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 first hierarchical level category label and a second hierarchical level category label,

wherein first and second hierarchical levels are different hierarchical levels within a hierarchical structure of a common attribute, and the second hierarchical level category label is a subcategory of the first hierarchical level category label; and

for each example document in the training dataset:

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

providing the numerical representation of the example document to a first hierarchical level attribute predictor to generate a predicted first hierarchical level category;

determining a predicted second hierarchical level category;

determining a first loss value based on the predicted first hierarchical level category and the first hierarchical level category label associated with the example document;

determining a second loss value based on the predicted second hierarchical level category and the second hierarchical level category label associated with the example document;

determining a combined loss value based on the first loss value and the second loss value; and

adjusting one or more weights of the numerical representation generation model based on the combined loss value; and

determining the numerical representation generation model and the first hierarchical level attribute predictor to be a trained transaction attribute prediction model.

2. The method of claim 1, wherein the first hierarchical level is a lower level in the hierarchical structure than the second hierarchical level.

3. The method of any one of the preceding claims, wherein determining a predicted second hierarchical level category comprises:

providing the numerical representation of the example document to a second hierarchical level attribute predictor to generate the predicted second hierarchical level category.

4. The method of any one of claims 1 to 2, wherein determining a predicted second hierarchical level category comprises:

querying a hierarchical chart of accounts, each account being associated with multiple hierarchical level categories, using the predicted first hierarchical level category; and

determining the predicted second hierarchical level category as a parent category of the predicted first hierarchical level category.

5. The method of any one of the preceding claims, wherein each example document of the training dataset is further associated with a third hierarchical level category label, wherein a third hierarchical level is different from the first and second hierarchical levels within the hierarchical structure of the common attribute, the method further comprising:

for each example document in the training dataset:

determining a predicted third hierarchical level category; and

determining a third loss value based on the predicted third hierarchical level category and the third hierarchical level category label associated with the example document;

wherein determining the combined loss value is further based on the third loss value.

6. The method of claim 5, wherein determining the predicted third hierarchical level category comprises:

providing the numerical representation of the example document to a third hierarchical level attribute predictor to generate the predicted third hierarchical level category.

7. The method of claim 5, wherein determining the predicted third hierarchical level category comprises:

querying a hierarchical chart of accounts, each account being associated with multiple hierarchical level categories, using the predicted first hierarchical level category; and

determining the predicted third hierarchical level category as a grandparent category of the predicted first hierarchical level category.

8. The method of any one of the preceding claims, wherein the first hierarchical level of the hierarchical structure of the common attribute is an account code.

9. The method of any one of the preceding claims, wherein the second hierarchical level of the hierarchical structure of the common attribute is an account type.

10. The method of any one of claims 5 to 9, wherein the third hierarchical level of the hierarchical structure of the common attribute is an account class.

11. The method of any one of claims 1 to 8, wherein the second hierarchical level of the hierarchical structure of the common attribute is an account class.

12. The method of any one of claims 1 to 7, wherein the first hierarchical level of the hierarchical structure of the common attribute is an account type.

13. The method of any one of claim 12, wherein the second hierarchical level of the hierarchical structure of the common attribute is an account class.

14. The method of any one of the preceding claims, wherein determining the combined loss value comprises:

applying a first weighting to the first loss value and applying a second weighting to the second loss value.

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

deploying the trained transaction attribute prediction model on an accounting system.

16. The method of any one of the preceding claims, wherein the hierarchical structure is a hierarchical classification structure of a chart of accounts.

17. 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 16.

18. 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 16.