

# Big Data + AI im Gesundheitswesen

## Seminar NTBD 2018

Jan Zickermann

Universität Hamburg

8. Januar 2018



Betreuer: Dr. Julian Kunkel

# Übersicht

- 1 Motivation
- 2 Bioinformatik
  - Drug Discovery
- 3 Medical Imaging
- 4 Umfassende Diagnosen/Prognosen
  - Pervasive Sensing
  - Medical Informatics
  - Public Health
- 5 Ausblick

# Deep Learning im Gesundheitswesen

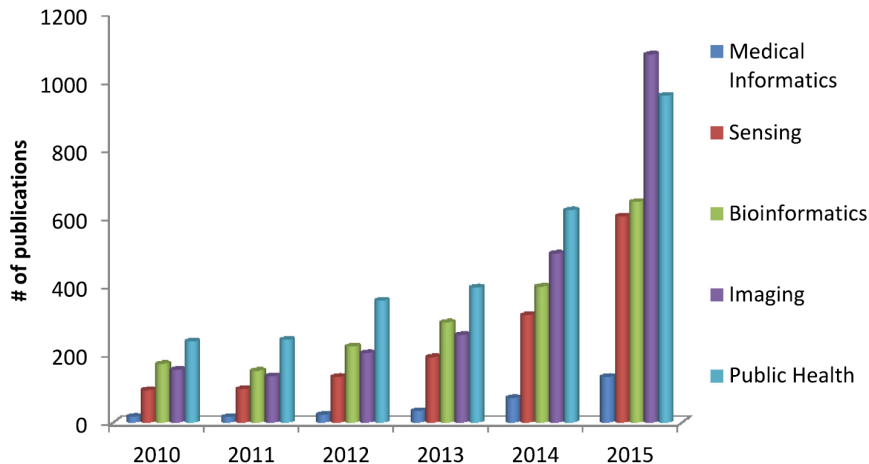


Abbildung: Publikationen, die *deep learning* mit *medical / health* enthalten (Ravi et al. [10])

# Bioinformatik

- Verwendete Daten
  - DNA, RNA, Moleküle und Proteinstrukturen
- **Datenknappheit** aufgrund von hohen Forschungskosten (trotz Fortschritte in DNA-Sequenzierung)
- Analyse von *compound-protein interaction*, genetische Variation, DNA-Methylierung
  - Vorhersagen biologischer Prozesse
  - Prävention von Krankheiten
  - Personalisierte Behandlung (*precision medicine*)

## Compound-Protein Interaction

- Keine Daten für **potentielle neue Medikamente** vorhanden
  - Fehlschläge: zu giftig, niedrige Wirksamkeit, schlechte Löslichkeit
  - Traditioneller Ansatz: aufwendige und teure Untersuchungen (mehrere Jahre)
  - *Virtual screening* mit Hilfe von *massively multitask networks*
- Medikamente nur wirksam mit 1-2% von Teststoffen (*rare events*)

## massively multitask networks

- *Softmax classifier* für jeden *Task*
- Höhere *performance* mit mehr *Tasks* und Daten
  - 40 Millionen Datenpunkte mit 249 *Tasks*
  - Erheblicher Performance-Zuwachs mit mehr *Tasks*
- *Hidden layers* generalisieren für mehrere *Tasks*
- Besondere Generalisierung im ersten *hidden layer* (*transfer learning* denkbar)

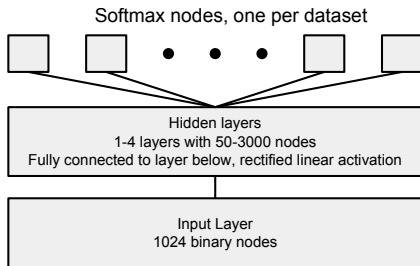


Abbildung: Multitask neural network (Bharath Ramsundar [3])

## *one-shot learning*

- Modifiziertes *matching network* für *one-shot learning*
  - Prinzip: Erlernen neuer Klassifikation mit einem oder wenigen Datenpunkten
  - Zu wenige Daten für *massively multitask networks (mmn)*
- Analoge, bekannte Stoffe mit (hoher Wirksamkeit und niedrigem Risiko) zur Hilfe nehmen
- **Sehr wenige Datenpunkte** für genaue Vorhersagen für *Verhalten von einem Molekül* in neuer Umgebung

