



(10) **DE 10 2018 210 003 A1** 2019.12.24

(12)

Offenlegungsschrift

(21) Aktenzeichen: **10 2018 210 003.9**

(22) Anmeldetag: **20.06.2018**

(43) Offenlegungstag: **24.12.2019**

(51) Int Cl.: **G06N 3/08** (2006.01)

(71) Anmelder:

Robert Bosch GmbH, 70469 Stuttgart, DE

(72) Erfinder:

Karatzoglou, Antonios, 70191 Stuttgart, DE

(56) Ermittelter Stand der Technik:

Goo, Wonjoon [et.al.]: Taxonomy-Regularized Semantic Deep Convolutional Neural Networks. Computer Vision -- ECCV 2016. Cham : Springer International Publishing, 2016. S. 86-101. - ISBN 978-3-319-46475-6. https://doi.org/10.1007/978-3-319-46475-6_6 [abgerufen am 19.02.2019]

He, X.; Peng, Y.: Fine-Grained Image Classification via Combining Vision and

Language. In: 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2017, S. 7332-7340. - ISSN 1063-6919. <https://doi.org/10.1109/CVPR.2017.775> [abgerufen am 19.02.2019]

Lei, J.; Guo, Z.; Wang, Y.: Weakly Supervised Image Classification with Coarse and Fine Labels. In: 2017 14th Conference on Computer and Robot Vision (CRV), 2017, S. 240-247. <https://doi.org/10.1109/CRV.2017.21> [abgerufen am 19.02.2019]

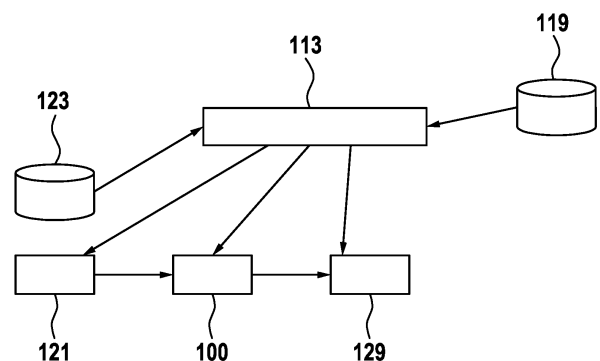
Tsai, Y.H.; Huang, L.; Salakhutdinov, R.: Learning Robust Visual-Semantic Embeddings. In: 2017 IEEE International Conference on Computer Vision (ICCV), 2017, S. 3591-3600. - ISSN 2380-7504. <https://doi.org/10.1109/ICCV.2017.386> [abgerufen am 19.02.2019]

Rechercheantrag gemäß § 43 PatG ist gestellt.

Die folgenden Angaben sind den vom Anmelder eingereichten Unterlagen entnommen.

(54) Bezeichnung: **Verfahren und Vorrichtung zum Anlernen eines Modells für maschinelles Lernen**

(57) Zusammenfassung: Die vorliegende Erfindung betrifft ein Verfahren zum Anlernen eines Modells für maschinelles Lernen, mit den Schritten eines Berechnens (S101) eines Lern-Istwertes durch das Modell für maschinelles Lernen auf Basis von Eingabedaten; eines Bestimmens (S102) eines numerischen Abweichungswertes, der eine semantische Abweichung zwischen dem berechneten Lern-Istwert und einem vorgegebenen Lern-Sollwert angibt; und eines Anpassen (S103) des Modells für maschinelles Lernen auf Basis des Abweichungswertes.



Beschreibung

[0001] Die vorliegende Erfindung betrifft ein Verfahren zum Anlernen eines Modells für maschinelles Lernen (Machine Learning Model), ein Computerprogramm und computerlesbares Speichermedium.

Stand der Technik

[0002] Aktuelle Verfahren für maschinelles Lernen (Machine Learning) sind im Vergleich zu wissensbasierten Systemen, die auf strukturierten Wissensmodellen aufbauen, datengetriebene Verfahren (Data-driven vs. Knowledge-driven Systems).

[0003] Datengetriebene Verfahren speisen das interne Modell mit Daten, um es für eine bestimmte Aufgabe zu trainieren. Dabei weiß das Modell nicht, d.h. der Machine Learner, um welche Daten es sich handelt und in welcher Beziehung die Daten zueinander stehen, mit denen es trainiert wird. Es fehlt die Eigenschaft einer semantischen Interpretation der Daten. Die Maschine versteht daher nicht genau, welcher Art die Daten sind.

