intelligence processors including the first artificial intelligence processor, wherein the model corresponds to a same model processed by the plurality of artificial intelligence processors during the first processing iteration of the data from the data set.
2. The system of claim 1 wherein the plurality of artificial intelligence processors other than the first artificial intelligence processor wait while the first artificial intelligence processor eliminates the error, and wherein the plurality of processors process data from the data set on a next processing iteration at the same time using a second same model generated from the same model used on said first processing iteration.
3. The system of claim 1 wherein the computing error is detected during a result aggregation phase of the first processing iteration, and wherein at least a portion of the plurality of artificial intelligence processors wait for the first artificial intelligence processor to produce a valid result during the aggregation phase before completing the result aggregation phase.
4. The system of claim 3 wherein the first artificial intelligence processor sends an invalid result indicator to the at least a portion of the plurality of artificial intelligence processors to trigger the wait.
5. The system of claim 3 wherein the result aggregation phase is an All-Reduce.
6. The system of claim 1 wherein said loading the model comprises loading different portions of the model in the one or more artificial intelligence processors including the first artificial intelligence processor, the method further comprising

processing a first portion of the data, received by the first artificial intelligence processor on the first processing iteration, in the plurality of the artificial intelligence processors including the first artificial intelligence processor.

7. The system of claim 1 wherein said loading the model comprises loading the model in the first artificial intelligence processor, the system further processing a first portion of the data, received by the first artificial intelligence processor on the first processing iteration, in the first artificial intelligence processor.

8. The system of claim 1 wherein the model is received in the first artificial intelligence processor from a controller.

9. The system of claim 1 wherein the model is received in the first artificial intelligence processor from one or more other processors of the plurality of artificial intelligence processors.

10. The system of claim 1 wherein the model is received in the first artificial intelligence processor from a local memory of the first artificial intelligence processor.

11. The system of claim 1 wherein the model comprises artificial intelligence parameters.

12. The system of claim 1 wherein the model comprises neural network weights.

13. The system of claim 1 wherein the data set is a training data set.

14. A computer readable storage medium having stored thereon program code executable by a computer system, the program code causing the computer system to:

detecting a computing error in a first artificial intelligence processor of a plurality of artificial intelligence processors during a first processing iteration of data from a data set;

eliminating the error from the first artificial intelligence processor; and

loading a model in one or more of the artificial intelligence processors including the first artificial intelligence processor, wherein the model corresponds to a same model processed by the plurality of artificial intelligence processors during the first processing iteration of the data from the data set.

15. An error recovery method comprising:

detecting a computing error in a first artificial intelligence processor of a plurality of artificial intelligence processors during a first processing iteration of data from

a data set;

eliminating the error from the first artificial intelligence processor; and

loading a model in one or more of the artificial intelligence processors including the first artificial intelligence processor, wherein the model corresponds to a same model processed by the plurality of artificial intelligence processors during the first processing iteration of the data from the data set.