

CLAIMS

1. A computer system comprising:
 - a processor operatively coupled to memory;
 - an artificial intelligence (AI) platform, in communication with the processor, having machine learning (ML) tools to employ deep learning techniques to fused data across modalities, the tools comprising:
 - a first data feed manager operatively coupled to a first data set, the first data set having a first modality in a first data format;
 - a second data feed manager operatively coupled to a second data set, the second data set having a second modality in a second data form, the second modality being different from the first modality;
 - the first data feed manager to encode the first data set into a first set of vectors;
 - the second data feed manager to encode the second data set into a second set of vectors;
 - an analyzer operatively coupled to the first and second data feed managers, the analyzer to leverage an artificial recurrent neural network (RNN) to analyze the encoded first and second data sets, including iteratively and asynchronously fuse encoded features from the first and second data modalities, the fusing including combining vectors from the first and second data sets representing correlated temporal behavior; and
 - the fused vectors returned as output data.
2. The system of claim 1, wherein a first input data from the first data set represents a time-stamped textual data feed and a second input data from the second data set represents time series data.
3. The system of claim 2, wherein the iterative and asynchronous fusing includes the RNN to correlate temporal behavior of the time series data of the second data set with representative and encoded vectors from the first data set.
4. The system of claim 2, wherein the iterative and asynchronous fusing includes the RNN to filter out one or more representative vectors from the first data set irrelevant to patterns ascertained in the encoded second data set.
5. The system of any of claims 2 to 4, wherein encoding a text based modality includes the first data feed manager to learn semantic dependencies among words and aggregate the text into a representative vector for each input text document.
6. The system of any of the preceding claims, wherein analysis of the encoded first and second data sets further comprises the RNN to explore and inter-relate information from at least two temporal sequences of different sampling frequencies.

7. A method for employing deep learning techniques to fused data across modalities, the method comprising:
receiving, by a computing device, a multi-modal data set, the multi-modal data set including data in different formats from two or more modalities, including a first data set having a first modality and a second data set having a second modality;
separately processing the first and second data sets, including encoding the first data set into one or more first vectors and encoding the second data set into one or more second vectors;
analyzing the processed multi-modal data set, including iteratively and asynchronously fusing encoded features from the first and second data modalities, the fusing including combining vectors from the first and second data sets representing correlated temporal behavior; and
returning the fused vectors as output data.
8. The method of claim 7, wherein a first input data from the first data set represents a time-stamped textual data feed and a second input data from the second data set represents time series data.
9. The method of claim 8, wherein the iterative and asynchronous fusing includes correlating temporal behavior of the time series data of the second data set with representative and encoded vectors from the first data set.
10. The method of claim 8, wherein the iterative and asynchronous fusing includes filtering out one or more representative vectors from the first data set irrelevant to patterns ascertained in the encoded second data set.
11. The method of any of claims 8 to 10, wherein encoding a text based modality includes learning semantic dependencies among words and aggregating the text into a representative vector for each input text document.
12. The method of any of claims 8 to 11, wherein analyzing the processed multi-modal data set further comprises exploring and interrelating information from at least two temporal sequences of different sampling frequencies.
13. A method of any of claims 9 to 12, wherein analyzing the processed multi-modal data set, including iteratively and asynchronously fusing encoded features from the first and second data modalities, the fusing including combining vectors from the first and second data sets representing correlated temporal behavior comprises analyzing the processed multi-modal data, including fusing encoded vectors from the first and second data modalities, the fusing including combining vectors from the first and second data sets representing correlated temporal behavior between performance behavior of data in the modalities included in the multi-modal data; and
returning the fused vectors as output data comprises returning the fused vectors encoding common behaviors.

14. The method of claim 13, wherein the first modality is textual data and the second modality is time series data, and wherein the fusing further comprises obtaining a probability mass of attention on the textual data for a current state of the time series modality.
15. The method of either of claims 13 or 14, wherein analyzing the processed multi-modal data set includes employing deep learning techniques to fuse data across the modalities.
16. The method of either of claims 14 or 15, wherein a first input data from the first data set represents a time-stamped textual data feed and a second input data from the second data set represents numerical time series data, and wherein fusing the encoded vectors further comprises referencing the numerical time series data against the time-stamped textual data.
17. The method of any of claim 13 to 16, wherein fusing the encoded vectors occurs unsupervised.
18. The method of any of claims 13 to 17, wherein the multi-modal data includes one or more of medical data, climate data, computer vision data, financial data, or a combination thereof.
19. A computer program product for employing deep learning techniques to fused data across modalities, the computer program product comprising:
a computer readable storage medium readable by a processing circuit and storing instructions for execution by the processing circuit for performing a method according to any of claims 7 to 18.
20. A computer program stored on a computer readable medium and loadable into the internal memory of a digital computer, comprising software code portions, when said program is run on a computer, for performing the method of any of claims 7 to 18.