# one-shot learning für Low Data Drug Discovery (I)

- Idee: Molekülstruktur  $\rightarrow$  Funktion
- Für jeden *Task* ein “support set”
  - $S = \{x_i, y_i\}_{i=1}^m, y_i \in \{0, 1\}$  (aktiv/inaktiv Beispiele)
- Sei  $x$  ein **neues Molekül** ohne experimentelle Untersuchung
- Vorhersage:  $x$  aktiv/inaktiv für *Task*
  - Funktion  $a$  berechnet die Ähnlichkeit von  $x$  zu bekannten Molekülen  $x_i$  mit Aktivität  $y_i$

$$h_S(x) = \sum_{i=1}^m a(x, x_i) y_i$$



## one-shot learning für Low Data Drug Discovery (II)

$$a(x, x_i) = \frac{k(f(x), g(x_i))}{\sum_{j=1}^m k(f(x), g(x_j))}$$

- $k$  als Distanzfunktion (Kosinus-Ähnlichkeit)
- *Molecular graph convolution networks*  $f'$ ,  $g'$  verwendet
  - Moleküle als Graph: *graph gather*, *graph pooling*
  - Komplexe Funktionen von Molekülstrukturen reduziert auf *local geometry*
- Kombination der Vektoren von  $f'$ ,  $g'$  mit *LSTM*
  - *RNN* mit Langzeitabhängigkeiten
  - *Attention* basiert auf Kontext (mit *attLSTM*), um Reihenfolgeabhängigkeit aufzulösen

# one-shot learning für Low Data Drug Discovery (III)

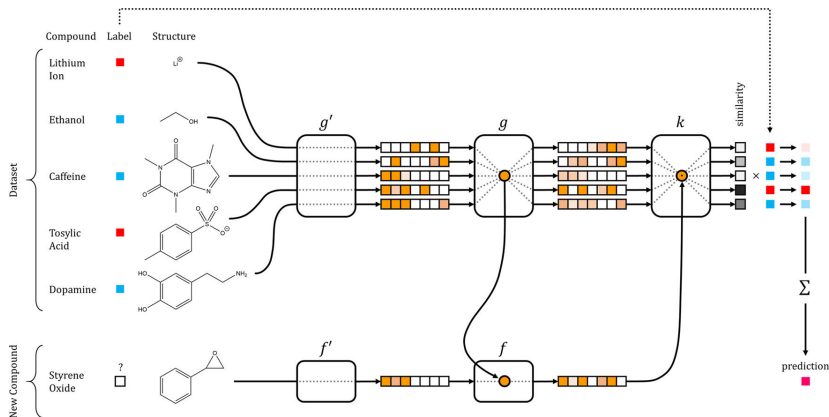


Abbildung: One-shot learning in drug discovery (Altae-Tran et al. [1])

$$f(x|S) = attLSTM(f'(x), \{g(x_i|S)\})$$

# Medical Imaging

- *Computer Aided Detection (CADe)*
  - **Potentielle Funde** werden Radiologen mitgeteilt
  - Beispiel *Brustkrebs*: Mikrokalk als helle Punkte in Mammographie sichtbar
  - Mikrokalkpixel werden erkannt (“*MC present/absent*“ classes)
- *Computer Aided Diagnosis (CADx)*
  - **Direkte Wahrscheinlichkeit** eines Befundes (ohne Radiologeneinfluss)
  - Mikrokalkbereiche automatisch erkennen (genauer als Radiologen)
  - Einschätzung, ob Bereiche bösartig sind durch NN
  - “*black box*“ Problem

# Automatische Bildqualitätsbestimmung

- Qualität bestimmt durch Diagnosefähigkeit des Betrachters nach *Luster 1970*
  - Sensitivität bei niedriger *false positive* Rate
  - *Area under curve (AUC)* von *receiver operating characteristic (ROC) curve*
  - Physikalische Genauigkeit nicht ausschlaggebend
- Optimales Weichzeichnen: **sicherere Klassifikation** durch Betrachter
- *Channelized Hotelling observer (CHO)* modelliert menschlichen Betrachter
  - Berechnet Zuversicht, ob eine Abnormität im Bild vorliegt
  - Erlernt durch menschliche Beispiele