[0004] Die Druckschrift DE 69 217 047 T2 betrifft adaptive Informationsverarbeitungssysteme und neurale Netzwerk-Systeme, die aus einem automatischen organisatorischen Apparat und Verfahren bestehen für die dynamische Aufzeichnung einer Eingabe in eine semantisch kongruente und gleichzeitig gültige korrekte Antwort.

[0005] Die Druckschrift DE 69 130 246 T2 betrifft ein Satzbildungsverfahren und -system zur Erzeugung von Sätzen in einer Sprache, wie beispielsweise Japanisch oder Englisch, aus einer semantischen Darstellung, die in das System eingegeben wurde. Dabei kann eine Anordnung angewendet werden, bei der ein neuronales Netz zur Durchführung des Lernens verwendet wird, indem semantische Darstellungen und ein Satz von diesen entsprechenden geeigneten Ausgabesätzen verwendet und eine Bewertung beruhend auf den Ergebnissen des Lernens durchgeführt wird.

[0006] Die Druckschrift DE 10 2016 005 407 A1 betrifft ein gemeinsames Tiefenschätzen und semantisches Bezeichnen eines einzelnen Bildes. Semantische Bezeichner und Tiefenwerte werden für jeden Pixel in einer jeweiligen lokalen Segmentierung ebenfalls unter Verwendung von maschinellem Lernen prädiiziert, und semantische Bezeichner werden gelernt, um Pixel mit dem Segment lokal mit einem Bezeichner zu versehen bzw. zu taggen.

[0007] Die Druckschrift EP 1 183 619 B1 betrifft ein semantisches Netz, in welchem ein Operieren innerhalb des semantischen Netzes abhängig von einer jeweiligen in dem semantischen Netz vorliegenden Si-

tuation durchgeführt wird. Das semantische Netz umfasst sowohl semantische Einheiten, die Beziehungsinhalte besitzen, als auch Verknüpfungseinheiten, die einen Beziehungsinhalt beschreiben, der jeweils zwei semantische Einheiten verknüpft.

Offenbarung der Erfindung**Vorteile der Erfindung**

[0008] Ein erster Aspekt betrifft ein Verfahren zum Anlernen eines Modells für maschinelles Lernen, mit den Schritten eines Berechnens eines Lern-Istwertes durch das Modell für maschinelles Lernen auf Basis von Eingabedaten; eines Bestimmens eines numerischen Abweichungswertes, der eine semantische Abweichung zwischen dem berechneten Lern-Istwert und einem vorgegebenen Lern-Sollwert angibt; und eines Anpassens des Modells für maschinelles Lernen auf Basis des Abweichungswertes. Dadurch kann das Anlernen eines Modells für maschinelles Lernen um ein Mehrfaches beschleunigt werden. Durch das Verfahren wird der technische Vorteil einer verbesserten Performanz in Hinsicht auf die Vorhersage-, Klassifikations-, Regressions- oder Erkennungsrate und Genauigkeit, einer geringen Trainingszeit des Modells erreicht.

[0009] In einer technisch vorteilhaften Ausführungsform des Verfahrens implementiert das Modell ein neuronales Netz. Dadurch wird beispielsweise der technische Vorteil erreicht, dass die Lern-Istwerte auf schnelle Art und Weise erzeugt werden können.

[0010] In einer weiteren technisch vorteilhaften Ausführungsform des Verfahrens wird eine Fehlerrückführung (Back Propagation) durch das neuronale Netz auf Basis des Abweichungswertes durchgeführt. Dadurch wird beispielsweise der technische Vorteil erreicht, dass eine effiziente Anpassung des Modells vorgenommen werden kann.

[0011] In einer weiteren technisch vorteilhaften Ausführungsform des Verfahrens wird eine Verlustfunktion des neuronalen Netzes durch den Abweichungswert angepasst. Dadurch wird beispielsweise der technische Vorteil erreicht, dass eine einfache und schnelle Anpassung des Modells vorgenommen werden kann.

[0012] In einer weiteren technisch vorteilhaften Ausführungsform des Verfahrens wird die Verlustfunktion mit dem Abweichungswert multipliziert. Dadurch wird beispielsweise der technische Vorteil erreicht, dass das Modell mit einem geringen Aufwand angepasst werden kann und sich das Anlernen des Modells beschleunigt.

