What Is Claimed Is:

1. A computer-implemented method for assisting a developer in improving an accuracy of a classification model, the computer implemented method comprising:

providing, by a computing device, the classification model including a plurality of features corresponding with a set of classes, wherein one or more features of the plurality of features correspond with at least one class of the set of classes;

selecting the one or more features of the plurality of features;

extracting one or more values for the one or more features selected;

determining at least one correlation of the one or more features with the set of classes respectively; and

generating at least one diagnostic example for the at least one correlation of the one or more features with the set of classes respectively;

wherein the at least one diagnostic example requires the developer to one of validate and invalidate a correctness of the at least one correlation of the one or more features with the set of classes respectively produced by the at least one diagnostic example,

when the developer invalidates the correctness of the at least one correlation of the one or more features with the set of classes respectively produced by the at least one diagnostic example, either automatically generating at least one additional diagnostic example for a training set that excludes the one or more features to suppress the at least one correlation, or automatically pulling from the training set identified examples in the training set with the at least one correlation.

2. The computer-implemented method as claimed in claim 1, wherein the classification model is created from a provided training set.

3. The computer-implemented method as claimed in claim 1, wherein the classification model is a pre-trained model.

4. The computer-implemented method as claimed in claim 1, wherein selecting the one or more features is based on, at least in part, a location of the one or more features in a neural network.

5. The computer-implemented method as claimed in claim 1, wherein selecting the one or more features is based on, at least in part, a received input selection from the developer.

6. The computer-implemented method as claimed in claim 1, wherein determining the at least one correlation includes computing the at least one correlation over a set of examples.

7. The computer-implemented method as claimed in claim 1, wherein determining the at least one correlation includes extracting the at least one correlation from the classification model.

8. The computer-implemented method as claimed in claim 1, further comprising selecting the at least one correlation among a plurality of correlations based on a highest correlation value.

9. The computer-implemented method as claimed in claim 8, further comprising using a neural network for determining the at least one correlation and using the neural network for selecting the at least one correlation used for generating the at least one diagnostic example.

10. The computer-implemented method as claimed in claim 1, wherein the at least one diagnostic example includes one of at least one text-based question, at least one image-based question, at least one audio-based question, at least one video-based question, and at least one data-based question for the developer to one of validate and invalidate the correctness of the correlation produced.

11. The computer-implemented method as claimed in claim 1, wherein generating the at least one diagnostic example comprises:

accessing a plurality of examples within a training set;

extracting the one or more features for each example of the plurality of examples; and

generating the at least one diagnostic example based upon, at least in part, the extracted features.

12. The computer-implemented method as claimed in claim 1, wherein each of the one or more features comprise at least one of a word, a part of a word, a phrase, a sentence, a

paragraph, a combination of words, a portion of an image, a portion of an audio, a portion of a video, and a portion of data.

13. The computer-implemented method as claimed in claim 1, further comprising:

receiving an input from the developer that one of validates and invalidates the correctness of the correlation in response to the at least one diagnostic example.

14. The computer-implemented method as claimed in claim 13, further comprising:

when the developer invalidates the correctness of the correlation selected, recommending the developer provide an additional set of examples used as training data for adjusting the classification model to suppress the correlation selected between the one or more features selected and the set of classes; and

receiving the additional set of examples.

15. The computer-implemented method as claimed in claim 13, further comprising:

when the developer invalidates the correctness of the correlation selected, automatically generating an additional set of examples used as training data for adjusting the classification model to suppress the correlation selected between the one or more features selected and the set of classes.

16. The computer-implemented method as claimed in claim 13, further comprising:

when the developer invalidates the correctness of the correlation selected, automatically generating an additional set of examples and recommending at least one of the developer revise and approve the additional set of examples such that the additional set of examples are used as training data for adjusting the classification model to suppress the correlation selected between the one or more features selected and the set of classes.

17. The computer-implemented method as claimed in claim 13, further comprising:

when the developer invalidates the correctness of the correlation selected, adjusting the classification model by modifying at least one parameter of the classification model.

