

Bericht zum Hackathon “hack4health”

an der Universität Osnabrück, 12. bis 16. Februar 2018

Fabian Eckelmann, Stéphane Ghozzi, Alexander Ullrich, RKI-FG31 / Signale

10 Juli 2018

Der Hackathon “hack4health - Cognitive Computing Hackathon on Epidemiology” wurde vom 12. bis 16. Februar an der Universität Osnabrück ausgetragen und hatte zum Ziel moderne Methoden des maschinellen Lernens auf klassische Probleme der Infektionsepidemiologie anzuwenden. Die Veranstaltung wurde gemeinschaftlich vom Robert Koch-Institut (RKI), dem Institut für Kognitionswissenschaften (IKW) der Universität Osnabrück und dem Gesundheitscampus Osnabrück organisiert. Auf Seite des RKI engagierten sich Fabian Eckelmann, Stéphane Ghozzi und Alexander Ullrich des Fachgebiets 31 für Datenmanagement, die Finanzierung stammte dabei aus dem BMG-geförderten Drittmittel Projekt “Signale 2.0”.

Motivation

Sowohl das RKI als auch die Abteilung 3 (Infektionsepidemiologie) wollen vermehrt in digitale Epidemiologie investieren und in diesem Bereich ihr Profil schärfen. Von der digitalen Epidemiologie werden sich neue Erkenntnisse erhofft, die so auf traditionellem Wege nicht oder nur schwer erreicht werden können. Allerdings ist es auch ein neues Gebiet und wirklich vielversprechende Anwendungen müssen erst noch erkundet werden. Dafür ist es unerlässlich neue Wege zu gehen und neue Ideen auszuprobieren und evaluieren. Ein Hackathon kann genau bei diesen Aufgaben helfen. Verschiedene Ideen können in kurzer Zeit als Prototypen realisiert und dann von Experten begutachtet werden.

Das Projekt “Signale 2.0” hat sich zum Ziel gesetzt moderne Methoden des maschinellen Lernens zu nutzen um die epidemiologische Lage bezüglich Infektionskrankheiten besser zu verstehen und für die Vorhersage von möglichen Ausbrüchen zu nutzen. Dabei sollen auch neuartige oder bisher vernachlässigte Datenquellen extrahiert und verwendet werden. Das IKW und seine Studenten besitzen große Erfahrung und Wissen auf dem Gebiet des cognitive computing und boten sich daher als Partner für solch einen Hackathon an. Über den Hackathon heraus können so Kooperationen zwischen RKI und Experten auf dem Gebiet der künstlichen Intelligenz geschaffen werden.

Das IKW, deren Studenten und andere Forscher auf dem Gebiet des maschinellen Lernens sind immer auf der Suche nach interessanten Datensätzen und relevanten Anwendungsgebieten. Beides kann das RKI bieten und war damit ein willkommener Partner für einen Hackathon.

Format

Im Vorfeld haben wir 4 Themenkomplexe vorgestellt aus denen die Teilnehmer Projektideen wählen bzw. selbst entwickeln konnten.

Die knapp 30 Teilnehmer haben sich am ersten Tag zu 8 Gruppen zu je 2-4 Personen gefunden und eine Projektidee entworfen die von uns und den Betreuern des IKW begutachtet wurde.

Von Montag Nachmittag bis Freitag früh konnten die Teams frei an der Umsetzung ihrer Projekte arbeiten. Die Teams bekamen dabei Räume und Rechner vom IKW gestellt, durften aber auch an anderer Stelle arbeiten. Die Betreuer von Seiten des IKW und wir haben den Teams mehrmals täglich Beratung angeboten und waren meist von 9 bis 18 Uhr vor Ort ansprechbar.

Die Ergebnisse wurden bis Freitag früh um 6 eingereicht. Erwünscht waren eine 3-minütige Präsentation der Ergebnisse in einem Video und die Bereitstellung des Programmcodes, der Analysen und Ergebnisse.

Die Jury zur Beurteilung der Projekte bestand aus den 6 Organistoren des RKI, IKW und Gesundheitscampus Osnabrück. Für die Auswahl der Gewinner wurden zunächst die Videos und anderweitig

eingereichten Ergebnisse begutachtet und anschliessend durch eine jeweils ca. 20-minütige Befragung der Teams eine abschließende Beurteilung vorgenommen. Basierend darauf wurden die 8 Teams platziert.

