

CLAIMS

1. ~~A method comprising:
receiving, at a computing device, a plurality of data records and a plurality of variables;
determining, for each attribute of each data record of a first subset of the plurality of data records, a numeric representation, wherein each data record of the first subset of the plurality of data records is associated with a label;
determining, for each attribute of each variable of a first subset of the plurality of variables, a numeric representation, wherein each variable of the first subset of the plurality of variables is associated with the label;
generating, by a first plurality of encoder modules, and based on the numeric representation for each attribute of each data record of the first subset of the plurality of data records, a vector for each attribute of each data record of the first subset of the plurality of data records;
generating, by a second plurality of encoder modules, and based on the numeric representation for each attribute of each variable of the first subset of the plurality of variables, a vector for each attribute of each variable of the first subset of the plurality of variables;
generating, based on the vector for each attribute of each data record of the first subset of the plurality of data records, and based on the vector for each attribute of each variable of the first subset of the plurality of variables, a concatenated vector;
training, based on the concatenated vector, a model architecture comprising a predictive model, the first plurality of encoder modules, and the second plurality of encoder modules; and
outputting the model architecture.~~
2. ~~The method of claim 1, wherein each attribute of each of the plurality of data records comprises an input sequence.~~
3. ~~The method of claim 1, wherein each data record of the plurality of data records is associated with one or more variables of the plurality of variables.~~
4. ~~The method of claim 1, wherein the model architecture is trained according to a first~~

set of hyperparameters associated with one or more attributes of the plurality of data records and one or more attributes of the plurality of variables.

5. ~~The method of claim 2, further comprising:
optimizing the model architecture based on a second set of hyperparameters and a cross-validation technique;~~
6. ~~The method of claim 1, wherein determining, for each attribute of each variable of the first subset of the plurality of variables, the numeric representation comprises:
determining, by a plurality of tokenizers, for at least one attribute of at least one variable of the first subset of the plurality of variables, a token.~~
7. ~~The method of claim 7, wherein the at least one attribute of the at least one variable comprises at least a non-numeric portion, and wherein the token comprises the numeric representation for the at least one attribute of the at least one variable.~~
8. ~~A method comprising:
receiving, at a computing device, a data record and a plurality of variables;
determining, for each attribute of the data record, a numeric representation;
determining, for each attribute of each variable of the plurality of variables, a numeric representation;
generating, by a first plurality of trained encoder modules, and based on the numeric representation for each attribute of the data record, a vector for each attribute of the data record;
generating, by a second plurality of trained encoder modules, and based on the numeric representation for each attribute of each variable of the plurality of variables, a vector for each attribute of each variable of the plurality of variables;
generating, based on the vector for each attribute of the data record, and based on the vector for each attribute of each variable of the plurality of variables, a concatenated vector; and
determining, by a trained predictive model, based on the concatenated vector, one or more of a prediction or a score associated with the data record.~~
9. ~~The method of claim 8, wherein the prediction comprises a binary label.~~

- ~~10. — The method of claim 8, wherein the score is indicative of a likelihood that a first label applies to the data record.~~
- ~~11. — The method of claim 8, wherein the first plurality of trained encoder modules comprises a plurality of neural network blocks.~~
- ~~12. — The method of claim 8, wherein the second plurality of trained encoder modules comprises a plurality of neural network blocks.~~
- ~~13. — The method of claim 8, wherein determining, for each attribute of each variable of the plurality of variables, the numeric representation comprises:~~
~~determining, by a plurality of tokenizers, for at least one attribute of at least one variable of the plurality of variables, a token.~~
- ~~14. — The method of claim 13, wherein the at least one attribute of the at least one variable comprises at least a non-numeric portion, and wherein the token comprises the numeric representation for the at least one attribute of the at least one variable.~~
- ~~15. — A method comprising:~~
~~receiving, at a computing device, a first plurality of data records and a first plurality of variables associated with a label;~~
~~determining, for each attribute of each data record of the first plurality of data records, a numeric representation;~~
~~determining, for each attribute of each variable of the first plurality of variables, a numeric representation;~~
~~generating, by a first plurality of trained encoder modules, and based on the numeric representation for each attribute of each data record of the first plurality of data records, a vector for each attribute of each data record of the first plurality of data records;~~
~~generating, by a second plurality of trained encoder modules, and based on the numeric representation for each attribute of each variable of the first plurality of~~

~~variables, a vector for each attribute of each variable of the first plurality of variables;~~

~~generating, based on the vector for each attribute of each data record of the first plurality of data records, and based on the vector for each attribute of each variable of the first plurality of variables, a concatenated vector; and~~
~~retraining, based on the concatenated vector, a trained predictive model, the first plurality of encoder modules, and the second plurality of encoder modules.~~

~~16. The method of claim 15, further comprising: outputting the retrained predictive model.~~

~~17. The method of claim 15, wherein the first plurality of trained encoder modules are trained based on a plurality of training data records associated with the label and a first set of hyperparameters, wherein the first plurality of data records are associated with a second set of hyperparameters that differ at least partially from the first set of hyperparameters.~~

