

WHAT IS CLAIMED IS:

1. A system of generating a hybrid sensor in industrial processing, comprising:
5 a hardware processor; and
a storage device coupled with the hardware processor;
the hardware processor configured to at least:
receive a first time series of upstream sensor data having forward dependence
to a target variable;
10 receive a second time series of downstream sensor data having backward
dependence to the target variable;
receive a time series of measured target variable data associated with the target
variable, the target variable having a measuring frequency which is lower than measuring
frequencies associated with the upstream sensor data and the downstream sensor data;
15 determine a first time window representing a lag of the forward dependence
between the upstream sensor data and the target variable;
determine a second time window representing a lag of the backward
dependence between the downstream sensor data and the target variable; and
train a machine learning model based on a training data set including at least
20 the first time series of upstream sensor data in the first time window, the second time series of
downstream sensor data in the second time window, and the time series of measured target
variable data to estimate a value of the target variable at a given time.
2. The system of claim 1, where in the hardware processor is further configured to
25 create a causality relationship data structure forward linking in time the first time series of the
upstream sensor data with the time series of measured target variable data, backward linking
in time the second time series of the downstream sensor data with the times series of
measured target variable data.
- 30 3. The system of claim 2, wherein the upstream sensor data and the downstream
sensor data are selected based on the causality relationship data structure.
4. The system of claim 1, wherein the trained machine learning model estimates the
value of the target variable at the given time, wherein based on the value of the target variable

estimated for the given time, the hardware processor further controls a set point in an industrial process.

- 5 5. The system of claim 4, wherein the industrial process includes oil sand processing, and the target variable includes a property associated with tailing.

- 10 6. The system of claim 1, wherein the machine learning model includes an aggregation of a first neural network trained to predict the value of the target variable based on the first time series of the upstream sensor data's forward dependence, a second neural network trained to predict the value of the target variable based on the second time series of the downstream sensor data's backward dependence, and a third neural network trained to predict the value of the target variable based on the times series of measured target variable data.

- 15 7. The system of claim 6, wherein a measured target variable's report time has a delay from the target variable's harvest time, and the training data set including the time series of measured target variable data includes data of the measured target variable determined at report time shifted to the measured target variable's corresponding harvest time, wherein the machine learning model at runtime estimates the value of the target variable at
20 the given time without the delay.

- 25 8. The system of claim 1, wherein the first time series of upstream sensor data includes a multi-dimensional time series data and the second time series of downstream sensor data includes a multi-dimensional time series data.

9. The system of claim 1, wherein the machine learning model includes a neural network model.

- 30 10. A computer program product comprising a computer readable storage medium having program instructions embodied therewith, the program instructions executable by a device to cause the device to:

receive a first time series of upstream sensor data having forward dependence to a target variable;

receive a second time series of downstream sensor data having backward dependence to the target variable;

receive a time series of measured target variable data associated with the target variable, the target variable having a measuring frequency which is lower than measuring
5 frequencies associated with the upstream sensor data and the downstream sensor data;

determine a first time window representing a lag of the forward dependence between the upstream sensor data and the target variable;

determine a second time window representing a lag of the backward dependence between the downstream sensor data and the target variable; and

10 train a machine learning model based on a training data set including at least the first time series of upstream sensor data in the first time window, the second time series of downstream sensor data in the second time window, and the time series of measured target variable data to estimate a value of the target variable at a given time.

15 11. The computer program product of claim 10, wherein the device is further caused to create a causality relationship data structure forward linking in time the first time series of the upstream sensor data with the time series of measured target variable data, and backward linking in time the second time series of the downstream sensor data with the times series of measured target variable data.

20

12. The computer program product of claim 11, wherein the upstream sensor data and the downstream sensor data are selected based on the causality relationship data structure.

25 13. The computer program product of claim 10, wherein the trained machine learning model estimates the value of the target variable at the given time, wherein based on the value of the target variable estimated for the given time, the hardware processor further controls a set point in an industrial process.

30 14. The computer program product of claim 13, wherein the industrial process includes oil sand processing, and the target variable includes a property associated with tailing.

15. The computer program product of claim 10, wherein the machine learning model includes an aggregation of a first neural network trained to predict the value of the target

variable based on the first time series of the upstream sensor data's forward dependence, a second neural network trained to predict the value of the target variable based on the second time series of the downstream sensor data's backward dependence, and a third neural network trained to predict the value of the target variable based on the times series of measured target
5 variable data.

16. The computer program product of claim 15, wherein a measured target variable's report time has a delay from the target variable's harvest time, and the training data set including the time series of measured target variable data includes data of the measured target
10 variable determined at report time shifted to the measured target variable's corresponding harvest time, wherein the machine learning model at runtime estimates the value of the target variable at the given time without the delay.

17. The computer program product of claim 10, wherein the first time series of
15 upstream sensor data includes a multi-dimensional time series data and the second time series of downstream sensor data includes a multi-dimensional time series data.

18. The system of claim 1, wherein the machine learning model includes a neural
20 network model.

19. A method of generating a hybrid sensor in industrial processing, comprising:
receiving a first time series of upstream sensor data having forward dependence to a
target variable;

receiving a second time series of downstream sensor data having backward
25 dependence to the target variable;

receiving a time series of measured target variable data associated with the target
variable, the target variable having a measuring frequency which is lower than measuring
frequencies associated with the upstream sensor data and the downstream sensor data;

determining a first time window representing a lag of the forward dependence
30 between the upstream sensor data and the target variable;

determining a second time window representing a lag of the backward dependence
between the downstream sensor data and the target variable; and

training a machine learning model based on a training data set including at least the
first time series of upstream sensor data in the first time window, the second time series of

downstream sensor data in the second time window, and the time series of measured target variable data to estimate a value of the target variable at a given time.

20. The method of claim 19, wherein the first time series of upstream sensor data
- 5 includes a multi-dimensional time series data and the second time series of downstream sensor data includes a multi-dimensional time series data, and wherein the machine learning model includes a neural network model.