Bei der Preisverleihung wurden sowohl eine persönliche Auszeichnung (Olivera Stojanovic, Einladung für 2-monatiges Praktikum am RKI) als auch Auszeichnungen für Teams vergeben. Die 4 Gewinner Teams erhielten Geldpreise (1 x 1500€, 3 x 500€) und die Einladung ans RKI zur Präsentation ihrer Ergebnisse vor Fachpublikum.

Vorbereitung

Der Gesundheitscampus Osnabrück hat sich um die Organisation der Räumlichkeiten, Bekanntmachung der Veranstaltung und Anmeldung der Teilnehmenden gekümmert. Das IKW organisierte die Vorbereitung der Teilnehmenden, der nötigen Infrastruktur während des Hackathons (Räumlichkeiten und Computer) und generell den reibungslosen Ablauf während der Implementierungsphase des Hackathons. Zusammen mit dem IKW haben wir die inhaltliche Vorbereitung (Formulierung und Präsentation der Projekte, Bereitstellung der notwendigen Daten) durchgeführt. Weiterhin haben wir uns um die Finanzierung der Räume (Tag 1 und 5) und des Caterings gekümmert.

Wir haben mehrere Themen gesammelt und uns dann entschieden, 4 Themenkomplexe zu wählen. Die Themen stammen zumeist aus Vorschlägen, Ideen und Wünschen von Mitarbeitern/innen der Abteilung 3. Die Projekte wurden so gewählt, dass sie ein für das RKI sinnvolles Produkt als Ziel haben und zugleich den Fähigkeiten der Teilnehmenden (hauptsächlich Studierende am IKW) entsprechen und für diese auch spannend sind.

Thema 1: Experten-/Supportsystem

Ein System, was Expert/innen dabei hilft, Entscheidungen zu treffen und dabei deren Feedback nutzt, um Fachwissen zu lernen um daraufhin bessere Vorschläge zu machen. Dabei werden zwei Anwendungen vorgeschlagen.

Anwendung A: Epidemic Intelligence (Idee mit FG32, Maria an der Heiden, Michaela Diercke, Bettina Rühle)

Für die wöchentlichen EpiLag-Runden werden immer wieder Themen gesucht. Dabei recherchieren mehrere Expert/innen interessante Berichte aus vielen verschiedenen Quellen und Sammlungen wie z.B. ProMed. Das Problem dabei ist die Fülle der Information. Systeme wie ProMed oder MediSys liefern hunderte von Beiträgen, die dann einzeln manuell eingeschätzt und priorisiert werden müssen. Ein System, das bei diesem Filtern und Priorisieren unterstützen könnte, wäre wünschenswert. Mit Hilfe von Clustering-Ansätzen und unter Einbeziehung von Kontextinformation könnte zunächst eine Vorauswahl getroffen werden. Die gefilterten/priorisierten Beiträge könnten in einer Oberfläche mit sortierbaren/filterbaren Tabellen und Karten dargestellt werden. Entscheidend ist dann die Einbeziehung und Abfrage des Fachwissen der Expert/innen. Jeder Vorschlag des Expertensystems kann bewertet werden und führt zu einem besseren Ergebnis beim nächsten Vorgang.

Anwendung B: Ausbruchserkennung

Bei den Signale-Berichten oder im zukünftigen Signale-Dashboard werden den Expert/innen Signale präsentiert, die eventuelle Ausbruchsevents signalisieren sollen. Die Expert/innen müssen nun durch ihre Erfahrung und Recherche entscheiden, ob das Signal tatsächlich ein Ausbruchsgeschehen darstellt. Das Expertensystem muss einen geeigneten Weg finden, dieses Wissen zu extrahieren, um es in späteren Iterationen einzubeziehen.

Thema 2: Daten-Crawler, Data Mining

Ein System, das potentiell epidemiologisch relevante Daten aus externen Quellen extrahiert, darstellt und auf deren Relevanz (Korrelation mit epidemiologischen Daten) untersucht. Auch hier sind zwei Anwendungsbeispiele vorgesehen.

Anwendung A: Todesanzeigen (Idee des FG36)

Zurzeit findet die Mortalitätssurveillance nur in Berlin und Frankfurt statt. Die Mortalitätsraten sind aber wichtig, um Exzessmortalität für Influenzawellen zu bestimmen. Hier sollen Todesanzeigen aus Lokalzeitungen analysiert werden, um Ort, Geburts- und Sterbedatum zu extrahieren und damit eine Art Mortalitätssurveillance zu gewährleisten.

Anwendung B: Versorgungsatlas oder Umweltatlas o.ä.