[0013] In einer weiteren technisch vorteilhaften Ausführungsform des Verfahrens gibt der Abweichungs-

wert eine semantische Ähnlichkeit oder semantische Verwandtheit zwischen dem Lern-Istwert und dem Lern-Sollwert an. Dadurch wird beispielsweise der technische Vorteil erreicht, dass das Modell unterschiedliche semantische Konzepte umsetzen kann.

[0014] In einer weiteren technisch vorteilhaften Ausführungsform des Verfahrens wird eine semantische Wissensbasis zum Bestimmen des numerischen Abweichungswertes verwendet. Dadurch wird beispielsweise der technische Vorteil erreicht, dass der Abweichungswert zuverlässig bestimmt werden kann.

[0015] In einer weiteren technisch vorteilhaften Ausführungsform des Verfahrens ist das Modell für maschinelles Lernen ein probabilistisches Modell und Übergangswahrscheinlichkeiten werden auf Basis des Abweichungswertes angepasst. Dadurch wird beispielsweise der technische Vorteil erreicht, dass ein effizientes Modell zum Berechnen des Lern-Istwertes herangezogen wird.

[0016] Ein zweiter Aspekt betrifft ein Computerprogramm, umfassend Befehle, die bei der Ausführung des Programms durch einen Computer diesen veranlassen, das Verfahren nach dem ersten Aspekt auszuführen. Dadurch werden die gleichen technischen Vorteile wie durch das Verfahren nach dem ersten Aspekt erreicht.

[0017] Ein dritter Aspekt betrifft ein computerlesbares Speichermedium, umfassend Befehle, die bei der Ausführung durch einen Computer diesen veranlassen, das Verfahren nach dem ersten Aspekt auszuführen. Dadurch werden die gleichen technischen Vorteile wie durch das Verfahren nach dem ersten Aspekt erreicht.

[0018] Ein vierter Aspekt betrifft eine Vorrichtung, die eingerichtet ist, das Verfahren nach dem ersten Aspekt auszuführen. Dadurch werden die gleichen technischen Vorteile wie durch das Verfahren nach dem ersten Aspekt erreicht.

[0019] Ausführungsbeispiele der Erfindung sind in den Zeichnungen dargestellt und in der nachfolgenden Beschreibung näher erläutert. Es zeigen

Fig. 1 eine schematische Ansicht eines Modells für maschinelles Lernen;

Fig. 2 eine schematische Ansicht einer semantischen Analyse; und

Fig. 3 ein Blockdiagramm des Verfahrens.

[0020] **Fig. 1** zeigt eine schematische Ansicht eines Modells **100** für maschinelles Lernen. Das Modell **100** umfasst ein neuronales Netz **101**. Das Modell **100** wird zu einem semantisch verbesserten maschinellen Lernen (Semantic Enhanced Machine Learning - SEML) durch eine semantisch modifizierte Back-

propagation (Semantic-Enhanced Backpropagation (SEBP)) verwendet. Dadurch wird eine höhere Transparenz und ein reduzierter Bedarf an Trainingsdaten erreicht. Das neuronale Netz **101** lernt auf diese Weise semantische Beziehungen zwischen den Daten. Intelligente, selbstlernende Systeme können damit schneller, besser und genauer lernen.

[0021] Das neuronale Netz **101** umfasst eine Eingabeschicht **103-1** (Input Layer), ein oder mehrere verarbeitende Schichten **103-2** (Hidden Layer) und eine Ausgabeschicht **103-M** (Output Layer). In die Eingabeschicht **103-1** wird ein Eingangsvektor **125** eingegeben, wie beispielsweise ein gegenwärtiger Aufenthaltsort eines Benutzers. Der Eingabevektor **125** wird mittels eines Encoders **109** erzeugt und stellt den Eingabewert **105** binär dar.

[0022] Die verarbeitenden Schichten **103-2** agieren als Neuronen mit Aktivierungsfunktion, wie beispielsweise einem Schwellwert (Signum-Funktion), einer Tangens-Hyperbolicus -Funktion, einer Sigmoid-Funktion oder einer ReLU-Funktion (ReLU - Rectified Linear Unit). Der binäre Ausgabevektor **107** des neuronalen Netzes **100** wird durch die Werte der Ausgabeschicht **101-M** erhalten. Aus dem Ausgabevektor **107** werden mittels des Decoders **111** die Ausgabedaten **127** erhalten.

