June 2022

METHOD FOR GENERATING COMPLEX-VALUED SCORE FOR DEEP LEARNING MODELS

Pei Yang
Visa

Peng Wu
Visa

Dan Wang
Visa

METHOD FOR GENERATING COMPLEX-VALUED SCORE FOR DEEP LEARNING MODELS

TECHNICAL FIELD

[001] The present disclosure relates generally for generating complex-valued score for deep learning models, and more particularly to construct deep neural network models that can generate complex-valued scores in FinTech.

BACKGROUND

[002] In the existing approaches, when modelling a binary classification problem with deep learning model, the model's output is a real number between 0 and 1. For example, deciding whether to approve or reject a request, predicting fraud or not-fraud, and so on use deep neural networks for predictions and/ or decisions, while the predicted output is a real number.

[003] For some challenging binary classification problems, because the interval [0, 1] is one dimensional (one degree of freedom), the score distributions of the two classes will overlap significantly, making it difficult to find a suitable cut-off for making accurate predictions or decisions.

[004] One of the existing technologies discloses generating the output score of deep neural network model. However, the generated output score is a real number in the interval [0.0, 1.0]. For binary classification, a cut-off value will be set, if the score is less than the cut-off, then it’s classified as class 0, otherwise it’s classified as class 1. Hence, the score is a real number and has only one degree of freedom.

[005] As a result, increasing the degree of freedom for the score is advantageous.

[006] The information disclosed in this background of the disclosure section is only for enhancement of understanding of the general background of the invention and should not be taken as an acknowledgement or any form of suggestion that this information forms the prior art already known to a person skilled in the art.

SUMMARY

[007] One of the common uses for deep learning is performing binary classification, which looks at an input and predicts which of two possible classes it belongs to. Practical uses include sentiment analysis, spam detection, and credit-card fraud detection. However, since the score value is a real number between 0 to1 and has one degree of freedom, the score distributions of the two classes will overlap significantly, making it difficult to find a suitable cut-off for making accurate predictions or decisions.

[008] The present invention aims at generating complex-valued score for deep learning models and increasing the degree of freedom for the score, while this approach finds a suitable cut-off for making accurate predictions or decisions. In one embodiment, a computer-implemented method of generating complex-valued score for deep learning models is proposed. The method comprises, mapping input features into real number and non-negative number using two neural network models. Further, considering real number and non-negative number to be real and imaginary parts of a point on the upper half-plane of the complex plane. Furthermore, mapping the point onto the unit disk on the complex plane using conformal mapping. Thereafter, considering complex-valued number (103) in the unit disk as the score of the deep learning model.

[009] The foregoing summary is illustrative only and is not intended to be in any way limiting. In addition to the illustrative aspects, embodiments, and features described above, further aspects, embodiments, and features will become apparent by reference to the drawings and the following detailed description.

BRIEF DESCRIPTION OF THE DRAWINGS

[0010] The example embodiment(s) of the present invention are illustrated by way of example, and not in way by limitation, in the figures of the accompanying drawings and in which like reference numerals refer to similar elements and in which:

[0010] Figures 1 depicts environmental diagram representing the generation of complex-valued score for deep learning models, in accordance with an embodiment of the present disclosure;

[0011] Figure 2 illustrates an exemplary flow of method of generating a complex-valued score for deep learning models, in accordance with an embodiment of the present disclosure;

[0012] Figure 3 depicts an exemplary conformal mapping to map the point onto the unit disk on the complex plane, in accordance with an embodiment of the present disclosure;

[0013] Figure 4 illustrates an exemplary flow diagram for generating a complex-valued score for deep learning models, in accordance with an embodiment of the present disclosure;

[0014] Figure 5 illustrates an exemplary shows a block diagram of a computer apparatus in accordance with an embodiment;

[0015] It should be appreciated by those skilled in the art that any block diagrams herein represent conceptual views of illustrative systems embodying the principles of the present subject matter. Similarly, it will be appreciated that any flow charts, flow diagrams, state transition diagrams, pseudo code, and the like represent various processes which may be substantially represented in computer readable medium and executed by a computer or processor, whether or not such computer or processor is explicitly shown. While each of the figures illustrates a particular embodiment for purposes of illustrating a clear example, other embodiments may omit, add to, reorder, and/or modify any of the elements shown in the figures.