Der Grad der Gesundheitsversorgung (Anzahl Ärzt/innen pro Einwohner) oder Umwelteinflüsse (Luftverschmutzung) können eine Rolle in der Verbreitung von Infektionskrankheiten spielen. Verschiedene Indikatoren über die gesundheitliche Versorgung oder Umwelteinflüsse verschiedener Regionen (Flächeneinheiten, Landkreise, ...) liegen bereits vor. Diese Daten müssten extrahiert, aufbereitet und dargestellt werden. Anschließend können die Daten für Analysen über den Zusammenhang mit epidemiologischen Ereignissen genutzt werden.

Thema 3: Parametrisierbare Vorhersage

Bisher wird nur die aktuelle epidemiologische Lage eingeschätzt, eine Vorhersage von Fallzahlen oder Ausbruchsgeschehen wäre wünschenswert, ist aber auf Grund der Komplexität des Geschehens sehr schwierig. Hilfreich wäre aber trotzdem eine Möglichkeit, Vorhersagen für verschiedene Szenarien zu machen, dazu benötigt man eine Vorhersage, die über verschiedene epidemiologisch relevante Parameter verfügt. Es soll eine Möglichkeit geschaffen werden, dass Expert/innen diese Parameter einstellen können und sofort die angepasste Vorhersage bekommen.

Thema 4: Health-Monitoring-App: Selbst-Berichterstattung und automatische Beratung

Bei teilnahmebasierten Anwendungen wie der GrippeWeb-App ist es wichtig, dass die Nutzer/innen dauerhaft Einträge übersenden. Ohne Nutzen für sie kann es zu hohen Abbruchraten kommen. Daher wäre es wünschenswert, Mehrwert anzubieten, indem man zum Beispiel personalisierte (automatische) Beratung oder motivierendes Feedback liefert. Das Feedback könnte aus Darstellung und Analyse der Nutzerdaten entstehen.

Verlauf

Tag 1 - Montag, 12. Februar 2018

InnovationsCentrum Osnabrück, Albert-Einstein-Str. 1, 49076 Osnabrück

09:00 Uhr: Registrierung (Foyer im 2. OG)

10:00 Uhr: Eröffnung (Konferenzraum Edison, 2. OG)

- Eröffnung der Veranstaltung (Prof. Wolfgang Lücke, Präsident der Universität Osnabrück)
- Grußwort per Video (Prof. Lothar H. Wieler, Präsident des Robert Koch-Instituts)
- Begrüßung der Teilnehmer und namentliche Vorstellung der Faculty (Prof. Kai-Uwe Kühnberger, Direktor des IKW)

10:30 Uhr: Faculty & Topics

- Vorstellung der Faculty / Jury und der Themen / Challenges des Hackathons (Prof. Gordon Pipa, Leiter des Bereichs Neuroinformatik am IKW gemeinsam mit der gesamten Faculty / Jury)

11:00 Uhr: Programm

- Führung der Teilnehmer durch den Ablauf des Hackathons, Locations, Regeln, Arbeitsmodus, Deadlines, etc. (Prof. Gordon Pipa)

11:30 Uhr: Team up (Foyer und Räume im 2. OG)

- Teambildung und Austausch mit der Faculty bei Fingerfood

14:45 Uhr: Vorstellung der Teams (Konferenzraum Edison, 2. OG)

Institut für Kognitionswissenschaft, Wachsbleiche 27, 49090 Osnabrück

16:00 Uhr: Briefing

- Einführung in IKW Infrastruktur und Teamwork (Prof. Gordon Pipa)

Tag 2 - Dienstag, 13. Februar 2018

Institut für Kognitionswissenschaft, Wachsbleiche 27, 49090 Osnabrück

09:00 Uhr: Hackathon

- Selbstständiges Arbeiten der Teams
- Beratung aller Teams durch RKI Team und IKW Team

Tag 3 - Mittwoch, 14. Februar 2018

Institut für Kognitionswissenschaft, Wachsbleiche 27, 49090 Osnabrück

09:00 Uhr: Hackathon

- Selbstständiges Arbeiten der Teams
- Beratung aller Teams durch RKI Team und IKW Team

Tag 4 - Donnerstag, 15. Februar 2018

Institut für Kognitionswissenschaft, Wachsbleiche 27, 49090 Osnabrück

09:00 Uhr: Hackathon

- Selbstständiges Arbeiten der Teams
- Beratung aller Teams durch RKI Team und IKW Team

18:00 Uhr: Ergebnis-Videos

- Beratung zur Videoaufnahme
- selbständiges Erstellen von Videos zur Präsentation der Ergebnisse

