FACULTY OF ENGINEERING AND ARCHITECTURE

IDLab INTERNET & DATA LAB

Clinical Reasoning over Tabular Data and Text with Bayesian Networks

Paloma Rabaey, Johannes Deleu, Stefan Heytens, Thomas Demeester

Motivation

Bayesian networks are ideally suited to **automate clinical reasoning**.

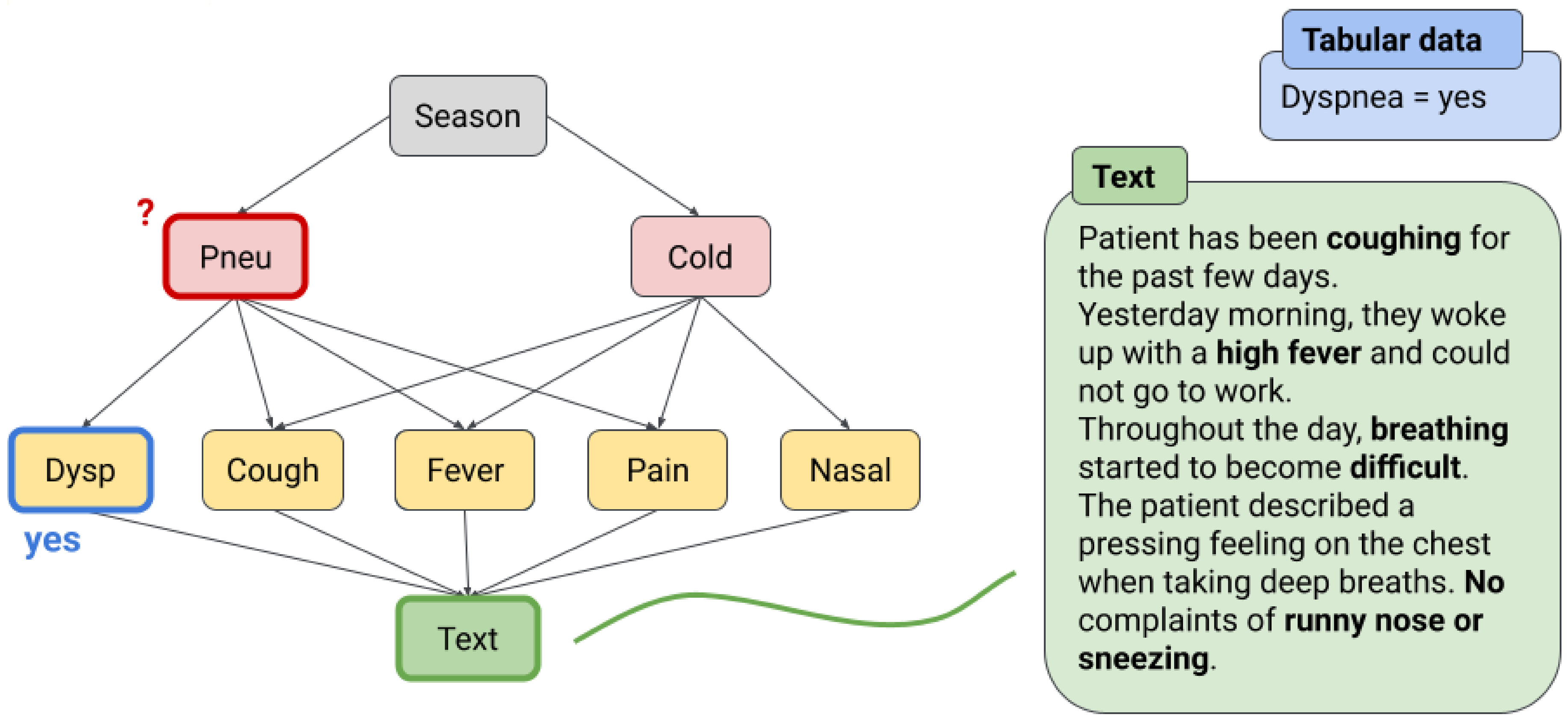
1. Model complex problems involving uncertainty
2. Interpretable graphical structure
3. Combine data and expert knowledge

However, they enjoy **limited adoption** in clinical practice due to their inability to deal with realistic medical data.

Electronic health records (EHRs) are a mix of…

- **Structured tabular data**: encoded symptoms and diagnoses (ICPC, ICD-10/11), medication codes, lab results…
- **Unstructured text**: detailed clinical notes providing further context on the structured portion of the data

Joint clinical reasoning over structured tabular data and unstructured text



Use-case and dataset

Use-case: prediction of **respiratory diseases** (pneumonia vs. common cold) in primary care.

Artificial dataset with 5000 patient records containing mixed data (tabular + text).

