

# Entscheidungsunterstützung mit Bayesschen Netzen - Modellierung einer COVID-19 Domäne mit HUGIN

Sebastian Tillack

Bachelorarbeit • Studiengang ACS • Fachbereich Informatik und Medien • 26.01.2021

## Aufgabenstellung

Bayessche Netze (BN) sind gut zur Modellierung von Unsicherheit geeignet. Ein aktuelles Beispiel für das Auftreten von Unsicherheit ist die COVID-19 Domäne, insbesondere die Zusammenhänge zwischen u.a. Symptomen, Analysen, Auswirkungen und Folgen. Nach einer kurzen Einführung in die Grundlagen der BN sollen die wesentlichen Konzepte der COVID-19 Domäne einschließlich ihrer Zusammenhänge dargestellt werden. Eine Analyse zum Stand der Forschung zu BN, die genau diese Domäne bereits als Anwendung haben, schließt sich an, ebenfalls eine eigene kurze Bewertung. Kern der Bachelorarbeit ist eine eigene Umsetzung mit Hilfe des HUGIN-Tools.

## Bayessche Netze

Meistens werden sie in Form eines gerichteten azyklischen Graphen dargestellt. Dabei stellen die Knoten Variablen dar und die Kanten beschreiben bedingte Abhängigkeiten zwischen den Knoten. Die Richtung der Kanten stellt häufig den kausalen Zusammenhang dar. Die mathematische Grundlage ist das Theorem von Bayes:

$$P(A|B) = \frac{P(B|A) \cdot P(A)}{P(B)},$$

welches sich aus der Definition der bedingten Wahrscheinlichkeit ableiten lässt.

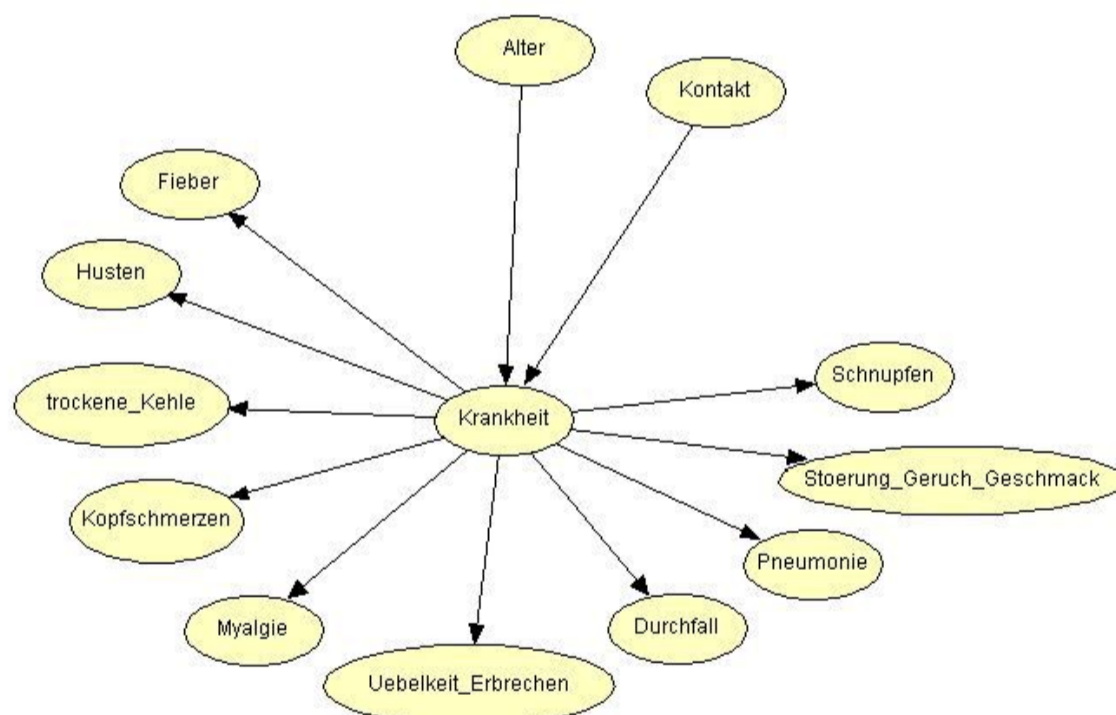


Abb. 1: Bayessches Netz in HUGIN

## Inferenz

Die Inferenz beschreibt im Wesentlichen, wie sich die Vertrauenswerte von Variablen verändern, wenn neue Evidenz bekannt wird. Man spricht auch von *Belief Updating*. Dabei unterscheidet man in exakte und approximative Inferenzverfahren. Damit sich alle Knoten aktualisieren können, werden Informationen zwischen benachbarten Knoten propagiert.

