

Diagnose per Computer: Gefährliche Krankheitserreger mithilfe maschinellen Lernens erkennen

HIRI-Forscher in Würzburg haben ein neues maschinelles Lern-Tool mitentwickelt, das helfen kann, gefährliche Krankheitserreger schneller zu entdecken

Per Computer erkennen, ob ein neu auftretender Salmonellenstamm eher eine gefährliche Infektion in der menschlichen Blutbahn oder eine Lebensmittelvergiftung verursacht – das kann ein neu entwickeltes maschinelles Lern-Tool. Die Software entstand im Rahmen eines internationalen Projekts von Wissenschaftlern des Wellcome Sanger Institute (Großbritannien), der University of Otago (Neuseeland) und des Helmholtz-Instituts für RNA-basierte Infektionsforschung (Würzburg), einem Tochterinstitut des Helmholtz-Zentrums für Infektionsforschung in Kooperation mit der Julius-Maximilians-Universität Würzburg. Die Forscher können nun deutlich schneller solche genetischen Veränderungen identifizieren, die typischerweise im Zusammenhang mit aggressiven Salmonellenstämmen stehen. Damit ist das neue Tool auch von hoher Relevanz für das Gesundheitswesen. Wie kürzlich das Fachjournal *PLOS Genetics* berichtete, könnte das Lern-Tool von großer Bedeutung sein, um gefährliche Krankheitserreger zu erkennen, bevor sie einen Krankheitsausbruch verursachen.

Seit Genom-Sequenzierungen erschwinglich geworden sind, setzen Wissenschaftler sie auf der ganzen Welt ein, um anhand der genetischen Eigenschaften von Bakterien besser zu verstehen, wie diese Erreger Infektionen verursachen, wie sich Krankheiten verbreiten, wie Bakterien resistent gegen Medikamente werden und welche Bakterienstämme Krankheitsausbrüche auslösen können. Die Methoden, die derzeit zur Verfügung stehen, um die genetischen Veränderungen neuer Erregerstämme bei einem Krankheitsausbruch zu identifizieren, sind allerdings sehr zeitaufwendig und oft mit dem manuellen Abgleichen des neuen Stammes mit vorhandenen Referenz-Sammlungen verbunden.

Die Bakteriengattung *Salmonella*, im deutschen Sprachgebrauch Salmonellen genannt, umfasst viele verschiedene Arten. Einige von ihnen verursachen Lebensmittelvergiftungen, bekannt als Magen-Darm-Salmonellen, während andere weit über den Darm hinausreichende Krankheiten hervorrufen – wie der Typhuserreger *Salmonella enterica* Serovar Typhi.

Um zu verstehen, wie bestimmte genetische Veränderungen dazu führen, dass ein neu auftretender Salmonellen-Stamm eine Lebensmittelvergiftung oder eine weitreichendere Infektion auslöst, haben die Forscher ein sogenanntes Machine-Learning-Tool entwickelt. Diese Software analysiert, welche Mutationen eine wichtige Rolle in der Krankheitsentwicklung spielen.

Das Entwicklerteam etablierte dafür ein Modell, das darauf trainiert wurde, wichtige krankheitsassoziierte Unterschiede zwischen verschiedenen Salmonellenstämmen gezielt zu erkennen. Als Trainingsdatensatz diente den Forschern die genetische Information von 13 bekannten und genetisch sehr unterschiedlichen *Salmonella enterica*-Stämmen mit unterschiedlichen Krankheitspotenzialen. Das Modell identifizierte dabei fast 200 Gene, die einen Einfluss darauf haben, ob ein Bakterium eher eine Lebensmittelvergiftung oder einen invasiven Infektionsprozess auslöst.

Dr. Nicole Wheeler, Ko-Hauptautorin der Studie am Wellcome Sanger Institute, sagte: „Wir haben ein neues Machine-Learning-Modell entwickelt, mit dem wir erkennen können, ob neu auftretende Bakterienstämme ein potenzielles Risiko für die öffentliche Gesundheit darstellen. Mit diesem Tool können wir große Datenmengen bewältigen und in Sekundenschnelle Ergebnisse erhalten. Dies wird zukünftig eine Form der Überwachung von krankheitserregenden Bakterien ermöglichen, die bisher noch nicht denkbar war – und das nicht nur auf einzelnen Krankenstationen, sondern weltweit.

