

```
# Crear un nuevo ejemplo para predecir
nuevo_dato = pd.DataFrame([
    {'Runtime': 120,          # Duración de la película en minutos
     'Released_Year': 2020,   # Año de lanzamiento
     'IMDB_Rating': 8.5,      # Calificación en IMDB
     'Meta_score': 80,        # Puntuación en Metascore
     'No_of_Votes': 500000    # Número de votos
    }])
```

Ingreso bruto predicho para el nuevo dato: \$206,121,391.10

El modelo de árbol de decisión utilizado permite estimar el ingreso bruto de una película a partir de distintas características, mostrando un desempeño aceptable según el gráfico de dispersión entre valores reales y predichos. Aunque existe una cierta dispersión, especialmente en valores altos, el modelo logra capturar la tendencia general, lo que indica una capacidad predictiva razonable, aunque no perfecta.

Del análisis de importancia de las variables, se concluye que el número de votos (**No_of_Votes**) es el factor más determinante en la predicción del ingreso bruto, seguido por el año de lanzamiento (**Released_Year**) y la calificación en IMDb (**IMDB_Rating**). Esto sugiere que la popularidad y el reconocimiento del público tienen un peso significativo en el éxito económico de una película, incluso más que métricas críticas como el metascore o la duración.

Al aplicar el modelo a un nuevo ejemplo con características específicas —una película de 120 minutos, lanzada en 2020, con calificación IMDb de 8.5, metascore de 80 y 500,000 votos— se obtuvo una predicción de ingreso bruto de aproximadamente **\$206,121,391.10**. Este resultado valida la utilidad del modelo para hacer estimaciones basadas en variables conocidas, destacando cómo una alta participación del público puede traducirse en ingresos elevados.