DETAILED DESCRIPTION

[0016] In the present document, the word "exemplary" is used herein to mean "serving as an example, instance, or illustration." Any embodiment or implementation of the present subject matter described herein as "exemplary" is not necessarily to be construed as preferred or advantageous over other embodiments.

[0017] While the disclosure is susceptible to various modifications and alternative forms, specific embodiment thereof has been shown by way of example in the drawings and will be described in detail below. It should be understood, however that it is not intended to limit the disclosure to the particular forms disclosed, but on the contrary, the disclosure is to cover all modifications, equivalents, and alternative falling within the scope of the disclosure.

[0018] The terms “comprises”, “comprising”, or any other variations thereof, are intended to cover a non-exclusive inclusion, such that a setup, device or method that comprises a list of components or steps does not include only those components or steps but may include other components or steps not expressly listed or inherent to such setup or device or method. In other words, one or more elements in a system or apparatus proceeded by “comprises… a” does not, without more constraints, preclude the existence of other elements or additional elements in the system or apparatus.

[0019] Deep learning techniques aids in the detection of scams in a variety of fields, including FinTech, Healthcare, and eCommerce. The amount of transactions made through online purchases utilizing credit and debit cards has increased in this digital era of information. Debit and credit card fraud detection is often thought of as a binary classification problem, with each transaction being classed as either valid or fraudulent. As a result, fraud detection has become a global concern. As a result, the FinTech industry has employed a variety of fraud detection approaches. Binary classification is a type of classification in deep learning that looks at an input and predicts which of two classes it belongs to. The score distributions of the two groups will overlap significantly, making it difficult to establish a reasonable cut-off for making accurate predictions or choices, because the score value is a real number between 0 and 1 and has one degree of freedom.

[0020] Embodiments of the present disclosure relate to method for generating complex-valued score for deep learning models. The complex-valued score increases the degree of freedom from one to two, which helps to avoid score overlapping in challenging binary classification tasks which are very typical in FinTech.

[0021] Figure 1 depicts environmental diagram representing the generation of complex-valued score for deep learning models, in accordance with an embodiment of the present disclosure. The environment (100) comprises a computer system (101). The computer system (101) comprises a processor and a memory (not shown). The processor is configured to obtain input features X. Additionally, the computer system (101) comprises two neural network models (102). The input features obtained by the processor are then passed on to the neural network models. The input features are then mapped into real number and non-negative number using two neural network models. In the neural network model 1 the linear activation function is used to define the output of a node given an input or set of input, while in the neural network model 2 the non-negative activation function is used. The activation functions are trained for different application to generate complex-valued score. Further, real number and non-negative number obtained as output are considered to be real and imaginary parts of a point on the upper half-plane of the complex plane. Furthermore, mapping the point onto the unit disk on the complex plane using conformal mapping. Thereafter, considering complex-valued number (103) in the unit disk as the score of the deep learning model.

[0022] Reference is now made to Figure 2. The following method describes the steps performed by the computer system (101). Figure 2 illustrates an exemplary flow of method of generating a complex-valued score for deep learning models, in accordance with an embodiment of the present disclosure.

[0023] As illustrated in Figure 2, the method (200) may comprise one or more steps. The method (200) may be described in the general context of computer executable instructions. Generally, computer executable instructions can include routines, programs, objects, components, data structures, procedures, modules, and functions, which perform particular functions or implement particular abstract data types.

[0024] The order in which the method (200) is described is not intended to be construed as a limitation, and any number of the described method blocks can be combined in any order to implement the method. Additionally, individual blocks may be deleted from the methods without departing from the scope of the subject matter described herein. Furthermore, the method can be implemented in any suitable hardware, software, firmware, or combination thereof.

[0025] The following steps are performed by the computer system (101) after obtaining the input features. The following steps are performed by the processor present in the computer system (101) to generate a complex-valued score within the unit disk on the complex plane.