Seinen ersten Praxiseinsatz hat das maschinelle Lern-Tool bereits erfolgreich bestanden: Beim Vergleich neuer *Salmonella*-Stämme, die derzeit südlich der Sahara auftreten, hat die Software aus der Gesamtheit häufig auftretender Infektionstypen (*Salmonella* Enteritidis und *Salmonella* Typhimurium) gezielt zwei Varianten des Erregers hervorgehoben, die besonders gefährlich sind und eine höhere Anzahl von Blutbahn-Infektionen auslösen. Diese Infektionen bedrohen vor allem Menschen mit einem geschwächten Immunsystem, wie es zum Beispiel HIV-Infizierte haben. Das maschinelle Lern-Tool zeigte deutlich die genetischen Veränderungen auf, die es diesen beiden *Salmonella*-Stämmen ermöglichen, sich an ihre Wirte anzupassen und damit invasiver zu werden.

Dr. Lars Barquist, Ko-Hauptautor am Helmholtz-Institut für RNA-basierte Infektionsforschung in Würzburg, sagte: „Das maschinelle Lern-Tool ist ein bedeutender methodischer Fortschritt, da es nicht nur nach Genen und Mutationen sucht, sondern auch nach den funktionellen Auswirkungen, die diese Mutationen haben. Es kann damit vorhersagen, welche Mutationen es den Krankheitserregern ermöglichen, sich über den Darm hinaus zu verbreiten und damit lebensbedrohliche Krankheiten auszulösen. Das wird zukünftig dabei helfen, effektivere Behandlungsmethoden zu entwickeln.“

Das maschinelle Lern-Tool ist nicht auf Salmonellen beschränkt, sondern vielfältig einsetzbar. Genauso könnte es zur Untersuchung anderer Faktoren, zum Beispiel entstehende Antibiotikaresistenzen bei verschiedensten Krankheitserregern, angewendet werden. Mit diesem Tool lassen sich gefährliche Bakterienstämme in Echtzeit identifizieren und Krankheitsausbrüche verhindern.

Dr. Nicholas Feasey von der Liverpool School of Tropical Medicine sagte dazu: „Wir nutzen diesen Ansatz bereits, um nach den wichtigsten Unterschieden zwischen asiatischen und afrikanischen *Salmonella enterica* (Serovar Typhi)-Stämmen zu suchen. Anstatt die Genome verschiedener Bakterienstämme in wochen- und monatelanger Kleinarbeit manuell zu vergleichen, sind wir nun in der Lage, die genetischen Veränderungen neuer Krankheitserreger sofort zu bestimmen. Das bietet uns die Möglichkeit, Krankheitsausbrüche sozusagen live zu verfolgen und damit schnell gesundheitspolitische Maßnahmen zur Kontrolle und Prävention ergreifen zu können.“

Originalpublikation:

Nicole E. Wheeler, Paul P. Gardner, Lars Barquist: Machine learning identifies signatures of host adaptation in the bacterial pathogen *Salmonella enterica*. *PLOS Genetics* 2018, DOI: 10.1371/journal.pgen.1007333

Diese Pressemitteilung und Bildmaterial finden Sie auch auf unserer Webseite unter dem Link https://www.helmholtz-hzi.de/de/aktuelles/news/ansicht/article/complete/diagnose_per_computer_gefaehrliche_krankheitserreger_mithilfe_maschinellen_lernens_erkennen/

Die Pressemitteilung des Wellcome Sanger Institutes finden Sie unter <https://www.sanger.ac.uk/news/view/machine-learning-flags-emerging-pathogens>

Förderung:

Die Forschungsarbeit wurde gefördert durch folgende Partner: Wellcome (grant 206194), University of Canterbury, Biomolecular Interaction Centre, Alexander von Humboldt-Stiftung sowie Rutherford Discovery Fellowship. Koordiniert wurde sie von der Royal Society of New Zealand.