- Some symptoms are partially encoded
- Some symptoms are never encoded in tabular form

➔ **Automated clinical reasoning needs to take text into account.**

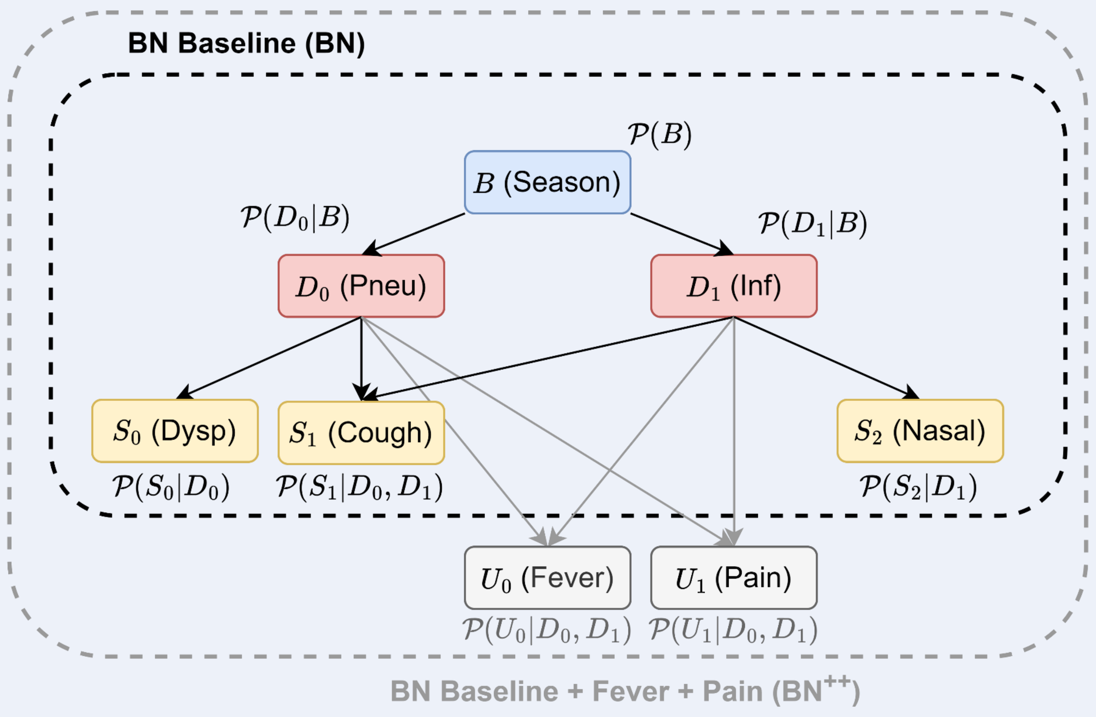
Generate tabular data using **expert-defined Bayesian network**

Ask **ChatGPT** to **generate a clinical note** given patient symptoms (no diagnoses!)

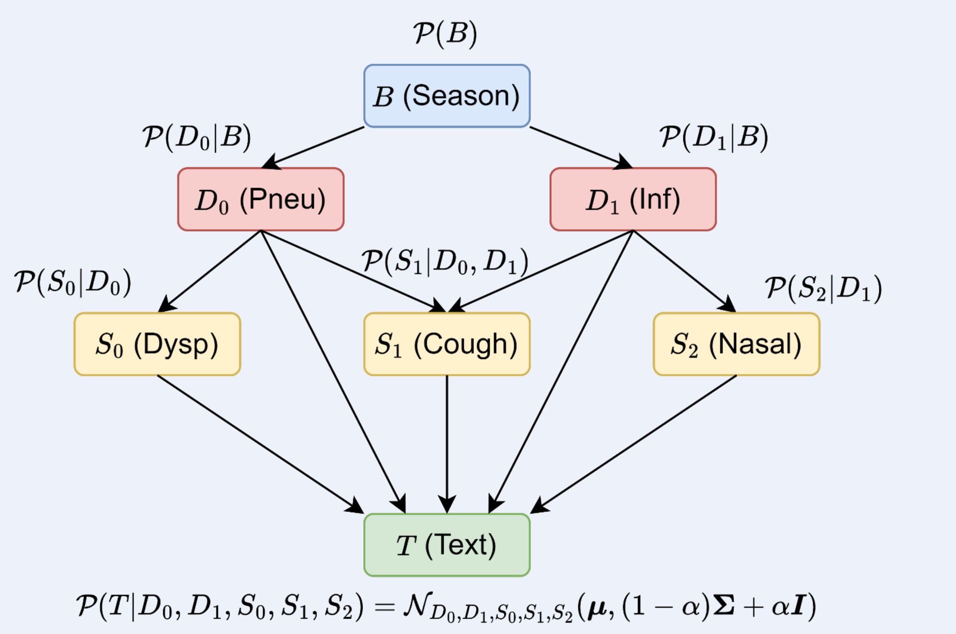
	Pneu	Cold	Season	Dysp	Cough	Fever	Pain	Nasal	Text
0	No	No	Summer	No	No	~~None~~	~~No~~	No	Patient complaining of abdominal pain after eating; referred for further evaluation.
1	No	No	Summer	/	/	~~/~~	~~/~~	/	Patient reports shortness of breath on exertion for the past two weeks. No associated cough or chest pain. Low-grade fever occasionally present.
2	No	Yes	Winter	No	Yes	~~None~~	~~No~~	Yes	Patient presents with a dry cough and sneezing.

