CLAIMS

1. A method comprising:

determining a training dataset for training a model, the training dataset comprising a plurality of example documents, each example document being associated with a document type label;

for each example document in the training dataset:

providing the example document to a numerical representation generation model to generate a numerical representation of an example;

providing the numerical representation of the example document to a document type attribute predictor to generate a predicted document type;

determining a loss value based on a score of the predicted document type and the document type label associated with the example; and

adjusting one or more weights of the document type attribute predictor based the determined loss value; and

determining the numerical representation generation model and the document type attribute predictor to be a trained document type prediction model.

2. The method of claim 1, wherein the document type label is one of: (i) a bill or invoice; (ii) a credit note; and (iii) a spend money transaction.

3. The method of claim 1 or claim 2, wherein the trained document type prediction model is a neural network.

4. The method of claim 3, wherein the numerical representation generation model comprise an embedding layer of the neural network.

5. The method of claim 4, wherein the example documents are accounting records.

6. The method of any one of the preceding claims, wherein the example documents are tokenised example documents.

7. The method of any one of the preceding claims, further comprising generating a corpus of tokens using the plurality of example documents of the training dataset, wherein each token of the corpus is associated with a respective vector.

8. The method of claim 7, wherein the numerical representation generation model generates numerical representations of the examples using the corpus.

9. The method of any one of the preceding claims, further comprising:

deploying the trained document type prediction model on an accounting system.

10. The method of any one of the preceding claims, wherein providing the numerical representation of the example document to a document type attribute predictor to generate a predicted document type comprises:

providing the numerical representation of the example document to recurrent neural network, wherein the recurrent neural network is configured to output a processed numerical representation comprising encoded sequence information about tokens of the example document; and

providing the processed numerical representation to the document type attribute predictor to generate the predicted document type.

11. The method of claim 10, wherein the recurrent neural network comprises a gated recurrent unit (GRU), a long short-term memory (LSTM), or a transformer with appropriate positional embeddings.

12. The method of any one of the preceding claims, wherein the trained document type prediction model comprises: (i) one or more binary classifiers; (ii) a multi-class classifier; or (iii) one or more binary classifiers and a multi-class classifier.

13. The method of any one of the preceding claims, wherein the trained document type prediction model comprises a neural network and the neural network comprises at least two

activation functions, and wherein one of the at least two activation functions is applied to each layer of the neural network.

14. The method of claim 13, wherein the at least two activation functions comprise a first activation function and a second activation function, and the first activation function and the second activation function are applied in an alternating pattern to layers of the neural network.

15. The method of claim 14, wherein the first or second activation function comprises an unbounded activation function and the second or first activation function comprises a bounded activation function.

16. A system comprising:

one or more processors; and

memory comprising computer executable instructions, which when executed by the one or more processors, cause the system to perform the method of any one of claims 1 to 15.

17. A computer-readable storage medium storing instructions that, when executed by a computer, cause the computer to perform the method of any one of claims 1 to 15.