[0023] Bei der Weiterleitung der Eingabedaten **105** an die nachfolgende verarbeitende Schicht **103-2** werden die jeweiligen Komponenten des Eingangsvektors mit entsprechenden Kantengewichten multipliziert. Dies geschieht ebenfalls bei jeder Weiterleitung an eine nachfolgende Schicht **103-M**. Die Kantengewichte einer Schicht **103-1**, ..., **103-M** sind beispielsweise Fließkommazahlen. Diese Vorwärtsprozessierung der Eingangsdaten durch das neuronale Netz wird als Forward-Propagation bezeichnet.

[0024] Die Ausgabedaten **127** in Reaktion auf die Eingabedaten **105** stellen den Lern-Istwert des neuronalen Netzes **101** dar. Sobald das neuronale Netz **101** den Lern-Istwert ermittelt hat, wird ein numerischer Abweichungswert **115** bestimmt, der eine semantische Abweichung zwischen dem berechneten Lern-Istwert des neuronalen Netzes **101** und dem vorgegebenen Lern-Sollwert angibt.

[0025] Bei der semantischen Vorhersage (Prädiktion) eines Aufenthaltsortes mittels des Modells **100** soll beispielsweise der semantische Typ (semantische Location) des nächstbesuchten Aufenthaltsortes eines Benutzers vorhergesagt werden. Wird ein gegenwärtiger Aufenthaltsort als Eingabe in das neuronale Netz **101** eingegeben, kann das angelernete neuronale Netz **101** beispielsweise den semantischen Typ des nächsten Aufenthaltsortes des Benutzers vorausberechnen.

[0026] Beispielsweise ist ein Benutzer zu einer bestimmten Uhrzeit zuhause. Das Modell **100** soll anhand von gelernten Daten vorhersagen, wo sich der Benutzer als nächstes aufhält. Der nächste Aufenthaltsort könnte beispielsweise ein Fast-Food-Restaurant **M** sein. Demnach wäre also Fast-Food-Restaurant **M** als Lokationsklasse der Instanz Fast-Food-Restaurant die korrekte Vorhersagelösung (Ground Truth).

[0027] Bei den heutigen Trainingsverfahren gibt es hierzu zwei Möglichkeiten. Erstens das Modell **100** schätzt richtig, d.h. Fast-Food-Restaurant. In diesem Fall lernt das Modell nichts und es wird mit der nächsten Vorhersage fortgefahren. Zweitens das Modell **100** schätzt falsch, wie beispielsweise „Fitness Studio“ als nächsten Aufenthaltsort des Benutzers. In diesem Fall wird das neuronale Netz **101** entsprechend bestraft, indem dessen Parameter entsprechend rekonfiguriert und aktualisiert werden. Obwohl es um eine semantische Information geht, verstehen aktuelle Vorhersagemodelle, die auf maschinellem Lernen basieren, nicht, worum es bei den semantischen Lokationen geht. Beispielsweise verstehen die Modelle nicht die Zusammenhänge und Bedeutung von „Zuhause“, „Arbeitsplatz“ oder „Restaurant“.

[0028] Falls das Modell **100** als nächsten Aufenthaltsort jedoch „Chinesisches Restaurant“ schätzen würde, würde es den nächsten Aufenthaltsort des Benutzers ebenfalls falsch vorhersagen. In derzeitigen Modellen **100** wird eine qualitativ gleiche Anpassung („Bestrafung“) des neuronalen Netzes **101** wie im Falle der falschen Vorhersage „Fitness Studio“ vorgenommen.

[0029] Allerdings liegt die Vorhersage „chinesisches Restaurant“ semantisch näher an „Fast-Food-Restaurant **M**“ als an „Fitness Studio“. Beide Vorhersagen gehören zu der Kategorie der „Fast Food Restaurants“. Ein Mensch würde diese Vorhersage des Modells intuitiver, unschärfer (fuzzier) als nah, aber nicht nah genug interpretieren.

[0030] Um dieses Wissen in das Modell **100** zurückzuführen, wird die Art der Interbeziehungen der Konzepte durch einen Abweichungswert **115** numerisch repräsentiert, der beispielsweise mit Hilfe einer semantischen Analyse **113** gewonnen wird. Der Abweichungswert **115** gibt den Grad einer semantischen Abweichung zwischen dem berechneten Lern-Istwert („Fitness-Studio“ oder „chinesisches Restaurant“) und dem Lern-Sollwert („Fast Food Restaurant“) an. Im Falle des Lern-Istwertes „Fitness-Studio“ ist der Abweichungswert **115** höher als im Falle des Lern-Istwertes „chinesisches Restaurant“. Auf diese Weise wird die semantische Lokation numerisch kodiert (z.B. binär) und ins Modell **100** eingeführt und weiterverarbeitet.

