

Evidencia Análisis



Jacques Emiliano López Andalco A01746203

Septiembre, 2023

Instituto Tecnológico y de Estudios Superiores de Monterrey

Concentración Inteligencia Artificial Avanzada

Evidencia

Introducción

Para esta evidencia, utilizaremos una base de datos médica denominada "Fetal Health Classification". El conjunto de datos proporcionado consta de 21 columnas de variables y una columna con las etiquetas que buscamos predecir. Entre nuestras características, se encuentran diversas medidas relacionadas con cuestiones cardíacas, con un énfasis significativo en las mediciones obtenidas a partir de histogramas cardíacos. A continuación, se presentan las 21 características que nuestra red neuronal deberá aprender a utilizar:

baseline value: Frecuencia cardíaca fetal basal (FHR)

accelerations: Número de aceleraciones por segundo

fetal_movement: Número de movimientos fetales por segundo

uterine_contractions: Número de contracciones uterinas por segundo

light_decelerations: Número de desaceleraciones leves por segundo

severe_decelerations: Número de desaceleraciones graves por segundo

prolonged_decelerations: Número de desaceleraciones prolongadas por segundo

abnormal_short_term_variability: Porcentaje de tiempo con variabilidad a corto plazo anormal

mean_value_of_short_term_variability: Valor medio de la variabilidad a corto plazo

percentage_of_time_with_abnormal_long_term_variability: Porcentaje de tiempo con variabilidad a largo plazo anormal

mean_value_of_long_term_variability: Valor medio de la variabilidad a largo plazo

histogram_width: Ancho del histograma realizado utilizando todos los valores de un registro

histogram_min: Valor mínimo del histograma

histogram_max: Valor máximo del histograma

histogram_number_of_peaks: Número de picos en el histograma del examen

histogram_number_of_zeroes: Número de ceros en el histograma del examen

histogram_mode: Moda del histograma

histogram_mean: Media del histograma

histogram_median: Mediana del histograma

histogram_variance: Varianza del histograma

histogram_tendency: Tendencia del histograma

Las etiquetas de la variable objetivo clasifican la salud del feto en tres categorías: normal (1), sospechosa (2) y patológica (3). La capacidad de predecir el estado de salud del feto en las etapas iniciales de su desarrollo puede tener un impacto significativo en su supervivencia. El objetivo de nuestras redes neuronales será predecir la clase de salud de nuevas instancias, es decir, de nuevos embarazos, utilizando esta información médica detallada.

Datos

El dataset de enfermedades fetales es apropiado para el algoritmo de aprendizaje automático y para demostrar la generalización en un modelo de red neuronal secuencial debido a varias razones. En primer lugar, el dataset contiene características médicas relevantes que se pueden utilizar para predecir la salud fetal, lo que lo hace adecuado para un enfoque de clasificación. Además, al dividir el conjunto de datos en conjuntos de entrenamiento, validación y prueba, se siguen buenas prácticas para evaluar la generalización del modelo. La elección de una red neuronal secuencial permite capturar relaciones complejas entre las características médicas, lo que es fundamental para la detección precisa de enfermedades fetales. La inclusión de capas densas y una función de pérdida adecuada, como la "sparse categorical cross-entropy", se alinea con la naturaleza de la tarea de clasificación multiclase. En conjunto, estos elementos hacen que el dataset sea apropiado para entrenar un modelo de aprendizaje

profundo y demostrar su capacidad para generalizar en datos no vistos, con el enfoque de división de conjuntos de entrenamiento, validación y prueba.



Fig1. Training Samples



Fig2. Test Samples



Fig3. Validation Samples

En los gráficos presentados con anterioridad divididos de la siguiente forma (60%, 20%, 20%). Podemos observar un problema de desbalance en la distribución de las clases por separación de los datos. Lo que significa que la división de datos de esta manera. Priva al modelo de conocer la clase tres, ya que esta solo se encuentra en los datos de prueba. Para solucionar esto, podemos aplicar un oversampling. Para poder tener las clases objetivo balanceadas y no sesgar el modelo.

