WHAT IS CLAIMED IS:

1. A system, comprising:

at least one processor programmed or configured to:

receive a time series of historical data points, wherein the historical data points include values of a plurality of transaction features with regard to a plurality of time intervals during an initial time period;

determine a historical time period, wherein the historical time period comprises a first time period of the initial time period;

determine a contemporary time period, wherein the contemporary time period comprises a second time period of the initial time period that is after the historical time period, wherein the historical time period is longer than the contemporary time period;

determine a first time series of data points associated with a historical transaction metric from the historical time period;

determine a second time series of data points associated with a historical target transaction metric from the historical time period;

determine a third time series of data points associated with a contemporary transaction metric from the contemporary time period;

generate a machine learning model, wherein the machine learning model is configured to provide an output that comprises a predicted time series of data points associated with a contemporary target transaction metric, wherein, when generating the machine learning model, the at least one processor is programmed or configured to:

train the machine learning model based on the first time series of data points associated with the historical transaction metric, the second time series of data points associated with the historical target transaction metric, and the third time series of data points associated with the contemporary transaction metric; and

generate an output of the machine learning model, wherein the output of the machine learning model comprises the contemporary target transaction metric, and wherein the contemporary target transaction metric comprises a value of a target transaction metric during a targeted prediction period.;

provide the first time series of data points associated with the historical transaction metric, the second time series of data points associated with the historical target transaction metric, and the third time series of data points associated with the contemporary transaction metric as inputs to a processing layer of the machine learning model, wherein the processing layer comprises a fast Fourier transform (FFT) layer;

provide an output of the processing layer as an input to a feature extraction component of the machine learning model;

provide an output of the feature extraction component of the machine learning model as an input to a dual-attention component of the machine learning model; and

provide an output of the dual-attention component of the machine learning model as an input to a learning and prediction component of the machine learning model.

~~2. The system of claim 1, wherein, when training the machine learning model, the at least one processor is programmed or configured to:~~

~~provide the first time series of data points associated with the historical transaction metric, the second time series of data points associated with the historical target transaction metric, and the third time series of data points associated with the contemporary transaction metric as inputs to a processing layer of the machine learning model, wherein the processing layer comprises a fast Fourier transform (FFT) layer;~~

~~provide an output of the processing layer as an input to a feature extraction component of the machine learning model;~~

~~provide an output of the feature extraction component of the machine learning model as an input to a dual-attention component of the machine learning model; and~~

~~provide an output of the dual-attention component of the machine learning model as an input to a learning and prediction component of the machine learning model.~~

~~3.~~2. The system of claim 1~~,2~~, wherein, when providing the first time series of data points associated with the historical transaction metric, the second time

series of data points associated with the historical target transaction metric, and the third time series of data points associated with the contemporary transaction metric as inputs to the processing layer of the machine learning model, the at least one processor is programmed or configured to:

provide each feature of a first plurality of features of the first time series of data points associated with the historical transaction metric as an input to the FFT layer to generate a real part and an imaginary part for each feature of the first plurality of features;

provide each feature of a second plurality of features of the second time series of data points associated with the historical target transaction metric as an input to the FFT layer to generate a real part and an imaginary part for each feature of the second plurality of features; and

provide each feature of a third plurality of features of the third time series of data points associated with the contemporary transaction metric as an input to the FFT layer to generate a real part and an imaginary part for each feature of the third plurality of features.

~~4~~3. The system of claim ~~3~~2, wherein, when training the machine learning model, the at least one processor is programmed or configured to:

combine the real part for each feature of the first plurality of features to generate a combined real part of the first plurality of features;

combine the imaginary part for each feature of the first plurality of features to generate a combined imaginary part of the first plurality of features;

combine the real part for each feature of the second plurality of features to generate a combined real part of the second plurality of features;

combine the imaginary part for each feature of the second plurality of features to generate a combined imaginary part of the second plurality of features;

combine the real part for each feature of the third plurality of features to generate a combined real part of the third plurality of features; and

combine the imaginary part for each feature of the third plurality of features to generate a combined imaginary part of the third plurality of features; and

wherein the output of the processing layer of the machine learning model comprises the combined real part of the first plurality of features, the combined imaginary part of the first plurality of features, the combined real part of the second

plurality of features, the combined imaginary part of the second plurality of features, the combined real part of the third plurality of features, and the combined imaginary part of the third plurality of features.