Models

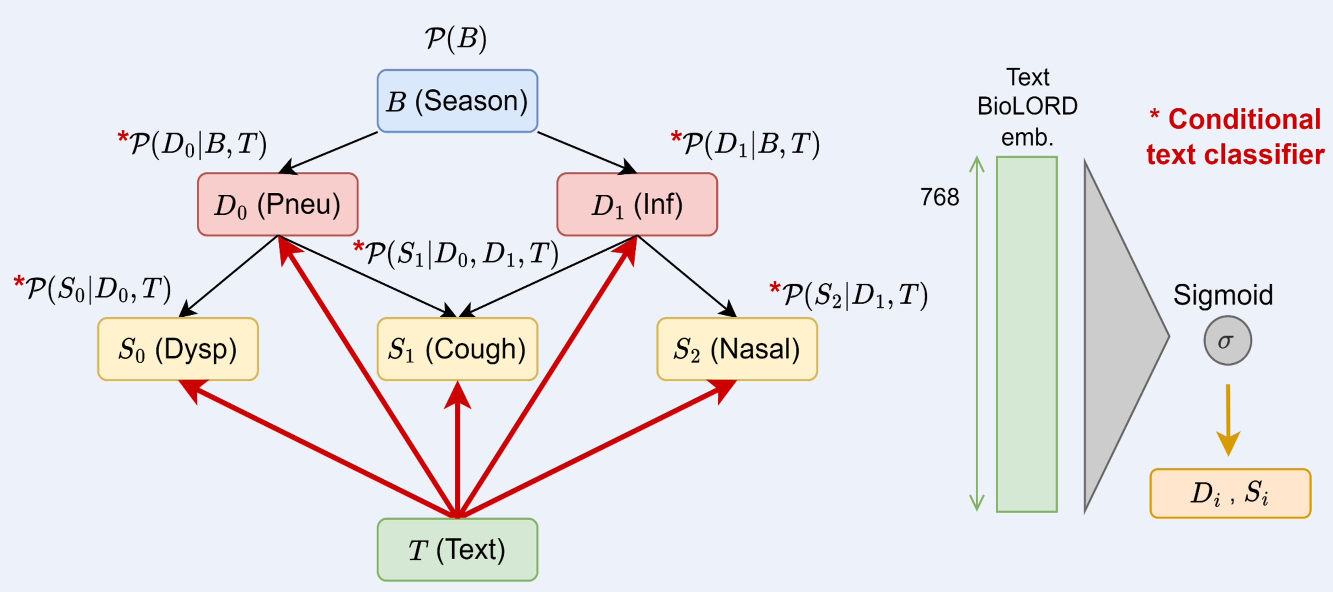
Baseline: Bayesian network (BN and BN^{++})



Bayesian network with text generator (BN-gen-text)



Bayesian network with text discriminator (BN-discr-text)



Results

Evaluate avg. precision of the **posterior diagnostic probability**, under 3 sets of evidence.

- $\boldsymbol{P(D_0|B, S_0, S_1, S_2, T)}$: predict the probability of pneumonia (D_0), given tabular background (B), all observed **tabular symptoms** (S_0, S_1, S_2) and the **text** (T) as evidence.
- $\boldsymbol{P(D_0|B, S_0, S_1, S_2)}$: predict the probability of pneumonia (D_0), given only tabular background (B) and all observed **tabular symptoms** (S_0, S_1, S_2) as evidence.
- $\boldsymbol{P(D_0|B, T)}$: predict the probability of pneumonia (D_0), given only tabular background (B) and **text** (T) as evidence

	$\boldsymbol{P(D_0\|B, S_0, S_1, S_2, T)}$	$\boldsymbol{P(D_0\|B, S_0, S_1, S_2)}$	$\boldsymbol{P(D_0\|B, T)}$
BN	-	0.0914 ($\pm$0.0000)	-
BN^{++}	-	0.8326 ($\pm$0.0000)	-
BN-gen-text	0.5870 ($\pm$0.0000)	0.0892 ($\pm$0.0007)	0.4434 ($\pm$0.0000)
BN-discr-text	0.7538 ($\pm$0.0323)	0.1079 ($\pm$0.0011)	0.6922 ($\pm$0.0000)

Main take-aways

1. BN-discr-text model manages to successfully **integrate text into the reasoning process** using neural classifiers, while maintaining the **interpretable modular structure** of the BN.
2. BN-discr-text model is able to **leverage information on the unobserved symptoms** fever and pain, without explicitly modeling these in the BN or ever having access to these in an encoded tabular format.
3. BN-text falls behind due to its **difficulty in modeling the distribution of the text directly**.

GHENT UNIVERSITY

fwo

paloma.rabaey@ugent.be
github.com/prabaey/bn-text
/in/paloma-rabaey-3a29091a2