[0031] Die Information dieser semantischen Ähnlichkeit zwischen dem vorhergesagten Lern-Istwert des Modells **100** und dem tatsächlichen Lern-Sollwert (Ground Truth) kann verwendet werden, um das Training und das Modell **100** schneller zu optimieren. Der Begriff der Semantik bezieht sich auf die Interpretation und das Verstehen der Konzepte und der Beziehungen zwischen diesen. In der Informatik ist die Bestimmung der Semantik in der Regel auf Graph-ähnliche, konzeptuelle und logische Konstrukte (z.B. Knowledge Graphs, Linked Data, Ontologien) angewiesen.

[0032] Der numerische Abweichungswert **115** wird verwendet, um beispielsweise eine Verlustfunktion (Loss Funktion) **117** oder die Gewichte oder andere Parameter des Modells **100** entsprechend anzupassen und aktualisieren. Bei einem höheren semantischen Abweichungswert **115** wird das Modell **100** stärker angepasst als bei einem niedrigeren Abweichungswert **115**. Die Verlustfunktion **117** ist die Funktion, die im Prozess der Modellanpassung minimiert wird. Die Verlustfunktion **117** stellt ein ausgesuchtes Maß der Diskrepanz zwischen den beobachteten Daten und den mit der angepassten Funktion „vorhergesagten“ Daten dar.

[0033] Durch das semantisch verbesserte maschinelle Lernen wird semantisches Wissen in das Modell **100** zurückgegeben. Dabei kann das semantische Wissen Standardwissen sein, Common Sense Knowledge, Expertenwissen, oder persönliches (personalisiertes) Wissen sein, das von den Erfahrungen und Präferenzen des Benutzers beeinflusst wird. Im letzten Fall ist das Verfahren besonders im Bereich der Benutzer-Modellierung (User Modeling) and der Mensch-Computer-Interaction (Human-Computer-Interaction) vorteilhaft.

[0034] Fig. 2 zeigt eine schematische Ansicht einer semantischen Analyse **113** zur Bestimmung des Abweichungswertes **115**. Für die Bestimmung des semantischen Abweichungswertes **115** kann eine semantische Wissensbasis (Knowledge Base) **119**, oder semantisches Wissen im Allgemeinen **123** als Informationsquelle in einem selbstlernenden Mechanismus verwendet werden.

[0035] Strukturiertes Wissen, strukturierte Daten oder Informationen können direkt aus externen oder internen Wissens- und Datenbasen verwendet werden. Unstrukturiertes Wissen wird zuerst zu einem strukturierten Konstrukt vorverarbeitet. Der aus den Wissensbasen **119** extrahierte semantische Abweichungswert **115** kann entweder an der Ausgabe **129** des Modells **100** eingesetzt werden, beispielsweise durch semantische Backpropagation. Der extrahierte semantische Abweichungswert **115** kann auch im Kern des Modells **100** eingesetzt werden, beispielsweise durch Anpassung von Modellparametern.

tern, wie Gewichten, Übergangswahrscheinlichkeiten oder Hyperparametern von neuronalen Netzen (Filtergröße bei einem Convolutional Neural Network oder Speichergröße eines langen Kurzzeitgedächtnisses (LSTM - Long Short-Term Memory)). Der extrahierte semantische Abweichungswert **115** kann auch an der Eingabe **121** des Modells **100** eingesetzt werden, beispielsweise kann das Modell **100** einen Aufenthaltsort und einen ähnlichsten Aufenthaltsort gleichzeitig als Eingabe haben. Dabei erfolgt eine Extraktion des Abweichungswertes **115** durch Ähnlichkeits- oder Verwandtheitsanalyse oder eine Analyse komplexer Relationen (binary and n-ary Object Relations).

[0036] Für die Bestimmung des Abweichungswertes **115** können mehrere unterschiedliche Metriken für die semantische Ähnlichkeit (Semantic Similarity) verwendet werden. Die semantische Ähnlichkeit kann als Mittel zur Bestimmung des Abweichungswertes **115** und zur Integration vom semantischen Wissen in das Modell **100** für maschinelles Lernen verwendet werden.