# Brain Mapping

- Änderung der räumlichen Hirnaktivität mit Antidepressivum untersucht
  - Placebo- und Medikamentenzustand verglichen
  - Dimensionen der Daten reduziert mit SVD
  - Aktivitätsunterschied farblich dargestellt

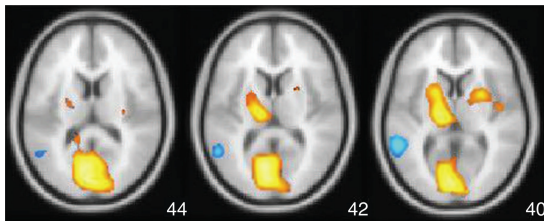


Abbildung: Räumliche Aktivierungsmuster im Gehirn (Miles N. Wernick [9])

# CNN für *Medical Imaging*

- Diagnosen für Hirn und Herz durch *MRI*-Daten
- *CNN* Herausforderungen:
  - ① 3D Daten mit variierenden *voxel* Größen
    - Bild verliert Schärfe durch Interpolation (andere Transformation notwendig)
  - ② Räumliche Abhängigkeiten von *CNN* nicht modelliert
    - Informationen von anliegenden/benachbarten Regionen fehlen
  - ③ *preprocessing* notwendig für einheitlichen Datensatz
    - Gleichzeitig hohe Entropie (Selektion nötig)
- **Erfolgreicheres Lernen:** vor-trainiertes *CNN* (mit natürlichen Bildern)
  - *Autoencoder* für *transfer learning*
- Größerer Datensatz durch *Crowdsourcing* (geringe Kosten, Qualitätsbedenken)

# CNN für *Medical Imaging*

- Diagnosen für Hirn und Herz durch *MRI*-Daten
- *CNN* Herausforderungen:
  - ① 3D Daten mit variierenden *voxel* Größen
    - Bild verliert Schärfe durch Interpolation (andere Transformation notwendig)
  - ② Räumliche Abhängigkeiten von *CNN* nicht modelliert
    - Informationen von anliegenden/benachbarten Regionen fehlen
  - ③ *preprocessing* notwendig für einheitlichen Datensatz
    - Gleichzeitig hohe Entropie (Selektion nötig)
- **Erfolgreicheres Lernen:** vor-trainiertes *CNN* (mit natürlichen Bildern)
  - *Autoencoder* für *transfer learning*
- Größerer Datensatz durch *Crowdsourcing* (geringe Kosten, Qualitätsbedenken)

# Gesundheits-Apps

- Weltweit **500 Millionen Nutzer** von *health-care apps* (2010-2015)
- Physiologische, biologische, Verhaltens- und Umgebungsdaten (auch Benutzerangaben)
- *Smartphone* oder Geräte von u.a. *Apple, Fitbit, Nike*
- **Eigenüberwachung** und Quantifizierung der körperlichen Leistung mit *Apps*
- Nützlich für medizinische Fachkräfte
  - Körperliche/geistige Gesundheit einschätzen, Medikamente verschreiben, klinische Entscheidungen



# Nutzen von *Pervasive Sensing*

- Bewegungs- und Standortdaten
  - Patienten lokalisieren: *GPS*
- Umgebungsdaten
  - Schlafmuster analysieren: Umgebungslichtsensor des *Smartphones*
- Physiologische Daten
  - Allgemeine Einschätzung: Einzelgeräte für elektrodermale Aktivität, Blutdruck, Elektroenzephalografie (*EEG*)

## Nutzen von *Pervasive Sensing* cont.

- Körperliche Gesundheit verbessern
  - Implantate: Reha-Patienten können Zuhause leben
  - *Wearables + apps*: Essen und Energieverbrauch überwachen (bei **Fettleibigkeit, Diabetes**)
  - *Abnormality detection*: schnelles Reagieren für schwer kranke Patienten
- Geistige Gesundheit einschätzen durch *Smartphone*
  - GPS, Umgebungslicht, Textnachrichten, Anrufverlauf und Gefühlslage (durch Benutzerangaben)
  - *ML* gestütztes Feedback (**Depression-Prognose, Suizidprävention**)

# Physiologische Sensoren



**Abbildung:** Ansammlung an gewerblich erhältlichen tragbare Sensoren für physiologische Einschätzung. (1) Aktivitätsüberwacher (Fitbit), (2) EEG Sensor (Emotive), (3) EKG Sensor, (4) PPG Sensor, (5) Blutdruckprüfer, (6) Atmungssensor (Spire). (David D. Luxton [5])