45. The system of claim 43, wherein, when training the machine learning model, the at least one processor is programmed or configured to:

provide the combined real part of the first plurality of features as an input to a one-dimensional feature extraction convolutional layer;

provide the combined imaginary part of the first plurality of features as an input to the one-dimensional feature extraction convolutional layer;

provide the combined real part of the second plurality of features as an input to the one-dimensional feature extraction convolutional layer;

provide the combined imaginary part of the second plurality of features as an input to the one-dimensional feature extraction convolutional layer;

provide the combined real part of the third plurality of features as an input to the one-dimensional feature extraction convolutional layer;

provide the combined imaginary part of the third plurality of features as an input to the one-dimensional feature extraction convolutional layer; and

generate the output of the feature extraction component of the machine learning model based on an output of the one-dimensional feature extraction convolutional layer.

56. The system of claim 54, wherein the input to the dual-attention component of the machine learning model comprises a real input portion and an imaginary input portion,

wherein the real input portion comprises:

a first real input based on an output of the one-dimensional feature extraction convolutional layer for the combined real part of the first plurality of features;

a second real input based on an output of the one-dimensional feature extraction convolutional layer for the combined real part of the second plurality of features; and

a third real input based on an output of the one-dimensional feature extraction convolutional layer for the combined real part of the third plurality of features; and

wherein the imaginary input portion comprises:

a first imaginary input based on an output of the one-dimensional feature extraction convolutional layer for the combined imaginary part of the first plurality of features;

a second imaginary input based on an output of the one-dimensional feature extraction convolutional layer for the combined imaginary part of the second plurality of features; and

a third imaginary input based on an output of the one-dimensional feature extraction convolutional layer for the combined imaginary part of the third plurality of features; and

wherein, when training the machine learning model, the at least one processor is programmed or configured to:

generate a real component attention matrix based on the real input portion, wherein the real component attention matrix has a number of attention vectors that is equal to a number of features of the plurality of transaction features; and

generate an imaginary component attention matrix based on the imaginary input portion, wherein the imaginary component attention matrix has a number of attention vectors that is equal to a number of features of the plurality of transaction features.

6~~7~~. The system of claim ~~6~~5, wherein the at least one processor is further programmed or configured to:

generate the real input portion of the input to the dual-attention component of the machine learning model; and

generate the imaginary input portion of the input to the dual-attention component of the machine learning model;

wherein, when generating the real input portion of the input to the dual-attention component of the machine learning model, the at least one processor is programmed or configured to:

provide the output of the one-dimensional feature extraction convolutional layer for the combined real part of the first plurality of features as an input to a dropout layer to generate a real contemporary transaction metric feature vector as the first real input;

provide the output of the one-dimensional feature extraction convolutional layer for the combined real part of the second plurality of features as an input to the dropout layer to generate a real historical transaction metric feature vector as the second real input; and

provide the output of the one-dimensional feature extraction convolutional layer for the combined real part of the third plurality of features as an input to the dropout layer to generate a real historical target transaction metric feature vector as the third real input; and

wherein, when, generating the imaginary input portion of the input to the dual-attention component of the machine learning model, the at least one processor is programmed or configured to:

provide the output of the one-dimensional feature extraction convolutional layer for the combined imaginary part of the first plurality of features as an input to the dropout layer to generate an imaginary contemporary transaction metric feature vector as the first imaginary input;

provide the output of the one-dimensional feature extraction convolutional layer for the combined imaginary part of the second plurality of features as an input to the dropout layer to generate an imaginary historical transaction metric feature vector as the second imaginary input; and

provide the output of the one-dimensional feature extraction convolutional layer for the combined imaginary part of the third plurality of features as an input to the dropout layer to generate an imaginary historical target transaction metric feature vector as the third imaginary input.