[0037] Allerdings kann auch eine semantische Verwandtheit (Semantic Relatedness) als Mittel zur Bestimmung des Abweichungswertes **115** verwendet werden. Im Konzept der semantischen Verwandtheit sind beispielsweise die Begriffe „Auto“ und „Fahrer“ miteinander eng verwandt, während im Konzept der semantischen Ähnlichkeit die Begriffe „Auto“ und „LKW“ eine hohe Ähnlichkeit aufweisen. Allerdings kann auch auf andere Methoden zum Bestimmen des Abweichungswertes **115** zurückgegriffen werden.

[0038] Fig. 3 zeigt ein Blockdiagramm des Verfahrens zum Anlernen des Modells **100** für maschinelles Lernen. Das Verfahren umfasst den Schritt **101** eines Berechnens des Lern-istwertes durch das Modell **100** für maschinelles Lernen auf Basis von Eingabedaten **105**. Anschließend wird in Schritt **S102** der numerische Abweichungswert **115** bestimmt, der eine semantische Abweichung zwischen dem berechneten Lern-istwert und einem Lern-Sollwert angibt. Im Schritt **S103** wird das Modell **100** für maschinelles Lernen auf Basis des Abweichungswertes angepasst.

[0039] Beispielsweise kann das Verfahren mit einem künstlichen neuronalen Netz (Artificial Neural Network - ANN), einer Fehlerrückführung (Backpropagation Learning) und einer Lokationsvorhersage (Location Prediction) durchgeführt werden. Im Allgemeinen können jedoch auch andere Modelle verwendet werden, wie beispielsweise probabilistische Modelle.

[0040] Das Modell **100** lernt somit anhand des Abweichungswertes **115** zusätzlich semantische Relationen zwischen den jeweiligen Konzepten, die hinter den vorhergesagten Daten stecken. Im Falle von

anderen Modellen kann es sich um andere Parameter handeln, wie beispielsweise die Übergangswahrscheinlichkeiten im Fall eines probabilistischen Modells (z.B. Markov oder Bayes).

[0041] Der Gegenstand eines „Autos“ ist optisch und semantisch ähnlicher zum einem „Bus“ als zu einem „Baum“ bei einer optischen Objekterkennung. Der Gegenstand eines „Autos“ ist zwar nicht optisch aber semantisch ähnlicher zum einem „Motorrad“ als zu einem „Baum“. Wenn das Modell **100** viele „Auto“- und wenige „Motorrad“-Trainingsdaten verwendet, hat ein herkömmliches System Probleme bei der Erkennung von Motorrädern. Mit dem verbesserten Verfahren für maschinelles Lernen anhand des semantischen Abweichungswertes ist dies nicht der Fall.

[0042] Das Verfahren kann auf alle mögliche Klassifikations-, Erkennungs-, Vorhersage- und Regressionsprobleme in der realen Welt übertragen werden. Dies können intelligente (selbst-)lernende Systeme jeglicher Art, unabhängig von einer bestimmten Domäne sein, wie beispielsweise autonomes Fahren, Szenenerkennung oder eine Benutzerinteraktion. Das Verfahren für verbessertes maschinelles Lernen (Semantic-Enhanced Machine Learning - SEML) kann universell verwendet werden, um entsprechende Algorithmen hinsichtlich der Performanz (Klassifikations-, Prädiktions-, Mustererkennungsrate) und der Transparenz zu optimieren.

[0043] Das Verfahren kann am Beispiel eines multi-label Klassifikationsproblems, wie beispielsweise der semantischen Ortsvorhersage (Semantic Location Prediction) im Rahmen einer smarten Navigation, beschrieben werden. Das semantische Wissen kann demnach direkt oder indirekt entweder in die Eingabe, in die Ausgabe, oder direkt in das Modell (Transitionsmatrix, Gate sin LSTMs) eingeführt werden.

[0044] Durch das Verfahren wird der technische Vorteil einer verbesserten Performanz in Hinsicht auf die Vorhersage-, Klassifikations-, Regressions- oder Erkennungsrate und Genauigkeit, einer geringen Trainingszeit erreicht.

[0045] Der Schutzbereich der vorliegenden Erfindung ist durch die Ansprüche gegeben und wird durch die in der Beschreibung erläuterten oder den Figuren gezeigten Merkmale nicht beschränkt.

