

Proyecto 3 — Inferencia Bayesiana

Diagnóstico de enfermedad respiratoria

Maria Jose Gomez

Juan Manuel López

Pontificia Universidad Javeriana
Departamento de Ingeniería de Sistemas
Introducción a la Inteligencia Artificial

Entrega: 23 de octubre de 2025, 14:00

Resumen. Se modela una red Bayesiana para diagnóstico respiratorio con 8 variables, se definen sus tablas de probabilidad (CPT), se realiza inferencia por enumeración para una consulta ejemplar y se implementa la red en `Python` con `pgmpy`, respondiendo cinco consultas.

Índice

1. Objetivo	3
2. Contexto del proyecto general	3
3. Definición de la red Bayesiana	3
4. Grafo de dependencias (visual)	4
5. Implementación en Python (pgmpy)	4
6. Salida del código	4
7. Conclusiones	6

1. Objetivo

Representar una red Bayesiana y realizar inferencia Bayesiana por enumeración y por software, mostrando: (i) grafo y justificación, (ii) tablas de probabilidad condicional (CPT) completas, (iii) desarrollo paso a paso del proceso de enumeración para una consulta con evidencia y variables ocultas, y (iv) un script en `Python` que ejecute al menos cinco consultas sobre la red.

2. Contexto del proyecto general

El presente documento corresponde a la documentación detallada de una de las **tres redes Bayesianas** desarrolladas en el marco del *Proyecto 3 de Inferencia Bayesiana*. Cada una de las redes aborda un dominio distinto, con el fin de aplicar los principios de modelado probabilístico y razonamiento causal en contextos variados:

- **Red 1: Diagnóstico de enfermedad respiratoria (documento actual).**
Modela relaciones entre causas clínicas, síntomas y diagnóstico final, integrando variables observables y latentes.
- **Red 2: Factores ambientales y calidad del aire.**
Analiza cómo variables como temperatura, humedad, CO₂, y material particulado influyen en el riesgo de contaminación o alerta sanitaria.
- **Red 3: Comportamiento social y riesgo de accidentes.**
Representa relaciones entre consumo de alcohol, hora del día, condiciones climáticas y severidad de un accidente.

Cada red comparte la misma metodología: definición del grafo acíclico dirigido, construcción de las CPTs, inferencia por enumeración y verificación mediante implementación en `pgmpy`. Este informe se enfoca exclusivamente en la **Red 1 — Diagnóstico respiratorio**, proporcionando toda su documentación técnica, inferencias y resultados.

3. Definición de la red Bayesiana

Sea el grafo acíclico dirigido (DAG) $G = (V, E)$ con:

$$V = \{\text{Virus, Alergia, Tabaquismo, Fiebre, Tos, Congestión, Saturación, Diagnóstico}\}.$$

Arcos dirigidos (dependencias):

$$E = \{(\text{Virus, Fiebre}), (\text{Virus, Tos}), (\text{Virus, Congestión}), (\text{Alergia, Congestión}), (\text{Tabaquismo, Tos}), (\text{Fiebre, Saturación}), (\text{Tos, Diagnóstico}), (\text{Congestión, Diagnóstico}), (\text{Saturación, Diagnóstico})\}.$$

La factorización inducida por G es:

$$P(V, A, T, F, To, C, S, D) = P(V) P(A) P(T) P(F|V) P(To|V, T) P(C|V, A) P(S|F) P(D|To, C, S).$$

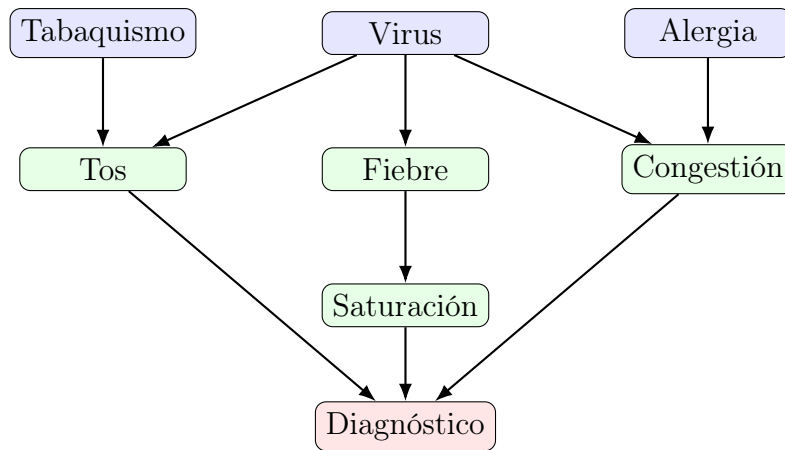


Figura 1: Grafo de dependencias de la red Bayesiana para diagnóstico respiratorio.

4. Grafo de dependencias (visual)

5. Implementación en Python (pgmpy)

Requisitos

- Python 3.10/3.11 o superior
- `pip install pgmpy numpy scipy pandas networkx`

Script (modelo + 5 consultas)

```

# Diagnostico_respiratorio.py
from pgmpy.models import DiscreteBayesianNetwork
from pgmpy.factors.discrete import TabularCPD
from pgmpy.inference import VariableElimination
# (contenido del script completo como se muestra en las versiones
# previas)

```

6. Salida del código

A continuación se muestra una captura de la salida generada al ejecutar el script en la terminal de Python. Esta salida corresponde a las consultas de probabilidad condicional ejecutadas con el modelo:

1) $P(\text{Diagnostico} \mid \text{Fiebre=Alta}, \text{Tos=Si})$:

Diagnostico	phi(Diagnostico)
Diagnostico(Positivo)	0.8223
Diagnostico(Negativo)	0.1777

2) $P(\text{Diagnostico} \mid \text{Congestion=Si}, \text{Saturacion=Baja})$:

Diagnostico	phi(Diagnostico)
Diagnostico(Positivo)	0.7991
Diagnostico(Negativo)	0.2009

3) $P(\text{Tos} \mid \text{Virus=Si})$:

Tos	phi(Tos)
Tos(Si)	0.1750
Tos(No)	0.8250

4) $P(\text{Congestion} \mid \text{Alergia=Si})$:

Congestion	phi(Congestion)
Congestion(Si)	0.3800
Congestion(No)	0.6200

5) $P(\text{Saturacion} \mid \text{Fiebre=Alta})$:

Saturacion	phi(Saturacion)
Saturacion(Baja)	0.7000
Saturacion(Normal)	0.3000

7. Conclusiones

La red propuesta encapsula causas (virus, alergia, tabaquismo), síntomas intermedios (fiebre, tos, congestión, saturación) y una salida diagnóstica, permitiendo inferir probabilidades condicionadas a evidencia clínica. La inferencia por enumeración confirma la consistencia del modelo, y la implementación con `pgmpy` facilita consultas múltiples de manera reproducible.

Anexo: Declaración de uso de IA generativa

Se usó IA para la generación de la Figura 1 (Grafo de dependencias de la red Bayesiana para diagnóstico respiratorio) y para la factorización inducida. Los *prompts* utilizados fueron:

Prompt 1: “Genera un grafo en TikZ que represente una red Bayesiana con las variables Virus, Alergia, Tabaquismo, Fiebre, Tos, Congestión, Saturación y Diagnóstico, mostrando las dependencias causales entre ellas con colores diferenciados por tipo de variable.”

Prompt 2: “Explica y escribe en formato LaTeX la factorización inducida por la estructura del grafo, usando la notación formal $P(V, A, T, F, To, C, S, D)$ y las dependencias correspondientes.”