[0026] At step 201, mapping input features into real number and non-negative number using two neural network models. The real numbers herein may be the values that fall in class 0 of real value domain. The non-negative numbers herein may be the values that fall in class 1 of real value domain. The input layer of a neural network is composed of artificial input neurons and brings the initial data into the system for further processing by subsequent layers of artificial neurons. The input features are mapped into real number and non-negative number depending upon activation function. Activation function is used to determine the output of node. The activation functions can be basically of two types: Linear activation function and non-negative activation function. Linear activation functions provide linear output, output value is same as the input value. Non-negative activation function makes the model to easily generalize or adapt with variety of data and to differentiate between the output. Examples of non-negative activation functions maybe hyperbolic tangent activation function, sigmoid activation function, ReLU (Rectified Linear Unit) activation function. At step 202, considering real number and non-negative number to be real and imaginary parts of a point on the upper half-plane of the complex plane. For example, if the real number is -2 and the non-negative number is 3, then the corresponding point on the upper half-plane of the complex plan is -2 + 3i (a complex number). Upon mapping the input features into real number and non-negative number using linear and non-negative activation functions, the real number and non-negative number are considered as the real and imaginary parts in the complex domain.

[0027] At step 203, mapping the point onto the unit disk on the complex plane using conformal mapping. Upon considering real number and non-negative number to be real and imaginary parts of a point, the real and imaginary parts of a point are mapped onto the unit disk of the complex plane. A conformal mapping, also called an angle-preserving transformation, is a transformation w=f(z) that preserves local angles. The complex plane comprises real parts on x-axis and imaginary parts on y-axis. The real and imaginary points form a complex number which is mapped onto the unit disk of the complex plane.

[0028] At step 204, considering complex-valued number in the unit disk as the score of the deep learning model. Upon mapping the point onto the unit disk on the complex plane using conformal mapping, the complex valued score is considered for deep learning model. The complex valued number within the unit disk of complex plane is considered as the complex-value score of the deep learning model. Thus, the score is increased from one degree of freedom to two degrees of freedom i.e., one is the complex norm, and the other is the complex phase. For example, Z = Cr + iCi, wherein z is the complex number, Cr is a real part and Ci is an imaginary part. Complex norm = √ Cr 2 + Ci2

Complex phase = $\tan^{-1}$(Ci/ Cr)

Consider an example, in figure 1, neural network 1 will provide output as a real number (Cr) and neural network 2 will provide output as a non-negative number (Ci). Let z = Cr + Ci. Therefore, f(z) may be the complex-valued score.

For example, if Cr = -2, Ci = 2, a = 2i

Then we have, z = -2 + 2i

So, f(z) = (z - 2i)/ (z + 2i)

= (-2 + 2i -2i)/ (-2 + 2i + 2i)

= (-2)/ (-2 + 4i)

= (-1)/ (-1 + 2i)

= (1 + 2i)/ 5

Therefore, the complex-valued score in this example is (1 + 2i)/ 5

[0029] Figure 3 depicts an exemplary conformal mapping to map the point onto the unit disk on the complex plane, in accordance with an embodiment of the present disclosure. Conformal mapping is basically used to transform an image from z-plane to w-plane via transformation function or transfer function.

$$Z = C_r + iC_i,\ f(z) = w = e^{r_w + i\theta_w} = e^{r_w} e^{i\theta_w},$$

$$f(z) = e^{i\theta} \frac{z - a}{z - \bar{a}},\ \theta \text{ is real and a is in upper half} - \text{plane.}$$

For example, a line in z-plane can be transformed to circle in w-plane. As shown in Figure 3, f(z) is a conformal mapping that maps all numbers in upper half-plane of z plane to unit disk of complex plane. The line in z-plane has real valued scores of class 0 and class 1, while the scores are overlapping as the scores are of real numbers between 0 and 1. Conformal mapping is applied to map the scores on to the complex domain. The complex valued number within the unit disk of complex plane is considered as the complex-value score of the deep learning model.