Evaluación

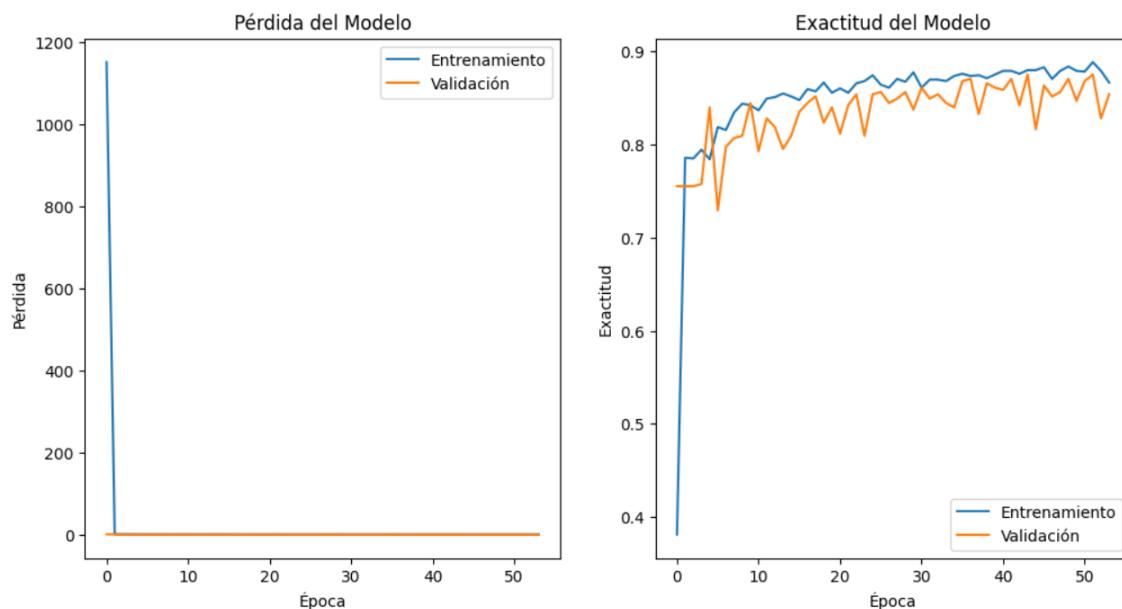


Fig. 4 Visualización de la función de pérdida así como de exactitud entrenamiento y validación

En la gráfica anterior, podemos observar claramente la generalización del modelo. En líneas azules, observamos los datos de entrenamiento y en amarillo los de validación. Tomamos dos métricas para realizar esta prueba, la función de métrica (izq) y la exactitud (der). Podemos ver como la línea amarilla, en ambos casos, tiende a ajustarse a los valores de entrenamiento, sin estar exactamente sobre ella. Razón por la cual, podemos observar la generalización del modelo.

```
# Análisis del ajuste del modelo
# Determinar si el modelo está subajustado
if sesgo > 0.2:
    ajuste_modelo = 'Subajustado'
elif sesgo <= 0.2 and varianza > 0.2:
    ajuste_modelo = 'Sobreajustado'
else:
    ajuste_modelo = 'Bien Ajustado'
```

```
Sesgo (Error de Entrenamiento): 0.1137
Varianza: 0.0416
Ajuste del Modelo: Bien Ajustado
```

De igual forma se calcula el sesgo y la varianza del modelo. Con estas métricas, estableceremos un rango, donde podremos posicionar al modelo y continuar con la evaluación. Como podemos observar, el modelo se encuentra bien ajustado según los límites que hemos establecido.

En cuestión de otras métricas, con recall, f1 y la matriz de confusión. Podemos ver que detalles le falta ajustar al modelo para seguir mejorando. En general, el modelo obtenido es bueno. Más que realizar un refinamiento del modelo para elevar su calificación. Se propone la construcción de una base de datos que represente de manera fidedigna y completa la realidad de los fenómenos en la salud fetal. De la información presentada a continuación, vale la pena resaltar lo que sucede con la etiqueta intermedia. Es la única en la que el modelo predice más otras clases que la verdadera.

```
Puntuación de Recuperación: 0.8568
Puntuación F1: 0.8455
Matriz de Confusión:
[[317  13   2]
 [ 34  22   3]
 [  4   5  26]]
```

Conclusión

La realización de este proyecto abre la puerta a numerosas incógnitas en relación al aprendizaje profundo. La diversidad de herramientas con sus respectivos hiper parámetros brinda un alto grado de versatilidad a las redes neuronales, permitiéndoles abordar una amplia variedad de desafíos. En lo que respecta al problema planteado, obtenemos valores destacados para la precisión de nuestro modelo. Es fundamental recordar que la métrica "precisión" simplemente representa el resultado de dividir las predicciones correctas entre el total de predicciones realizadas. Dado que nuestra base de datos presentaba sesgo, no está de más realizar una matriz de confusión para identificar en qué áreas el modelo comete errores en sus predicciones. Como se presentó en este reporte. En resumen, es crucial hacer un uso eficiente de la amplia gama de herramientas a nuestra disposición. Ya sea que estas herramientas nos ahorren tiempo, eviten el "sobreajuste" o incluso costos computacionales, comprender su funcionamiento y aplicación es esencial. En los modelos que presenté, uno de ellos no considera los hiper parámetros, mientras que el segundo incorpora nuevas herramientas y se amplía con muchas más capas ocultas, lo que resulta en un modelo superior en comparación con el primero presentado.