7~~8~~. The system of claim ~~7~~6, wherein, when generating the real component attention matrix, the at least one processor is programmed or configured to:

generate a dot product of the real contemporary transaction metric feature vector and the real historical transaction metric feature vector;

provide the dot product of the real contemporary transaction metric feature vector and the real historical transaction metric feature vector as an input to a softmax layer to generate a real attention feature vector; and

generate a dot product of the real attention feature vector and the real historical target transaction metric feature vector to generate the real component attention matrix;

wherein, when generating the imaginary component attention matrix, the at least one processor is programmed or configured to:

generate a dot product of the imaginary contemporary transaction metric feature vector and the imaginary historical transaction metric feature vector;

provide the dot product of the imaginary contemporary transaction metric feature vector and the imaginary historical transaction metric feature vector as a input to a softmax layer to generate an imaginary attention feature vector; and

generate a dot product of the imaginary attention feature vector and the imaginary historical target transaction metric feature vector to generate the imaginary component attention matrix.

~~9. The system of claim 7, wherein, when generating the real component attention matrix, the at least one processor is programmed or configured to:~~

~~perform an addition function on the real contemporary transaction metric feature vector and the real historical transaction metric feature vector;~~

~~provide a result of the addition function of the real contemporary transaction metric feature vector and the real historical transaction metric feature vector as an input to a tanh activation function;~~

~~provide an output of the tanh activation function as a input to a softmax layer to generate a real attention feature vector; and~~

~~generate a dot product of the real attention feature vector and the real historical target transaction metric feature vector to generate the real component attention matrix;~~

~~wherein, when, generating the imaginary component attention matrix, the at least one processor is programmed or configured to:~~

perform an addition function on the imaginary contemporary transaction metric feature vector and the imaginary historical transaction metric feature vector;

provide the result of the addition function of the imaginary contemporary transaction metric feature vector and the imaginary historical transaction metric feature vector as an input to a tanh activation function;

provide an output of the tanh activation function as an input to a softmax layer to generate an imaginary attention feature vector; and

generate a dot product of the imaginary attention feature vector and the imaginary historical target transaction metric feature vector to generate the imaginary component attention matrix.

10. The system of claim 7, wherein, when providing the output of the dual-attention component of the machine learning model as the input to the learning and prediction component of the machine learning model, the at least one processor is programmed or configured to:

concatenate the real component attention matrix and the real contemporary transaction metric feature vector to generate a first concatenation;

provide the first concatenation as an input to a feature learning one-dimensional convolutional layer;

concatenate the imaginary component attention matrix and the imaginary contemporary transaction metric feature vector to generate a second concatenation;

provide the second concatenation as an input to the feature learning one-dimensional convolutional layer;

provide an input to a reshape layer based on an output of the feature learning one-dimensional convolutional layer for the first concatenation to generate a real component feature vector; and

provide an input to the reshape layer based on an output of the feature learning one-dimensional convolutional layer for the second concatenation to generate an imaginary component feature vector.

11. The system of claim 10, wherein, when generating the output of the machine learning model, the at least one processor is programmed or configured to:

~~generate a complex number feature vector based on the real component feature vector and the imaginary component feature vector; and~~

~~provide the complex number feature vector as an input to an inverse FFT layer to generate the contemporary target transaction metric.~~

~~1~~28. A method, comprising:

receiving, with at least one processor, a time series of historical data points, wherein the historical data points include values of a plurality of transaction features with regard to a plurality of time intervals during an initial time period;

determining, with at least one processor, a historical time period, wherein the historical time period comprises a first time period of the initial time period;

determining, with at least one processor, a contemporary time period, wherein the contemporary time period comprises a second time period of the initial time period that is after the historical time period, wherein the historical time period is longer than the contemporary time period;

determining, with at least one processor, a first time series of data points associated with a historical transaction metric from the historical time period;

determining, with at least one processor, a second time series of data points associated with a historical target transaction metric from the historical time period;

determining, with at least one processor, a third time series of data points associated with a contemporary transaction metric from the contemporary time period;

generating, with at least one processor, a machine learning model, wherein the machine learning model is configured to provide an output that comprises a predicted time series of data points associated with a contemporary target transaction metric, wherein generating the machine learning model comprises:

training the machine learning model based on the first time series of data points associated with the historical transaction metric, the second time series of data points associated with the historical target transaction metric, and the third time series of data points associated with the contemporary transaction metric; and

generating an output of the machine learning model, wherein the output of the machine learning model comprises the contemporary target transaction metric,

and wherein the contemporary target transaction metric comprises a value of a target transaction metric during a targeted prediction period.;

providing the first time series of data points associated with the historical transaction metric, the second time series of data points associated with the historical target transaction metric, and the third time series of data points associated with the contemporary transaction metric as inputs to a processing layer of the machine learning model, wherein the processing layer comprises a fast Fourier transform (FFT) layer;

providing an output of the processing layer as an input to a feature extraction component of the machine learning model;

providing an output of the feature extraction component of the machine learning model as an input to a dual-attention component of the machine learning model; and

providing an output of the dual-attention component of the machine learning model as an input to a learning and prediction component of the machine learning model.