[0046] Alle in Verbindung mit einzelnen Ausführungsformen der Erfindung erläuterten und gezeigten Merkmale können in unterschiedlicher Kombination in dem erfindungsgemäßen Gegenstand vorgesehen sein, um gleichzeitig deren vorteilhafte Wirkungen zu realisieren.

ZITATE ENTHALTEN IN DER BESCHREIBUNG

Diese Liste der vom Anmelder aufgeführten Dokumente wurde automatisiert erzeugt und ist ausschließlich zur besseren Information des Lesers aufgenommen. Die Liste ist nicht Bestandteil der deutschen Patent- bzw. Gebrauchsmusteranmeldung. Das DPMA übernimmt keinerlei Haftung für etwaige Fehler oder Auslassungen.

Zitierte Patentliteratur

- DE 69217047 T2 [0004]
- DE 69130246 T2 [0005]
- DE 102016005407 A1 [0006]
- EP 1183619 B1 [0007]

Patentansprüche

1. Verfahren zum Anlernen eines Modells (100) für maschinelles Lernen, mit den Schritten:
 - Berechnen (S101) eines Lern-Istwertes durch das Modell (100) für maschinelles Lernen auf Basis von Eingabedaten (105);
 - Bestimmen (S102) eines numerischen Abweichungswertes (115), der eine semantische Abweichung zwischen dem berechneten Lern-Istwert und einem vorgegeben Lern-Sollwert angibt; und
 - Anpassen (S103) des Modells (100) für maschinelles Lernen auf Basis des Abweichungswertes (117).
2. Verfahren nach Anspruch 1, wobei das Modell (100) ein neuronales Netz implementiert.
3. Verfahren nach Anspruch 2, wobei eine Fehler-rückführung durch das neuronale Netz auf Basis des Abweichungswertes (115) durchgeführt wird.
4. Verfahren nach Anspruch 3, wobei eine Verlust-funktion (117) des neuronalen Netzes durch den Abweichungswert (115) angepasst wird.
5. Verfahren nach Anspruch 4, wobei die Verlust-funktion (117) mit dem Abweichungswert (115) multipliziert wird.
6. Verfahren nach einem der vorangehenden Ansprüche, wobei der Abweichungswert (115) eine semantische Ähnlichkeit oder semantische Verwandtheit zwischen dem Lern-Istwert und dem Lern-Sollwert angibt.
7. Verfahren nach einem der vorangehenden Ansprüche, wobei eine semantische Wissensbasis zum Bestimmen des numerischen Abweichungswertes (115) verwendet wird.
8. Verfahren nach einem der vorangehenden Ansprüche, wobei das Modell (100) für maschinelles Lernen ein probabilistisches Modell (100) ist und Übergangswahrscheinlichkeiten auf Basis des Abweichungswertes (115) angepasst werden.
9. Computerprogramm, umfassend Befehle, die bei der Ausführung des Programms durch einen Computer diesen veranlassen, das Verfahren nach einem der Ansprüche 1 bis 8 auszuführen.
10. Computerlesbares Speichermedium, umfassend Befehle, die bei der Ausführung durch einen Computer diesen veranlassen, das Verfahren nach einem der Ansprüche 1 bis 8 auszuführen.
11. Vorrichtung, die eingerichtet ist, das Verfahren nach einem der Ansprüche 1 bis 8 auszuführen.