# Medical Informatics

- Analyse von sehr großen Ansammlungen von Behandlungsverläufen
- *Electronic Health Records (EHR)*
  - Diagnosen, Untersuchungen, Medikamente, Behandlungspläne, Impfungen, Allergien, Röntgenbilder
  - Sonstiges: *EEG* Sensoren oder Labortests
- Komplexität von *EHR*
  - Irreguläre Datenlängen mit Lücken (von Klinik und Personen abhängig)
  - Viele Abhängigkeiten in Patientendaten (Generalisierung erschwert)

# EHR Verwendung

- Automatische Erkennung von *features* mit *denoising autoencoder*
  - Wahrscheinlichkeit Krankheit zu entwickeln:  
Diabetes, Schizophrenie, Krebs
- Wiederaufnahme und Aufenthaltsdauer vorhersagen
  - Studien mit *LSTM* für zeitabhängige Muster
- Optimierung von **Klinikorganisation**  
(Zeitplanung/Patientenentlassungen)
- Personalisierte Behandlungen
- Neue *features* → neue Forschungsfragen

# Public Health

- Meist komplexe Krankheitsverbreitungsmodelle
  - Expertenwissen erforderlich: Physik, Chemie, Biologie
- *Online learning* Beispiele
  - **Luftverschmutzungsvorhersagen**: basierend auf Sensoren in 52 Städten in Japan
  - **Krankheitsverbreitung**: basierend auf 150.000 *Tweets* (412 manuelle *label*)
- *Tracking* durch Metadaten und *Social Media* Aktivität
  - Depressionen vorhersagen (von Müttern nach der Geburt)

# Google Flu - Suchaktivität „Grippe“

Flu search activity (standard deviation from baseline) ?

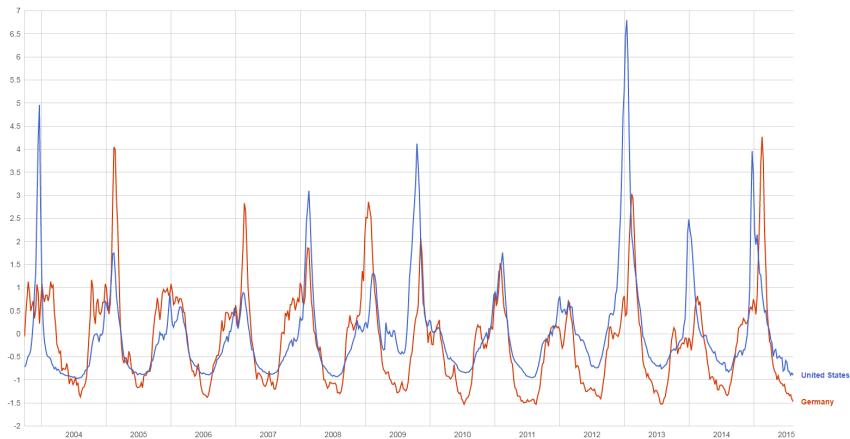


Abbildung: Google Flu Trendeinschätzung (Google [6])

# Durkheim Projekt

- **Suizidprävention** bei US-Veteranen (2011-2015)
- Basierend auf *EHR* und *Social Media*
  - Suizidgedanken werden nicht gemeldet (Stigmatisierung)
  - Ärzte können individuell Risiko einschätzen (für Intervention)
  - Meldungen für Patienten
- *Opt-in: Social Media* über *Facebook*
- **Vertrauen:** in Kompetenz und Datenschutz



# Durkheim Projekt Modell

- Auf Sprache basiertes Modell für Suizid-Prognose
- Variante von *decision tree forest* (*supervised learning*)
  - Benutzt: *genetic* (Vielfalt) + *maximum entropy* (Größe minimieren)
- 3 Gruppen (Suizid,  $\neg$ Suizid + Behandlung,  $\neg$ Suizid) je 70 Personen
- Verwendete **Wortgruppen analysiert** für Gruppenzuordnung
  - *Accuracy* > 67%
- **Echtzeit**: Intervention falls Risiko über Schwellwert (*Facebook, E-Mail*)



