

PROPUESTA DE PROYECTO: Planificador de Dosis de Insulina con Aprendizaje por Refuerzo

1. Descripción del Problema

Contexto: La diabetes mellitus tipo 1 requiere un control constante de los niveles de glucosa mediante insulina. Ajustar la dosis correcta es complejo porque cada paciente responde de forma diferente y la insulina tiene efectos retardados. Si la dosis es baja, la glucosa sube (hiperglucemia); si es alta, puede provocar hipoglucemia, que es peligrosa.

Problema técnico: Diseñar un sistema que aprenda a recomendar la dosis óptima de insulina en función del estado actual del paciente. Esto implica equilibrar eficacia y seguridad, manejando un espacio de decisiones amplio y con múltiples variables clínicas.

Relevancia: Controlar la glucosa en el rango saludable (70–180 mg/dL) es clave para reducir complicaciones. Actualmente, este control depende de la experiencia del paciente o del médico. Un agente de aprendizaje por refuerzo (RL) puede ofrecer un apoyo automatizado y adaptativo.

Fuentes:

- [American Diabetes Association \(2023\). Glycemic Targets: Standards of Care in Diabetes.](#)
- [Bergman, R. N. et al. \(1979\). Quantitative estimation of insulin sensitivity.](#)

2. Datos / Entorno del Agente

Entorno simulado: Se utilizará un entorno virtual basado en modelos médicos validados (como el modelo de Bergman). El agente aprenderá a administrar insulina a un paciente simulado observando los niveles de glucosa y otros factores.

Estados:

- Glucosa: <70, 70–180, 180–250, >250 mg/dL
- Insulina activa: 0, 5, 10, 15, 20 unidades
- Tiempo desde la última dosis: 0–60, 60–120, 120–240, >240 minutos
- Sensibilidad del paciente: baja, normal, alta

(Total: 240 combinaciones posibles)

Acciones: Cantidad de insulina a administrar: [0, 5, 10, 15] unidades.

Recompensa:

```
if 70 <= glucosa <= 180: return +10 # Rango ideal
elif glucosa < 60: return -20 # Hipoglucemia
elif glucosa > 250: return -15 # Hiperglucemia
elif dosis > 20: return -5 # Exceso
else: return -1 # Penalización leve
```

Objetivo del agente: Aprender una política que mantenga la glucosa dentro del rango óptimo el mayor tiempo posible.

Métrica	Meta
Tiempo en rango (70–180 mg/dL)	>70%
Hipoglucemias (<70 mg/dL)	<5%
Hiperglucemias (>180 mg/dL)	<25%

3. Modelo Seleccionado: Q-Learning con Función de Aproximación

Justificación: Q-Learning es simple, estable y permite entender las decisiones del agente con facilidad, ideal para una primera versión del proyecto. La función de aproximación ayuda a generalizar el aprendizaje a distintos pacientes sin requerir demasiados datos.

Simulador médico ↔ Agente RL (Q-Learning) ↔ Ambiente de diabetes

Ventajas:

- Permite entender por qué el agente toma cada decisión.
- Tiene bajo costo computacional.
- Facilita pruebas y validación con distintas configuraciones.

4. Consideraciones Éticas y Limitaciones

- El proyecto no busca reemplazar decisiones médicas reales, sino demostrar el uso del RL en un entorno simulado.
- Se usarán datos sintéticos basados en parámetros médicos reales.
- Se garantizará transparencia y trazabilidad en los resultados.

5. Referencias

1. [Bergman, R. N. et al. \(1979\). Quantitative estimation of insulin sensitivity.](#)
2. [American Diabetes Association. \(2023\). Glycemic Targets: Standards of Care in Diabetes.](#)
3. [Man, C. D. et al. \(2014\). The UVA/PADOVA Type 1 Diabetes Simulator: New Features.](#)
4. [Fox, I. et al. \(2019\). Deep reinforcement learning for closed-loop blood glucose control.](#)
5. [Hovorka, R. et al. \(2004\). Nonlinear model predictive control of glucose concentration.](#)