## SARS-CoV-2

Im Dezember 2019 wurde in China ein neuartiges Virus entdeckt, welches später SARS-CoV-2 (severe acute respiratory syndrome coronavirus type 2) genannt wurde. Die dadurch verursachte Krankheit COVID-19 (corona virus disease 2019) hat sich schnell weltweit verbreitet, sodass sie am 11. März 2020 von der Weltgesundheitsorganisation (WHO) als Pandemie eingestuft wurde. Die Symptome sind ähnlich zu denen anderer Erkältungskrankheiten oder Influenza, was eine Diagnose erschwert.

## Modellierung der Domäne mit HUGIN

Eine Domäne kann entweder aus Expertenwissen manuell modelliert oder aus großen Datensätzen automatisch erlernt werden. Bei einer manuellen Entwicklung werden zuerst die relevanten Variablen bestimmt. Daraufhin werden die möglichen Werte bzw. Zustände, die eine Variable annehmen kann, festgelegt. Als nächstes werden die Zusammenhänge der Variablen durch gerichtete Kanten dargestellt. Schlussendlich müssen die Tabellen der bedingten Wahrscheinlichkeiten aller Knoten ausgefüllt werden. Dafür können Fakten, Studienergebnisse, Expertenbefragungen, aber auch Schätzungen und Erfahrungswerte genutzt werden.

Kontakt	Ja		Nein	
	0-19	20+	0-19	20+
Covid-19	0,047	0,107	0,03	0,03
SARS	0,0001	0,0001	0,0001	0,0001
MERS	0,0001	0,0001	0,0001	0,0001
Influenza	0,001	0,001	0,001	0,001
keine Krankheit	0,9518	0,9518	0,9518	0,9518

Abb. 2: Bedingte Wahrscheinlichkeiten der Variable *Krankheit*

## Ergebnisse

Die entstandene Anwendung ermöglicht es, die Wahrscheinlichkeit einer Erkrankung an COVID-19, SARS, MERS oder Influenza zu bestimmen. Dafür werden die beobachteten Symptome dem Netz als Evidenz bekannt gemacht. Das heißt, der Wert der entsprechenden Variable wird festgelegt und ist nicht mehr abhängig von der ursprünglichen Wahrscheinlichkeit. Es lässt sich zeigen, dass spezifische Symptome, wie die Störung des Geschmacks und/oder Geruchssinns, die A-posteriori-Wahrscheinlichkeiten der Krankheiten stärker beeinflussen als häufige Symptome wie Husten.

## Fazit

In dieser Arbeit wurde gezeigt, dass Bayessche Netze beim Treffen von Entscheidungen unter Unsicherheit hilfreich sein können. Damit eine COVID-19 Domäne in der Praxis eingesetzt werden kann, muss jedoch die Krankheit und insbesondere der Einfluss von Vorerkrankungen genauer erforscht werden. Außerdem müssen möglicherweise ortsspezifische Änderungen vorgenommen werden, um verschiedene Prävalenzen und demografische Unterschiede zu berücksichtigen

## Literatur

### Bayessche Netze:

Boersch, Ingo ; Heinsohn, Jochen ; Socher, Rolf: *Wissensverarbeitung: Eine Einführung in die Künstliche Intelligenz für Informatiker und Ingenieure*. 2. Auflage, Spektrum Akademischer Verlag, 2007.

### HUGIN Tutorial:

<https://hugin.com/wp-content/uploads/2016/05/Building-a-BN-Tutorial.pdf>

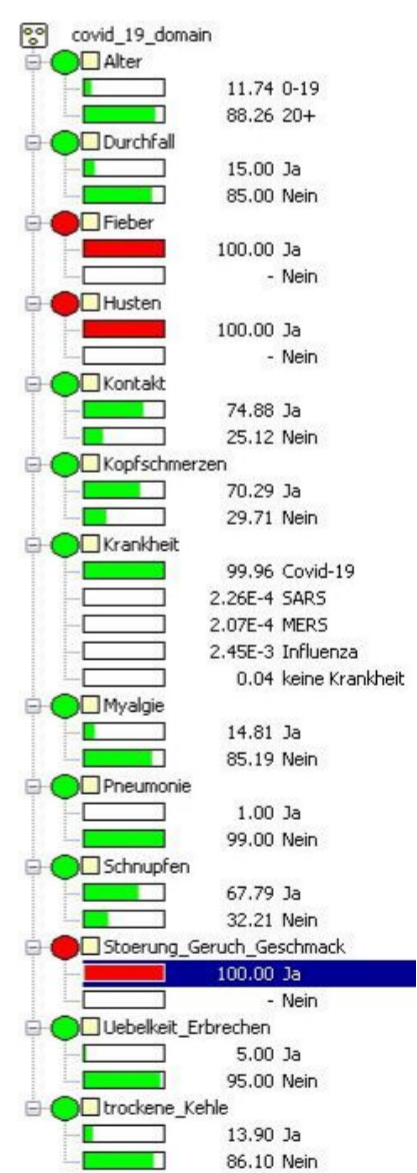


Abb. 3: rot: Evidenz; grün: berechnete Wahrscheinlichkeiten