# Fotos auf Instagram - Depressionsmerkmale

- 166 *Instagram* Nutzer mit 43.950 Fotos
- Depression entdecken ohne klinische Diagnose (70% korrekt klassifiziert)

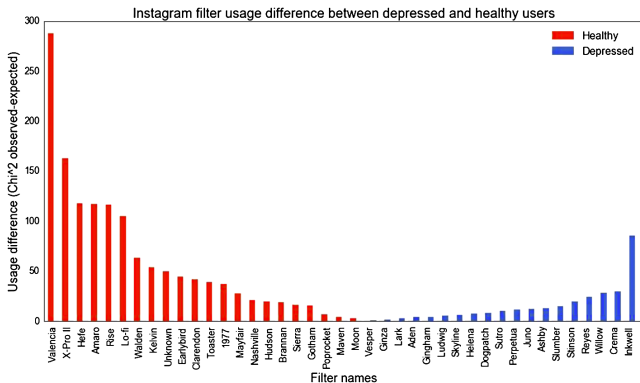


Abbildung: Instagram Filterbenutzung der Teilnehmer (Andrew G Reece [2])

# Prognose des physiologischen Alters

- Physiologisches mit chronologischem Alter vergleichen (allgemeine Gesundheit)
- 377.686 *EMR* von Patienten (18-85 Jährigen)
- 3 *hidden layers* (je 1000 Neuronen) mit 40% *dropout*
- **Schätzfehler**: Standardabweichung von 7 Jahren
  
- *Genome-wide association study (GWAS)*: neue Genvarianzen in Verbindung mit dem Altern
- **Höherer Werte** im Alter: z.B. Blutdruck, Glukose, Stickstoff im Urin
- Einige Werte besitzen Wendepunkt ab 50-60 möglicherweise durch *survival bias*
- Ungenauere Vorhersagen mit Alter 40-55

# Prognose des physiologischen Alters

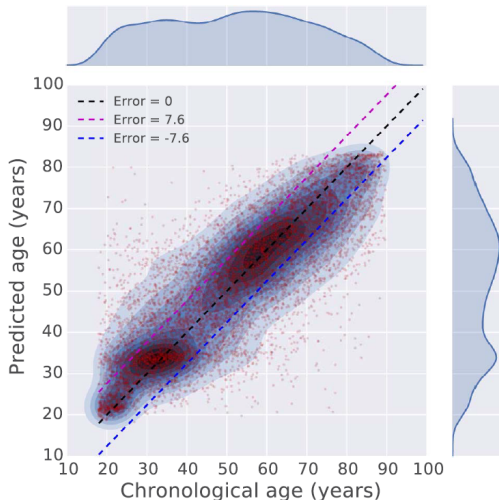


Abbildung: Vergleich chronologisches/vorhergesagtes Alter. (Zichen Wang [12])

# Ausblick

- Ärzte müssen Umgang mit *Big Data* lernen
  - Verkürzte Verarbeitungszeiten
  - Qualitätsverbesserungen
- Hunderte Biomarker, Bildmaterial von Millionen, aggregierte Forschungsergebnisse und *EHR*
- AI gründlichere Verarbeitung ohne **menschliche kognitive Begrenzung**

[8, 7]

# Ausblick - Potential

- *Medical Imaging*
  - **Großes Entwicklungspotential:** erweitertes Bildverständnis (wie *Facebook*)
  - Mehr Daten (Auflösung, 3D) → mehr/verbesserte AI
  - Dermatologen gleiche Leistung und effektiver als Radiologen
  - AI könnte Radiologen ersetzen (medizinisches Gebiet)
- Verbesserte *scores* (mehr Variablen)
  - Für Prognosen, Wiederaufnahmen, Sterberaten
- Ärzte ↔ Patienten (Kommunikation)
  - Intelligente Medikamentenauswahl: Kostendeckung durch Krankenkasse, Vorlieben des Patienten
  - Überbuchung, Zeitplanung
- Sonstiges: *sensing*, *precision medicine*

# Ausblick - Unternehmen

- **IBM-Watson**: Krebsdiagnose und Behandlungsvorschläge (und DNA-Analyse)
- **Google DeepMind**: Differenzierung von Augenkrankheiten (und Effizienz bei Strahlentherapie)
- „Dr. AI“ von **HealthTag**
  - Basierend auf ärztlichem Wissen + EHR
  - Patientensymptome → Einschätzung und Wegweisungen