~~13. The method of claim 12, wherein training the machine learning model comprises:~~

~~providing the first time series of data points associated with the historical transaction metric, the second time series of data points associated with the historical target transaction metric, and the third time series of data points associated with the contemporary transaction metric as inputs to a processing layer of the machine learning model, wherein the processing layer comprises a fast Fourier transform (FFT) layer;~~

~~providing an output of the processing layer as an input to a feature extraction component of the machine learning model;~~

~~providing an output of the feature extraction component of the machine learning model as an input to a dual-attention component of the machine learning model; and~~

~~providing an output of the dual-attention component of the machine learning model as an input to a learning and prediction component of the machine learning model.~~

~~14~~9. The method of claim ~~8~~13, wherein providing the first time series of data points associated with the historical transaction metric, the second time series of data points associated with the historical target transaction metric, and the third time series of data points associated with the contemporary transaction metric as inputs to the processing layer of the machine learning model comprises:

providing each feature of a first plurality of features of the first time series of data points associated with the historical transaction metric as an input to the FFT layer to generate a real part and an imaginary part for each feature of the first plurality of features;

providing each feature of a second plurality of features of the second time series of data points associated with the historical target transaction metric as an input to the FFT layer to generate a real part and an imaginary part for each feature of the second plurality of features; and

providing each feature of a third plurality of features of the third time series of data points associated with the contemporary transaction metric as an input to the FFT layer to generate a real part and an imaginary part for each feature of the third plurality of features.

~~15~~10. The method of claim ~~14~~9, wherein training the machine learning model comprises:

combining the real part for each feature of the first plurality of features to generate a combined real part of the first plurality of features;

combining the imaginary part for each feature of the first plurality of features to generate a combined imaginary part of the first plurality of features;

combining the real part for each feature of the second plurality of features to generate a combined real part of the second plurality of features;

combining the imaginary part for each feature of the second plurality of features to generate a combined imaginary part of the second plurality of features;

combining the real part for each feature of the third plurality of features to generate a combined real part of the third plurality of features; and

combining the imaginary part for each feature of the third plurality of features to generate a combined imaginary part of the third plurality of features; and

wherein the output of the processing layer of the machine learning model comprises the combined real part of the first plurality of features, the combined

imaginary part of the first plurality of features, the combined real part of the second plurality of features, the combined imaginary part of the second plurality of features, the combined real part of the third plurality of features, and the combined imaginary part of the third plurality of features.

~~16~~11. The method of claim ~~15~~10, wherein training the machine learning model comprises:

providing the combined real part of the first plurality of features as an input to a one-dimensional feature extraction convolutional layer;

providing the combined imaginary part of the first plurality of features as an input to the one-dimensional feature extraction convolutional layer;

providing the combined real part of the second plurality of features as an input to the one-dimensional feature extraction convolutional layer;

providing the combined imaginary part of the second plurality of features as an input to the one-dimensional feature extraction convolutional layer;

providing the combined real part of the third plurality of features as an input to the one-dimensional feature extraction convolutional layer;

providing the combined imaginary part of the third plurality of features as an input to the one-dimensional feature extraction convolutional layer; and

generating the output of the feature extraction component of the machine learning model based on an output of the one-dimensional feature extraction convolutional layer.

~~17~~12. The method of claim ~~16~~11, wherein the input to the dual-attention component of the machine learning model comprises a real input portion and an imaginary input portion,

wherein the real input portion comprises:

a first real input based on an output of the one-dimensional feature extraction convolutional layer for the combined real part of the first plurality of features;

a second real input based on an output of the one-dimensional feature extraction convolutional layer for the combined real part of the second plurality of features; and

a third real input based on an output of the one-dimensional feature extraction convolutional layer for the combined real part of the third plurality of features; and

wherein the imaginary input portion comprises:

a first imaginary input based on an output of the one-dimensional feature extraction convolutional layer for the combined imaginary part of the first plurality of features;

a second imaginary input based on an output of the one-dimensional feature extraction convolutional layer for the combined imaginary part of the second plurality of features; and

a third imaginary input based on an output of the one-dimensional feature extraction convolutional layer for the combined imaginary part of the third plurality of features; and

wherein training the machine learning model comprises:

generating a real component attention matrix based on the real input portion, wherein the real component attention matrix has a number of attention vectors that is equal to a number of features of the plurality of transaction features; and

generating an imaginary component attention matrix based on the imaginary input portion, wherein the imaginary component attention matrix has a number of attention vectors that is equal to a number of features of the plurality of transaction features.