18. The computer-implemented method as claimed in claim 13 further comprising:

adjusting the classification model; and

re-determining the at least one correlation of the one or more features selected upon adjusting the classification model.

19. The computer-implemented method as claimed in claim 1, wherein the classification model is a neural network classification model.

20. The computer-implemented method as claimed in claim 1, further comprising:

iteratively generating another diagnostic example for the developer for another correlation selected from the at least one correlation, wherein the another diagnostic example requires the developer to one of validate and invalidate the correctness of another correlation produced by the another diagnostic example.

21. A computer implemented method for assisting a developer in improving an accuracy of a classification process, the computer implemented method comprising:

providing an example set;

extracting a plurality of features from the example set, wherein one or more features of the plurality of features correspond with at least one class of a set of classes;

selecting the one or more features of the plurality of features;

determining at least one correlation of the one or more features with the set of classes respectively; and

generating at least one diagnostic example for the at least one correlation of the one or more features with the set of classes respectively;

wherein the at least one diagnostic example requires the developer to one of validate and invalidate a correctness of the at least one correlation of the one or more features with the set of classes respectively produced by the at least one diagnostic example,

when the developer invalidates the correctness of the at least one correlation of the one or more features with the set of classes respectively produced by the at least one diagnostic example, either automatically generating at least one additional diagnostic example for a training set that excludes the one or more features to suppress the at least one correlation, or automatically pulling from the training set identified examples in the training set with the at least one correlation.

22. A system for assisting a developer in improving an accuracy of a classification model, the system comprising:

the classification model including a plurality of features corresponding with a set of classes, wherein one or more features of the plurality of features correspond with at least one class of the set of classes;

a feature selector configured to select the one or more features of the plurality features;

a feature extractor for extracting one or more values for the one or more features;

a correlation engine configured to determine at least one correlation of the one or more features with the set of classes respectively; and

a diagnostic engine configured to generate at least one diagnostic example for the at least one correlation of the one or more features with the set of classes respectively;

wherein the at least one diagnostic example requires the developer to one of validate and invalidate a correctness of the at least one correlation of the one or more features with the set of classes respectively produced by the at least one diagnostic example,

when the developer invalidates the correctness of the at least one correlation of the one or more features with the set of classes respectively produced by the at least one diagnostic example, either automatically generating at least one additional diagnostic example for a training set that excludes the one or more features to suppress the at least one correlation, or automatically pulling from the training set identified examples in the training set with the at least one correlation.

23. The system as claimed in claim 22, wherein the diagnostic engine is further configured to:

iteratively generate another diagnostic example for the developer for another correlation selected from the at least one correlation, wherein the another diagnostic example requires the developer to at least one of validate and invalidate the correctness of another correlation produced by the another diagnostic example.

24. The system as claimed in claim 22, wherein the diagnostic engine is further configured to:

generate at least one of a text-based question, at least one image-based question, at least one audio-based question, at least one video-based question, and at least one data-based question for the developer to one of validate and invalidate the correctness of the correlation.

25. The system as claimed in claim 22, wherein each feature of the one or more features comprises at least one of a word, a part of a word, a phrase, a sentence, a paragraph, a

combination of words, a portion of an image, a portion of an audio, a portion of a video, and a portion of data.

26. The system as claimed in claim 22, further comprising:

an input interface for receiving an input from the developer that one of validates and invalidates the correctness of the correlation in response to the at least one diagnostic example.

27. The system as claimed in claim 22, further comprising:

a recommendation engine configured to recommend the developer to provide an additional set of a plurality of examples used as training data for adjusting the classification model to suppress the correlation selected between the one or more features selected and the set of classes when the developer invalidates the correctness of the correlation selected.

28. The system as claimed in claim 22, further comprising:

a recommendation engine configured to automatically generate an additional set of a plurality of examples used as training data for adjusting the classification model to suppress the correlation selected between the one or more features selected and the set of classes when the developer invalidates the correctness of the correlation selected.