By this research work approach, overlapping is avoided. The model is trained with loss function L. The model is trained to minimize the value of L. On basis of the loss value can be updated with training data. Lesser the value of loss value the better will be the result. For example, if we let scores for class 0 be closer to origin and scores for class 1 be closer to the unit circle, then a loss function can be:

$L = (1 - y)e^{r_w} + \mathrm{y}(1 - e^{r_w})$ (y=0 for class 0 and y=1 for class 1)

[0030] Figure 4 illustrates an exemplary flow diagram for generating a complex-valued score for deep learning models, in accordance with an embodiment of the present disclosure. The processor's input features are subsequently sent along to the neural network models. Two neural network models (102) are used to map the input features into real and non-negative numbers. The linear activation function is used in neural network model 1 to describe the output of a node given an input or set of inputs, whereas the non-negative activation function is used in neural network model 2. Further, the real and non-negative numbers obtained as output are viewed as the real and imaginary parts of a point on the complex plane's upper half-plane, respectively. Furthermore, in the complex plane, mapping the point onto the unit disc using conformal mapping. Thereafter, considering complex-valued number in the unit disk as the score of the deep learning model.

[0031] Figure 5 illustrates a block diagram of an exemplary computer system (500) for implementing embodiments consistent with the present disclosure. In an embodiment, the computer system (500) is used to implement the method for authorising transactions in the platform (500). The computer system (500) may comprise a central processing unit (“CPU” or “processor”) (502). The processor (502) may comprise at least one data processor for executing program components for dynamic resource allocation at run time. The processor (502) may include specialized processing units such as integrated system (bus) controllers, memory management control units, floating point units, graphics processing units, digital signal processing units, etc.

[0032] The processor (502) may be disposed in communication with one or more input/output (I/O) devices (not shown) via I/O interface (501). The I/O interface (501) may employ communication protocols/methods such as, without limitation, audio, analog, digital, monoaural, RCA, stereo, IEEE-1394, serial bus, universal serial bus (USB), infrared, PS/2, BNC, coaxial, component, composite, digital visual interface (DVI), high-definition multimedia interface (HDMI), RF antennas, S-Video, VGA, IEEE 802.n /b/g/n/x, Bluetooth, cellular (e.g., code-division multiple access (CDMA), high-speed packet access (HSPA+), global system for mobile communications (GSM), long-term evolution (LTE), WiMax, or the like), etc.

[0033] Using the I/O interface (501), the computer system (500) may communicate with one or more I/O devices. For example, the input device (510) may be an antenna, keyboard, mouse, joystick, (infrared) remote control, camera, card reader, fax machine, dongle, biometric reader, microphone, touch screen, touchpad, trackball, stylus, scanner, storage device, transceiver, video device/source, etc. The output device 511 may be a printer, fax machine, video display (e.g., cathode ray tube (CRT), liquid crystal display (LCD), light-emitting diode (LED), plasma, Plasma display panel (PDP), Organic light-emitting diode display (OLED) or the like), audio speaker, etc.

[0034] In some embodiments, the computer system (500) is connected to the service operator through a communication network (509). The processor (502) may be disposed in communication with the communication network (509) via a network interface (503). The network interface (503) may communicate with the communication network (509). The network interface (503) may employ connection protocols including, without limitation, direct connect, Ethernet (e.g., twisted pair 10/100/1000 Base T), transmission control protocol/Internet protocol (TCP/IP), token ring, IEEE 802.11a/b/g/n/x, etc. The communication network (509) may include, without limitation, a direct interconnection, e-commerce network, a peer to peer (P2P) network, local area network (LAN), wide area network (WAN), wireless network (e.g., using Wireless Application Protocol), the Internet, Wi-Fi, etc. Using the network interface (503) and the communication network (509), the computer system 400 may communicate with the one or more service operators.

[0035] In some embodiments, the processor (502) may be disposed in communication with a memory (505) (e.g., RAM, ROM, etc) via a storage interface (504). The storage interface (504) may connect to memory (505) including, without limitation, memory drives, removable disc drives, etc., employing connection protocols such as serial advanced technology attachment (SATA), Integrated Drive Electronics (IDE), IEEE-1394, Universal Serial Bus (USB), fibre channel, Small Computer Systems Interface (SCSI), etc. The memory drives may further include a drum, magnetic disc drive, magneto-optical drive, optical drive, Redundant Array of Independent Discs (RAID), solid-state memory devices, solid-state drives, etc.