~~18~~13. The method of claim ~~17~~12, wherein the method further comprises:

generating the real input portion of the input to the dual-attention component of the machine learning model; and

generating the imaginary input portion of the input to the dual-attention component of the machine learning model;

wherein generating the real input portion of the input to the dual-attention component of the machine learning model comprises:

providing the output of the one-dimensional feature extraction convolutional layer for the combined real part of the first plurality of features as an input to a dropout layer to generate a real contemporary transaction metric feature vector as the first real input;

providing the output of the one-dimensional feature extraction convolutional layer for the combined real part of the second plurality of features as an input to the dropout layer to generate a real historical transaction metric feature vector as the second real input; and

providing the output of the one-dimensional feature extraction convolutional layer for the combined real part of the third plurality of features as an input to the dropout layer to generate a real historical target transaction metric feature vector as the third real input; and

wherein generating the imaginary input portion of the input to the dual-attention component of the machine learning model comprises:

providing the output of the one-dimensional feature extraction convolutional layer for the combined imaginary part of the first plurality of features as an input to the dropout layer to generate an imaginary contemporary transaction metric feature vector as the first imaginary input;

providing the output of the one-dimensional feature extraction convolutional layer for the combined imaginary part of the second plurality of features as an input to the dropout layer to generate an imaginary historical transaction metric feature vector as the second imaginary input; and

providing the output of the one-dimensional feature extraction convolutional layer for the combined imaginary part of the third plurality of features as an input to the dropout layer to generate an imaginary historical target transaction metric feature vector as the third imaginary input.

~~19~~14. The method of claim ~~18~~13, wherein generating the real component attention matrix comprises:

generating a dot product of a real contemporary transaction metric feature vector and the real historical transaction metric feature vector;

providing the dot product of the real contemporary transaction metric feature vector and the real historical transaction metric feature vector as an input to a softmax layer to generate a real attention feature vector; and

generating a dot product of the real attention feature vector and the real historical target transaction metric feature vector to generate the real component attention matrix;

wherein generating the imaginary component attention matrix comprises:

generating a dot product of the imaginary contemporary transaction metric feature vector and the imaginary historical transaction metric feature vector;

providing the dot product of the imaginary contemporary transaction metric feature vector and the imaginary historical transaction metric feature vector as a input to a softmax layer to generate an imaginary attention feature vector; and

generating a dot product of the imaginary attention feature vector and the imaginary historical target transaction metric feature vector to generate the imaginary component attention matrix.

~~20. The method of claim 18, wherein generating the real component attention matrix comprises:~~

~~performing an addition function on the real contemporary transaction metric feature vector and the real historical transaction metric feature vector;~~

~~providing the result of the addition function of the real contemporary transaction metric feature vector and the real historical transaction metric feature vector as an input to a tanh activation function;~~

~~providing an output of the tanh activation function as a input to a softmax layer to generate a real attention feature vector; and~~

~~generating a dot product of the real attention feature vector and the real historical target transaction metric feature vector to generate the real component attention matrix;~~

~~wherein generating the imaginary component attention matrix comprises:~~

~~performing an addition function on the imaginary contemporary transaction metric feature vector and the imaginary historical transaction metric feature vector;~~

~~providing the result of the addition function of the imaginary contemporary transaction metric feature vector and the imaginary historical transaction metric feature vector as an input to a tanh activation function;~~

providing an output of the tanh activation function as a input to a softmax layer to generate an imaginary attention feature vector; and

generating a dot product of the imaginary attention feature vector and the imaginary historical target transaction metric feature vector to generate the imaginary component attention matrix.