# Ausblick - Schwierigkeiten für AI

- Anspruchsvolle/komplexe Unterhaltung (**Vertrauen** der Patienten)
- **Kritisches Denken** (z.B. Neurologie)
- Neue Diagnosen und Entscheidungen
- Schuldfrage bei Versagen von AI

# Übersicht

- Bioinformatik - *Drug Discovery*
  - *Massively multitask networks* und *one-shot learning*
- *Medical Imaging*
  - Bildqualität und *MRI*
- Umfassende Diagnosen/Prognosen
  - *Pervasive Sensing, Medical Informatics, Public Health*
  - Suizidprävention, Altersprognose
- Potential von AI im Gesundheitswesen

- Wachsender Markt mit neuen Unternehmen, jedoch noch **nicht weit entwickelt**
- Bei geistige Beeinträchtigung z.B. Diagnose/Behandlung Autismus, als Bewegungsaufmunterung, gegen Einsamkeit
- Pflegekräfte unterstützen und nicht sichtbare geistige Beeinträchtigungen erkennen
- *Zu wenige klinische Untersuchungen* (teilweise kein Mehrwert festgestellt)
- **Vertrauensproblem:** Aussehen (Mensch/Cartoon), Verhalten (Kulturunterschiede)

[11]

# References

- [1] Han Altae-Tran, Bharath Ramsundar, Aneesh S. Pappu, and Vijay Pande. Low data drug discovery with one-shot learning. *ACS Central Science*, 3(4):283–293, 2017. doi: 10.1021/acscentsci.6b00367. URL <http://dx.doi.org/10.1021/acscentsci.6b00367>.
- [2] Christopher M Danforth Andrew G Reece. Instagram photos reveal predictive markers of depression. *EPJ Data Science*, 2017. URL <https://epjdatascience.springeropen.com/track/pdf/10.1140/epjds/s13688-017-0110-z?site=epjdatascience.springeropen.com>.
- [3] Patrick Riley Dale Webster David Konerding Vijay Pande Bharath Ramsundar, Steven Kearnes. Massively multitask networks for drug discovery. 2015. URL <https://arxiv.org/abs/1502.02072>.
- [4] Craig Bryan Chris Poulin, Paul Thompson. Chapter 9 - public health surveillance: Predictive analytics and big data. In *Artificial Intelligence in Behavioral and Mental Health Care*, pages 205–230. 2015.
- [5] Akane Sano Timothy Bickmore David D. Luxton, Jennifer D. June. Chapter 6 – intelligent mobile, wearable, and ambient technologies for behavioral health care. In *Artificial Intelligence in Behavioral and Mental Health Care*, pages 137–162. 2015.
- [6] Google. Google flu trends estimates - google public data explorer. [https://www.google.com/publicdata/explore?ds=z3bsqef7ki44ac\\_](https://www.google.com/publicdata/explore?ds=z3bsqef7ki44ac_). Accessed: 2018-01-04.
- [7] Ezekiel Emanuel Katie Chockley. The end of radiology? three threats to the future practice of radiology. *Journal of the American College of Radiology*, 3:1415–1420, 2016.
- [8] C. Krittanawong. The rise of artificial intelligence and the uncertain future for physicians. *European Journal of Internal Medicine*, 2017.
- [9] Jovan G. Brankov Grigori Yourganov Stephen C. Strother Miles N. Wernick, Yongyi Yang. Machine learning in medical imaging. 2010. URL <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4220564>.
- [10] D. Ravi, C. Wong, F. Deligianni, M. Berthelot, J. Andreu-Perez, B. Lo, and G. Z. Yang. Deep learning for health informatics. *IEEE Journal of Biomedical and Health Informatics*, 21(1):4–21, Jan 2017. ISSN 2168-2194. doi: 10.1109/JBHI.2016.2636665.
- [11] Laurel D. Riek. Chapter 8 - robotics technology in mental health care. In *Artificial Intelligence in Behavioral and Mental Health Care*, pages 185–203. 2015.
- [12] Benjamin S.Glicksberg Ariel Israel Joel T. Dudley Avi Ma'ayan Zichen Wang, Li Li. Predicting age by mining electronic medical records with deep learning characterizes differences between chronological and physiological age. *Journal of Biomedical Informatics*, 2015.