~~18. The method of claim 17, wherein the second plurality of trained encoder modules are trained based on a plurality of training variables associated with the label and the first set of hyperparameters, wherein the first plurality of variables are associated with the second set of hyperparameters.~~

~~19. The method of claim 17, wherein retraining the first plurality of encoder modules comprises: retraining, based on the second set of hyperparameters, the first plurality of encoder modules.~~

~~20. The method of claim 17, wherein retraining the second plurality of encoder modules comprises: retraining, based on the second set of hyperparameters, the second plurality of encoder modules.~~

1. A method comprising:
receiving, at a computing device, a plurality of data records and a plurality of
variables;

determining, for each attribute of each data record of a first subset of the plurality of data records, a numeric representation, wherein each data record of the first subset of the plurality of data records is associated with a label;

determining, for each attribute of each variable of a first subset of the plurality of variables, a numeric representation, wherein each variable of the first subset of the plurality of variables is associated with the label;

generating, by a first plurality of encoder modules, and based on the numeric representation for each attribute of each data record of the first subset of the plurality of data records, a vector for each attribute of each data record of the first subset of the plurality of data records;

generating, by a second plurality of encoder modules, and based on the numeric representation for each attribute of each variable of the first subset of the plurality of variables, a vector for each attribute of each variable of the first subset of the plurality of variables;

generating, based on the vector for each attribute of each data record of the first subset of the plurality of data records, and based on the vector for each attribute of each variable of the first subset of the plurality of variables, a concatenated vector;

training, based on the concatenated vector, a model architecture comprising: a predictive model, the first plurality of encoder modules, and the second plurality of encoder modules; and

outputting the model architecture.

2. The method of claim 1, wherein each attribute of each data record of the plurality of data records comprises an input sequence.
3. The method of any one of claims 1 or 2, wherein each data record of the plurality of data records is associated with one or more variables of the plurality of variables.
4. The method of any one of the preceding claims, wherein the first plurality of encoder modules comprises a first plurality of neural network blocks and the second plurality of encoder modules comprises a second plurality of neural network blocks.
5. The method of any one of the preceding claims, wherein the model architecture is

trained according to a first set of hyperparameters associated with one or more attributes of the plurality of data records and one or more attributes of the plurality of variables.

6. The method of claim 5, further comprising:
optimizing the model architecture based on a second set of hyperparameters and a cross-validation technique.
7. The method of any one of the preceding claims, wherein determining, for each attribute of each variable of the first subset of the plurality of variables, the numeric representation comprises:
determining, by a plurality of tokenizers, for at least one attribute of at least one variable of the first subset of the plurality of variables, a token.
8. The method of claim 7, wherein the at least one attribute of the at least one variable comprises at least a non-numeric portion, and wherein the token comprises the numeric representation for the at least one attribute of the at least one variable.
9. The method of any one of the preceding claims, further comprising:
receiving an input data record and a plurality of input variables;
determining, for each attribute of the input data record, an input numeric representation;
determining, for each attribute of each input variable of the plurality of input variables, an input numeric representation;
generating, by the first plurality of encoder modules, and based on the input numeric representation for each attribute of the input data record, an input vector for each attribute of the input data record;
generating, by the second plurality of encoder modules, and based on the input numeric representation for each attribute of each input variable of the plurality of input variables, an input vector for each attribute of each input variable of the plurality of input variables;
generating, based on the input vector for each attribute of the input data record, and based on the input vector for each attribute of each input variable of the plurality of input variables, an input concatenated vector; and

determining, by the predictive model, based on the input concatenated vector, one or more of a prediction or a score associated with the input data record.

10. The method of claim 9, wherein the prediction comprises a binary label.

11. The method of any one of claims 9 or 10, wherein the score is indicative of a likelihood that a first label applies to the input data record.

12. The method of any one of the preceding claims, further comprising: retraining, based on an additional concatenated vector determined based on a vector for each attribute of each data record of an additional plurality of data records, and based on a vector for each attribute of each additional variable of a plurality of additional variables: the predictive model, the first plurality of encoder modules, and the second plurality of encoder modules.

13. The method of claim 12, wherein the first plurality of encoder modules are initially trained based on the plurality of data records being associated with a label and based on a first set of hyperparameters, wherein the additional plurality of data records are associated with a second set of hyperparameters that differ at least partially from the first set of hyperparameters.

14. The method of claim 13, wherein the second plurality of encoder modules are initially trained based on the plurality of variables being associated with the label and based on the first set of hyperparameters, wherein the additional plurality of variables are associated with the second set of hyperparameters.

15. The method of claim 14, wherein retraining the first plurality of encoder modules comprises retraining, based on the second set of hyperparameters, the first plurality of encoder modules and wherein retraining the second plurality of encoder modules comprises retraining, based on the second set of hyperparameters, the second plurality of encoder modules.