interact with another component. For example, in a serial architecture where an output from one model flows directly into an input of another, addressing when or if a given component should interact with the next component does not create a technical challenge. As the architecture evolves into a non-serial web of complex and/or structured data layers, determining when or if a given component should interact with another component creates a fundamental technical challenge to the functioning of the modular architecture. To overcome this technical problem, the modular architecture uses a gating network. For example, the gating network may determine when to use one layer and/or component (or an input/output thereof) versus another layer and/or component (or an input/output thereof).

[0009] Having solved the technical problems of allowing a modular component to interact as well as determine when to interact, the system may introduce multiple different components and/or layers that are each specifically trained to perform specific functions, trained on specific data and/or use specific algorithms. For example, the modular architecture may use a chatbot application component to receive requests for complex data modeling systems, a normalization layer to interpret the request, a gating network to select one or more data modeling components to analyze data to service the requests, and/or the same (or different chatbot application) to provide a response based on the analyzed data.

[0009a] According to an aspect, the present disclosure refers to a system for generating dynamic human-like conversational responses using layered data models with gated neural networks, the system comprising:

one or more processors; and

a non-transitory computer-readable media comprising of instructions that, when executed by the one or more processors, cause operations comprising:

receiving, at an Application Programming Interface endpoint layer, a dynamic human-like conversational request for a database query;

receiving a first portion of non-normalized source layer data for a first data model, wherein the first data model comprises an aggregated subset of additional data models;

receiving a second portion of the non-normalized source layer data for a second data model, wherein the second data model is trained on a first set of training data;

receiving a third portion of the non-normalized source layer data for a third data model, wherein the third data model is trained on a first set of training data, wherein the third data model is trained on a second set of training data;

determining, by processing each respective portion through a first normalization layer, a first feature input for the first data model based on the first portion, a second feature input for the second data model based on the second portion, and a third feature input for the third data model based on the third portion;

inputting the first feature input into the first data model, the second feature input into the second data model, and the third feature input into the third data model to generate a respective outputs;

retrieving a first configuration file for a gating network from a first configuration layer, wherein the first configuration file defines one or more parameters for the gating network;

inputting the respective outputs into the gating network to generate a gating network output based on the one or more parameters;

determining, based on the gating network output, a database query; and

generating a dynamic human-like conversational response based on the database query.

[0009b] According to an aspect, the present disclosure refers to a method for generating dynamic human-like conversational responses using layered data models with gated neural networks, the method comprising:

receiving a first portion of non-normalized source layer data for a first data model, wherein the first data model comprises an aggregated subset of additional data models;

receiving a second portion of the non-normalized source layer data for a second data model;

receiving a third portion of the non-normalized source layer data for a third data model;

determining, by processing each respective portion through a first normalization layer, a first feature input for the first data model based on the first portion, a second feature input for the second data model based on the second portion, and a third feature input for the third data model based on the third portion;

inputting the first feature input into the first data model, the second feature input into the second data model, and the third feature input into the third data model to generate a respective outputs;

retrieving a first configuration file for a gating network from a first configuration layer, wherein the first configuration file defines one or more parameters for the gating network;

inputting the respective outputs into the gating network to generate a gating network output based on the one or more parameters;

determining, based on the gating network output, a database query; and

generating a dynamic human-like conversational response based on the database query.

[0009c] According to an aspect, the present disclosure refers to a non-transitory, computer-readable medium comprising of instructions that, when executed by one or more processors, cause operations comprising:

receiving a first portion of non-normalized source layer data for a first data model, wherein the first data model comprises an aggregated subset of additional data models;

receiving a second portion of the non-normalized source layer data for a second data model;

determining, by processing each respective portion through a first normalization layer, a first feature input for the first data model based on the first portion and a second feature input for the second data model based on the second portion;

inputting the first feature input into the first data model and the second feature input into the second data model to generate a respective outputs;

retrieving a first configuration file for a gating network from a first configuration layer, wherein the first configuration file defines one or more parameters for the gating network;

inputting the respective outputs into the gating network to generate a gating network output based on the one or more parameters;

determining, based on the gating network output, a database query; and

generating a dynamic human-like conversational response based on the database query.

[0010] In some examples, systems and methods for generating dynamic human-like conversational responses using a modular architecture featuring layered data models with gated neural networks are described. For example, the system may receive a first portion of non-normalized source layer data for a first data model, wherein the first data model comprises an aggregated subset of additional data models. The system may receive a second portion of the non-

normalized source layer data for a second data model. The system may receive a third portion of the non-normalized source layer data for a third data model. The system may determine, by processing each respective portion through a first normalization layer, a first feature input for the first data model based on the first portion, a second feature input for the second data model based on the second portion, and a third feature input for the third data model based on the third portion. The system may input the first feature input into the first data model, the second feature input into the second data model, and the third feature input into the third data model to generate respective outputs. The system may retrieve a first configuration file for a gating network from a first configuration layer, wherein the first configuration file defines one or more parameters for the gating network. The system may