Über Maschinelles Lernen:

Machine Learning steht für maschinelles Lernen als ein Teilgebiet der Informatik, das ausgehend von der Mustererkennung und dem Computerlernen eine neue Perspektive auf die künstliche Intelligenz setzen möchte. Machine Learning untersucht die Konstruktion von Algorithmen, die Lernprozesse simulieren können. Ziel ist es, Computerprogramme so weit zu bringen, dass sie annähernd eigenständig kognitive Prozesse umsetzen können. Die Methoden des maschinellen Lernens werden in der Anwendung häufig als Predictive Analytics bzw. Predictive Modelling bezeichnet, um Vorhersagen möglicher Varianten für die Zukunft zu erhalten. Maschinelles Lernen kann als Oberbegriff bei der künstlichen Generierung von Erfahrungswissen bestimmt werden. Ein Computersystem lernt an praktischen Beispielen, interaktive Eingriffe nach jedem Lernschritt geben dann eine Rückmeldung, womit sich das System selbstständig verbessern kann. Enge Berührungspunkte gibt es zum Data-Mining: Zusammenhänge in großen Datenmengen sollen mit maschinellern Lernen besser bestimmt werden.

Helmholtz-Zentrum für Infektionsforschung:

Am Helmholtz-Zentrum für Infektionsforschung (HZI) untersuchen Wissenschaftler die Mechanismen von Infektionen und ihrer Abwehr. Was Bakterien oder Viren zu Krankheitserregern macht: Das zu verstehen soll den Schlüssel zur Entwicklung neuer Medikamente und Impfstoffe liefern. Das HZI ist Mitglied im Deutschen Zentrum für Infektionsforschung (DZIF).

www.helmholtz-hzi.de

Helmholtz-Institut für RNA-basierte Infektionsforschung:

Das Helmholtz-Institut für RNA-basierte Infektionsforschung wurde im Mai 2017 als gemeinsame Institution durch das Helmholtz-Zentrum für Infektionsforschung (HZI) und die Julius-Maximilians-Universität Würzburg (JMU) gegründet. HIRI ist die weltweit erste Forschungseinrichtung, die die Bereiche RNA-Biologie und Infektionsforschung vereint. www.helmholtz-hiri.de

Biomolekulares Interaktionszentrum, University of Canterbury, Neuseeland:

Das Biomolekulare Interaktionszentrum ist ein multidisziplinäres Forschungszentrum, das sich der Erforschung molekularer Wechselwirkungen widmet, die für die biologische Funktion entscheidend sind. Das Verständnis dieser Wechselwirkungen ist von zentraler Bedeutung für die Grundlagenforschung, für neue Therapieansätze und eine breite Palette hochfunktioneller Produkte. www.canterbury.ac.nz/bic

Wellcome Sanger Institute:

Das Wellcome Sanger Institute ist eines der weltweit führenden Genomzentren. Durch seine Fähigkeit, Forschung in großem Maßstab zu betreiben, ist es in der Lage, sich an neuartigen und langfristigen Forschungsprojekten zu beteiligen, die darauf abzielen, die medizinische Wissenschaft weltweit zu beeinflussen und zu stärken. Die Forschungsergebnisse des Instituts, die durch eigene Forschungsprogramme und durch seine führende Rolle in internationalen Konsortien gewonnen werden, fließen in die Entwicklung neuer Diagnostika und Therapien für Krankheiten des Menschen ein. Zur Feier seines 25-jährigen Bestehens im Jahr 2018 sequenziert das Institut 25 neue Genome von Arten in Großbritannien. Erfahren Sie mehr unter www.sanger.ac.uk oder folgen Sie

@sangerinstitute auf Twitter.

Wellcome:

Wellcome existiert, um die Gesundheit für alle zu verbessern, indem es großen Ideen hilft zu gedeihen. Wellcome ist eine weltweit agierende, gemeinnützige Stiftung und sowohl politisch als auch finanziell unabhängig. Sie unterstützt Wissenschaftler und Forscher, stellt sich großen gesellschaftlichen Herausforderungen, treibt die Fantasie an und entfacht Debatten.

www.wellcome.ac.uk