[0036] The memory (505) may store a collection of program or database components, including, without limitation, user interface (506), an operating system (507), web server (508) etc. In some embodiments, computer system (500) may store user/application data (506), such as the data, variables, records, etc. as described in this disclosure. Such databases may be implemented as fault-tolerant, relational, scalable, secure databases such as Oracle or Sybase.

[0037] The operating system (507) may facilitate resource management and operation of the computer system (500). Examples of operating systems include, without limitation, Apple Macintosh OS X, Unix, Unix-like system distributions (e.g., Berkeley Software Distribution (BSD), FreeBSD, NetBSD, OpenBSD, etc.), Linux distributions (e.g., Red Hat, Ubuntu, Kubuntu, etc.), IBM OS/2, Microsoft Windows (XP, Vista/7/8, 10 etc.), Apple iOS, Google Android, Blackberry OS, or the like.

[0038] In some embodiments, the computer system (500) may implement a web browser (508) stored program component. The web browser (508) may be a hypertext viewing application, such as Microsoft Internet Explorer, Google Chrome, Mozilla Firefox, Apple Safari, etc. Secure web browsing may be provided using Secure Hypertext Transport Protocol (HTTPS), Secure Sockets Layer (SSL), Transport Layer Security (TLS), etc. Web browsers (508) may utilize facilities such as AJAX, DHTML, Adobe Flash, JavaScript, Java, Application Programming Interfaces (APIs), etc. In some embodiments, the computer system (500) may implement a mail server stored program component. The mail server may be an Internet mail server such as Microsoft Exchange, or the like. The mail server may utilize facilities such as ASP, ActiveX, ANSI C++/C#, Microsoft .NET, CGI scripts, Java, JavaScript, PERL, PHP, Python, WebObjects, etc. The mail server may utilize communication protocols such as Internet Message Access Protocol (IMAP), Messaging Application Programming Interface (MAPI), Microsoft Exchange, Post Office Protocol (POP), Simple Mail Transfer Protocol (SMTP), or the like. In some embodiments, the computer system (500) may implement a mail client stored program component. The mail client may be a mail viewing application, such as Apple Mail, Microsoft Entourage, Microsoft Outlook, Mozilla Thunderbird, etc.

[0039] In an embodiment, the computer system (500) is a directory server providing services for facilitating transactions between a merchant associated with an acquirer system, and an issuer system. In an embodiment, the computer system (500) is connected to the entities comprising the merchant, acquirer system, issuer system.

[0040] The terms "an embodiment", "embodiment", "embodiments", "the embodiment", "the embodiments", "one or more embodiments", "some embodiments", and "one embodiment" mean "one or more (but not all) embodiments of the invention(s)" unless expressly specified otherwise.

[0041] The terms "including", "comprising", “having” and variations thereof mean "including but not limited to", unless expressly specified otherwise.

[0042] The enumerated listing of items does not imply that any or all of the items are mutually exclusive, unless expressly specified otherwise. The terms "a", "an" and "the" mean "one or more", unless expressly specified otherwise.

[0043] A description of an embodiment with several components in communication with each other does not imply that all such components are required. On the contrary a variety of optional components are described to illustrate the wide variety of possible embodiments of the invention.

[0044] When a single device or article is described herein, it will be readily apparent that more than one device/article (whether or not they cooperate) may be used in place of a single device/article. Similarly, where more than one device or article is described herein (whether or not they cooperate), it will be readily apparent that a single device/article may be used in place of the more than one device or article, or a different number of devices/articles may be used instead of the shown number of devices or programs. The functionality and/or the features of a device may be alternatively embodied by one or more other devices which are not explicitly described as having such functionality/features. Thus, other embodiments of the invention need not include the device itself.

[0045] Finally, the language used in the specification has been principally selected for readability and instructional purposes, and it may not have been selected to delineate or circumscribe the inventive subject matter. It is therefore intended that the scope of the invention be limited not by this detailed description, but rather by any claims that issue on an application based here on. Accordingly, the disclosure of the embodiments of the invention is intended to be illustrative, but not limiting, of the scope of the invention, which is set forth in the following claims.

