A337 Examen Parcial parte II

Usted y su compañero de equipo trabajan para para una obra social que está estudiando qué factores inciden sobre la diabetes, y desean tener un predictor de la probabilidad de que una persona tenga o desarrolle diabetes a partir de una serie de preguntas sobre sus hábitos y estado de salud.

En el centro disponen de la base de datos de la Universidad de California Irvine

<https://archive.ics.uci.edu/dataset/891/cdc+diabetes+health+indicators>

y quieren en base a esos indicadores que le armen un modelo logístico y uno de redes neuronales feedforward ( que deberán diseñar ) para predecir el desarrollo de la enfermedad.

Ambos modelos deben ser construidos usando la plataforma keras.

Se espera que provean un informe comparativo de la performance y métricas de bondad de clasificación ( accuracy, recall, precisión, F1 score, AUC, KS ) en ambos modelos.

Del modelo logístico, se espera que puedan identificar qué factores son los más determinantes en la incidencia de diabetes. Busque evidencia respaldatoria en otros trabajos de investigación o publicaciones periodísticas que respalden sus hallazgos.