

CLAIMS

1. A computer-implemented method for prediction modelling, the method comprising:
receiving, using a processor, a plurality of data inputs from a process;
selecting, using the processor, a data subset from the plurality of data inputs by solving linear programming to obtain a solution;
building, using the processor, an optimal decision tree ("ODT") based on the data subset; and
alerting, using the processor, a user when a prediction of the optimal decision tree is greater than a threshold value.
2. The computer-implemented method of claim 1 further comprising recommending, using the processor, a control action when the prediction of the optimal decision tree is greater than the threshold value.
3. The computer-implemented method of either of the preceding claims, wherein selecting, using the processor, a data subset from the plurality of data inputs comprises performing, using the processor, a hyperplane data-selection to select the data subset.
4. The computer-implemented method of any of the preceding claims, further comprising performing, using the processor, data selection within each leaf node of the ODT to select the data subset.
5. The computer-implemented method of any of the preceding claims, further comprising using, using the processor, a set of cutting planes to strengthen the ODT.
6. The computer-implemented method of any of the preceding claims, wherein selecting, using the processor, a data subset from the plurality of data inputs comprises selecting the data subset such that points inside a convex hull of the data subset are maximized.
7. The computer-implemented method of any of the preceding claims, wherein the ODT comprises a multivariable decision tree.
8. A system for prediction modelling, the system comprising:
a memory;
a processor communicatively coupled to the memory, the processor operable to execute instructions stored in the memory, the instructions causing the processor to:
receive a plurality of data inputs from a process;
select a data subset from the plurality of data inputs by solving linear programming to obtain a solution;
build an optimal decision tree ("ODT") based on the data subset; and

alert a user when a prediction of the optimal decision tree is greater than a threshold value.

9. The system of claim 8, wherein the instructions further cause the processor to recommend a control action when the prediction of the optimal decision tree is greater than the threshold value.
10. The system of either of claims 8 or 9, wherein selecting a data subset from the plurality of data inputs comprises performing a hyperplane data-selection to select the data subset.
11. The system of any of claims 8 to 10, wherein the instructions further cause the processor to perform data selection within each leaf node of the ODT to select the data subset.
12. The system of claims 8 to 11, wherein the instructions further cause the processor to use a set of cutting planes to strengthen the ODT.
13. The system of claims 8 to 12, wherein selecting a data subset from the plurality of data inputs comprises selecting the data subset such that points inside a convex hull of the data subset are maximized.
14. The system of claims 8 to 13, wherein the ODT comprises a multivariable decision tree.
15. A computer program product for prediction modelling, the computer program product comprising:
a computer readable storage medium readable by a processing circuit and storing instructions for execution by the processing circuit for performing a method according to any of claims 1 to 7.
16. A computer program stored on a computer readable medium and loadable into the internal memory of a digital computer, comprising software code portions, when said program is run on a computer, for performing the method of any of claims 1 to 7.