21. The method of claim 17, wherein providing the output of the dual-attention component of the machine learning model as the input to the learning and prediction component of the machine learning model comprises:

concatenating the real component attention matrix and the real contemporary transaction metric feature vector to generate a first concatenation;

providing the first concatenation as an input to a feature learning one-dimensional convolutional layer;

concatenating the imaginary component attention matrix and the imaginary contemporary transaction metric feature vector to generate a second concatenation;

providing the second concatenation as an input to the feature learning one-dimensional convolutional layer;

providing an input to a reshape layer based on an output of the feature learning one-dimensional convolutional layer for the first concatenation to generate a real component feature vector; and

providing an input to the reshape layer based on an output of the feature learning one-dimensional convolutional layer for the second concatenation to generate an imaginary component feature vector.

22. The method of claim 21, wherein generating the output of the machine learning model comprises:

generating a complex number feature vector based on the real component feature vector and the imaginary component feature vector; and

providing the complex number feature vector as an input to an inverse FFT layer to generate the contemporary target transaction metric.

~~23.~~15. A computer program product comprising at least one non-transitory computer-readable medium including one or more instructions that, when executed by at least one processor, cause the at least one processor to:

receive a time series of historical data points, wherein the historical data points include values of a plurality of transaction features with regard to a plurality of time intervals during an initial time period;

determine a historical time period, wherein the historical time period comprises a first time period of the initial time period;

determine a contemporary time period, wherein the contemporary time period comprises a second time period of the initial time period that is after the historical time period, wherein the historical time period is longer than the contemporary time period;

determine a first time series of data points associated with a historical transaction metric from the historical time period;

determine a second time series of data points associated with a historical target transaction metric from the historical time period;

determine a third time series of data points associated with a contemporary transaction metric from the contemporary time period;

generate a machine learning model, wherein the machine learning model is configured to provide an output that comprises a predicted time series of data points associated with a contemporary target transaction metric, wherein, the one or more instructions that cause the at least one processor to generate the machine learning model, cause the at least one processor to:

train the machine learning model based on the first time series of data points associated with the historical transaction metric, the second time series of data points associated with the historical target transaction metric, and the third time series of data points associated with the contemporary transaction metric; and

generate an output of the machine learning model, wherein the output of the machine learning model comprises the contemporary target transaction metric, and wherein the contemporary target transaction metric comprises a value of a target transaction metric during a targeted prediction period~~.~~;

provide the first time series of data points associated with the historical transaction metric, the second time series of data points associated with the historical

target transaction metric, and the third time series of data points associated with the contemporary transaction metric as inputs to a processing layer of the machine learning model, wherein the processing layer comprises a fast Fourier transform (FFT) layer;

provide an output of the processing layer as an input to a feature extraction component of the machine learning model;

provide an output of the feature extraction component of the machine learning model as an input to a dual-attention component of the machine learning model; and

provide an output of the dual-attention component of the machine learning model as an input to a learning and prediction component of the machine learning model.

~~24.~~. ~~The computer program product of claim 23, wherein, the one or more instructions that cause the at least one processor to train the machine learning model, cause the at least one processor to:~~

~~provide the first time series of data points associated with the historical transaction metric, the second time series of data points associated with the historical target transaction metric, and the third time series of data points associated with the contemporary transaction metric as inputs to a processing layer of the machine learning model, wherein the processing layer comprises a fast Fourier transform (FFT) layer;~~

~~provide an output of the processing layer as an input to a feature extraction component of the machine learning model;~~

~~provide an output of the feature extraction component of the machine learning model as an input to a dual-attention component of the machine learning model; and~~

~~provide an output of the dual-attention component of the machine learning model as an input to a learning and prediction component of the machine learning model.~~

~~25~~16. The computer program product of claim ~~24~~15, wherein, the one or more instructions that cause the at least one processor to provide the first time series of data points associated with the historical transaction metric, the second time

series of data points associated with the historical target transaction metric, and the third time series of data points associated with the contemporary transaction metric as inputs to the processing layer of the machine learning model, cause the at least one processor to:

provide each feature of a first plurality of features of the first time series of data points associated with the historical transaction metric as an input to the FFT layer to generate a real part and an imaginary part for each feature of the first plurality of features;

provide each feature of a second plurality of features of the second time series of data points associated with the historical target transaction metric as an input to the FFT layer to generate a real part and an imaginary part for each feature of the second plurality of features; and

provide each feature of a third plurality of features of the third time series of data points associated with the contemporary transaction metric as an input to the FFT layer to generate a real part and an imaginary part for each feature of the third plurality of features.