Tag 5 - Freitag, 16. Februar 2018

Campus Westerberg

09:00 Uhr: Team Q&A Session

- Lernlandschaft im SL-Gebäude („Frosch“) Teams verteilen sich in der Lernlandschaft, tauschen sich aus und stehen für Fragen der Jury bereit

Botanischer Garten

13:30 Uhr: Preisverleihung / Awards (Helikoniensaal, Bohnenkamp-Haus)

- Begrüßung (Prof. Wolfgang Lücke, Präsident der Universität Osnabrück)
- Keynote zur Digitalisierung in der Gesundheitsversorgung (Staatssekretär Stefan Muhle, 2. Vorsitzender von GewiNet – Kompetenzzentrum Gesundheitswirtschaft)
- Preisverleihung / Awards (Prof. Kai-Uwe Kühnberger, Direktor des IKW, gemeinsam mit der Jury)
 - Verleihung der individuellen Preise
 - Präsentation der Pitch-Videos
 - Auszeichnung der vier Gewinner-Teams.
- Verabschiedung / Closing (Prof. Wolfgang Lücke)

– Einladung zum lockeren Ausklang der Veranstaltung bei Getränken

15:00 Uhr: Ende der Veranstaltung

Ergebnisse

1. Platz - Team “Die Influenza”

Prototyp einer App, wo jede Nutzerin oder jeder Nutzer Symptome eingeben kann und Fallzahlen (vergangene oder vorhergesagte) visualisieren kann. Die eingetragenen Symptome werden zusammen mit SurvStat-Daten ausgewertet, um eine Vorhersage der Krankheitsverbreitung in jedem Landkreis zu machen. Für die Vorhersage werden drei unterschiedliche neuronale Netzwerke verwendet. Ein Twitter-Bot berichtet darüber, welche drei Landkreise am auffälligsten sind.

2. Platz - Team “404”

Alle deutsche Online-Zeitungen werden automatisch gelesen, Traueranzeigen gefunden, daraus eine Statistik zu Mortalität mit Ort, Geschlecht, Alter gemacht. Es scheint sehr gut zu funktionieren, mit Verzug zum Todesdatum von etwa 2 bis 3 Tagen. Eine Vielzahl von Informationen, u.a. zu Umständen des Todes, können durch Bildanalyse (“optical character recognition”) der Traueranzeigen gewonnen werden.

2. Platz - Team “Healthy Herbert”

“Hierarchical temporal memory”, ein neues maschinelles Lernverfahren, wird verwendet, um Fallzahlen unter Berücksichtigung von verschiedenen Datenquellen vorherzusagen und Auffälligkeiten zu erkennen.

2. Platz - Team “Mercura”

Eine App wird prototypisch entwickelt, mit der durch einfaches Reden eine Ärztin oder ein Arzt einen Bericht zur syndromischen Lage machen kann und Bürgerinnen und Bürger sich informieren können. Maschinelles Lernverfahren werden auf SurvStat-Daten trainiert und analysieren die eingegebenen Daten, mit Einbindung von anderen Quellen.

Team “Ni”

Ein multivariates Regressionsmodell, das Variablen für verschiedene räumliche, zeitliche und raumzeitliche Effekte integriert. Das Modell war bereits sehr ausgereift und könnte Aufschluss über die Art der Verbreitung einer Krankheit geben und den Vergleich zwischen Krankheiten erlauben.

Team MedPred

Ein neuronales Netz zur Vorhersage von Krankheitsfällen und Ausbrüchen für die Folgewochen.

Team “EpiNoLag”

Ein System zur Erkennung von weltweiten infektionsepidemiologisch interessanten Gefahrenlagen. Dabei werden Promed Einträge gecrawled, Schlagwörter in den Einträgen gesucht und deren Auffälligkeit bewertet. Die Ergebnisse werden auf einer Website angezeigt.

Team “Mobile EEG”

Eine App welche die Daten eines mobile EEG Geräts per Bluetooth einliest, dann als Zeitreihen darstellt und einfache Analysen ermöglicht. Das Ziel der Analysen ist die Erkennung von Epilepsie-Patienten in Entwicklungsländern. Die App bietet zudem die Möglichkeit sich über Epilepsie zu informieren indem ein Chatbot befragt wird.

Kosten

Von unserer Seite sind Kosten für Catering während der gesamten Woche, Mieten der Räumlichkeiten an Tag 1 und 5, sowie die Reisekosten für 3 Mitarbeiter über 5 Tage.

Das Catering umfasste ein Mittagsbuffet und Getränke für Montag, ein Sektempfang am Freitag, belegte Brötchen, Kuchenbuffet und Getränke an allen Tagen. Das Catering kostete für die 5 Tage ca. 4000€.