Es folgen 2 Seiten Zeichnungen

Anhängende Zeichnungen

Fig. 1

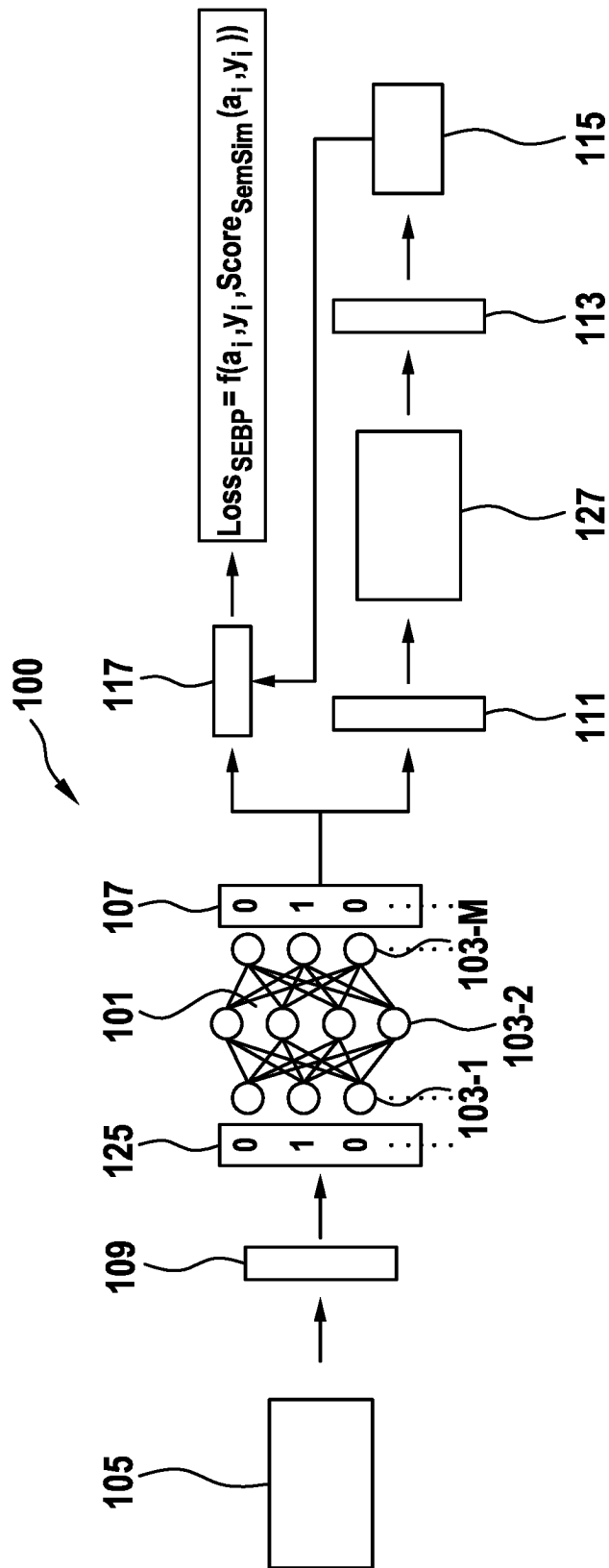


Fig. 2

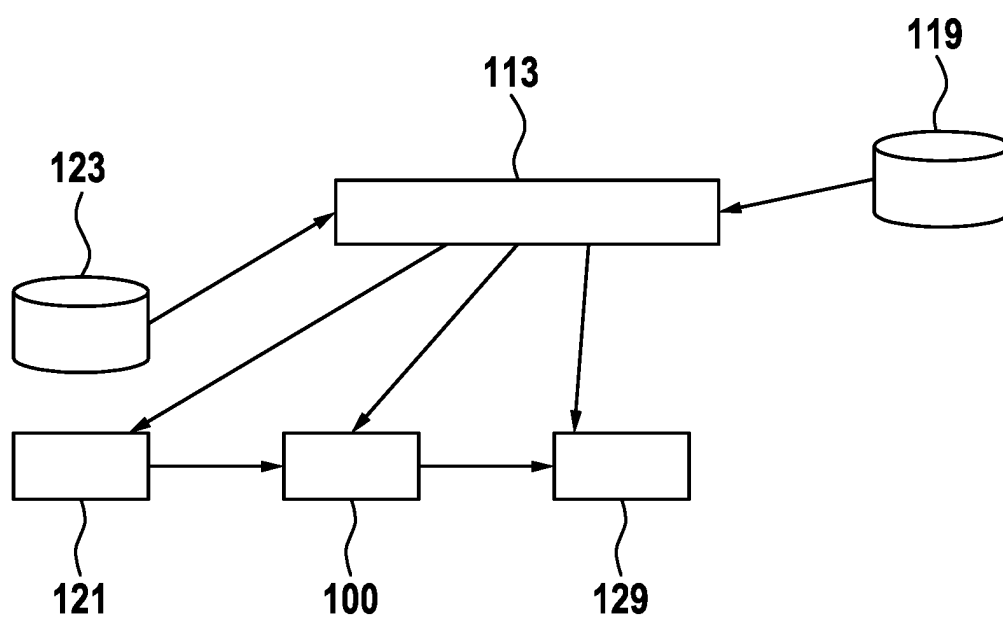


Fig. 3