~~26~~17. The computer program product of claim ~~25~~16, wherein, the one or more instructions that cause the at least one processor to train the machine learning model, cause the at least one processor to:

combine the real part for each feature of the first plurality of features to generate a combined real part of the first plurality of features;

combine the imaginary part for each feature of the first plurality of features to generate a combined imaginary part of the first plurality of features;

combine the real part for each feature of the second plurality of features to generate a combined real part of the second plurality of features;

combine the imaginary part for each feature of the second plurality of features to generate a combined imaginary part of the second plurality of features;

combine the real part for each feature of the third plurality of features to generate a combined real part of the third plurality of features; and

combine the imaginary part for each feature of the third plurality of features to generate a combined imaginary part of the third plurality of features; and

wherein the output of the processing layer of the machine learning model comprises the combined real part of the first plurality of features, the combined

imaginary part of the first plurality of features, the combined real part of the second plurality of features, the combined imaginary part of the second plurality of features, the combined real part of the third plurality of features, and the combined imaginary part of the third plurality of features.

~~27~~18. The computer program product of claim ~~26~~17, wherein, the one or more instructions that cause the at least one processor to train the machine learning model, cause the at least one processor to:

provide the combined real part of the first plurality of features as an input to a one-dimensional feature extraction convolutional layer;

provide the combined imaginary part of the first plurality of features as an input to the one-dimensional feature extraction convolutional layer;

provide the combined real part of the second plurality of features as an input to the one-dimensional feature extraction convolutional layer;

provide the combined imaginary part of the second plurality of features as an input to the one-dimensional feature extraction convolutional layer;

provide the combined real part of the third plurality of features as an input to the one-dimensional feature extraction convolutional layer;

provide the combined imaginary part of the third plurality of features as an input to the one-dimensional feature extraction convolutional layer; and

generate the output of the feature extraction component of the machine learning model based on an output of the one-dimensional feature extraction convolutional layer.

~~28~~19. The computer program product of claim ~~27~~18, wherein the input to the dual-attention component of the machine learning model comprises a real input portion and an imaginary input portion,

wherein the real input portion comprises:

a first real input based on an output of the one-dimensional feature extraction convolutional layer for the combined real part of the first plurality of features;

a second real input based on an output of the one-dimensional feature extraction convolutional layer for the combined real part of the second plurality of features; and

a third real input based on an output of the one-dimensional feature extraction convolutional layer for the combined real part of the third plurality of features; and

wherein the imaginary input portion comprises:

a first imaginary input based on an output of the one-dimensional feature extraction convolutional layer for the combined imaginary part of the first plurality of features;

a second imaginary input based on an output of the one-dimensional feature extraction convolutional layer for the combined imaginary part of the second plurality of features; and

a third imaginary input based on an output of the one-dimensional feature extraction convolutional layer for the combined imaginary part of the third plurality of features; and

wherein, the one or more instructions that cause the at least one processor to train the machine learning model, cause the at least one processor to:

generate a real component attention matrix based on the real input portion, wherein the real component attention matrix has a number of attention vectors that is equal to a number of features of the plurality of transaction features; and

generate an imaginary component attention matrix based on the imaginary input portion, wherein the imaginary component attention matrix has a number of attention vectors that is equal to a number of features of the plurality of transaction features.

~~29~~20. The computer program product of claim ~~28~~19, wherein the one or more instructions further cause the at least one processor to:

generate the real input portion of the input to the dual-attention component of the machine learning model; and

generate the imaginary input portion of the input to the dual-attention component of the machine learning model;

wherein, the one or more instructions that cause the at least one processor to generate the real input portion of the input to the dual-attention component of the machine learning model, cause the at least one processor to:

provide the output of the one-dimensional feature extraction convolutional layer for the combined real part of the first plurality of features as an input to a dropout layer to generate a real contemporary transaction metric feature vector as the first real input;

provide the output of the one-dimensional feature extraction convolutional layer for the combined real part of the second plurality of features as an input to the dropout layer to generate a real historical transaction metric feature vector as the second real input; and

provide the output of the one-dimensional feature extraction convolutional layer for the combined real part of the third plurality of features as an input to the dropout layer to generate a real historical target transaction metric feature vector as the third real input; and