Gemietet wurden am Montag ein großer Seminarraum und 3 kleinere Seminarräume, sowie am Freitag ein Saal im Botanischen Garten. Insgesamt wurden hierfür 280€ fällig.

Die Reisekosten für 3 Personen und 5 Tage ergaben insgesamt $3 \times 295\text{€}$ (Hotel) + $3 \times 80\text{€}$ (Bahn) = 1125€.

Die Gesamtkosten belaufen sich somit auf ca. 5500€.

Medien

Bereits im Vorfeld ist die Redaktion des NDR auf den Hackathon und uns aufmerksam geworden. Daraufhin wurde eine Recherche durchgeführt indem Stéphane Ghozzi und Ute Rexroth per Telefon über den Hackathon, die Signale Gruppe und Surveillance allgemein befragt wurden. Mit der Pressestelle wurden dann Statements abgesprochen die am Montag vor der Kamera geäußert werden sollten. Der erste Tag des Hackathons wurde komplett von einem Fernsehteam sowie einer Mitarbeiterin des NDR Info Radioteams begleitet. Dabei wurden sowohl Stéphane Ghozzi, als auch Mitarbeiter des IKW und die Teilnehmer interviewt. Dies wurde zu einem Beitrag im NDR Fernsehen zusammengefasst, siehe [NDR-Beitrag und Interview mit Prof. Pipa](#).

Wir standen während der gesamten Woche in Kontakt mit den Fernseh- und Radioteams des NDR bzw. NDR Info und unterrichteten über Verlauf des Hackathons und interessante Projekte bzw. Teams. Zur Preisverleihung war dann wieder das Team von NDR Info vor Ort und interviewte noch einmal Stéphane Ghozzi und die Gewinner Teams. Ein [Bericht](#) über das Gewinnerteam wurde dann am Samstag im NDR Radio ausgestrahlt. Das Event wurde im Kurznachrichtendienst Twitter mehrfach angekündigt und auch während und nach Durchführung kommentiert.

Aussicht

Das Projekt von Team “404” soll mittels eines Werkvertrags fortgesetzt werden um ein produktionsreifes Tool zu erhalten und damit eine Mortalitätssurveillance durch Crawlen von Traueranzeigen zu etablieren. Die Anforderungen wurden bereits zusammen mit dem relevanten Fachgebiet FG36 zusammengetragen. Ebenso wird gerade ein Datenschutzkonzept mit dem Datenschutzbeauftragten Jörg Lekschas erarbeitet.

Es wurde bereits einen Vertrag über 2 Monate mit einer Teilnehmerin geschlossen, um die Projektarbeit des Teams “Ni” fortzusetzen, mit dem Ziel ein produktionsreifes Analysemodell zu entwickeln und die Erkenntnisse in einer wissenschaftlichen Publikation festzuhalten.

Es ist geplant mehrere der Teilnehmer durch mehrmonatige Werkstudentenverträge die Chance zu geben ihre Projekte fortzuführen oder andere interessante Themen der automatischen Signalerkennung zu bearbeiten.

Fazit

Auch wenn eine abschließende Bewertung noch nicht vollständig möglich ist, empfinden wir die Durchführung des Hackathons trotz des hohen Organisationsaufwands als sehr lehr- und ertragreiche Veranstaltung. Wir haben interessante Ergebnisse erzielt, die eventuell durch verschiedene Praktika und Werkverträge fortgesetzt werden und damit in Produkte verwandelt werden, die am RKI routinemäßig verwendet werden. Auf jeden Fall sind Ideen entstanden, die uns in der Fortführung des Signale Projekts sehr helfen werden. Ebenfalls haben wir durch den Hackathon unsere Beziehung zum IKW, und damit Experten auf dem Gebiet des Cognitive Computing, gestärkt. Durch die enge Zusammenarbeit mit den Teilnehmern und die eventuellen weiteren Zusammenarbeiten können wir weitere Kontakte in diesem Bereich knüpfen und vielleicht sogar zukünftig als Mitarbeiter gewinnen. Zusätzlich haben wir öffentlichkeitswirksam unser Profil bezüglich digitaler Epidemiologie geschärft und werden nun auch mehr im Bereich des maschinellen Lernens als möglicher Kooperationspartner und Arbeitgeber wahrgenommen.

Ob weitere Hackathons sinnvoll sind muss auf Grund des hohen Aufwands noch besprochen werden. Dafür sollten die Ergebnisse der Praktika und Werkverträge, sowie eigene Umsetzung der Hackathon-Ideen abgewartet werden.