[0046] While various aspects and embodiments have been disclosed herein, other aspects and embodiments will be apparent to those skilled in the art. The various aspects and embodiments disclosed herein are for purposes of illustration and are not intended to be limiting, with the true scope being indicated by the following claims.

ABSTRACT

METHOD FOR GENERATING COMPLEX-VALUED SCORE FOR DEEP LEARNING MODELS

The present disclosure discloses a method to construct deep neural network models (102) that can generate complex-valued scores in FinTech. The method employs conformal mapping, which maps points in the upper half-plane onto the unit disc while using the complex-valued number in the unit disc as the deep learning model's score. This gives the score two degrees of freedom, which helps to avoid the score overlapping problem that happens when using a real-valued score ([0, 1] interval). For training deep learning models, the approach also includes a loss function. On the basis of the Loss value, the deep learning model is updated to get the best results.

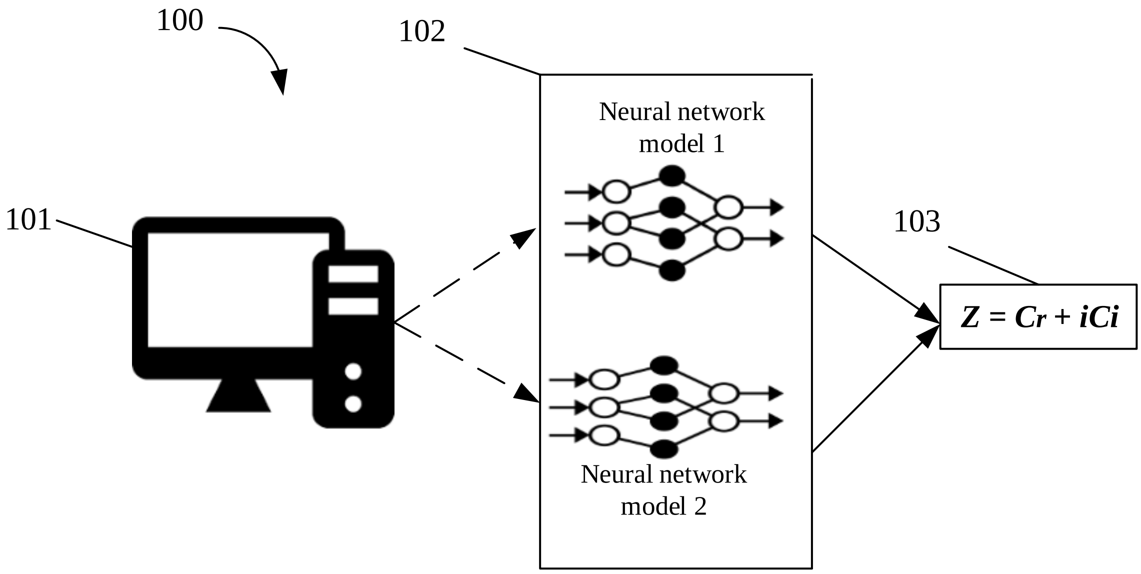

Figure 1

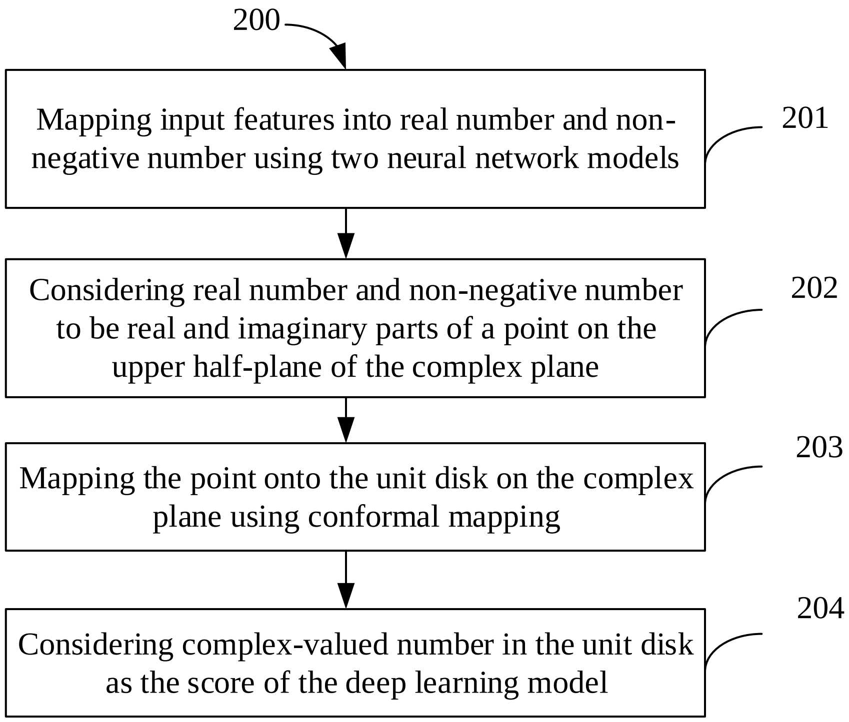

Figure 2

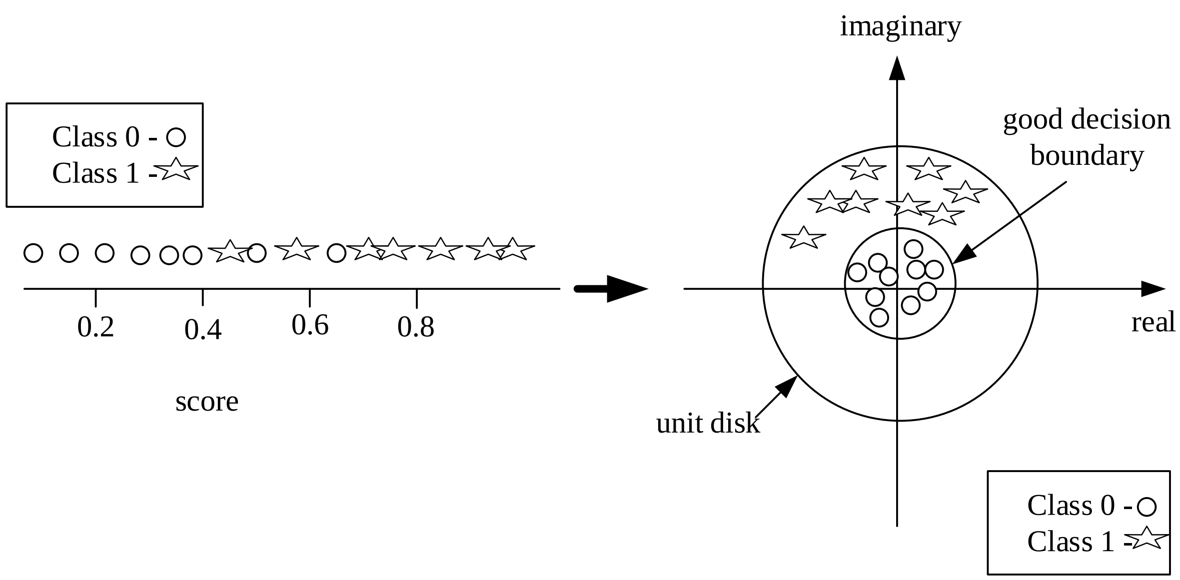

Figure 3

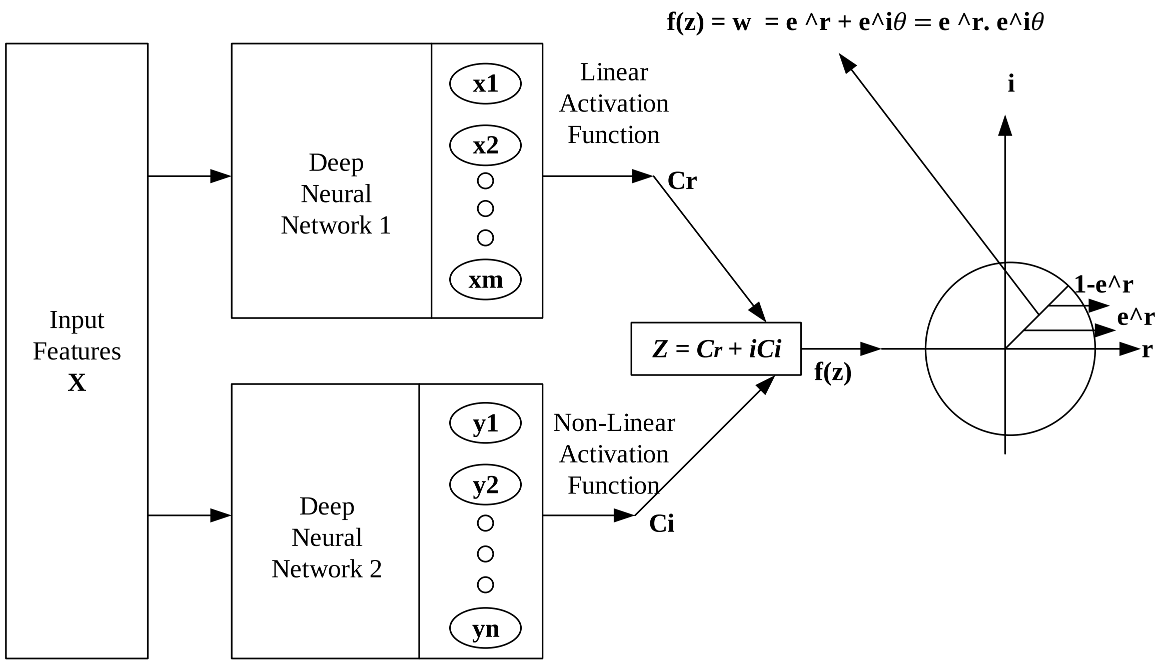

Figure 4

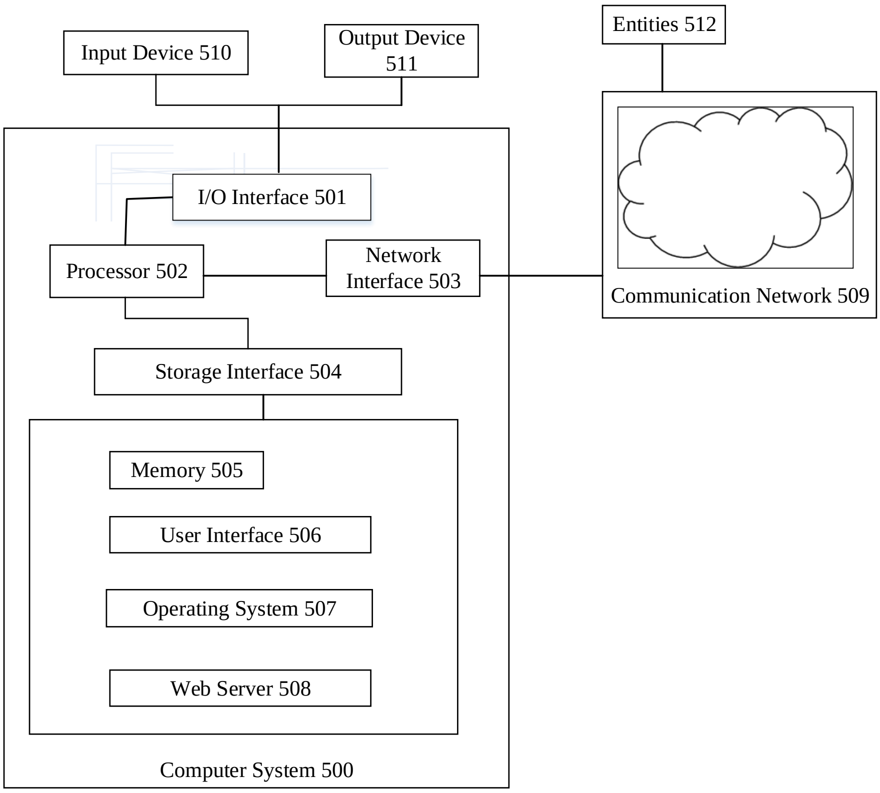
Input Device 510
Output Device 511
Entities 512
I/O Interface 501
Processor 502
Network Interface 503
Communication Network 509
Storage Interface 504
Memory 505
User Interface 506
Operating System 507
Web Server 508
Computer System 500

Figure 5