wherein, the one or more instructions that cause the at least one processor to generate the imaginary input portion of the input to the dual-attention component of the machine learning model, cause the at least one processor to:

provide the output of the one-dimensional feature extraction convolutional layer for the combined imaginary part of the first plurality of features as an input to the dropout layer to generate an imaginary contemporary transaction metric feature vector as the first imaginary input;

provide the output of the one-dimensional feature extraction convolutional layer for the combined imaginary part of the second plurality of features as an input to the dropout layer to generate an imaginary historical transaction metric feature vector as the second imaginary input; and

provide the output of the one-dimensional feature extraction convolutional layer for the combined imaginary part of the third plurality of features as an input to the dropout layer to generate an imaginary historical target transaction metric feature vector as the third imaginary input.

~~30. The computer program product of claim 29, wherein, the one or more instructions that cause the at least one processor to generate the real component attention matrix, cause the at least one processor to:~~

~~generate a dot product of the real contemporary transaction metric feature vector and the real historical transaction metric feature vector;~~

~~provide the dot product of the real contemporary transaction metric feature vector and the real historical transaction metric feature vector as an input to a softmax layer to generate a real attention feature vector; and~~

~~generate a dot product of the real attention feature vector and the real historical target transaction metric feature vector to generate the real component attention matrix;~~

~~wherein, the one or more instructions that cause the at least one processor to generate the imaginary component attention matrix, cause the at least one processor to:~~

~~generate a dot product of the imaginary contemporary transaction metric feature vector and the imaginary historical transaction metric feature vector;~~

~~provide the dot product of the imaginary contemporary transaction metric feature vector and the imaginary historical transaction metric feature vector as a input to a softmax layer to generate an imaginary attention feature vector; and~~

~~generate a dot product of the imaginary attention feature vector and the imaginary historical target transaction metric feature vector to generate the imaginary component attention matrix.~~

~~31. The computer program product of claim 29, wherein, the one or more instructions that cause the at least one processor to generate the real component attention matrix, cause the at least one processor to:~~

~~perform an addition function on the real contemporary transaction metric feature vector and the real historical transaction metric feature vector;~~

~~provide a result of the addition function of the real contemporary transaction metric feature vector and the real historical transaction metric feature vector as an input to a tanh activation function;~~

~~provide an output of the tanh activation function as an input to a softmax layer to generate a real attention feature vector; and~~

~~generate a dot product of the real attention feature vector and the real historical target transaction metric feature vector to generate the real component attention matrix;~~

~~wherein the one or more instructions that cause the at least one processor to generate the imaginary component attention matrix, cause the at least one processor to:~~

perform an addition function on the imaginary contemporary transaction metric feature vector and the imaginary historical transaction metric feature vector;

provide the result of the addition function of the imaginary contemporary transaction metric feature vector and the imaginary historical transaction metric feature vector as an input to a tanh activation function;

provide an output of the tanh activation function as a input to a softmax layer to generate an imaginary attention feature vector; and

generate a dot product of the imaginary attention feature vector and the imaginary historical target transaction metric feature vector to generate the imaginary component attention matrix.

32. The computer program product of claim 29, wherein, the one or more instructions that cause the at least one processor to provide the output of the dual-attention component of the machine learning model as the input to the learning and prediction component of the machine learning model, cause the at least one processor to:

concatenate the real component attention matrix and the real contemporary transaction metric feature vector to generate a first concatenation;

provide the first concatenation as an input to a feature learning one-dimensional convolutional layer;

concatenate the imaginary component attention matrix and the imaginary contemporary transaction metric feature vector to generate a second concatenation;

provide the second concatenation as an input to the feature learning one-dimensional convolutional layer;

provide an input to a reshape layer based on an output of the feature learning one-dimensional convolutional layer for the first concatenation to generate a real component feature vector; and

provide an input to the reshape layer based on an output of the feature learning one-dimensional convolutional layer for the second concatenation to generate an imaginary component feature vector.

~~33. The computer program product of claim 32, wherein, the one or more instructions that cause the at least one processor to generate the output of the machine learning model, the at least one processor is programmed or configured to:~~

~~generate a complex number feature vector based on the real component feature vector and the imaginary component feature vector; and~~

~~provide the complex number feature vector as an input to an inverse FFT layer to generate the contemporary target transaction metric.~~